GLOBAL AIRBORNE BACTERIA AND ANTIBIOTIC RESISTOME FROM NATURAL DISTRIBUTION TO ANTHROPOGENIC IMPACTS: COMMUNITY STRUCTURE, BIOGEOGRAPHY, DRIVING MECHANISM, AND HUMAN HEALTH

ZHAO JUE

PhD

The Hong Kong Polytechnic University

2023

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

Global Airborne Bacteria and Antibiotic Resistome from Natural Distribution to Anthropogenic Impacts: Community Structure, Biogeography, Driving Mechanism, and Human Health

Zhao Jue

A thesis submitted in partial fulfilment of the requirements for

the degree of Doctor of Philosophy

March 2023

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

Zhao Jue (Name of student)

Abstract

Air pollution has become one of the top environmental issues across the globe, particularly bioaerosols, playing a crucial role in the interactions between ecosystems, climate, and public health. Airborne bacteria and potentially hazardous components therein, such as antibiotic resistance genes (ARGs), are part of the bioaerosols. It is imperative to identify their biogeographic patterns, driving mechanisms, source contributions, and potential impact on human health. However, a gap remains in the understanding of these key interactions with regard to the systematic surveillance of airborne bacterial community and quantified antibiotic resistance (AMR) risk at a global scale. To address this scientific challenge, the present PhD study used the 16S rRNA gene dataset and metagenomic dataset in the global atmosphere to establish the first atlas of global airborne bacteria from background sites and urban areas with a wide geographic and altitudinal range.

The data obtained in this study revealed the maximum microbial richness in the intermediate latitudinal regions and dynamic airborne bacterial community structure by encompassing the bacterial community among the three largest ecosystems on the Earth's surface (*i.e.*, atmospheric, oceanic, and terrestrial systems). Among the thousands of bacteria detected in the global atmosphere, the core set (*n=24*) and key taxa (*n=19*) were identified and confirmed to impact the entire complex community network of interconnected bacteria. Both the uniform latitudinal bacterial diversity pattern and similarities in compositions and inferred functions of key taxa across various ecosystems suggested their potential linkages. Thereby, the Earth's bacterial co-occurrence network involving 23 habitats indicated the importance of airborne bacteria to the planetary

microbiomes and the great contributions from surface environments to airborne bacteria compositions (46.25%).

Anthropogenic activities also affected the airborne bacterial community structure and corresponding phenotypes via reduced environmental filtering effects and elevated human-related source contributions. Notably, the higher abundance and diversity of airborne pathogens in urban areas indicated a higher exposure risk of pathogens in densely populated regions. In terms of driving mechanisms, although airborne bacterial communities assembled in more stochastic processes (72.4%), the deterministic processes, such as biotic interactions and environmental filtering, showed significant impacts on global airborne bacterial community structure. Both biotic and abiotic factors could influence the structure and distribution of global airborne bacterial communities; however, the most determinant mechanism was the environmental filtering process. Even though atmosphere is a highly dynamic and flowing ecosystem, its bacterial community was discovered to be largely affected by local environments, particularly for the source contributions and air pollutants from human activities, indicating that global climate change and impacted air quality may cause alterations of airborne bacterial abundance, community structure, and diversity.

To further explore the exposure risks in urban air, the airborne AMR was investigated at a global scale, especially for the estimation of quantified ARG-related health risks. The differentia in the composition and abundance of airborne ARGs and pathogens (carrying ARGs) between urban and background areas illustrated the anthropogenic impacts on airborne AMR exposure risks: the ARGs associated with major urban sources (*e.g.*, wastewater treatment plants (WWTPs), hospitals, and landfills) contributed more than 30%

of the urban airborne ARGs in total, which could lead to a higher risk rank of airborne ARGs in urban areas than background areas, while the abundance of urban indicator ARGs was largely affected by antibiotic consumptions, particularly aminoglycoside, tetracycline, and beta-lactam, raising the importance of the appropriate use of these drugs. The signatures of ARGs and mobile genetic elements (MGEs) co-occurrence were much more frequent in urban air, exhibiting the higher mobility of airborne AMRs. A genome-resolved "panorama" of AMR was also revealed, and humans were found to inhale more potential antibiotic-resistant pathogens daily (averagely 789.75 ± 586.80 cells) in cities than in background areas (105.6 ± 82.05 cells). Furthermore, the *Staphylococcus aureus* genome identified in the MAGs generated from the global urban air samples showed close genetic relatedness to those strains of Methicillin-resistant *Staphylococcus aureus* (MRSA) isolated from nosocomial infections and was shown to carry *mecC*, which suggested airborne transmission as a possible route in the prevalent community acquisition of MRSA and other types of resistant infections alike. The higher proportion of horizontally transferred ARGs in urban air also provided early warning for the rapid spread of AMR, and their dissemination from sources to human inhalation contributed to predicting future threats and improving public health management.

To sum up, this integrated study conducted comprehensive research on the bacterial (pathogenic) community, ARGs profiles, and AMR risks in the global atmosphere with quantitative estimations of source contributions, shaping mechanisms, and related potential health risks. These findings will provide key reference for predicting the evolution of airborne bacteria (and pathogens) and ARGs in a changing climate condition and highlights

the urgent need to involve biological parameters, such as airborne microbiome and AMR, in the current and future air quality standards on public health.

Acknowledgements

The PhD study has been a truly life-changing experience for me, and it would not be achievable without the guidance, encouragement, and corporation from many people in the study period.

I would like to first convey my genuine gratitude to my supervisor, Xiang-dong Li, for all the insightful guidance and continued support he gave me, during the years I spent undertaking my PhD study. He played as not only a supervisor in my study but also a mentor in my life. His overwhelming passion for science and rigorous approach to research impressed and motivated me a lot. Besides, he always gave me a psychological lift after a setback and helped me using his accumulated wisdom through years of experience.

My heartfelt appreciation goes out to the research group members: Dr. Ling Jin, Dr. Jia-wen Xie, Dr. Dong Wu, Dr. Tang-tian He, Xiao-hua Zhang, Chang-chao Li, and Yi-hua Wang, for their constructive suggestions and methodological guidance. I am also grateful to the scientific officers and technicians, including Mr. Hang Liu, Mr. C.S. Chan, Mr. W.S. Lam, Ms. Carman Ip, and Ms. Celine Che for their assistance with my lab work. I greatly appreciate the cooperations from Prof. James M. Tiedjej, Prof. Xiao-san Luo, Prof. Jun Li, Prof. Xue-wu Fu, Prof. Zhi-yuan Cong, Prof. Ping-qing Fu, Prof. Yang Zhang, Prof. Xin-bin Feng, and Prof. Gan Zhang and many more, who provided us the achieved samples for the first comprehensive global study on airborne bacterial community.

I gratefully acknowledge the PolyU PhD studentship and the research funding from the National Natural Science Foundation of China (92043302), the Strategic Priority Research

Program of the Chinese Academy of Sciences (XDB40020102), and the Hong Kong Research Grants Council (15210618, 15203920 and T21-705/20-N).

Finally, I am indebted to my parents, husband, and parents-in-law for always believing in me and supporting me wholeheartedly. I would like to give special thanks to my son, and his coming made me stronger, better, and more fulfilled than I could have ever imagined.

Table of Content

Abbreviations

Abbreviations	Full name
16SPIP	16S Pathogenic Identification Process
AIC	Akaike information criterion
AMR	antimicrobial resistance
AOR	abundance-occupancy relationship
AP	air pressure
AQI	Air Quality Index
ARB	antibiotic-resistant bacteria
ARG(s)	antibiotic resistance gene(s)
AT	air temperature
CFI	comparative fit index
CFU	colony-forming units
DDR	distance-decay relationship
EC	element carbon
EMP	Earth Microbiome Project
ESKAPE	*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumonia, Acinetobacter baumanni, Pseudomonas aeruginosa and Enterobacter* spp.
FCM	flow cytometry
GDP	gross domestic product
GTDB	Genome Taxonomy Database
HBP(s)	human bacterial pathogen(s)
HGT(s)	horizontal gene transfer(s)
HT-ARGs	horizontally transferred ARG(s)
HVF(s)	human virulent factor(s)
HVF-PARB	virulent potential antibiotic resistant bacteria
ICP-OES	inductively coupled plasma - optical emission spectrometry
IGBP	International Geosphere – Biosphere Programme

Abbreviations	Full name
*intI*1	the class 1 integrase gene
LEfSe	linear discriminant analysis effect size
LGD	latitudinal gradient of diversity
MAG(s)	metagenome-assembled genome(s)
MGE(s)	mobile genetic element(s)
MODIS	Moderate-resolution imaging spectroradiometer
MRSA	methicillin-resistant *Staphylococcus aureus*
MSSA	methicillin-sensitive *Staphylococcus aureus*
NCBI	National Center for Biotechnology Information
NMDS	non-metric multidimensional scaling
NOAA	National Oceanic and Atmospheric Administration
NPB	non-pathogenic bacteria
NST	normalized stochasticity ratio
OC	organic carbon
ORF(s)	open reading frame(s)
OUT(s)	operational taxonomic unit(s)
PARB	potential antibiotic-resistant bacteria
PC1	the first principal coordinates in PCoA
PCoA	Principal Coordinate Analysis
PM(s)	particulate matter(s)
PM_{10}	particles less than 10 μm in diameter
$PM_{2.5}$	fine particles with a diameter of 2.5 microns or less
qPCR	quantitative polymerase chain reaction
RDA	redundancy analysis
RH	relative humidity
RMSEA	root means square error of approximation
rRNA	ribosomal ribonucleic acid
SEM	structure equation modeling
SRA	Sequence Read Archive

Abbreviations	Full name
SRMR	standardized root means square residual
S. aureus	*Staphylococcus aureus*
sub-MICs	sub-minimum inhibitory concentrations
TLI	Tucker-Lewis index
TOR	thermal optical reflectance
TSP	total suspended particulates
VFDB	Virulence Factor Database
VPA	variation partitioning analysis
WD	wind direction
WHO	World Health Organization
WS	wind speed
WWTP(s)	wastewater treatment plant(s)

Chapter 1 Introduction

1.1 Background

In recent years, air pollution has become an increasingly serious problem and the top environmental concern worldwide (Landrigan, 2017), particularly bioaerosols, showing the great importance of atmospheric pollution and critical impacts on human health (Fröhlich-Nowoisky *et al.*, 2016). Bacteria, as ubiquitous biogenic aerosol particles, are also one of the key components of the global microbiome community and are of paramount importance to the ecosystem on a global scale (Burrows *et al.*, 2009a). Ambient air plays a crucial role in bacteria dissemination in the environment (Burrows *et al.*, 2009b) and is inhaled directly and constantly by humans and animals without any purification virtually. All the above facts show that airborne microbiomes are inextricably linked to ecological conditions and human health.

Comprehensive and systematic studies on the global microbial community in soil (Bahram *et al.*, 2018), ocean (Sunagawa *et al.*, 2015), and wastewater treatment plants (WWTPs) (Wu *et al.*, 2019) have been systematically conducted. The findings show that each ecological habitat harbors unique microbiomes and reveals interconnections between microbiomes in the atmosphere and other surface environments. Nevertheless, air has usually been recognized as purely a conduit for terrestrial and aquatic microbial life (Berendonk *et al.*, 2015); however, it is actually also a habitat of microorganisms (Womack *et al.*, 2010), with over 1×10^4 bacterial cells/m^3 (Burrows *et al.*, 2009a) and hundreds of unique taxa (Brodie *et al.*, 2007). Globally, there has been limited documentation of airborne microbiomes, including their community structures, biogeography, anthropogenic

impacts, and interconnections with other microbiomes on the Earth. Conducting a comprehensive and systematic large-scale study can help to reveal the crucial role of the atmosphere in contributing to microbial habitats on Earth. Moreover, such a study can aid in predicting how ecosystems may respond to environmental changes such as climate, air quality, land use, and human activities (Smets *et al.*, 2016).

Microbes never live in isolation, irrespective of the intra-environment (atmosphere) or inter-environments, accompanied by multifarious ecological relationships, such as mutualism and competition (Faust and Raes, 2012). Theoretically, these interrelationships based on an ultra-large sample size could be mathematically modelled as an adjacent matrix in recent studies (Proulx *et al.*, 2005; Reshef *et al.*, 2011; Faust and Raes, 2012), such as network structures, for soil (Barberán *et al.*, 2012), plant (Agler *et al.*, 2016), and marine ecosystems (Gilbert *et al.*, 2012), as well as for the human microbiome (Ma *et al.*, 2020a). However, the pivotal mode of dissemination, namely the atmospheric environment, remains unresolved. Moreover, investigations pertaining to Earth's bacterial co-occurrence network still represent a significant gap in the scholarly comprehension of the function of aerial bacterial communities within the wider microbial domain and their interconnections with various microbiomes in diverse ecosystems.

Increasing evidence suggests that human activities are affecting airborne bacterial communities (Burrows *et al.*, 2009a; Fröhlich-Nowoisky *et al.*, 2016; Li *et al.*, 2019), but the specific alterations due to urbanization from a global view and the related mechanisms remain elusive, which is essential for determining the interaction between human activities and existing airborne microbiomes and for comprehending the delicate balance between humans and the natural world.

In addition to the airborne microbes and their metabolites as biological pollutants, antibiotic resistance genes (ARGs) have attracted global attention since they were clearly considered as a kind of emerging environmental contaminant in the last 2-3 decades (Pruden *et al.*, 2006). The propagation of antimicrobial resistance (AMR) is an increasing global threat to human health (Sugden *et al.*, 2016), which annually caused about 1.27 million deaths worldwide in 2019 (Murray *et al.*, 2022), and this may increase to 10 million deaths a year by 2050 if no action is taken to control it (May, 2021). The emergency of clinically relevant ARGs, antibiotic-resistant pathogens, and "superbugs" released from anthropogenic sources, together with the overuse and misuse of antibiotics across the globe, have been viewed by many as serious environmental problems (Berendonk *et al.*, 2015). The ARGs are commonly considered to be contaminants (Zhang *et al.*, 2022) and AMR can be developed, transmitted, and prevail extensively in natural environments (*e.g.*, water (Zhu *et al.*, 2017), soil (Bahram *et al.*, 2018), and air (Xie *et al.*, 2019), and engineered systems (*e.g.*, wastewater systems (Hendriksen *et al.*, 2019) and landfill sites (Wu *et al.*, 2017)), as well as clinical settings (Chng *et al.*, 2020) via environmental pathways. Thereby, a quintessential "One Health" concept emphasizing the interdependence of human, animal, and environmental health, has been set up to tackle AMR (Vikesland *et al.*, 2017; Hernando-Amado *et al.*, 2019). During this process, air plays a vital role in channelling the transmission of ARGs, because ARGs could be released together with aerosols into the atmosphere from typical AMR contributors, (*e.g.*, hospitals (Wu *et al.*, 2022), WWTPs (Xie *et al.*, 2022), and animal husbandries (McEachran *et al.*, 2015)), and subsequently disseminate worldwide via atmospheric circulation (Fröhlich-Nowoisky *et al.*, 2016).

Of all the environmental compartments that facilitate the persistence of ARGs, ambient air shows closer and more pervasive interconnections with human beings, because ambient air is capable of being inhaled into human body virtually without any treatment, compared with processed food and water resources before ingestion (Morawska *et al.*, 2022). The inhaled ARGs enduringly exist and can be greatly enriched in the atmosphere (approximately $10^{2\text{-}3}$ copies/m^3 quantified on genetic levels) (Xie *et al.*, 2019), and pathogenic antibiotic-resistant bacteria (ARB) increase the risks of exposure and resistant infections through air inhalation (Wu *et al.*, 2022).

The ARGs carried by these bacterial contaminants can cause self-inheritance, be horizontally transferred to other host bacteria through horizontal gene transfer (HGT) or be transmitted from environments to host bacteria with the help of mobile genetic elements (MGEs) (Kruse and Sørum, 1994; Layeghifard *et al.*, 2017). However, the major ARG hosts in ambient air and their roles in impacting human health and driving the transmission of ARGs remain unknown but show great importance (Li *et al.*, 2018). Fortunately, technical improvements in DNA extraction from airborne particles and the availability of metagenomic library preparation and shotgun sequencing analysis with low biomass contributed to drawing a complete picture of airborne ARG profiles (Jiang *et al.*, 2015). The identification of ARG-hosting bacteria at the species level and their pathogenicity judgment could better estimate the real risks of dissemination and the exposure of inhalable ARGs from environmental reservoirs to human pathogens, particularly considering clinically relevant ARGs (Zhang *et al.*, 2021) and ESKAPE nosocomial pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter*

Baumann, Pseudomonas aeruginosa, and Enterobacter spp.) exhibiting multidrug resistance and virulence (Mulani *et al.*, 2019).

Recent regional and global surveys on airborne ARGs revealed disparities in ARG profiles and risks across countries and regions (Xie *et al.*, 2018; Xie *et al.*, 2019). This study has revealed greater variations in the abundance of bacteria and ARGs in temperate regions compared to subtropical urban areas, with distinct ARG profiles observed regardless of land-use gradients within each region. Additionally, airborne $PM_{2.5}$ exhibits a higher level of ARG enrichment across various environmental and human media. The findings of the regional and global patterns of airborne ARGs are echoing current clinical observations on AMR diseases. According to a report by the Hospital Authority of Hong Kong, the proportions of AMR in *Staphylococcus aureus* (*S. aureus*) resistant to methicillin and *E. coli* resistant to third-generation cephalosporins and fluoroquinolones are both at a high level in Asia, (*e.g.*, China and Hong Kong), but at a low level in Northern Europe, (*e.g.*, Norway and Sweden) (Protection, 2017). Moreover, the lowest age-standardized mortality rates per 100,000 population for *E. coli*, *K. pneumoniae*, and *S. aureus* were also observed in Northern Europe (Mestrovic *et al.*, 2022). The regional disparities in AMR profile and health risks could be ascribed to the locally independent source contributions, antibiotic consumptions, land-use gradients, and others (Li *et al.*, 2018; Xie *et al.*, 2022). Meanwhile, global-scale studies assessing health risks of ARGs from aquatic, soil, engineered, and human-related habitats have been systematically documented (Bahram *et al.*, 2018; Hendriksen *et al.*, 2019; Cuadrat *et al.*, 2020; Zhang *et al.*, 2022). ARGs can disseminate, in the case of both the intra-environment (atmosphere) and inter-environments (across media); nevertheless, the interrelationships between various habitats and contributions

from other ecosystems remain less understood. Hence, studying the global profile of airborne AMR could not only reflect the regional differences in AMR burden but also understand the driving forces of environmental dissemination of antibiotic resistance towards the environmental-clinical dialogues on AMR risks. Humans have made great efforts to intervene in the arise and spread of AMR; however, coordinated action at national or international level is largely absent, due to the lack of global surveillance of AMR situation (Laxminarayan *et al.*, 2013).

Collectively, systematic surveillances of spatiotemporal variations of the airborne bacterial community and quantified AMR risks at a global scale would contribute to an integrated understanding of anthropogenic impacts on air quality from (micro)biological perspectives. Furthermore, determining the pathogenicity of airborne bacteria and conducting exposure risk assessments for AMR could yield valuable insights for health-focused pollution control that is tailored to local-specific conditions. This approach could help mitigate global threats to planetary and human health.

1.2 Research Objectives

The aim of this thesis is to conduct a comprehensive and systematic investigation into the airborne bacterial community and AMR in the global atmosphere. The study seeks to establish the linkage among emission sources, the underlying driving mechanisms, and the resultant health risks. The global airborne bacterial dataset and the global airborne metagenomic dataset were established, respectively, based on both background and urban air samples with a wide geographic range across the globe, coupled with16S ribosomal ribonucleic acid (rRNA) sequencing data (*n=5,000*) from the Earth Microbiome Project (EMP) (Thompson *et al.*, 2017) and metagenomic data (*n=700*) retrieved from the National Center for Biotechnology Information (NCBI) as a supplement, to address the aforementioned knowledge gaps. The objectives of the study are shown as follows (Figure **1-1**):

a) To establish a comprehensive atlas of global airborne bacteria by deciphering their biogeographic patterns and structures;

b) To further understand the role of airborne bacterial communities in the Earth's microbiomes and their potential sources across the globe;

c) To explore the mechanisms driving biogeographic patterns of global airborne bacteria and to quantify the importance of anthropogenic impacts and environmental filtering in the process;

d) To systematically characterize ARG profiles and distributions worldwide and to estimate the potential sources of airborne ARGs; and

e) To estimate the risks to public health caused by AMR through inhalation across the globe, particularly in urban areas.

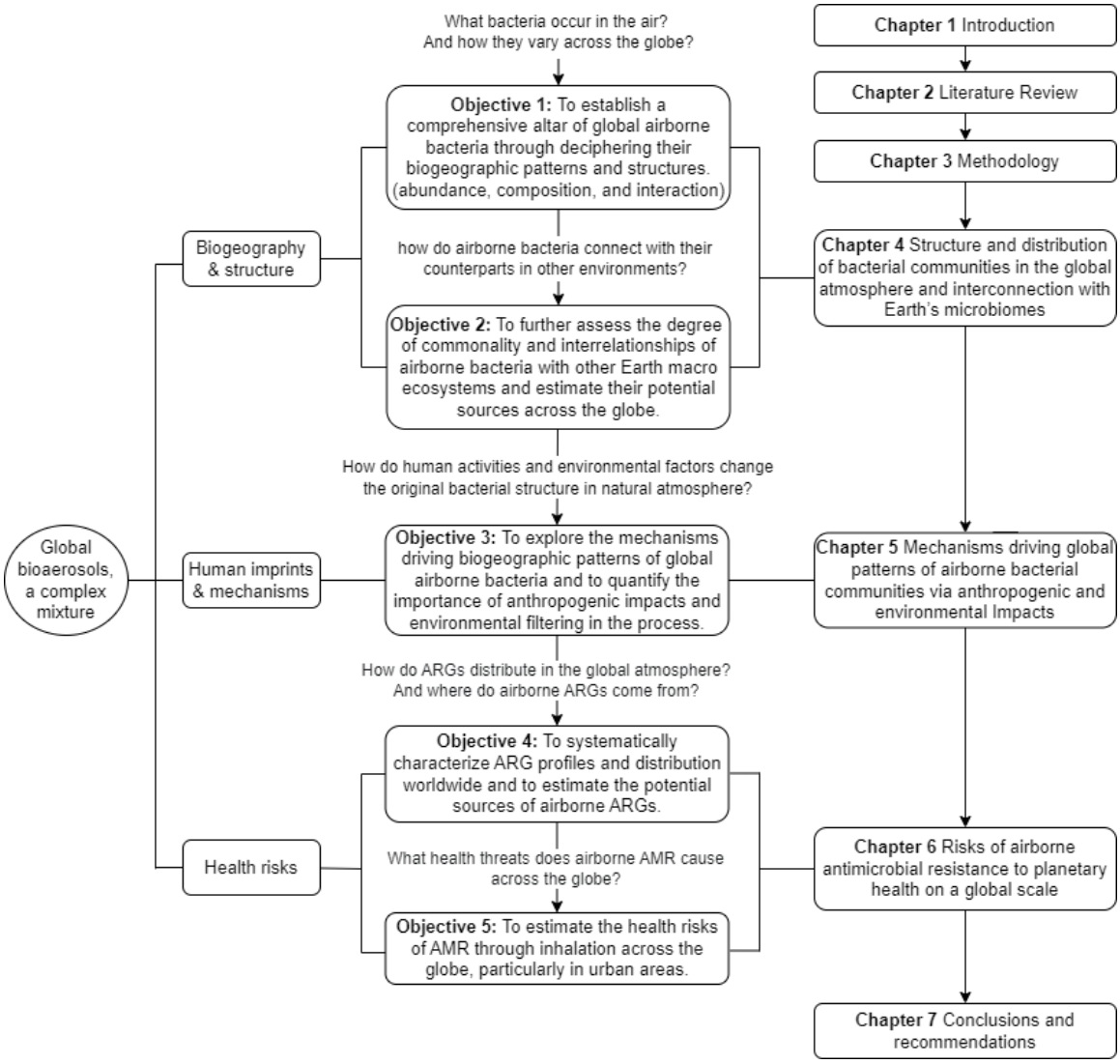


Figure 1-1 A flowchart that outlines the objectives and structure of the study.

1.3 Organizations

This thesis consists of seven chapters (Figure **1-1**), including the "Introduction", "Literature review", "Methodology", three chapters on the results and discussions, and the closing chapter of "Conclusions and Recommendations".

Chapter 1 "Introduction" provides a brief overview of the background information on bioaerosols, the threat of AMR, the research background, and knowledge gaps. It also states the research objectives of this study and outlines the structure of this thesis to help readers visualize the contents.

Chapter 2 "Literature review" provides a brief overview of the background information on bioaerosols, the threat of AMR, the research background, and knowledge gaps. It also states the research objectives of this study and outlines the structure of this thesis to help readers visualize the contents.

Chapter 3 "Methodology" outlines the data sources, sampling methods, and chemical and biological analysis technologies used in the study. It also describes the sequencing, data processing, and statistical calculations applied in this thesis.

Chapter 4 investigates the structure and biogeographic patterns of the global airborne bacterial community and its interconnections with the Earth's microbiomes, as well as the potential source contribution from other habitats to global airborne bacteria.

Chapter 5 explores the underlying mechanisms shaping the structure and distribution patterns of the global airborne bacterial community and emphasizes the importance of environmental filtering and anthropogenic impacts in this process.

Chapter 6 reveals the distribution and dissemination of ARGs as well as their hosts detected in global ambient air and quantitatively surveys the potential exposure risk of AMR through inhalation across the globe.

Chapter 7 “Conclusions and future suggestions” summarizes the major findings of the study and emphasizes its novelty and scientific significance. Additionally, the chapter provides suggestions for future research in this area.

Chapter 2 Literature Review

Given that bioaerosols are complex mixtures with the probability of causing serious health concerns through inhalation, this chapter provides an overview of airborne bacteria and ARGs along with advancements in detection and analysis techniques. Adopting the "One Health" approach, this review concentrates on two key biological components, bacteria and ARGs, and extends further to their biogeographical patterns, driving mechanisms, source contributions, and exposure risk assessments, thereby increasing comprehension of the existing research gaps and assisting in the resolution of scientific concerns.

2.1 Ambient air pollution, a global health-threatening environmental issue

With the rapid industrialization and urbanization worldwide, air pollution problems have more frequently emerged in recent years, which requires rethinking assessment methods of health risks and regulatory measures for reducing related risks (Li *et al.*, 2019). The global issue becomes more complicated under climate change, resulting in frequent episodes of serious air pollution driven by extreme weather events (Kan *et al.*, 2012). For instance, haze and smog have occurred frequently in recent years at a global scale, particularly in highly developed and densely populated urban areas (Zhang and Cao, 2015). These pollution events are long-lasting and broadly covered with a high intensity of pollution, thereby causing the hazards of air quality, climate conditions, ecosystem interactions, and human health at regional and global scales (Fröhlich-Nowoisky *et al.*, 2016). According to the World Health Organization (WHO), approximately 99% of the world’s population was

in regions where the air quality exceeded the WHO's air quality guideline levels (WHO, 2021). Moreover, it was estimated that more than 4 million people die prematurely due to ambient air pollution from both urban and rural areas every year by the WHO (WHO, 2021). Notably, the even more health-damaging particles are fine particles with a diameter of 2.5 microns or less ($PM_{2.5}$), which can penetrate the lung barrier and enter the blood system, enhancing the risk of developing cardiovascular and respiratory diseases, as well as lung cancer (Fröhlich-Nowoisky *et al.*, 2016; Smets *et al.*, 2016; Chen *et al.*, 2017; Cohen *et al.*, 2017).

It has been demonstrated by mounting evidence that geographical locations could cause differences in health risks related to air pollution (Zhang and Cao, 2015; Chen *et al.*, 2017; Cohen *et al.*, 2017). Although the concentration of $PM_{2.5}$ was positively related to the death rate from air pollution worldwide (Figure **2-1**), it was not the only driving factor to dominate the risk assessment of air quality (Hannah Ritchie and Roser, 2019). For instance, Europeans and North Americans presumably die from heart diseases or acute respiratory disorders more than Chinese people with exposure to similar $PM_{2.5}$ concentrations (Chen *et al.*, 2017). These above health issues center on the key features of particulate matters (PMs) as complex physicochemical and biological mixtures evolving with regional characteristics and temporal dynamics (West *et al.*, 2016). As a result, it is necessary to make the health threats of complex PM mixtures clear and to quantify the various components in PMs, especially the major causative agents, from the perspective of effective public health interventions (Jin *et al.*, 2016).

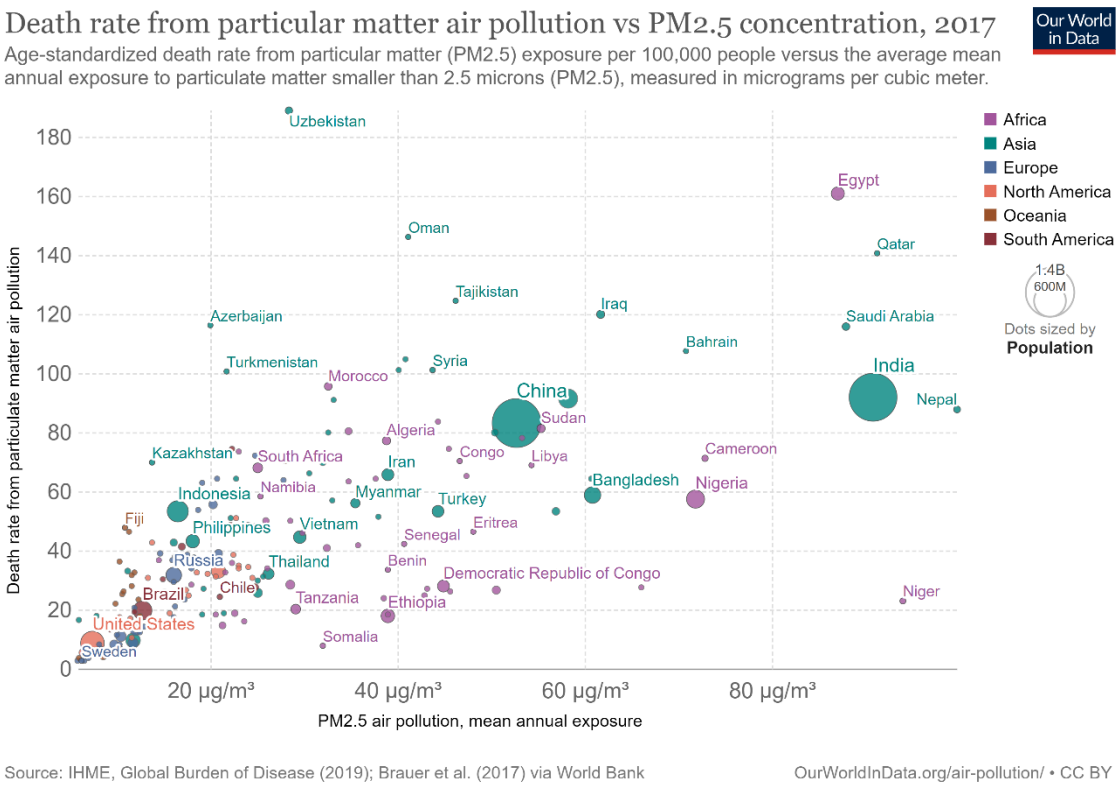


Figure 2-1 Death rate from particular matter air pollution *vs.* $PM_{2.5}$ concentration (Hannah Ritchie and Roser, 2019).

2.2 Bioaerosols in the Earth system

Bioaerosols play an important role in the Earth system, showing tight junctions of atmosphere (Ariya and Amyot, 2004), biosphere (Pöschl and Shiraiwa, 2015), climate (Morris *et al.*, 2013), and public health (Ghosh *et al.*, 2022). Aerosols of biological origin, such as airborne bacteria, pollen, and fungal spores, are necessary for the spread and propagation of human, animal, and plant species, across different ecosystems, which can also cause diseases in these organisms (Fröhlich-Nowoisky *et al.*, 2016). Moreover,

bioaerosols can affect the hydrological cycle and climate change by serving as nuclei for cloud droplets, ice crystals, and precipitation (Huffman *et al.*, 2013).

2.2.1 Characterization of bioaerosols

Bioaerosols, a subset of atmospheric particles, are directly released from the biosphere into ambient air. They contain a mixture of both alive and dead microorganisms (*e.g.*, archaea, bacteria, fungus, virus), dispersal units (*e.g.*, fungal spores and plant pollen), and various excretions or organism fragments (*e.g.*, plant debris and brochosomes) (Madelin, 1994; Matthias-Maser *et al.*, 1995; Rogerson and Detwiler, 1999; Graham *et al.*, 2003; Womack *et al.*, 2010; Després *et al.*, 2012; Tesson *et al.*, 2016). The diameters of bioaerosols range from nanometers to about one-tenth of a millimeter. The upper limit of this range is determined by rapid sedimentation, as larger particles are too heavy to remain airborne for long periods (Pöschl, 2005).

Research has shown that urban air contains several nanograms of DNA per cubic meter, indicating that a human adult may inhale up to ~0.1–1 μg of DNA per day, equal to 10^7-10^8 bacterial genomes or 10^4–10^5 human genomes (Després *et al.*, 2007). Additionally, materials of biological origin accounted for approximately 25% of atmospheric aerosol, and bioaerosols contributed as much as 5–10% of atmospheric PMs (Jaenicke, 2005). Thereby, as an important atmospheric component, biological particles may show important effects on atmospheric events and are influential on human health (Fröhlich-Nowoisky *et al.*, 2016).

2.2.2 Emission and transport of bioaerosols

Different models have been widely used to enhance the understanding of the emission and transport of bioaerosols in previous studies at global or regional scales (Heald and Spracklen, 2009; Burrows *et al.*, 2009a; Burrows *et al.*, 2009b; Hoose *et al.*, 2010; Sesartic *et al.*, 2012; Burrows *et al.*, 2013; Ansari *et al.*, 2015; Hummel *et al.*, 2015). Nevertheless, the results of these modelling studies suffer from large uncertainties, originating from biological processes in the ecosystems, such as seasonal variation, life cycling, climate change, meteorological conditions, changes in microbial populations, and interspecific relationships, mainly because the models do not consider the biological consequences for the microbiomes such as survival, vitality, and metabolic activity. Consequently, the models estimating the bioaerosol emission and the interactions of bioaerosols with atmospheric transportation require further improved quantification (Elbert *et al.*, 2007).

Previous studies mainly focused on ground-based measurements of bioaerosols, and observations at 50 m above ground level are very infrequent. Concretely, there were only few surveillances from aircraft (Fulton, 1966; Gruber *et al.*, 1998; Andreeva *et al.*, 2001; Kourtev *et al.*, 2011; Zweifel *et al.*, 2012; Ziemba *et al.*, 2016; Cáliz *et al.*, 2018) or high-altitude stations are available (Matthias-Maser *et al.*, 2000b; Hallar *et al.*, 2011; Gabey *et al.*, 2013; Crawford *et al.*, 2016; Du *et al.*, 2017). Moreover, global and regional model estimations regarding potential impacts of bioaerosols on atmospheric processes also suffer from the uncertainties caused by altitude, which requires better constrain by observations with diverse altitude coverage in the atmosphere. Notably, it is vital for reliable estimations of the impacts of bioaerosols on the atmospheric condition to quantify bioaerosol emissions and understand their sources and subsequent transformation processes.

2.2.3 Bioaerosol ecosystem interactions

Terrestrial ecosystems are considered major sources of bioaerosols in the atmosphere. For instance, pollen and spores can be produced and released from vascular plants and fungi to the atmosphere during reproduction, and the debris of plants and fungi can be emitted into ambient air during decay processes (Matthias-Maser *et al.*, 2000a; Jaenicke, 2005; Després *et al.*, 2012). Besides, the huge coverage surface area of cryptogamic (around 8.5×10^7 km^2), even larger than the summation of surface area of Africa and the Americas, leads to the large diversity and high abundance of microorganisms and bryophytes in various natural and human-impacted environments (Morris and Kinkel, 2002; Lindow and Brandl, 2003; Yadav *et al.*, 2005; Hantsch *et al.*, 2013). Moreover, the formation of bioaerosols and the impacts of bioaerosols on plants are affected by climate change and habitat conditions. It is well known that changes in land use cause severe effects on both the formation and dispersal of bioaerosols. In particular, the vegetation and cryptogamic covers of natural lands are often destroyed with conversion to human use (*e.g.*, agriculture and construction), which may enhance the frequency and damage of dust storms, because biological soil crusts contribute to drastically stabilizing the soil surface, largely reducing the erosive impacts of wind (Belnap and Gillette, 1998; Eldridge and Leys, 2003). For example, the dust load levels increased 5 times during the settlement in the western United States in the 19th century (Neff *et al.*, 2008). Also, the more frequent dust storm events recently occurring in China are considered to be caused by converting natural lands to human use (Hill *et al.*, 2014). Consequently, the bioaerosol emissions from terrestrial ecosystems may be significantly affected by surrounding environments and should be further explored in

regional and global studies of the interactions between atmosphere, biosphere, climate change, and anthropogenic impacts.

In contrast to terrestrial ecosystems, more knowledge gaps concerning the contributions of marine ecosystems to bioaerosols still remain, even though the ocean covers more than 70% of the surface of our planet (Seifried *et al.*, 2015; Xia *et al.*, 2015; Mayol *et al.*, 2017; Ma *et al.*, 2020b). On the one hand, the sources from terrestrial ecosystems and the long-distance transport of airborne microbes, such as plant fragments and human pathogens, can affect the bioaerosols over the oceans to a certain extent (Brown and Hovmøller, 2002; Cho and Hwang, 2011; Sharoni *et al.*, 2015). On the other hand, emissions from oceans are considered the major source of aerosol particles, which contributes to low cloud formation over the oceanic regions (Matthias-Maser *et al.*, 1999; Pósfai *et al.*, 2003; Aller *et al.*, 2005; Leck and Bigg, 2005; Amato *et al.*, 2007; Després *et al.*, 2012; DeLeon-Rodriguez *et al.*, 2013; Fahlgren *et al.*, 2015; McCluskey *et al.*, 2019). It was estimated that the bacterial concentrations in aquatic environments (*i.e.*, marine or freshwater) were around 10^6 cells/mL (Cho and Hwang, 2011; O'Dowd *et al.*, 2015), which can be emitted to ambient air through the burst of rising bubbles from deep layers to the sea-surface microlayer (Blanchard, 1975; Blanchard *et al.*, 1981; Aller *et al.*, 2005; Hultin *et al.*, 2011; Veron, 2015; Wilson *et al.*, 2015). Compared with the emissions of microbiomes from subsurface water to aerosols, the bubble-erupting leads to an enrichment of microbes in the ambient air (Aller *et al.*, 2005; Wilson *et al.*, 2015). According to the bacterial concentration in the aquatic ecosystems and related enrichment factors, global marine was estimated to emit roughly 2,000-10,000 Gg/a to bioaerosols (Burrows *et al.*, 2009a). Since most studies on airborne microbiomes were conducted in urban, rural, or background regions, the quite

small number of measurements concerning marine bioaerosols led to many uncertainties in this estimation.

Because only a few seawater bacteria (around 0.001 – 0.1%) are culturable (Amann *et al.*, 1995), thus, the high-throughput sequencing method, such as 16S rRNA gene pyrosequencing, were applied in recent studies to further investigate the bacterial communities of marine bioaerosol samples and better cover the whole marine microbial communities (DeLeon-Rodriguez *et al.*, 2013; Seifried *et al.*, 2015; Uetake *et al.*, 2020). Moreover, real-time quantitative polymerase chain reaction (qPCR) was also used to quantify selected health-relevant microorganisms, such as cyanobacteria and dinoflagellate species, in marine bioaerosols (Casabianca *et al.*, 2013; Lang-Yona *et al.*, 2014). Despite the crucial role of marine bioaerosols in the global atmosphere, there is still a knowledge gap of the interconnections of them with aquatic ecosystems and their source contributions, which limited our understanding of the overview of the global airborne microbial world.

2.3 Microorganisms in airborne particulates

Airborne microorganisms are widely distributed with metabolic activity in atmospheric environments, taking up an essential portion of airborne particles (Sattler *et al.*, 2001; Jaenicke, 2005). It has been documented that plenty of pathogens and allergic fungi were scattered in the air, influencing human health, wildlife, and vegetation directly or indirectly (Cao *et al.*, 2014; Barberán *et al.*, 2015). Moreover, the microorganisms in airborne particulates enable particles to be transported over a long distance through the airflow, thus further exerting their impacts on the Earth's ecosystems (Maki *et al.*, 2011). As a result, it

is essential to explore the structure and distribution of airborne microorganisms and to investigate the dominant impacting factors, especially on a global scale.

2.3.1 Particle size distribution of airborne microorganisms

The exposure to microorganisms in airborne particulates presents negative effects on public health (Fernstrom and Goldblatt, 2013), which were also influenced by the size of particles (Miller *et al.*, 1988; Cheng, 2003; Cho *et al.*, 2005; Sturm, 2012). For instance, the penetration depth of bioaerosol particles is highly determined by their size in the respiratory process, and the larger biological particles would be stuck in the nose or mouth, yet only smaller particles can enter deeply into the alveolar region of the lung (Clauß, 2015). In particular, inhalable bioaerosol particles, especially $PM_{2.5}$, show an appreciable impact on public health (Fröhlich-Nowoisky *et al.*, 2016; Smets *et al.*, 2016), thus, $PM_{2.5}$ is gradually recognized as an air pollutant and hazard by the public (*e.g.*, toxins, allergens, pathogens, and ARGs) (Cao *et al.*, 2014; Li *et al.*, 2018; Yue *et al.*, 2018; Zhou *et al.*, 2018).

It has been well documented by previous research that the structures of airborne microbial communities varied across aerosol particle size fractions (Polymenakou *et al.*, 2008; Franzetti *et al.*, 2011; Bertolini *et al.*, 2013; Bowers *et al.*, 2013; Lu *et al.*, 2018; Yan *et al.*, 2018; Stern *et al.*, 2021). According to the results of high-throughput sequencing, the most dominant microbiological component of bioaerosols is bacteria. In spite of the significantly higher diversity (*i.e.*, richness) of airborne bacteria in coarse particles than in smaller particles, their proportion in the total microbiomes was generally higher in $PM_{2.5}$ than in larger particles (Bowers *et al.*, 2013; Cao *et al.*, 2014; Stern *et al.*, 2021). Some researchers

proposed that the greater bacterial diversity in coarser fractions might be caused by the more abundant nutrients and more effective shelter from environmental stress (*e.g.*, ultraviolet radiation, high temperature, and droughts) in larger particles (Burrows *et al.*, 2009a). In addition, the majority of eukaryotes, such as fungi, algae, and plant debris, generally adhere to particles less than 10 μm in diameter (PM_{10}) and total suspended particulates (TSP).

2.3.2 Concentration, community structure, and biogeography of airborne bacteria

The concentrations of culturable bacteria in the atmosphere were first measured in 1883 (De Varigny, 1883); since then, more and more measurements of both culturable and unculturable airborne bacteria were conducted in diverse areas worldwide. The results indicated significant spatial and temporal variations of the airborne bacteria concentrations. Near-ground measurements have already been reviewed by some researchers concerning specific problems, such as the long-range transport of bioaerosols (Bovallius *et al.*, 1980; Petroselli *et al.*, 2021), temporal (annual and diurnal) variation (Lighthart, 1997; Gusareva *et al.*, 2019), the effects of dust storms or hazes on airborne bacterial abundance (Kellogg and Griffin, 2006; Barberán *et al.*, 2015; Gat *et al.*, 2017; Lu *et al.*, 2018), emission mechanisms and especially the impacts of meteorological factors (Jones and Harrison, 2004). The regional differentia can impact the concentrations of airborne bacteria by orders of magnitude in the near-surface measurements of ambient air, most of them ranging from 10^3 – 10^6 copies/m^3 of the 16S rRNA gene quantified by qPCR (Harrison *et al.*, 2005; Bowers *et al.*, 2009; Lee *et al.*, 2010; Cho and Hwang, 2011; Bowers *et al.*, 2012; Bertolini

et al., 2013; Murata and Zhang, 2014; Yamaguchi *et al.*, 2014; Barberán *et al.*, 2015; Tanaka *et al.*, 2015; Deng *et al.*, 2016a; Gat *et al.*, 2017; Genitsaris *et al.*, 2017; Gou *et al.*, 2017; Innocente *et al.*, 2017; Xu *et al.*, 2017; Zhen *et al.*, 2017; Gao *et al.*, 2017a; Šantl-Temkiv *et al.*, 2018; Xie *et al.*, 2018; Shen *et al.*, 2019; Tignat-Perrier *et al.*, 2019; Xu *et al.*, 2019), while a few of them ranging from 10^4 – 10^6 cells/m^3 identified by cell-based methods (*e.g.*, flow cytometry (FCM) and fluorescence microscopy) (Bowers *et al.*, 2011b; Murata and Zhang, 2016). Also, the measurements above the canopy in the megacity of cities (Du *et al.*, 2017) and even in the middle-to-upper troposphere (8-15 km altitude) were conducted to quantify the bacterial concentration in the atmosphere (DeLeon-Rodriguez *et al.*, 2013).

Although the diversity and profiles of the airborne bacterial community vary across the globe (Bowers *et al.*, 2011a; Bowers *et al.*, 2013; Gandolfi *et al.*, 2015; Zhang and Cao, 2015; Yee *et al.*, 2020), there are still some similarities and affinities of airborne bacterial communities. For instance, it has been confirmed that Actinobacteria, Firmicutes, Proteobacteria, and Bacteroidetes represent the most dominant phyla in bioaerosols with various particle sizes (Bertolini *et al.*, 2013; Xie *et al.*, 2019; Liang *et al.*, 2020a). Moreover, sequences assigned to the Actinobacteria (Bertolini *et al.*, 2013; Bowers *et al.*, 2013) and Actinomycetes (Cao *et al.*, 2014; Gandolfi *et al.*, 2015) were also prevalently detected in previous studies on diverse sampling sites, probably because these taxa mainly emanate from terrestrial ecosystems and show dominate roles in the whole planet bacterial community.

Seasonal variation and the typical diurnal cycling of airborne bacteria have generally been verified in certain sites (Andreeva *et al.*, 2001; Bertolini *et al.*, 2013; Bowers *et al.*, 2013; Gusareva *et al.*, 2019), and the airborne bacterial communities also showed differentia across the geographical locations (Després *et al.*, 2007; Bowers *et al.*, 2011a; Bowers *et al.*, 2011b; Deng *et al.*, 2016a; Tignat-Perrier *et al.*, 2019). The spatiotemporal patterns of airborne bacteria should be controlled by multiple impacting factors, containing meteorological conditions, air quality, contribution sources, and so on. For example, meteorological variables, such as air temperature (AT), relative humidity (RH), wind speed (WS), wind direction (WD), air pressure (AP), and ultraviolet flux, have generally been shown to have significant impacts on shaping the bacterial community by exerting selection pressure (Bertolini *et al.*, 2013; Gandolfi *et al.*, 2015). Moreover, some studies show the significant correlations of airborne bacteria with environmental variables, such as air pollutants, as these chemical components could be considered to be either nutritional supplies or toxic materials to microorganisms in airborne particles (Schulze, 1989; Gandolfi *et al.*, 2015). Due to the highly dynamic air ecosystem and the complex interactions among these potential impacting factors, it becomes fairly difficult for the comprehensive understanding of driving mechanisms and the quantification of the contributions of each factor to microbial community variations in the atmospheric environments, which requires more robust approaches like variation partitioning analysis (VPA), structure equation modelling (SEM), redundancy analysis (RDA), random forest analysis and so on (Zhen *et al.*, 2017; Romano *et al.*, 2020). As a result, a comprehensive understanding of the concentration and distribution patterns of airborne bacteria, as well as

the related mechanism, is needed on a global scale to further explore the invisible microbial world.

2.3.3 Source tracking and transportation of bacteria in the ambient air

It has been previously shown that variations in the abundance and profiles of airborne bacteria were significantly dependent on the changes in emission contributions from different sources (land covers), as most airborne bacteria are aerosolized from the surfaces, containing topsoil, aquatic surfaces, and aerial plant parts, and are removed from these surfaces by winds like leaf shaking or pounding surf (Jones and Harrison, 2004). For example, the total bacterial loadings in the air are commonly lower in the winter but increase in the following spring with the highest values in the summer, consistent with the seasonal variations in the growing status of vegetation, particularly in agricultural and forest areas (Bertolini *et al.*, 2013; Bowers *et al.*, 2013; Xie *et al.*, 2018).

Numerous studies have aimed to track the potential sources of airborne bacteria and further quantify their contributions, which revealed that topsoil, terrestrial plants, human and animal feces, and local or long-transport dust were the dominant terrestrial origins of airborne bacteria in near ground (Bowers *et al.*, 2011a; Bowers *et al.*, 2011b; Bowers *et al.*, 2013; Gat *et al.*, 2017). As for bacteria in marine air, the aquatic ecosystems, *i.e.*, seawater and freshwater, presented remarkable contributions to them (Cao *et al.*, 2014). In human-impacted areas, the anthropogenic sources, such as WWTPs (Bauer *et al.*, 2002a), hospitals (Gilbert *et al.*, 2010), landfills (Breza-Boruta, 2016), and farms (Bakutis *et al.*, 2004), were non-negligible to contribute to airborne bacteria.

To further quantify the relative contributions from different putative sources, new approaches have been developed on the basis of the sequencing results in recent studies. For example, some scientists assigned DNA sequences of the object samples with the counterparts of samples representing various sources based on the comparisons of indicative taxa identified in these samples, thus, resulting in the proportion of each putative source (Bowers *et al.*, 2013; Cao *et al.*, 2014). Moreover, an improved program, SourceTracker2, using a Bayesian approach to estimate the assignments of sink sequences to all source samples, has been recently applied in airborne microbial studies to create a joint distribution of those assignments (Knights *et al.*, 2011; Yang *et al.*, 2019; Uetake *et al.*, 2020; Wu *et al.*, 2022; Xie *et al.*, 2022).

Upon entering the air from various sources, bacteria can be transported upwards by airflow, remain in the ambient air for around a few days, and even transport over geographic barriers and long distances (Burrows *et al.*, 2009b; Fröhlich-Nowoisky *et al.*, 2016), which may have implications for human health, agriculture, cloud formation, and microbial biogeography (Burrows *et al.*, 2009a). Models of global circulation and air mass back trajectories have been widely applied in exploration of the long-range dispersal of airborne microorganisms, such as the source region of culturable microbes (Andreeva *et al.*, 2002) and biological ice nuclei (Pratt *et al.*, 2009).

2.3.4 Survival, viability, and health issues of airborne microbes

Air has long been recognized as a harsh environment for microorganisms compared with terrestrial and aquatic systems, however, the higher than expected airborne bacterial

viability in various environments (averagely 60%) was demonstrated by using fluorescent microscopy coupled with fluorescent staining (can both detect culturable and non-culturable bacterial cells without cultivations) in recent studies (Hara and Zhang, 2012; Murata and Zhang, 2014; Yuan *et al.*, 2017). Also, dead bacteria and even cell fragments can cause environmental changes like cloud development and hydrological cycles (Roszak and Colwell, 1987). The fact that most bacteria are viable in natural environments but with only a 1% probability of culture (Christner *et al.*, 2008) further emphasizes that the overall perspective of the airborne bacteria could only be excavated through genome sequencing and be considered as a necessary step towards mechanistic macroecology.

Due to the lack of nutrients, extremely dry conditions, ultraviolet radiation, and various gaseous taxon compounds, the atmosphere is a quite severe environment for the survival of most microbiomes. Nevertheless, previous studies revealed the viability of some airborne microbes in ambient air (Fang *et al.*, 2007; Gao *et al.*, 2015) (Fang et al., 2007; Gao et al., 2015). For example, airborne bacteria could prevent the damage of solar radiation to themselves by pigmentation (Imshenetsky *et al.*, 1978). Also, microbes can survive in the harsh atmospheric environment by morphologically transforming into spores to present the high resistance to environmental stress, as there will be a large reduction in metabolic rates with the cessation of cell division in this process (Horneck *et al.*, 1994; Bär *et al.*, 2002). Notably, once the spores come back to suitable environments, they can restore their activity and live in the new conditions, which was documented in a previous study where microbes from Asia were detected in the free troposphere in North America and were still culturable after a cross-continental storm transport (Smith *et al.*, 2012).

In recent studies, most of the microbes in the bioaerosols were identified with low or no pathogenicity to humans; however, the relative abundance of the allergens and pathogens has been shown to increase with more severe air pollution, like smog events (Cao *et al.*, 2014; Sun *et al.*, 2020) (Cao et al., 2014; Sun et al., 2020). These discoveries further suggested a significant increase of the anthropogenic impacts on (opportunistic) pathogens in the urban atmospheric environments.

2.4 Methodology for bioaerosol characterization

Bioaerosols are airborne particles that originate from living organisms, including bacteria, viruses, fungi, and other microorganisms (Fröhlich-Nowoisky *et al.*, 2016). These particles can have a significant impact on human health, as they can cause respiratory diseases, allergies, and other health problems (Brown and Hovmøller, 2002; Brodie *et al.*, 2007). Therefore, it is important to develop effective methods for characterizing bioaerosols, which can help to identify the types of microorganisms present and their potential health effects. Several methods have been commonly used for characterizing bioaerosols, including culture-dependent techniques, DNA sequencing, and flow cytometry (Franchitti *et al.*, 2020).

Culturing is a traditional method for characterizing bioaerosols, which involves the collection of bioaerosols onto a growth medium, where the microorganisms can grow and form visible colonies that can be counted and identified (Crook *et al.*, 1991; Mériaux *et al.*, 2006; Dutil *et al.*, 2008). These methods are relatively simple and inexpensive, and they can provide valuable information about the types and concentrations of microorganisms

present in a sample, as well as their viability and growth characteristics (Jalili *et al.*, 2021; Li *et al.*, 2021b). However, culturing has several limitations, including its inability to detect non-viable microorganisms, the requirement for specific growth conditions for different types of microorganisms, and the potential for selective growth of certain types of microorganisms (Amann *et al.*, 1995). Additionally, culturing can be time-consuming and labor-intensive, and may not provide a comprehensive picture of the microbial community present in a sample (Peccia and Hernandez, 2006).

In recent years, studies using molecular methods for the detection of airborne microorganisms have revealed that non-culturable microorganisms are considerable constituents of bioaerosols (Blais-Lecours *et al.*, 2015). DNA sequencing is a powerful tool for detecting a wider range of microorganisms including non-culturable or slow-growing microorganisms, as well as viruses and fungi, and can provide information on the genetic diversity and functional potential of the microorganisms present in the air, allowing for a more comprehensive understanding of their potential impacts. PCR amplification followed by Sanger sequencing (Boreson *et al.*, 2004; Fierer *et al.*, 2008) or high-throughput sequencing (Andersson *et al.*, 2008; Hamady *et al.*, 2008) of the amplified DNA is a common approach to DNA sequencing for bioaerosol characterization, allowing for targeted sequencing of specific genes or regions of the genome, such as the 16S rRNA gene for bacterial identification or the ITS region for fungal identification (Bowers *et al.*, 2011a; Bowers *et al.*, 2013; DeLeon-Rodriguez *et al.*, 2013; Cao *et al.*, 2014; Womack *et al.*, 2015). Furthermore, metagenomic sequencing is another approach that involves sequencing the entire DNA content of the air sample without any prior amplification or targeting (Fouladi *et al.*, 2020; Drautz-Moses *et al.*, 2022; Wu *et al.*, 2022; Xie *et al.*, 2022).

This approach allows for a more comprehensive analysis of the microorganisms present in the air, but it can also be more challenging due to the larger amount of data generated (Shamarina *et al.*, 2017). Despite its advantages, DNA sequencing also has some limitations. It can be more complex and expensive than culture-based methods (Dubuis *et al.*, 2017), and it may require more advanced expertise in bioinformatics and data analysis. Additionally, DNA sequencing can be affected by contamination, which can lead to false positive results (Schmieder and Edwards, 2011; Glassing *et al.*, 2016). As such, it is important to carefully consider the sampling and sequencing approach in the context of the research question and available resources.

Flow cytometry is a method for characterizing bioaerosols that involves analyzing the physical and chemical properties of individual particles (Chen and Li, 2005). Flow cytometry can provide information about the size, shape, and fluorescence properties of particles, which can be used to distinguish between different types of microorganisms (Davey, 2003). Flow cytometry can also provide information about the viability and metabolic activity of microorganisms, as well as their sensitivity to environmental stressors (Negron *et al.*, 2020; Yu *et al.*, 2021). However, flow cytometry has several limitations, including its inability to identify specific microorganisms and the potential for interference from non-biological particles (Veal *et al.*, 2000). Additionally, flow cytometry requires specialized equipment and expertise, and may not be suitable for analyzing complex or heterogeneous samples (Vives-Rego *et al.*, 2000).

In addition to these methods, other techniques are also available for characterizing bioaerosols, including immunological methods and microscopy-based methods (Cruz and Buttner, 2016). Immunological methods involve detecting specific antibodies or antigens

to identify the types of microorganisms present (Lindsley *et al.*, 2006). Microscopy-based methods involve visualizing individual particles using light or electron microscopy to identify their size, shape, and other physical characteristics (Gong *et al.*, 2018). Each of these methods has its advantages and limitations, and the choice of method will depend on the specific research question and the available resources.

2.5 ARGs as an emerging environmental pollutant

Antibiotics have been regarded as one of the most effective classes of antimicrobial drugs for treating human and animal diseases. Nevertheless, the excessive and improper use of antibiotics may cause the emergence and development of ARGs and antibiotic-resistant bacteria (ARB), facilitating the dissemination of them in both clinical and nonclinical environments. Moreover, HGT and MGEs like plasmids, transposons, and integrons, are also conducive to accelerating the dissemination of ARGs among various environments. The consumption of antibiotics was used for not only human therapy but also the agriculture and breeding industry, leading to a widespread of ARGs and ARB in WWTPs, landfill, agricultural soil, farms, and hospital waste, so they are gradually treated as environmental pollutants, which can contaminate the natural environment.

2.5.1 Origin, distribution, and development of ARGs in the environment

ARGs can be loosely defined as genetic elements encoding resistance to antibiotics. It has been shown that the concentration of ARGs was much larger in modern environments than

that has been previously discovered (D'Costa *et al.*, 2006; Dantas *et al.*, 2008; Sommer *et al.*, 2009). Moreover, the various homologs of recognized ARGs have been confirmed to broadly spread throughout the environments in recent metagenomic research, suggesting that there might be a richer and longer natural history of ARGs (Wright, 2007). In fact, antibiotic resistance inherently exists in natural environments (Chen *et al.*, 2016) and has existed since geological times (Hall and Barlow, 2004; Baltz, 2005), which indicated that ARGs should be very ancient (D'Costa *et al.*, 2011). For example, a large variety of ARGs are carried by microbes that inhabit various niches without anthropogenic impacts, such as deep oceans (Toth *et al.*, 2010; Chen *et al.*, 2013), isolated caves (Bhullar *et al.*, 2012), and deep terrestrial subsurface (Brown and Balkwill, 2009). Many lines of evidence indicate that antibiotics could be secreted by microbes as a competitive mode between them (Newman and Cragg, 2016), and the microbial synthetic pathways of antibiotics have evolved over millions of years, which is much earlier than the use of human-made antibiotics in medicine (Baltz, 2008).

Intrinsic ARGs are a characteristic feature of certain environmental microorganisms, which provide them with a defensive mechanism against external aggression and the ability to produce metabolites that inhibit the growth of other microbes in their immediate environment (Martinez *et al.*, 2009). Moreover, some soil microbes can acquire nutrients by breaking down surrounding antibiotics via their intrinsic ARGs (Allen *et al.*, 2010). Such intrinsic ARGs are ubiquitously distributed on the Earth as functional genes in different ecosystems. For other microbes, their ARGs can be acquired by spontaneous HGT or mutation (Martinez and Baquero, 2000).

However, human activities can also drive the selection of genes from environmental or cellular sources, which can subsequently confer antibiotic resistance (Czekalski *et al.*, 2015). Antibiotics were used in healthcare and illness treatment to kill or inhibit the growth of pathogenic microorganisms, which has been recognized as the largest medical revolution in the 20th century (Quinn, 2013). After penicillin was first isolated and used in the clinic in 1929 (Fleming, 1929), increasing numbers of antibiotics were discovered successively and introduced as a medicine in microbial infection treatments of humans, animals, and agriculture (Årdal *et al.*, 2020). The remaining antibiotics in the environment can drive gene selection, which can subsequently confer antibiotic resistance, thereby, ARBs and ARGs gradually emerge in the environment. ARGs have been detected in various environments, including natural (aquatic (Zhang *et al.*, 2020b), soil (Zhu *et al.*, 2017), and atmospheric systems (Xie *et al.*, 2018)), and engineered (WWTPs (Machado *et al.*, 2023), landfill (Wu *et al.*, 2017), traffic systems (Kang *et al.*, 2018; Danko *et al.*, 2021), and clinical habitats (Chng *et al.*, 2020))). The increasing use of antibiotics enhanced the development and dissemination of ARGs in various environments, which would be considered as a global threat to human, animal, and environmental health (Zhang *et al.*, 2022).

ARGs occur as emerging environmental contaminants, because bacteria develop resistance against the antibiotics used in human and animal therapy, agriculture, and husbandry (Riesenfeld *et al.*, 2004; Baquero *et al.*, 2008; Zhang *et al.*, 2009a; Allen *et al.*, 2010; Chen *et al.*, 2013). Recently, researchers have paid more attention to ARG-hosting bacteria in the environments, as ARGs are widely distributed in the environments (Zhang *et al.*,

2009a), and most of the pathogens carrying ARGs are derived from environmental bacteria (Martínez, 2008).

2.5.2 Crucial role of air in the dissemination of ARGs

ARGs participate in atmospheric circulation through evaporation processes and wind (Tripathi and Cytryn, 2017) and could be disseminated worldwide with air pollutant transportation (Kellogg and Griffin, 2006; Li and Osada, 2007). After a long-term dissemination, ARGs could be completely mixed in the atmosphere. ARGs and ARB adhered to PMs in the air have been detected in hospitals or places close to animal feeding operations, where amounts of antibiotics were used (Gibbs *et al.*, 2006; Gilbert *et al.*, 2010; Fan *et al.*, 2014; McEachran *et al.*, 2015). Besides, scientists have also isolated ARB in PM samples collected in outdoor environments (Gandolfi *et al.*, 2011; Ling *et al.*, 2013). Consequently, atmosphere could be regarded as a gene bank of ARGs, especially close to key emission sources, like hospitals (Wu *et al.*, 2022) and WWTPs (Xie *et al.*, 2022). ARGs and ARB could be detected several kilometers downwind of the emission sources, indicating that they could be transmitted via airflows and were probably viable during the transport process (McEachran *et al.*, 2015). Also, more evidence showed the important role of air in ARG presence and transmission. For instance, multidrug-resistant airborne pathogens were frequently detected in hospital-related environments (Dijkshoorn *et al.*, 2007), which might cause the infection of patients living in hospitals and even pose a severe health risk to inhabitants in the surrounding areas via air movements (Lis *et al.*, 2009).

In recent years, it has been a topic of increasing interest to detect ARGs and MGE in PM and quantify their abundance in various human-related areas (Ling *et al.*, 2013; Pal *et al.*, 2016; Hu *et al.*, 2018), such as hospitals (Wu *et al.*, 2022), livestock farming (McEachran *et al.*, 2015; Gao *et al.*, 2017b), WWTPs (Li *et al.*, 2016a), and urban environments (Echeverria-Palencia *et al.*, 2017; Xie *et al.*, 2018). For example, ARGs conferring resistance to tetracycline, *tet*X, and *tet*W, were detected in inhabited indoor environments with estimated concentrations of 100–200 copies/m^3 and 100–400 copies/m^3, respectively in Colorado, USA (Ling *et al.*, 2013). Pal et al. used known gene databases to estimate the diversity of ARGs across the samples, revealing the higher richness in Beijing smog (64.4 different ARG types) than in other external environments, like pharmaceutically polluted environments (38.9), wastewater/sludge (19.4), animals (11.8), humans (1.0–16.6), and other terrestrial sources (1.6 – 3.3) (Pal *et al.*, 2016). In a comprehensive investigation of airborne ARGs inside a typical WWTP in Beijing, *sul*2 and *intI*1 (the class 1 integrase gene) were detected (Li *et al.*, 2016a). Also, airborne ARGs, such as bla_{SHA} and *sul*1 were discovered at up to 10^2 copies/m^3 and 10^3 copies/m^3 in an urban park in Califonia, USA (Echeverria-Palencia *et al.*, 2017). The above observations suggested that atmospheric transmission may play a crucial role in spread of ARGs across diverse environments. Because the fine PMs can be suspended in the air for a long time, the long-distance dispersal of ARGs and ARB attached to PMs may happen in a regional scale and even a global scale, which can be promoted by strong winds or storms. Shortly after being transported to a new destination, ARB and the local microbes that receive the newcome ARGs may occupy new niches, posing a threat to the balance of the local ecosystem (Griffin, 2007).

Moreover, recent studies revealed that the abundance of airborne ARGs showed obvious variations over time and space (Li *et al.*, 2018; Xie *et al.*, 2018; Liang *et al.*, 2020a). However, research into the spatiotemporal differentiation of airborne ARGs, particularly in terms of their absolute abundance and exposure risks, remains still unknown on a global scale. It has been verified by regional studies that the spatial and temporal variations of airborne ARGs were influenced by their hosting bacteria, environmental changes, and human activities (Liang *et al.*, 2020a). In particular, it has been shown that the abundance and richness of airborne ARGs were positively correlated with PM pollution by analyzing the differentia in hazy and non-hazy days in same places, highlighting the increasing exposure risk of AMR in pollutant air (Sun *et al.*, 2020; Zhao *et al.*, 2020). In summary, spatiotemporally dynamic variations of aerosol-associated ARB and ARGs require a comprehensive study on long-term human exposures via inhalation on a large scale, particularly for a global scope.

2.5.3 Relevance between ARGs and pathogens

Some researchers named the small but important group, *i.e.*, *Enterococcus faecium, S. aureus, Klebsiella pneumonia, Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Enterobacter* spp., as the ESKAPE bugs, because not only do they account for the majority of hospital-acquired infections, but they also serve as prime examples of pathogenesis, transmission, and resistance (Rice, 2008; Mulani *et al.*, 2019). The emergence of antibiotic resistance among gram-negative pathogens, which cause diseases in humans and animals, has been strongly correlated with the consumption of antibiotics over the past fifty years.

The correlation between the β-lactam class of antibiotics and their inactivating enzymes, the β-lactamases, is evident, as highlighted by previous studies (Bush and Jacoby, 2010; Poirel *et al.*, 2010). *Pseudomonas aeruginosa*, a pathogen commonly acquired in hospitals, poses a significant threat to patients with cystic fibrosis (Horrevorts *et al.*, 1990; Fajardo *et al.*, 2008), with antibiotic resistance mechanisms evolving alongside the introduction of new antibiotic derivatives, leading to reduced effectiveness of potent treatments like β-lactams and aminoglycosides. The most notorious superbug currently is the gram-positive organism S. aureus, which has a close association with humans and is carried as a nasal commensal in 30% of the population (Lindsay and Holden, 2004). Despite the development of methicillin, the first designer antibiotic for resistance in 1959, as a reliable defence against penicillinases, the emergence of methicillin-resistant *Staphylococcus aureus* (MRSA) within three years led to the appearance of other multi-antibiotic-resistant variants (Enright *et al.*, 2002). Resistance mechanisms are widespread among commensals (Marshall *et al.*, 2009) and pathogens, and they can be disseminated by various gene transfer mechanisms.

Nowadays, a large body of research has demonstrated the quantity and species of common ARGs; however, there is still a large gap in ARB understanding. Some researchers predicted a recent database from available genome sequences, which lists the existence of more than 20,000 potential resistance genes of nearly 400 different types (Liu and Pop, 2009; Davies and Davies, 2010); however, it is still unknown how many would have been found using a wider range of expression systems and hosts. So far, studies on the airborne dissemination of Angstroms of the ambient environment to human bodies are very limited (Dickson and Huffnagle, 2015), especially with a lack of understanding of whether

pathogenic hosts in airborne microbial communities and antibiotic resistance transfer in the clinical context. We have very little or even no evidence that any of the putative resistance genes identified in these environmental studies have been mobilized into pathogenic bacteria and expressed as resistance phenotypes. However, these gaps are crucial to controlling the antibiotic-resistant pathogens and superbugs, because ARGs take efforts only when they host pathogens and transcribe them into proteins.

As an active embedded MGE, *intI*1 has been found in a wide variety of pathogenic and nonpathogenic bacteria and is closely linked to genes conferring resistance to antibiotics (Gillings *et al.*, 2015), disinfectants, and heavy metals, leading to speculation that *intI*1 could serve as a proxy for anthropogenic pollution including ARGs (Griffin *et al.*, 2018). Nevertheless, a recent study found that the statistical correlation between ARGs and the MGE, *intI*1, weakened from rural to urban and industrial sites, indicating the heterogeneity of airborne dissemination of ARGs (Xie *et al.*, 2018). Improved knowledge of clinical elevators and hosts (*e.g.*, pathogens) shared between ambient air and human airways is critical to discern and address the key health implications (Dickson and Huffnagle, 2015). It has been known for some time that bacterial strains resistant to antibiotics can be isolated by planting environmental bacteria on antibiotic-containing media in the laboratory.

Chapter 3 Methodology

This chapter provides sample collection, global dataset generation, and detailed descriptions of the biological and statistical analyses applied in this study.

3.1 Air sample collection

In this study, a total of 803 air subsamples in the regions covering human-impacted and background areas were collected for an annual cycle in Asia, and the geographic locations of the sampling sites are shown in Figure **3-1**.

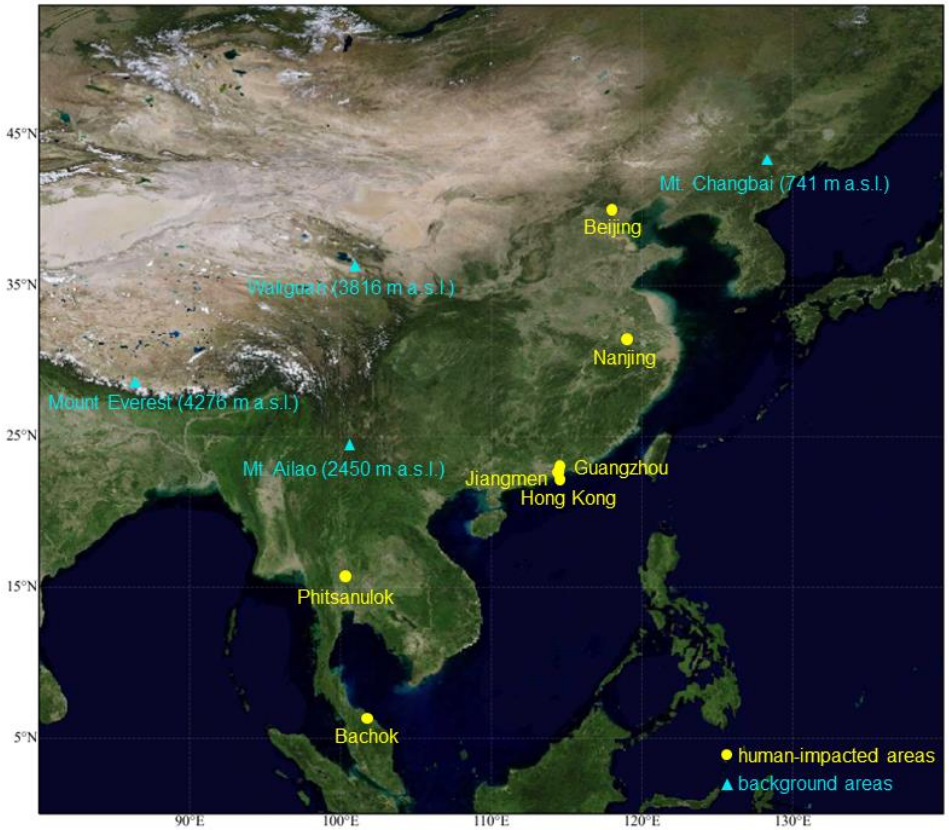


Figure 3-1 Geographic location of sampling sites. The yellow circles represent the human-impacted areas, and the blue triangles represent the background areas. The satellite map was obtained from National Platform for Common Geospatial Information Services (https://www.tianditu.gov.cn/)

3.1.1 Description of sampling strategy

In this study, air sample collection was first conducted at multiple sites with different land use impacts in selected Chinese typical cities, Beijing, Nanjing, and Guangzhou, to investigate the regional differences in air pollution from biological perspectives. Hong Kong, one of the most densely populated cities in the world and an important transition between terrestrial and offshore areas, was also included in the study. Compared with the extensive study on the mechanisms driving chemical pollution, the mechanistic theories on airborne microbiomes and anthropogenic impacts on the biological fractions have been less explored. To address this gap, air samples in corresponding background areas (*i.e.*, areas that are far less impacted by humans, such as the studied sites in remote mountains, offshore environments, and the polar regions) were collected simultaneously in Tibet and at Waliguan, Mt. Changbai, and Mt. Ailao. Considering the climate factors influencing the atmospheric microbe, the sampling work was conducted in temperate and subtropical zones, as well as tropical regions, including Phitsanulok in Thailand and Bachok in Malaysia.

Table **3-1** shows the sampling frequency and number of samples collected at each site, along with detailed descriptions of the surrounding environment, geographic information,

and demographic, traffic, and economic characteristics of the districts in which the sampling sites are located.

Table 3-1 Sampling frequency and number of air subsamples and description of sampling sites in this study.

Location	District (Abbreviated site name)	Sampling period	Sampling Frequency	Sample size	Sampling site	Site description
Beijing (northern China)	Haidian, Beijing (PKU)	Apr 2016 - May 2017	weekly	61	The campus of Peking university	Rooftop sampling on an academic building at the main part of the campus, 15 m above ground level; around 200 m away from a major road with heavy traffic; approximately 500 m away from the fourth-ring road.
	Haidian, Beijing (IAP)	Dec 2016 - Nov 2017	weekly	44	Institute of Atmospheric Physics, Chinese Academy of Sciences	Rooftop sampling on a two-floor building; located between the 3rd and 4th ring roads of the city; surrounded by heavy traffic, buildings ranging from 30 to 60 m in height, and a public park; approximately 1 km away from a highway; covered with some vegetation.
	Huairou, Beijing (HR)	Aug 2018- May 2019	weekly	33	Huairou campus of the University of Chinese Academy of Sciences	Rooftop sampling on the first teaching building (30 m above-ground) in a peri-urban village; located near a lake, mountains with vegetation, and a national road; with no apparent sources of industrial manufacturing pollution.
Nanjing (eastern China)	Pukou, Nanjing (PK)	Mar 2016- May 2017	Every 7-10 days	46	Nanjing University of Information Science and Technology	Rooftop sampling on a twelve-floor building; near chemical and metallurgical industries; surrounded by petrochemical plants, steel plants, and highways.
	Xuanwu, Nanjing (XW)	Mar 2016 - May 2017	Every 7-10 days	48	Institute of Soil Science, Chinese Academy of Sciences	Rooftop sampling on a five-floor building; downtown area surrounded by schools, parks, and residential and commercial buildings, with heavy traffic.
	Lishui, Nanjing (LS)	Apr 2016 - May 2017	Around every month	18	A botanical garden	Sampling site on the ground; distant from main roads and industrial pollution sources, with extensive vegetation cover; with low population density.

Location	District (Abbreviated site name)	Sampling period	Sampling Frequency	Sample size	Sampling site	Site description
Guangzhou (southern China)	Tianhe, Guangzhou (TH)	Mar 2016 - May 2017	Around every 3 days	122	Guangzhou Institute of Geochemistry, Chinese Academy of Sciences	Rooftop Sampling on a five-floor building; near two expressways; surrounded by schools and residential buildings.
	Conghua, Guangzhou (CH)	Mar 2016 - May 2017	Around weekly	52	Tianhu Park	Ground-level sampling site located on a hill; covered by massive vegetation; situated in a recreation area around 60 km away from Guangzhou downtown areas.
Jiangmen (southern China)	Heshan, Jiangmen (HS)	Mar 2016 - May 2017	weekly	65	Guangdong Atmospheric Monitoring Supersite of China	Rooftop sampling on a four-floor building; located on a hill covered by vegetation; surrounded by farmlands and country roads; with minimal industrial and traffic pollution from outer space.
Hong Kong	Hung Hom, Hong Kong (HK)	Jan 2016 – Nov 2016	Around every month	16	The campus of the Hong Kong Polytechnic University	Sampling site in the hanging garden on the 11th floor of a twelve-story building; located in the largely urbanized and densely populated downtown center with heavy traffic pollution; surrounded by commercial and residential buildings.
	Hok Tsui, Hong Kong (HT)	Jan 2016 – Dec 2016	Around weekly	37	Hok Tsui Background Air Monitoring Station	Ground-level sampling located on a hill; situated at the southern tip of Hong Kong; with the least human-related pollution and environmental impact; approximately 15 km away from large anthropogenic impacts and heavy traffic in the urban center (Victoria Harbour)
Phitsanulok, Thailand	Mueang Phitsanulok District, Phitsanulok (TL)	Dec 2015 – Nov 2016	Around weekly	40	The campus of IOES University Malaya	Sampling site at the top of an atmospheric observation tower (18 m height); located on the coastline of the South China Sea within 100 m of the shore; near to the main road with traffic emissions.

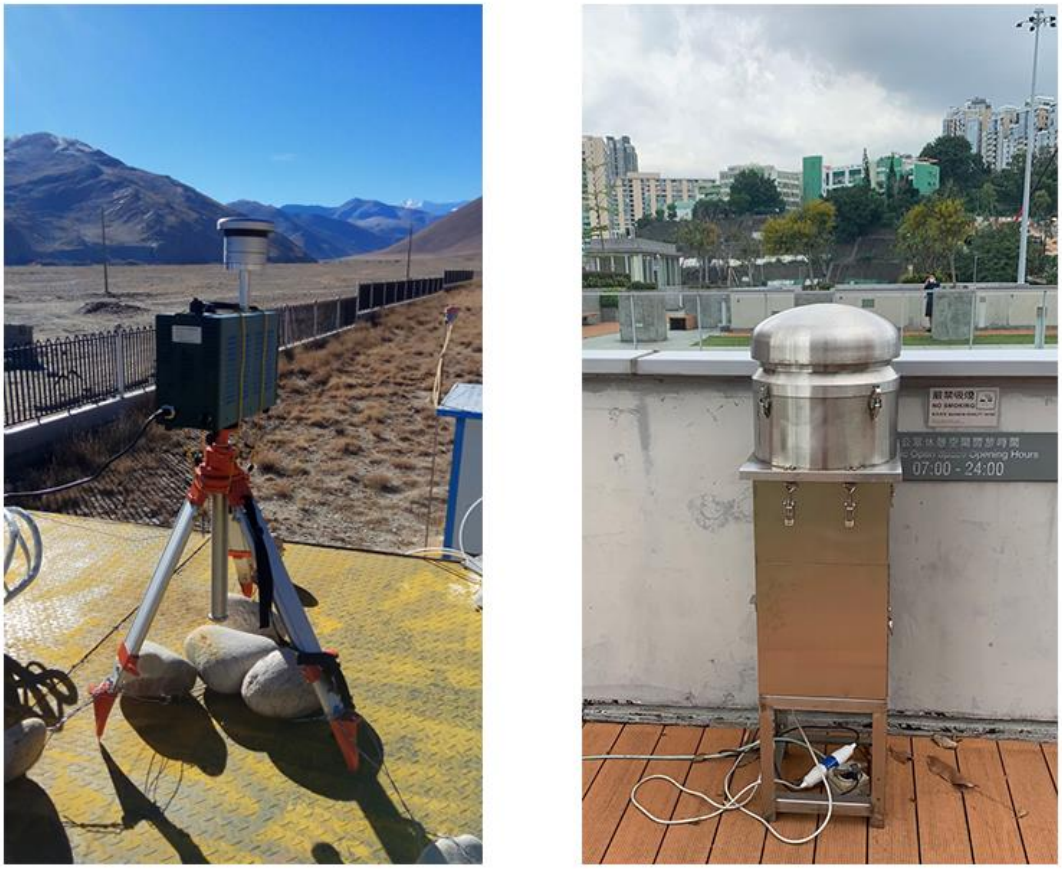
Location	District (Abbreviated site name)	Sampling period	Sampling Frequency	Sample size	Sampling site	Site description
Bachok, Malaysia	Bachok, Kelantan (ML)	Nov 2017 – Feb 2019	Daily or weekly	122	The campus of Naresuan University	Sampling site on the ground near a college campus; surrounded by a lush green landscape and flowing rivers; situated near a road and residential buildings.
Tibet	Mount Everest (ME)	Oct 2014 – Jun 2015	Around weekly	22	Qomolangma Station for Atmospheric and Environmental Observation and Research	Sampling site on the Tibetan Plateau at a high altitude; far away from anthropogenic emission sources, industrial utilities, and traffic pollution.
Waliguan	Gonghe County, Qinghai (WL)	Sep 2013 – Jul 2014	Around weekly	23	the Waliguan Baseline Observatory	Sampling site at the summit of Waliguan; located at the northeast edge of the Tibetan Plateau; naturally preserved and featured arid/semi-arid lands and grasslands; isolated from major industrial sources and populated centers.
Changbai Mountain	Antu County, Jilin (CB)	Apr 2014 – May 2015	Around weekly	36	Ailao Mountain National Nature Reserve	Sampling site on the hilly terrain; near the boundary between China and North Korea; dominantly covered by temperate broadleaf and Korean pine mixed forests; without large industrial emissions nearby.
Ailao Mountain	Fengqing County, Yunnan (AL)	Nov 2013 – Aug 2014	Around weekly	17	Changbai Mountain National Nature Reserve	Sampling site at the summit of the northern edge of Mt. Ailao; dominantly covered by evergreen broadleaf primary forests; without large anthropogenic emission sources in the surrounding areas.

3.1.2 Sampling method and instrument

For accurate air quality measurements, a filter was placed in the air sampler at each sampling site but not operating during the entire sampling process, which could be served as a blank control. Quartz microfiber filters (8 × 10 in, PALL, United States) were pre-baked at 500 ^{o}C for 5 h to eliminate any contamination from carbonaceous material. The filters were weighed before and after sampling separately, with a sensitivity of ± 0.0001 g. Prior to each weighing, the filters were required to equilibrate at 25 ^{o}C with 40-50% RH for at least 24 h.

In urban areas, including Beijing, Nanjing, Guangzhou, and Hong Kong, the $PM_{2.5}$ samples were collected using high-volume (1000 L/min) samplers (TH-1000C II, Wuhan Tianhong Instruments Co., Ltd.) for 24 h (Figure **3-2**). The $PM_{2.5}$ samples at Mt. Changbai, Waliguan, and Mt. Ailao, were collected onto preheated quartz filters using high-volume $PM_{2.5}$ samplers. The sampling durations at Waliguan and Mt. Changbai were 96 h, and 48−72 h, respectively, at a volumetric flow rate of 1000L/min. Besides, air samples were collected at Mt. Ailao for 168 h at a volumetric flow rate of 300L/min. The $PM_{2.5}$ samples in Tibet were collected using the Ambient Air Sampling Instrument (Lao Ying 2034, LAOYING Institute, China) at a flow rate of 100 L/min for 23.5 h (Figure **3-2**). TSP in Thailand and Malaysia were also collected on quartz microfiber filters using a high-volume sampler, with a sampling duration of 24 h and a flow rate of 300 L/min. A filter placed without operation was served as blank control, which contribution was removed for more precise quantitative analysis. Bacterial concentration in the air was calculated using a standard

volume. All filter subsamples were combined into 76 seasonal samples and stored at -20 ℃ before further analysis (No. 1-76 in Appendix 1).



Ambient air sampling instrument (Lao Ying 2034, LAOYING Institute, China), Tibet

High-volume samplers (TH-1000C II, Wuhan Tianhong Instruments Co., Ltd.), Hong Kong

Figure 3-2 Air sampling instruments applied in this study.

3.2 DNA extraction and sequencing

This section presents a comprehensive description of the molecular biological analyses employed in this study, containing the pretreatment of air samples, DNA extraction, quantification of total bacterial loads, library generation, 16S rRNA sequencing, and metagenomic sequencing (Figure **3-3**).

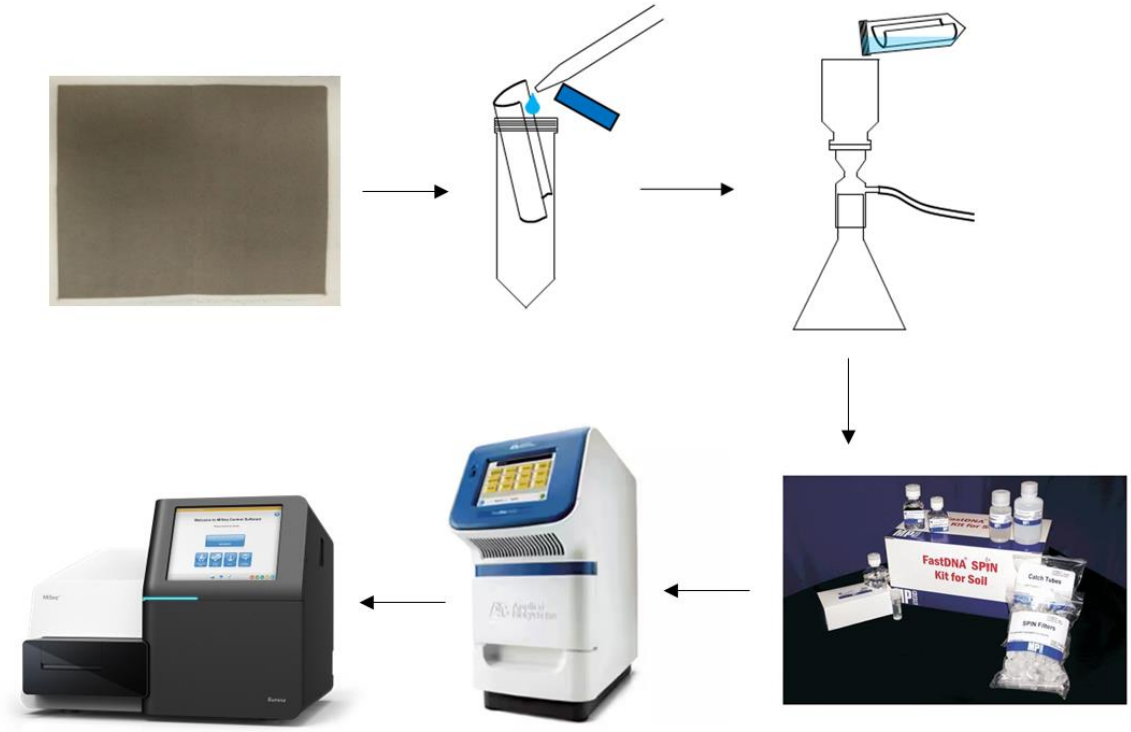


Figure 3-3 Flow chart of molecular biological analyses.

3.2.1 DNA extraction

In order to increase the amount of DNA obtained, field subsamples obtained from the same site and season were combined. Each filter sample was then cut into fragments approximately 8 cm × 10 cm in size and subjected to ultrasonic extraction with 1× phosphate-buffered saline in 50 mL centrifuge tubes. All equipment used in the extraction

process was sterilized, and the 1× phosphate-buffered saline was pretreated at 120 ºC for 20 min. After extraction, the resulting solution was filtered using a 0.2 μm PES membrane disc filter (47 mm, Pall) to concentrate the airborne microbiome, and the filter was then used for DNA extraction without delay.

The standard protocol for isolating DNA using the FastDNA SPIN Kit for Soil (MP Biomedicals) was followed for the remaining steps, except for the column purification step. The latter was replaced with magnetic bead purification (Agencourt AMPure XP, Beckman, CA, U.S.) to enhance the yield. All procedures were carried out on a clean bench. Following extraction, all DNA solution samples were kept at −80 °C until needed for qPCR quantification, 16S rRNA gene amplicon sequencing, and metagenomic sequencing.

3.2.2 Real-Time qPCR Quantification of Targeted Genes

The total bacterial loading was approximated by the concentration of 16S rRNA gene copies in the air. The 16S rRNA gene was amplified on a StepOnePlus Real-Time qPCR System (Applied Biosystems) with the following primer sequences: 5'-TCCTACGGGAGGCAGCAGT-3' as the forward primer and 5'-GGACTACCAGGGTATCTAATCCTGTT-3' as the reverse primer. To determine the absolute number of 16S rRNA gene copies, a seven-point standard curve (including a blank standard) in a 10-fold serial dilution was run with samples. Samples, standards, and blanks were analyzed in triplicate with an application efficiency between 90% - 110%. The specificity of the amplicons was confirmed through melt curve analysis. The 20-μL qPCR

reaction mixture was composed of 10 µL of Power SYBR™ Green PCR Master Mix (Life Technologies, CA, USA), 1 µL of template DNA, 0.5 µL of each primer (100 nM), and RNAse-free water to complete the final 20 µL volume. The 16S rRNA gene was amplified according to the following protocol: an initial step at 95 °C for 10 min for enzyme activation, then 40 cycles of 10 s at 95°C, and 1 min at 60°C for hybridizations and elongations. The amplicon length was around 400-500 bp. In addition, to reduce the variations in the 16S rRNA gene copies caused by particle size, the mean ratios of bacterial loadings in $PM_{2.5}$ with PM_{10} (1: 1.56 ± 0.74) and with TSP (1: 7.44 ± 3.86) collected at the same sites during the same sampling period were used to modify the data (Figure **3-4**).

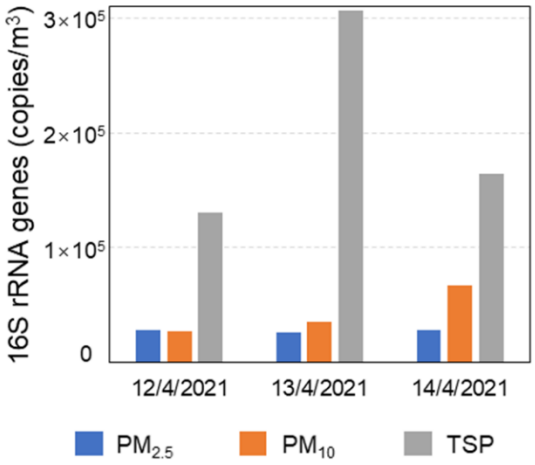


Figure 3-4 Effect of particulate matter size on total airborne bacterial biomass. Comparison of airborne bacterial biomass in $PM_{2.5}$, PM_{10}, and TSP collected in the hanging garden on the 11th Floor of a twelve-story building at the Hong Kong Polytechnic University (22.31N, 114.18E) in three consecutive days.

3.2.3 Library generation and DNA sequencing

3.2.3.1 16S rRNA gene amplicon sequencing

The 16S rRNA gene is widely used for bacterial identification and classification. Amplification, barcoding, pooling, and library preparation for sequencing the V3-V4 hypervariable region of the 16S rRNA gene followed the Illumina protocol (Caporaso *et al.*, 2012). KAPA HiFi HotStart ReadyMix (Kapa Biosystems) was used to amplify the 16S rRNA gene with degenerate PCR primers, 341F (5'-ACTCCTACGGGAGGCAGCAG-3'), and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Takahashi *et al.*, 2014), both of which were tagged with an Illumina adapter, pad, and linker sequences. PCR enrichment was performed in a 50 μL mixture containing a 30-ng template, a fusion PCR primer, and a PCR master mix. Thermal cycling consisted of an initial denaturation at 94 °C for 3 min, followed by 30 cycles of 94 °C for 30 s, annealing at 56 °C for 45 s, and elongation at 72 °C for 45 s, with a final extension for 10 min at 72 °C. The PCR products were purified with AMPure XP beads and eluted in an Elution buffer. Libraries were qualified by the Agilent 2100 bioanalyzer Agilent (USA). Sequencing was conducted on an Illumina MiSeq platform, generating 2 × 300 bp paired-end reads. The DNA sequence data were deposited in the Sequence Read Archive (SRA) database of the NCBI (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA757592).

3.2.3.2 Metagenomic sequencing

The concentration of genomic DNA was quantified using a fluorescent dsDNA-binding dye assay (Qubit Fluorometer, Life Technologies, CA, U.S.). Low-input library generation was performed using limited DNA solution (~50 ng). Then the required length of DNA

fragments (150-bp paired-end read length) was collected by electrophoresis, and cluster preparation was performed by adding joints. Finally, whole-genome shotgun sequencing of ambient air samples was performed on an Illumina Hiseq X Ten platform. Details of the DNA quantity and data size were provided in Appendix 3.1, and an overview of metagenomic sequencing including the key steps was shown in Figure **3-5**. The raw reads were filtered using fastp (v0.21.0 with default parameters) to remove low-quality reads (Chen *et al.*, 2018). After quality control procedures, the clean data (~870 GB) of 71 samples were uploaded to the SRA database of NCBI with the accession number PRJNA858396 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA858396).

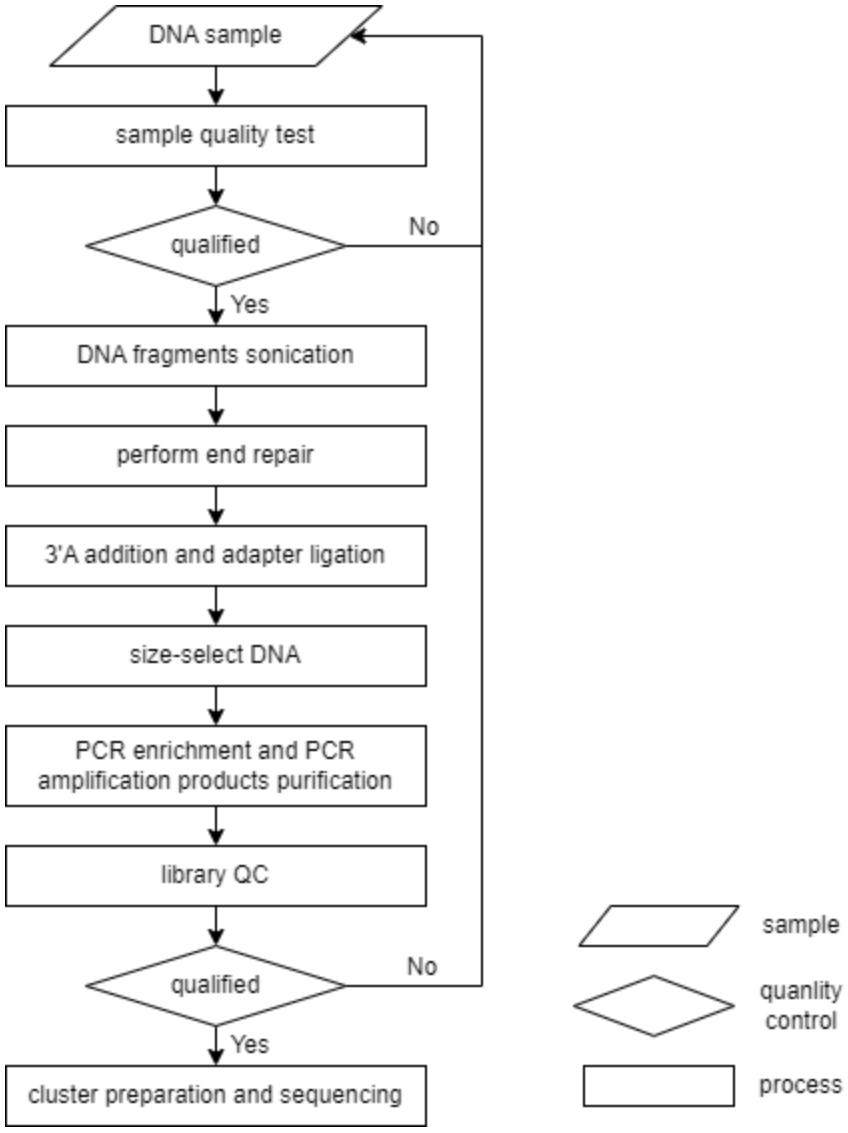


Figure 3-5 Flow chart of the metagenomic sequencing.

3.3 Global airborne bacterial dataset generation and analyses

A dataset of airborne bacteria on a global scale was compiled using 76 newly collected air samples (a combination of 803 weekly samples listed as No. 1-76 in Appendix 1), along with an additional 294 samples from reliable studies conducted across 63 sites worldwide.

The sampling sites were diverse, ranging from ground level (1.5 - 2 m high) to rooftops (5 - 25 m high), to high-altitude mountain peaks (5,380 m asl), and included densely populated urban centers and even the Arctic Circle. This broader coverage of altitudes and geographical regions provides greater diversity than has previously been attempted. Additionally, 16S rRNA gene sequencing data were gathered from EMP (Thompson *et al.*, 2017), involving 5,000 samples from 23 different surface environments, to explore the relationships between airborne bacteria and other microbiomes. The section also includes details on environmental data acquisition, chemical analysis, and various bioinformatic analyses.

3.3.1 Metadata collection

In order to broaden the scope of the study on airborne bacterial communities, only studies that utilized a filter-based flow sampler, total DNA extraction, high-throughput sequencing on an Illumina platform, and 16S rRNA gene sequence data were considered for air sample selection. This led to the identification of 294 air samples from literature (listed as No. 77-370 in Appendix 1) that met the quality standards and were downloaded and uniformly processed. Despite variations in flow rate and sampling time, the quantifications per unit volume of each sample were calculated, resulting in a global airborne bacterial dataset consisting of 370 air samples with different particle sizes (68 $PM_{2.5}$, 171 PM_{10}, and 131 TSP) covering 63 sites worldwide including a wide range of latitudes (65.53°S - 81.57°N), altitudes (0 – 5380 m asl), climates (15 climatic types following the Köppen−Geiger climate classification system (Peel *et al.*, 2007) (Figure **3-6**), anthropogenic impacts (*e.g.*,

urban, terrestrial background, and offshore areas), and land cover types. Detailed information of climate types and proportion of land cover types at each sampling site was shown in Table **3-2**, and detailed information of each sample was listed in Appendix 1. Additionally, the quantification of land covering type was described in section 3.3.3.2. Moreover, the rarefaction curve of global airborne bacterial community was constructed to verify that the global dataset covered the diversity in the air (Figure **3-7**).

We also obtained a global topsoil 16S rRNA gene sequence dataset (Sunagawa *et al.*, 2015) (n = 65, PRJEB19856) and a global metagenomic dataset on the surface seawater layer (Bahram *et al.*, 2018) (n = 62, PRJEB7988) from the NCBI to compare with the airborne microbial communities.

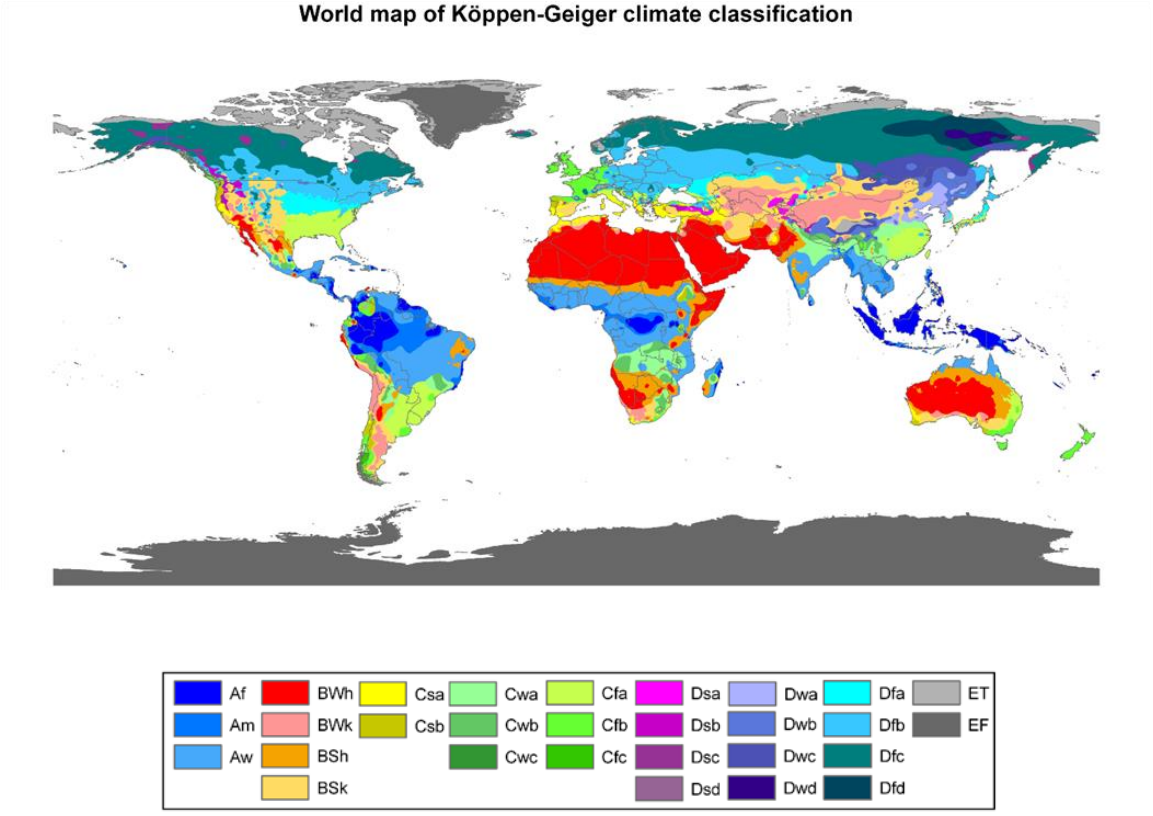


Figure 3-6 World map of Köppen-Geiger climate classification (Peel *et al.*, 2007).

The Köppen climate classification divides climates into five main climate groups, with each group being divided based on seasonal precipitation and temperature patterns. The five main groups are A (tropical), B (dry), C (temperate), D (continental), and E (polar).

Color indicates climate types: Af, Tropical rainforest climate; Am, Tropical monsoon climate; Aw, Tropical savanna climate; BWh, Hot desert climate; BWk, Cold desert climate; BSh, Hot semiarid climate; BSk, Cold semi-arid climate; Csa, Hot-summer Mediterranean climate; Csb, Warm summer Mediterranean climate; Cwa, Humid subtropical climate; Cwb, Subtropical highland climate; Cwc, Cold subtropical highland climate; Cfa, Humid subtropical climate; Cfb, Temperate oceanic climate; Cfc, Subpolar oceanic climate; Dsa, Mediterranean-influenced hot summer humid continental climate; Dsb, Mediterranean-influenced warm-summer humid continental climate; Dsc, Mediterranean-influenced subarctic climate; Dsd, Mediterranean influenced extremely cold, subarctic climate; Dwa, Monsoon-influenced hot-summer humid continental climate; Dwb, Monsoon-influenced warm-summer humid continental climate; Dwc, Monsoon-influenced subarctic climate; Dwd, Monsoon-influenced extremely cold subarctic climate; Dfa, Hot-summer humid continental climate; Dfb, Warm-summer humid continental climate; Dfc, Subarctic climate; Dfd, Extremely cold subarctic climate; ET, Tundra climate; and EF, Ice cap climate.

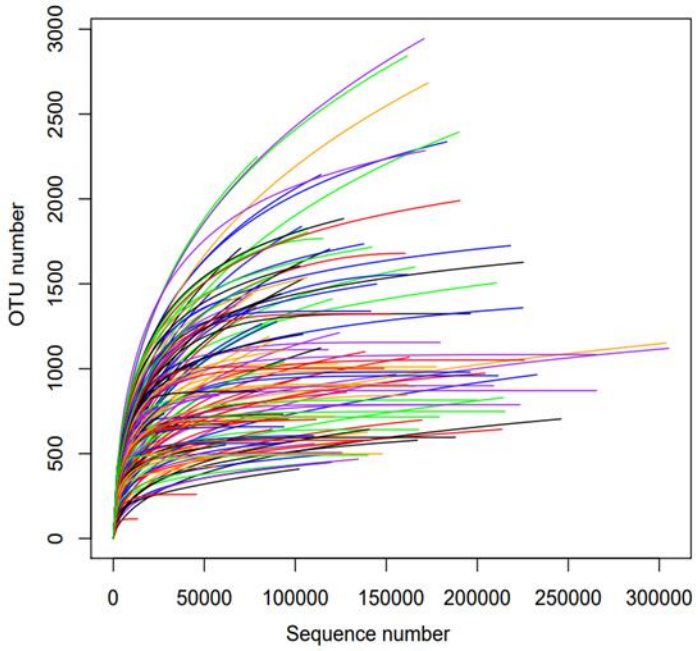


Figure 3-7 Rarefaction curve of all sequences at a global level. The sequence number of each sample ranged from 9,206 to 316,150, with a mean value of 74,914 reads.

Table 3-2 The overview of climate types and land cover types of 370 air samples in 63 sampling sites.

Sites	Climate system	Sample size	Land coverage type composition (%)					
			Water/Seas	Forest	Shrubs	Grassland	Cropland	Built-up
Bachok, Malaysia	Af	4	0.00	0.00	3.21	10.29	84.88	1.62
Phitsanulok, Thailand	Aw	4	54.27	5.69	29.12	7.02	3.52	0.38
Tsogt-Ovoo, Mongolia	BWk	12	0.00	0.00	59.78	0.00	0.00	40.22
Nanjing_ LS, China	Cfa	5	19.08	0.09	24.14	17.68	39.00	0.00
Nanjing_ XW, China	Cfa	5	10.97	0.61	19.39	21.79	43.22	4.02
Nanjing_ PK, China	Cfa	5	3.91	0.60	19.01	29.72	46.23	0.53
Ohio 1, USA	Cfa	6	34.85	0.95	33.04	1.02	10.25	19.88
Ohio 2, USA	Cfa	6	1.45	1.77	12.02	10.71	11.70	62.34
Ljungbyhed, Sweden	Cfb	46	1.50	39.00	26.79	1.23	31.39	0.10
Wiesbaden, Germany	Cfb	43	2.12	18.37	16.49	4.29	54.26	4.48
Cape Point, South Africa	Csb	7	93.55	0.04	5.74	0.18	0.03	0.47
Hok Tsui, China	Cwa	4	89.82	0.86	6.10	2.16	0.00	1.07
Mt. Ailao, China	Cwa	4	0.00	83.19	11.61	1.15	4.05	0.00
Guangzhou_ TH, China	Cwa	5	7.44	0.02	14.11	44.69	30.63	3.11
Guangzhou_ HS, China	Cwa	5	9.37	0.00	22.47	26.89	40.86	0.41
Guangzhou_ CH, China	Cwa	5	2.03	15.09	36.29	5.96	40.62	0.00
Hong Kong, China	Cwa	4	67.01	2.31	8.08	12.81	8.59	1.19
Chacaltaya, Bolivia	Cwb	16	0.00	2.05	52.86	40.97	4.11	0.00
STP, USA	Dfa	4	0.12	0.00	39.92	10.82	1.83	47.32
Colorado 1, USA	Dfa	2	0.26	0.00	9.95	46.81	31.78	11.21
Colorado 2, USA	Dfa	1	0.00	65.36	20.40	11.39	2.52	0.33
Colorado 3, USA	Dfa	1	0.23	0.00	8.22	37.82	49.67	4.06

Sites	Climate system	Sample size	Land coverage type composition (%)					
			Water/Seas	Forest	Shrubs	Grassland	Cropland	Built-up
Colorado 4, USA	Dfa	1	0.23	0.97	16.92	33.48	45.18	3.23
Colorado 5, USA	Dfa	1	0.00	0.00	49.63	29.55	16.82	4.00
Colorado 6, USA	Dfa	1	0.00	36.76	51.88	9.77	1.59	0.00
Colorado 7, USA	Dfa	1	0.00	6.74	14.76	32.33	43.99	2.18
Colorado 8, USA	Dfa	8	0.00	60.62	24.01	3.29	12.09	0.00
Colorado 9, USA	Dfa	1	1.18	6.33	16.04	43.38	31.09	1.99
Michigan 1, USA	Dfa	5	12.49	0.00	2.82	2.58	23.51	58.59
Michigan 2, USA	Dfa	3	0.73	1.13	51.31	0.78	45.53	0.52
Pic-du-Midi, France	Dfb	12	0.00	24.29	10.52	0.00	65.01	0.18
Kiruna, Sweden	Dfb	46	9.45	3.62	84.53	1.38	0.00	1.02
Grenoble, France	Dfc	10	0.67	61.40	26.86	3.08	7.48	0.51
PuydeDôme, France	Dfc	12	0.70	31.72	46.33	2.55	18.32	0.38
Beijing_ PKU, China	Dwa	5	0.25	0.14	5.80	48.67	30.80	14.33
Beijing_ IAP, China	Dwa	5	0.25	0.14	5.80	48.67	30.80	14.33
Beijing_HR, China	Dwa	4	4.16	6.09	60.54	3.85	25.36	0.00
Mt. Changbai, China	Dwb	4	0.11	36.04	55.91	0.91	6.98	0.05
Waliguan, China	Dwc	4	0.00	0.00	44.85	34.66	19.24	1.24
Mt. Everest	ET	4	0.00	0.00	46.38	2.92	0.00	50.71
Namco, China	ET	8	22.63	0.00	14.45	62.74	0.00	0.18
Station-Nord, Greenland	ET	13	100.00	0.00	0.00	0.00	0.00	0.00
Amsterdam-Island, France	Offshore	9	100.00	0.00	0.00	0.00	0.00	0.00
South Ocean (19 sites)	Offshore	19	100.00	0.00	0.00	0.00	0.00	0.00

3.3.2 Sequence processing

In order to ensure consistent processing of the data, all of the global air data that was collected - a total of 27,719,673 V3-V4 hypervariable regions of 16S rRNA gene amplicon reads from 370 air samples combined - were processed using mothur (v1.42) (Schloss *et al.*, 2009). To begin, the VSEARCH tool was used to remove chimeric sequences, employing the UCHIME algorithm for quality control (Rognes *et al.*, 2016). Next, the sequences were separated into operational taxonomic units (OTUs) at a 97% similarity threshold using the UPARSE pipeline and taxonomically annotated using SILVA (v123) as the reference database, with an 80% confidence cut-off (Pruesse *et al.*, 2007). To identify bacterial pathogens at the species level, raw sequences for each sample were processed against pathogenic sequences using the 16SPIP pipeline, with a criterion of 99% or greater similarity (Miao *et al.*, 2017). This method has been tested and found to be effective in identifying pathogens using paired reads of the V3-V4 region of the 16S gene from Beijing hospital samples, identified through culture and whole-genome shotgun metagenomic analyses (Miao *et al.*, 2017). Multivariate data analysis from METAGENassist was employed to generate phenotypic information covering various functional categories, such as genome size, oxygen requirements, energy sources, and preferred temperature range (Arndt *et al.*, 2012). Additionally, the same procedure was used to reanalyze the global topsoil 16S rRNA gene sequences.

3.3.3 Acquisition of abiotic factors

In this section, the chemical components commonly found in air, such as organic carbon (OC) and element carbon (EC), heavy metals, and soluble ions, were detected and analyzed. Besides, the meteorological condition and air quality at each sampling site during sampling periods were collected from various official websites and datasets. The land cover types were estimated using satellite data.

3.3.3.1 Chemical Analysis

Filter-based air samples (size: 5cm * 7.5cm) were cut and submerged in a mixture of HNO_3 and $HClO_4$ with a volumetric ratio of 4:1 in test tubes. The solution was then digested in a heating block with a progressive temperature-raising program up to 190 ^oC until dried. A 10-mL aliquot of 5% (v/v) HNO_3 was added to the test tubes at a temperature of 70 ^oC. After pretreatment of the filter samples, the final solution was analyzed by inductively coupled plasma-optical emission spectrometry (ICP-OES) to detect the total concentrations of various elements.

The concentrations of OC and EC were analyzed by the DRI Thermal/Optical Carbon Analyzer (Model 2001) using the thermal optical reflectance (TOR) protocol (Chow *et al.*, 1993). Soluble anions (NO_3^- and SO_4^{2-}) and cations (NH_4^+) infiltrates were analyzed by Ion Chromatography.

3.3.3.2 Acquisition of Environmental Data

In this study, all environmental variables were categorized into three major groups: meteorological condition, air quality, and earth surface type. The hourly meteorological data for each site during the sampling duration, containing AT, AP, RH, WS, and WD, were downloaded from the official website of the National Climatic Data Center (ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-lite/) and Weather Underground (https://www.wunderground.com/). Since there is still no globally uniform air quality monitoring system like the NOAA (National Oceanic and Atmospheric Administration) for air pollutant records, the corresponding air quality data including Air Quality Index (AQI), PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3, and CO, were retrieved from various sources, including the United States Environmental Protection Agency (https://www.epa.gov/), China National Environmental Monitoring Centre (http://www.cnemc.cn/), European Environmental Agency (https://www.eea.europa.eu/), the Royal Thai Pollution Control Department (http://aqmthai.com/aqi.php), the South Africa Air Quality Information System (http://saaqis.environment.gov.za/), the Malaysian Department of the Environment (https://www.doe.gov.my/portalv1/en/), the Government of Peru (https://www.gob.pe/), and Environment and Climate Change Canada (http://data.ec.gc.ca/data/).

To quantify the types of land cover within a 50 km diameter range of sampling sites, a moderate-resolution imaging spectroradiometer (MODIS) land cover approach was utilized with a 5′ × 5′ resolution. The International Geosphere–Biosphere Programme (IGBP) system (MCD12Q1-1) was performed to describe the different MODIS land coverages (Friedl *et al.*, 2002). The relative contribution of each type of landscape to the aerial emission of bacterial cells was predicted by weighting these relative surfaces by

their associated bacterial cell concentration, as reported earlier (Table **3-3**) (Burrows *et al.*, 2009b). The maps of the land cover type at each site were shown in Appendix 2.

Table 3-3 Estimates of total mean bacterial aerosol concentration in near-surface air above various land cover types.

Earth Surface	Best Estimate (cells/m^3)	Reference
Water/Sea	1×10^4	(Bauer *et al.*, 2002b; Harrison *et al.*, 2005; Griffin *et al.*, 2006)
Forest	5.6×10^4	(Shaffer and Lighthart, 1997)
Cropland	1.1×10^5	(Harrison *et al.*, 2005)
Shrubs	3.5×10^5	(Bauer *et al.*, 2002b)
Grassland	1.1×10^5	(Harrison *et al.*, 2005)
Urban	1.2×10^5	(Harrison *et al.*, 2005)
Tundra	1.2×10^4	(Burrows *et al.*, 2009b)

3.3.4 Bioinformatic and statistical analyses

3.3.4.1 Core bacteria identification

The determination of a global core set of airborne bacteria was based on their abundance and occupancy in regional and temporal variations with reference to multiple reported methods. First, 166 OTUs with a high mean relative abundance (> 0.01%) across 370 samples were selected as overall abundant OTUs (Saunders *et al.*, 2016). Then, 68 OTUs among the overall abundant OTUs with an occurrence frequency in all samples of more than 80% were filtered out as widespread OTUs (Székely and Langenheder, 2014). Finally,

the dominant OTUs in more than half of the samples were further filtered out and identified as the final core airborne bacterial community. In sum, this process involved sorting OTUs based on their abundance in each sample, selecting the top 80% of bacterial communities in each sample with an occurrence in more than half of the samples to form the final core bacteria.

3.3.4.2 Diversity analyses and correlations with environmental factors

To compare the structure of the bacterial community across different regions, α-diversities and β-diversities were computed using the "Picante" package (Kembel *et al.*, 2010) in R based on the original OTU table generated by sequence processing. To minimize deviations in the number of bacterial taxa (*i.e.*, richness) caused by different particle sizes, the mean ratio of richness in $PM_{2.5}$ with PM_{10} (1: 1.73 ± 0.63) and TSP (1: 2.14 ± 0.84) collected in the same sites during the same sampling period was used to modify the richness data drawn from the literatures (Gou *et al.*, 2016; Lu *et al.*, 2018; Yan *et al.*, 2018; Stern *et al.*, 2021) (Figure **3-8**). The Bray–Curtis dissimilarity matrix for the airborne bacterial community structure (OTU abundance-based) between pairs of samples was calculated to estimate the taxonomic β-diversity using the "vegdis" function in the "vegan" R package (Jari Oksanen *et al.*, 2018). The geographic distance between any two sampling sites across the globe was calculated using the "geosphere" R package based on geographic coordinates (Robert J. Hijmans *et al.*, 2015).

For each environmental variable, a partial Mantel test with 999 permutations was conducted to examine the correlation (Pearson's rank correlation) between the

environmental variable matrix and the composition of the bacterial community using the "vegan" R package (Jari Oksanen *et al.*, 2018). Akaike information criterion (AIC) was used to select the best model to fit the relationship between two variables. AIC has been widely used for statistical inference and gradually forms the basis of a paradigm for the foundations of statistics (Stoica and Selen, 2004).

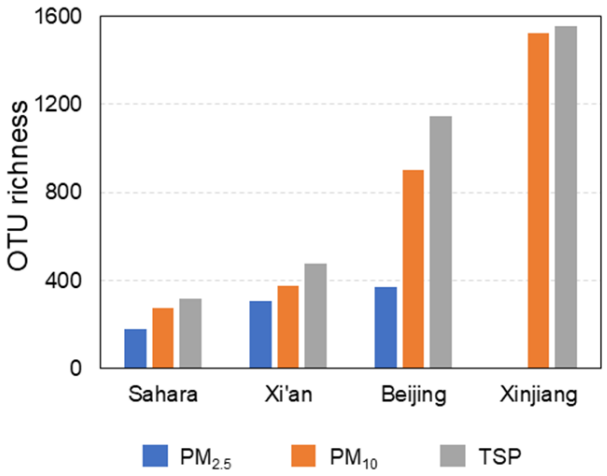


Figure 3-8 Effect of particulate matter size on airborne bacterial richness. Comparison of airborne bacterial richness in $PM_{2.5}$, PM_{10}, and TSP collected in the same site during the same period(Gou *et al.*, 2016; Yan *et al.*, 2018; Yue *et al.*, 2018; Stern *et al.*, 2021).

3.3.4.3 Network construction, topological property calculation, and key taxa identification

To construct the co-occurrence network, OTUs with a mean relative abundance of over 0.1% in all samples and a relative abundance of over 0.5% in any one sample were retained

while rare OTUs were filtered out. The network was based on the Spearman correlation matrix performed with the WGCNA package (Langfelder and Horvath, 2012), and only statistically significant relationships ($p < 0.01$, $|R| > 0.6$) were included (Junker, 2008). The correlation cut-off threshold for network construction was mathematically defined by a random matrix theory-based approach (Deng *et al.*, 2012) with the "RMThreshold" R package (Menzel and Menzel, 2016). The nodes in this network represent OTUs, and the edges (*i.e.*, connections) represent significant associations between OTUs.

The topological properties of each node in the resulting network were calculated using the "igraph" R package (Patrick R. Amestoy *et al.*, 2020), including degree, betweenness centrality, closeness centrality, and transitivity. Betweenness centrality reveals the role of a node as a bridge between components of a network, while degree reveals the role of a node with direct connections with other OTUs in the whole community. Thus, the two important indexes were normally selected as critical criteria for the identification of key taxa in the overall co-occurrence network (Banerjee *et al.*, 2018). OTUs with high degree (> 72, 75% of the highest degree) and low betweenness centrality scores (< 2000, 10% of the highest betweenness centrality) were identified as the key taxa (Berry and Widder, 2014). Besides, the topological properties were also calculated for the entire co-occurrence network to better understand the interconnected bacteria (Berry and Widder, 2014). These topological indexes contained the number of nodes (interconnected objects, namely bacteria), the number of edges (links between each pair of nodes), the average degree (groups of objects with tight interconnections), the average shortest path length (average network distance between all pairs of nodes), the average connectivity (maximum of

pairwise distances between every two nodes), and the average clustering coefficient (the degree to which nodes tend to cluster together).

The “smallworldness” index, which represents the correlation compactness of individuals in a network, was computed using the “qgraph” R package based on the transitivity (any pairs of nodes with direct or indirect connections could be transitive) and the average shortest path length (the average number of steps along the shortest paths for all possible pairs of network nodes). A network can be recognized as "small-world" if its “smallworldness” is higher than one (a stricter rule is “smallworldness” ≥ 3) (Humphries and Gurney, 2008). Additionally, the network is inspected to ensure that its transitivity is substantially higher than that of comparable random networks and that its average shortest path length is similar to or higher (but not many times higher) than that of random networks (Watts and Strogatz, 1998). Edge weights, signs, and directions are not taken into account in the computation of the indices.

3.3.4.4 Estimation of the total abundance and richness of global microbiomes

Scaling laws describe the functional relationship between two physical quantities, *i.e.*, the total abundance (N_T) and the abundance of the most abundant species (N_{max}), which scale with each other over a significant interval, underpin unifying theories of biodiversity, and are among the most predictively powerful relationships in biology. Scaling laws were used to predict global airborne bacterial richness (S) based on the lognormal species abundance model (Locey and Lennon, 2016). S could be predicted in terms of N_T, N_{max}, and the assumption that the rarest species is a singleton ($N_{min} = 1$), and by using Equation 3-1:

$$S = \frac{\sqrt{\pi}}{a} exp\left\{\left[alog_2\left(\sqrt{\frac{N_{max}}{N_{min}}}\right)\right]^2\right\}$$

Equation 3-1

where *a* could be numerically solved by Equation 3-2 and Equation 3-3:

$$N_T = \frac{\sqrt{\pi N_{max} N_{min}}}{2a} exp\left\{\left[alog_2\left(\sqrt{\frac{N_{max}}{N_{min}}}\right)\right]^2\right\} exp\left\{\left[\frac{\ln(2)}{2a}\right]^2\right\}$$

Equation 3-2

$$\left\{erf\left[alog_2\left(\sqrt{\frac{N_{max}}{N_{min}}} - \frac{\ln(2)}{2a}\right)\right] + erf\left[alog_2\left(\sqrt{\frac{N_{max}}{N_{min}}} + \frac{\ln(2)}{2a}\right)\right]\right\}$$

Equation 3-3

The total airborne bacterial abundance (N_T) was estimated by using qPCR quantification results of 16S rRNA gene copies from this study and published data. Thus, the N_T (global airborne bacterial abundance in the troposphere) is about 2.69×10^{25}. Details of the calculation process are given below. In addition, the value of N_{max} was inferred based on the proportion of the typically most abundant genus or using the dominance-abundance scaling law (Equation 3-4):

$$N_{max} = 0.4 \times {N_T}^{0.93}$$

Equation 3-4

The most dominant taxonomic unit (based on a 97% 16S rRNA sequence similarity) in the troposphere was typically predicted to be a member of the *Bacillus* genus and accounted for around 1.2% of 16S rRNA gene reads in the whole dataset. N_{max} would then be approximately 3.2×10^{23}, which was close to the prediction by the scaling law, $N_{max} = 1.78 \times 10^{23}$. The same method was also applied to microbial communities in global soil, global freshwater, and global leaf surfaces.

The approximate values of N_T and N_{max} for global Microbiomes were calculated as follows:

(1) Troposphere. The exact upper boundary of the atmosphere (biosphere) was a Gordian knot for estimating the total number of airborne bacteria in the whole biosphere. The traditionally cited highest altitude for aerobiology is 77 km, due to the detection of microbes on the surface of one rocket; however, many researchers doubted that the microbes came from the rocket itself (in that study there was no detailed description of sterilized operations, or of the steps taken to prevent contamination) or from the flying soil caused by the rocket making landfall (Smith, 2013). Nevertheless, it is certain that the troposphere contains approximately 80% of the total mass of the atmosphere; also, the temperature was -56 ^{o}C and the humidity was nearly zero at the frontier between troposphere and stratosphere, where microbiomes can hardly survive (Horneck *et al.*, 2010). In addition, human activities mainly proceeded in the troposphere. As a result, the study focused on airborne bacterial communities in the troposphere, beginning at the land surface and extending to between 17 km at the equator and 7 km at the poles, with a mean altitude of 12 km (Horneck *et al.*, 2010). First, the earth was assumed as a sphere with a radius of 6,371 km and overlooked the surface effects like mountains and valleys. Secondly, the troposphere was divided into three circles based on elevation: 0-1 km, 1-8 km, and 8-12

km. The mean bacterial density in the three circles was predicted: 4.8×10^5 cells/m^3 (0-1 km); 6.8×10^5 cells/m^3 (1-8 km); 5.1×10^3 cells/m^3 (8-12 km) based on the measurements in this study and published studies. And the locations of these air samples were shown in a global map (Figure **3-9**). It was assumed that the air was kept in the same conditions within cycles, ignoring the intra-circle variations in AT and AP. The total bacterial loading in each circle was calculated by multiplying the total air volume and mean bacterial density in the corresponding circle. Finally, the values were added up to determine the total abundance of airborne bacteria in the troposphere. The N_T (global airborne bacterial abundance in the troposphere) was estimated to be about 2.69×10^{25}, and N_{max} was inferred as 3.2×10^{23} or 1.78×10^{23} respectively by the proportion of the most dominant genus and scaling law (Locey and Lennon, 2016).

(2) Global soil. The most dominant genus-level candidate in global soil is Mycobacterium, with an estimated proportion of 0.61% in each sample on average based on a 97% 16S rRNA sequence similarity (Bahram *et al.*, 2018). The total number of bacteria cells in the global soil was estimated using literature data, which suggests that there are about 9.4×10^{28} microbial cells in the global soil ecosystems (Whitman *et al.*, 1998). The detailed steps were as follows: First, the global soil was categorized into 12 classes according to their ecosystem types, and the total areas of each class on the Earth's surface were calculated. Then, the mean bacterial densities were calculated with different depths (0-1 m and 1-8 m) based on as many as possible measurements globally. Next, boreal forest and tundra and alpine soils were assumed to be 1 m deep, but other classes of soil were 8 m deep. The soil volume could be calculated by multiplying surface areas and depth. Finally, the sum of the results of 12 classes of soils suggested that N_{max} of the global soil would be approximately

5.7×10^{26}, which is close to the estimated value of N_{max} using the dominance scaling law of 3.5×10^{26} (Locey and Lennon, 2016).

(3) Global freshwater. the total number of microbial cells in global freshwater, including both rivers and lakes, was estimated to be 4.7×10^{25} (Whitman *et al.*, 1998; Kallmeyer *et al.*, 2012). In addition, the most abundant taxonomic unit (based on a 97% sequence similarity in 16S rRNA reads) in global freshwater is typically a member of the Pseudomonas genus, accounting for around 1.57% of the 16S rRNA gene reads in a sample based on the EMP database (Thompson *et al.*, 2017); the N_{max} of the global freshwater was estimated to be 7.4×10^{23} or using the scaling law (Locey and Lennon, 2016), as 3.0×10^{23}.

Figure 3-9 Map showing the geographical locations of the sampling sites for estimating the mean airborne bacterial density from this study and literature (Harrison *et al.*, 2005; Bowers *et al.*, 2009; Lee *et al.*, 2010; Cho and Hwang, 2011; Bowers *et al.*, 2012; Bertolini *et al.*, 2013; DeLeon-Rodriguez *et al.*, 2013; Murata and Zhang, 2014; Barberán *et al.*,

2015; Tanaka *et al.*, 2015; Gou *et al.*, 2016; Deng *et al.*, 2016a; Gat *et al.*, 2017; Genitsaris *et al.*, 2017; Innocente *et al.*, 2017; Xu *et al.*, 2017; Zhen *et al.*, 2017; Gao *et al.*, 2017a; Šantl-Temkiv *et al.*, 2018; Shen *et al.*, 2019; Tignat-Perrier *et al.*, 2019; Xu *et al.*, 2019; Maki *et al.*, 2022).

3.3.4.5 Interconnections of bacterial community composition of air with other bacterial habitats

This study utilized a bacterial abundance table containing 5,000 global samples from multiple habitats, which was obtained from the EMP database (Thompson *et al.*, 2017). The same standard workflow employed by the EMP was also followed to analyze airborne bacterial sequence data, using closed reference against Greengenes 13.8 in Qiime2, which picked 68.1% of the sequences for higher quality control. (Bolyen *et al.*, 2019). The resulting OTU table was merged with the EMP OTU table, which included samples from soil, rhizosphere, freshwater, ocean, air, human and animal-associated habitats. To compare the microbial community compositions across habitats, non-metric multidimensional scaling (NMDS) analysis was performed based on the Bray–Curtis dissimilarity matrix. In addition, the derived OTU table was used as the input file to estimate the proportion of each airborne bacterial sample attributable to various habitats on the genus level by using "SourceTracker" (Knights *et al.*, 2011). To explore the patterns of bacterial community interconnection and coexistence across various habitats at the global scale, the Earth's metacommunity co-occurrence network was also constructed using a communal catalog with 5,189 samples with robust correlations ($\rho > 0.7$, $p < 0.01$).

3.3.4.6 Quantifying ecological stochasticity in the microbial community assembly process

The normalized stochasticity ratio (NST) was used to quantitatively assess ecological stochasticity in microbial community assembly process, *i.e.*, the degree of randomness or unpredictability in ecological systems (Ning *et al.*, 2019). The NST was calculated by comparing the variance of observed species abundances to the expected variance under a neutral model of community assembly, which assumed that species were distributed randomly, based on "NST" R package (Ning and Ning, 2021). A high NST indicates that ecological stochasticity is playing a large role in community assembly, while a low NST suggests that deterministic factors such as competition or environmental filtering are more important (Ning *et al.*, 2019). The NST was used to compare stochasticity across different communities or ecosystems.

3.3.4.7 Multivariate analysis

Principal Coordinate Analysis (PCoA) was utilized to visualize the spatial and temporal differences in key taxa and core bacterial communities between samples based on the Euclidean index and to assess the impacts of environmental variables on these communities. Additionally, RDA was used to identify the relationships between key taxa or core bacterial genera and soluble ions or heavy metals in $PM_{2.5}$ through the "vegan" R package (Jari Oksanen *et al.*, 2018).

To quantify the relative contributions of the environmental and distance effects (separately or jointly) on β-diversity, a VPA was performed based on the RDA algorithm. To reduce the number of explanatory variables entering the analysis while keeping the variance explained by them to the maximum, a subset of explanatory variables from the set of all variables for constrained ordination was selected using forward selection (adding explanatory variables one by one) and backward selection (starting from the full model and deleting variables of which the least decreases the total explained variance). In addition, the initial set of explanatory groups included three explanatory groups: air quality (AQI, PM_{10}, $PM_{2.5}$, NO_2, SO_2, and CO), meteorological conditions (AT, AP, WS, WD, and RH), and land cover type (water/sea, forest, shrubs, grassland, cropland, and built-up areas). After forward selection of environmental factors, the remained variables for VPA analysis are: NO_2, CO, PM_{10}, $PM_{2.5}$, WS, AT, water/sea, cropland, and grassland (core bacteria); SO_2, NO_2, CO, O_3, PM_{10}, $PM_{2.5}$, AP, AT, RH, WS, WD, water/sea, forest, shrubs, cropland, grassland, and built-up areas (key taxa).

Furthermore, to explore the direct and indirect relationships among geographic locations, environmental variables, and bacterial communities, SEM was built using the "lavaan" package (Rosseel, 2012). The prior model included all hypothesized reasonable indirect and direct links among the variables based on their pairwise correlations. The non-significant relationships and variables were subsequently removed, and new links between other terms were also created, *i.e.*, the post hoc model modification, until all quantitative indices met the overall goodness of fit. The composition of airborne key taxa and core bacterial communities was indicated by the first principal coordinates in PCoA (PC1). The SEM evaluation is based on the fit indices for the test of a non-significant chi-square test

(p > 0.05), the root means a square error of approximation (RMSEA) < 0.08, the standardized root means square residual (SRMR) < 0.05, the Tucker-Lewis index (TLI) > 0.90, and the comparative fit index (CFI) > 0.95.

3.4 Global airborne metagenomic dataset generation and analyses

A metagenomic dataset containing 262 air samples collected by high-volume samplers was organized to gain insight into the global distribution and dissemination of airborne ARGs as well as their potential hosts. This dataset was then used to construct a comprehensive atlas of global airborne AMR, which evaluated the exposure risks to humans based on abundance, mobility, and host pathogenicity from both regional and seasonal perspectives.

3.4.1 Metagenomic data collection

To extend airborne antibiotic resistome into a global scale, the studies were included if they met specific standards: a) using a high-volume air sampler near the ground, b) pretreatment with PBS and PES membrane disc, c) total DNA extraction, d) whole-genome shotgun sequencing on Illumina platform, and e) sequencing size over 2 GB. In total, 191 air samples incorporated from the literature (No. 72-262 in Appendix 3.1) that met the selection criteria were included and processed with the 71 air samples collected in this study (No. 1-71 in Appendix 3.1) uniformly, resulting in a global metagenomic dataset (~ 1.33 TB) of 262 air samples with different particle sizes (116 $PM_{2.5}$, 53 PM_{10}, and 93 TSP), covering 32 sites worldwide, including both background and urban areas, with a wide

geographic range of latitudes (77.52°S – 69.13°N) and altitudes (6 – 4276 m asl) (Figure **3-10**). Detailed information of each sample in the global airborne metagenomic dataset was shown in Appendix 3.1. The rarefaction curve of the global metagenomic dataset indicated that the sequencing depth was deep enough for determining the diversity of global air samples (Figure **3-11**).

Additionally, to better understand the airborne antibiotic resistome, the global metagenomic dataset from other environmental media, such as topsoil (Bahram *et al.*, 2018) (*n=65*, PRJEB19856) and surface seawater layer (Sunagawa *et al.*, 2015) (*n=62*, PRJEB7988), were also obtained from the NCBI and connected to the airborne dataset.

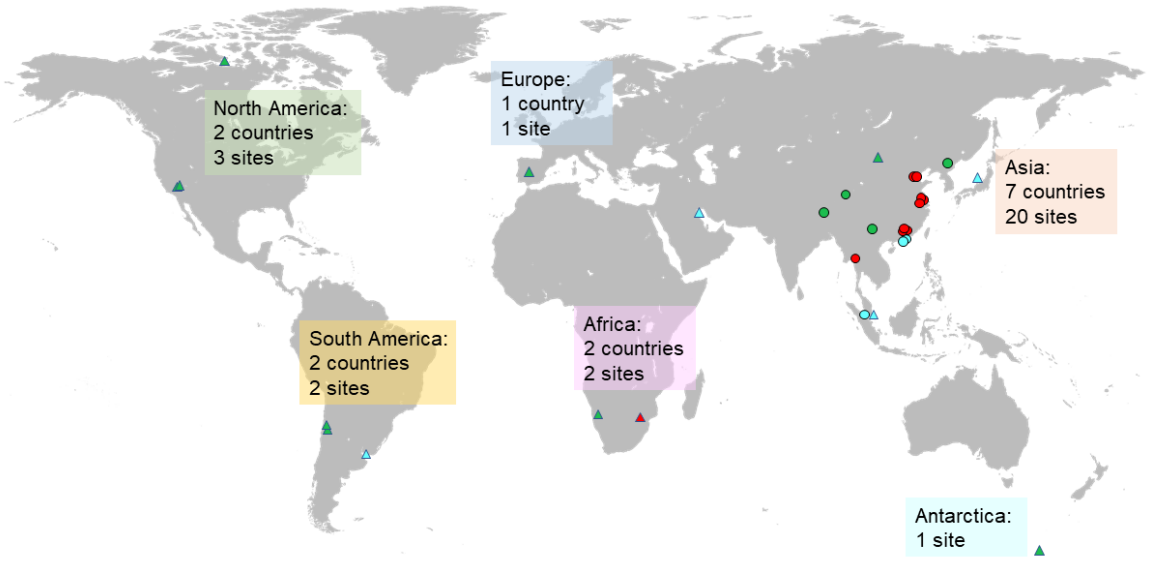


Figure 3-10 Locations where air samples and environmental data were collected across the globe.

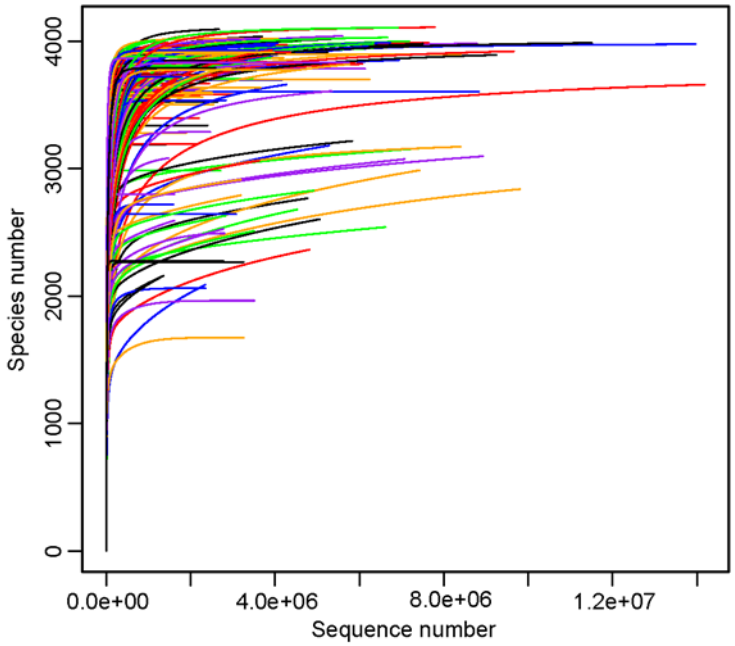


Figure 3-11 Rarefaction curve of all sequences from the metagenomic dataset at a global level. The sequence number of each sample ranged from1.02×10^6 to 1.72×10^7, with a mean value of 3.89×10^6 reads.

3.4.2 Taxonomic and ARG profiling

The filtered sequencing data were taxonomically profiled using Kraken 2 (v2.0.8-beta) (Wood *et al.*, 2019) and Bracken (v2.5.0) (Lu *et al.*, 2017) using the standard Kraken 2 database. Human pathogens, especially the nosocomial ESKAPE pathogens and WHO-identified priority pathogens, which urgently need new antibiotics (Table **3-4**), were identified according to a list summarized for broad-spectrum monitoring of bacterial pathogens in various environmental samples in a previous study (Li *et al.*, 2015). ARGs were annotated using the DeepARG pipeline with a 70% sequence similarity criteria (Arango-Argoty *et al.*, 2018), and an NMDS analysis was performed based on Bray-Curtis dissimilarity matrix to differentiate the bacterial and ARG profiles among various air samples. Subsequently, linear discriminant analysis effect size (LEfSe) was used to determine the biomarkers of airborne ARGs in urban and background air (Segata *et al.*, 2011).

Table 3-4 List of WHO-identified priority pathogens urgently need new antibiotics.

Level	Pathogen	Antimicrobial resistance type
Critical	*Acinetobacter baumannii*	carbapenem-resistant (beta-lactam)
	Pseudomonas aeruginosa	carbapenem-resistant (beta-lactam)
	Enterobacteriaceae	carbapenem-resistant (beta-lactam), ESBL-producing (Extended spectrum beta-lactamases)
High	*Enterococcus faecium*	vancomycin-resistant (glycopeptide)
	Staphylococcus aureus	methicillin-resistant (beta-lactam), vancomycin-intermediate and resistant (glycopeptide)
	Helicobacter pylori	clarithromycin-resistant (MLS)
	Campylobacter spp.	fluoroquinolone-resistant
	Salmonellae	fluoroquinolone-resistant
	Neisseria gonorrhoeae	cephalosporin-resistant (beta-lactam), fluoroquinolone-resistant
Medium	*Streptococcus pneumoniae*	penicillin-non-susceptible (beta-lactam)
	Haemophilus influenzae	ampicillin-resistant (beta-lactam)
	Shigella spp.	fluoroquinolone-resistant

3.4.3 Potential mobility and hosts of ARGs

The high-quality cleaned reads were de novo assembled using Megahit v1.13 with default parameters (Li *et al.*, 2016b), and the assembled contigs with lengths $\geqslant$ 500bp were kept. On the basis of the assembled contigs with lengths more than 1000 bp, both chromosomes or plasmids were predicted by using PlasFlow (Krawczyk *et al.*, 2018), and full-length plasmid sequences were identified by another new versatile metagenomic assembler, metaplasmidSPAdes (Antipov *et al.*, 2019). These output plasmid sequences from PlasFlow and metaplasmidSPAdes were further clustered by using CD-HIT (v4.6) (Fu *et al.*, 2012) with 90% clustering threshold. To evaluate the co-occurrence patterns of ARGs,

MGEs, and bacteria, contigs and plasmid sequences were aligned with existing ARG (v1.1.1.A.fasta; https://bench.cs.vt.edu/ftp/argminer/release/) and MGE databases (https://bench.cs.vt.edu/ftp/data/databases/) using DIAMOND (--id 50 --E-value 1e-7 --query-cover 70) (Buchfink *et al.*, 2015). ARGs colocalized with MGEs or located on the plasmids were considered potentially mobile. Taxonomic information was annotated using Centrifuge with default parameters (Kim *et al.*, 2016), and potential for dissemination of ARGs into human pathogens was estimated in each sample using MetaCompare based on the co-occurrence patterns of ARGs, MGEs, and pathogen-like sequences on assembled contigs(Oh *et al.*, 2018). Besides, the health risk ranking was evaluated based on the human-associated-enrichment, gene mobility, and host pathogenicity in the environments (Zhang *et al.*, 2021). The risk ranking assessment of AMR from both the genetic locations and universal understanding of ARGs would be more accurate and comprehensive.

3.4.4 Assembly of metagenomic bins and relevant resistomes analysis

To improve taxonomic classification, a genome-resolved "panorama" of airborne antibiotic resistomes was generated by clustering numerous contigs belonging to an individual genome (Liang *et al.*, 2020b) and producing high-quality individual metagenome-assembled genomes (MAGs) using metaWRAP (v1.2.1) (Uritskiy *et al.*, 2018). Firstly, the filtered clean metagenomic reads were co-assembled by sampling regions (urban or background sites) using megahit (Li *et al.*, 2016b) in the Assembly module with default parameters. After binning the output contigs by MetaBAT 2 (Kang *et al.*, 2019), MaxBin 2.0 (Wu *et al.*, 2016), and CONCOCT (Alneberg *et al.*, 2014), the constructed bacterial

genomes were further refined to generate high-quality MAGs in the built-in refining module with a criterion of contamination <10% and completeness criterion of >50%. At last, the taxonomic classifications of acquired MAGs were annotated using Genome Taxonomy Database (GTDB; v1.4.0) (Chaumeil *et al.*, 2019), and the quantification of MAGs (genome copies per million reads) was conducted by the Quant_bin module of metaWRAP (Uritskiy *et al.*, 2018).

The MAGs were individually processed to predict open reading frames (ORFs) by Prodigal (v2.6.3; -c -p meta mode) (Hyatt *et al.*, 2010). ARGs and MGEs were identified using the same method applied to contigs. MAGs containing at least one queried scaffold with ARGs were identified as potential antibiotic-resistant bacteria (PARB). To identify human virulent factor (HVF) genes, MAGs were mapped against the Virulence Factor Database (VFDB) (Chen *et al.*, 2005) using blastx with the E-value $\leqslant 1\times10^{-7}$. The identified PARB containing HVF genes could be treated as the virulent potential antibiotic-resistant bacteria (HVF-PARB), helping to assess environmental risks to human health (Wu *et al.*, 2022). A phylogenetic tree of these MAGs and 135 *S. aureus* isolate genomes in clinic infections (Manara *et al.*, 2018) was built using PhyloPhlAn (v3.0.51) (Asnicar *et al.*, 2020) and visualized using iTOL v6. HGTs among members of a microbial community at various taxonomic levels were detected using MetaCHIP v1.10.10 (Song *et al.*, 2019). Moreover, horizontally transferred ARGs (HT-ARGs) were enumerated and analyzed for transfer directionality to facilitate environmental monitoring and targeted controls of antibiotic resistance (Song *et al.*, 2019).

3.4.5 Bacterial cell quantification based on FCM method

The obtained 0.2 μm filtered extracts from 1× sterile PBS impingement medium with filter fragments were analyzed in three-parallel experiments by using a BD Accuri C6 Flow Cytometer (Figure **3-12**a). To detect bacterial viability, the samples were stained with the LIVE/DEAD BacLight Bacterial Viability Kit containing SYTO 9 and propidium iodide nucleic acid stains. After a 10-min incubation period in the dark at room temperature, the live bacteria with intact cell membranes emitted green fluorescence, while dead bacteria with compromised membranes emitted red fluorescence. The number of live cells and microspheres was counted, and the cell was quantified by the Equation 3-5:

$$\text{Cell count in samples (cells/ml)} = X/Y \times N/V$$

Equation 3-5

Where *X* represented the number of active cells detected by FCM; *Y* represented the number of microspheres detected by FCM; *N* represented the setting number of microspheres in the test system (*N=50000*); *V* represented the volume of samples.

To account for noise and ensure the accuracy of the bacterial counts, PBS solution blanks were also pre-treated and analyzed in the same way as the samples. In addition, ATCC Strain #25922 E. Coli (Migula) Castellani and Chalmers were used as standard bacteria, which were cultured in LB solution for 24 h, and serial dilutions were made using the E. Coli suspension (Figure **3-12**b). The absolute bacterial cell number in each solution was estimated using plate counting method. The known bacterial density of 10-fold serial dilutions was used as a standard to determine the absolute bacterial cell counts. The results

of FCM and colony counting were found to be remarkably consistent, validating the accuracy of the results (R^2=0.963, Figure **3-12**c).

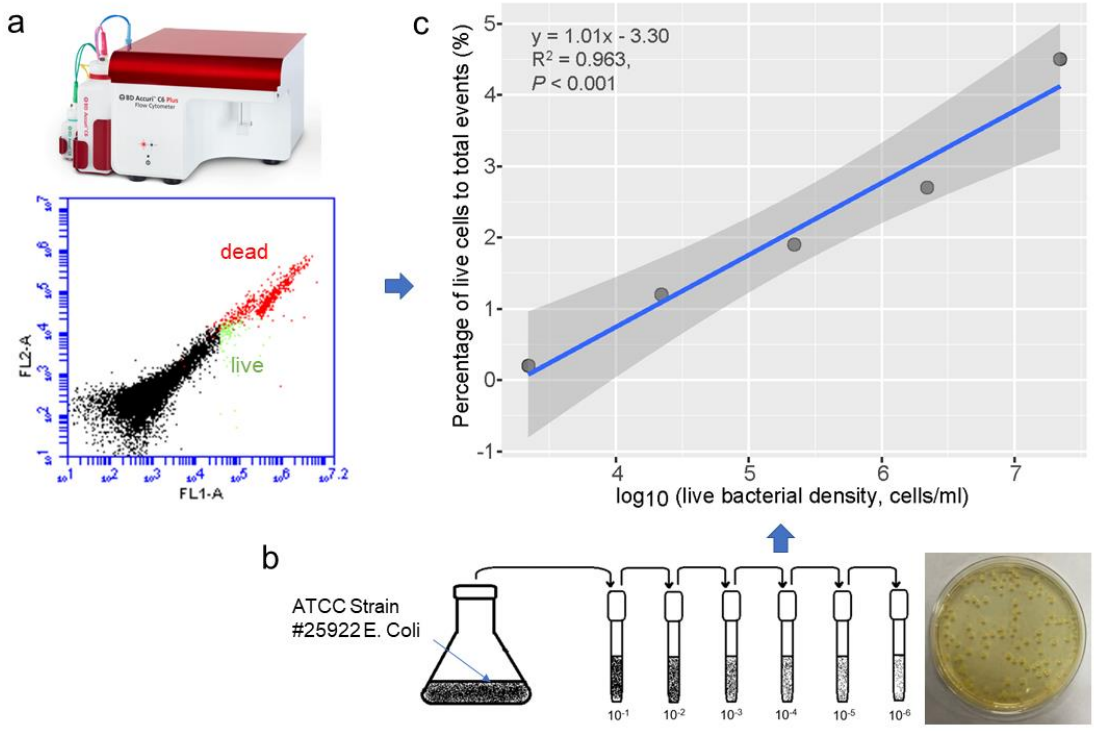


Figure 3-12 A FCM method for bacterial quantification.

3.4.6 Quantification of MAGs and potential hazards

The quantification of the constructed MAGs (genome copies per million reads) was conducted by the Quant_bin module of metaWRAP (Uritskiy *et al.*, 2018) and was presented in genome copies per million sequencing reads. The cell density of the MAGs (copy/m^3-air) in each sample was calculated according to Equation 3-6 (Liang *et al.*, 2020b):

$$cell\ density\ of\ MAGs = \frac{Absolute\ cell\ count \times Map.r \times Ab.i}{\sum_{1}^{n=i} Ab.i}$$

Equation 3-6

where *Map.r* represented the percentage of clean sequencing reads that have been mapped onto the MAG reference (*i.e.*, BAM files) by using Bowtie2 (Langmead and Salzberg, 2012) in each sample. *Ab.i* indicated the relative abundance of each MAG in each sample, and i referred to the number of MAGs.

The coverages of intragenomic ARGs in the host MAGs (ARG-cov/ppm) were quantified according to Equation 3-7:

$$intragenomic\ ARGs = Abundance_{MAG} \times Coverage_{ARG}$$

Equation 3-7

where $Abundance_{MAG}$ represented the relative abundance of the constructed MAG quantified by Quant_bin module. $Coverage_{ARG}$ referred to the mean coverage of ARGs located on the scaffolds of MAGs and was calculated using the built-in pipelines of BBmap (version 38.96) in default.

3.4.7 Source tracking of global airborne ARGs

The database of putative sources of airborne ARGs, which contained 700 metagenomes representing different sample types across the globe, was retrieved from NCBI (Details information in Appendix 3.2). To map the profile of antibiotic resistomes with 70% identity, the unified standard pipeline employed by the short-read module of DeepARG was

followed(Arango-Argoty *et al.*, 2018). The processed airborne ARG table was merged with the acquired ARG table, including samples collected from various habitats globally, including soil, ocean, freshwater, drinking water, plants, human-associated habitats, farms, WWTPs, and hospitals. The NMDS analysis was performed based on the Bray–Curtis dissimilarity matrix for the comparison of ARG profiles across habitats. Additionally, the derived ARG table was input to estimate the proportion of airborne ARGs in each sample attributable to various habitats by using "SourceTracker" (Knights *et al.*, 2011). This program uses Bayesian methods to evaluate all assignments of sink sequences (ARG sequences in air samples in this case) to all source samples, including an unknown source, and generates a joint distribution of those assignments.

Chapter 4 Structure and Distribution of Bacterial Communities in the Global Atmosphere and Interconnection with Earth's Microbiomes

The present study presents a comprehensive overview of both airborne bacterial communities across the globe. To achieve this, air samples were newly collected, and data from reliable studies were incorporated. The sampling locations encompassed a range of altitudes and geographic regions, from ground level to rooftops and high mountains exceeding 5,000 m asl, and from highly populated urban centers to the Arctic and Antarctic Circles. This approach allowed for a diverse representation of airborne bacterial communities and ARG profiles and enabled the use of genomic analysis to describe their community structure, biogeographic patterns, and ARG profiles in macro ecosystems. Our findings indicate that the complexity and dynamics of airborne communities are comparable to bacterial assemblages in soil and ocean environments and are strongly linked to local environmental conditions.

4.1 Structure of global airborne bacterial communities

The study generated a global airborne bacterial dataset comprising 370 air samples covering 63 sites worldwide. and the dataset was analyzed to construct an airborne bacterial reference catalog that contained over 27 million nonredundant 16S rRNA gene sequences. This study revealed the presence of 10,897 taxa across the samples, with the most bacterial sequences belonging to phyla (and sub-phyla), Firmicutes (24.8), Alpha-proteobacteria

(19.7), Gamma-proteobacteria (18.4), Actinobacteria (18.1), and Bacteroidetes (8.6) (Fig. **4-1**). While the relative abundance of different taxa varied significantly, the average abundance of each taxon was consistent with its highest abundance in one sample, emphasizing the differential among samples was less prominent than variations of bacterial abundance. Thereby, this chapter primarily focused on the global airborne microbiomes from a macroecological perspective.

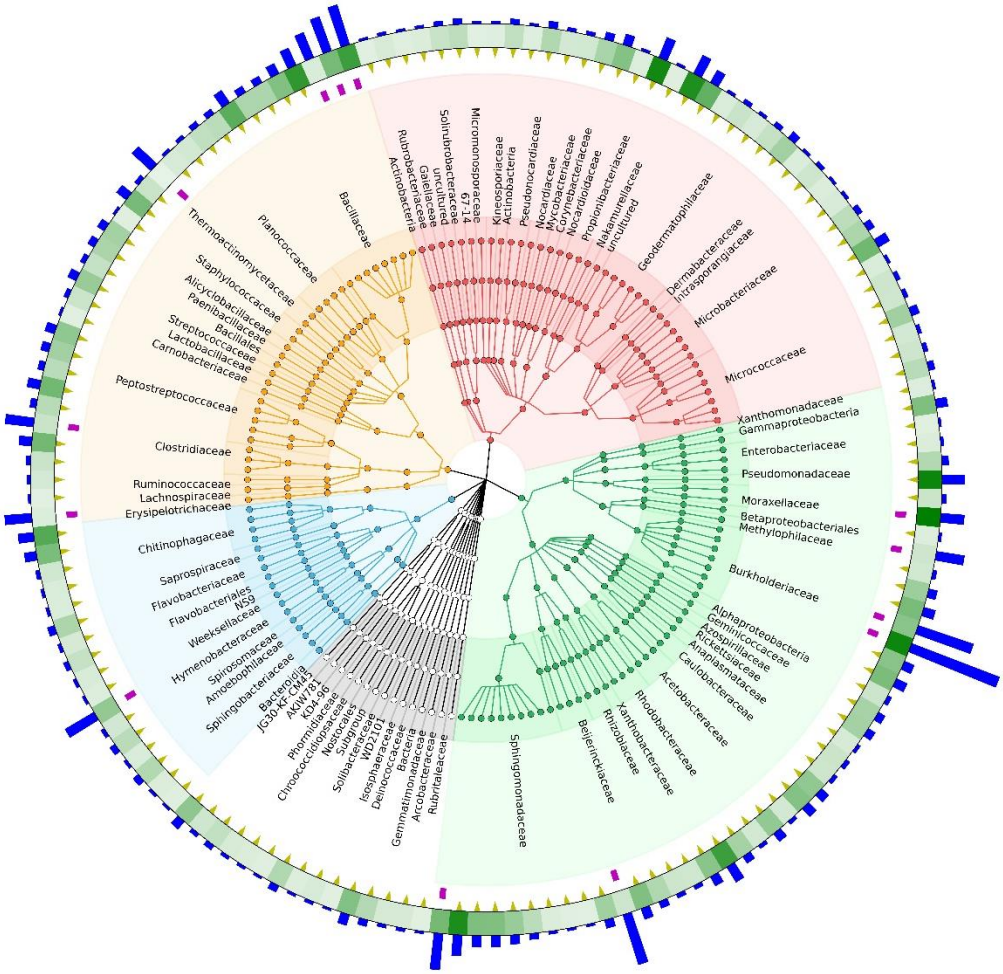


Figure 4-1 Phylogenetic tree of dominant airborne bacterial OTUs. The center is a phylogenetic tree of bacteria abundance in the global atmosphere. The middle ring corresponds to body sites at which the various taxa are abundant. The OTUs are color-coded at the phyla level. The majority of airborne bacteria belonged to four phyla: Proteobacteria (green), Firmicutes (yellow), Actinobacteria (red), and Bacteroidetes (blue). In the external middle ring, the relatively abundant bacteria (mean relative abundance ≥ 1) are indicated by purple rectangles, and the relatively inadequate bacteria (mean relative abundance < 1) are indicated by yellow triangles. The heights of blue bars outside the circle correspond to the abundance of taxa at the body site of greatest prevalence.

4.1.1 Global core airborne bacterial community

Species abundance and distributions are two fundamental questions in ecology, and the two key variables are often positively correlated (Gaston *et al.*, 2000), which has been well documented in various environments across a wide range of taxa (Shade and Stopnisek, 2019; Stopnisek and Shade, 2021), however, little is known about the such pattern in the global airborne bacterial communities. Here, a Sigmoid curve was observed in the abundance-occupancy relationship (AOR) between the average abundance of a bacterial taxon within the global air and the number of samples it occupies. This finding represents a significant expansion of macroecological theory to a new environment and offers great potential for further research and development of mechanistic macroecological theories (Shade and Stopnisek, 2019).

The AOR concept was employed to identify the core subset of bacteria in the atmosphere, which consists of abundant and widely-distributed bacteria indicated by both high abundance and high occupancy (Stopnisek and Shade, 2021) (Figure **4-2**a). A hyper-dominant pattern was observed globally (Shade and Stopnisek, 2019), with 24 OTUs (0.22 of the total number of OTUs) accounting for 18.5% of the total detected sequences (Figure **4-2**b, Table **4-1**). These exceptionally abundant taxa could dominate and strongly influence the whole communities in ecology (Fauset *et al.*, 2015). Similar to the profile of the whole global airborne bacterial community, most (92%) of the core community members belonged to Proteobacteria (*n=9*), Firmicutes (*n=7*), and Actinobacteria (*n=6*) (Figure **4-2**c). The most abundant OTU, accounting for 1.90 ± 0.39 of the sequence abundance in air samples with nearly 99% occupancy in all samples, belonged to Methylobacterium, a dominant genus ubiquitously distributed in the air (Weon *et al.*, 2008). The second most abundant OTU with 1.79.10 ± 2.68 relative abundance and occurring in roughly 93% air samplers, belonged to *Bacillus*, which has been detected as common predominant species in the air by both culture-dependent and culture-independent methods (Yoo *et al.*, 2019) and shows a wide range of physiologic abilities to live in every natural environment (Christie and Setlow, 2020). The core communities in marine and topsoil habitats based on the global datasets (Sunagawa *et al.*, 2015; Bahram *et al.*, 2018) were also identified, and no overlaps within the three largest ecosystems revealed a unique core community in each ecosystem (Table **4-2**).

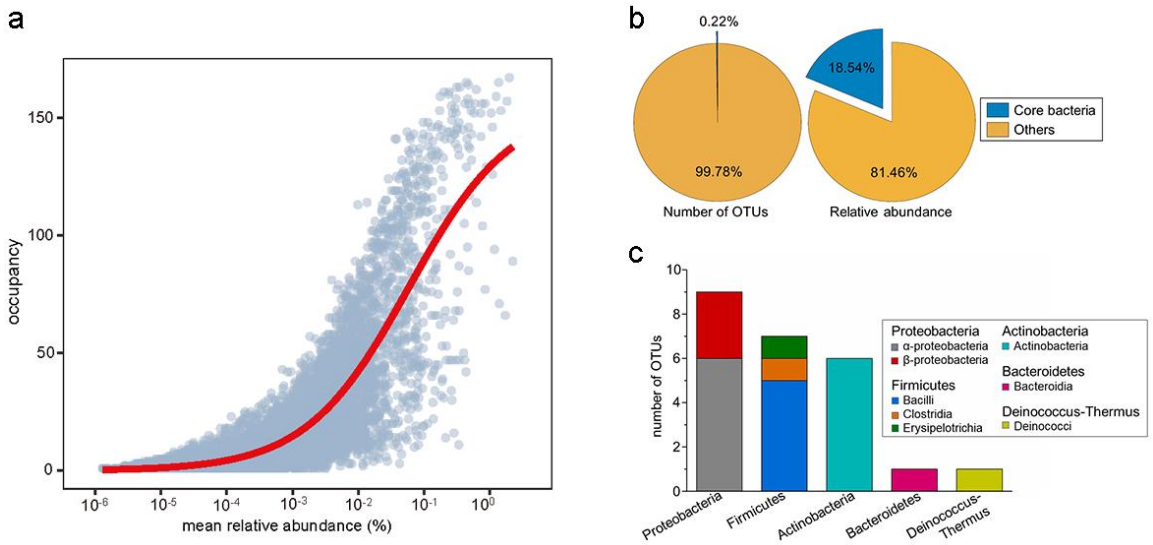


Figure 4-2 Identification of core subset in the global airborne bacterial community. (a) AOR: mean relative abundance (x-axis) and occupancy (y-axis) plot after combining the OTUs with the same annotation (n = 10,897). The mean relative abundances were estimated by averaging the relative abundances of each OTU in all samples; occupancy represents the number of samples in which the OTU was detected. The fitted model (sigmoid curve) was occupancy versus logarithm of abundance for each species, and the red solid line is the global fit to all species and samples. (b) The number proportion and relative abundance of the global core OTUs compared with those of the remaining bacterial OTUs. (c) The taxonomic composition of the global core bacteria at the phylum and class level. More specific taxonomic information of 24 global airborne core bacteria can be found in **Table 4-1**.

Table 4-1 The taxonomic composition of 24 global airborne core bacteria (phylum, class, order, family, and genus level).

OTU number	Mean abundance (0.01)	Phylum	Class	Order	Family	Genus
OTU4	204.75	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Methylobacterium*
OTU28	153.25	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sphingomonas*
OTU25	150.75	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*
OTU221	142.00	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	*Hymenobacter*
OTU14	114.00	Firmicutes	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	*Turicibacter*
OTU18	107.98	Firmicutes	Bacilli	Bacillales	Thermoactinomycetaceae	*Thermoactinomyces*
OTU6	96.16	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*
OTU22	87.82	Actinobacteria	Actinobacteria	Frankiales	uncultured	*uncultured_ge*
OTU7	77.39	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*
OTU36	74.88	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rubellimicrobium*
OTU167	63.12	Firmicutes	Bacilli	Bacillales	Bacillaceae	*unclassified*
OTU12	60.82	Actinobacteria	Actinobacteria	Frankiales	Geodermatophilaceae	*Geodermatophilus*
OTU41	54.44	Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	*Terrisporobacter*
OTU32	53.83	Actinobacteria	Actinobacteria	Micrococcales	Microbacteriaceae	*unclassified*
OTU17	50.39	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	*Staphylococcus*
OTU38	48.42	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*1174-901-12*
OTU410	48.07	Deinococcus-Thermus	Deinococci	Deinococcales	Deinococcaceae	*Deinococcus*
OTU94	47.88	Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	*Nocardioides*
OTU115	46.71	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Craurococcus*
OTU85	45.67	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	*unclassified*
OTU60	32.34	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*unclassified*
OTU189	31.95	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	*Paenibacillus*
OTU90	30.92	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	*Corynebacterium_1*
OTU159	30.20	Actinobacteria	Actinobacteria	Pseudonocardiales	Pseudonocardiaceae	*Pseudonocardia*

Table 4-2 Taxonomic information and mean relative abundance (0.01%) of core communities in air, marine, and soil habitats based on the global datasets (Sunagawa *et al.*, 2015; Bahram *et al.*, 2018).

Global air		Global marine		Global soil	
taxonomy	**abundance**	**taxonomy**	**abundance**	**taxonomy**	**abundance**
g_*Methylobacterium*	190.34	f_Surface 1	1784.32	g_*Subgroup_6_ge*	417.06
g_*Bacillus*	179.10	f_SAR86 clade	853.58	g_*Candidatus_Udaeobacter*	339.68
g_*Sphingomonas*	153.25	o_SAR11 clade	789.24	f_Xanthobacteraceae	215.98
g_*Hymenobacter*	142.00	g_*Prochlorococcus*	747.39	g_*Acidothermus*	191.05
g_*Turicibacter*	114.00	g_*AEGEAN-169*	331.23	f_Gemmataceae	175.25
g_*Thermoactinomyces*	107.98	f_Rhodobacteraceae	318.25	f_Xanthobacteraceae	164.15
g_*Acinetobacter*	96.16	f_SAR116 clade	283.27	g_*RB41*	156.39
o_Frankiales	87.82	f_OCS155 marine group	269.20	g_*67-14_ge*	155.74
g_*Pseudomonas*	77.39	f_SAR406 clade	266.45	g_*uncultured_ge*	137.16
g_*Rubellimicrobium*	74.88	f_Marine Group II	228.16	f_Xanthobacteraceae	127.60
f_Bacillaceae	63.12	g_*Candidatus Pelagibacter*	225.97	o_Acidobacteriales	118.46
g_*Geodermatophilus*	60.82	f_Surface 2	212.47	g_*WD2101_soil_group_ge*	115.17
g_*Terrisporobacter*	54.44	g_*NS5 marine group*	130.67	g_*Acidibacter*	114.52
f_Microbacteriaceae	53.83	g_*NS4 marine group*	115.99	g_*KD4-96_ge*	111.71
g_*Staphylococcus*	50.39	g_*Marinicella*	80.53	o_Gaiellales	110.52
g_*1174-901-12*	48.42	f_Surface 4	73.32	g_*Mycobacterium*	100.56
g_*Deinococcus*	48.07			o_Gaiellales	99.99
g_*Nocardioides*	47.88			g_*Candidatus_Solibacter*	80.32
g_*Craurococcus*	46.71			g_*Bryobacter*	80.02
f_Burkholderiaceae	45.67			g_*Reyranella*	70.93
f_Acetobacteraceae	32.34			c_Alphaproteobacteria	68.74
g_*Paenibacillus*	31.95			g_*Haliangium*	67.63
g_*Corynebacterium_1*	30.92			g_*Solirubrobacter*	65.08
g_*Pseudonocardia*	30.20			g_*IMCC26256_ge*	61.50
				g_*Chthoniobacter*	61.36
				g_*TK10_ge*	51.65
				g_*MND1*	49.24

4.1.2 Networked global airborne bacterial community and key taxa identification

Microorganisms do not exist in isolation but generate complex ecological interrelationship webs on the Earth (Faust and Raes, 2012), and thus a global airborne community co-occurrence network was constructed to explore the interconnections (Figure **4-3**). The network encompassed 5,038 significant correlation relationships (Spearman's $\rho > 0.6$, $p < 0.01$) among 482 connected OTUs (around 21 edges per node, Table **4-2**). Similarly, the global interrelationship webs of bacterial communities were also established based on global topsoil and marine datasets (Figure **4-3**). In general ecology, real networks, including biological networks, have been proven to have the "small-world" property (Zhang and Zhang, 2009b), which means that individuals are more connected to each other than in a random network. As the first attempt to construct an airborne bacterial network, the "smallworldness" index was also computed relying on the global transitivity of the network and its average shortest path length (Humphries and Gurney, 2008). The "smallworldness" index of the airborne bacterial community (0.51) was less than 1, which indicated that the global airborne bacterial community network was not a "small-world" network (Figure **4-4**). Conversely, the other two global bacterial datasets on topsoil and marine ecosystems both met the properties of a "small-world" network (soil "smallworldness" index=5.82> 3 for a stricter rule; marine "smallworldness" index =1.21> 1 for the general rule). The "smallworldness" index of the bacterial community network showed a decreased gradient from soil, marine, to air habitats, which was consistent with the variations in other topological properties (Table **4-3**), such as the average shortest path length (3.03< 3.97< 5.24), diameter (9< 10<15), and clustering coefficient (0.48<0.58<0.67). All the findings indicated that the clustering approach in the airborne

bacterial community appeared to be more random, and the topological characteristics showed low resistance to changes, such as the loss of nodes (airborne bacteria). This could be interpreted as decreasing the speed of the response of the network to perturbations, finally leading to a less stable community network structure. Hence, the identified remote associations and dispersed clusters in the network indicate that the airborne bacterial community is prone to disruption due to environmental factors that typically result in significant alterations in bacterial makeup.

Based on the co-occurrence network, node-level topological features, including degree, betweenness centrality, and closeness centrality (Figure **4-3**) were examined to represent the different functions and roles of each node (bacterium) in the whole microbial communities. Additionally, the validated power-law degree distribution of nodes stated that the overall network roughly conformed to a scale-free degree distribution (Figure **4-3**). This meant that the majority of bacteria showed low-degree values, while only a small number of hub nodes exhibited high-degree values (Barabási and Albert, 1999). Moreover, 19 clustering hub nodes were identified among these interlaced nodes that functioned as the root of a power-law degree distribution network, which was based on the acknowledged criteria with high degree (>100) and low betweenness centrality values (<5,000) in co-occurrence networks (Banerjee *et al.*, 2018) (Figure **4-5**a and Table **4-4**). These hub nodes showed a tightly clustered distribution with a mean correlation coefficient of 0.903, and each of them displayed significant correlations with 15-18 of the OTUs in the entire network (Figure **4-5**b), indicating the presence of a tightly connected community that may be a crucial module in the global network. The key taxa within this community exerted a dominant influence on the overall topological characteristics of the network. Due to their

significantly higher connection efficiency, these hub nodes may be considered key taxa that played a vital role in maintaining the structure of a microbial community relative to their abundance (Banerjee *et al.*, 2018).

The structure of the airborne bacterial community exhibited significant differentiation from those in other ecosystems. Consequently, a core set of 24 bacteria and 19 key taxa were exclusively determined in such unique and huge airborne bacterial communities (Tables **4-1** and **4-4**). In addition, OTU22, OTU94, and OTU159 were found to be both core bacteria and key taxa (Figure **4-5**b and Figure **4-5**c), indicating their crucial roles in the whole community due to hyper dominance and strong connections with other members, and thereby recognized as the top three essential species. Importantly, all three species belong to the same phylum, Actinobacteria, and are gram-positive bacteria.

Furthermore, the comparison of the community composition structures of key, core, and all OTUs was constructed. Firmicutes, known for their ability to produce endospores that withstand extreme desiccation and survive in extreme conditions (Wunderlin *et al.*, 2016), are one of the most widespread phyla in airborne bacterial communities (comprising 24.8% of whole communities) However, none of Firmicutes were identified as key taxa due to their limited interconnections with communities. On the other hand, Actinobacteria, which did not exhibit overwhelming abundance (18.1% in whole communities and 19.6% in the core set), showed a close association with the whole airborne bacterial communities and even surpassed Proteobacteria. This was evident from the finding that Actinobacteria showed a great proportion (72.2%) in key taxa (Figure **4-5**d and Table **4-4**). In summary, the composition of the global airborne bacterial community was comparable to that of the core set, while both were markedly different from key taxa, which exhibited substantial

variations.

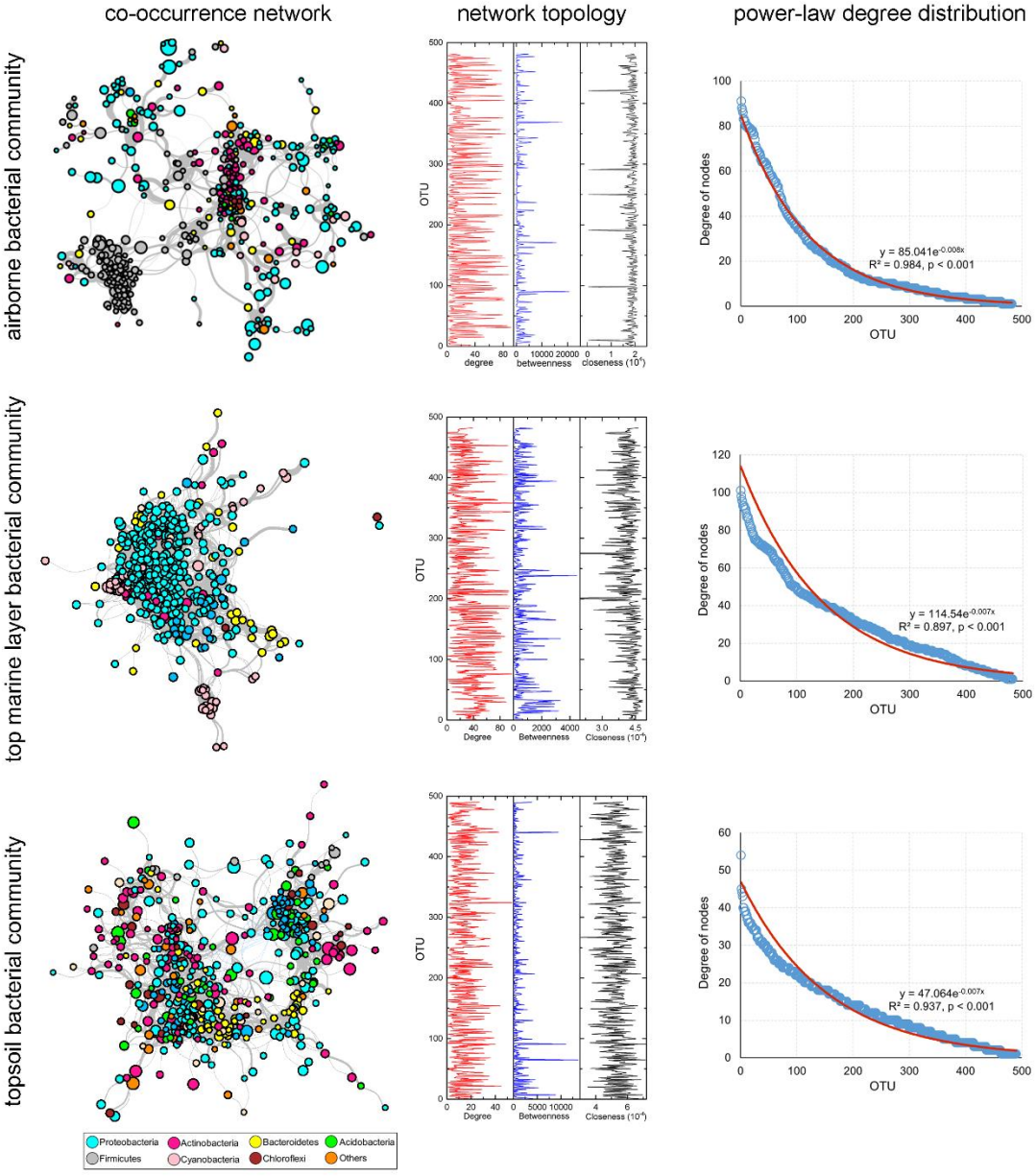


Figure 4-3 Networked bacterial communities in the atmosphere, topsoil, and top marine layer at the global scale. Co-occurrence network: The connection (edges) stands for a strong (Spearman's $\rho > 0.6$) and significant ($p < 0.01$) correlation. The nodes represent the

combined OTUs with a unique genus-level annotation in the datasets. The size of each node is proportional to the mean relative abundance across all samples. Nodes are color-coded according to the phyla of the bacteria. Network topology: The degree and centrality (betweenness and closeness) of each node from the networks were measured. Degree represents the number of direct connections of a node with other OTUs in the whole community. Betweenness centrality reveals the role of a node as a bridge between components of a network. Closeness centrality measures the average shortest distance from each node to each other node. Power-law degree distribution: The node-degree distribution shows a power-law behavior for an airborne bacterial community co-occurrence network in airborne bacterial communities (R^2 = 0.984, p < 0.001), marine layer bacterial communities (R^2 = 0.897, p < 0.001), and topsoil bacterial communities (R^2 = 0.937, p < 0.001), respectively.

Table 4-3 The topological properties of the co-occurrence network.

Topological property	Air	Soil	Marine
number of nodes	482	490	482
number of edges	5,038	3,323	7,629
number of clusters	3	1	2
average shortest path length	5.24	3.03	3.97
diameter	15	9	10
clustering coefficient	0.67	0.48	0.58

Notes (the definition and/or functions of each topological property):

1. number of nodes: interconnected objects (bacteria)

2. number of edges: links between each pair of nodes

3. number of clusters: groups of objects with tight interconnections

4. average shortest path length: average network distance between all pairs of nodes

5. diameter: maximum of pairwise distances between every two nodes

6. clustering coefficient: the degree to which nodes tend to cluster together.

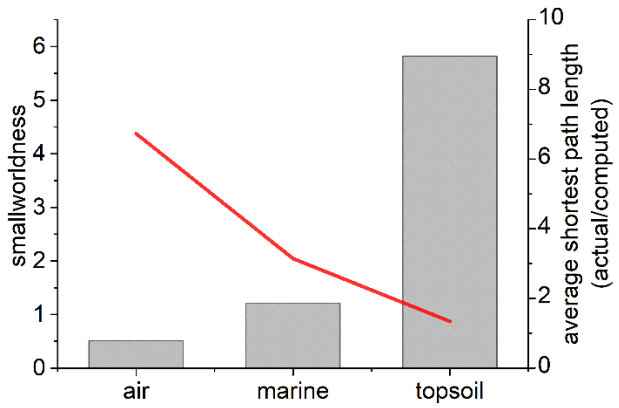


Figure 4-4 Identification of "small-networks" was performed using a "small-worldness" index and the average shortest path length of the bacterial community network in the global air, marine, and topsoil environments.

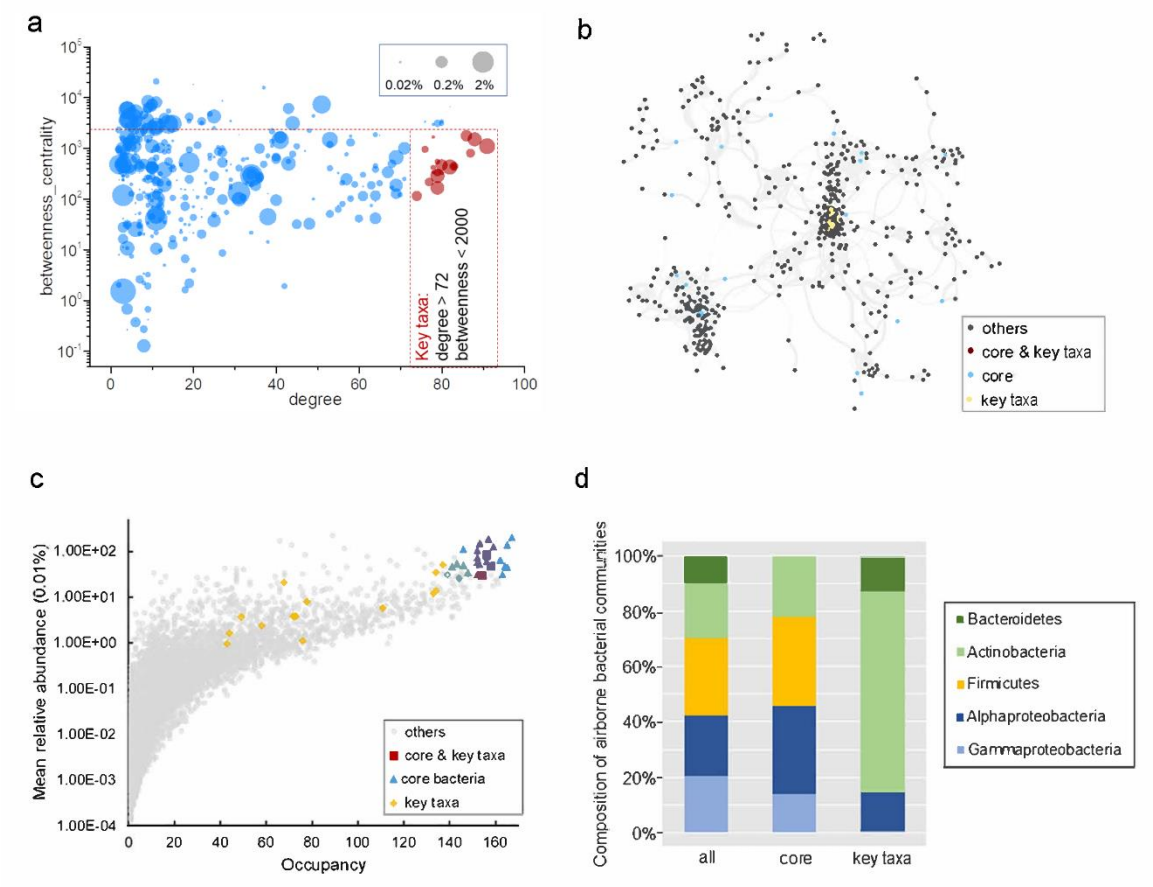


Figure 4-5 Role and composition of the core bacterial communities and key taxa in the atmosphere on a global scale. (a) Plot of degree and betweenness centrality for each node in the co-occurrence network. The nodes colored in red are viewed as key taxa. The size of the nodes corresponds to the relative proportions of the OTUs in the total microbiome. (b) The position and role of core and key bacterial communities in a community network and (c) AOR pattern. (d) The taxonomic composition of bacterial communities at the phylum level (class level for Proteobacteria) for whole communities, core bacteria, and key taxa.

Table 4-4 The taxonomic composition of 19 global airborne key taxa (phylum, class, order, family, and genus level).

OTU number	Mean abundance (0.01)	Phylum	Class	Order	Family	Genus
OTU94	47.88	Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	Nocardioides
OTU22	87.82	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured_ge
OTU40	8.27	Actinobacteria	Actinobacteria	Micromonosporales	Micromonosporaceae	unclassified
OTU125	1.70	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Citricoccus
OTU80	3.87	Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	Marmoricola
OTU158	2.51	Actinobacteria	Actinobacteria	Kineosporiales	Kineosporiaceae	unclassified
OTU19	22.16	Actinobacteria	Actinobacteria	Frankiales	Geodermatophilaceae	Blastococcus
OTU284	1.01	Actinobacteria	Actinobacteria	Micrococcales	Bogoriellaceae	Georgenia
OTU159	30.20	Actinobacteria	Actinobacteria	Pseudonocardiales	Pseudonocardiaceae	Pseudonocardia
OTU372	12.27	Actinobacteria	Thermoleophilia	Gaiellales	Gaiellaceae	Gaiella
OTU721	13.63	Actinobacteria	Thermoleophilia	Gaiellales	uncultured	uncultured_ge
OTU290	26.05	Actinobacteria	Thermoleophilia	Solirubrobacterales	67-14	67-14_ge
OTU133	51.66	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Segetibacter
OTU342	5.87	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Flavisolibacter
OTU144	4.04	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	unclassified
OTU55	30.65	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Microvirga
OTU63	3.71	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	uncultured
OTU35	35.30	Proteobacteria	Alphaproteobacteria	Azospirillales	Azospirillaceae	Skermanella
OTU250	1.15	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	unclassified

4.1.3 Key taxa in airborne bacteria associated with evolutional and ecological functions

Moreover, this study revealed similarities among key taxa with regard to their compositions and inferred functions, across atmospheric, aquatic, and terrestrial ecosystems (Table **4-5**). This suggested potential associations between airborne bacterial communities and other surface microbial habitats.

According to the rich-gets-richer preferential attachment process of growth in a scale-free network (Barabási, 2009), the highly connected nodes, namely, the key taxa, could acquire more links, thereby contributing to the establishment of the entire microbial network. Therefore, the nodes of key taxa are recognized as initial components in networks. In evolutionary terms, this suggests that key taxa emerged earlier than other species, and their lineages may have a longer evolutionary history in microbial co-occurrence networks (Berry and Widder, 2014). This has important implications for exploring the origins of microbes in the atmosphere and other ecosystems. For example, the important key order, Frankiales (OTU19 and OTU22 in Table **4-4**), was thought to have an adaptable ancestral bacterium that evolved to occupy a variety of ecological niches, including the root nodules of woody dicots, hot springs, rocky surfaces, gamma-irradiated substrates, activated sludge, compost, and soils (Parte *et al.*, 2012). Thus, Frankiales could be assumed to be the crucial ancestor of many airborne bacteria.

The concept of key taxa was first developed in the context of food-web ecology (Paine, 1966), where interspecies relationships, including mutualism, commensalism, parasitism, competition, and others, were mainly viewed in terms of trophic relationships. Accordingly, the functions of most key taxa were related to nutrition and metabolism, as was reflected in this study. For instance, Rhizobiales (OTU63 and OTU55 in Table **4-4**) serve a unique physiological function in biological

nitrogen-fixation and have been recognized as the key player in global nitrogen cycling (Andreote *et al.*, 2009). Similarly, Burkholderiaceae (OTU250 in Table **4-4**) produce secondary metabolites that significantly affect microbial interconnections within the network (Nazaries *et al.*, 2013), and Gaiellales (OTU372 and OTU721 in Table **4-4**) utilize several organic compounds (Albuquerque and da Costa, 2014) and play an important role in the overall nutrition cycling process of microbial communities. The activities and abundance of these individual populations, key taxa, have been shown to profoundly affect the integrity and stability of the overall community over time (Cottee-Jones *et al.*, 2012).

Moreover, for insight into the functions of key taxa, the microbial key taxa across various ecosystems as reported from other literature were also summarized (Table **4-5**). It is noteworthy that Rhizobiales and Burkholderiales orders contain diverse members that have consistently been recognized as key taxa in a range of ecosystems and habitats, including soil, aquatic systems, and environments from equatorial to polar habitats (Table **4-5**). Rhizobiales for example, not only comprise nitrogen-fixing bacteria (*Rhizobium* spp. and *Bradyrhizobium* spp.), but also methanotrophs (*Methylobacterium*) with high abundance in the phyllosphere (Andreote *et al.*, 2009). In addition, Burkholderiales encompasses several well-known pathogenic bacteria, including species of *Bordetella, Ralstonia, Oxalobacter,* and *Burkholderia,* one of the most versatile and widespread terrestrial microbial groups (Nazaries *et al.*, 2013). Despite being recognized as playing an important role in key taxa in the natural world, not all members of Rhizobiales and Burkholderiales could be considered as key taxa through computational identification. Plenty of subordinate taxa in the two orders had no significant impacts on community composition or function. However, their frequent roles of key taxa across diverse habitats could be attributed to their large abundance and wide occupancy in natural environments.

Nevertheless, there is a possibility that members of Rhizobiales and Burkholderiales could be identified as key taxa, and it is important for future studies to evaluate their roles in microbial functions and interactions.

Table 4-5 Summary of key taxa reported in the literature on different ecosystems and habitats.

Ecosystem or habitat		Key taxa	Reference
Aquatic ecosystems	shallow lake	Cyanobacteria	(Zhao *et al.*, 2016)
	freshwater sediment	Planctomycetes, Proteobacteria, Nitrospira, Clostridia, Actinobacteria, Bacilli	(Ji *et al.*, 2016)
	nearshore	Verrucomicrobiaceae, Crenarchaeaceae, Candidatus OP3, Chloroflexi,Chloracidobacteria	(Graham *et al.*, 2017)
	seawater_cluture	Rhodobacteraceae, Alteromonadaceae	(Geng *et al.*, 2016)
Forests	Brazilian forest	Burkholderiales, Acidimicrobiaceae, Rhodospirillaceae	(Lupatini *et al.*, 2014)
	Chinese forest	Rubrobacteriales, Gaiellales, Rhizobiales, Gaiellales, Solirubrobacteriales	(Ma *et al.*, 2016)
Agricultural lands	soybean soil	Rhizobiales, Methylophilaceae, Acetobacteraceae	(Lupatini *et al.*, 2014)
	sugarcane soil	Burkholderiales, Caulobacter, Kineosporia, Sporichthya	(Lupatini *et al.*, 2014)
	blueberry soil	Frankiales, Pseudomonas, Burkholderia	(Jiang *et al.*, 2017)
	paddy soil	Rhodobiaceae, Hydrogenophilaceae, Comamonadaceae, Alcaligenaceae	(Wang *et al.*, 2017)
	Pinus plantation	Sporichthya, Burkholderiales	(Lupatini *et al.*, 2014)
Arctic and Antarctic ecosystems	Antarctic lake	Acidobacteriaceae	(Vick-Majors *et al.*, 2014)
	Arctic soil	Burkholderia, Bradyrhizobium, Rhodoplanes, Pseudomonas	(Hill *et al.*, 2016)
	permafrost	Burkholderiales, Actinomycetales, Rhizobiales, Puniceicoccaceae	(Comte *et al.*, 2016)
	permafrost soils	Burkholderia, Bradyrhizobium, Phenylobacterium	(Gokul *et al.*, 2016)
	Arctic ice cap	Actinobacteria	(Gokul *et al.*, 2016)

4.2 Biogeographic Distribution of Global Airborne Bacteria

The intermediate latitudinal regions showed the highest microbial diversity ($R^2 = 0.25$, $p < 10^{-15}$, Figure **4-6**), which was consistent with patterns observed in soil (Bahram *et al.*, 2018) and water (Sunagawa *et al.*, 2015), but differed significantly from the typical latitudinal gradient of diversity (LGD) pattern observed in macroscopic organisms (Lomolino and Brown, 2009). The key driving factors of latitudinal diversity patterns in other ecosystems are pH and temperature of soil (Fierer and Jackson, 2006; Bahram *et al.*, 2018), and salinity and temperature of water (Herlemann *et al.*, 2011; Sunagawa *et al.*, 2015). Because AT was directly relevant to latitude (Figure **4-7**a), temperature could be considered as the crucial factor driving the distribution of latitudinal diversity ($R^2 = 0.064$, $p < 0.005$, Figure **4-7**b). This finding is consistent with the role of AT in determining diversity in a vertical stratification study of airborne microorganisms (Du *et al.*, 2017). Thus, it could be hypothesized that temperature is the primary factor responsible for the uniform parabolic latitudinal diversity patterns observed in microbial worlds across the three largest ecosystems on Earth (*i.e.*, atmosphere, ocean, and terrestrial systems).

As a summary, the latitudinal diversity pattern of global airborne bacterial communities showed a clear trend of richness peaking at intermediate latitudes and declining towards the equator and the poles (Figure **4-6**). However, there were some deviations from this pattern, especially in the mid-latitudinal regions (35^o - 45^o). In order to explain these deviations, mid-latitudinal samples (*n = 64*) were further studied. It was discovered that the bacterial richness was strongly correlated with both PM_{10} concentration ($R^2 = 0.549$, $p < 1 \times 10^{-10}$) and $PM_{2.5}$ concentration ($R^2 = 0.517$, $p < 1 \times 10^{-7}$, Figure **4-8**), with richness being higher in moderately polluted air but lower in both good air quality and heavily polluted areas. This could be attributed to pollutants serving as a source of nutrients for the microbiome. In addition, populations are distributed unevenly across the globe, a

situation that is much more obvious in mid-latitudinal regions. In particularly developed cities and economic circles, populations are very dense, while deserts, polar regions, and high mountains are very sparsely populated. For instance, Hong Kong has a high population density of 6,544 individuals per square kilometer, which far exceeds that of Mt. Ailao located at a similar latitude. However, there was no significant difference in richness between the two sites. In conclusion, from a global perspective, airborne bacterial diversity follows a downward opening parabola-shaped latitudinal pattern and is hardly influenced by human distribution and the intensity of human activities.

Another widely used relationship in spatial biodiversity studies is the distance-decay relationship (DDR) (Morlon *et al.*, 2011), in which the similarity in species composition between two communities decreases with the increasing geographic distance separating them. Consistent with results in other domains (Martiny *et al.*, 2011), local environments generated a DDR concerning the similarities of the airborne bacterial communities across the globe ($R^2 = 0.13$, $p < 10^{-9}$, Figure **4-9**), despite atmosphere breaks geographic barriers and facilitates transport of over long distances (Burrows *et al.*, 2009b). Since the DDR reflected spatial distribution and autocorrelation of the global airborne bacterial community, it might be sensitive to crucial ecological processes, such as DL, and develop as an effective tool for testing mechanistic ecological theories during airborne bacterial community assembly (Condit *et al.*, 2002). Together, the findings align with the pronounced biogeographic patterns of atmospheric microbiomes observed in other ecosystems (Bahram *et al.*, 2018; Wu *et al.*, 2019).

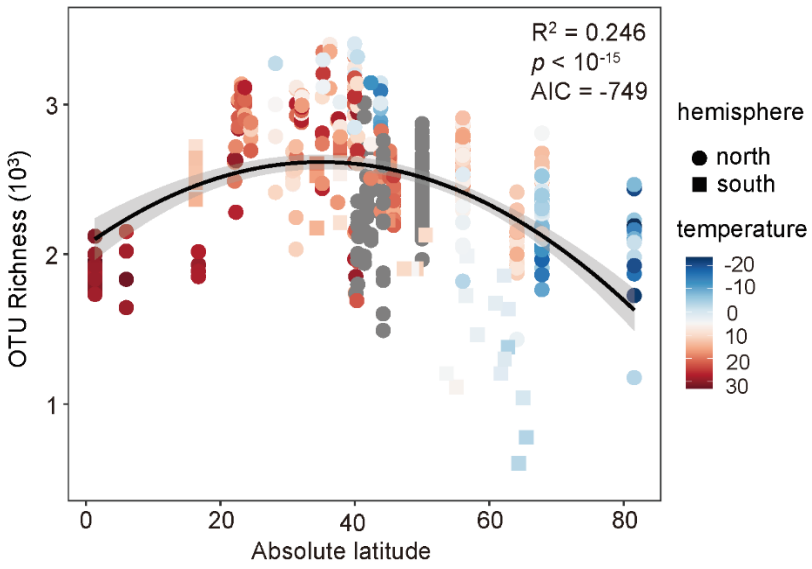


Figure **4-6** The distribution of airborne bacterial diversity across latitudes was analyzed using 455 biologically independent samples. A second-order polynomial fit based on ordinary least squares regression was used to determine the best fit for the datasets in this study. The corrected AIC was used to select the best polynomial fit. The results revealed that the maximum microbial diversity was observed in the intermediate latitudinal regions ($R^2 = 0.246$, $p < 10^{-15}$), with the color gradient indicating the corresponding AT for each sample. The symbols used to denote the origin of a sample, i.e., Northern Hemisphere (circle) or Southern Hemisphere (square), are also displayed.

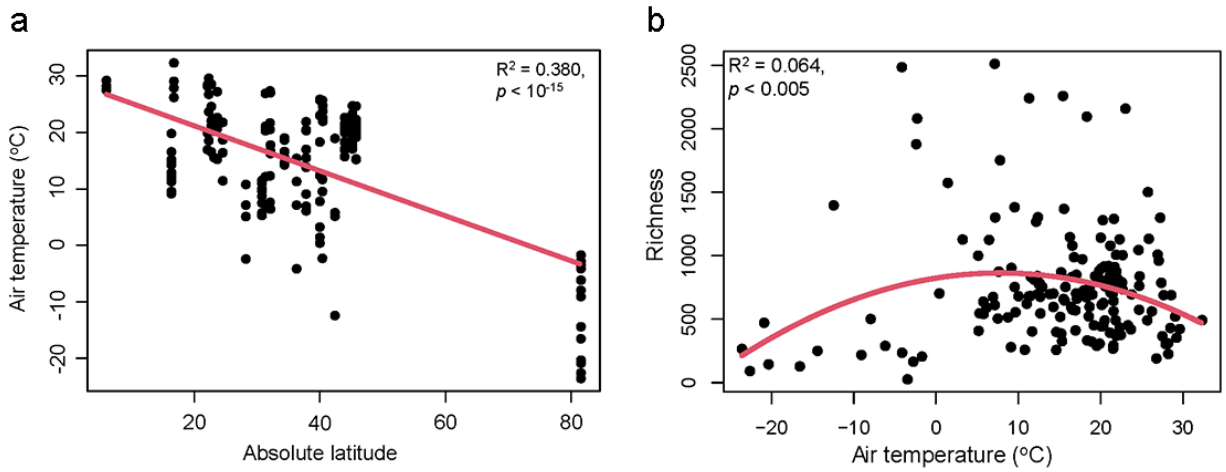


Figure **4-7** The role of AT in the latitudinal diversity pattern of global airborne bacterial community. (a) Direct relationship of AT with absolute latitude, namely distance to the equator. (b) Impact of AT on richness of airborne bacterial communities.

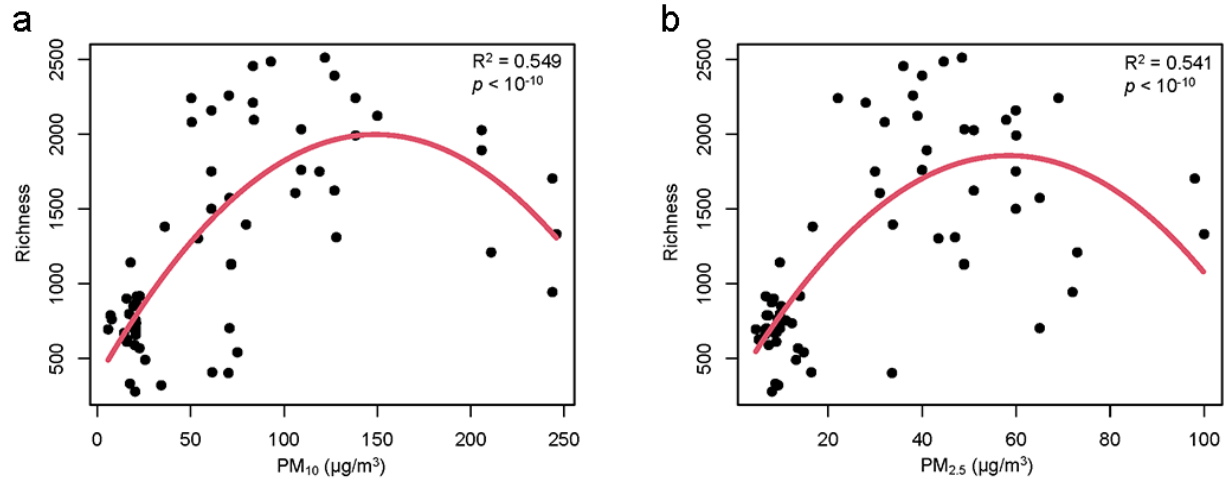


Figure **4-8** Airborne bacterial richness (*y-axis*) and PM concentration (*x-axis*) plots of samples collected in the mid-latitudinal regions (35° - 45°). The richness was strongly correlated with PM_{10}

concentration (R^2=0.549, $p < 1 \times 10^{-10}$) and $PM_{2.5}$ concentration (R^2=0.541, $p < 1 \times 10^{-10}$), explaining the large deviation from the fitting line of Figure **4-6** in mid-latitudinal areas.

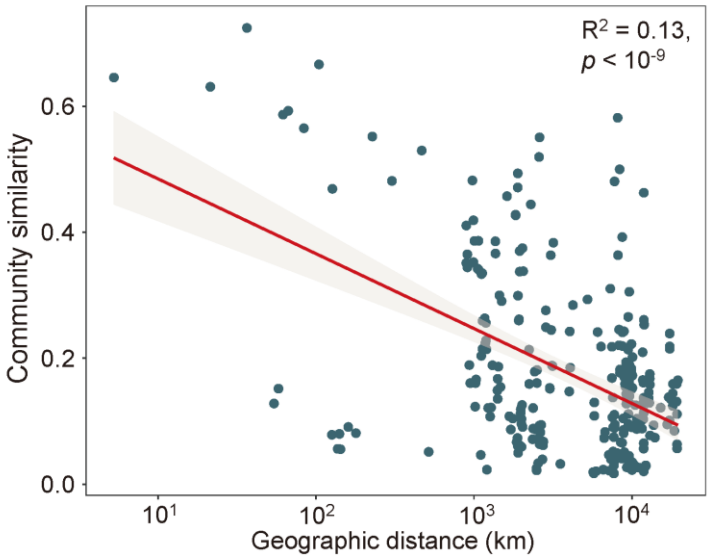


Figure **4-9** The Bray-Curtis pairwise similarity of microbial communities, which is based on the relative abundance of OTUs, increases as the geographic distance between sampling sites increases. The least-squares linear regression is represented by the red line.

4.3 Global Airborne Bacteria Linked with Other Habitats

Bacterial co-occurrence networks in specific environments, such as global atmosphere, have been elaborated to explore these complex systems in section 4.1.2 and previous studies, but their interconnections across microbiomes in various habitats, as well as the role of airborne bacteria in the Earth's microbial world, have yet to be explored. Furthermore, the uniform biogeographic pattern and shared key taxa in the three largest ecosystems, namely air, marine, and soil ecosystems, suggest interrelationships among bacterial communities in various habitats.

4.3.1 Bacterial biomass and richness in the troposphere and other typical Earth habitats

Estimation of species numbers in different global ecosystems can provide insight into the commonness and rarity of taxa and the connections across scales of space, time, and abundance (Curtis *et al.*, 2002; Locey and Lennon, 2016). Using the lognormal model, microbial richness was predicted based on the total abundance of individuals (N) and the quantity of the most dominant taxonomic unit (N_{max}) according to all available data (Locey and Lennon, 2016). Despite the total abundance of global airborne bacteria (1.72×10^{24} cells) being 1-3 orders of magnitude lower than other habitats, such as soil (9.36×10^{28} cells), freshwater (4.70×10^{25} cells), and marine (4.68×10^{28} cells), the estimated bacterial richness in the atmosphere (4.71×10^{8} ~ 3.08×10^{9}) was comparable to that in the hydrosphere (Figure **4-10**). Since the atmosphere is less favorable to microorganisms than surface habitats, the comparable diversity and complexity of aerial microbial communities suggest that they receive contributions from surrounding ecosystems, indicating interconnections with microbiomes in surface ecosystems.

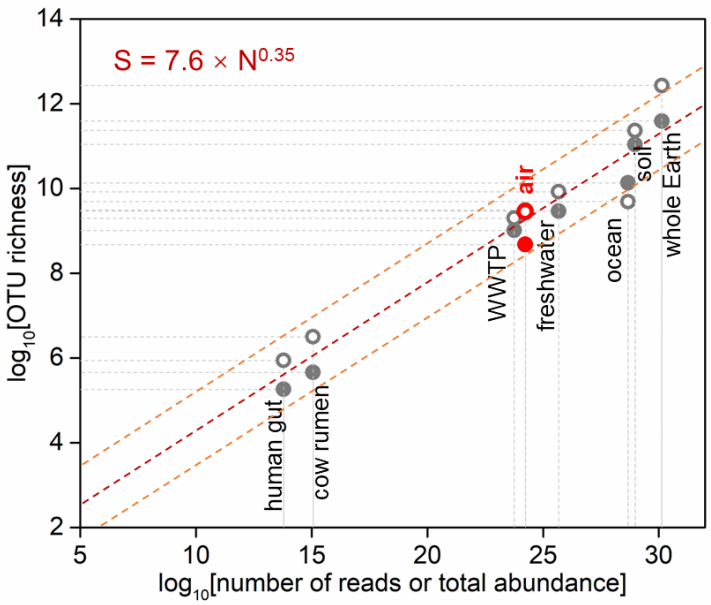


Figure **4-10** The estimation of global microbial abundance and richness in various habitats. The global richness (*S*) and the total abundance (*N*) in the corresponding habitats show a scaling relationship (the dashed orange line is the 95% prediction interval). Richness was predicted using the lognormal model, with N_{max} derived either from our sequencing data (filled circles) or from the dominance-scaling law (open circles). It should be noted that the estimated *S* and *N* for each habitat represent a global sum, which includes data from previous studies (Locey and Lennon, 2016; Wu *et al.*, 2019).

4.3.2 Interconnections of airborne bacteria with their counterparts in other habitats

The bacterial co-occurrence network in the global atmospheric environment has been deeply explored, which revealed that microbial interconnections could affect the structure and function of microbial communities (Figure **4-3**). First, NMDS was employed to compare the level of similarity among bacterial communities in 23 major habitats on the Earth, using 5,000 samples obtained from

the EMP (Thompson *et al.*, 2017). Interestingly, these habitats were not entirely distinct and exhibited varying degrees of overlaps (connections) with each other. Moreover, terrestrial air exhibited a greater similarity with environments associated with humans and animals, while offshore air bored a closer relation to oceanic systems (Figure **4-11**). Additionally, each type of atmosphere (urban, background, and marine surface) was found to harbor a unique microbiome compared to others, prompting to divide air samples into three distinct groups for more in-depth investigation.

To further examine the interconnections between airborne bacteria and their counterparts in other environments, a bacterial co-occurrence network for Earth was constructed using the hierarchical agglomeration algorithm (Park and Bader, 2011). The resulting network categorized the 23 habitats into three groups: human- and animal-associated environments (*Group I*), terrestrial natural environments (*Group II*), and aquatic environments (*Group III*) (Figure **4-12**). The network demonstrated gradual transitions and connections among the different environments, with marine systems linked to freshwater systems, which were in turn connected to soil and rhizosphere habitats, and finally to human- and animal-associated environments. Notably, the airborne bacterial communities appeared to have close associations with their surrounding environments, with more pronounced influences observed in settings with larger surface areas of contact with air (Figure **4-12**), such as seawater ($\rho = 0.70$, $p < 0.01$), animal surfaces ($\rho = 0.72$, $p < 0.01$), and human surfaces (nasopharyngeal: $\rho = 0.71$, $p < 0.01$; skin: $\rho = 0.75$, $p < 0.01$).

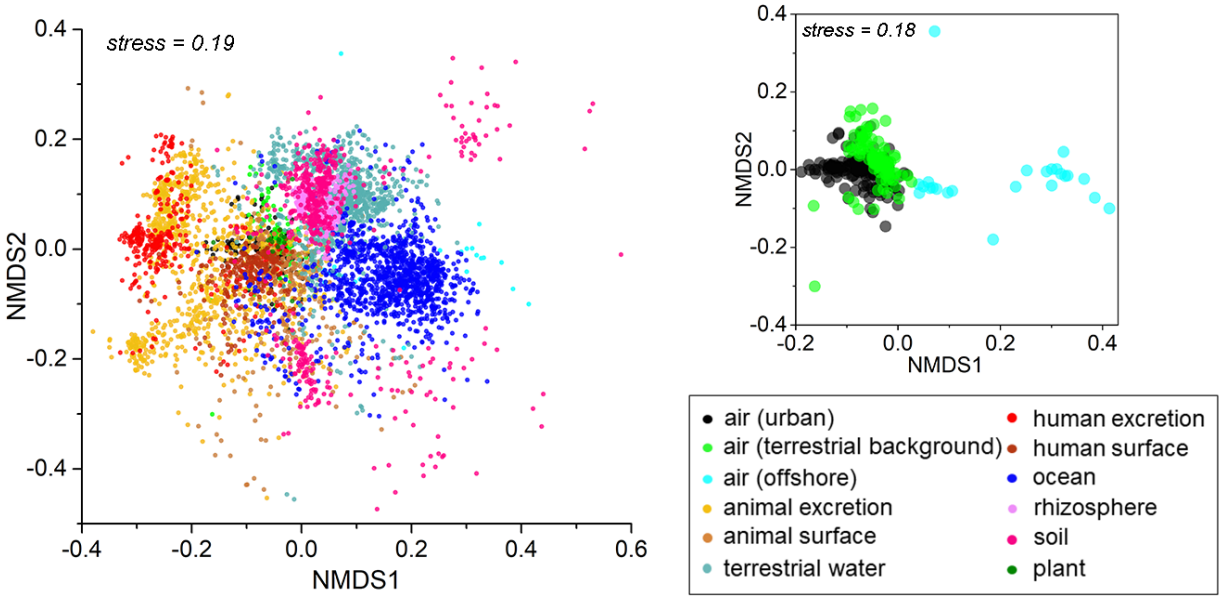


Figure **4-11** Bray-Curtis based NMDS plot showing that different microbial habitats harbor different bacterial communities on the Earth. Bray–Curtis distance was calculated to represent the dissimilarity in bacterial community compositions.

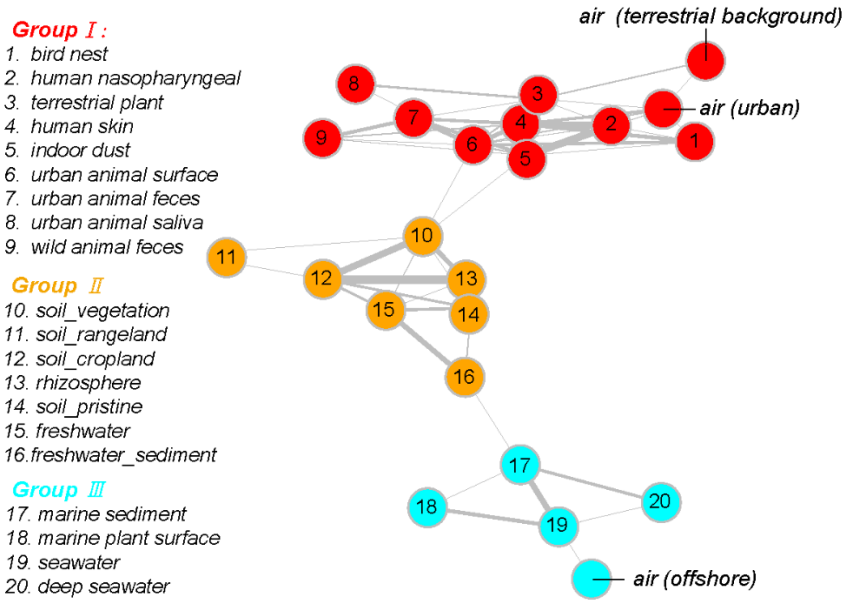


Figure **4-12** The network diagram presented illustrates the relationships and interconnections among 23 major microbial habitats on Earth. The connections, represented by edges, indicate a

strong correlation (Spearman's $\rho > 0.7$) and statistical significance ($p < 0.01$). The thickness of lines reflects the strength of the correlation. The habitats were categorized into three groups based on their similarities, represented by different colors.

4.3.3 Analysis of the Sources of Global Airborne Bacteria

In order to identify potential sources of airborne bacterial communities at the genus level in different regions, SourceTracker2 was utilized (Knights *et al.*, 2011). This software uses Bayesian methods to assess all possible sources of sink sequences (16S rRNA marker gene sequences in air samples), including an unknown source, and produces a joint distribution of these assignments. The source datasets were obtained from the EMP (ftp://ftp.microbio.me/emp/) (Thompson *et al.*, 2017). The distribution was then sampled to estimate the probability that a sequence in an air sample originated from a specific source (Glickman and Dyk, 2007). This analysis contradicted previous assumptions, which were based only on aerosol modeling in surface ecosystems, that airborne bacteria primarily came from grasslands, shrubs, and crops (Burrows *et al.*, 2009b). Instead, it was discovered that the dominant sources of airborne bacteria were determined by the characteristics of the corresponding surface environments. For offshore sites, the primary sources were oceanic (56.3 ± 36.3), while for onshore sites, human-related sources (23.2 ± 31.5) contributed significantly to airborne bacteria in urban areas, overshadowing plant-related sources (22.6 ± 25.2), which were dominant in less populated areas (Figure **4-13**). The contributions of human-related sources and terrestrial plants to airborne bacteria onshore varied considerably (Figure **4-14**a), mainly due to local population density (Figure **4-14**b) and vegetation coverage (Figure **4-14**c), respectively. Notably, despite being the most abundant ($\sim 10^{29}$) and diverse ($\sim 10^{11}$)

microbiological environment on the Earth (Locey and Lennon, 2016) (Figure **4-10**), soil was found to have a negligible contribution (<1%), possibly due to the limited contact between topsoil and air. The surface area of the global soil (1.21×10^8 km^2) (Jackson *et al.*, 1997) is smaller than that of the marine surface (3.62×10^8 km^2) (Eakins and Sharman, 2010) and leaf surfaces (5.09×10^8 km^2) (Vorholt, 2012), coupled with the crashing of waves (Wilson *et al.*, 2015) and the shaking of leaves (Burrows *et al.*, 2009b), resulted in greater exchanges between airborne bacteria and microbiomes in other bacterial habitats than with soil. Despite the lack of an advantage in surface areas with air interactions, humans and animals' frequent activities and constant respiration increase their exposure to air, resulting in human- and animal-associated habitats being the dominant source of airborne bacteria (Hospodsky *et al.*, 2012), especially in urban settings, which was overlooked in earlier emission modeling studies (Burrows *et al.*, 2009b).

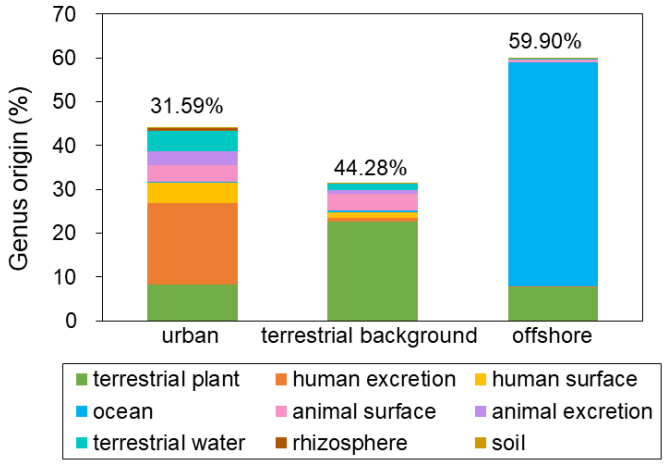


Figure **4-13** Analysis of global sources of airborne bacteria: The percentage contribution of potential bacterial genera from different environments to airborne bacterial communities in urban, terrestrial background, and offshore areas on a global scale.

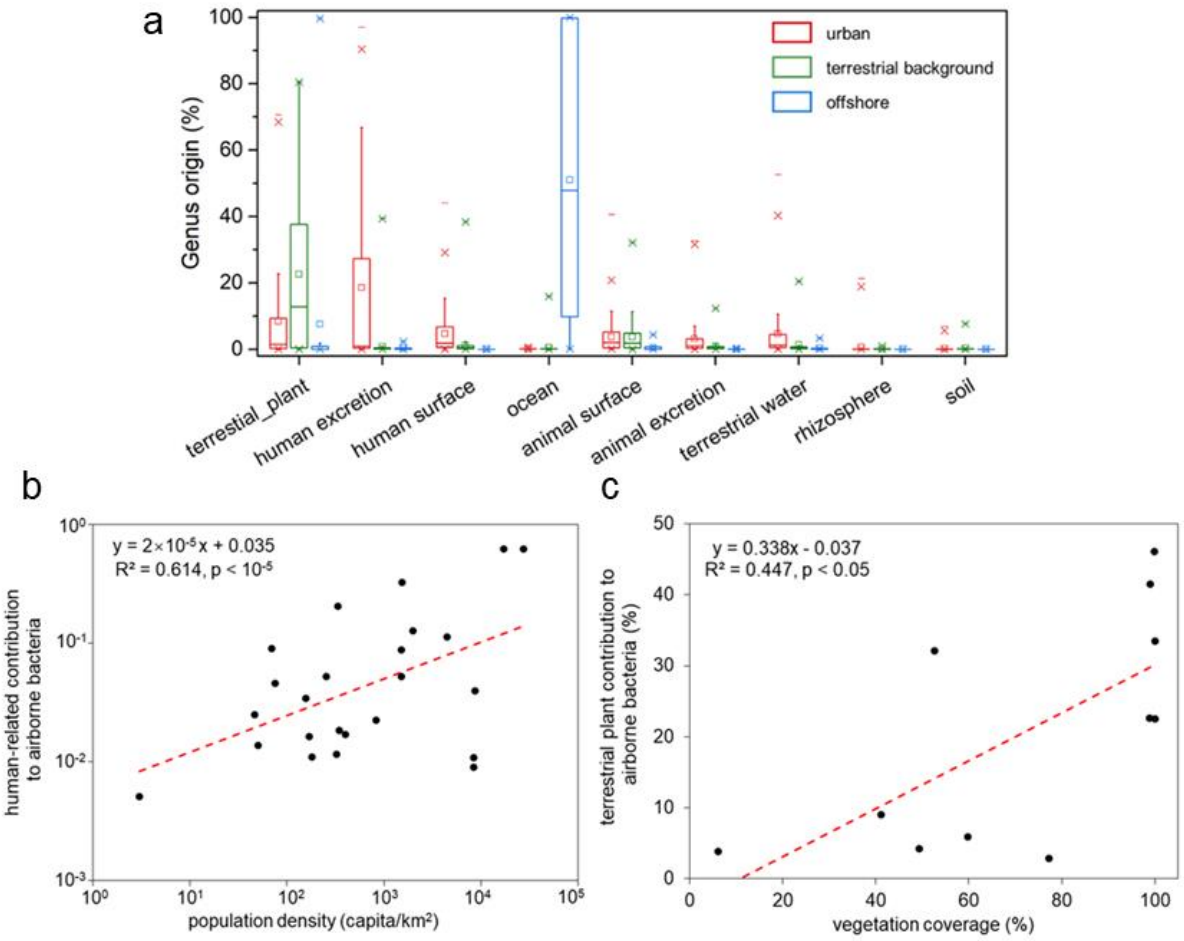


Figure **4-14** Global airborne bacterial source contribution. (a) Potential contributions of various environments to airborne bacteria (at the genus level) in urban, terrestrial background, and offshore areas respectively at the global scale. (b) Relationship of the human-related contribution to airborne bacteria and population density in urban areas. (c) Relationship of the terrestrial plant contribution to airborne bacteria and vegetation coverage in background areas.

4.4 Summary

This chapter established a comprehensive atlas of global airborne bacterial communities across 75 locations from the sub-Antarctic to the Arctic and compared it with similar atlases for the other two largest ecosystems, namely ocean and terrestrial systems. Illumina sequencing and analysis of 16S rRNA genes were conducted to draw a complete understanding of the taxonomic compositions and structures of global airborne bacterial communities, and to further identify the core bacterial sets in the three largest microbial habitats. Networked communities were used to explore the interconnections among airborne bacteria and to recognize the key taxa dominating the topological characteristics of the whole network. Besides, the bacterial biogeographic patterns, particularly diversity patterns in macro ecosystems, were presented. Moreover, the Earth bacterial co-occurrence network containing 23 various environments based on the EMP was structured to explore the interconnections of airborne bacterial communities with other microbiomes. The major findings of this chapter are listed as follows.

a) The global airborne bacterial community structure followed AOR pattern, a hyper-dominant pattern in which 24 OTUs (0.22% of the total OTU number) accounted for 18.54% of total detected sequences. These exceptionally abundant taxa could dominate and strongly influence the whole airborne bacterial communities in ecology.

b) The co-occurring bacterial communities in the global atmosphere displayed more distant relationships among OTUs, which could impede the rapid response of airborne bacteria to perturbations and result in a less stable network structure than those found in soil and marine ecosystems.

c) A group of highly connected and centrally clustering key taxa were identified, each of which displayed significant correlations with 15-18% of the OTUs in the entire network, indicating

a concentrated distribution (mean correlation coefficient = 0.903). Additionally, the similarities in composition and inferred functions among key taxa in atmospheric, aquatic, and terrestrial ecosystems suggested potential associations between airborne bacterial communities and other surface microbial habitats.

d) The uniform parabolic latitudinal diversity patterns of microbial worlds were observed in the three largest ecosystems on Earth (*i.e.*, atmosphere, ocean, and terrestrial systems), which differed from the typical latitudinal gradient of diversity found in macroscopic organisms. Temperature was identified as the crucial factor driving the distribution of diversity.

e) The complexity and dynamics of global airborne bacterial communities were found to be comparable to those of bacterial assemblages in soil and ocean environments, owing to the similar estimated bacterial richness in the global atmosphere to that of the hydrosphere. Furthermore, the essential role of airborne bacteria in the Earth's microbial world was broadly established based on the close interrelationships with bacteria in 23 major habitats and the finding that almost half of the contributions of airborne bacteria came from other ecosystems.

In summary, this chapter presented the first systematic study of global airborne bacterial communities, constructing a reference catalog of over 27 million nonredundant 16S rRNA gene sequences. The biomass, diversity, biogeographic patterns, compositions and functions of core and key taxa, and networked community structure of airborne bacterial communities were explored and compared with counterparts in the other two largest ecosystems on Earth: marine and soil. This chapter revealed that airborne bacterial communities were as complex and dynamic as bacterial assemblages in other environments, with a unique set of dominant taxa (*n=24*) but a structure that appeared to be more easily perturbed. Despite having lower total abundance, the

airborne bacterial richness (4.71×10^8 ~ 3.08×10^9) is comparable to that in the hydrosphere, and its distribution followed a similar latitudinal pattern with other ecosystems. Additionally, this chapter highlighted the great contribution of surface microbiomes to airborne bacteria (averaging 46.3%). Nevertheless, only the structure and distribution of global airborne bacterial communities were discussed in this chapter. To achieve a better comprehension of the underlying mechanisms, it is imperative to explore the impacts of environmental variables on airborne bacterial communities at a global scale. This will facilitate the forecasting of ecosystem responses to potential environmental changes in the future but has yet to be profiled. Consequently, in the upcoming chapter (Chapter 5), the direct and indirect impacts of biotic factors and abiotic factors on the composition, diversity, core bacteria, key taxa, and abundance of each OUT in the airborne bacterial communities worldwide will be analyzed and quantified to generate a comprehensive unscrambling of the driving mechanisms responsible for the biogeographic patterns observed in global airborne bacterial communities. The finding of enhancing human-related source contributions to airborne bacteria in urban areas indicated the anthropogenic impacts on airborne bacterial communities. However, the specific alterations due to urbanization from a global view and the related mechanisms remain elusive, and it is essential to pinpoint the interplay between human activities and the existing airborne microbiomes to comprehend the harmony between humans and nature. Finally, in Chapter 6, there will be a further focus on risks of airborne AMR to human health on a global scale, given the human imprints on airborne bacteria, particularly pathogens.

Chapter 5 Mechanisms Driving Global Patterns of Airborne Bacterial Communities via Anthropogenic and Environmental Processes

Following up on the higher sources of airborne bacteria from human-related microbial habitats in urban areas, the anthropogenic impacts on global airborne bacterial communities, with a particular focus on structure, genotypes, and pathogenicity, were detailed discussed in this chapter. The biogeographic patterns indicated that environmental filtering played a significant role in global airborne bacterial community, and this will be further investigated by considering both direct and indirect factors from three perspectives: diversity and total biomass, key taxa and core bacterial sets, and each OTU in the whole bacterial community. Moreover, community assembly processes were also investigated to uncover the specific mechanisms driving the global airborne bacterial community structure. To better understand the underlying mechanisms shaping structure and biogeographic patterns of global airborne bacterial community, a comprehensive analysis was conducted based on a range of meteorological conditions, air qualities, and source contributions.

5.1 Human Imprints on Airborne Bacterial Communities

5.1.1 Anthropogenic impacts on global airborne bacterial community structure

The dissimilar structures of airborne bacterial communities between urbanized and less human-impacted sites emphasize the importance of human influence on airborne bacterial communities (Figure **4-11**). However, within the same latitude range, there were no significant differences in

the bacterial community richness between urban and background areas (*i.e.*, areas that are much less affected by humans, such as remote mountain sites, offshore environments, and the Arctic region) (Figure **5-1**). This finding is consistent with the discovery in section 4.2 that human distribution and activities minimally affect airborne bacterial diversity. These results suggest that geographic location is the primary determinant of airborne bacterial richness rather than anthropogenic influences. Notably, the OTU richness of airborne bacterial communities did not significantly differ between urban and background areas, indicating that humans inhale a comparable number of bacterial species (Figure **5-2**a). However, the evenness of bacterial communities was substantially lower in urban areas (Figure **5-2**b), indicating a significant increase in the abundance of some bacterial types. For example, the relative abundance of two typical commensal bacteria, *Burkholderia* and *Pseudomonas* (including some pathogenic species), was significantly higher in urban areas (5.56% and 2.50% respectively) than in background areas (1.44% and 1.11%).

In terms of community composition, both urban and background areas harbored unique bacterial types exclusive to their respective areas (713 and 2,835), while more than half of total bacterial types (4,352) were found in both areas (Figure **5-1**b). Additionally, the contribution of bacterial mass to PM mass was much lower in urban areas than in natural areas (Figure **5-2**c), indicating that urbanization increased the proportion of non-biological particulates, *e.g.*, dust and soot in air PM.

To discern the differences in the microbial co-occurrence patterns of the airborne bacterial communities in urban and background areas, two networks were generated using $|\rho| > 0.6$ and $p < 0.05$ (Figure **5-3**a and Figure **5-3**b). The biotic interconnection network of urban bacterial communities ($n=62$) revealed 134 nodes with 383 connections (edges), and the average degree

was 2.9. In contrast, the background bacterial communities, consisting of 178 nodes, namely OTUs identified from 88 individual background air samples, demonstrated 1,372 strong and significant edges. The higher average degree (7.7) indicated that the biotic interconnections within background bacterial communities were more complicated. In addition, the general topological characteristics, such as average shortest path length and transitivity, illuminated the impact of human activities on the structure and stability of microbial networks. In particular, the decreased transitivity and heightened average shortest path length in the co-occurrence network of airborne bacterial communities in urban areas indicated that human influences disrupted the network structure (Figure **5-3**c).

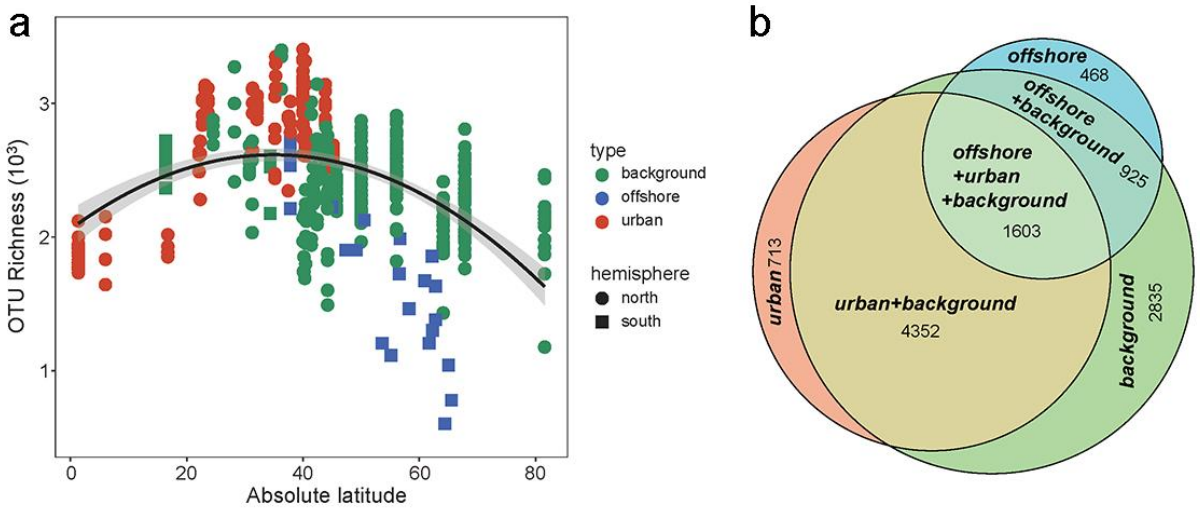


Figure 5-1 Anthropogenic impacts on the composition of airborne bacterial communities. (a) Latitudinal distribution of airborne bacterial α-diversity (richness). Plotting OTU richness against the absolute latitude of sampling locations shows that there is no significant disparity in richness between urban and background areas. (b) Number of exclusive and shared airborne bacteria in urban, terrestrial background, and offshore areas.

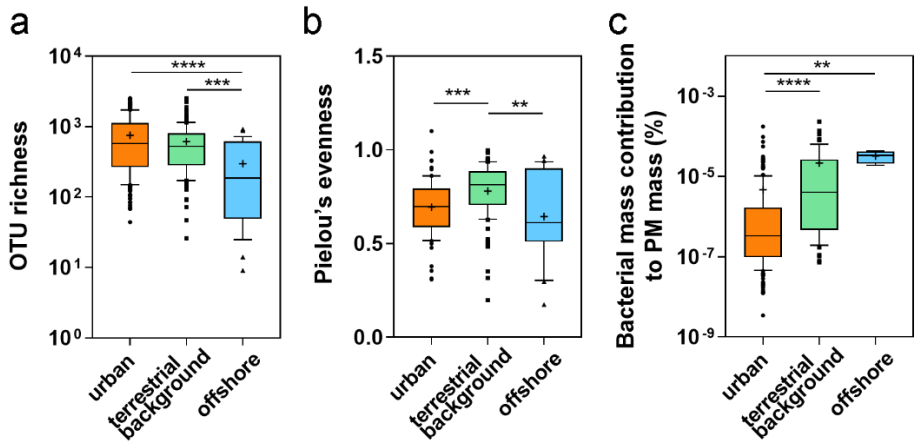


Figure 5-2 A comparison of diversity indexes (*i.e.*, richness (a) and evenness (b)) and bacterial mass contributions to PM mass (c) in urban, terrestrial background, and offshore areas.

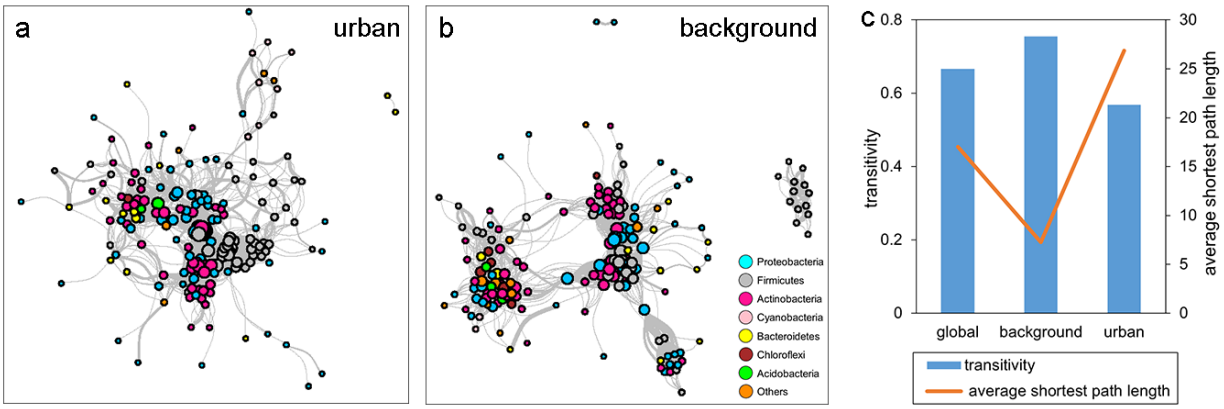


Figure 5-3 Anthropogenic impacts on the network structure of airborne bacterial communities in urban and background areas. Co-occurrence networks of airborne bacterial communities in urban (a) and background (b) areas (including terrestrial background and offshore areas), respectively. Different node colors indicate different phyla. (c) Comparison of network topological characteristics in urban areas, background areas, and the whole global dataset.

5.1.2 Impacts of urbanization on airborne bacterial genotypes

The alteration of the airborne bacterial taxonomic composition due to urbanization also brought a corresponding change to some phenotypic characters. For instance, the structure (*i.e.*, cell shape and cell management) of airborne bacterial cells changes a great deal due to urbanization as can be seen from the finding that the percentage of Bacilli and the bacteria existing in clusters was higher in urban areas than in terrestrial background and offshore areas (Figure **5-4**). Bacilli are known for their resistance to various environmental stresses, including heat, radiation, and chemicals such as antibiotics, which makes them capable of surviving for long periods in controlled environments (Christie and Setlow, 2020). The higher proportion of Bacilli in urban air might increase AMR risks. Additionally, it is well known that AT is mainly driven by latitude, yet the relatively higher AT in urban areas than in terrestrial background and offshore areas in general caused there to be a correspondingly higher optimal temperature range for airborne bacteria in urban areas (Figure **5-5**a). For instance, the relative abundance of thermophilic bacteria was higher in the urban areas of Guangzhou (5.25%) than in a background area with a similar latitude, Mt. Ailao (2.39%). Nevertheless, urbanization did not affect the subsistence state of airborne bacteria; there were no significant differences in the environmental tolerance caused by sporulation and the surviving modes (symbiotic or free-living) of bacteria between urban and background air (Figure **5-5**b and Figure **5-5**c). The motility of airborne bacteria was higher in urban areas than in natural areas (terrestrial background and offshore areas), as reflected in the higher ratio of bacteria carrying flagella, a motility organelle enabling movement and chemotaxis, in urban air (Figure **5-6**). Although the flagellum has traditionally been recognized only as a motility organelle, recent evidence suggests that it could be a potential virulence factor due to its adhesive and invasive

properties (Haiko and Westerlund-Wikström, 2013), which may increase the risk of pathogens invasion in urban air.

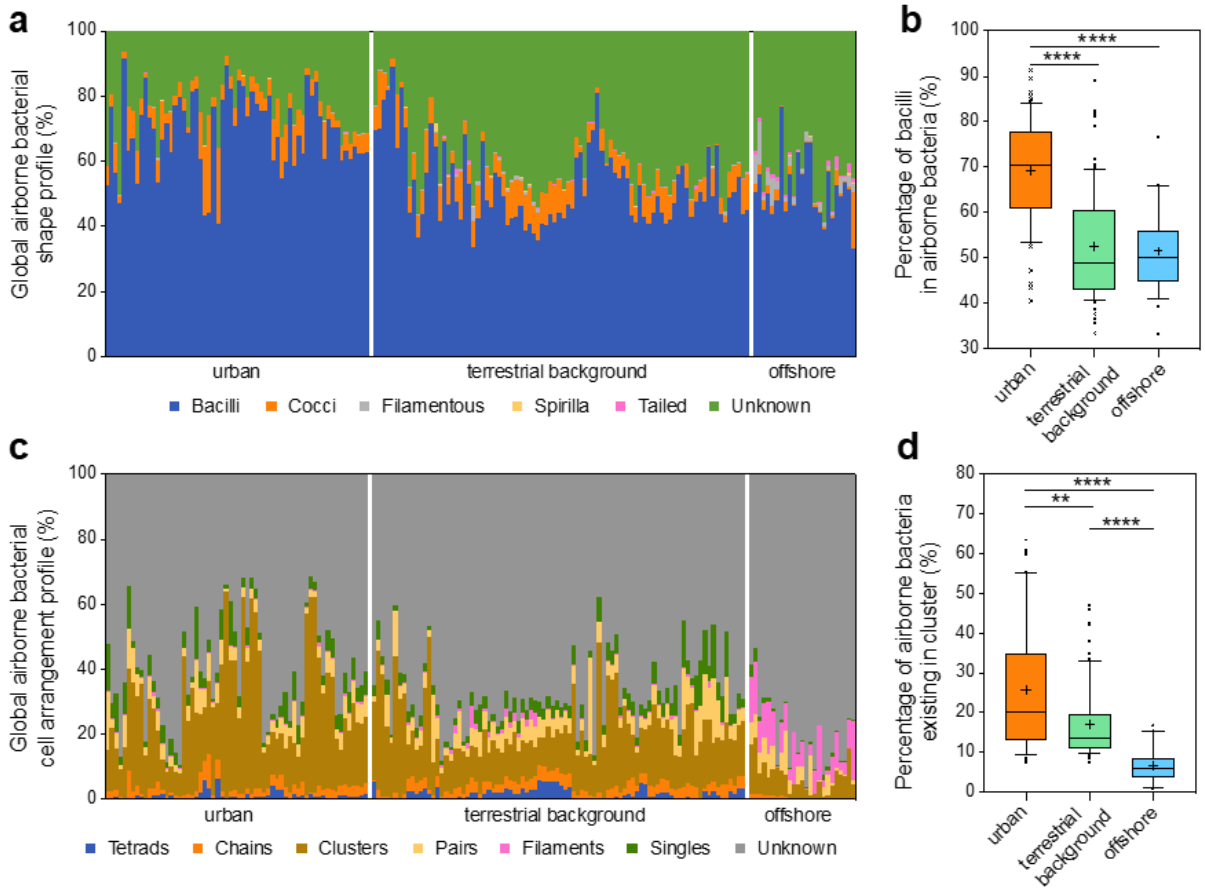


Figure 5-4 Structure of global airborne bacterial cells in urban, terrestrial background, and offshore areas. (a) Global airborne bacterial shape profile. (b) Percentage of Bacilli in airborne bacteria. (c) Global airborne bacterial cell arrangement profile. (d) Percentage of airborne bacteria existing in clusters.

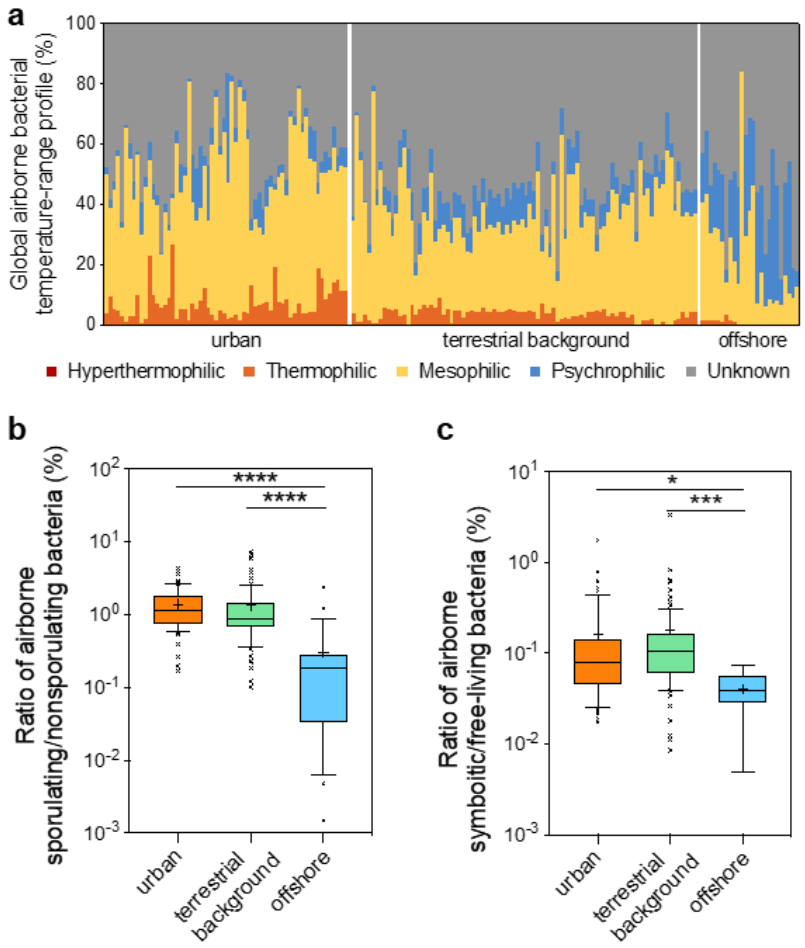


Figure 5-5 Adaptation of global airborne bacteria to local environments in urban, terrestrial background, and offshore areas. (a) Optimal temperature range of airborne bacteria. (b) Sporulation of airborne bacteria. (c) Ratio of airborne symbiotic/free-living bacteria.

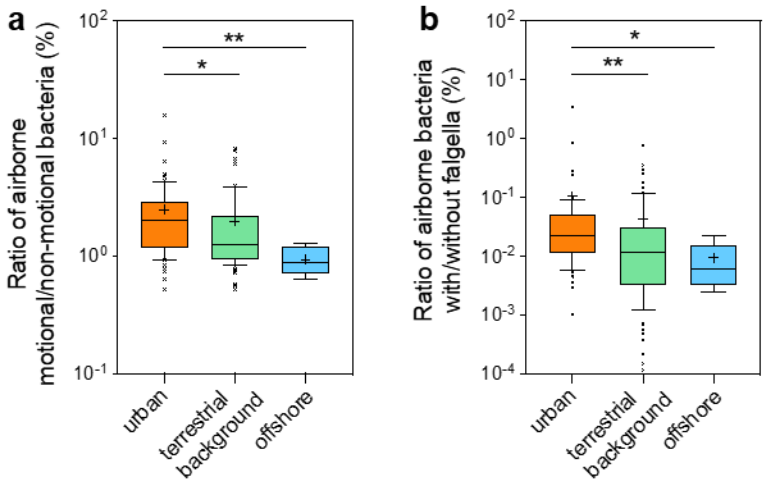


Figure 5-6 Motility of global airborne bacteria in urban, terrestrial background, and offshore areas. (a) Ratio of airborne motional bacteria to non-motional bacteria. (b) Ratio of airborne bacteria with flagella to bacteria without flagella.

5.1.3 Anthropogenic forcing on global airborne bacterial pathogenicity

The conventional culture methods based on phenotypes have lower sensitivity and accuracy making them unsuitable for comparing potential airborne pathogens in urban and background air (Manaka *et al.*, 2017) Therefore, the 16SPIP (Miao *et al.*, 2017), a comprehensive pipeline designed for clinical samples but also applicable to diverse environmental samples (Li *et al.*, 2020; Li *et al.*, 2021a), was used to overcome these limitations. The results showed that despite lower total bacterial loading (Figure **5-7**a), urban air has a significantly higher relative abundance of potential pathogens (Figure **5-7**b), particularly the ESKAPE pathogens with the highest risk of mortality (Mulani *et al.*, 2019) (Figure **5-8**a). Humans inhaled less abundant airborne bacteria; nevertheless, there is a risk that various pathogenic infections might increase in cities, with 22.4% of identified airborne pathogens (*n = 37*) occurring exclusively in urban areas (Figure **5-8**b). An

additional metagenomic analysis confirmed the accuracy of composition and abundance of potential pathogens identified from 16S rRNA gene sequencing data (Figure **5-8**b), although more accurate quantitative diagnostic methods are still required in future studies, *e.g.*, pathogen-specific real-time qPCR analysis (Rinttilä *et al.*, 2004). We hypothesized that the elevated abundance and diversity of airborne pathogens in urban areas might have originated from human-related sources, as previously suggested in section 4.3.3. In addition, no great difference was seen in the ratio of gram-positive to gram-negative airborne bacteria (and pathogens) in urban, terrestrial background, and offshore areas, further illustrating that urbanization did not alter the likelihood of human illness related to gram-negative and/or gram-positive pathogens (Figure **5-9**). However, the higher mobility of airborne bacteria and higher ratio of bacteria carrying flagella, which was verified in section 5.1.2, could increase the spread of airborne pathogens and enhance the risk of human infections in urban regions (Figure **5-6**).

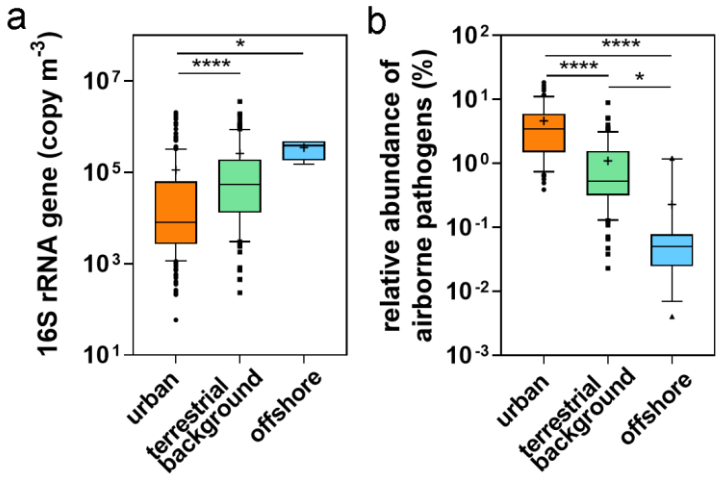


Figure 5-7 A comparison of total airborne bacterial loadings (a) and the relative abundance of pathogens (b) in urban, terrestrial background, and offshore areas.

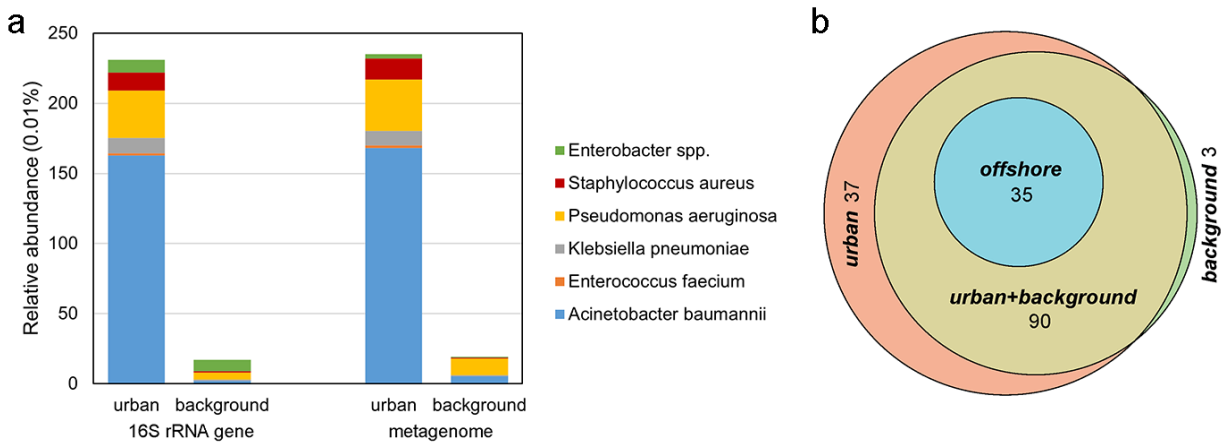


Figure 5-8 Anthropogenic impacts on the composition of airborne pathogenic communities. (a) Ratio of the abundance of airborne ESKAPE pathogens in urban areas to background areas based on two datasets, 16S rRNA gene and metagenomic sequencing data, respectively. (b) Number of exclusive and shared airborne pathogens in urban, terrestrial background, and offshore areas.

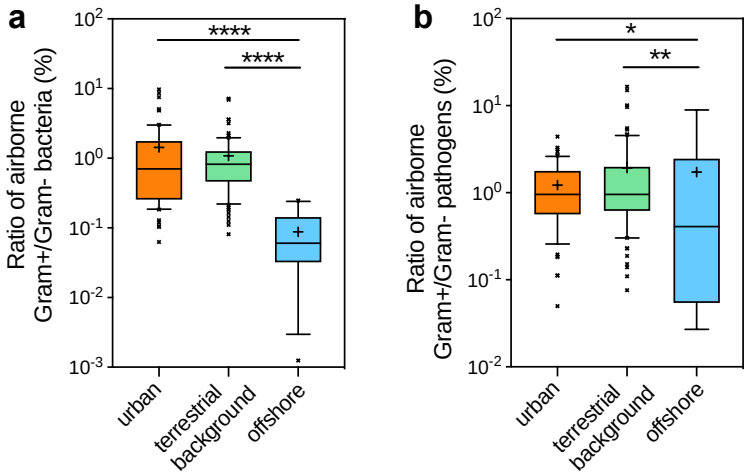


Figure 5-9 Gram types of global airborne bacteria in urban, terrestrial background, and offshore areas. (a) Ratio of airborne gram-positive to gram-negative bacteria. (b) Ratio of airborne gram-positive to gram-negative pathogens.

5.2 Environmental impacts on global airborne bacterial communities

Although many regional correlations between total airborne bacterial concentrations and environmental variables have been examined in a previous review (Burrows et al., 2009), few correlations were found at the global scale (Figure **5-10**). This section addressed the effects of environmental filtering on global airborne bacterial communities, analyzing the direct and indirect impacts from three perspectives: diversity and total biomass, key taxa and core bacterial sets, and each OTU in the whole bacterial community.

5.2.1 Relationships between environmental factors and airborne bacterial diversity and biomass

To explore the relationships between environmental variables and airborne bacterial diversity, the data was displayed as bivariate plots (Figure **5-10**), which exhibited several patterns. The bacterial richness as a function of variable environmental factors revealed considerable variation. In these patterns, most coefficients of correlation (*i.e.*, R) were modest, despite the high significant significance due to the large sample size. Thus, there was still a large unexplained variation between the environmental variables and bacterial richness.

Despite the variation, significant patterns were still evident. For example, the variations in the diversity of global airborne bacteria with latitude closely matched the variations in the AQI scores, as well as concentrations of PM_{10}, $PM_{2.5}$, SO_2, and CO, and other meteorological parameters like AT, AP, WD, WS, and RH (Figure **5-10**). In contrast to its strong correlations with latitude,

richness was not as well correlated with other geographic locations. A pairwise correlation analysis suggested that the relationships between richness and altitude (R^2 = 0.019, p = 0.209) or distance to coast (R^2 = 0.019, p = 0.209) were either weak or nonexistent. Besides, different pollutants showed different relationships with bacterial diversity, including parabolic fitting curve associations (AQI, $PM_{2.5}$, PM_{10}), positive correlations (SO_2, CO), and independent relationships (NO_2, O_3). Among the various meteorological conditions, RH was the best-predicted factor of bacterial richness (R^2 = 0.190, $p < 10^{-7}$).

Similar to bacterial diversity, the correlations between the total airborne bacteria biomass (*i.e.*, the number of 16S rRNA gene copies) and environmental factors were also analyzed (Figure **5-11**). Three factors were significantly related to bacterial biomass: NO_2 concentration (R^2 = 0.109, $p < 10^{-5}$), WS (R^2 = 0.173, $p < 10^{-8}$), and WD (R^2 = 0.189, $p < 10^{-6}$). The parabolic fitting curve relationship showed that the biomass was highest with southerly winds compared to other directions. At this point, it was hypothesized that warm temperatures and low latitudes might be hospitable to bacterial survival and diversity, because most samples were collected in the northern hemisphere, and southerly winds could increase the local AT to a certain extent. However, this hypothesis conflicted with another discovery that there were no relationships between biomass and either latitude or AT and was soon disavowed. As a result, there must be unknown mechanisms linking WD to airborne bacterial biomass, which might be related to atmospheric circulation or other geographic factors affecting microbial communities.

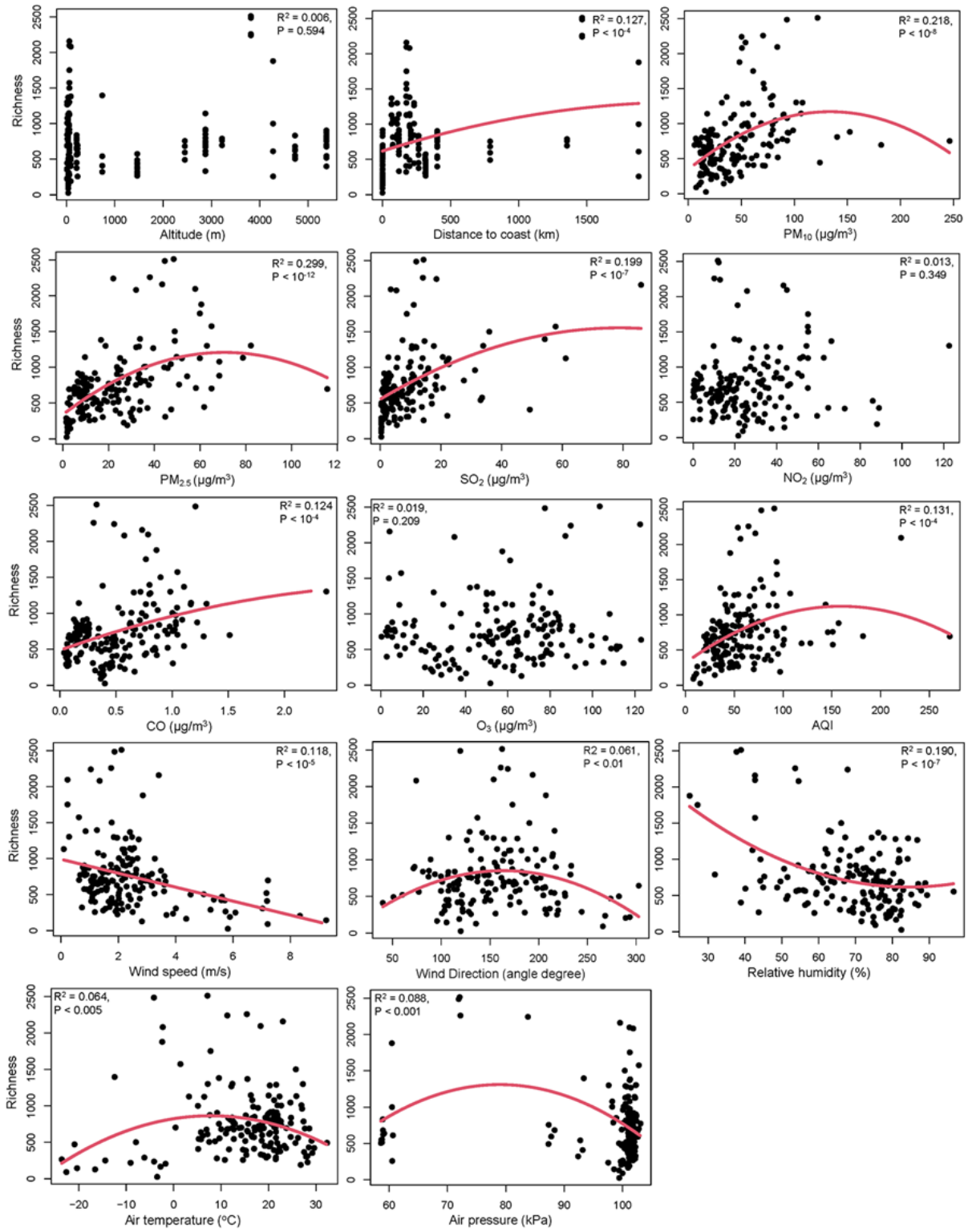


Figure 5-10 Relationships of the global airborne bacterial diversity index (OTU richness) with air quality and meteorological conditions.

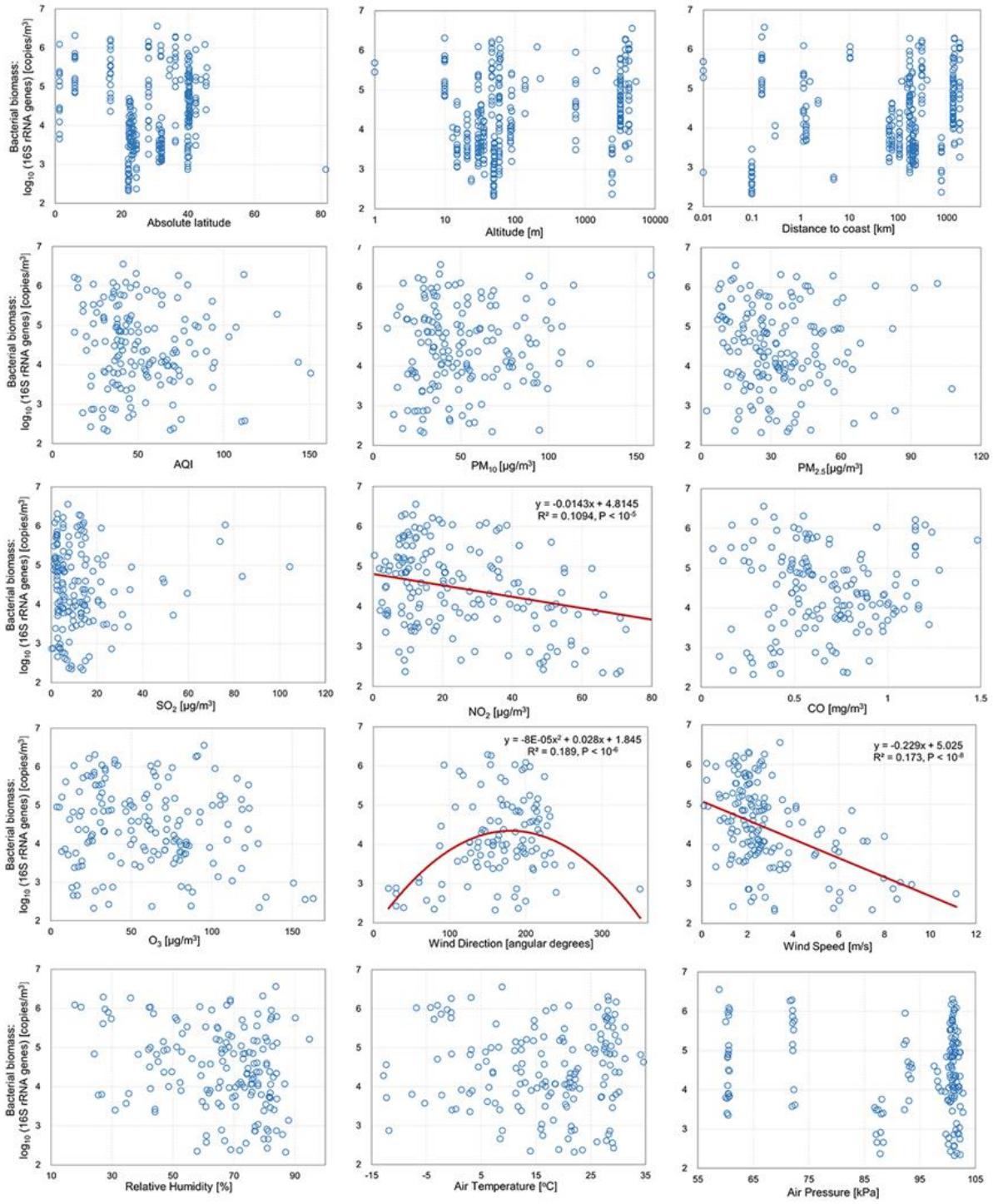


Figure 5-11 Relationships of global airborne bacterial biomass (16S rRNA gene copy number) with geographic locations, air quality, and meteorological conditions.

5.2.2 Impacts of environmental variables on key taxa and core bacterial communities

To investigate the variability of two crucial airborne bacterial communities (24 core OTUs and 19 key taxa), PCoA analysis was performed, which revealed that the samples were significantly differentiated by their geographic locations (Figure **5-12**a and Figure **5-12**b). Moreover, PCoA visualization showed that variations in core bacterial community composition were strongly associated with differences in latitudinal regions, similarly, the significantly linear relationship between PC1 of PCoA scores and absolute latitude (R^2 = 0.316, $p < 10^{-14}$) confirmed the determinant effect of latitude on the spatial and temporal distribution of core bacteria (Figure **5-12**c). In the PCoA plot, the 370 global air samples were clearly clustered into three groups: low latitudinal samples were located in the first quadrant, intermediate latitudinal samples were mainly located in the third quadrant, and samples collected in high latitudinal regions were distributed in the second quadrant (Figure **5-12**a). Similarly, the importance of geographic effect was also verified in key taxa, demonstrated by the significantly negative correlation between PC1 score in key taxa and altitude (R^2 = 0.481, $p < 10^{-24}$, Figure **5-12**d). Additionally, AT and AP were directly driven by latitude and altitude respectively, thus, the two typical environmental factors could contribute to explaining these two biogeographic patterns occurring in core and key taxa. There were indeed linear relationships between the core and key taxa with AT and AP respectively (Figure **5-12**e and Figure **5-12**f), however, these correlations were weaker than those observed with geographic locations (AT with PC1_core bacterial communities: R^2 = 0.085 < 0.316; AP with PC1_key taxa: R^2 = 0.290 < 0.481). As a result, while AT and AP could partially explain the biogeographic pattern, there were still great unknown driving mechanisms that require further exploration.

To identify the determinant factors respectively driving biogeographic distribution of global core and key taxa in the atmospheric environment, multiple regression analyses were performed to assess their correlations with other potential environmental factors. The results showed that all the key taxa and most of the core bacteria were significantly ($p < 0.05$) correlated with at least one of the 11 environmental factors tested (Figure **5-13**a and Figure **5-13**b). Further analysis using VPA showed that air quality affected the key taxa most with a subset of air pollutants, including SO_2, NO_2, CO, O_3, PM_{10}, and $PM_{2.5}$, together explaining a substantial amount (53.3%) of the structural variations, which was substantially higher than the figures for meteorological condition (26.54%) and landscape coverage type (18.38%) (Figure **5-14**b). Land coverage type contributed far less to the variations in core bacteria (15.93%), compared to the other two groups (36.53% and 31.99%) (Figure **5-14**a). Notably, all three groups had a significant impact on the whole communities, explaining over 70% of the variations in both the core (70.63%) and key (80.11%) taxa.

Various forms and species of heavy metals have been well known and widely used as antimicrobial agents (Mittapally *et al.*, 2018), thus, exploring the effects of heavy metals in PM on airborne bacteria could provide insight into mechanisms of bacterial survival and community organization. Despite the divergent genetic construction and response to environmental factors of key taxa, all key taxa were negatively affected by heavy metals, as suggested by the associations between heavy metals and key taxa in RDA (Figure **5-15**a). The negative impacts of metals on bacteria could be attributed to the ability of high heavy metal concentrations to destroy cell membrane structure, change enzyme specificity, impair cell function, change protein structure, damage DNA stands, and even cause cell death (Silver and Phung le, 2005). On the other hand, heavy metals could also have positive impacts on bacterial growth by providing essential nutrients (such as K, Na, Ca, Mg, Fe, Cu, and Zn) (Gadd, 2010), and required energy for microbial metabolism in oxidation-

reduction reactions (such as As^{3+}/As^{5+}, Fe^{2+}/Fe^{3+}, Mn^{2+}/Mn^{4+}, V^{4+}/V^{5+}, Se^{4+}/Se^{6+} and U^{4+}/U^{6+}) (Lengke and Southam, 2006). For example, around half of the core OTUs were positively correlated with heavy metals (Figure **5-15**b), perhaps because these bacteria were too abundant and adaptable in atmospheric environments to be damaged by metal toxicity. However, the remaining core bacteria showed weak or negative correlations with heavy metals, indicating the complex effects of metals on bacteria that depend on specific conditions. To further explore the link between key taxa and nutrition cycling, RDA was conducted. Water soluble ions (NH^{4+}, NO^{3-}) in PM were also a kind of nitrogen source for bacteria, and the correlations of the two ions with key taxa, particularly the nitrogen-fixing bacteria *(Rhizobiales:* OTU63 and OTU55), were stronger than with other OTUs (Figure **5-15**c). Another key taxon (Frankiales: OTU19) with a close relationship with soluble ions was commonly adaptable and occupied different ecological niches, consistent with the ecological functions of key taxa discussed in section 4.1.3. In contrast, only OTU6 (*Acinetobacter*) and OTU25 (*Bacillus*) were correlated with soluble ions, possibly due to their ability to survive in a broad array of environments and even extreme conditions (Figure **5-15**d). Moreover, these soluble ions (NH^{4+}, NO^{3-}, SO^{2-}) could be connected with gaseous pollutants (NO_2, SO_2) due to their shared elements. Additionally, the metabolic activity of airborne bacteria with various environmental conditions has been documented in previous studies (Sattler *et al.*, 2001; Meola *et al.*, 2015; Klein *et al.*, 2016; Stone *et al.*, 2016).

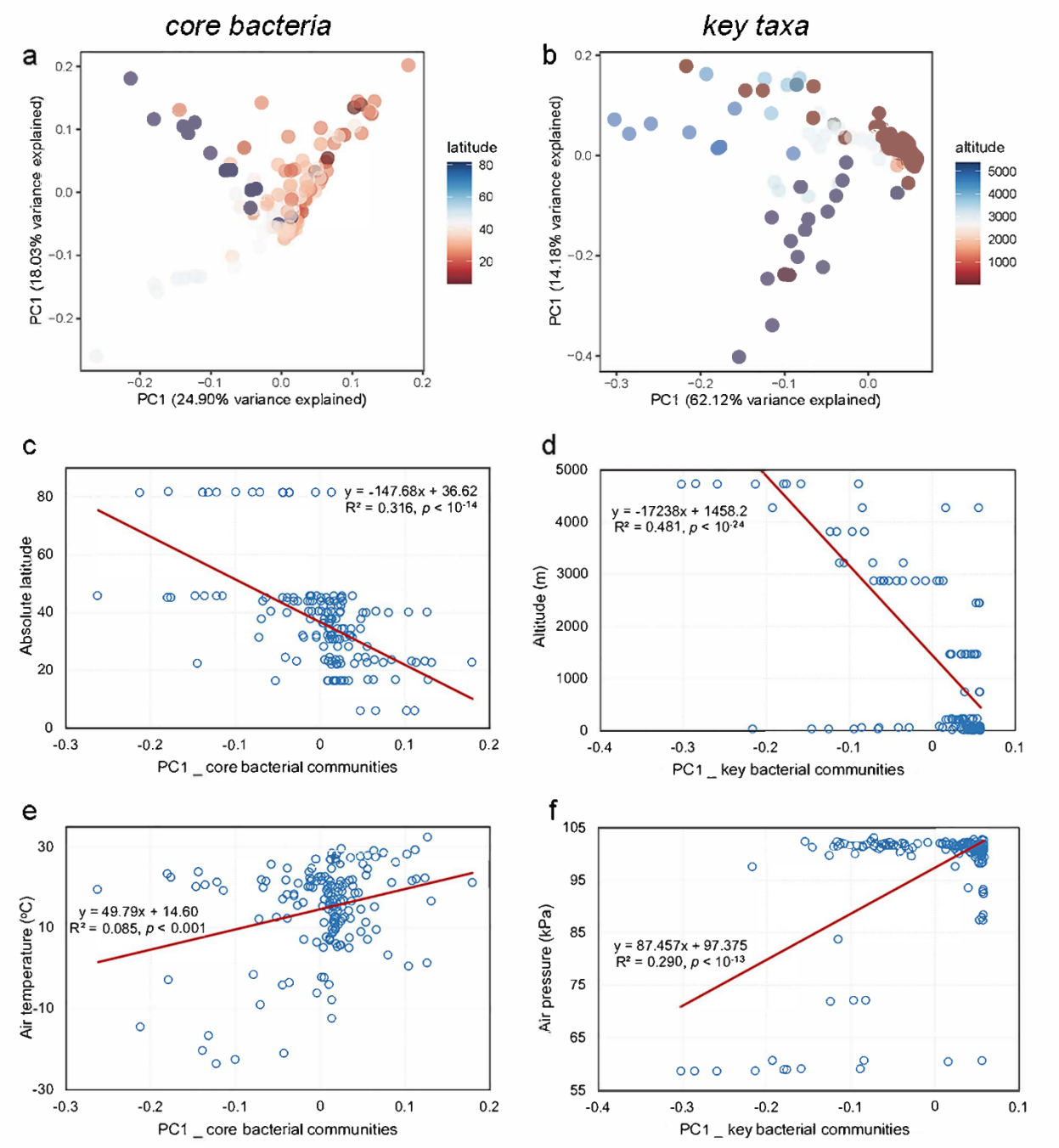


Figure 5-12 Environmental factors driving global airborne core and key taxa. PCoA performed on global airborne core (a) and key (b) bacterial community composition dissimilarities on the basis of "Euclidean" index of 370 air samples. Samples are color-coded according to the corresponding geographic locations. Impacts of geographic locations on global airborne core (c) and key (d)

bacterial community structure. Environmental factors driving global airborne core (e) and key (f) bacterial community structure.

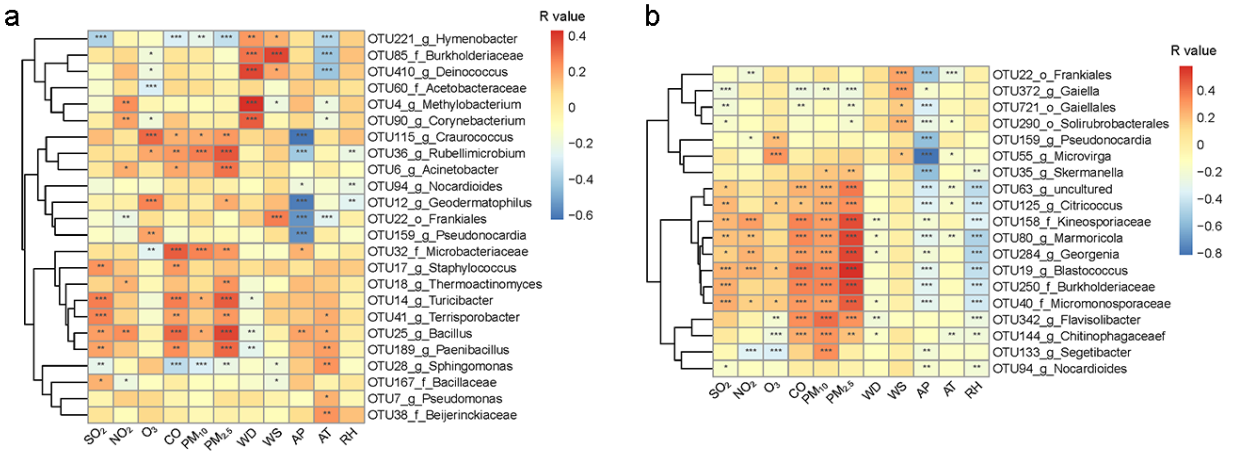


Figure 5-13 The correlations of 24 core bacteria (a) and 19 key taxa (b) and environmental variables. The right color gradient indicates Spearman's rank correlation coefficients (*i.e.*, R), with dark blue indicating stronger positive correlations and dark red indicating stronger negative correlations. The number of asterisks indicates the significance levels (two-sided) of Spearman's rank correlation coefficients (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). Each row represents the correlation of a specific OTU and environmental variables in 370 biologically independent air samples, which are clustered based on Spearman's rank correlation coefficients.

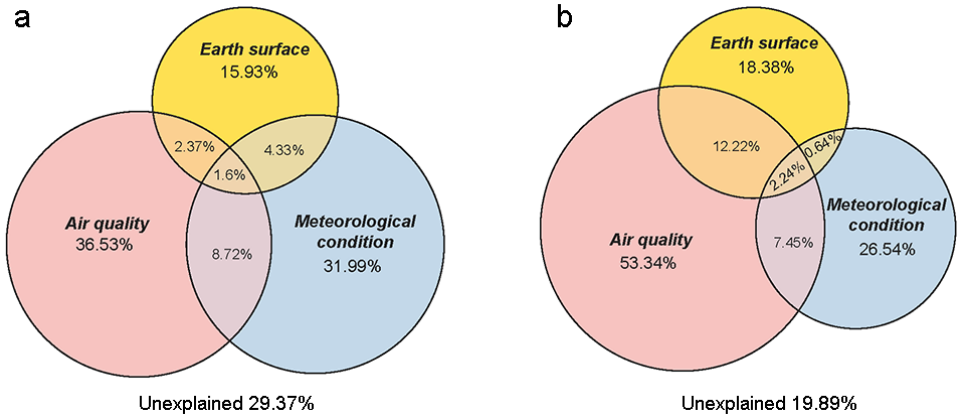


Figure 5-14 VPA showing relative contributions of air quality (PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3, and CO), meteorological condition (AT, AP, RH, WS, and WD), and landscape coverage type (water/sea, urban, grassland, cropland, forest, and shrubs) to the community variations of core bacteria (a) and key taxa (b). The overlap represents the joint effect explained by two or three factor groups together, while the percentage number below each group name represents the variance explained by one group alone. "Unexplained" denotes a variance that could not be explained by any one of these three groups.

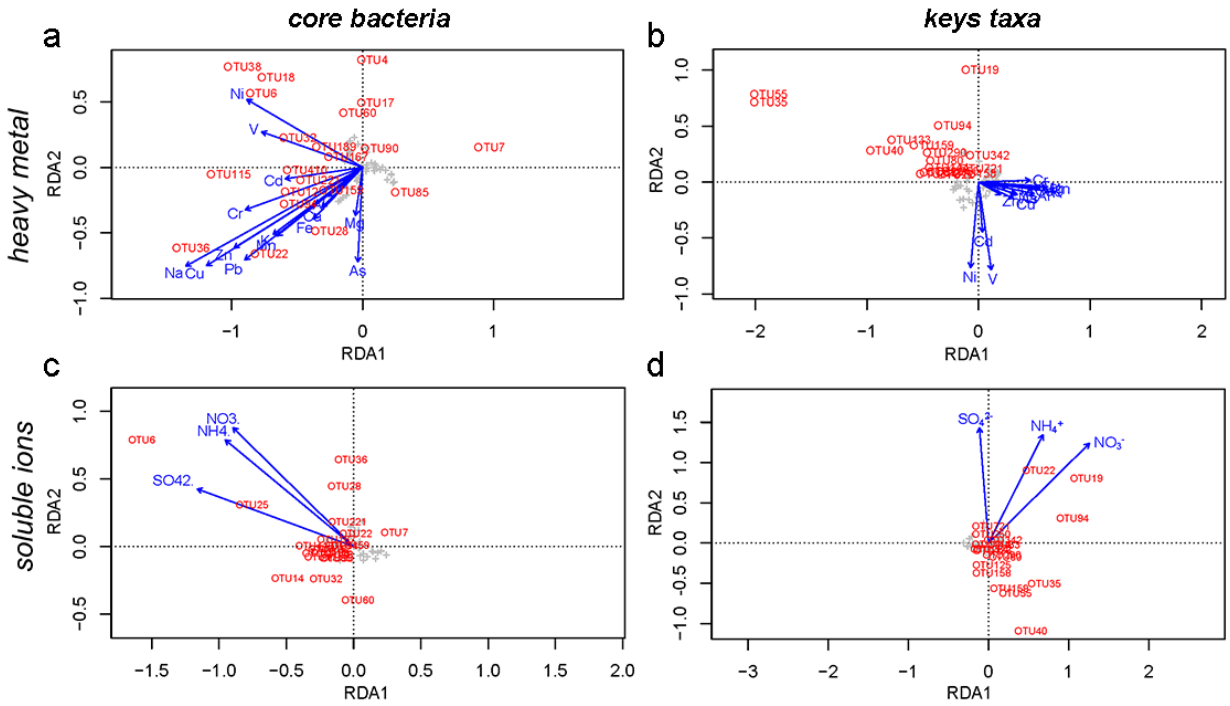


Figure 5-15 RDA identifying the relationships between core and key taxa with soluble ions and heavy metals in air samples. Blue arrows represent the environmental factors, with their length indicating the strength of the correlations between the environmental factors and the sample distributions. The longer the line is, the stronger the corresponding correlation will be. The included angle between the arrows denotes the correlation between environmental factors. An

acute angle means a positive correlation and an obtuse angle a negative correlation, while a smaller included angle corresponds to a stronger correlation. Red labels indicate the OTU number and gray markers indicate the distribution of the samples.

5.2.3 Environmental filtering impacting each OTU in whole global airborne bacterial communities

To further understand the impacts of environmental factors on whole airborne bacterial communities, a thorough investigation into the correlations between11 typical environmental variables (*i.e.*, concentration of PM_{10}, $PM_{2.5}$, SO_2, CO, NO_2, and O_3, AT, AP, WD, WS, and RH) and a total of 10,897 OTUs was conducted. The findings exhibited that more than half of the OTUs (57.7%) showed no discernible connections with any of the environmental variables. For the few OTUs displaying relationships, t the vast majority (97.4%) exhibited weak correlations (absolute Spearman's $\rho < 0.5$), indicating that the impacts of environmental filtering on the abundance of each specific OTU were negligible (Figure **5-16**a). In addition, the total number of significant correlations of each OTU with air pollutants (*n = 6,440*) was much larger than with meteorological conditions (*n = 2,284*), particularly for CO, PM_{10}, and $PM_{2.5}$, which were respectively correlated with 14.4%, 12.9%, 12.5% members of whole airborne bacterial communities (Figure **5-16**b). Moreover, most of the impacts of air pollutants on airborne bacteria (78.0% in key taxa, 72.5% in core bacteria, and 69.0% in all OTUs) were positive (Figure **5-16**c), which indicated that it was easier for the structure of airborne bacterial communities to be positively influenced by moderate air pollution. AP and RH also contributed to shaping community structure and could directly affect

9.8% and 8.2% of bacteria, respectively. However, other meteorological conditions had almost no impacts on bacterial abundance.

Abiotic factors, namely environmental variables, affected airborne bacterial richness in various regions but had less influence on the abundance of each specific OTU, with a low ratio (7.28%) of the number of existing significant relationships ($p < 0.05$) to the number of all possible connections between OTUs and environmental variables. Thus, there must be some unknown connections between the environmental variables and the overall community structure. Here, the significant effect frequencies of environmental factors on each core and key OTU were quantified using multiple regression analysis, showing the ratios of 24.62% and 44.02%, respectively (Figure **5-16**c). This revealed that the strength of the impact of environmental filtering on these two crucial bacterial communities largely outweighed those of other normal OTUs. Key taxa and core bacteria played a key role in shaping the composition of the whole community and were much more affected by the environmental filtering process. In addition, the abundance of each key OTU was more frequently related to concentration of air pollutants (26.3%) but less related to meteorological conditions (17.7%) (Figure **5-16**a), which was consistent with the VPA results in section 5.2.2 (Figure **5-14**b).

Although environmental factors had less of an effect on the bacterial abundance, bacterial richness was associated with most air pollutants and climatic conditions (Figure **5-10**), which has been documented in section 5.2.1. The patterns of diversity of airborne bacterial communities worldwide were considerably influenced by environmental variables, still, some of which were interrelated (Figure **5-17**). Latitude was highly correlated with the average AT ($R = -0.560$, $p < 10^{-14}$), and there were no significant relationships of latitude with other meteorological conditions, including AP, RH, WS, and WD. Latitude showed significant negative correlations with

concentrations of four types of air pollutants, *i.e.*, SO_2 ($R = -0.337$, $p < 10^{-4}$), CO ($R = -0.405$, $p < 10^{-6}$), $PM_{2.5}$ ($R = -0.461$, $p < 10^{-8}$), PM_{10} ($R = -0.437$, $p < 10^{-8}$), which might be caused by high population density and more human activities effects in low and intermediate latitudinal regions. In addition, the meteorological conditions showed a mutual independent relationship, while the correlations between PMs and gaseous pollutants are overall positive: $PM_{2.5}$ with PM_{10} ($R = 0.811$, $p < 10^{-37}$), CO with $PM_{2.5}$ ($R = 0.773$, $p < 10^{-31}$) and PM_{10} ($R = 0.740$, $p < 10^{-27}$), SO2 with $PM_{2.5}$ ($R = 0.382$, $p < 10^{-6}$) and PM_{10} ($R = 0.416$, $p < 10^{-7}$), NO_2 with $PM_{2.5}$ ($R = 0.458$, $p < 10^{-8}$). There is a great deal of research indicating that most emissions of SO_2, NO_2, and CO are along with the emission of PMs, which in turn contributed to the formation of secondary aerosols by decreasing particle acidity, and these secondary aerosols account for a large proportion of $PM_{2.5}$ and PM_{10} (Streets and Waldhoff, 2000; Berglen *et al.*, 2004; Tie *et al.*, 2006; Yee *et al.*, 2020). On the other hand, the relationship between O_3 and PMs is unrelated and different from that of other gaseous pollutants, perhaps because O_3 not being a primary aerosol (Rypdal *et al.*, 2009). As a result, the environmental filtering on global airborne bacterial communities was intricate and was dominated by multiple factors.

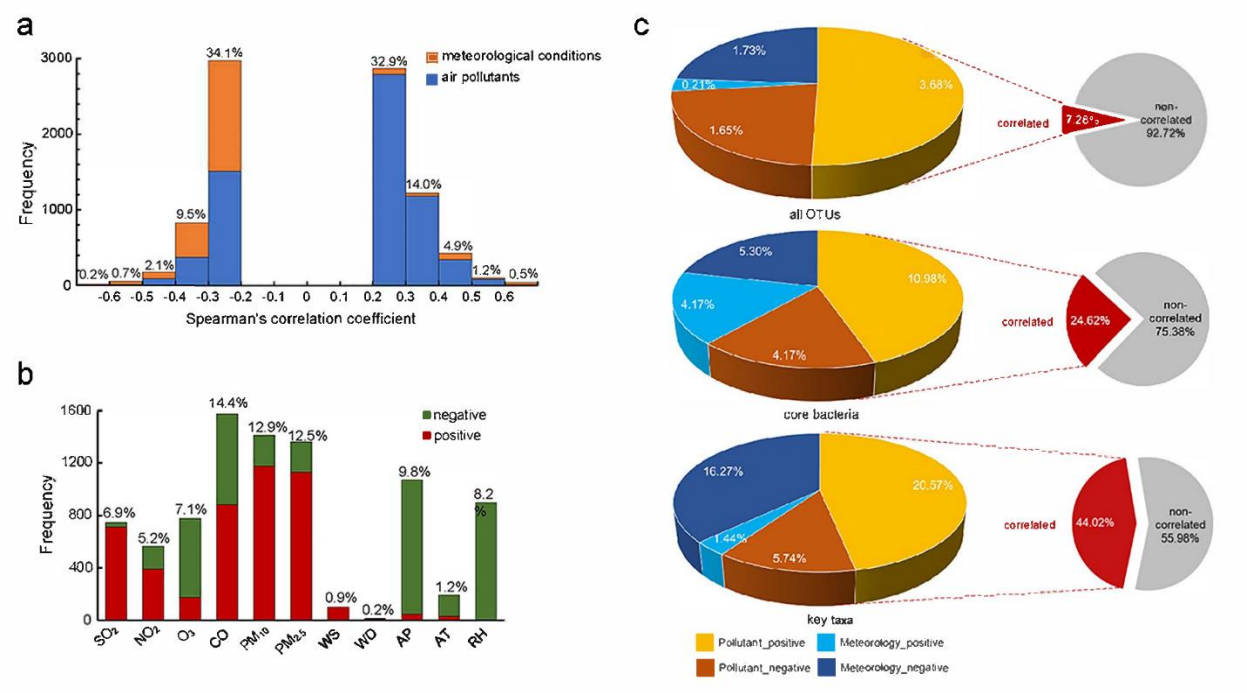


Figure 5-16 Correlation distribution between airborne bacterial community composition and environmental factors (air pollutants and meteorological conditions). (a) Histogram of the correlation frequency showing the distribution of Spearman's correlation coefficients (*i.e.*, R) between the whole global airborne bacterial communities (10,897 OTUs) and environmental factors. Only significant correlations ($p < 0.05$) were counted in the histogram. The percentage number above each pillar indicates the proportion of significant relationships that fall within that section. (b) Histogram of correlation frequency showing the distribution of significant correlations ($p < 0.05$) with diverse environmental factors in the whole global airborne bacterial communities (10,897 OTUs). The percentage number above each pillar indicates the probability of the occurrence of a significant correlation of that factor with OTUs. (c) Pie charts of correlation frequencies showing the probability of the occurrence of significant relationships (positive or negative, affected by air pollutants or meteorological conditions) in the whole airborne bacterial communities, key taxa, and core bacterial sets. The right panel highlights the ratio of the number

of existing significant relationships ($p < 0.05$) to the number of all possible connections between OTUs and environmental variables. The left panel classifies all of these existing significant correlations into four groups and gives an indication of the strength of the correlation through a percentage value.

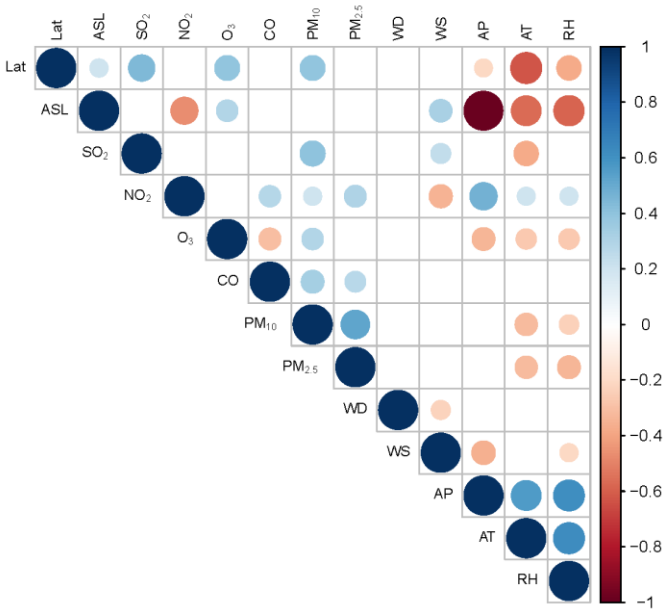


Figure 5-17 The correlation matrix among geographic locations, air quality, and meteorological conditions. The color gradient on the right indicates Spearman's rank correlation coefficients (*i.e.*, R), with more positive values (dark blue), indicating stronger positive correlations and more negative values (dark red) indicating stronger negative correlations. The sizes of the colored circles represent correlation strengths. (*n = 370 biologically independent samples*).

5.3 Mechanisms driving global airborne bacterial communities

The human imprints and environmental impacts on global airborne bacterial communities have been verified in section 5.1 and section 5.2, respectively, and the related and comprehensive driving mechanisms would be explored in this section.

5.3.1 The weakened importance of deterministic processes to microbial community assembly in high-mobility and human-impacted environments

Understanding the ecological drivers that influence community assembly is a key issue. There are two mechanisms involved in community assembly: niche-based deterministic mechanisms (including environmental filtering (*e.g.*, pH, temperature, moisture, and salinity) and various biological interactions (*e.g.*, competition, facilitation, mutualisms, and predation)) and neutral-based stochastic (including birth/death, speciation/extinction, and immigration) (Bahram *et al.*, 2016). To better quantify the relative importance of stochastic process in airborne community assembly, an index, NST was calculated based on a general null model-based framework (Ning *et al.*, 2019). This allowed for further exploration of the mechanisms shaping microbial communities and biogeographic patterns. As shown in Figure **5-18**, the importance of stochastic processes increased from topsoil (53.2%) to marine (61.3%) and air (72.4%) ecosystems.

Bioaerosols are widespread, and large particles remain airborne for only a short time, reducing the contact periods of bacterial cells with environmental elements (*i.e.*, polycyclic aromatic hydrocarbon, endotoxin, heavy metals, etc.) and other microbial cells in the air (Fröhlich-Nowoisky *et al.*, 2016). This results in fewer impacts on airborne bacteria from environmental factors and species interactions, leading to less significant effects from deterministic processes in shaping airborne bacterial communities. On the other hand, environmental gradients in the

atmosphere are inconspicuous due to constant airflow, reducing the selection pressure of environmental variables on airborne bacteria (Burrows *et al.*, 2009b), further weakening the influence of deterministic processes on airborne bacterial community assembly.

Furthermore, it was found that human-affected areas exhibited less influence from deterministic processes on microbial community assembly compared to natural areas. This was observed in atmospheric environments where urban areas had less control than offshore and terrestrial backgrounds, as well as in topsoil ecosystems where croplands had less control than grasslands and forests (Figure **5-18**b). This disturbance by frequent human activities reduces the natural environmental gradients (Stamenković *et al.*, 2019), thereby weakening the effects of environmental factors on microbes. In addition, a disrupted microbial structure network could lead to an increase in stochastic community assembly, regardless of atmospheric type (Figure **5-3**c) or different habitats (Figure **4-4**), as documented in previous sections. As seen in the terrestrial atmosphere, it could be hypothesized that coastal airborne microbiomes might also be less affected by environmental filtering and bacterial interactions compared to ocean areas.

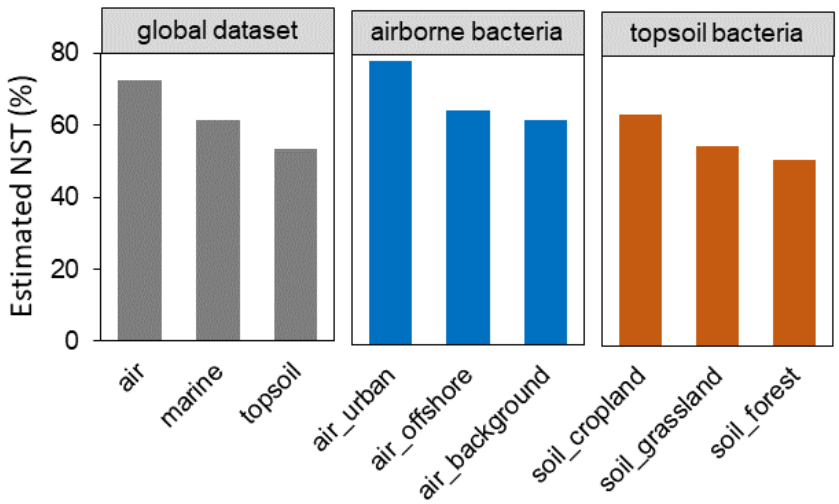


Figure 5-18 The relative importance of stochasticity represented by NST (%) in bacterial community assembly processes of the global airborne (*n = 370*), marine (*n = 62*), and topsoil (*n = 65*) bacterial communities, respectively.

5.3.2 Mechanisms shaping airborne bacterial communities

Extensive analysis was conducted on the direct impacts of 20 different environmental factors on bacterial communities, including diversity (Figure **5-10**), biomass (Figure **5-11**), core bacterial set (Figure **5-14**a), key taxa (Figure **5-14**b), and even the abundance of each OTU (Figure **5-16**) in section 5.2. This analysis revealed that the distribution of global airborne bacteria may have been influenced by geographic locations, meteorological parameters, and air quality conditions. However, these environmental factors were found to be interrelated (Figure **5-17**), making the mechanisms shaping global airborne bacterial communities more complex. In order to explore the direct and indirect relationships between environmental factors and bacterial communities, as well as the causalities among these variables and the overall contributions of each factor, SEM was applied to the global airborne bacterial dataset (Bahram *et al.*, 2018; Wu *et al.*, 2019). The findings indicated that bacterial communities are influenced by numerous factors (Figure **5-19**). These included geographical locations, which could have direct or indirect impacts on airborne bacteria due to their influence on environmental factors. Furthermore, biotic interactions such as key taxa, core communities, and bacterial richness had significant effects on microbial communities. The study also calculated the overall effects of environmental filtering (β = 3.06) and bacterial interactions (β = 0.25) on shaping communities. T Overall, deterministic processes involving a

combination of biotic and abiotic factors played a crucial role in determining the structure and distribution of microbial communities, with environmental filtering being the primary determinant.

The results of the study demonstrate that stochastic processes have a strong influence on global bacterial communities, with a relative importance of 72.4%, 61.3%, and 53.2% observed in atmosphere, ocean, and soil ecosystems, respectively (Figure **5-18**). Additionally, it was found that almost half of the airborne bacteria (46.3% on average) originated from other environments (Figure **4-13**), which further highlights the role of stochastic processes in shaping community assembly. To investigate the mechanisms responsible for shaping global airborne bacterial communities, a VPA was conducted that considered both environmental filtering (deterministic processes) and source contribution (stochastic processes) (Figure **5-20**). The analysis revealed that airborne bacterial source profiles had the greatest impact on communities, accounting for 43.68% of the structural variations, which was substantially higher than that of air quality (29.41%) and meteorological conditions (25.78%). Due to the dynamic nature of the air ecosystem, certain key environmental variables suffered from considerable uncertainties, which increased the significance of neutral processes in driving airborne bacterial communities (Fröhlich-Nowoisky *et al.*, 2016). The study also found that human activities had a notable impact on air quality and airborne bacterial source profiles, explaining around 60% of the variation in community structures. This finding corroborates the view that humans mainly influence airborne bacteria by reducing environmental filtering effects and increasing human-related source contributions. Importantly, the three primary factor groups significantly influenced the entire community, explaining more than 80% of the variations. Hence, it can be concluded that global airborne bacterial communities are primarily influenced by atmospheric environments and bacterial communities in the surrounding ecosystems.

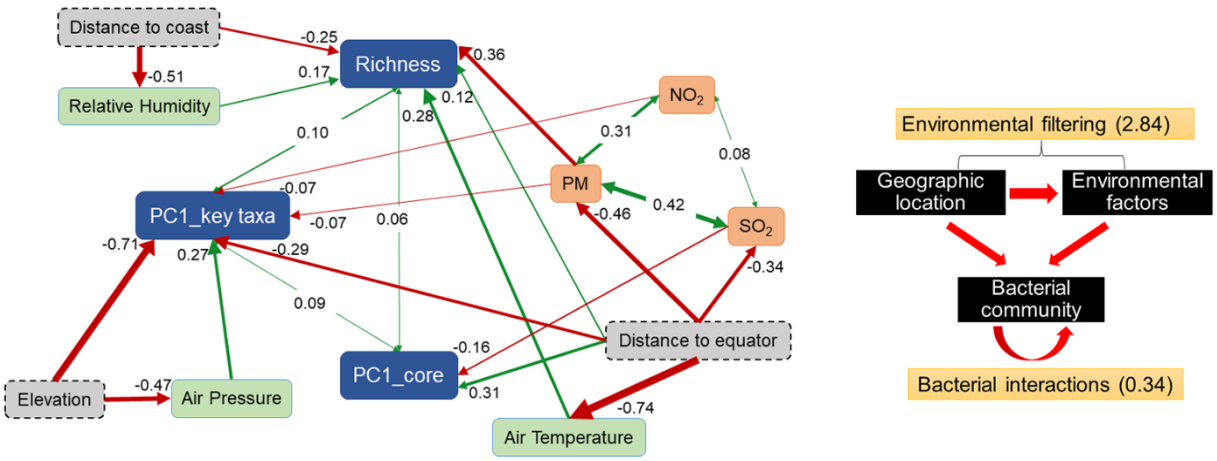


Figure 5-19 SEM showing the direct and indirect relationships among geographic locations, environmental variables, bacterial interactions, and community composition. The diversity of global airborne bacteria can be explained by a combination of biotic and abiotic factors. The key taxa and core bacterial communities are represented by the PC1 from the Euclidean similarity index principal coordinate analysis. "PM" represents the collected particle mass concentration of each sample. The one-way arrows represent causal relationships, while the two-way arrows represent mutual effects. The thickness of the lines shows the strength of the association (green – positive; red – negative). Standardized path coefficients (β) all appear near the corresponding pathways. The goodness of fit was acceptable: Model χ^2 = 23.008, d.f. = 16, P (chi-square test) = 0.114, RMSEA = 0.049, SRMR = 0.028, CFI = 0.973, TLI = 0.950, n = 370 biologically independent samples. The right panel summarizes the total effects of environmental filtering and bacterial interactions on shaping communities.

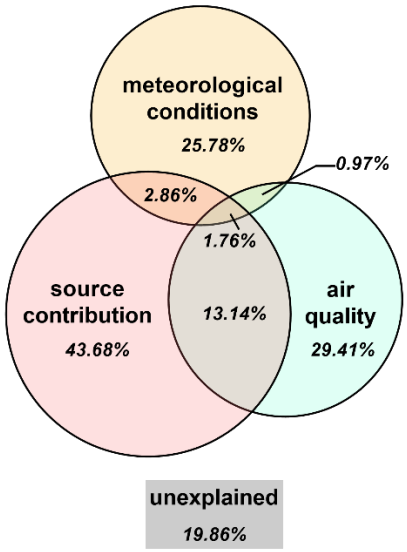


Figure 5-20 A VPA was conducted to determine the relative contributions of air quality, meteorological conditions, and source contributions to the variations observed in global airborne bacterial communities. The overlapping areas in the VPA represent the joint effect explained by two or three factor groups combined, while the percentage below each group name reflects the variance explained by a single group. The "Unexplained" category denotes the variance that could not be attributed to any of these three groups.

5.4 Summary

This chapter presented a comprehensive understanding of the underlying mechanisms driving global airborne bacterial communities. It was revealed that atmospheric environments and bacterial communities in the surrounding areas were the main factors impacting global airborne bacterial communities. This expanded on the previous understanding of the biogeography of

airborne microbiomes, and the significant contribution of human-related sources to urban airborne bacteria, providing a basis for predicting changes in airborne bacterial dynamics at global or regional scales in response to environmental changes, air pollution conditions, and human activities. The major findings of this chapter can be summarized as follows:

a) Urbanization did not impact the airborne bacterial richness but altered the community composition and structure. Concretely, the taxonomic composition of airborne bacteria was significantly impacted by urbanization, leading to changes in some phenotypic characteristics. Anthropogenic impacts destabilized the network structure of urban airborne bacterial communities by reducing transitivity and increasing the average shortest path length. In urban areas, humans inhaled less abundant airborne bacteria but more abundant and diverse pathogens, especially human-related ESKAPE pathogens that pose a high risk of mortality, increasing the risk of pathogenic infection through inhalation.

b) Geographic locations impact airborne bacteria directly or indirectly through typical environmental factors, while biotic interactions also affect microbial communities, with key taxa, core communities, and bacterial richness showing significant interactions. Global airborne bacterial community structure and distribution were attributed to various biotic and abiotic factors, with the most determinant part being environmental filtering.

c) The airborne bacterial community assembly was more driven by stochastic processes than that in other environments and showed a decreased gradient from urban areas to pristine areas. Human activities disturbed original environmental conditions, reduced environmental filtering impacts, and increased the human-related bacterial contribution to air, finally leading to enhancing importance of stochastic processes in community assembly.

d) The variations in air qualities, meteorological conditions, and source profiles significantly affect whole communities and explain over 80% of the community variations, indicating that global airborne bacterial communities are mainly impacted by atmospheric environments and bacterial communities in the surrounding ecosystems.

In summary, this chapter extensively explored the mechanisms driving airborne bacterial communities at a global scale from anthropogenic and environmental perspectives. Despite the highly dynamic and mobile nature of air ecosystems, the structure and distribution of these communities are largely influenced by local environments, especially in terms of source contributions and air quality conditions dominated by human activities. Despite the highly dynamic and mobile nature of air ecosystems, the structure and distribution of these communities are largely influenced by local environments, especially in terms of source contributions and air quality conditions dominated by human activities. Anthropogenic impacts on airborne bacteria are reflected in fewer biomass loadings, greater pathogenic abundance, destabilized network structures, and phenotypic differentiation, primarily due to the reduced environmental filtering effects and increased human-related source contributions. The increased presence of pathogenic bacteria in urban air poses a potential risk of inhalation infections, although the threat of antibiotic-resistant pathogens remains uncertain and requires further exploration. Consequently, the next chapter analyzes a vast amount of metagenomic data to investigate the distribution and dissemination of airborne antibiotic resistance genes and their potential hosts, providing a comprehensive understanding of the global airborne AMR risks.

Chapter 6 Risks of Airborne Antimicrobial Resistance to Human Health at a Global Scale

AMR happens when microbes develop the ability to defeat the drugs designed to kill them, which makes infections harder to treat and increases the risk of disease spread, severe illness, and death. Beyond the pathogens that cause infection and health risks, antibiotic-resistant pathogens and "superbugs" represent one of the most critical threats to public health and modern health care. Following up on the different profiles of airborne pathogens between urban and background areas and the higher infection risks of inhaled pathogens in urban areas in the last chapter, we conducted an airborne AMR investigation at the global scale and in association with other environmental compartments and explored the global standard leading to the differences in the burden of AMR risks across countries and regions. A global metagenomic dataset based on air samples was established to first attempt the global airborne AMR, especially with regard to the evaluation of quantified ARG-related health risks. This chapter revealed the global patterns of 747 identified airborne ARGs as well as their potential hosting bacteria and even pathogens. Furthermore, the exposure risks were estimated based on the abundance, mobility, and host pathogenicity of each ARG in diverse air samples from a regional perspective to construct a comprehensive atlas of global airborne AMR. Overall, we aim to understand the source, dissemination, and health risks of airborne AMR pathogens and contribute to the global burden of AMR diseases in association with human exposure risks.

6.1 Broad-spectrum profile of the global ambient air resistome

The natural and engineered environment itself contains multiple functional ARGs (Biswas *et al.*, 2021); bioaerosol is one of the largest resources and habitats of ARGs (Chen *et al.*, 2022). A dataset of 262 metagenomic samples was organized and the metagenomic reads mapped to a total of 747 ARGs that potentially confer resistance to 25 drug classes of antibiotics. From an overview perspective of the global airborne ARGs, the abundance and occupancy of each ARG in this dataset were calculated to further identify the core ARGs. Moreover, the global pattern of airborne ARG distribution was also demonstrated from regional variations.

6.1.1 Identification of core ARGs at a global scale

The abundance and occupancy of each ARG were estimated in the global airborne ARG dataset, and a core ARG set was determined based on this AOR pattern (Figure **6-1**a), in which 39 ARGs (0.05% of total number of ARGs) accounted for 61.2 ± 11.8% (mean ± s.e.m.) of the sequences in air samples (Figure **6-1**b). To further identify AMR threats, the risks of airborne ARGs were evaluated and ranked based on a published framework considering anthropogenic enrichment, the mobility of genes, and the pathogenicity of hosting bacteria (Zhang *et al.*, 2021). Although the core ARGs showed an overwhelming advantage in abundance, the risk ranking of the core ARGs was lower than the whole airborne ARG ranking, which could be embodied in the lower proportion of Rank I (5%) – mobile and human-associated ARGs that are already present among pathogens and Rank II (0%) – mobile and anthropogenic enriched ARGs with novel resistance emerging from non-pathogens from all airborne ARG across the world (Rank I, 7%; Rank II, 2%) (Figure **6-1**c).

Two-thirds of global core ARGs (26/39) in the air overlapped with core ARGs in marine and topsoil habitats, as determined based on global datasets (Sunagawa *et al.*, 2015; Bahram *et al.*,

2018) (Table **6-1**). In contrast, there were no shared core species among the global bacterial communities in the air, ocean, and topsoil ecosystems, suggesting the more general applicability to various environments and the higher transferability of ARGs than microbiomes. Also, the ARG profiles on drug resistance types in the global atmospheric, aquatic, and terrestrial ecosystems were very similar (Figure **6-2**a), and more than 80% of the 747 ARGs were shared by marine and topsoil ecosystems (Figure **6-2**b); together, these illustrated the high communicability and universality of ARGs in natural environments. Moreover, ARGs were more diverse in the air than in aquatic and terrestrial habitats at a global scale, and airborne ARGs conferring multidrug resistance showed a higher proportion. Here, we hypothesized that air played a role as the medium in which ARGs were carried from one habitat to another on the Earth's surface and gradually enriched the diversity and risks of AMR during this process. To further determine the crucial role of the "One Health" framework, we also estimated the potential sources of global airborne ARGs in section 6.2.2.

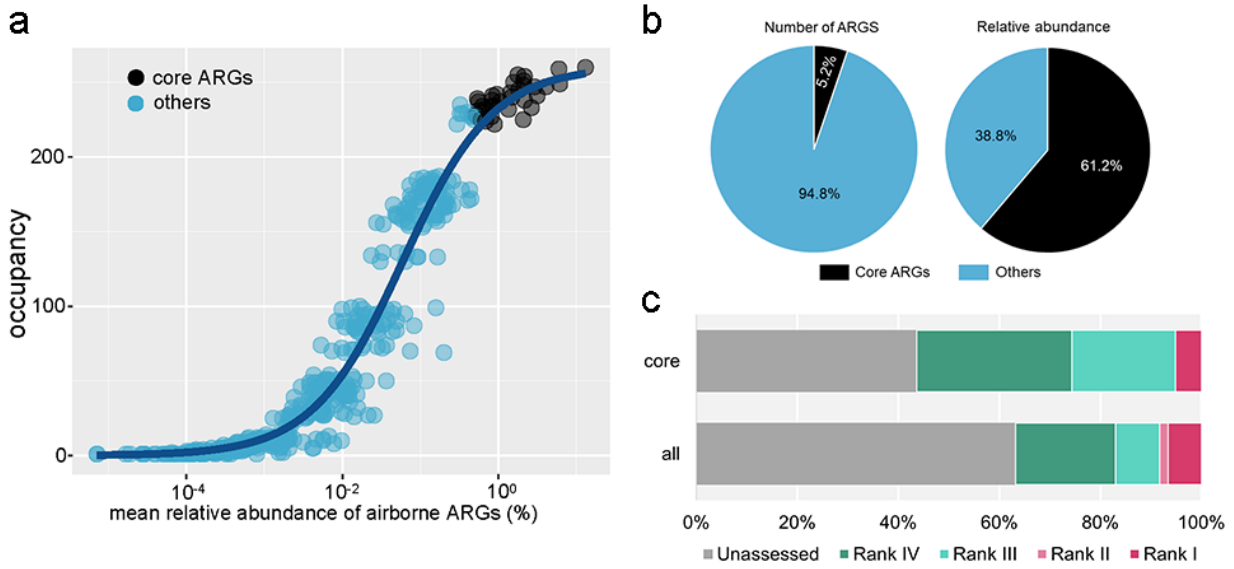


Figure 6-1 Identification of core ARGs in the global atmosphere. (a) AOR: mean relative abundance (*y-axis*) and occupancy (*x-axis*) plot. The mean relative abundances were estimated by

averaging the relative abundances of each ARG in all air samples; occupancy represents the proportion of samples (*n=262*) in which the ARG was detected. The fitted model (sigmoid curve) was occupancy *vs.* logarithm of abundance for each species, and the dark blue solid line is the global fit to all ARGs and samples. (b) The number proportion and relative abundance of the global core ARGs compared with those of the remaining ARGs. (c) Risk ranking of (core) ARGs detected in the global atmosphere according to an omics-based framework based on the three criteria: enrichment in human-associated environments, gene mobility, and host pathogenicity (Zhang *et al.*, 2021). Rank I represents the highest risk and Rank IV represents the lowest risk.

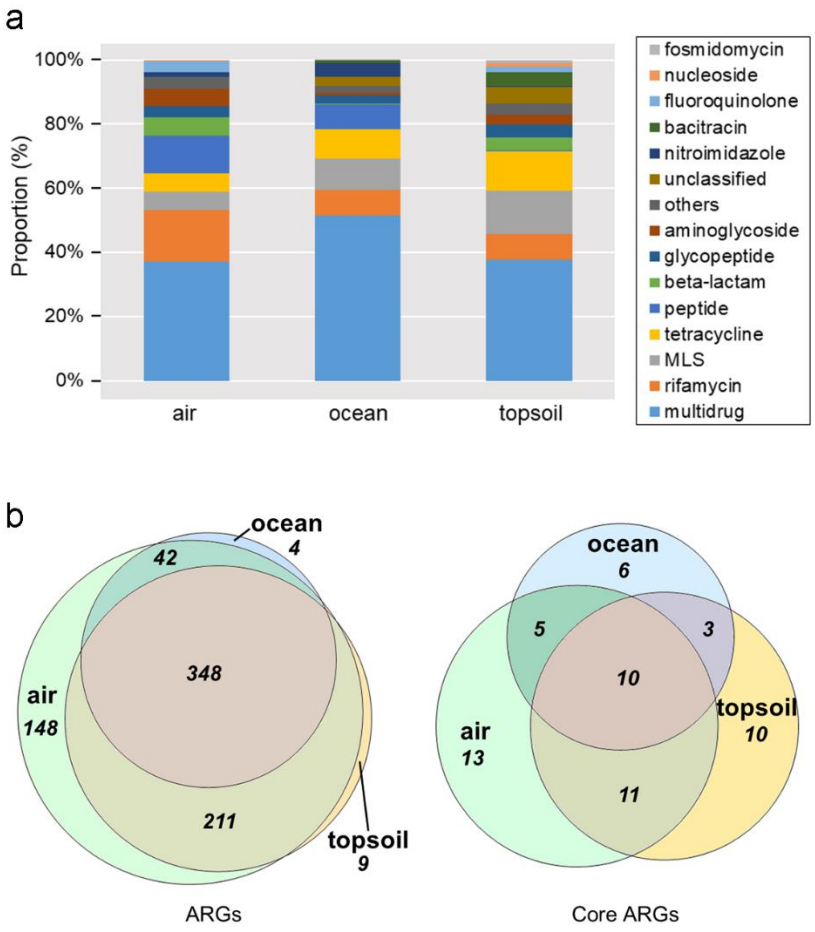


Figure 6-2 Comparison of ARG profiles in the three largest ecosystems, namely atmospheric, aquatic, and terrestrial ecosystems. (a) Composition of the antibiotic resistome in the three largest ecosystems globally. Air, ocean, and topsoil ecosystems contained 262, 72, and 74 independent samples, respectively. (b) Number of exclusive and shared (core) ARGs in air, marine, and topsoil across the globe.

Table 6-1 Summary of core ARGs in the global air, marine, and topsoil ecosystems. The core ARGs in the global air shared with counterparts in marine or topsoil ecosystems were emphasized in bold.

TYPE	Global air ARGs and relative abundance (%)		Global marine ARGs and relative abundance (%)		Gloal topsoil ARGs and relative abundance (%)	
multidrug	**MULTIDRUG_ABC_TRANSPORTER**	0.059	MULTIDRUG_ABC_TRANSPORTER	0.022	MULTIDRUG_ABC_TRANSPORTER	2.098
	OMPR	0.031	OMPR	0.007	OMPR	0.533
	TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR	0.029	TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR	0.009	TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR	0.807
	ACRB	0.022	ACRB	0.002	ACRB	0.176
	MEXF	0.021	MEXF	0.003	MEXF	0.994
	MSBA	0.016	MSBA	0.019		
	MEXW	0.005	MEXW	0.001		
	MTRA	0.041			MTRA	0.378
	EMRB	0.011			EMRB	0.200
	MEXE	0.009			MEXE	0.307
	OLEC	0.008			OLEC	0.659
	EMRA	0.008			EMRA	0.299
	MAJOR FACILITATOR SUPERFAMILY TRANSPORTER	0.007			MAJOR_FACILITATOR_SUPERFAMILY_TRANSPORTER	0.732
	MUXB	0.005			MUXB	0.162
	MEXT	0.006	MEXT	0.003	TRUNCATED_PUTATIVE_RESPONSE_REGULATOR_ARLR	0.619
	ABCA	0.026	MDSB	0.002	EMRB-QACA_FAMILY_MAJOR_FACILITATOR_TRANSPORTER	0.161
	MEXB	0.008	BAER	0.001		
	MEXK	0.008	SMER	0.001		
	CEOB	0.007				
	BPEF	0.006				
	OPRM	0.005				
rifamycin	**RPOB2**	0.131	RPOB2	0.039	RPOB2	0.914
	RIFAMPIN_MONOOXYGENASE	0.010			RIFAMPIN_MONOOXYGENASE	0.399
	RPHB	0.005				
	ADP-RIBOSYLATING_TRANSFERASE_ARR	0.013				
bacitracin	**BACA**	0.061	BACA	0.002	BACA	0.424
					BCRA	1.108

TYPE	Global air ARGs and relative abundance (%)		Global marine ARGs and relative abundance (%)		Gloal topsoil ARGs and relative abundance (%)	
polymyxin	**ROSA**	0.018	ROSA	0.329		
	ROSB	0.021				
	ARNA	0.006				
tetracycline	**TETA(48)**	0.014	TETA(48)	0.002	TETA(48)	0.756
	TETA	0.021	TETP	0.001	TETP	0.209
					OTRC	0.197
beta_lactam	OXA	0.029			PBP-1A	0.205
glycopeptide	**VANR**	0.022	VANR	0.006	VANR	1.032
	VANS	0.006			VANS	0.158
			VANRI	0.002	VANRI	0.261
					VANH	0.241
MLS	**MACB**	0.018	MACB	0.014	MACB	1.753
					VGBC	0.156
aminoglycoside	**KDPE**	0.009			KDPE	0.516
	APH(3')-I	0.007				
peptide	**UGD**	0.015	UGD	0.029		
			CAMP-REGULATORY_PROTEIN	0.002		
kasugamycin	**KASUGAMYCIN_RESISTANCE_PROTEIN_KSGA**	0.005	KASUGAMYCIN_RESISTANCE_PROTEIN_KSGA	0.002		
quinolone					PATA	0.512
puromycin					PUROMYCIN_RESISTANCE_PROTEIN	0.371
triclosan					TRIC	0.193
trimethoprim			DFRA3	0.001	DFRA3	0.162
pleuromutilin			TAEA	0.003		
mupirocin			STAPHYLOCOCCUS_MUPA_CONFERRING_RESISTANCE_TO_MUPIROCIN	0.002		

6.1.2 Global patterns of airborne ARG distribution

Of the 747 ARGs detected in the air, 253 conferred multiple drug resistances, and 494 were genes with resistance to only one drug class (Appendix 4.1); the three most abundant ARG classes were the multidrug, rifamycin, and peptide classes, which accounted for 64.1% of total ARG sequences detected in the entire global airborne metagenomic dataset (Figure **6-2**a). A heatmap was constructed to show the relative abundance of ARG classes (normalized to the 16S rRNA gene) in the air across 22 sampling areas worldwide. As shown in Figure **6-3**, the genes providing resistance to multidrug (0.3±0.39, normalized by 16S rRNA gene, hereafter), rifamycin (0.16±0.18), and peptide (0.12±0.13) were the top three most abundant types of ARGs in the global atmospheric ecosystem, followed by MLS (0.06±0.07), beta-lactam (0.06±0.08), tetracycline (0.05±0.05), aminoglycoside (0.05 ± 0.06), glycopeptide (0.03 ± 0.06), fluoroquinolone (0.03 ± 0.04), and nitroimidazole (0.01 ± 0.01). Moreover, according to the hierarchical clustering results in the heatmap, the variation trends in the relative abundance of ARG classes in the ambient air were similar across the sampling areas, showing that the ARG class profiles were relatively consistent in the air.

The sampling sites involved in the global airborne metagenomic dataset were clustered into two groups, urban (including cities, industrial regions, and rural areas) and background areas (including remote mountains, offshore environments, and the Antarctica region), according to the degree of anthropogenic impacts. Of the 262 metagenomic sequencing datasets, urban and background air samples exhibited considerable variance of ARG composition based on Bray−Curtis dissimilarity ($p < 0.05$, Figure **6-4**a), indicating the human imprints on AMR in global ambient air. Besides, the airborne ARG compositions and abundance on subtype level were significantly impacted by the

sampling sites and countries, with a concrete manifestation in the cluster of air samples collected in the same area in an NMDS analysis (Figure **6-4**b). Additionally, particle size also impacted AMR profiles across the globe, and the ARG compositions were similar between $PM_{2.5}$ and PM_{10}, distinguishing them from the TSP samples (Figure **6-4**c), which was in line with a previous study showing no significant differences in ARG number and reads between $PM_{2.5}$ and PM_{10} samples in Beijing (Qin *et al.*, 2020). In summary, human activities, geographic locations, and particle sizes significantly contributed to ARG distributions and profiles in the global atmosphere.

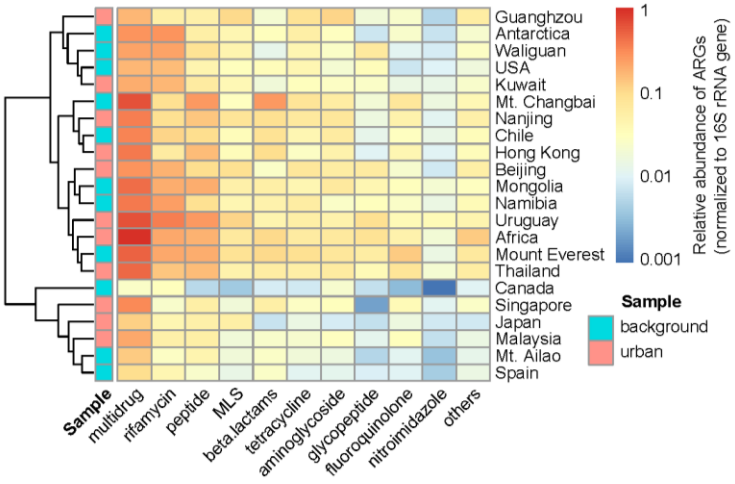


Figure 6-3 Heatmap showing the relative abundance profile of ARG classes (normalized to the 16S rRNA gene) in the air across 22 sampling areas worldwide.

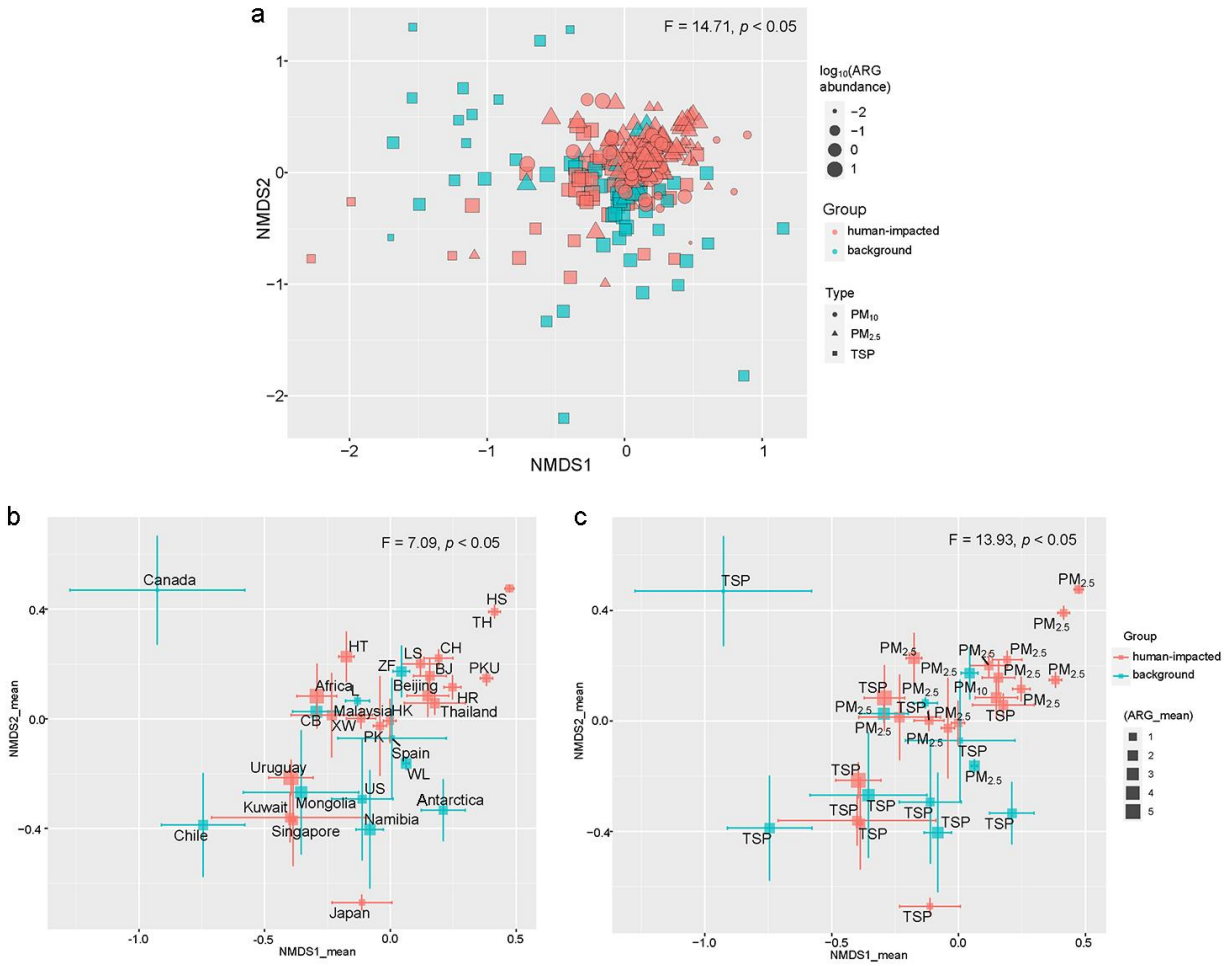


Figure **6-4** NMDS analysis based on Bray−Curtis dissimilarity showing the disparities of ARG profiles between urban and background air (a), and indicating that the global ambient air with different sampling sites (b) and particle sizes (c) harbors various ARG composition on subtype level.

6.2 Anthropogenic impacts on global airborne ARGs

6.2.1 Anthropogenic impacts on variations and enrichment of global airborne ARGs

As shown in Figure **6-5**, there were no significant differences in airborne ARG type abundance in urban and background areas ($p = 0.103$); however, the anthropogenically enriched ARGs identified according to a previous study (Zhang *et al.*, 2021) showed higher abundance in urban air than background air at a global scale ($p < 0.05$). Besides, urban and background air samples exhibited considerable variance of ARG composition based on Bray−Curtis dissimilarity ($p < 0.05$, Figure **6-4**a). A series of indicator ARGs that effectively discriminated the airborne ARG profiles between urban and background areas were further identified based on LEfSe analysis (Figure **6-6**a). The indicator ARGs in urban air belonged to various drug resistance types including *vatE*, *ermC*, *ermF*, *ermX*, *erm(TR)*, and *erm(36)*, conferring resistance to MLS; *aph(3')-I, aac(3)-II, aad(9)*, and *ant*(9)-I, conferring resistance to aminoglycoside; *tet*Z, *tet*P, *tet*S, and *tet*(K), conferring resistance to tetracycline; *emrB*, *qacG*, and *tolC*, conferring resistance to multidrug; *mecA* and *m*ecR1, conferring resistance to beta-lactam; and ble_{MBL}, *catB*, *fosB*, and *sul1*, conferring resistance to other antibiotics (LDA score > 3.0, $p < 0.05$). At the same time, 9 indicator ARGs in background air belonged to beta-lactam, multidrug, peptide, and polymyxin resistance types. In general, the dominant urban airborne ARG indicators were the genes conferring resistance to pervasively used antibiotics, such as tetracycline, MLS, and aminoglycoside, which indicated that human activities, especially the use of antibiotics, contributed to the distribution and dissemination of ARGs (Figure **6-6**b).

It was observed that air samples clustered according to their sampling sites and countries, showing different ARG compositions and abundance (Figure **6-4**b). A report on global antibiotic use revealed that different countries owned different levels of antibiotic consumption and dependence (Klein *et al.*, 2018) (Figure **6-7**). Besides, the long-term exposure of bacteria to sub-minimum inhibitory concentrations (sub-MICs) of antibiotics led to a stepwise increase in their inhibitive

antibiotic concentrations (Marvasi *et al.*, 2017), consistent with the finding that the relative abundances of airborne ARG types referred to five common antibiotic types were positively correlated with the global antibiotic consumption of corresponding antibiotic types in the previous study (Li *et al.*, 2018). Given this perspective, the abundance of indicator ARGs in the global urban air should be higher in the countries with relevant antibiotic use. Concretely, we found three antibiotic consumptions that exhibited significantly positive impacts on the total urban abundances of urban indicator ARGs conferring resistance to the related antibiotics, respectively (aminoglycosides: $R^2 = 0.554$, $p < 0.001$; tetracyclines: $R^2 = 0.526$, $p < 0.001$; beta-lactam: $R^2 = 0.568$, $p < 0.001$), while the other indicator ARGs in global urban air which conferred resistance to multidrug, MLS, and chloramphenicol were negligibly affected by related antibiotic use but were driven by more complicated factors. (Figure **6-8**). In particular, multidrug and MLS involve a variety of drugs, which cannot be directly reflected by antibiotic use. Therefore, the usage of aminoglycosides, tetracyclines, and beta-lactam in the clinic, agriculture, aquaculture, and other fields should be more prudent to reduce the exposure concentration and selective pressure of related indicator ARGs.

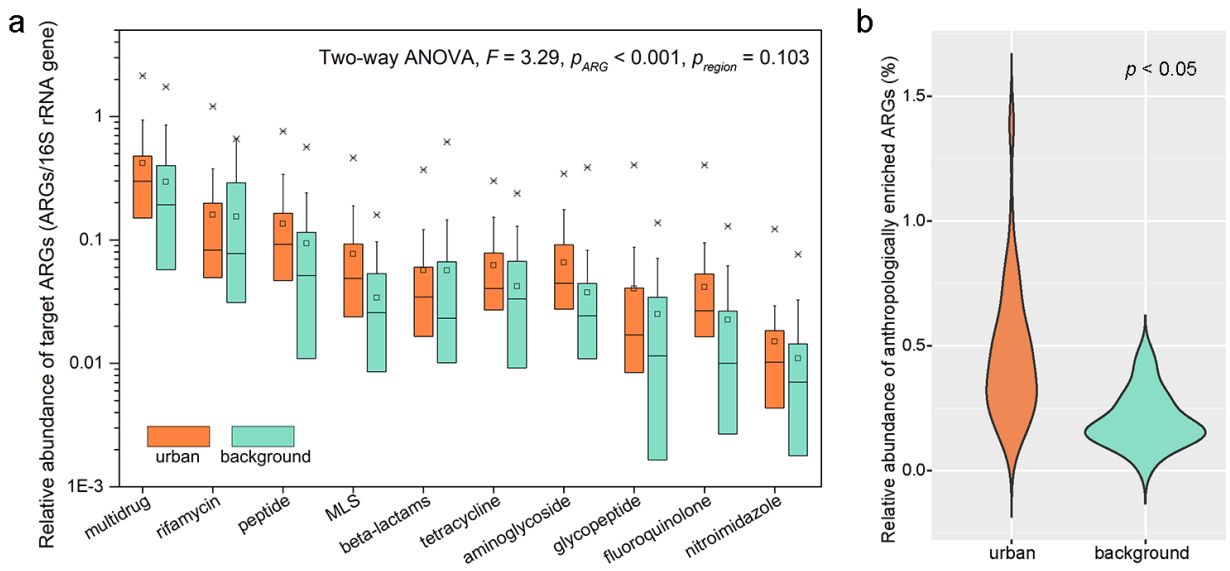


Figure 6-5 Comparison of airborne ARG abundance between urban and background areas across the globe. (a) Relative abundance of ARGs on class level across the globe. No significant differentiation in airborne ARG type abundance in urban and background areas (p_{region} = 0.103). (b) Relative abundance of anthropogenically enriched ARGs.

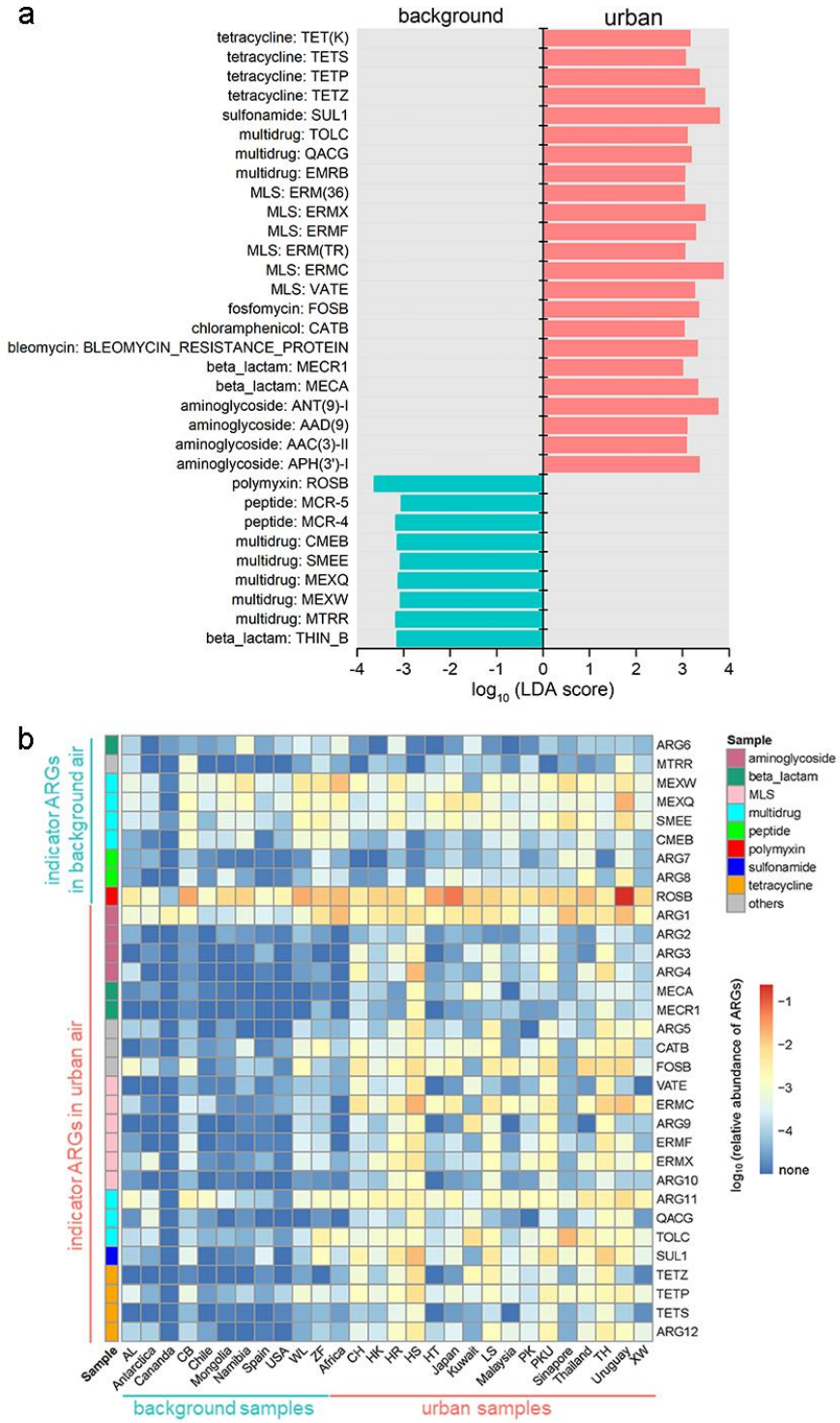


Figure 6-6 Indicator ARGs in urban and background air across the globe. (a) Major significantly discriminative ARGs and drug resistance types between urban air and background air samples as

revealed by an LEfSe analysis (LDA score [log_{10}] > 3.0). (b) Heatmap showing the relative abundance profile of indicator ARG subtypes (normalized to the 16S rRNA gene, log_{10} transformed) in urban and background air across global sampling locations.

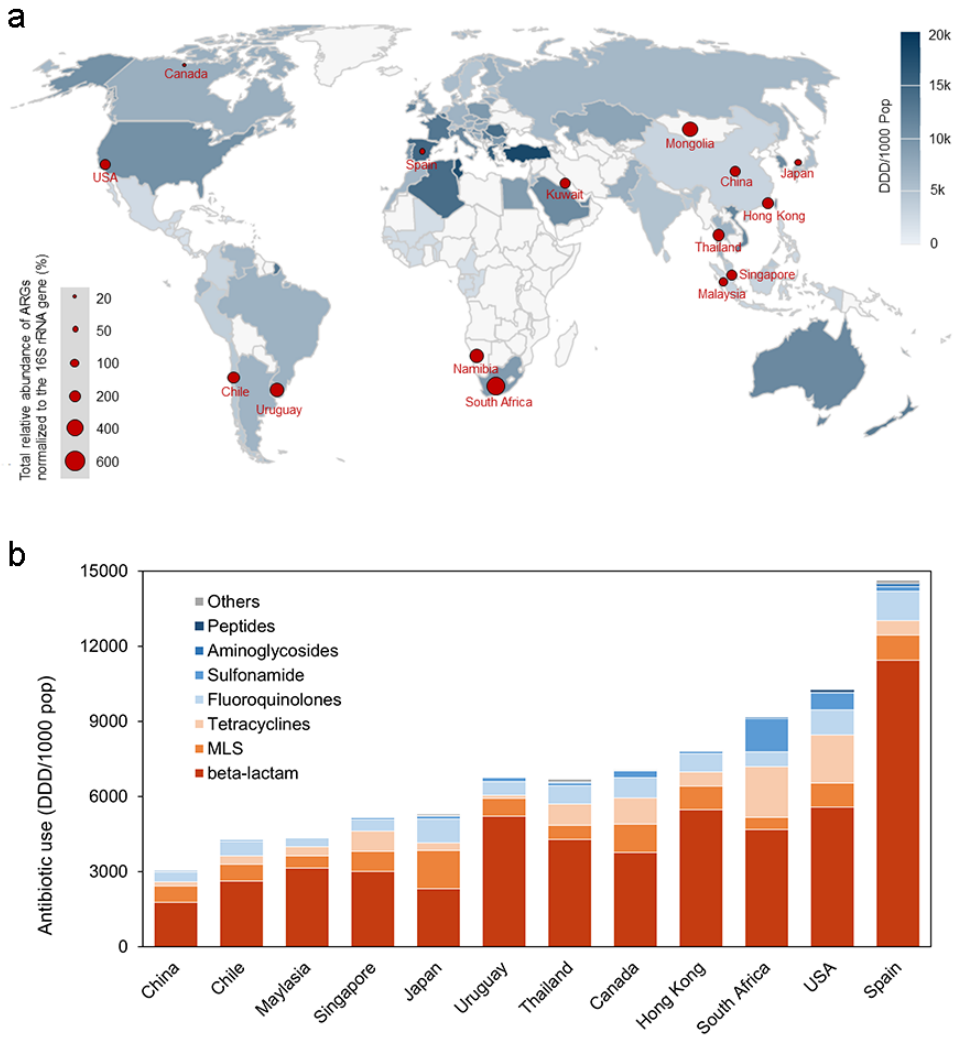


Figure 6-7 Geographic convergence of antibiotic use and total airborne ARG abundance by country at a global scale. (a) The total antibiotic use map and the relative abundance of total ARGs in the atmosphere. The amount of annually used antibiotics (defined daily doses/1000 people,

DDDs/1000 Pop) in 2015 in each country was proportional to the opacity of blue color. Each point indicated one country, rounded to the nearest degree, with point size reflecting the total abundances of airborne ARGs normalized to the 16S rRNA gene. (b) Global consumption and usage of antibiotics belonging to eight major classes in humans (Klein *et al.*, 2018).

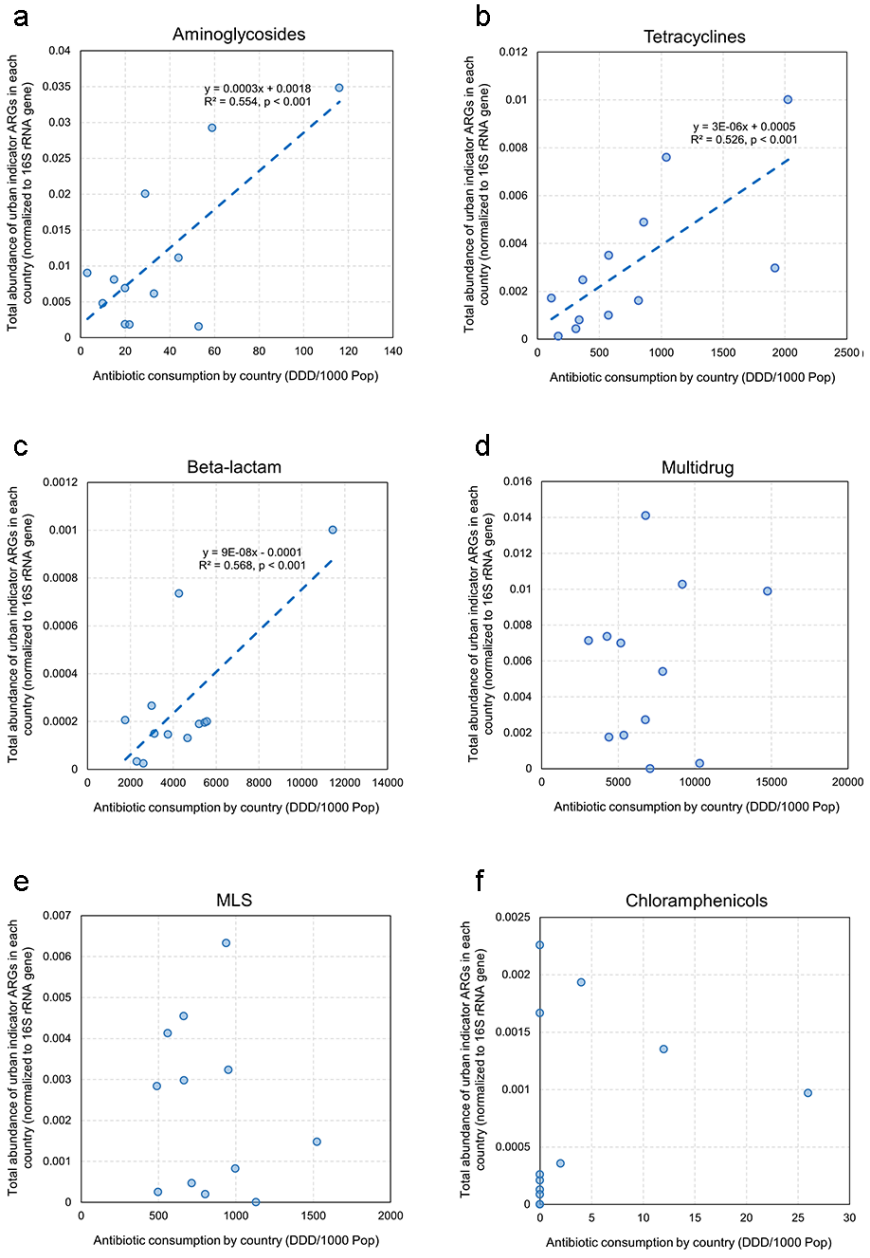


Figure 6-8 The linear regressions of total abundance of urban airborne indicator ARGs conferred resistance to various antibiotics with the corresponding antibiotic consumption by country. The significant correlation (Pearson, $p < 0.001$) is depicted using a dashed line.

6.2.2 Source tracking of global airborne ARGs showing the importance of anthropogenic sources in urban air

To further explore the influencing factors on airborne ARGs, the potential sources of airborne ARGs in various regions were predicted by SourceTracker2 (Knights *et al.*, 2011) with the following assumptions: (1) there were no changes in the air dissimilation processes, and (2) there might not be other processes that could alter the original source profiles. Among the 16 major habitats, urban air exhibited more similarity to urban facilities (WWTPs and hospitals), while background air bore a closer relation to plant ecosystems (Figure **6-9**a). Echoing the distinguished airborne ARG profiles between the urban and background sites (Figure **6-4**a), their corresponding sources were also relatively different (Figure **6-9**b). There appeared to be a dominant contribution of airborne ARGs from soil (19.1±11.3% in background and 11.4±9.2% in urban air) and plants (10.9±11.3% in background and 10.2±8.9% in urban air) among the identified sources. Hospitals (including air and sewage) and WWTPs (including air, sewage, and sludge) contributed a lot to urban airborne ARGs: 13.2±11.8% and 15.3±12.5%, respectively. As acknowledged ARG hotspots in urban areas (Wu *et al.*, 2022; Xie *et al.*, 2022), air particles emitted from WWTPs and hospitals could be major anthropogenic sources of ARGs in global human-impacted atmospheric environments, which may explain the higher abundance of anthropogenically-enriched ARGs in urban air (Figure **6-5**b). In contrast to 55.72% and 68.41% of unknown sources of global airborne

bacteria in urban and background areas (Figure **4-13**), only 11.6% and 31.9% of ARG sources were undetermined in global urban and background air, respectively, illustrating that source contributions showed the main determinant role in shaping the global distribution of airborne ARGs. Although human-related sources played relatively more important roles in shaping airborne bacterial communities (23.2%) than other potential sources, the ARG hotpots, including WWTPs, hospitals, and landfills, contributed an average of 38.25% to urban airborne ARGs, suggesting the higher importance of anthropogenic impacts on airborne ARGs. The inequality between the source profiles of bacteria and ARGs in the atmosphere might be caused by the different enriching abilities of bacteria with specific taxonomy from various sources to ARGs.

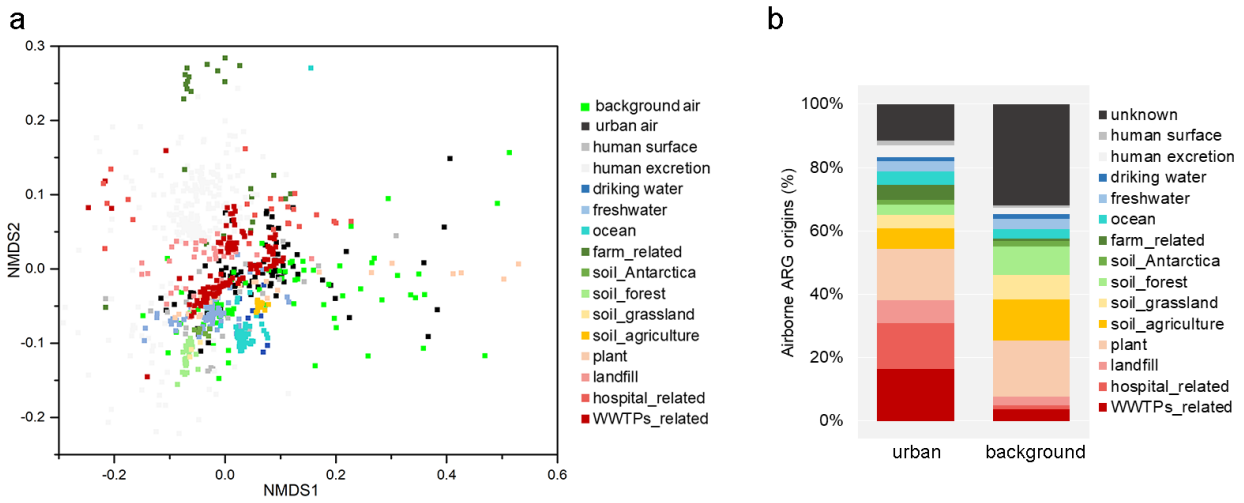


Figure 6-9 Comparing ARG profiles in ambient air and with other habitats across the globe. (a) NMDS analysis of ARG profile at the ARG subtype level of different sample types across the globe used in source tracking. For comparison, we merged our ARG subtype table (*n = 156 air samples*) with the database of putative sources of airborne ARGs from various habitats such as the human surface (*n = 29*), human excretion (*n = 241*), drinking water (*n = 7*), freshwater (*n = 75*), ocean (*n = 67*), farm-related (*n = 24*), Antarctica soil (*n = 6*), forest soil (*n= 27*), grassland soil (*n*

= 3), agricultural soil (*n = 10*), landfill-related (*n = 40*), hospital-related (*n = 25*), plant (*n = 24*), and WWTPs-related (*n = 122*). Bray–Curtis distance was calculated to represent the dissimilarity of ARG profiles. (b) The percentage of airborne ARGs attributable to various potential sources determined by SourceTracker2 (Knights *et al.*, 2011).

6.3 Co-occurrence patterns of ARGs, MGEs, and potential hosts to indicate the environmental resistome risks

To explore the genetic location of ARGs, 1.93 million generated contigs were de novo assembled from the high-quality cleaned reads in the global airborne metagenomic dataset. Moreover, the ARGs, MGE, and HVF were identified from the contigs, and the taxonomic information of contigs carrying ARGs was further annotated to recognize the potential ARG hosts.

An analysis of 1.93 million assembled contigs across the global atmosphere showed that plasmids were found to be the dominant genetic locations for urban air ARG samples (Figure **6-10**a). To sum up, the proportion of identified ARGs co-occurring with MGEs (*i.e.*, integrons/transposons) on the same contig in the urban air (29.06%) was a little higher than background air (24.51%) (Figure **6-10**b), in line with the previous study that interactions between ARGs and MGEs were enhanced by increasing air pollution (Zhu *et al.*, 2021). Echoing the limited mobility of atmospheric ARGs, the corresponding potential mobile ARG profiles exhibited different patterns from all airborne ARGs detected in urban and background areas, respectively (Figure **6-11**). The urban indicator ARGs showed higher mobility in urban air than background air, such as *sul1, tet(K)*, and *mecA*, which co-occurred with MGEs in 67.6% of urban samples but only 18.7% of background samples. Besides ARGs conferring multidrug resistance playing commonly

preponderant roles in the two patterns, ARGs conferring resistance to rifamycin, and peptides dominated in the global atmospheric ARGs, while ARGs related to MLS and tetracycline collectively showed great importance in potential mobile ARG compositions. Moreover, among these potential mobile ARGs, the proportion of Rank I–II ARGs as the highest risk was much higher in urban air than in background air worldwide (Figure **6-12** and Appendix 4.2). For instance, *bacA, lnuA, ErmB,* and *ErmC,* reported to cause problems in hospitals, were frequently detected in potential mobile ARGs identified in urban air. These enriching ARGs with potential mobility and high risk to human health accelerated AMR dissemination and exacerbated the global threat.

As shown in Figure **6-13**, the taxonomic compositions of airborne bacterial community were significantly correlated with atmospheric AMR; nevertheless, the correlations showed different patterns between urban and background air. Additionally, the proportion of pathogens in the entire bacterial community was higher in urban air than the counterpart in background air, particularly the ESKAPE pathogens featuring the highest risk of mortality (Figure **5-8**a), which perhaps further heighten AMR risks to human health in cities. To further explore the associations between airborne ARGs and bacteria, particularly human-related pathogens, we identified the potential ARG hosts based on assembled contigs. Proteobacteria, Actinobacteria, and Firmicutes were highly connected to airborne ARGs encountered in background sites, carrying 36.88%, 19.09%, and 16.05% of the total identified ARGs in the airborne bacterial community, respectively, while Firmicutes were observed to be the phylum that is most closely associated with airborne ARGs (44.12% of total identified ARGs) in human-impacted areas (Figure **6-14**). At the species level, 17.3% (632/3654) of potential ARG hosts were identified as human pathogens, while only 5.0% (162/3232) of contigs located with ARGs showed potential carriage by pathogenic bacteria. Notably, those human pathogens that developed resistance to common antibiotic drugs, like ESKAPE and other

WHO priority pathogens in urgent need of new treatments, were frequently detected from these putative ARG carriers in urban air. For instance, *S. aureus*, *Pseudomonas aeruginosa*, and *Acinetobacter baumannii* were found to carry airborne ARGs in more than half of urban sampling sites, the exposure of which was of particular concern to human infection.

To further rank resistome risks among samples, assembled contigs were processed by MetaCompare based on the co-occurrence of ARGs, MGEs, and putative pathogens annotated on the assembled contigs. Urban air (21.32±3.98) generally exhibited a higher resistome risk score than background air (19.35±1.17) (Figure **6-15**), indicating that AMR in the urban samples was featured with higher abundance and the mobility of environmental resistomes and was hosted more by human pathogens than in background air samples. We hypothesized that the potential emission sources with higher anthropogenic enrichments of ARGs and pathogens could affect the surroundings and lead to a higher health risk. Moreover, any pathogenic bacteria harboring ARGs dwelling in the atmosphere could cause exposure and infection risks with inhalation, with a particular occurrence in urban areas; therefore, there is an urgent need to assess human health risks of potential environmental sources and control the leakage from the artificial sources with high AMR risks.

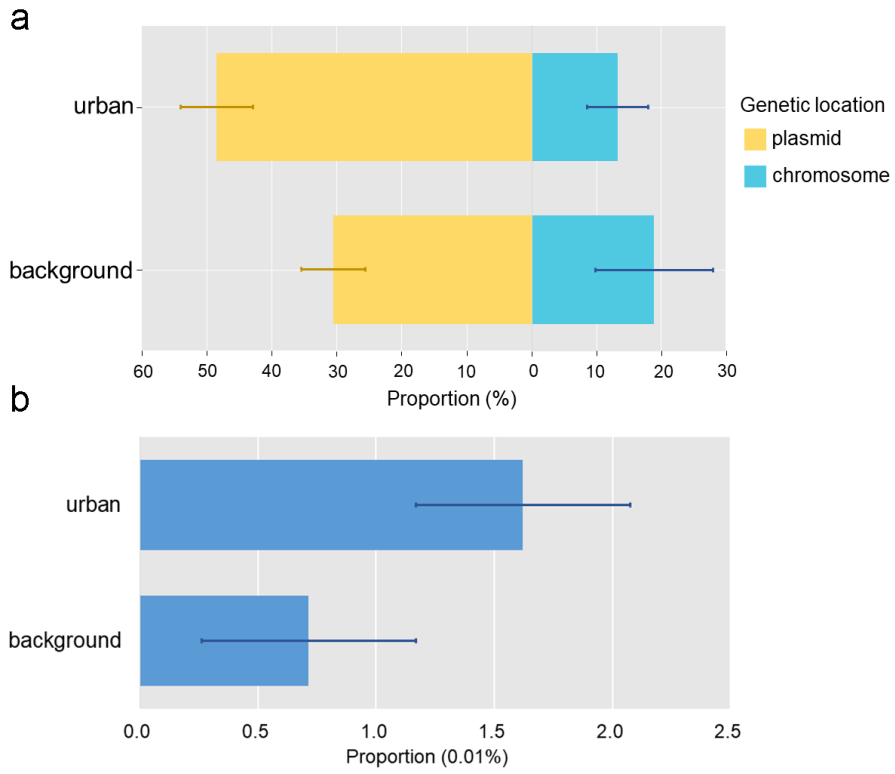


Figure 6-10 Mobility of ARGs across the global atmospheric environment. (a) The genetic locations of ARGs in bacterial genomes (unclassified fraction not shown). (b) Proportion of identified ARGs co-occurring with MGEs on the same contig.

Figure 6-11 Proportion and composition of ARGs and potential mobile ARGs (co-located with MGEs on one assembled contig) across the global air samples.

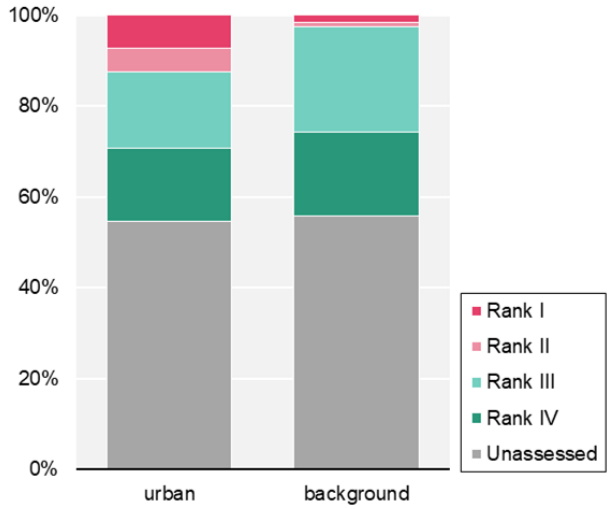


Figure 6-12 Proportion of potential mobile ARGs with different health risk ranks in urban and background air. The risk ranks of each ARG were identified from a previous study (Zhang *et al.*, 2021).

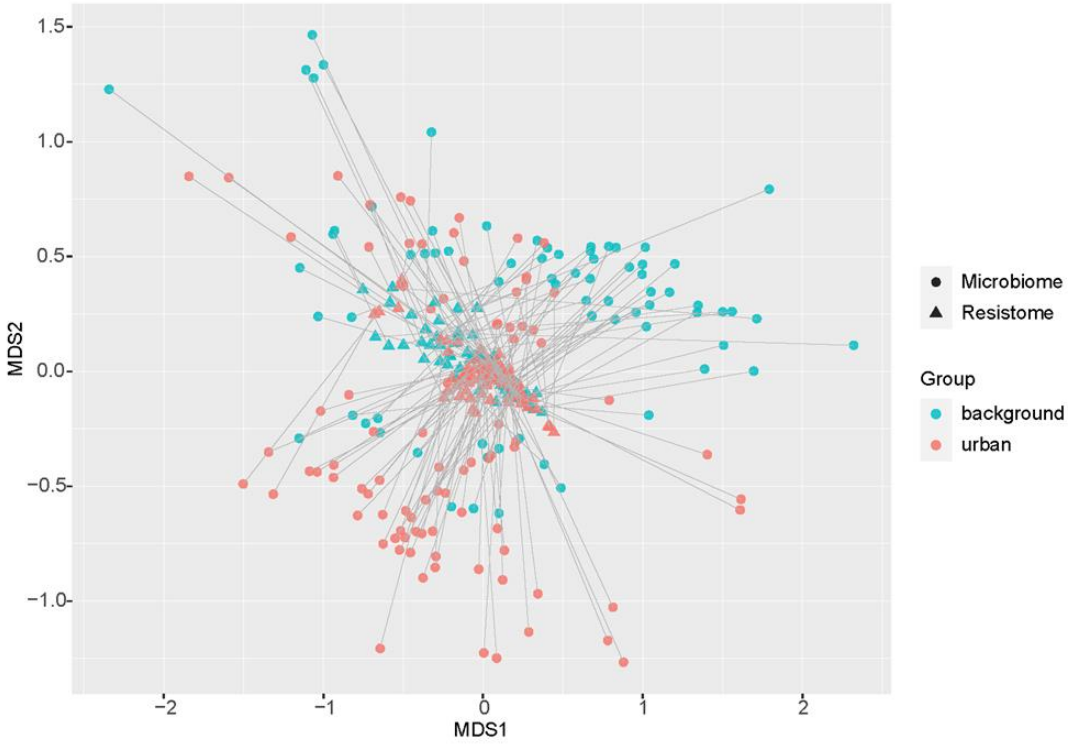


Figure 6-13 The correlation of community structure between microbiome and ARGs via a Procrustes analysis (NMDS-based).

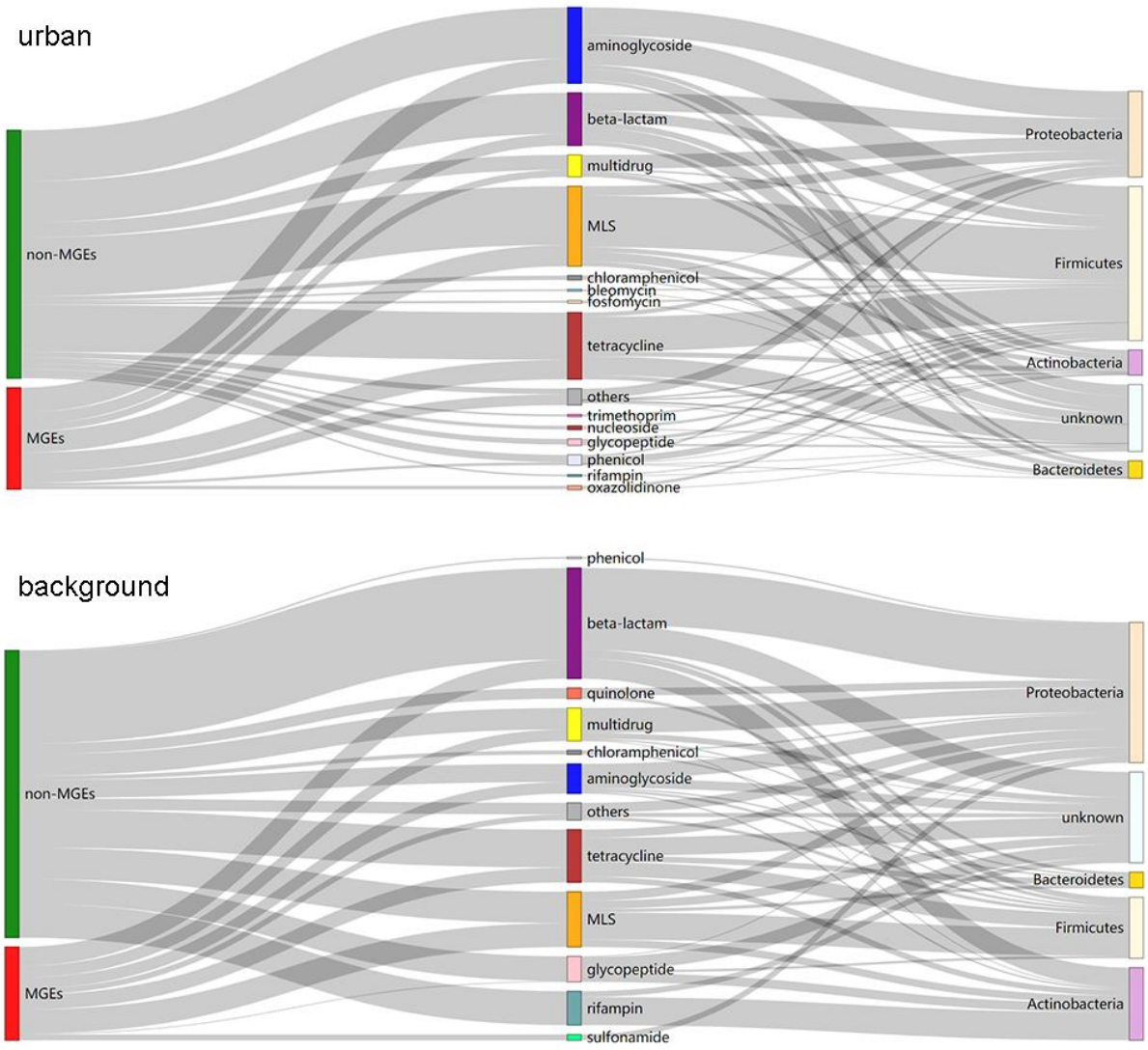


Figure 6-14 Detailed co-occurrence patterns of airborne ARGs and potentially mobile ARGs with their possible hosts in urban and background areas in Sankey diagrams, respectively. The left column shows whether the identified ARGs are co-localized with MGEs on the assembled contigs, and the middle and right columns are ARG classes and their potential hosts, respectively.

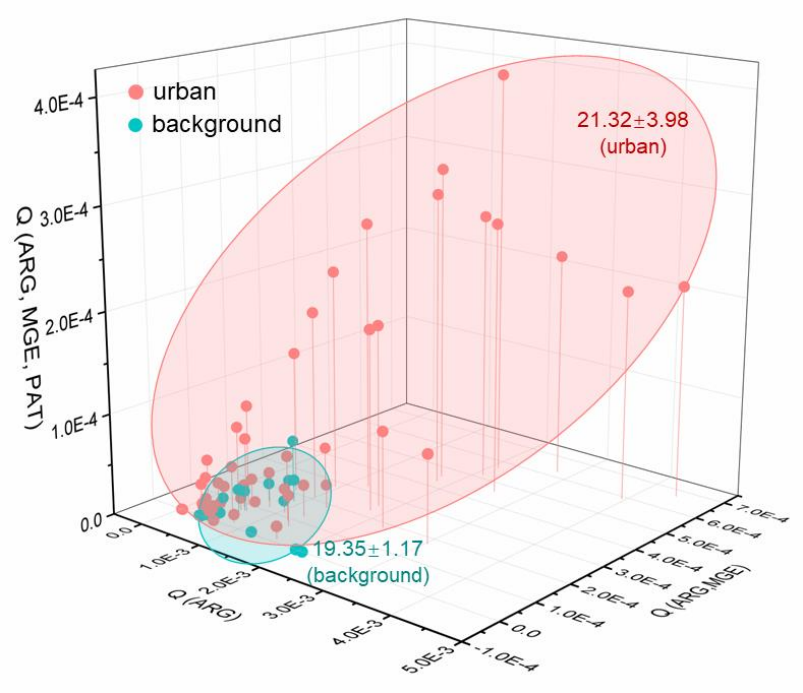


Figure 6-15 The AMR risk scores were estimated by using MetaCompare. The x, y, and z axes represent the portions of contigs concerning the ARGs, MGE-associated ARGs, and pathogen-hosting ARGs to the total assembled contigs. The vertex indicates the theoretically highest AMR risk, and the mean relative-risk scores of urban and background samples were calculated respectively.

6.4 Global AMR exposure hazards and risk ranking in the atmospheric environment

To further clarify the relationship between airborne bacteria and ARGs, 156 and 72 non-redundant MAGs with high quality were generated from urban and background atmospheric environments across the globe (Figure **6-16**), and the relative abundance of MAGs in global air samples varied from 1 to 1×10^5 genome copies per million reads on average and generally exhibited a higher value in TSP samples (Pairwise t-test, $p < 0.05$). This is reasonable, because the greater abundance and

diversity of bacteria in larger particles compared to other particle sizes has been demonstrated in previous studies (Lu *et al.*, 2018; Stern *et al.*, 2021). Accordingly, the abundance of PARB and HVF-PARB also decreased with the atmospheric particle size fraction. Nevertheless, lower abundance cannot be identified with lower hazards, because only those airborne bacteria associated with inhalable particles (*e.g.*, $PM_{2.5}$) can be transmitted from the air to the human respiratory system via inhalation and cause potential health risks.

In urban air, there were 78 MAGs identified as PARB. Among them, Proteobacteria accounted for 83.8% of the total abundance of all PARB metagenomic bins (44703.9 genome copies per million reads), and almost all Proteobacteria PARB bins were also identified as carrying HVFs, which aggravated the AMR risks in global urban air. The MAGs became less diverse in background air, and the proportion of PARB (30.5%) was lower than values in urban areas (50%). Particularly, the total abundance of PARB (urban=44703.9 vs. background=152.3 genome copies per million reads) and HVF-PARB (urban=37464.6 vs. background=98.0 genome copies per million reads) decreased by 2–3 orders of magnitude, emphasizing the importance of anthropogenic impacts to ARG enrichment and relevant risks.

To label the importance of PARB more clearly to AMR risks for human beings, the concentrations of HVF-PARB detected in the global atmosphere were estimated. The mean concentration of HVF-PARB in the urban air (3.35 ± 3.57 genome copies/m^3) was seven times higher than the estimation in background air (0.47 ± 0.56 genome copies/m^3) (Figure **6-17**a). The potential exposure risks of airborne pathogens were significantly higher in urban air than in background air (One-way ANOVA, F =6.4, $p < 0.05$, Figure **6-17**a), which was consistent with the higher density of HVF-PARB in the urban air samples (52.65 ± 39.12 cells/m^3) than background air samples (7.04 ± 5.47 cells/m^3) (Figure **6-17**b). The values corresponded to the total amounts of

789.75 ± 586.80 and 105.6 ± 82.05 cells of HVF-PARB that are transferred daily through human inhalation (15 m^3/day) in urban and background areas, respectively. Notably, roughly 31% of the urban airborne bacteria were contributed by hospitals and WWTPs emerging as hotpots for ARGs (Rice *et al.*, 2020) (Figure **6-9**b), aggravating the potential AR hazards in cities. In summary, the higher daily intake rate of HVF-PARB cells in urban air was taken into consideration to conclude that humans were potentially infected by PARB at heightened risk in cities.

The cell density of each HVF-PARB in the ambient air was quantified in each sampling site and showed regional differentiations (Figure **6-18**), and each HVF-PARB showed a higher abundance in urban air than in background air. In particular, *Ralstonia insidiosa* and *Acinetobacter junii* were detected in all air samples, and their concentrations in urban air were significantly higher than in background air. It is noteworthy that *Rothia dentocariosa, Acinetobacter johnsonii, S. aureus, Rickettsia felis,* and *Staphylococcus epidermidis* only occurred in urban air, implying that various pathogenic infections might increase in cities. Despite the lowest mean concentration in all detected HVF-PARB, *S. aureus* could acquire resistance to most antibiotics in the clinic until now and was defined as a high-priority pathogen to support the research and development of new effective drugs by the WHO (Tacconelli *et al.*, 2018). To construct the genotype-phenotype relationships and environmental-health linkage, we reconstructed a strain-level phylogenic tree based on MAG identified as *S. aureus* in this study and 182 new *S. aureus* isolate genomes, including both methicillin-resistant *Staphylococcus aureus* (MRSA) and methicillin-sensitive *Staphylococcus aureus* (MSSA) from the seven most common sites of nosocomial infections (Manara *et al.*, 2018) by using PhyloPhlAn 3.0 (Asnicar *et al.*, 2020) (Figure **6-19**). The MAG generated from ambient air samples was closely connected to 20 clinical MRSA isolates with distances less than 0.001 in the phylogenetic tree, which further verified the infection risks of

airborne *S. aureus* carrying ARGs. Moreover, *mecC*, one ARG with a surprisingly high prevalence in MRSA strains (García-Álvarez *et al.*, 2011; Bengtsson *et al.*, 2017), was detected in this MAG, heightening the threat of AMR inhalation risks. Also, the MAG generated from urban air showed closer genetic relationships with MRSA isolates retrieved from sputum samples, indicating a higher risk of chronic lung infections caused by airborne antibiotic-resistant *S. aureus.* Assuming that an adult inhales 15 m^3/day air, the estimated inhalation amount of *S. aureus* varied from 0 to 5700 cells per day (Figure **6-18**). This is much lower than the infective dose of *S. aureus*, *i.e.*, exceeds 100,000 cells/g, in food to produce an intoxication amount of enterotoxins (United States Food and Drug Administration) and the previous exposure doses of *S. aureus*, *i.e.*, ranging from 10^5 to 10^7 colony-forming units (CFU), to mice by intranasal instillation (Wang *et al.*, 2016; Ashley *et al.*, 2020). From a clinical point of view, however, the infective dose is also subject to the sensitivity of people, mode of transmission (*e.g.*, direct contact with wounds), and virulence factors (*e.g.*, enterotoxin serotypes). Nevertheless, *S. aureus* can also be carried both chronically and intermittently (Chmielowiec-Korzeniowska *et al.*, 2020), together with continuous respiration, which might increase the respiratory infection risks with long-term exposure to *S. aureus* distributed in urban air. Thus, the low-dose and long-term inhalation exposures of *S. aureus* in mice *in vivo* are suggested in a chamber model to further identify safe concentrations of *S. aureus* in the air.

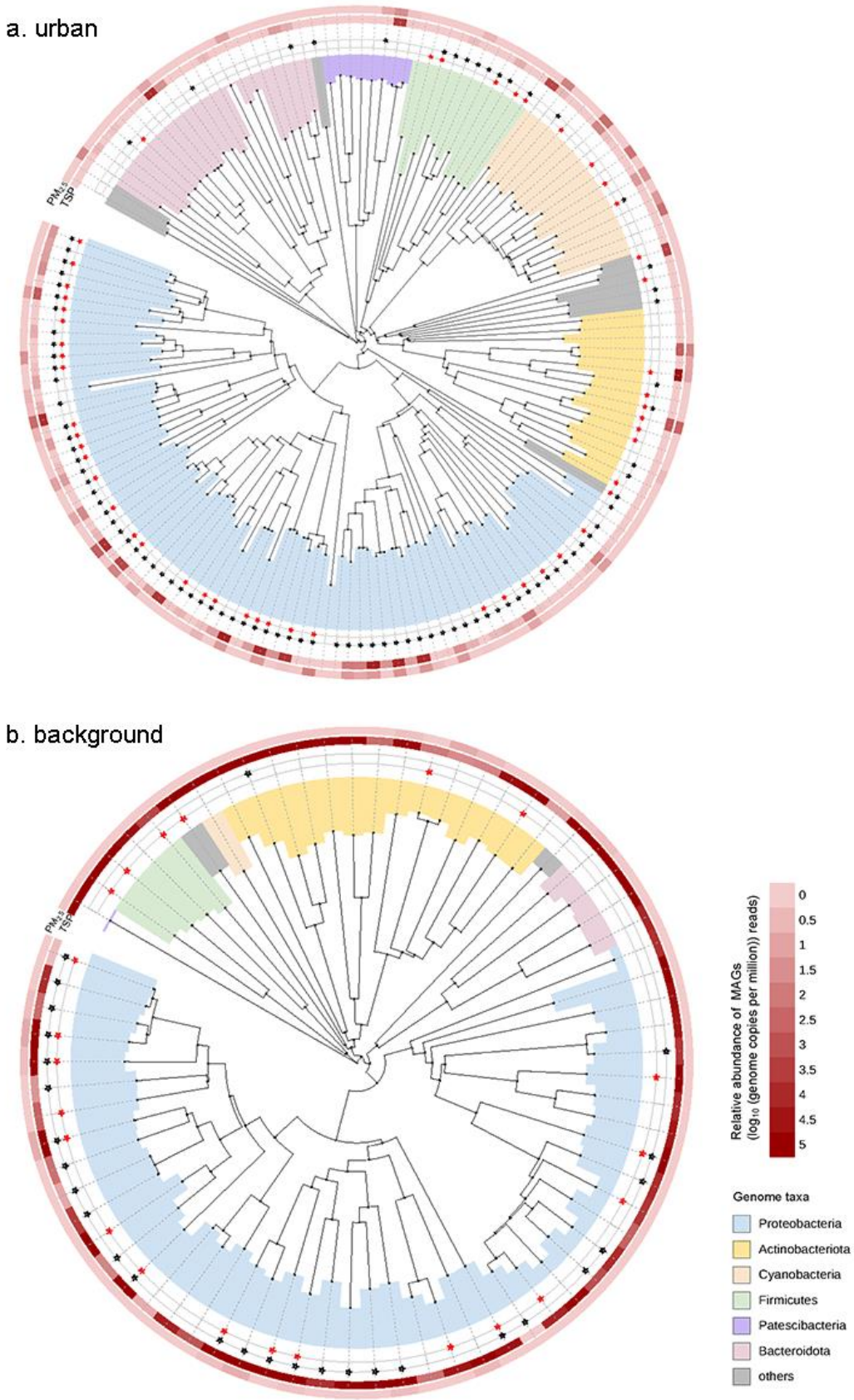

a. urban
b. background
PM2.5
TSP
Relative abundance of MAGs
(log10 (genome copies per million)) reads)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Genome taxa
Proteobacteria
Actinobacteriota
Cyanobacteria
Firmicutes
Patescibacteria
Bacteroidota
others

Figure 6-16 Phylogenetic trees and distribution of the MAGs in urban (a) and background (b) atmosphere worldwide. Each MAG is color-coded at the phyla level. Among them, the bacterial taxa carrying high-risk ARGs are depicted in the shapes of red stars, and the genomes carrying HVF genes are marked using black stars. The darkness of the colors in the heatmaps outside the circle corresponds to the relative abundances of the MAGs (genome copies/per million reads) at the body site of greatest prevalence.

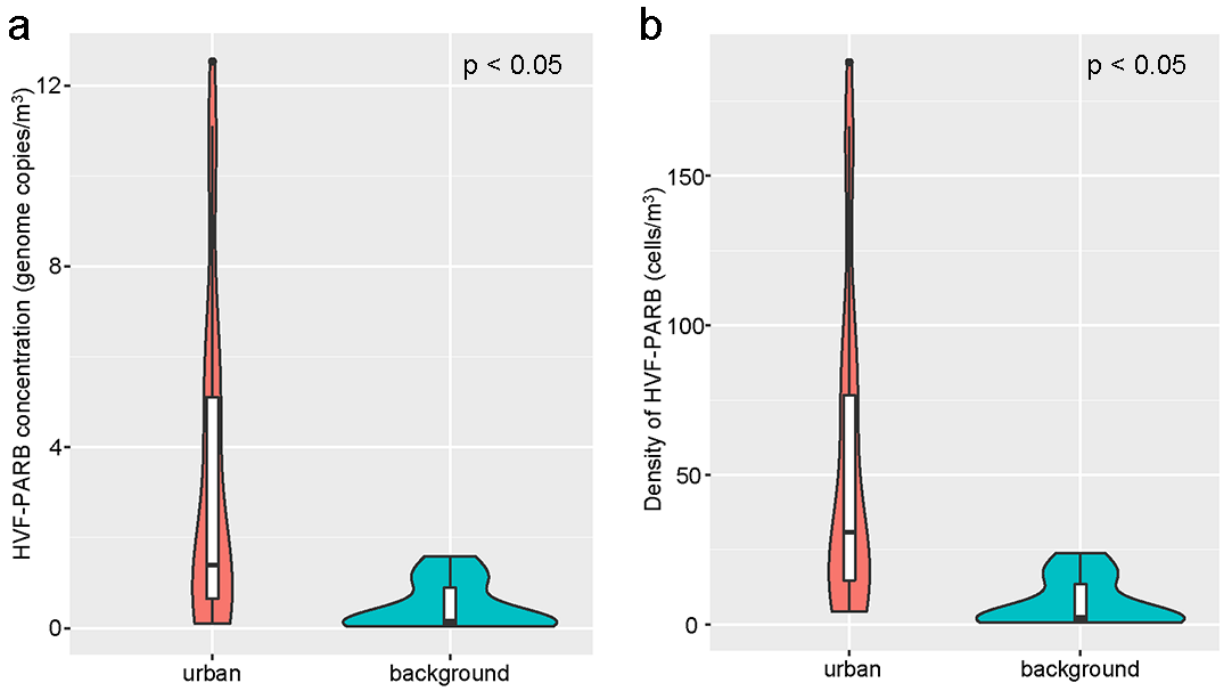


Figure 6-17 (a) Comparison of HVF-PARB concentration (genome copies/m^3) in urban and background air samples. (b) Comparison of cell densities of HVF-PARB (cells/m^3) in the urban and background air samples in this study.

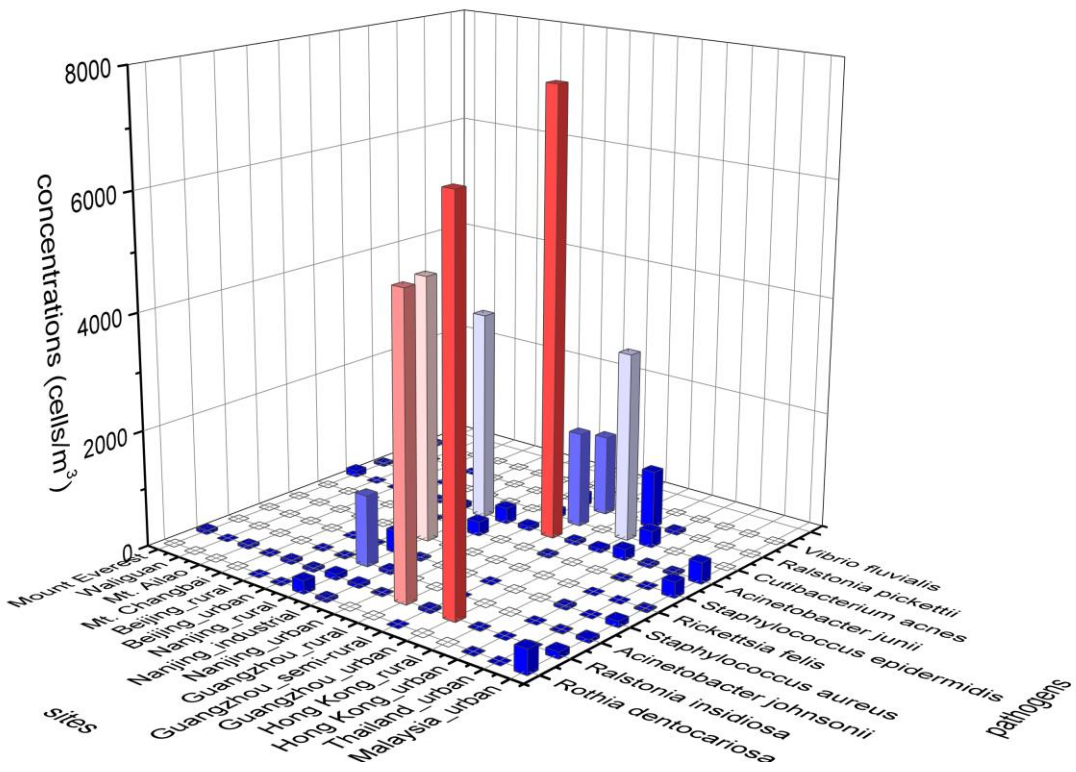


Figure 6-18 Regional comparisons of the concentrations (cells/m^3) of identified HVF-PARB in the MAGs generated from global air samples.

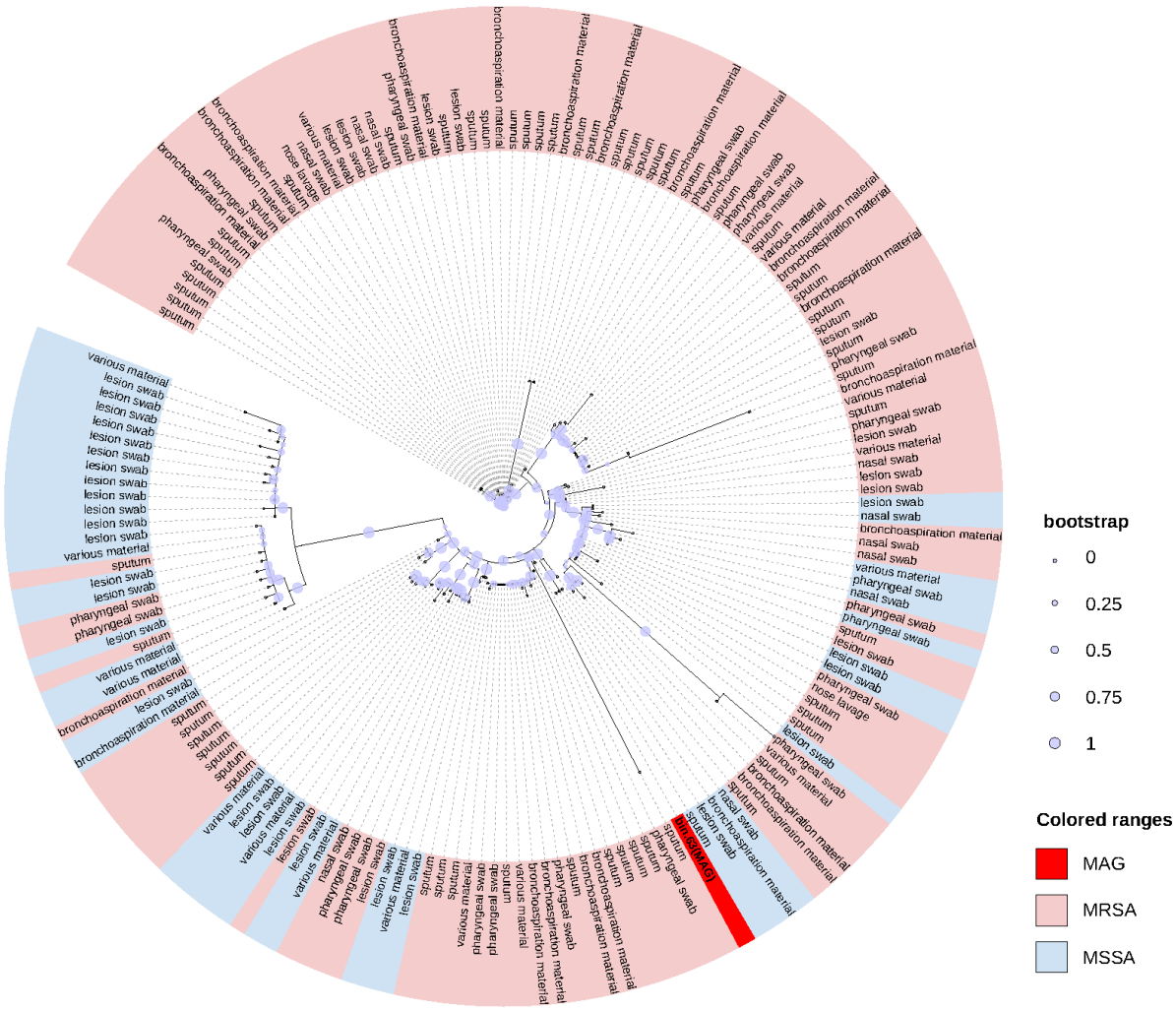


Figure 6-19 Phylogenetic tree of 182 *S. aureus* strains from a pediatric hospital (Manara *et al.*, 2018) and MAG identified as *S. aureus* in this study reconstructed by PhyloPhlAn 3.0 (Asnicar *et al.*, 2020). Each isolate is color-coded by types of resistant antibiotics, including both MRSA and MSSA. The label of each branch represents the seven most common sites of hospital infections for *S. aureus*, including airways (bronchial aspirates, sputum, or oropharyngeal and nasal swabs) or soft-tissue and skin lesions. The size of purple circles internal to the phylogeny represents the bootstrap of subtrees.

6.5 Community-level identification of HT-ARGs

HGT between organisms is thought to show the importance of driving microbial evolution and adaptation, such as the development of antibiotic resistance and virulence (Ochman *et al.*, 2000). In this study, we identified community-level HGT through the combination of best-match and phylogenetic approaches by using MetaChip within two global airborne bacterial communities (156 and 72 genome bins in urban and background air) (Song *et al.*, 2019); therefore, 951 and 463 gene flows were predicted with specific taxonomic information and transfer direction in urban and background air, respectively (Figure **6-20**a and Figure **6-20**b). Not surprisingly, the number of HGT detected in the two communities was proportional to the number of MAGs it contained, showing a similar probability of occurrence. High rates of HGT within the pathogen-associated orders *Rhizobiales, Burkholderiales* (Riera-Ruiz *et al.*, 2014), and *Sphingomonadales* (Glaeser and Kämpfer, 2014) were described in global airborne bacterial communities. Moreover, HGT was likely to occur between two closely related taxa, which was generally consistent with previous findings (Bolotin and Hershberg, 2017). For example, a genome bin of the *Burkholderiales* was also discovered to more frequently share HGT with *Rhizobiales* and *Pseudomonadales*.

To identify and track the directionality of HT-ARGs, we performed the annotation of AMR-related genes identified in the HGT analysis based on DeepARG databases (Das *et al.*, 2022). We found that HGT associated with ARGs occurred more frequently within bacteria associated with human virulent factors, especially in urban air, accounting for 88.5% of recipient bacteria involved in each HGT event, which posed an elevated risk to human health (Figure **6-20**c and Figure **6-20**d). Additionally, the urban airborne bacteria estimated to be engaged in ARG transfer were found to be enriched with *Burkholderiales, Cyanobacteriales,* and *Sphingomonadales,* and showed a greater proportion in the whole community than the counterpart in background air (2.4 and 0.1

genome copies/m^3 in urban and background areas, respectively). As expected, urban air tended to be detected a larger portion of HT-ARGs among MetaCHIP-predicted HGTs (5.5%) than background air (2.3%) (Figure **6-20**e and Figure **6-20**f). We further classified and quantified these HT-ARGs according to their risk ranking and corresponding resistant antibiotics, and the two datasets represented remarkably consistent profiles of HT-ARGs, with dominant classes of multidrug, MLS, and glycopeptides. Still, the highest-risk ARGs, such as *bacA* resistant to bacitracin, were only detected in urban airborne HT-ARGs, revealing an increasing health risk to humans. The comprehensive analysis of HT-ARGs and the identification of related donors and recipients provided substantial value in targeting environmental monitoring efforts towards tracking mobile ARGs and contributed to stopping the spread of antibiotic resistance directly.

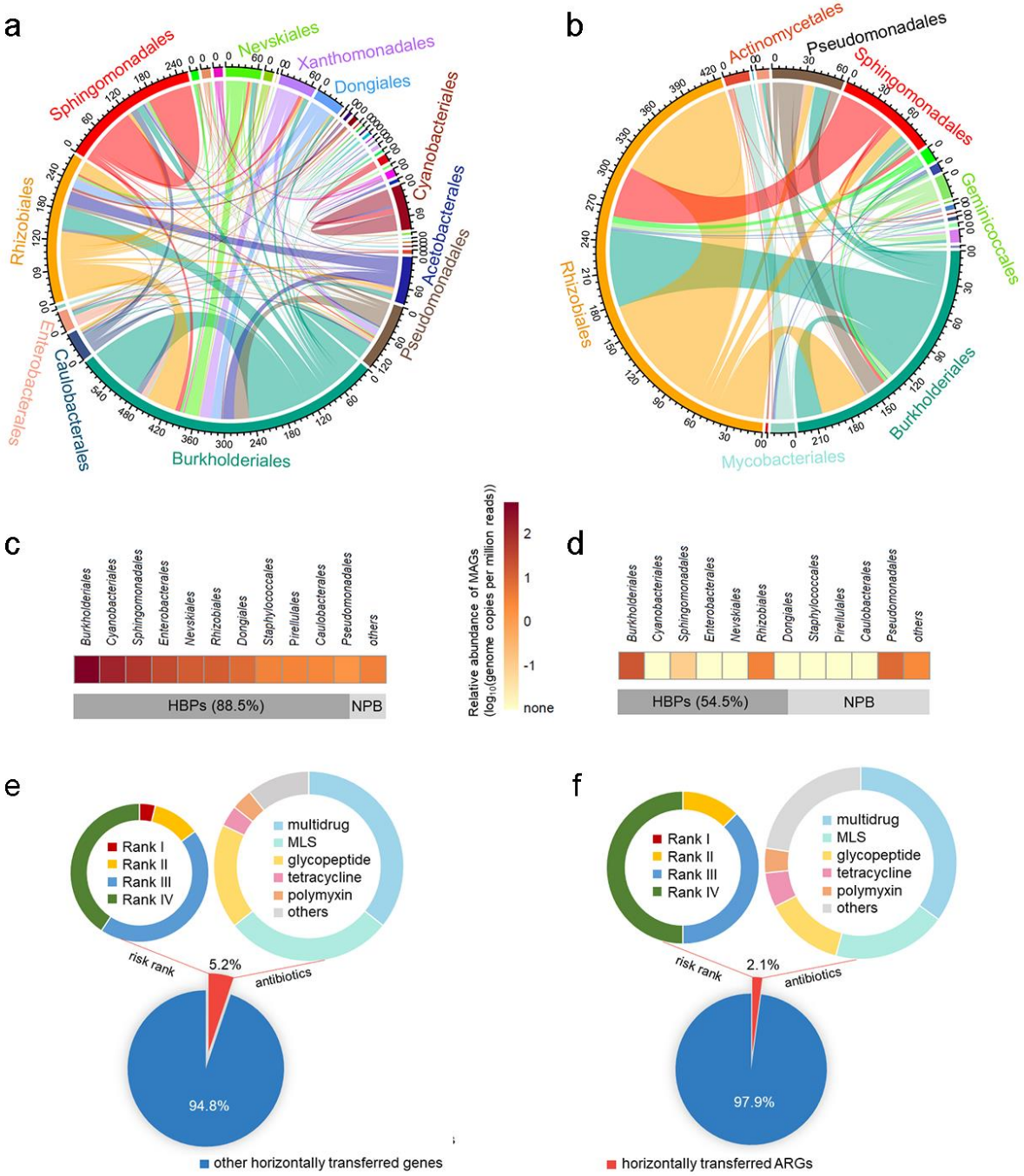


Figure 6-20 Community-level identification and the analysis of HT-ARGs. Predicted gene flow within the urban (a) and background (b) bacterial communities. Bands connect donors and recipients, with the width of the band correlating to the number of HGTs and the color

corresponding to the donors. Heatmaps showing the mean relative abundance of bacterial orders estimated to be engaged in ARG transfer across the samples collected in urban (c) and background air (d). Bacterial recipients involved in HGT were counted by HBPs and NPB, respectively. The percentage of HT-ARGs detected among urban (c) and background (d) HGT events. The HT-ARGs were classified and quantified based on risk ranking and resistant antibiotic types.

6.6 Summary

To the best of our knowledge, this study provides the first comprehensive atlas of global ARG distribution and assessment of AMR risks in atmospheric environments, which would contribute to public health management and improve decision-making at the intersection of clinics and environments. In particular, the risks to public health were evaluated based on the abundance, mobility, and host pathogenicity of ARGs from a regional perspective. Also, 88.4% and 68.1% of airborne ARGs were contributed by other ecosystems in global urban and background areas, respectively, illustrating that the airborne ARG distributions were mainly impacted by source contributions. Furthermore, a genome-resolved "panorama" of AMR was described, in which the bacterial density and daily inhalation of each HVF-PARB were quantified. Notably, the *S. aureus* MAG carrying *mecC* was detected in urban air samples and showed close genetic relationships with MRSA isolated from the clinic. Additionally, HGT associated with ARGs between organisms was further identified to benefit from targeting atmospheric monitoring efforts towards tracking mobile ARGs. The major results of this chapter are provided below.

a) In this chapter, 747 ARGs (28 drug classes of antibiotics) were detected in the global ambient air; most of them (*n=565*) conferred multiple drug resistances. A core ARG set (*n=39*) was determined based on the abundance and occupancy in the global dataset and accounted for 61.2±11.8% of the sequences in all air samples. Moreover, the similar ARG profiles on drug resistance types, together with the significant overlap of (core) ARGs in the global atmospheric, aquatic, and terrestrial ecosystems, illustrated the high communicability and universality of ARGs in natural environments.

b) The different profiles and abundance of ARGs and pathogens between urban and background air indicated the importance of anthropogenic effects on airborne AMR risks. The abundance of urban indicator ARGs was largely affected by relevant antibiotic consumptions, particularly aminoglycoside, tetracycline, and beta-lactam, raising the importance of the appropriate use of these drugs.

c) The major contributors of ARGs in urban air were common ARG hotpots, like WWTPs (16.4%), hospitals (14.6%), and landfills (7.2%), which partially explained the higher risk rank of urban airborne ARGs than the counterparts in background air. This finding highlights the dissemination of ARGs and their bacterial hosts from typical ARG hotpots to the ambient atmosphere, which might result in potential health implications through inhalation.

d) The bacterial density of HVF-PARB in the urban air (52.65±39.12 cells/m^3) was around 7 times higher than the counterpart in background air (7.04±5.47 cells/m^3); thus, humans living in cities were potentially infected by PARB at heightened risk. Notably, the *mecC*-hosting *S. aureus* genome closely associated with strain isolates in hospital infections was frequently detected in urban air, providing a warning that AMR threats would be higher than expected.

e) The urban airborne bacteria estimated to be engaged in ARG transfer were found to be enriched with *Burkholderiales, Cyanobacteriales,* and *Sphingomonadales* and showed a greater abundance in the entire community (2.4 genome copies/m^3) than the counterpart in background air (0.1 genome copies/m^3).

This chapter conducted a global metagenomic dataset (~1.33 TB) based on 262 air samples worldwide to reveal the distribution and dissemination of 747 airborne ARGs as well as their potential hosts. The results showed that anthropogenic impacted air was being polluted by ARGs, particularly those emitted from surrounding hospitals and WWTPs, which are known as AMR urban hotspots. Moreover, different areas were challenged with varying infected health risks related to airborne ARGs or PARB via inhalation, indicating the targeted control countermeasures to specific ARGs and pathogens from urban sources, environmental dissemination, to human inhalation risks. In light of the global disparities in airborne AMR risks and the associated AMR burden, this study highlighted the global significance of the airborne transmission of ARGs and the urgent need to involve biological parameters such as ARGs or ARB in current air quality standards on public health.

Chapter 7 Conclusions and Recommendations

7.1 Overall summary and major conclusions

As the first attempt of systematic study concerning bacterial communities, ARG profiles, and ARG hosting bacteria and pathogens, in the global atmosphere, this thesis specifically investigated (1) the biogeographic patterns of global airborne bacterial communities and their interconnections with Earth's microbiomes; (2) the mechanisms driving the global patterns of airborne bacterial structure and distribution via environmental and anthropogenic perspectives; and (3) the comprehensive atlas of AMR in global ambient air and the related risks to human health. The main findings of the thesis are as follows:

a) By encompassing global bacterial communities of the three largest ecosystem types (*i.e.*, atmospheric, oceanic, and terrestrial systems), the uniform latitudinal bacterial diversity pattern was discovered to show the maximin microbial richness in the mid-latitude regions of the world. Besides, this study also revealed that the atmosphere harbors a unique core bacterial community, but a less stable structure compared to terrestrial and marine ecosystems.

b) The global airborne bacterial communities showed as complex and dynamic structure as bacterial assemblages in the other two largest ecosystems, terrestrial and aquatic environments, and the importance of airborne bacteria in the Earth's microbiomes was ascertained according to the close interrelationships of bacteria in various habitats and nearly half of airborne bacterial contributions from other ecosystems.

c) Even though air is a highly flowing ecosystem enabling long transport and dynamic procedures over geographic barriers, the structure of airborne bacterial communities appeared

to be driven more by local environments, especially the potential source contribution and air quality dominated by human activities, which is in contrast to conventional wisdom. The anthropogenic impacts on airborne bacteria were mainly reflected in fewer biomass loadings, higher pathogenic abundance, a less stabilized network structure, and bacterial phenotype differentia, with the mechanisms of reduced environmental filtering effects and elevated human-related source contributions.

d) This thesis conducted a systematic and targeted surveillance of quantified antibiotic resistance risks across the global ambient air for both spatial and seasonal patterns, revealing that urban air is being polluted by ARGs, particularly those emitted from surrounding hospitals and WWTPs, which are known as urban AMR sources. Furthermore, each area (country) inevitably faced challenges with varying threats related to airborne ARGs or PARB via inhalation, contributing to the public policy of targeted control to specific ARGs and pathogens frequently occurring in specific regions.

e) This study emphasized the anthropogenic impacts on AMR and revealed the more closely related infection risks in urban air based on the higher mobility of ARGs and more inhaled HVF-PARB (averagely 53 genome copies) daily in cities. Besides, *S. aureus* genomes associated with *mecC* identified in urban air showed closer genetic relationships with strains isolated from hospital infections, further verifying the inhalation threats of AMR to public health.

7.2 Limitations of the current study and future perspectives

Although the structures and distributions of the airborne bacterial community and ARG profiles were well documented on a global scale and extended to the ecological mechanisms shaping biogeographic patterns and risks of AMR to human health, which could be more integrated than the previous regional studies that mainly explored the dynamics of airborne bacteria and ARGs, the current results are restricted by certain aspects like sample amount, sampling period, models and experimental techniques applied, experimental settings, etc. The limitations of the current study and the relevant suggestions for future perspective are briefly discussed below:

a) This study mainly focused on the ecological mechanisms shaping biogeographic patterns, however, the evolutionary mechanisms were not mentioned due to the short span of sampling time, which was around one year. Evolutionary processes also drive microbial biogeography and even take place on ecological scales due to changing environments. For instance, microbial responses are classified into two groups: non-evolutionary changes (*e.g.*, phenotypic plasticity, shifting metabolic states, and microbial interactions) and evolutionary changes (*i.e.*, adaptation), and adaptation can be regarded as the dynamic evolutionary process, which fits microbes to their living environments (Bradford *et al.*, 2008). When climate changes or other external disturbances cause the departure of microbes from their beneficial conditions or tolerance zones, microbes may develop new characteristics in terms of cell structure, physiology, or behaviors to shift their survival limitation on an evolutionary timescale. The adaptation occurs on an evolutionary scale and emphasizes functional changes over multiple generations(Gunde-Cimerman *et al.*, 2018). As a result, for bridging micro-scale mechanisms and macro-scale biogeographic patterns of microbial abundance, diversity, and functions, systematic sampling work with a longer period (around 10 years) should be conducted, which will benefit a comprehensive understanding of mechanisms shaping various macroecological

patterns and bacterial community structures in the future.

b) The composition and abundance of potential pathogens have been analyzed and quantified based on two mainstream methods: 16S rRNA gene sequencing data via 16SPIP and metagenomic sequencing data according to the pathogen list (species level) summarized by previous studies, which revealed similar results. However, more accurate quantitative diagnostic methods are still required in future studies. Besides, it is well-known that conventional culture-based methods could only provide an incomplete and biased view of the biodiversity of the microbiomes, as the majority of the pathogenic species are not able to be cultivated. To overcome the above disadvantages, culture-independent molecular methods, such as pathogen-specific real-time PCR analysis, are recommended as a sensitive and precise technique for an extensive quantitative evaluation of pathogens in ambient air for a future perspective.

c) This study constructed a co-occurrence network of global airborne bacterial community; however, it was still unclear how accurately the constructed network represented the real-world systems. To quantitatively disentangle direct and indirect relationships, iDIRECT is suggested to be applied in network analysis for the effective removal of spurious links and overcoming of overestimated direct association strength caused by indirect influences (Xiao *et al.*, 2022). Additionally, association networks are also recommended to be performed to assess the relative importance of biotic interactions, abiotic filtering, and dispersal in shaping airborne bacterial community diversity and dynamics (Yuan *et al.*, 2021).

d) This study only focused on the community and function prediction level and AMR-related risks; for future perspectives, we suggested probing the metabolically active bacterial

community with respect to their function in the atmosphere (*e.g.*, ice-nucleating, pollutant transformation, etc.) and impacts on human health from multiple aspects (*e.g.*, benefits arise from healthful bacteria communities/secondary metabolites, and/or hazardous effects such as allergens, etc.), while expanding the scope of airborne microbiome (*e.g.*, fungi, viruses, etc.) to enrich the field of bioaerosols and atmospheric science in relation to human health as well as contributing to the moving frontier of Earth's microbiome.

e) The evaluation of ARG mobility and dissemination should also consider the extracellular fraction of airborne ARGs (He *et al.*, 2021) as well as the phage-mediated process (Sun *et al.*, 2022). As increasing evidence suggests that phages are a potential reservoir of ARG and serve as vectors in diverse environments, it is necessary to assess the magnitude of their relative HGT contributions and further prioritize associated mitigation and risk management efforts.

f) Although PARB has been identified based on both assembled contigs and MAGs, this is only a projection. As a result, culture-based methods are required to confirm the existence and viability of PARB identified in this study and the expression of antibiotic resistance, as well as to verify the existing transmission of ARGs via HGT in future studies. The results of culture-dependent experiments will contribute to providing solid evidence beyond inferences derived from metagenomic data and a better understanding of ARG linkage to specific host species.

g) Based on estimations of the intake rate of AMR material, future risk assessments of AMR threats should consider the disparities in relation to human immunological responses to multiple AMR exposure pathways (respiration vs. digestion systems), particularly from the "One Health" perspective. Furthermore, the observation of respiratory infections needs to be

verified by connecting airborne AMR materials and human lung epithelial cells.

List of Appendix

Appendix 1

This section provides the basic sampling information, sequence number, and bacterial richness of air samples used to organize the global airborne bacterial dataset in this study.

No.	Sample ID	Location	Lat	Lng	ASL	Time	Temp.	Primer	Type	Ref.	Rich.	Seq.
1	CB13aut	Mt. Changbai, China	42.4	128.1	741	2013_aut	5.16	341F/806R	$PM_{2.5}$	This study	409	163641
2	CB13win	Mt. Changbai, China	42.4	128.1	741	2013_win	-12.44	341F/806R	$PM_{2.5}$	This study	1397	420614
3	CB14spr	Mt. Changbai, China	42.4	128.1	741	2014_spr	5.78	341F/806R	$PM_{2.5}$	This study	542	171567
4	CB14sum	Mt. Changbai, China	42.4	128.1	741	2014_sum	18.94	341F/806R	$PM_{2.5}$	This study	322	63488
5	HT13win	Hong Kong, China	22.12	114.15	50	2013_win	16.94	341F/806R	$PM_{2.5}$	This study	412	190718
6	HT14spr	Hong Kong, China	22.12	114.15	50	2014_spr	19.87	341F/806R	$PM_{2.5}$	This study	308	99493
7	HT14sum	Hong Kong, China	22.12	114.15	50	2014_sum	28.47	341F/806R	$PM_{2.5}$	This study	431	191164
8	HT14aut	Hong Kong, China	22.12	114.15	50	2014_aut	28.18	341F/806R	$PM_{2.5}$	This study	309	89063
9	AL14sum	Mt. Ailao, China	24.53	101.02	2450	2014_sum	21.8	341F/806R	$PM_{2.5}$	This study	492	213867
10	AL14aut	Mt. Ailao, China	24.53	101.02	2450	2014_aut	16.41	341F/806R	$PM_{2.5}$	This study	595	303164
11	AL14win	Mt. Ailao, China	24.53	101.02	2450	2014_win	11.43	341F/806R	$PM_{2.5}$	This study	683	232175
12	AL15spr	Mt. Ailao, China	24.53	101.02	2450	2015_spr	18.68	341F/806R	$PM_{2.5}$	This study	757	161985
13	WL14aut	Waliguan, China	36.29	100.9	3816	2014_aut	11.3	341F/806R	$PM_{2.5}$	This study	2242	433152
14	WL14win	Waliguan, China	36.29	100.9	3816	2014_win	-4.16	341F/806R	$PM_{2.5}$	This study	2487	427621
15	WL15spr	Waliguan, China	36.29	100.9	3816	2015_spr	7.12	341F/806R	$PM_{2.5}$	This study	2513	404701
16	WL15sum	Waliguan, China	36.29	100.9	3816	2015_sum	15.41	341F/806R	$PM_{2.5}$	This study	2260	285751
17	ML15win	Bachok, Malaysia	6.01	102.43	10	2015_win	27.49	341F/807R	TSP	This study	366	41619
18	ML16spr	Bachok, Malaysia	6.01	102.43	10	2016_spr	29.21	341F/808R	TSP	This study	354	48699
19	ML16sum	Bachok, Malaysia	6.01	102.43	10	2016_sum	28.21	341F/809R	TSP	This study	226	50541
20	ML16aut	Bachok, Malaysia	6.01	102.43	10	2016_aut	27.65	341F/810R	TSP	This study	690	49536
21	TL15win	Phitsanulok, Thailand	16.75	100.2	47	2015_win	26.2	341F/811R	TSP	This study	562	44645
22	TL16spr	Phitsanulok, Thailand	16.75	100.2	47	2016_spr	32.33	341F/812R	TSP	This study	493	44285
23	TL16sum	Phitsanulok, Thailand	16.75	100.2	47	2016_sum	29.06	341F/813R	TSP	This study	521	48017
24	TL16aut	Phitsanulok, Thailand	16.75	100.2	47	2016_aut	27.88	341F/814R	TSP	This study	305	50152
25	TH16spr	Guangzhou, China	23.15	113.36	36	2016_spr	21.88	341F/806R	$PM_{2.5}$	This study	1007	17480
26	TH16sum	Guangzhou, China	23.15	113.36	36	2016_sum	22.29	341F/806R	$PM_{2.5}$	This study	1130	11805

27	TH16aut	Guangzhou, China	23.15	113.36	36	2016_aut	20.26	341F/806R	$PM_{2.5}$	This study	1281	14151
28	TH16win	Guangzhou, China	23.15	113.36	36	2016_win	15.54	341F/806R	$PM_{2.5}$	This study	1370	19615
29	TH17spr	Guangzhou, China	23.15	113.36	36	2017_spr	21.6	341F/806R	$PM_{2.5}$	This study	1291	15310
30	CH16spr	Guangzhou, China	23.65	113.62	21	2016_spr	21.17	341F/806R	$PM_{2.5}$	This study	1079	17764
31	CH16sum	Guangzhou, China	23.65	113.62	21	2016_sum	27.23	341F/806R	$PM_{2.5}$	This study	1300	17960
32	CH16aut	Guangzhou, China	23.65	113.62	21	2016_aut	22.65	341F/806R	$PM_{2.5}$	This study	1005	19623
33	CH16win	Guangzhou, China	23.65	113.62	21	2016_win	15.24	341F/806R	$PM_{2.5}$	This study	851	20994
34	CH17spr	Guangzhou, China	23.65	113.62	21	2017_spr	20.71	341F/806R	$PM_{2.5}$	This study	788	24043
35	HS16spr	Guangzhou, China	22.71	112.92	33	2016_spr	22.06	341F/806R	$PM_{2.5}$	This study	817	22347
36	HS16sum	Guangzhou, China	22.71	112.92	33	2016_sum	28.57	341F/806R	$PM_{2.5}$	This study	690	6292
37	HS16aut	Guangzhou, China	22.71	112.92	33	2016_aut	24.59	341F/806R	$PM_{2.5}$	This study	1046	12331
38	HS16win	Guangzhou, China	22.71	112.92	33	2016_win	16.55	341F/806R	$PM_{2.5}$	This study	1080	22590
39	HS17spr	Guangzhou, China	22.71	112.92	33	2017_spr	21.19	341F/806R	$PM_{2.5}$	This study	711	16773
40	LS16spr	Nanjing, China	31.33	119.01	15	2016_spr	20.3	341F/806R	$PM_{2.5}$	This study	446	43618
41	LS16sum	Nanjing, China	31.33	119.01	15	2016_sum	26.92	341F/806R	$PM_{2.5}$	This study	1009	45075
42	LS16aut	Nanjing, China	31.33	119.01	15	2016_aut	12.16	341F/806R	$PM_{2.5}$	This study	1268	59441
43	LS16win	Nanjing, China	31.33	119.01	15	2016_win	7.17	341F/806R	$PM_{2.5}$	This study	1302	51711
44	LS17spr	Nanjing, China	31.33	119.01	15	2017_spr	21.09	341F/806R	$PM_{2.5}$	This study	844	44464
45	XW16spr	Nanjing, China	32.06	118.8	30	2016_spr	20.54	341F/806R	$PM_{2.5}$	This study	760	51098
46	XW16sum	Nanjing, China	32.06	118.8	30	2016_sum	27.36	341F/806R	$PM_{2.5}$	This study	787	47002
47	XW16aut	Nanjing, China	32.06	118.8	30	2016_aut	12.31	341F/806R	$PM_{2.5}$	This study	840	56961
48	XW16win	Nanjing, China	32.06	118.8	30	2016_win	7.61	341F/806R	$PM_{2.5}$	This study	875	44859
49	XW17spr	Nanjing, China	32.06	118.8	30	2017_spr	21.69	341F/806R	$PM_{2.5}$	This study	912	42833
50	PK16spr	Nanjing, China	32.12	118.42	51	2016_spr	16.25	341F/806R	$PM_{2.5}$	This study	1147	52325
51	PK16sum	Nanjing, China	32.12	118.42	51	2016_sum	27.05	341F/806R	$PM_{2.5}$	This study	962	69117
52	PK16aut	Nanjing, China	32.12	118.42	51	2016_aut	17.76	341F/806R	$PM_{2.5}$	This study	973	62165
53	PK16win	Nanjing, China	32.12	118.42	51	2016_win	6.43	341F/806R	$PM_{2.5}$	This study	1126	60214
54	PK17spr	Nanjing, China	32.12	118.42	51	2017_spr	16.76	341F/806R	$PM_{2.5}$	This study	989	40831
55	IAP16spr	Beijing, China	39.98	116.38	56	2016_spr	18.31	341F/806R	$PM_{2.5}$	This study	2097	28792
56	IAP16sum	Beijing, China	39.98	116.38	56	2016_sum	25.85	341F/806R	$PM_{2.5}$	This study	1134	40471
57	IAP16aut	Beijing, China	39.98	116.38	56	2016_aut	12.38	341F/806R	$PM_{2.5}$	This study	1304	40121
58	IAP16win	Beijing, China	39.98	116.38	56	2016_win	0.4	341F/806R	$PM_{2.5}$	This study	704	40369
59	IAP17spr	Beijing, China	39.98	116.38	56	2017_spr	7.77	341F/806R	$PM_{2.5}$	This study	1753	17545

60	PKU16win	Beijing, China	39.99	116.36	56	2016_win	0.4	341F/806R	$PM_{2.5}$	This study	1575	38329
61	PKU17spr	Beijing, China	39.99	116.36	56	2017_spr	3.23	341F/806R	$PM_{2.5}$	This study	2160	34987
62	PKU17sum	Beijing, China	39.99	116.36	56	2017_sum	23	341F/806R	$PM_{2.5}$	This study	1502	39208
63	PKU17aut	Beijing, China	39.99	116.36	56	2017_aut	25.74	341F/806R	$PM_{2.5}$	This study	1128	40123
64	PKU17win	Beijing, China	39.99	116.36	56	2017_win	0.45	341F/806R	$PM_{2.5}$	This study	2541	39560
65	HK17win	Hong Kong, China	22.31	114.18	87	2017_win	18.55	341F/806R	$PM_{2.5}$	This study	522	276284
66	HK18spr	Hong Kong, China	22.31	114.18	87	2018_spr	23.67	341F/806R	$PM_{2.5}$	This study	420	297849
67	HK18sum	Hong Kong, China	22.31	114.18	87	2018_sum	29.6	341F/806R	$PM_{2.5}$	This study	423	204011
68	HK18aut	Hong Kong, China	22.31	114.18	87	2018_aut	26.76	341F/806R	$PM_{2.5}$	This study	191	77145
69	ME17win	Mt. Everest, China	28.21	86.56	4276	2017_win	-2.42	341F/806R	$PM_{2.5}$	This study	1880	378773
70	ME18spr	Mt. Everest, China	28.21	86.56	4276	2018_spr	5.1	341F/806R	$PM_{2.5}$	This study	1001	310189
71	ME18sum	Mt. Everest, China	28.21	86.56	4276	2018_sum	10.78	341F/806R	$PM_{2.5}$	This study	259	204085
72	ME18aut	Mt. Everest, China	28.21	86.56	4276	2018_aut	7.13	341F/806R	$PM_{2.5}$	This study	612	408329
73	HR18sum	Beijing, China	40.4	116.69	90	2018_sum	25.67	341F/806R	$PM_{2.5}$	This study	492	414332
74	HR18aut	Beijing, China	40.4	116.69	90	2018_aut	9.54	341F/806R	$PM_{2.5}$	This study	1383	444116
75	HR18win	Beijing, China	40.4	116.69	90	2018_win	-2.3	341F/806R	$PM_{2.5}$	This study	2082	458648
76	HR19spr	Beijing, China	40.4	116.69	90	2019_spr	11.67	341F/806R	$PM_{2.5}$	This study	404	418055
77	AMS S1	Amsterdam-Island, France	-37.81	77.55	59	10/9/2016	20.33	338F/518R	PM_{10}	Romie et al. 2019	628	35253
78	AMS S2	Amsterdam-Island, France	-37.81	77.55	59	17/9/2016	21.9	338F/518R	PM_{10}	Romie et al. 2019	876	28328
79	AMS S3	Amsterdam-Island, France	-37.81	77.55	59	24/9/2016	20.46	338F/518R	PM_{10}	Romie et al. 2019	917	47777
80	AMS S4	Amsterdam-Island, France	-37.81	77.55	59	1/10/2016	17	338F/518R	PM_{10}	Romie et al. 2019	703	23164
81	AMS S5	Amsterdam-Island, France	-37.81	77.55	59	8/10/2016	15.49	338F/518R	PM_{10}	Romie et al. 2019	660	46143
82	AMS S6	Amsterdam-Island, France	-37.81	77.55	59	14/10/2016	13.81	338F/518R	PM_{10}	Romie et al. 2019	697	31435
83	AMS S7	Amsterdam-Island, France	-37.81	77.55	59	21/10/2016	9.09	338F/518R	PM_{10}	Romie et al. 2019	280	48994
84	AMS S8	Amsterdam-Island, France	-37.81	77.55	59	28/10/2016	6.92	338F/518R	PM_{10}	Romie et al. 2019	676	43961
85	AMS S9	Amsterdam-Island, France	-37.81	77.55	59	4/11/2016	6.11	338F/518R	PM_{10}	Romie et al. 2019	591	15877
86	CAP S1	Cape Point, South Africa	-34.36	18.5	230	14/10/2016	14.6	338F/518R	PM_{10}	Romie et al. 2019	258	18629
87	CAP S2	Cape Point, South Africa	-34.36	18.5	230	21/10/2016	14.25	338F/518R	PM_{10}	Romie et al. 2019	596	72367
88	CAP S3	Cape Point, South Africa	-34.36	18.5	230	28/10/2016	16.56	338F/518R	PM_{10}	Romie et al. 2019	577	69838
89	CAP S4	Cape Point, South Africa	-34.36	18.5	230	4/11/2016	15.61	338F/518R	PM_{10}	Romie et al. 2019	647	85746
90	CAP S5	Cape Point, South Africa	-34.36	18.5	230	11/11/2016	18.64	338F/518R	PM_{10}	Romie et al. 2019	598	63292
91	CAP S6	Cape Point, South Africa	-34.36	18.5	230	18/11/2016	16.61	338F/518R	PM_{10}	Romie et al. 2019	575	64081
92	CAP S7	Cape Point, South Africa	-34.36	18.5	230	25/11/2016	18.99	338F/518R	PM_{10}	Romie et al. 2019	695	71044

93	CHC S1	Chacaltaya, Bolivia	-16.35	-68.13	5380	1/7/2016	9.56	338F/518R	PM_{10}	Romie et al. 2019	754	53454
94	CHC S10	Chacaltaya, Bolivia	-16.35	-68.13	5380	9/9/2017	11.96	338F/518R	PM_{10}	Romie et al. 2019	811	112160
95	CHC S11	Chacaltaya, Bolivia	-16.35	-68.13	5380	16/9/2016	12.6	338F/518R	PM_{10}	Romie et al. 2019	781	4567
96	CHC S12	Chacaltaya, Bolivia	-16.35	-68.13	5380	23/9/2016	15.09	338F/518R	PM_{10}	Romie et al. 2019	519	11374
97	CHC S13	Chacaltaya, Bolivia	-16.35	-68.13	5380	30/9/2016	15.2	338F/518R	PM_{10}	Romie et al. 2019	870	112541
98	CHC S14	Chacaltaya, Bolivia	-16.35	-68.13	5380	14/10/2016	14.6	338F/518R	PM_{10}	Romie et al. 2019	400	21405
99	CHC S15	Chacaltaya, Bolivia	-16.35	-68.13	5380	21/10/2016	14.25	338F/518R	PM_{10}	Romie et al. 2019	702	52131
100	CHC S16	Chacaltaya, Bolivia	-16.35	-68.13	5380	28/10/2016	16.56	338F/518R	PM_{10}	Romie et al. 2019	834	8336
101	CHC S2	Chacaltaya, Bolivia	-16.35	-68.13	5380	8/7/2016	19.82	338F/518R	PM_{10}	Romie et al. 2019	883	109139
102	CHC S3	Chacaltaya, Bolivia	-16.35	-68.13	5380	15/7/2016	12.51	338F/518R	PM_{10}	Romie et al. 2019	774	53274
103	CHC S4	Chacaltaya, Bolivia	-16.35	-68.13	5380	29/7/2016	12.67	338F/518R	PM_{10}	Romie et al. 2019	546	14004
104	CHC S5	Chacaltaya, Bolivia	-16.35	-68.13	5380	5/8/2016	12.35	338F/518R	PM_{10}	Romie et al. 2019	698	51261
105	CHC S6	Chacaltaya, Bolivia	-16.35	-68.13	5380	12/8/2016	11.29	338F/518R	PM_{10}	Romie et al. 2019	681	95248
106	CHC S7	Chacaltaya, Bolivia	-16.35	-68.13	5380	19/8/2016	12.65	338F/518R	PM_{10}	Romie et al. 2019	677	53277
107	CHC S8	Chacaltaya, Bolivia	-16.35	-68.13	5380	26/08/2016	9.14	338F/518R	PM_{10}	Romie et al. 2019	905	68796
108	CHC S9	Chacaltaya, Bolivia	-16.35	-68.13	5380	2/9/2016	12.9	338F/518R	PM_{10}	Romie et al. 2019	754	28162
109	GRE S16	Grenoble, France	45.19	5.76	210	3/7/2017	23.32	338F/518R	PM_{10}	Romie et al. 2019	452	16179
110	GRE S17	Grenoble, France	45.19	5.76	210	10/7/2017	22.45	338F/518R	PM_{10}	Romie et al. 2019	377	21143
111	GRE S18	Grenoble, France	45.19	5.76	210	17/7/2017	21.56	338F/518R	PM_{10}	Romie et al. 2019	563	19996
112	GRE S19	Grenoble, France	45.19	5.76	210	24/7/2017	20.23	338F/518R	PM_{10}	Romie et al. 2019	803	11923
113	GRE S20	Grenoble, France	45.19	5.76	210	31/07/2017	24.73	338F/518R	PM_{10}	Romie et al. 2019	839	23281
114	GRE S21	Grenoble, France	45.19	5.76	210	7/8/2017	17.91	338F/518R	PM_{10}	Romie et al. 2019	725	13277
115	GRE S22	Grenoble, France	45.19	5.76	210	14/8/2017	21.56	338F/518R	PM_{10}	Romie et al. 2019	644	25325
116	GRE S23	Grenoble, France	45.19	5.76	210	21/8/2017	21.72	338F/518R	PM_{10}	Romie et al. 2019	770	15392
117	GRE S24	Grenoble, France	45.19	5.76	210	28/8/2017	19.32	338F/518R	PM_{10}	Romie et al. 2019	839	22239
118	GRE S25	Grenoble, France	45.19	5.76	210	4/9/2017	17.09	338F/518R	PM_{10}	Romie et al. 2019	662	12470
119	NAM S1	Namco, China	30.78	91	4730	17/5/2017	5.28	338F/518R	PM_{10}	Romie et al. 2019	547	14153
120	NAM S2	Namco, China	30.78	91	4730	25/5/2017	5.74	338F/518R	PM_{10}	Romie et al. 2019	638	15715
121	NAM S3	Namco, China	30.78	91	4730	2/6/2017	7.49	338F/518R	PM_{10}	Romie et al. 2019	506	15944
122	NAM S4	Namco, China	30.78	91	4730	13/6/2017	9.65	338F/518R	PM_{10}	Romie et al. 2019	555	14988
123	NAM S5	Namco, China	30.78	91	4730	20/6/2017	8.72	338F/518R	PM_{10}	Romie et al. 2019	515	7830
124	NAM S7	Namco, China	30.78	91	4730	7/7/2017	10.01	338F/518R	PM_{10}	Romie et al. 2019	681	14015
125	NAM S8	Namco, China	30.78	91	4730	14/7/2017	11.53	338F/518R	PM_{10}	Romie et al. 2019	833	21432

126	NAM S9	Namco, China	30.78	91	4730	21/7/2017	11.01	338F/518R	PM_{10}	Romie et al. 2019	622	16796
127	PDD S1	PuydeDôme, France	45.77	2.97	1465	29/6/2016	20.09	338F/518R	PM_{10}	Romie et al. 2019	444	28056
128	PDD S10	PuydeDôme, France	45.77	2.97	1465	1/9/2016	21.89	338F/518R	PM_{10}	Romie et al. 2019	369	17925
129	PDD S11	PuydeDôme, France	45.77	2.97	1465	7/9/2016	22.28	338F/518R	PM_{10}	Romie et al. 2019	434	14885
130	PDD S12	PuydeDôme, France	45.77	2.97	1465	14/9/2016	15.2	338F/518R	PM_{10}	Romie et al. 2019	387	26865
131	PDD S13	PuydeDôme, France	45.77	2.97	1465	21/9/2016	15.3	338F/518R	PM_{10}	Romie et al. 2019	327	17547
132	PDD S2	PuydeDôme, France	45.77	2.97	1465	6/7/2016	21.55	338F/518R	PM_{10}	Romie et al. 2019	293	24593
133	PDD S3	PuydeDôme, France	45.77	2.97	1465	13/7/2016	19.53	338F/518R	PM_{10}	Romie et al. 2019	291	20552
134	PDD S4	PuydeDôme, France	45.77	2.97	1465	20/7/2016	21.5	338F/518R	PM_{10}	Romie et al. 2019	268	45354
135	PDD S6	PuydeDôme, France	45.77	2.97	1465	3/8/2016	19.17	338F/518R	PM_{10}	Romie et al. 2019	452	18791
136	PDD S7	PuydeDôme, France	45.77	2.97	1465	10/8/2016	21.13	338F/518R	PM_{10}	Romie et al. 2019	393	20903
137	PDD S8	PuydeDôme, France	45.77	2.97	1465	17/8/2016	20.48	338F/518R	PM_{10}	Romie et al. 2019	490	16592
138	PDD S9	PuydeDôme, France	45.77	2.97	1465	24/8/2016	24.67	338F/518R	PM_{10}	Romie et al. 2019	575	27527
139	PDM S1	Pic-du-Midi, France	43.94	0.14	2876	20/6/2016	20.01	338F/518R	PM_{10}	Romie et al. 2019	1143	2558
140	PDM S10	Pic-du-Midi, France	43.94	0.14	2876	23/8/2016	22.63	338F/518R	PM_{10}	Romie et al. 2019	737	18493
141	PDM S11	Pic-du-Midi, France	43.94	0.14	2876	13/9/2016	15.72	338F/518R	PM_{10}	Romie et al. 2019	755	31115
142	PDM S12	Pic-du-Midi, France	43.94	0.14	2876	20/9/2016	16.94	338F/518R	PM_{10}	Romie et al. 2019	570	9206
143	PDM S13	Pic-du-Midi, France	43.94	0.14	2876	6/9/2016	21.47	338F/518R	PM_{10}	Romie et al. 2019	703	22909
144	PDM S14	Pic-du-Midi, France	43.94	0.14	2876	27/9/2641	17.05	338F/518R	PM_{10}	Romie et al. 2019	851	11037
145	PDM S2	Pic-du-Midi, France	43.94	0.14	2876	29/6/2016	18.36	338F/518R	PM_{10}	Romie et al. 2019	333	9510
146	PDM S4	Pic-du-Midi, France	43.94	0.14	2876	12/7/2016	18.28	338F/518R	PM_{10}	Romie et al. 2019	672	33527
147	PDM S5	Pic-du-Midi, France	43.94	0.14	2876	19/7/2016	20.97	338F/518R	PM_{10}	Romie et al. 2019	919	17819
148	PDM S6	Pic-du-Midi, France	43.94	0.14	2876	26/7/2016	19.97	338F/518R	PM_{10}	Romie et al. 2019	799	15674
149	PDM S7	Pic-du-Midi, France	43.94	0.14	2876	2/8/2016	20.1	338F/518R	PM_{10}	Romie et al. 2019	902	22386
150	PDM S9	Pic-du-Midi, France	43.94	0.14	2876	16/8/2016	20.12	338F/518R	PM_{10}	Romie et al. 2019	613	27042
151	STN S1	Station-Nord, Greenland	81.57	16.64	37	27/3/2017	-22.61	338F/518R	PM_{10}	Romie et al. 2019	91	5384
152	STN S10	Station-Nord, Greenland	81.57	16.64	37	29/5/2017	-3.48	338F/518R	PM_{10}	Romie et al. 2019	26	155
153	STN S11	Station-Nord, Greenland	81.57	16.64	37	5/6/2017	-2.78	338F/518R	PM_{10}	Romie et al. 2019	165	10926
154	STN S12	Station-Nord, Greenland	81.57	16.64	37	12/6/2017	-4.13	338F/518R	PM_{10}	Romie et al. 2019	236	15440
155	STN S13	Station-Nord, Greenland	81.57	16.64	37	19/6/2017	-6.18	338F/518R	PM_{10}	Romie et al. 2019	290	28385
156	STN S2	Station-Nord, Greenland	81.57	16.64	37	3/4/2017	-16.56	338F/518R	PM_{10}	Romie et al. 2019	128	970
157	STN S3	Station-Nord, Greenland	81.57	16.64	37	10/4/2017	-23.59	338F/518R	PM_{10}	Romie et al. 2019	266	8738
158	STN S4	Station-Nord, Greenland	81.57	16.64	37	17/4/2017	-20.35	338F/518R	PM_{10}	Romie et al. 2019	145	1315

159	STN S5	Station-Nord, Greenland	81.57	16.64	37	24/4/2017	-20.9	338F/518R	PM_{10}	Romie et al. 2019	472	5856
160	STN S6	Station-Nord, Greenland	81.57	16.64	37	1/5/2016	-14.42	338F/518R	PM_{10}	Romie et al. 2019	251	2343
161	STN S7	Station-Nord, Greenland	81.57	16.64	37	8/5/2017	-9.08	338F/518R	PM_{10}	Romie et al. 2019	218	1744
162	STN S8	Station-Nord, Greenland	81.57	16.64	37	15/5/2017	-7.96	338F/518R	PM_{10}	Romie et al. 2019	502	47026
163	STN S9	Station-Nord, Greenland	81.57	16.64	37	22/5/2017	-1.72	338F/518R	PM_{10}	Romie et al. 2019	206	1951
164	STP S1	Colorado, USA	40.46	106.74	3220	14/7/2017	24.68	338F/518R	PM_{10}	Romie et al. 2019	763	38363
165	STP S2	Colorado, USA	40.46	106.74	3220	21/7/2017	23.75	338F/518R	PM_{10}	Romie et al. 2019	696	36037
166	STP S3	Colorado, USA	40.46	106.74	3220	28/7/2017	21.97	338F/518R	PM_{10}	Romie et al. 2019	789	43072
167	STP S7	Colorado, USA	40.46	106.74	3220	25/8/2017	22.59	338F/518R	PM_{10}	Romie et al. 2019	790	27320
168	L_06_W1	Ljungbyhed, Sweden	56.08	13.22	43	2/1/2006	-1.1	515F/806R	TSP	Karlsson et al. 2020	152	13575
169	L_06_W3	Ljungbyhed, Sweden	56.08	13.22	43	16/1/2006	-5.4	515F/806R	TSP	Karlsson et al. 2020	153	14475
170	L_06_W5	Ljungbyhed, Sweden	56.08	13.22	43	30/1/2006	-3.5	515F/806R	TSP	Karlsson et al. 2020	154	16504
171	L_06_W9	Ljungbyhed, Sweden	56.08	13.22	43	27/2/2006	-1.7	515F/806R	TSP	Karlsson et al. 2020	380	135245
172	L_06_W15	Ljungbyhed, Sweden	56.08	13.22	43	10/4/2006	4.8	515F/806R	TSP	Karlsson et al. 2020	721	80022
173	L_06_W17	Ljungbyhed, Sweden	56.08	13.22	43	24/4/2006	8.1	515F/806R	TSP	Karlsson et al. 2020	685	80319
174	L_06_W19	Ljungbyhed, Sweden	56.08	13.22	43	8/5/2006	14.5	515F/806R	TSP	Karlsson et al. 2020	870	23318
175	L_06_W21	Ljungbyhed, Sweden	56.08	13.22	43	22/5/2006	10.3	515F/806R	TSP	Karlsson et al. 2020	1657	62957
176	L_06_W23	Ljungbyhed, Sweden	56.08	13.22	43	5/6/2006	12.1	515F/806R	TSP	Karlsson et al. 2020	1412	15768
177	L_06_W25	Ljungbyhed, Sweden	56.08	13.22	43	19/6/2006	16.7	515F/806R	TSP	Karlsson et al. 2020	1334	64601
178	L_06_W27	Ljungbyhed, Sweden	56.08	13.22	43	3/7/2006	20.1	515F/806R	TSP	Karlsson et al. 2020	974	59396
179	L_06_W29	Ljungbyhed, Sweden	56.08	13.22	43	17/7/2006	20	515F/806R	TSP	Karlsson et al. 2020	1010	101982
180	L_06_W31	Ljungbyhed, Sweden	56.08	13.22	43	31/7/2006	17.9	515F/806R	TSP	Karlsson et al. 2020	1185	52203
181	L_06_W33	Ljungbyhed, Sweden	56.08	13.22	43	14/8/2006	16.1	515F/806R	TSP	Karlsson et al. 2020	1140	30067
182	L_06_W35	Ljungbyhed, Sweden	56.08	13.22	43	21/8/2006	15.2	515F/806R	TSP	Karlsson et al. 2020	1188	4137
183	L_06_W37	Ljungbyhed, Sweden	56.08	13.22	43	4/9/2006	14.9	515F/806R	TSP	Karlsson et al. 2020	992	54173
184	L_06_W39	Ljungbyhed, Sweden	56.08	13.22	43	18/9/2006	14.9	515F/806R	TSP	Karlsson et al. 2020	987	69492
185	L_06_W41	Ljungbyhed, Sweden	56.08	13.22	43	2/10/2006	11.8	515F/806R	TSP	Karlsson et al. 2020	589	65199
186	L_06_W43	Ljungbyhed, Sweden	56.08	13.22	43	16/10/2006	6	515F/806R	TSP	Karlsson et al. 2020	458	40972
187	L_06_W45	Ljungbyhed, Sweden	56.08	13.22	43	30/10/2006	7.4	515F/806R	TSP	Karlsson et al. 2020	782	121054
188	L_06_W47	Ljungbyhed, Sweden	56.08	13.22	43	13/11/2006	6.3	515F/806R	TSP	Karlsson et al. 2020	528	150335
189	L_06_W49	Ljungbyhed, Sweden	56.08	13.22	43	27/11/2006	6.6	515F/806R	TSP	Karlsson et al. 2020	555	92478
190	L_06_W51	Ljungbyhed, Sweden	56.08	13.22	43	11/12/2006	4.1	515F/806R	TSP	Karlsson et al. 2020	484	79238
191	L_07_W1	Ljungbyhed, Sweden	56.08	13.22	43	1/1/2007	4.5	515F/806R	TSP	Karlsson et al. 2020	380	55779
192	L_07_W3	Ljungbyhed, Sweden	56.08	13.22	43	15/1/2007	3.8	515F/806R	TSP	Karlsson et al. 2020	267	145172
193	L_07_W5	Ljungbyhed, Sweden	56.08	13.22	43	29/1/2007	2.9	515F/806R	TSP	Karlsson et al. 2020	379	153957
194	L_07_W7	Ljungbyhed, Sweden	56.08	13.22	43	12/2/2007	0.2	515F/806R	TSP	Karlsson et al. 2020	703	44513
195	L_07_W9	Ljungbyhed, Sweden	56.08	13.22	43	26/2/2007	1.4	515F/806R	TSP	Karlsson et al. 2020	895	168463

196	L_07_W11	Ljungbyhed, Sweden	56.08	13.22	43	12/3/2007	6.3	515F/806R	TSP	Karlsson et al. 2020	1727	114836
197	L_07_W13	Ljungbyhed, Sweden	56.08	13.22	43	26/3/2007	6.9	515F/806R	TSP	Karlsson et al. 2020	1211	104429
198	L_07_W15	Ljungbyhed, Sweden	56.08	13.22	43	9/4/2007	8.7	515F/806R	TSP	Karlsson et al. 2020	1079	71782
199	L_07_W17	Ljungbyhed, Sweden	56.08	13.22	43	23/4/2007	10.8	515F/806R	TSP	Karlsson et al. 2020	1001	56465
200	L_07_W19	Ljungbyhed, Sweden	56.08	13.22	43	7/5/2007	10.3	515F/806R	TSP	Karlsson et al. 2020	1017	9729
201	L_07_W21	Ljungbyhed, Sweden	56.08	13.22	43	21/5/2007	13.1	515F/806R	TSP	Karlsson et al. 2020	1885	1420
202	L_07_W27	Ljungbyhed, Sweden	56.08	13.22	43	2/7/2007	15.2	515F/806R	TSP	Karlsson et al. 2020	1456	171726
203	L_07_W29	Ljungbyhed, Sweden	56.08	13.22	43	16/7/2007	16.1	515F/806R	TSP	Karlsson et al. 2020	1406	19383
204	L_07_W31	Ljungbyhed, Sweden	56.08	13.22	43	30/7/2007	17.3	515F/806R	TSP	Karlsson et al. 2020	660	88278
205	L_07_W33	Ljungbyhed, Sweden	56.08	13.22	43	13/8/2007	16.4	515F/806R	TSP	Karlsson et al. 2020	843	16296
206	L_07_W35	Ljungbyhed, Sweden	56.08	13.22	43	27/8/2007	12.1	515F/806R	TSP	Karlsson et al. 2020	765	344154
207	L_07_W37	Ljungbyhed, Sweden	56.08	13.22	43	10/9/2007	11.8	515F/806R	TSP	Karlsson et al. 2020	895	78139
208	L_07_W39	Ljungbyhed, Sweden	56.08	13.22	43	24/9/2007	10	515F/806R	TSP	Karlsson et al. 2020	1001	132941
209	L_07_W41	Ljungbyhed, Sweden	56.08	13.22	43	8/10/2007	6.6	515F/806R	TSP	Karlsson et al. 2020	1062	40041
210	L_07_W43	Ljungbyhed, Sweden	56.08	13.22	43	22/10/2007	6.2	515F/806R	TSP	Karlsson et al. 2020	703	89564
211	L_07_W45	Ljungbyhed, Sweden	56.08	13.22	43	5/11/2007	2.6	515F/806R	TSP	Karlsson et al. 2020	258	148393
212	L_07_W47	Ljungbyhed, Sweden	56.08	13.22	43	19/11/2007	3.9	515F/806R	TSP	Karlsson et al. 2020	257	106251
213	L_07_W49	Ljungbyhed, Sweden	56.08	13.22	43	3/12/2007	1.4	515F/806R	TSP	Karlsson et al. 2020	791	110585
214	K_06_W1	Kiruna, Sweden	67.84	20.42	393	2/1/2006	-10.2	515F/806R	TSP	Karlsson et al. 2020	135	11755
215	K_06_W3	Kiruna, Sweden	67.84	20.42	393	16/1/2006	-14.2	515F/806R	TSP	Karlsson et al. 2020	231	30518
216	K_06_W5	Kiruna, Sweden	67.84	20.42	393	30/1/2006	-12.1	515F/806R	TSP	Karlsson et al. 2020	170	23807
217	K_06_W11	Kiruna, Sweden	67.84	20.42	393	13/3/2006	-2.6	515F/806R	TSP	Karlsson et al. 2020	528	136288
218	K_06_W13	Kiruna, Sweden	67.84	20.42	393	27/3/2006	-10.9	515F/806R	TSP	Karlsson et al. 2020	214	12729
219	K_06_W15	Kiruna, Sweden	67.84	20.42	393	10/4/2006	-4.1	515F/806R	TSP	Karlsson et al. 2020	257	82917
220	K_06_W17	Kiruna, Sweden	67.84	20.42	393	24/4/2006	2	515F/806R	TSP	Karlsson et al. 2020	415	47724
221	K_06_W19	Kiruna, Sweden	67.84	20.42	393	8/5/2006	6.1	515F/806R	TSP	Karlsson et al. 2020	249	18576
222	K_06_W25	Kiruna, Sweden	67.84	20.42	393	19/6/2006	14.5	515F/806R	TSP	Karlsson et al. 2020	354	331
223	K_06_W27	Kiruna, Sweden	67.84	20.42	393	3/7/2006	13.3	515F/806R	TSP	Karlsson et al. 2020	782	8973
224	K_06_W29	Kiruna, Sweden	67.84	20.42	393	17/7/2006	11.2	515F/806R	TSP	Karlsson et al. 2020	940	19911
225	K_06_W31	Kiruna, Sweden	67.84	20.42	393	31/7/2006	14.3	515F/806R	TSP	Karlsson et al. 2020	795	71493
226	K_06_W33	Kiruna, Sweden	67.84	20.42	393	14/8/2006	13.4	515F/806R	TSP	Karlsson et al. 2020	1097	55849
227	K_06_W35	Kiruna, Sweden	67.84	20.42	393	21/8/2006	11.8	515F/806R	TSP	Karlsson et al. 2020	362	10919
228	K_06_W37	Kiruna, Sweden	67.84	20.42	393	4/9/2006	7.2	515F/806R	TSP	Karlsson et al. 2020	214	20445
229	K_06_W39	Kiruna, Sweden	67.84	20.42	393	18/9/2006	1.1	515F/806R	TSP	Karlsson et al. 2020	222	23707
230	K_06_W41	Kiruna, Sweden	67.84	20.42	393	2/10/2006	2	515F/806R	TSP	Karlsson et al. 2020	494	23020
231	K_06_W43	Kiruna, Sweden	67.84	20.42	393	16/10/2006	-5.9	515F/806R	TSP	Karlsson et al. 2020	537	84649
232	K_06_W45	Kiruna, Sweden	67.84	20.42	393	30/10/2006	-7.2	515F/806R	TSP	Karlsson et al. 2020	344	106306
233	K_06_W47	Kiruna, Sweden	67.84	20.42	393	13/11/2006	-4.7	515F/806R	TSP	Karlsson et al. 2020	362	83662

234	K_06_W49	Kiruna, Sweden	67.84	20.42	393	27/11/2006	-5.4	515F/806R	TSP	Karlsson et al. 2020	240	136655
235	K_06_W51	Kiruna, Sweden	67.84	20.42	393	11/12/2006	-5.9	515F/806R	TSP	Karlsson et al. 2020	188	80180
236	K_07_W1	Kiruna, Sweden	67.84	20.42	393	1/1/2007	-7.8	515F/806R	TSP	Karlsson et al. 2020	354	9460
237	K_07_W3	Kiruna, Sweden	67.84	20.42	393	15/1/2007	-12.7	515F/806R	TSP	Karlsson et al. 2020	266	18418
238	K_07_W5	Kiruna, Sweden	67.84	20.42	393	29/1/2007	-15.7	515F/806R	TSP	Karlsson et al. 2020	170	112631
239	K_07_W7	Kiruna, Sweden	67.84	20.42	393	12/2/2007	-17.9	515F/806R	TSP	Karlsson et al. 2020	214	6420
240	K_07_W9	Kiruna, Sweden	67.84	20.42	393	26/2/2007	-12.4	515F/806R	TSP	Karlsson et al. 2020	249	14854
241	K_07_W11	Kiruna, Sweden	67.84	20.42	393	12/3/2007	-3.2	515F/806R	TSP	Karlsson et al. 2020	310	47980
242	K_07_W13	Kiruna, Sweden	67.84	20.42	393	26/3/2007	1.9	515F/806R	TSP	Karlsson et al. 2020	433	77292
243	K_07_W15	Kiruna, Sweden	67.84	20.42	393	9/4/2007	-0.5	515F/806R	TSP	Karlsson et al. 2020	450	28788
244	K_07_W17	Kiruna, Sweden	67.84	20.42	393	23/4/2007	1.8	515F/806R	TSP	Karlsson et al. 2020	310	57908
245	K_07_W19	Kiruna, Sweden	67.84	20.42	393	7/5/2007	1.7	515F/806R	TSP	Karlsson et al. 2020	292	36187
246	K_07_W21	Kiruna, Sweden	67.84	20.42	393	21/5/2007	7.5	515F/806R	TSP	Karlsson et al. 2020	318	41729
247	K_07_W23	Kiruna, Sweden	67.84	20.42	393	4/6/2007	12.5	515F/806R	TSP	Karlsson et al. 2020	887	16880
248	K_07_W29	Kiruna, Sweden	67.84	20.42	393	2/7/2007	12.9	515F/806R	TSP	Karlsson et al. 2020	956	68997
249	K_07_W31	Kiruna, Sweden	67.84	20.42	393	16/7/2007	13.9	515F/806R	TSP	Karlsson et al. 2020	703	36535
250	K_07_W33	Kiruna, Sweden	67.84	20.42	393	30/7/2007	13.1	515F/806R	TSP	Karlsson et al. 2020	581	26335
251	K_07_W35	Kiruna, Sweden	67.84	20.42	393	13/8/2007	4.4	515F/806R	TSP	Karlsson et al. 2020	1491	9628
252	K_07_W37	Kiruna, Sweden	67.84	20.42	393	27/8/2007	5.8	515F/806R	TSP	Karlsson et al. 2020	800	11683
253	K_07_W39	Kiruna, Sweden	67.84	20.42	393	10/9/2007	5.7	515F/806R	TSP	Karlsson et al. 2020	721	96393
254	K_07_W41	Kiruna, Sweden	67.84	20.42	393	24/9/2007	0.5	515F/806R	TSP	Karlsson et al. 2020	616	57206
255	K_07_W43	Kiruna, Sweden	67.84	20.42	393	8/10/2007	1.4	515F/806R	TSP	Karlsson et al. 2020	502	66301
256	K_07_W45	Kiruna, Sweden	67.84	20.42	393	22/10/2007	-4.7	515F/806R	TSP	Karlsson et al. 2020	484	129464
257	K_07_W47	Kiruna, Sweden	67.84	20.42	393	5/11/2007	-4.4	515F/806R	TSP	Karlsson et al. 2020	467	153631
258	K_07_W49	Kiruna, Sweden	67.84	20.42	393	19/11/2007	-5.3	515F/806R	TSP	Karlsson et al. 2020	573	38664
259	K_07_W51	Kiruna, Sweden	67.84	20.42	393	3/12/2007	-5.9	515F/806R	TSP	Karlsson et al. 2020	773	131477
260	SA1	South Ocean	-44.96	146.31	23	12/1/2018	17.8	27F/519R & 515F/926R	TSP	Uetake et al. 2020	389	17243
261	SA10	South Ocean	-65.04	139.85	23	30/1/2018	-0.4	27F/519R & 515F/926R	TSP	Uetake et al. 2020	26	9244
262	SA11	South Ocean	-65.53	147.47	23	2/2/2018	-3.7	27F/519R & 515F/926R	TSP	Uetake et al. 2020	14	6361
263	SA12	South Ocean	-64.4	150	23	4/2/2018	-3.3	27F/519R & 515F/926R	TSP	Uetake et al. 2020	9	5356
264	SA13	South Ocean	-62.8	149.07	23	5/2/2018	-4.1	27F/519R & 515F/926R	TSP	Uetake et al. 2020	56	6041
265	SA14	South Ocean	-62.82	144.24	23	7/2/2018	1.6	27F/519R & 515F/926R	TSP	Uetake et al. 2020	100	9491
266	SA15	South Ocean	-62.21	138.56	23	8/2/2018	1.7	27F/519R & 515F/926R	TSP	Uetake et al. 2020	168	10353

267	SA16	South Ocean	-62.31	133.31	23	10/2/2018	0.9	27F/519R & 515F/926R	TSP	Uetake et al. 2020	47	5328
268	SA18	South Ocean	-60.96	132.06	23	13/2/2018	1.9	27F/519R & 515F/926R	TSP	Uetake et al. 2020	109	15869
269	SA19	South Ocean	-58.27	132.19	23	14/2/2018	2.7	27F/519R & 515F/926R	TSP	Uetake et al. 2020	68	11328
270	SA2	South Ocean	-47.34	144.92	23	16/1/2018	10.3	27F/519R & 515F/926R	TSP	Uetake et al. 2020	186	20463
271	SA20	South Ocean	-56.7	138.1	23	16/2/2018	3.7	27F/519R & 515F/926R	TSP	Uetake et al. 2020	226	24527
272	SA21	South Ocean	-56.55	141.49	23	18/2/2018	2.9	27F/519R & 515F/926R	TSP	Uetake et al. 2020	123	25198
273	SA22	South Ocean	-53.64	142.97	23	19/2/2018	4.3	27F/519R & 515F/926R	TSP	Uetake et al. 2020	37	8096
274	SA23	South Ocean	-48.44	144.48	23	19/2/2018	5.1	27F/519R & 515F/926R	TSP	Uetake et al. 2020	186	4145
275	SA3	South Ocean	-49.14	144.16	23	16/1/2018	11.3	27F/519R & 515F/926R	TSP	Uetake et al. 2020	186	25981
276	SA4	South Ocean	-50.54	143.48	23	18/1/2018	9.9	27F/519R & 515F/926R	TSP	Uetake et al. 2020	310	19694
277	SA7	South Ocean	-55.12	140.94	23	22/1/2018	5.6	27F/519R & 515F/926R	TSP	Uetake et al. 2020	31	11394
278	SA9	South Ocean	-61.72	139.85	23	27/1/2018	1.6	27F/519R & 515F/926R	TSP	Uetake et al. 2020	38	17461
279	14To_1	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	16/3/2014	0.9	515F/806R	TSP	Maki et al. 2017	510	109202
280	14To_2	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	16/3/2014	0.9	515F/806R	TSP	Maki et al. 2017	909	17436
281	14To_3	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	17/3/2014	1.01	515F/806R	TSP	Maki et al. 2017	1332	216602
282	14To_4	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	18/3/2014	-16.85	515F/806R	TSP	Maki et al. 2017	603	36723
283	15To_1	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	7/3/2015	-8.28	515F/806R	TSP	Maki et al. 2017	174	35657
284	15To_2	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	7/3/2015	-8.28	515F/806R	TSP	Maki et al. 2017	379	12433
285	15To_3	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	9/3/2015	-10.4	515F/806R	TSP	Maki et al. 2017	72	10251
286	15To_4	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	9/3/2015	-10.4	515F/806R	TSP	Maki et al. 2017	155	27620
287	15To_5	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	10/3/2015	-7.73	515F/806R	TSP	Maki et al. 2017	93	23554
288	15To_6	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	10/3/2015	-7.73	515F/806R	TSP	Maki et al. 2017	1097	192268
289	15To_8	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	26/4/2015	16.46	515F/806R	TSP	Maki et al. 2017	987	171156
290	15To_9	Tsogt-Ovoo, Mongolia	44.23	105.17	1256	27/4/2015	9.65	515F/806R	TSP	Maki et al. 2017	387	121260
291	C1	Colorado, USA	40	-105.05	1624		27.7	515F/806R	PM_{10}	EMP	160	20924
292	AG1	Colorado, USA	40	-105.25	1624		16.4	515F/806R	PM_{10}	EMP	241	95919
293	F1	Colorado, USA	40.05	-105.52	1522		15.7	515F/806R	PM_{10}	EMP	95	73550
294	C2	Colorado, USA	40.15	-105.03	1519		22.5	515F/806R	PM_{10}	EMP	250	85366

295	AG2	Colorado, USA	40.16	-105.1	1519		29.3	515F/806R	PM_{10}	EMP	122	143824
296	F2	Colorado, USA	40.18	-105.31	1522		0	515F/806R	PM_{10}	EMP	95	455573
297	F3	Colorado, USA	40.3	-105.52	1783		22.6	515F/806R	PM_{10}	EMP	84	54234
298	AG3	Colorado, USA	40.42	-105.07	1519		23.7	515F/806R	PM_{10}	EMP	155	81007
299	SPL_Wint9	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	269	19978
300	SPL_Sp29	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	169	43823
301	SPL_Sp25	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	174	49460
302	SPL_Sum5	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	389	42694
303	SPL_Sum9	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	274	23823
304	SPL_Fall1	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	305	23064
305	SPL_Sp21	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	109	13277
306	SPL_Sum1	Colorado, USA	40.45	-106.74	2553		0	515F/806R	PM_{10}	EMP	159	40426
307	C3	Colorado, USA	40.59	-105.08	1525		0	515F/806R	PM_{10}	EMP	225	29267
308	Cle_W3	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	1426	65904
309	Cle_W2	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	1430	101161
310	Cle_W1	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	1763	96018
311	Cle_4	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	248	24643
312	Cle_3	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	166	11543
313	Cle_1	Ohio, USA	41.42	-81.87	199		0	515F/806R	PM_{10}	EMP	149	44446
314	Chi_W1	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	572	133870
315	Chi_4	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	333	54774
316	Chi_W2	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	869	133150
317	Chi_W3	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	461	40371
318	Chi_2	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	305	21464
319	Chi_1	Ohio, USA	41.78	-87.75	181		0	515F/806R	PM_{10}	EMP	389	13324
320	Det_2	Michigan, USA	42.42	-83.02	183		0	515F/806R	PM_{10}	EMP	457	174943
321	Det_W6	Michigan, USA	42.42	-83.02	183		0	515F/806R	PM_{10}	EMP	335	18241
322	Det_4	Michigan, USA	42.42	-83.02	183		0	515F/806R	PM_{10}	EMP	232	114385
323	Det_1	Michigan, USA	42.42	-83.02	183		0	515F/806R	PM_{10}	EMP	213	46459
324	Det_W7	Michigan, USA	42.42	-83.02	183		0	515F/806R	PM_{10}	EMP	343	16239
325	Mayv_W5	Michigan, USA	43.49	-88.55	283		0	515F/806R	PM_{10}	EMP	741	62164
326	Mayv_2	Michigan, USA	43.49	-88.55	283		0	515F/806R	PM_{10}	EMP	334	59895
327	Mayv_W2	Michigan, USA	43.49	-88.55	283		0	515F/806R	PM_{10}	EMP	526	93134

328	MZ_22a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	1279	192266
329	MZ_103b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	385	38025
330	MZ_15a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	1038	208582
331	MZ_41a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	419	160058
332	MZ_41b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	787	172427
333	MZ_50b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	362	132845
334	MZ_62b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	584	203976
335	MZ_59b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	321	192637
336	MZ_90a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	426	49718
337	MZ_88b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	280	80481
338	MZ_82b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	349	91901
339	MZ_82a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	1138	97248
340	MZ_52b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	269	136590
341	MZ_90b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	279	108073
342	MZ_50a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	466	99998
343	MZ_67a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	654	188900
344	MZ_31a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	545	170726
345	MZ_47a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	569	250299
346	MZ_26a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	983	141743
347	MZ_93a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	901	256244
348	MZ_62a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	736	222441
349	MZ_103a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	676	125351
350	MZ_93b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	304	117535
351	MZ_54a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	469	216768
352	MZ_81b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	222	134733
353	MZ_23a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	277	86496
354	MZ_15b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	383	165908
355	MZ_22b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	230	92854
356	MZ_101b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	260	52731
357	MZ_47b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	654	20496
358	MZ_81a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	309	109200
359	MZ_54b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	219	69811
360	MZ_74a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	473	147015

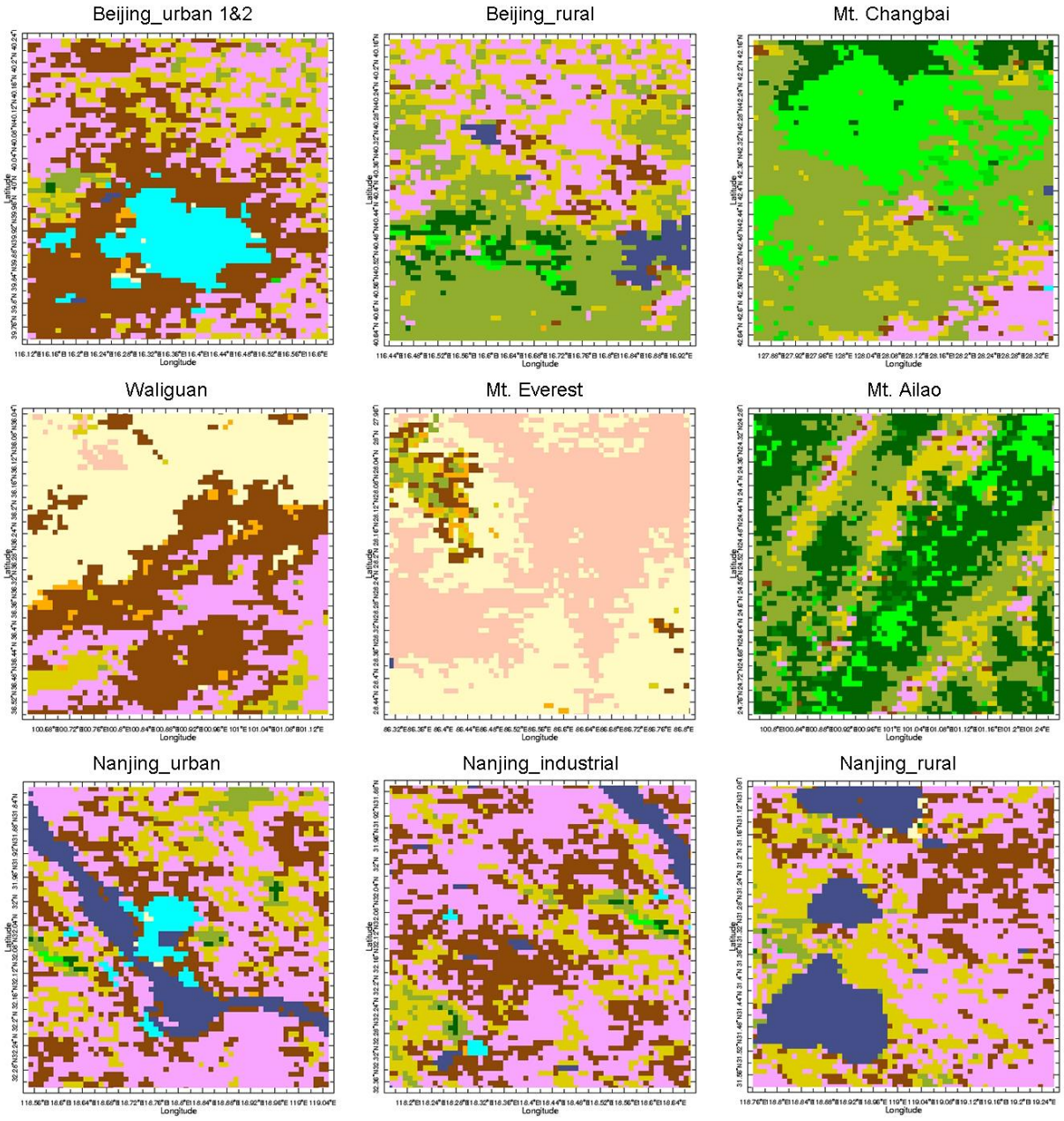
361	MZ_101a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	320	114214
362	MZ_11b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	432	22328
363	MZ_88a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	359	143909
364	MZ_26b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	347	108270
365	MZ_23b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	158	43745
366	MZ_59a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	518	235048
367	MZ_31b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	280	34135
368	MZ_67b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	224	84427
369	MZ_74b	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	246	62122
370	MZ_52a	Wiesbaden, Germany	50.01	8.23	95		0	515F/806R	PM_{10}	EMP	242	197428
371	GL_A1	low-Arctic Greenland	64.12	-54.37	500	26/7/2013	3.13	908F/1075R	PM_{10}	Temkiv et al. 2018	47	
372	GL_R2	low-Arctic Greenland	64.12	-54.37	500	26/7/2013	7.89	908F/1075R	PM_{10}	Temkiv et al. 2018	156	
373	GL_R3	low-Arctic Greenland	64.12	-54.37	500	27/7/2013	8.44	908F/1075R	PM_{10}	Temkiv et al. 2018	139	
374	GL_R1	low-Arctic Greenland	64.12	-54.37	500	25/7/2013	8.5	908F/1075R	PM_{10}	Temkiv et al. 2018	212	
375	GL_A2	low-Arctic Greenland	64.12	-54.37	500	27/7/2013	8.82	908F/1075R	PM_{10}	Temkiv et al. 2018	279	
376	GL_A3	low-Arctic Greenland	64.12	-54.37	500	27/7/2013	8.82	908F/1075R	PM_{10}	Temkiv et al. 2018	295	
377	GL_A4	low-Arctic Greenland	64.12	-54.37	500	28/7/2013	9.53	908F/1075R	PM_{10}	Temkiv et al. 2018	193	
378	GL_A16	low-Arctic Greenland	64.12	-54.37	500	1/8/2013	9.53	908F/1075R	PM_{10}	Temkiv et al. 2018	148	
379	GL_A17	low-Arctic Greenland	64.12	-54.37	500	1/8/2013	10.97	908F/1075R	PM_{10}	Temkiv et al. 2018	282	
380	GL_A15	low-Arctic Greenland	64.12	-54.37	500	31/7/2013	11.57	908F/1075R	PM_{10}	Temkiv et al. 2018	334	
381	GL_A14	low-Arctic Greenland	64.12	-54.37	500	31/7/2013	11.73	908F/1075R	PM_{10}	Temkiv et al. 2018	129	
382	GL_A12	low-Arctic Greenland	64.12	-54.37	500	31/7/2013	12.07	908F/1075R	PM_{10}	Temkiv et al. 2018	448	
383	GL_A6	low-Arctic Greenland	64.12	-54.37	500	28/7/2013	12.15	908F/1075R	PM_{10}	Temkiv et al. 2018	147	
384	GL_A13	low-Arctic Greenland	64.12	-54.37	500	31/7/2013	12.7	908F/1075R	PM_{10}	Temkiv et al. 2018	338	
385	GL_A5	low-Arctic Greenland	64.12	-54.37	500	28/7/2013	13.14	908F/1075R	PM_{10}	Temkiv et al. 2018	246	
386	GL_A9	low-Arctic Greenland	64.12	-54.37	500	30/7/2013	14.15	908F/1075R	PM_{10}	Temkiv et al. 2018	241	
387	GL_A7	low-Arctic Greenland	64.12	-54.37	500	29/7/2013	14.91	908F/1075R	PM_{10}	Temkiv et al. 2018	172	
388	GL_A8	low-Arctic Greenland	64.12	-54.37	500	29/7/2013	14.92	908F/1075R	PM_{10}	Temkiv et al. 2018	277	
389	GL_A10	low-Arctic Greenland	64.12	-54.37	500	30/7/2013	15.87	908F/1075R	PM_{10}	Temkiv et al. 2018	289	
390	GL_A11	low-Arctic Greenland	64.12	-54.37	500	30/7/2013	15.87	908F/1075R	PM_{10}	Temkiv et al. 2018	226	
391	UR_J10	Urumqi, China	43.83	87.62	835	2014_win	-0.8	319F/806R	PM_{10}	Gou et al. 2016	2393	27400
392	UR_M10	Urumqi, China	43.83	87.62	835	2014_win	-10.25	319F/806R	PM_{10}	Gou et al. 2016	1333	22155
393	UR_SM10	Urumqi, China	43.83	87.62	835	2014_win	-7.64	319F/806R	PM_{10}	Gou et al. 2016	1624	28112

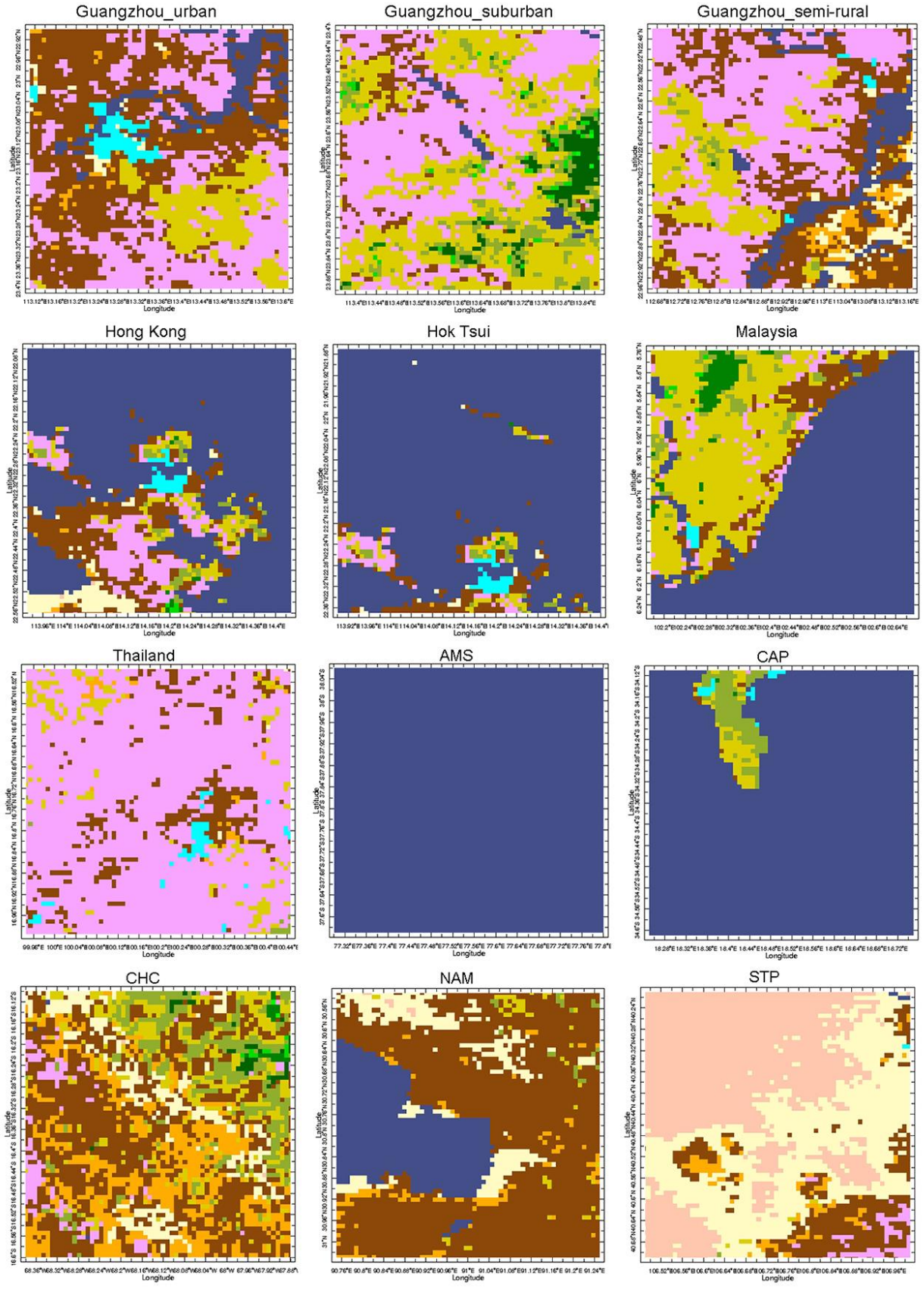
394	UR_T10	Urumqi, China	43.83	87.62	835	2014_win	-15.1	319F/806R	PM_{10}	Gou et al. 2016	2124	36715
395	UR_TS10	Urumqi, China	43.83	87.62	835	2014_win	-11.17	319F/806R	PM_{10}	Gou et al. 2016	1211	28156
396	UR_X10	Urumqi, China	43.83	87.62	835	2014_win	1.06	319F/806R	PM_{10}	Gou et al. 2016	1752	38516
397	Tai	Mt. Tai, China	36.25	117.1	1534	7/2014-8/2015	15.5	515F/926R	$PM_{2.5}$	Xu et al. 2019	1282	7044
398	XX_spr	Central China	35.3	113.92	71	2017_spr	19.63	515F/806R	$PM_{2.5}$	Li et al. 2019	2243	
399	XX_sum	Central China	35.3	113.92	71	2017_sum	25	515F/806R	$PM_{2.5}$	Li et al. 2019	1608	
400	XX_aut	Central China	35.3	113.92	71	2017_sut	3.67	515F/806R	$PM_{2.5}$	Li et al. 2019	1312	
401	XX_win	Central China	35.3	113.92	71	2017_win	1.8	515F/806R	$PM_{2.5}$	Li et al. 2019	1993	
402	P_win	Madrid, Spain	40.44	-3.69	657	2/3/2015	10.3	341F/806R	PM_{10}	Nunez et al. 2019	1050	187059
403	P_spr	Madrid, Spain	40.44	-3.69	657	21/4/2015	14.3	341F/806R	PM_{10}	Nunez et al. 2019	800	246262
404	P_sum	Madrid, Spain	40.44	-3.69	657	20/7/2015	29.2	341F/806R	PM_{10}	Nunez et al. 2019	1520	200235
405	P_aut	Madrid, Spain	40.44	-3.69	657	23/11/2015	9.1	341F/806R	PM_{10}	Nunez et al. 2019	1680	132682
406	GD	Beijing, China	39.97	116.37	64	15/10/2014	11.98	338F/806R	$PM_{2.5}$	Du et al. 2018	1762	11245
407	GN	Beijing, China	39.97	116.37	64	15/10/2014	11.98	338F/806R	$PM_{2.5}$	Du et al. 2018	2034	13521
408	MD	Beijing, China	39.97	116.37	64	17/10/2014	11.83	338F/806R	$PM_{2.5}$	Du et al. 2018	2028	15300
409	MN	Beijing, China	39.97	116.37	64	17/10/2014	11.83	338F/806R	$PM_{2.5}$	Du et al. 2018	1894	15823
410	HD	Beijing, China	39.97	116.37	64	18/10/2014	14.53	338F/806R	$PM_{2.5}$	Du et al. 2018	945	11790
411	HN	Beijing, China	39.97	116.37	64	18/10/2014	14.53	338F/806R	$PM_{2.5}$	Du et al. 2018	1705	12620
412	AD	Beijing, China	39.97	116.37	64	5/11/2017	7.96	338F/806R	$PM_{2.5}$	Du et al. 2018	2212	15268
413	AN	Beijing, China	39.97	116.37	64	5/11/2014	7.96	338F/806R	$PM_{2.5}$	Du et al. 2018	2457	14764
414	DVE2_1	Singapore	1.35	103.68	30	May-16	29.1	341F/806R	TSP	Gusareva et al. 2019	168	
415	DVE2_2	Singapore	1.35	103.68	30	May-16	31.2	341F/806R	TSP	Gusareva et al. 2019	144	
416	DVE2_3	Singapore	1.35	103.68	30	May-16	30.9	341F/806R	TSP	Gusareva et al. 2019	148	
417	DVE2_4	Singapore	1.35	103.68	30	May-16	29.5	341F/806R	TSP	Gusareva et al. 2019	132	
418	DVE2_5	Singapore	1.35	103.68	30	May-16	30.7	341F/806R	TSP	Gusareva et al. 2019	139	
419	DVE3_1	Singapore	1.35	103.68	30	Aug-16	29.5	341F/806R	TSP	Gusareva et al. 2019	233	
420	DVE3_2	Singapore	1.35	103.68	30	Aug-16	30.8	341F/806R	TSP	Gusareva et al. 2019	306	
421	DVE3_3	Singapore	1.35	103.68	30	Aug-16	28.5	341F/806R	TSP	Gusareva et al. 2019	126	
422	DVE3_4	Singapore	1.35	103.68	30	Aug-16	29.1	341F/806R	TSP	Gusareva et al. 2019	158	
423	DVE3_5	Singapore	1.35	103.68	30	Aug-16	29.7	341F/806R	TSP	Gusareva et al. 2019	126	
424	DVE4_1	Singapore	1.35	103.68	30	Dec-16	29.1	341F/806R	TSP	Gusareva et al. 2019	151	
425	DVE4_2	Singapore	1.35	103.68	30	Dec-16	29.5	341F/806R	TSP	Gusareva et al. 2019	202	
426	DVE4_3	Singapore	1.35	103.68	30	Dec-16	33.1	341F/806R	TSP	Gusareva et al. 2019	206	
427	DVE4_4	Singapore	1.35	103.68	30	Dec-16	32	341F/806R	TSP	Gusareva et al. 2019	189	
428	DVE4_5	Singapore	1.35	103.68	30	Dec-16	31.5	341F/806R	TSP	Gusareva et al. 2019	182	

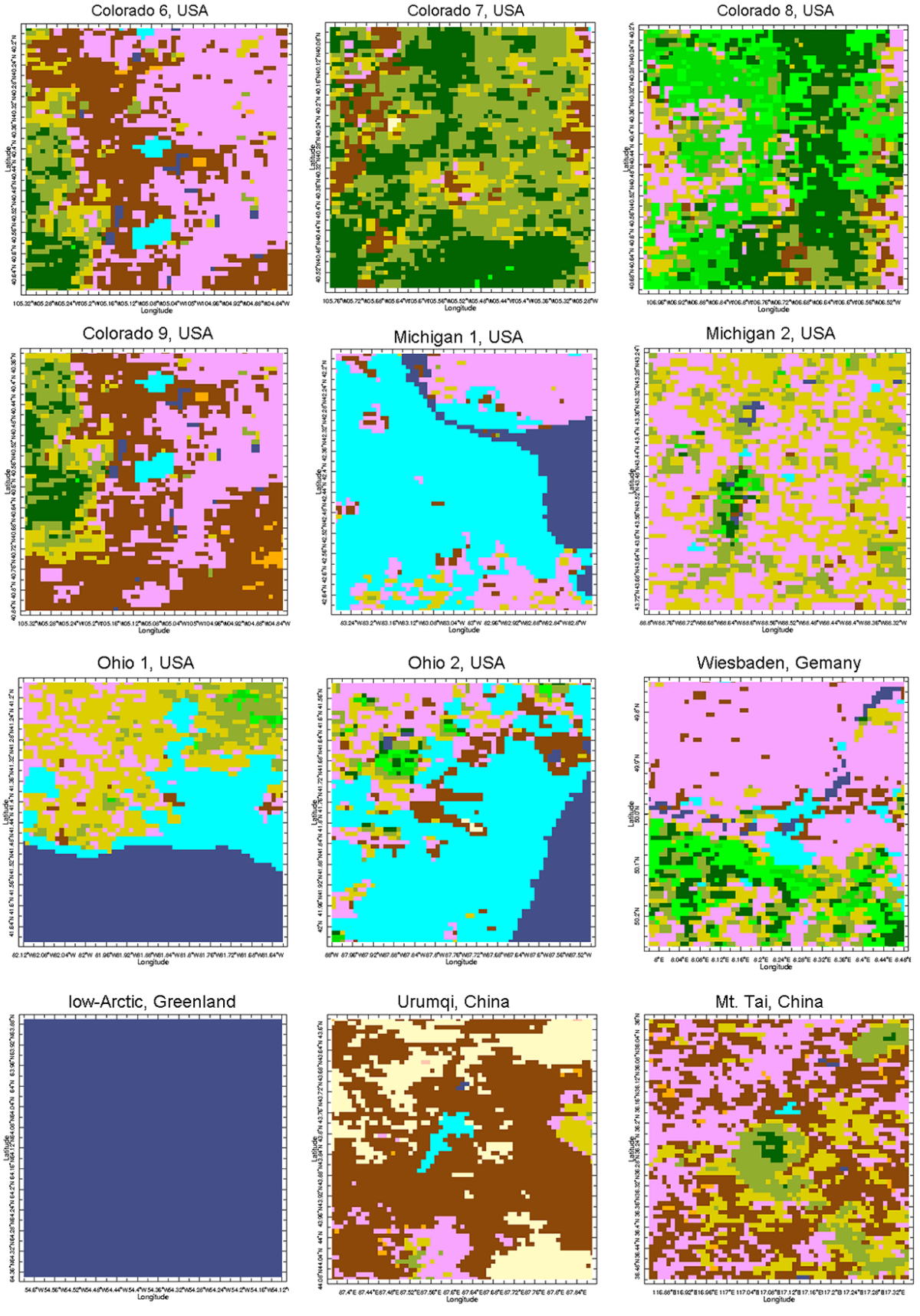
429	DVE5_1	Singapore	1.35	103.68	30	Feb-17	28.1	341F/806R	TSP	Gusareva et al. 2019	175	
430	DVE5_2	Singapore	1.35	103.68	30	Feb-17	27.9	341F/806R	TSP	Gusareva et al. 2019	191	
431	DVE5_3	Singapore	1.35	103.68	30	Feb-17	28.6	341F/806R	TSP	Gusareva et al. 2019	138	
432	DVE5_4	Singapore	1.35	103.68	30	Feb-17	29.3	341F/806R	TSP	Gusareva et al. 2019	194	
433	DVE5_5	Singapore	1.35	103.68	30	Feb-17	28.5	341F/806R	TSP	Gusareva et al. 2019	144	
434	Ev12Euk	Be’er Sheva, Israel	31.25	34.8	300	20/12/2012	13.7	8F/907R	TSP	Katra et al. 2014	409	
435	Ev13Euk	Be’er Sheva, Israel	31.25	34.8	300	20/12/2012	11.2	8F/907R	TSP	Katra et al. 2014	251	
436	Ev12Bac	Be’er Sheva, Israel	31.25	34.8	300	7/1/2013	13.7	8F/907R	TSP	Katra et al. 2014	1214	
437	Ev13Bac	Be’er Sheva, Israel	31.25	34.8	300	7/1/2013	11.2	8F/907R	TSP	Katra et al. 2014	869	
438	BC2009	Beijing, China	39.98	116.32	61	28/8/2009	24.3	27F/357R	PM_{10}	An et al. 2015	2353	10895
439	BC2010	Beijing, China	39.98	116.32	61	27/2/2010	-3.3	27F/357R	PM_{10}	An et al. 2015	603	3386
440	BC2011	Beijing, China	39.98	116.32	61	7/2/2011	-3.7	27F/357R	PM_{10}	An et al. 2015	3419	31959
441	BS2010	Beijing, China	39.98	116.32	61	27/2/2010	26.7	27F/357R	PM_{10}	An et al. 2015	649	6462
442	BS2011a	Beijing, China	39.98	116.32	61	18/3/2011	11	27F/357R	PM_{10}	An et al. 2015	3025	22323
443	BS2011b	Beijing, China	39.98	116.32	61	1/5/2011	21.8	27F/357R	PM_{10}	An et al. 2015	2676	17960
444	GC2009	Gwangju, Korea	35.16	126.95	208	21/8/2009	26.1	27F/357R	PM_{10}	An et al. 2015	791	4545
445	GC2010	Gwangju, Korea	35.16	126.95	208	8/3/2010	26.7	27F/357R	PM_{10}	An et al. 2015	321	3505
446	GC2011	Gwangju, Korea	35.16	126.95	208	17/5/2011	19.4	27F/357R	PM_{10}	An et al. 2015	1915	27054
447	GS2010	Gwangju, Korea	35.16	126.95	208	19/3/2010	9.6	27F/357R	PM_{10}	An et al. 2015	1227	3837
448	GS2011	Gwangju, Korea	35.16	126.95	208	13/5/2011	17.8	27F/357R	PM_{10}	An et al. 2015	1012	22893
449	IC2011	Incheon, Korea	37.46	126.71	24	17/5/2011	17.3	27F/357R	PM_{10}	An et al. 2015	506	23466
450	IS2011a	Incheon, Korea	37.46	126.71	24	1/5/2011	15.3	27F/357R	PM_{10}	An et al. 2015	2666	22047
451	IS2011b	Incheon, Korea	37.46	126.71	24	3/5/2011	15.7	27F/357R	PM_{10}	An et al. 2015	2637	47026
452	SC2009	Seoul, Korea	37.57	126.32	0	18/8/2009	27.8	27F/357R	PM_{10}	An et al. 2015	1315	8383
453	TC2011	Taiyuan, China	37.83	112.54	783	7/2/2011	2.7	27F/357R	PM_{10}	An et al. 2015	3244	21603
454	TS2011a	Taiyuan, China	37.83	112.54	783	18/3/2011	7.7	27F/357R	PM_{10}	An et al. 2015	4026	49755
455	TS2011b	Taiyuan, China	37.83	112.54	783	1/5/2011	16.2	27F/357R	PM_{10}	An et al. 2015	2092	20139

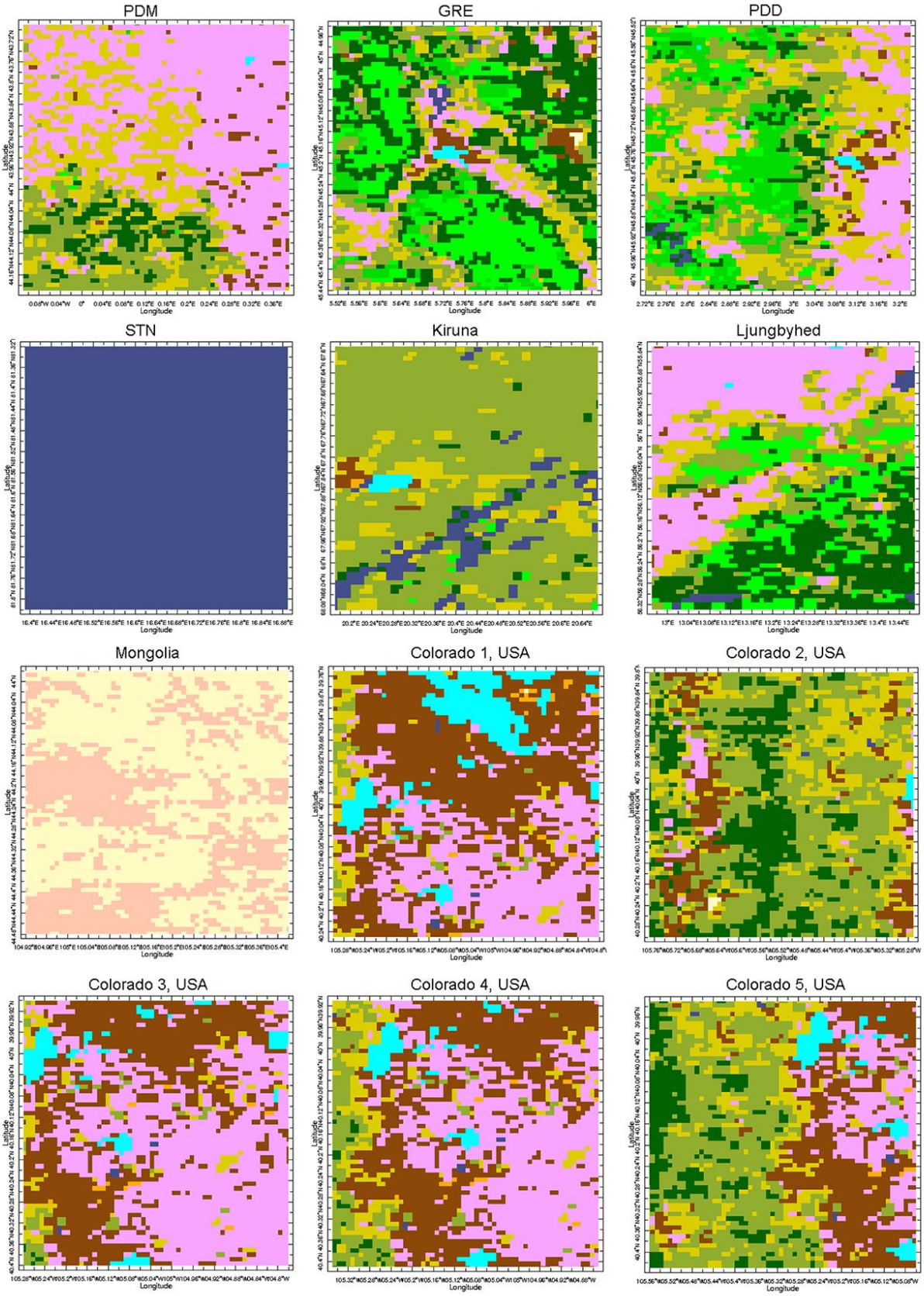
Appendix 2

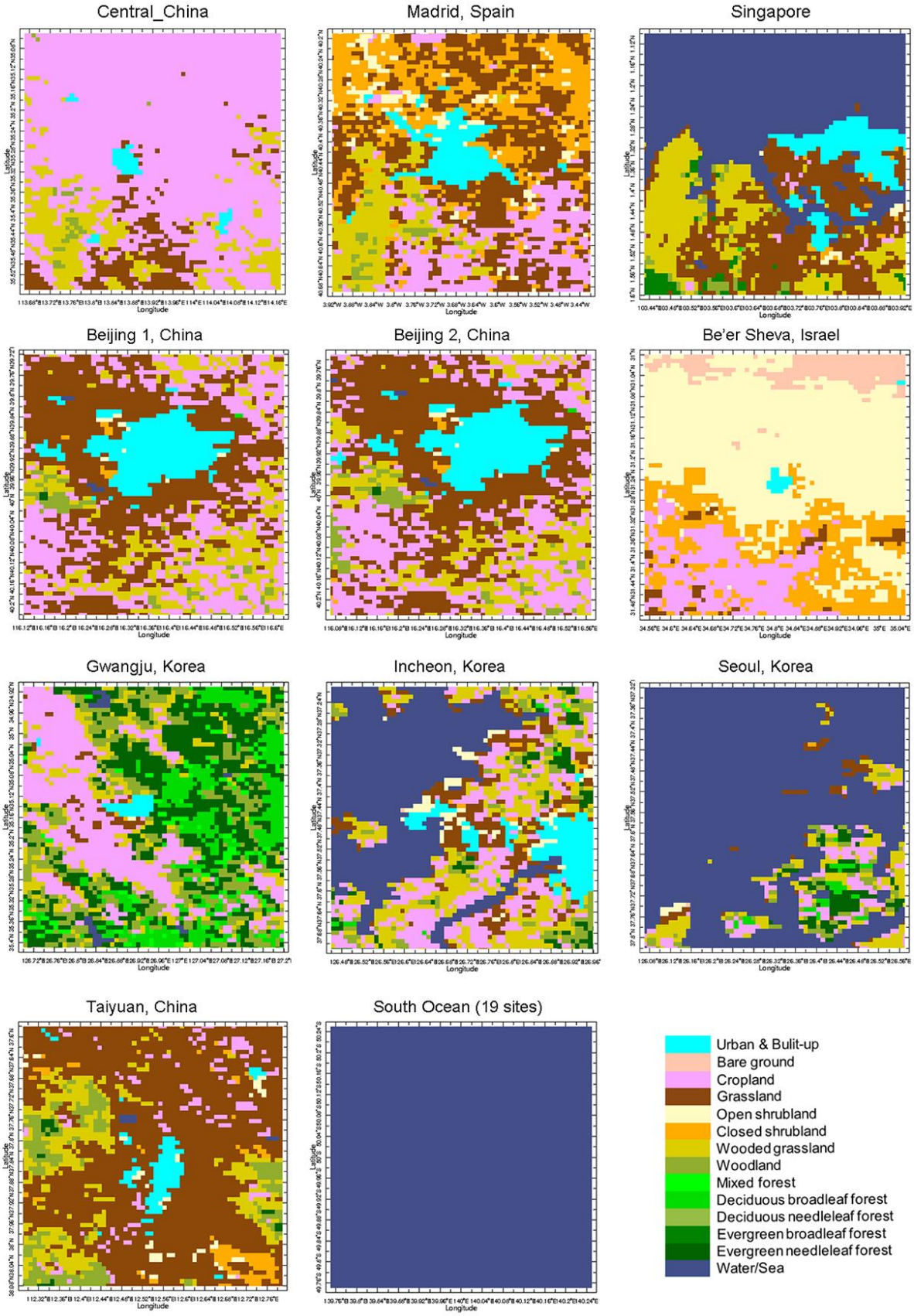
This section provides the maps of land cover type in a diameter range (50 km) of the sampling sites (*n=74*) performed with MODIS land cover approach (5′ × 5′ resolution).




Guangzhou_urban
Guangzhou_suburban
Guangzhou_semi-rural
Hong Kong
Hok Tsui
Malaysia
Thailand
AMS
CAP
CHC
NAM
STP
Latitude
Longitude


Colorado 6, USA
Colorado 7, USA
Colorado 8, USA
Colorado 9, USA
Michigan 1, USA
Michigan 2, USA
Ohio 1, USA
Ohio 2, USA
Wiesbaden, Gemany
low-Arctic, Greenland
Urumqi, China
Mt. Tai, China
Latitude
Longitude


PDM
GRE
PDD
STN
Kiruna
Ljungbyhed
Mongolia
Colorado 1, USA
Colorado 2, USA
Colorado 3, USA
Colorado 4, USA
Colorado 5, USA
Latitude
Longitude


Central_China
Madrid, Spain
Singapore
Beijing 1, China
Beijing 2, China
Be'er Sheva, Israel
Gwangju, Korea
Incheon, Korea
Seoul, Korea
Taiyuan, China
South Ocean (19 sites)
Urban & Bulit-up
Bare ground
Cropland
Grassland
Open shrubland
Closed shrubland
Wooded grassland
Woodland
Mixed forest
Deciduous broadleaf forest
Deciduous needleleaf forest
Evergreen broadleaf forest
Evergreen needleleaf forest
Water/Sea

Appendix 3

This section provides information of metagenomic data to generate the global airborne dataset and track the source contribution to airborne ARGs.

1. Detailed information of air samples in the global airborne metagenomic dataset in this study.

No.	Name	Site	season	Lat	Lng	ALS	Temp.	Type	Size (GB)	Run name
1	AL1	Mt. Ailao, China	spring	24.53	101.02	2450	18.68	$PM_{2.5}$	6.51	
2	AL2	Mt. Ailao, China	summer	24.53	101.02	2450	21.8	$PM_{2.5}$	6.03	
3	AL3	Mt. Ailao, China	autumn	24.53	101.02	2450	16.41	$PM_{2.5}$	5.72	
4	AL4	Mt. Ailao, China	winter	24.53	101.02	2450	11.43	$PM_{2.5}$	6.54	
5	CB1	Mt. Changbai, China	spring	42.40	128.10	741	5.78	$PM_{2.5}$	6.47	
6	CB2	Mt. Changbai, China	summer	42.40	128.10	741	18.94	$PM_{2.5}$	5.81	
7	CB3	Mt. Changbai, China	autumn	42.40	128.10	741	5.16	$PM_{2.5}$	5.78	
8	CB4	Mt. Changbai, China	winter	42.40	128.10	741	-12.44	$PM_{2.5}$	6.48	
9	CH1	Guangzhou, China	spring	23.65	113.62	21	21.17	$PM_{2.5}$	6.55	
10	CH2	Guangzhou, China	summer	23.65	113.62	21	27.23	$PM_{2.5}$	5.64	
11	CH3	Guangzhou, China	autumn	23.65	113.62	21	22.65	$PM_{2.5}$	6.59	
12	CH4	Guangzhou, China	winter	23.65	113.62	21	15.24	$PM_{2.5}$	6.26	
13	CH5	Guangzhou, China	spring	23.65	113.62	21	20.71	$PM_{2.5}$	6.57	
14	HK1	Hong Kong	spring	22.31	114.18	87	23.67	$PM_{2.5}$	6.43	
15	HK2	Hong Kong	summer	22.31	114.18	87	29.6	$PM_{2.5}$	6.52	
16	HK3	Hong Kong	autumn	22.31	114.18	87	26.76	$PM_{2.5}$	6.56	
17	HK4	Hong Kong	winter	22.31	114.18	87	18.55	$PM_{2.5}$	6.33	
18	HR1	Beijing, China	spring	40.40	116.69	90	11.67	$PM_{2.5}$	6.38	
19	HR2	Beijing, China	summer	40.40	116.69	90	25.67	$PM_{2.5}$	5.75	
20	HR3	Beijing, China	autumn	40.40	116.69	90	9.54	$PM_{2.5}$	5.78	
21	HR4	Beijing, China	winter	40.40	116.69	90	-2.3	$PM_{2.5}$	6.35	
22	HS1	Guangzhou, China	spring	22.71	112.92	33	22.06	$PM_{2.5}$	6.46	
23	HS2	Guangzhou, China	summer	22.71	112.92	33	28.57	$PM_{2.5}$	5.75	
24	HS3	Guangzhou, China	autumn	22.71	112.92	33	24.59	$PM_{2.5}$	5.78	
25	HS4	Guangzhou, China	winter	22.71	112.92	33	16.55	$PM_{2.5}$	6.35	
26	HS5	Guangzhou, China	spring	22.71	112.92	33	21.19	$PM_{2.5}$	6.41	
27	HT1	Hong Kong	spring	22.12	114.15	50	19.87	$PM_{2.5}$	5.65	

28	HT2	Hong Kong	summer	22.12	114.15	50	28.47	$PM_{2.5}$	6.49	
29	HT3	Hong Kong	autumn	22.12	114.15	50	28.18	$PM_{2.5}$	6.08	
30	HT4	Hong Kong	winter	22.12	114.15	50	16.94	$PM_{2.5}$	6.44	
31	LS1	Nanjing, China	spring	31.33	119.01	15	20.3	$PM_{2.5}$	6.44	
32	LS2	Nanjing, China	summer	31.33	119.01	15	26.92	$PM_{2.5}$	5.56	
33	LS3	Nanjing, China	autumn	31.33	119.01	15	12.16	$PM_{2.5}$	6.18	
34	LS4	Nanjing, China	winter	31.33	119.01	15	7.17	$PM_{2.5}$	6.56	
35	LS5	Nanjing, China	spring	31.33	119.01	15	21.09	$PM_{2.5}$	6.63	
36	M1	Bachok, Malaysia	spring	6.01	102.43	10	29.21	TSP	4.42	
37	M2	Bachok, Malaysia	summer	6.01	102.43	10	28.21	TSP	6.39	
38	M3	Bachok, Malaysia	autumn	6.01	102.43	10	27.65	TSP	6.48	
39	M4	Bachok, Malaysia	winter	6.01	102.43	10	27.49	TSP	6.55	
40	PK1	Nanjing, China	spring	32.12	118.42	51	16.25	$PM_{2.5}$	6.45	
41	PK2	Nanjing, China	summer	32.12	118.42	51	27.05	$PM_{2.5}$	7.2	
42	PK3	Nanjing, China	autumn	32.12	118.42	51	17.76	$PM_{2.5}$	6.86	
43	PK4	Nanjing, China	winter	32.12	118.42	51	6.43	$PM_{2.5}$	6.41	
44	PK5	Nanjing, China	spring	32.12	118.42	51	16.76	$PM_{2.5}$	6.43	
45	PKU1	Beijing, China	spring	39.99	116.36	56	3.23	$PM_{2.5}$	6.21	
46	PKU2	Beijing, China	summer	39.99	116.36	56	23	$PM_{2.5}$	6.06	
47	PKU3	Beijing, China	autumn	39.99	116.36	56	25.74	$PM_{2.5}$	6.57	
48	PKU4	Beijing, China	winter	39.99	116.36	56	0.45	$PM_{2.5}$	6.43	
49	PKU5	Beijing, China	spring	39.99	116.36	56	0.4	$PM_{2.5}$	6.59	
50	T1	Phitsanulok, Thailand	spring	16.75	100.20	47	32.33	TSP	5.82	
51	T2	Phitsanulok, Thailand	summer	16.75	100.20	47	29.06	TSP	5.41	
52	T3	Phitsanulok, Thailand	autumn	16.75	100.20	47	27.88	TSP	6.42	
53	T4	Phitsanulok, Thailand	winter	16.75	100.20	47	26.2	TSP	6.47	
54	TH1	Guangzhou, China	spring	23.15	113.36	36	21.88	$PM_{2.5}$	6.43	
55	TH2	Guangzhou, China	summer	23.15	113.36	36	22.29	$PM_{2.5}$	5.41	
56	TH3	Guangzhou, China	autumn	23.15	113.36	36	20.26	$PM_{2.5}$	6.42	
57	TH4	Guangzhou, China	winter	23.15	113.36	36	15.54	$PM_{2.5}$	6.47	
58	TH5	Guangzhou, China	spring	23.15	113.36	36	21.6	$PM_{2.5}$	6.22	
59	WL1	Waliguan, China	spring	36.29	100.90	3816	7.12	$PM_{2.5}$	8.48	
60	WL2	Waliguan, China	summer	36.29	100.90	3816	15.41	$PM_{2.5}$	6.92	
61	WL3	Waliguan, China	autumn	36.29	100.90	3816	11.3	$PM_{2.5}$	6.86	
62	WL4	Waliguan, China	winter	36.29	100.90	3816	-4.16	$PM_{2.5}$	6.88	
63	XW1	Nanjing, China	spring	32.06	118.80	30	20.54	$PM_{2.5}$	6.37	
64	XW2	Nanjing, China	summer	32.06	118.80	30	27.36	$PM_{2.5}$	6.92	
65	XW3	Nanjing, China	autumn	32.06	118.80	30	12.31	$PM_{2.5}$	6.86	
66	XW4	Nanjing, China	winter	32.06	118.80	30	7.61	$PM_{2.5}$	6.88	
67	XW5	Nanjing, China	spring	32.06	118.80	30	21.69	$PM_{2.5}$	5.61	
68	ZF1	Mount Everest, China	spring	28.21	86.56	4276	5.1	$PM_{2.5}$	6.24	
69	ZF2	Mount Everest, China	summer	28.21	86.56	4276	10.78	$PM_{2.5}$	4.87	
70	ZF3	Mount Everest, China	autumn	28.21	86.56	4276	7.13	$PM_{2.5}$	6.93	

71	ZF4	Mount Everest, China	winter	28.21	86.56	4276	-2.42	$PM_{2.5}$	5.61	
72	Antarctica1B	Antarctica		-77.49	161.81	134		TSP	8.26	SRR13553597
73	Antarctica2A	Antarctica		-77.52	161.82	882		TSP	6.39	SRR13553599
74	Antarctica3A	Antarctica		-77.52	161.82	882		TSP	5.26	SRR13553600
75	Antarctica4A	Antarctica		-77.52	161.82	882		TSP	6.74	SRR13553601
76	Canada1	Cambridge Bay, Canada		69.13	-105.06	6		TSP	1.31	SRR13553518
77	Canada2	Cambridge Bay, Canada		69.13	-105.06	6		TSP	2.44	SRR13553519
78	Canada3	Cambridge Bay, Canada		69.13	-105.06	6		TSP	1.35	SRR13553520
79	Canada4	Cambridge Bay, Canada		69.13	-105.06	6		TSP	3.50	SRR13553521
80	Canada5	Cambridge Bay, Canada		69.13	-105.06	6		TSP	2.69	SRR13553522
81	Canada6	Cambridge Bay, Canada		69.13	-105.06	6		TSP	3.25	SRR13553523
82	Canada7	Cambridge Bay, Canada		69.13	-105.06	6		TSP	4.07	SRR13553524
83	Chile1A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	8.61	SRR13553551
84	Chile1B	Copiapo, Chile		-27.29	-70.25	1053		TSP	8.05	SRR13553549
85	Chile2A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	4.36	SRR13553552
86	Chile3A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	5.09	SRR13553553
87	Chile4A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	3.63	SRR13553554
88	Chile5A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	5.81	SRR13553555
89	Chile6A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	2.22	SRR13553556
90	Chile7A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	8.72	SRR13553588
91	Chile8A	Antofagasta, Chile		-24.10	-70.02	1063		TSP	5.98	SRR13553589
92	Mongolia2	Zagiin us, Mongolia		44.57	105.65	1238		TSP	1.41	SRR13553489
93	Mongolia3	Zagiin us, Mongolia		44.57	105.65	1238		TSP	5.27	SRR13553490
94	Mongolia4	Zagiin us, Mongolia		44.57	105.65	1238		TSP	1.01	SRR13553491
95	Mongolia5	Zagiin us, Mongolia		44.57	105.65	1238		TSP	4.84	SRR13553492
96	Mongolia6	Zagiin us, Mongolia		44.57	105.65	1238		TSP	1.95	SRR13553493
97	Mongolia7	Zagiin us, Mongolia		44.57	105.65	1238		TSP	3.68	SRR13553494
98	Mongolia8	Zagiin us, Mongolia		44.57	105.65	1238		TSP	7.37	SRR13553496
99	Mongolia9	Zagiin us, Mongolia		44.57	105.65	1238		TSP	5.19	SRR13553497
100	Namibia1	desert, Namibia		-23.60	15.04	455		TSP	8.79	SRR13553567
101	Namibia10	desert, Namibia		-23.60	15.04	455		TSP	3.22	SRR13553577
102	Namibia2	desert, Namibia		-23.60	15.04	455		TSP	9.01	SRR13553568
103	Namibia3	desert, Namibia		-23.60	15.04	455		TSP	6.52	SRR13553569
104	Namibia4	desert, Namibia		-23.60	15.04	455		TSP	6.16	SRR13553570
105	Namibia5	desert, Namibia		-23.60	15.04	455		TSP	1.67	SRR13553571
106	Namibia6	desert, Namibia		-23.60	15.04	455		TSP	3.76	SRR13553573
107	Namibia7	desert, Namibia		-23.60	15.04	455		TSP	6.04	SRR13553574
108	Namibia8	desert, Namibia		-23.60	15.04	455		TSP	5.01	SRR13553575
109	Namibia9	desert, Namibia		-23.60	15.04	455		TSP	3.03	SRR13553576
110	Spain1A	Madrid, Spain		40.82	-3.96	1814		TSP	1.75	SRR13553508
111	Spain1B	Toledo, Spain		39.62	-4.13	816		TSP	10.05	SRR13553510
112	Spain2A	Madrid, Spain		40.82	-3.96	1814		TSP	2.16	SRR13553509
113	Spain3A	Madrid, Spain		40.82	-3.96	1814		TSP	7.46	SRR13553511

114	Spain4A	Madrid, Spain		40.82	-3.96	1814		TSP	4.73	SRR13553512
115	Spain5A	Madrid, Spain		40.82	-3.96	1814		TSP	4.23	SRR13553513
116	Spain6A	Madrid, Spain		40.82	-3.96	1814		TSP	6.79	SRR13553514
117	Spain7A	Madrid, Spain		40.82	-3.96	1814		TSP	6.46	SRR13553515
118	Spain8A	Madrid, Spain		40.82	-3.96	1814		TSP	5.93	SRR13553516
119	USA1A	California, USA		35.14	-116.10	284		TSP	5.00	SRR13553525
120	USA1B	Nevada, USA		36.26	-115.51	1672		TSP	4.26	SRR13553529
121	USA2A	California, USA		35.14	-116.10	284		TSP	7.45	SRR13553526
122	USA3A	California, USA		35.14	-116.10	284		TSP	7.85	SRR13553527
123	USA4A	California, USA		35.14	-116.10	284		TSP	3.25	SRR13553530
124	USA5A	California, USA		35.14	-116.10	284		TSP	4.55	SRR13553531
125	USA6A	California, USA		35.14	-116.10	284		TSP	2.67	SRR13553532
126	USA7A	California, USA		35.14	-116.10	284		TSP	5.61	SRR13553533
127	USA8A	California, USA		35.14	-116.10	284		TSP	7.66	SRR13553534
128	Japan1	Ishikawa, Japan		37.31	137.23	6		TSP	6.11	SRR13553486
129	Japan2	Ishikawa, Japan		37.31	137.23	6		TSP	7.51	SRR13553487
130	Kuwait10	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	5.10	SRR13553585
131	Kuwait11	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	5.34	SRR13553586
132	Kuwait12	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	4.46	SRR13553587
133	Kuwait4	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	4.23	SRR13553501
134	Kuwait5	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	2.49	SRR13553502
135	Kuwait6	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	3.50	SRR13553503
136	Kuwait7	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	3.34	SRR13553504
137	Kuwait8	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	3.17	SRR13553505
138	Kuwait9	Shalayhat Mina Abdullah, Kuwait		28.95	48.19	9		TSP	1.79	SRR13553507
139	Singapore1	Singapore		1.31	103.77	18		TSP	6.34	SRR13553485
140	Singapore2	Singapore		1.31	103.77	18		TSP	6.24	SRR13553603
141	Singapore3	Singapore		1.31	103.77	18		TSP	7.36	SRR13553604
142	Uruguay10	Colonia, Uruguay		-34.35	-57.24	30		TSP	4.10	SRR13553581
143	Uruguay11	Colonia, Uruguay		-34.35	-57.24	30		TSP	6.54	SRR13553582
144	Uruguay12	Colonia, Uruguay		-34.35	-57.24	30		TSP	7.20	SRR13553584
145	Uruguay4	Colonia, Uruguay		-34.35	-57.24	30		TSP	7.75	SRR13553560
146	Uruguay5	Colonia, Uruguay		-34.35	-57.24	30		TSP	9.29	SRR13553562
147	Uruguay6	Colonia, Uruguay		-34.35	-57.24	30		TSP	8.68	SRR13553563

148	Uruguay7	Colonia, Uruguay		-34.35	-57.24	30		TSP	5.11	SRR13553564
149	Uruguay8	Colonia, Uruguay		-34.35	-57.24	30		TSP	9.98	SRR13553565
150	Uruguay9	Colonia, Uruguay		-34.35	-57.24	30		TSP	5.67	SRR13553566
151	Africa1	Pretoria, South Africa		-25.73	28.26	1317		TSP	7.72	SRR13553495
152	Africa2	Pretoria, South Africa		-25.73	28.26	1317		TSP	5.59	SRR13553578
153	Africa3	Pretoria, South Africa		-25.73	28.26	1317		TSP	4.21	SRR13553579
154	Africa4	Pretoria, South Africa		-25.73	28.26	1317		TSP	3.40	SRR13553583
155	Africa5	Pretoria, South Africa		-25.73	28.26	1317		TSP	5.40	SRR13553594
156	Africa6	Pretoria, South Africa		-25.73	28.26	1317		TSP	4.38	SRR13553605
157	BJ001	Beijing, China		40.00	116.32	60	18	PM_{10}	3.43	SRR10613586
158	BJ002	Beijing, China		40.00	116.32	60	18	$PM_{2.5}$	2.50	SRR10613585
159	BJ003	Beijing, China		40.00	116.32	60	17	PM_{10}	5.62	SRR10613568
160	BJ004	Beijing, China		40.00	116.32	60	17	$PM_{2.5}$	2.88	SRR10613557
161	BJ005	Beijing, China		40.00	116.32	60	16	PM_{10}	6.37	SRR10613508
162	BJ006	Beijing, China		40.00	116.32	60	16	$PM_{2.5}$	6.13	SRR10613497
163	BJ007	Beijing, China		40.00	116.32	60	9	PM_{10}	3.33	SRR10613486
164	BJ008	Beijing, China		40.00	116.32	60	9	$PM_{2.5}$	2.76	SRR10613545
165	BJ009	Beijing, China		40.00	116.32	60	12	PM_{10}	5.48	SRR10613534
166	BJ010	Beijing, China		40.00	116.32	60	12	$PM_{2.5}$	5.12	SRR10613523
167	BJ011	Beijing, China		40.00	116.32	60	9	PM_{10}	2.90	SRR10613584
168	BJ012	Beijing, China		40.00	116.32	60	9	$PM_{2.5}$	3.38	SRR10613577
169	BJ013	Beijing, China		40.00	116.32	60	8	PM_{10}	3.13	SRR10613576
170	BJ014	Beijing, China		40.00	116.32	60	8	$PM_{2.5}$	2.90	SRR10613575
171	BJ015	Beijing, China		40.00	116.32	60	6	PM_{10}	6.05	SRR10613574
172	BJ016	Beijing, China		40.00	116.32	60	6	$PM_{2.5}$	2.54	SRR10613573
173	BJ017	Beijing, China		40.00	116.32	60	7	PM_{10}	3.22	SRR10613572
174	BJ018	Beijing, China		40.00	116.32	60	7	$PM_{2.5}$	3.29	SRR10613571
175	BJ019	Beijing, China		40.00	116.32	60	6	PM_{10}	2.97	SRR10613570
176	BJ020	Beijing, China		40.00	116.32	60	6	$PM_{2.5}$	3.15	SRR10613569
177	BJ021	Beijing, China		40.00	116.32	60	4	PM_{10}	3.18	SRR10613567
178	BJ022	Beijing, China		40.00	116.32	60	4	$PM_{2.5}$	5.36	SRR10613566
179	BJ023	Beijing, China		40.00	116.32	60	4	PM_{10}	13.85	SRR10613565
180	BJ024	Beijing, China		40.00	116.32	60	4	$PM_{2.5}$	3.52	SRR10613564
181	BJ025	Beijing, China		40.00	116.32	60	5	PM_{10}	3.42	SRR10613563
182	BJ026	Beijing, China		40.00	116.32	60	5	$PM_{2.5}$	3.16	SRR10613562
183	BJ027	Beijing, China		40.00	116.32	60	3	PM_{10}	2.73	SRR10613561
184	BJ028	Beijing, China		40.00	116.32	60	3	$PM_{2.5}$	5.17	SRR10613560
185	BJ029	Beijing, China		40.00	116.32	60	2	PM_{10}	6.54	SRR10613559
186	BJ030	Beijing, China		40.00	116.32	60	2	$PM_{2.5}$	2.00	SRR10613558
187	BJ031	Beijing, China		40.00	116.32	60	-7	PM_{10}	4.31	SRR10613556
188	BJ032	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	3.19	SRR10613555
189	BJ033	Beijing, China		40.00	116.32	60	-6	PM_{10}	2.79	SRR10613554
190	BJ034	Beijing, China		40.00	116.32	60	-6	$PM_{2.5}$	2.79	SRR10613553

191	BJ035	Beijing, China		40.00	116.32	60	-1	PM_{10}	4.93	SRR10613552
192	BJ036	Beijing, China		40.00	116.32	60	-1	$PM_{2.5}$	3.85	SRR10613551
193	BJ037	Beijing, China		40.00	116.32	60	-1	PM_{10}	3.26	SRR10613512
194	BJ038	Beijing, China		40.00	116.32	60	-1	$PM_{2.5}$	3.34	SRR10613511
195	BJ039	Beijing, China		40.00	116.32	60	-3	PM_{10}	2.61	SRR10613510
196	BJ040	Beijing, China		40.00	116.32	60	-3	$PM_{2.5}$	5.57	SRR10613509
197	BJ041	Beijing, China		40.00	116.32	60	-4	PM_{10}	3.45	SRR10613507
198	BJ042	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	3.03	SRR10613506
199	BJ043	Beijing, China		40.00	116.32	60	-11	PM_{10}	3.36	SRR10613505
200	BJ044	Beijing, China		40.00	116.32	60	-11	$PM_{2.5}$	6.22	SRR10613504
201	BJ045	Beijing, China		40.00	116.32	60	-7	PM_{10}	2.62	SRR10613503
202	BJ046	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	3.24	SRR10613502
203	BJ047	Beijing, China		40.00	116.32	60	-10	PM_{10}	2.72	SRR10613501
204	BJ048	Beijing, China		40.00	116.32	60	-10	$PM_{2.5}$	4.33	SRR10613500
205	BJ049	Beijing, China		40.00	116.32	60	-7	PM_{10}	3.03	SRR10613499
206	BJ050	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	7.04	SRR10613498
207	BJ051	Beijing, China		40.00	116.32	60	-7	PM_{10}	5.13	SRR10613496
208	BJ052	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	3.06	SRR10613495
209	BJ053	Beijing, China		40.00	116.32	60	-7	PM_{10}	3.18	SRR10613494
210	BJ054	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	2.65	SRR10613493
211	BJ055	Beijing, China		40.00	116.32	60	-7	PM_{10}	3.32	SRR10613492
212	BJ056	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	9.01	SRR10613491
213	BJ057	Beijing, China		40.00	116.32	60	-9	PM_{10}	5.43	SRR10613490
214	BJ058	Beijing, China		40.00	116.32	60	-9	$PM_{2.5}$	5.67	SRR10613489
215	BJ059	Beijing, China		40.00	116.32	60	-4	PM_{10}	3.21	SRR10613488
216	BJ060	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	5.67	SRR10613487
217	BJ061	Beijing, China		40.00	116.32	60	-7	PM_{10}	5.08	SRR10613485
218	BJ062	Beijing, China		40.00	116.32	60	-7	$PM_{2.5}$	5.71	SRR10613484
219	BJ063	Beijing, China		40.00	116.32	60	-4	PM_{10}	3.84	SRR10613483
220	BJ064	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	3.45	SRR10613482
221	BJ065	Beijing, China		40.00	116.32	60	-6	PM_{10}	4.38	SRR10613481
222	BJ066	Beijing, China		40.00	116.32	60	-6	$PM_{2.5}$	3.38	SRR10613550
223	BJ067	Beijing, China		40.00	116.32	60	-4	PM_{10}	3.29	SRR10613549
224	BJ068	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	2.96	SRR10613548
225	BJ069	Beijing, China		40.00	116.32	60	-3	PM_{10}	9.21	SRR10613547
226	BJ070	Beijing, China		40.00	116.32	60	-3	$PM_{2.5}$	4.76	SRR10613546
227	BJ071	Beijing, China		40.00	116.32	60	-6	PM_{10}	6.88	SRR10613544
228	BJ072	Beijing, China		40.00	116.32	60	-6	$PM_{2.5}$	7.02	SRR10613543
229	BJ073	Beijing, China		40.00	116.32	60	-5	PM_{10}	2.77	SRR10613542
230	BJ074	Beijing, China		40.00	116.32	60	-5	$PM_{2.5}$	3.30	SRR10613541
231	BJ075	Beijing, China		40.00	116.32	60	-1	PM_{10}	3.34	SRR10613540
232	BJ076	Beijing, China		40.00	116.32	60	-1	$PM_{2.5}$	3.18	SRR10613539
233	BJ077	Beijing, China		40.00	116.32	60	-5	PM_{10}	3.54	SRR10613538

234	BJ078	Beijing, China		40.00	116.32	60	-5	$PM_{2.5}$	5.00	SRR10613537
235	BJ079	Beijing, China		40.00	116.32	60	-1	PM_{10}	2.88	SRR10613536
236	BJ080	Beijing, China		40.00	116.32	60	-1	$PM_{2.5}$	3.16	SRR10613535
237	BJ081	Beijing, China		40.00	116.32	60	-2	PM_{10}	2.88	SRR10613533
238	BJ082	Beijing, China		40.00	116.32	60	-2	$PM_{2.5}$	2.84	SRR10613532
239	BJ083	Beijing, China		40.00	116.32	60	-4	PM_{10}	5.52	SRR10613531
240	BJ084	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	2.77	SRR10613530
241	BJ085	Beijing, China		40.00	116.32	60	-6	PM_{10}	3.00	SRR10613529
242	BJ086	Beijing, China		40.00	116.32	60	-6	$PM_{2.5}$	3.92	SRR10613528
243	BJ087	Beijing, China		40.00	116.32	60	-1	PM_{10}	11.55	SRR10613527
244	BJ088	Beijing, China		40.00	116.32	60	-1	$PM_{2.5}$	10.63	SRR10613526
245	BJ089	Beijing, China		40.00	116.32	60	-4	PM_{10}	7.16	SRR10613525
246	BJ090	Beijing, China		40.00	116.32	60	-4	$PM_{2.5}$	6.33	SRR10613524
247	BJ091	Beijing, China		40.00	116.32	60	0	PM_{10}	2.60	SRR10613522
248	BJ092	Beijing, China		40.00	116.32	60	0	$PM_{2.5}$	2.75	SRR10613521
249	BJ093	Beijing, China		40.00	116.32	60	1	PM_{10}	3.34	SRR10613520
250	BJ094	Beijing, China		40.00	116.32	60	1	$PM_{2.5}$	3.11	SRR10613519
251	BJ095	Beijing, China		40.00	116.32	60	7	PM_{10}	4.96	SRR10613518
252	BJ096	Beijing, China		40.00	116.32	60	7	$PM_{2.5}$	5.98	SRR10613517
253	BJ097	Beijing, China		40.00	116.32	60	4	PM_{10}	1.33	SRR10613516
254	BJ098	Beijing, China		40.00	116.32	60	4	$PM_{2.5}$	3.74	SRR10613515
255	BJ099	Beijing, China		40.00	116.32	60	8	PM_{10}	3.43	SRR10613514
256	BJ100	Beijing, China		40.00	116.32	60	8	$PM_{2.5}$	3.26	SRR10613513
257	BJ101	Beijing, China		40.00	116.32	60	6	PM_{10}	3.43	SRR10613583
258	BJ102	Beijing, China		40.00	116.32	60	6	$PM_{2.5}$	3.37	SRR10613582
259	BJ103	Beijing, China		40.00	116.32	60	7	PM_{10}	2.57	SRR10613581
260	BJ104	Beijing, China		40.00	116.32	60	7	$PM_{2.5}$	2.90	SRR10613580
261	BJ105	Beijing, China		40.00	116.32	60	6	PM_{10}	2.58	SRR10613579
262	BJ106	Beijing, China		40.00	116.32	60	6	$PM_{2.5}$	3.20	SRR10613578

2. Database of putative sources of airborne ARGs.

type	Run name
drinking water (*n=7*)	SRR6797136, SRR6797141, SRR6797149, SRR6797150, SRR6797151, SRR6986810, SRR6986811
farm-related (*n=24*)	ERR1135410, ERR1135409, ERR1135406, ERR1135408, ERR1135417, ERR1135419, ERR1135443, ERR1135472, ERR2027889, ERR2027890, ERR2530127, ERR2530126, ERR1135281, ERR1135282, ERR1135178, ERR1135180, ERR1135179, ERR1193301, ERR1193332, ERR1193300, ERR1193331, ERR1193299, ERR1193298, ERR1193297
freshwater (*n=75*)	ERR1193292, ERR1193293, ERR1193294, ERR1725854, ERR1726436, ERR1726985, ERR1726987, ERR1726993, SRR5431081, SRR5431137, SRR5431138, SRR5431152, SRR5433937, SRR5468423, SRR5468427, SRR5468431, SRR5468433, SRR5468437, SRR5535881, SRR5535882, SRR5747948, SRR6481338, SRR6481359, SRR6481365, SRR6481371, SRR6481372, SRR7760390, SRR7760391, SRR7760393, SRR7760396, SRR1047946, SRR1047948, SRR1047951, SRR1047952, SRR1047954, SRR3321505, SRR3321803, SRR3322106, SRR3322355, SRR14576896, SRR14576897, SRR14576898, SRR14576899, SRR14576900, SRR14576901, SRR14576902, SRR14576903, SRR14576904, SRR14576905, SRR14576906, SRR14576907, SRR14576908, SRR14576909, SRR14576910, SRR14576911, SRR14576912, SRR14576913, SRR14576914, SRR14576915, SRR14576916, SRR14576917, SRR14576918, SRR14576919, SRR14576920, SRR14576921, SRR14576922, SRR14576923, SRR14576924, SRR14576925, SRR14576926, SRR14576927, SRR14576928, SRR14576929, SRR14576930, SRR14576931
forest soil (*n=27*)	ERR1877665, ERR1877670, ERR1877678, ERR1877680, ERR1877684, ERR1877686, ERR1877689, ERR1877694, ERR1877701, ERR1877702, ERR1877704, ERR1877718, ERR1877723, ERR1877735, ERR1877758, ERR1877770, ERR1877785, ERR1877849, ERR1877852, ERR1877877, ERR1877880, ERR1877882, ERR1877884, ERR1877887, ERR1877890, ERR1877906, ERR1877911
agricultural soil (*n=10*)	SRR1190306, SRR1190308, SRR1190311, SRR1190316, SRR1190334, SRR1190336, SRR1190349, SRR1190350, SRR1190383, SRR1190384
Antarctica soil (*n=6*)	ERR5891570, ERR5891571, ERR5891572, ERR5891573, ERR5891574, ERR5891575
grassland soil (*n=3*)	ERR1877796, ERR1877921, ERR1877926
hospital-related (*n=25*)	ERR1191817, ERR1191818, ERR1191819, ERR1191820, ERR1191821, SRR14413839, SRR14413840, SRR14413841, SRR14413842, SRR14413843, SRR14413844, SRR14413845, SRR14413846, SRR14413847, SRR14413848, SRR14413849, SRR14413850, SRR14413851, SRR14413852, SRR14413853, SRR14413854, SRR14413855, SRR14413856, SRR14413857, ERR1191822
human surface (*n=29*)	SRR11242339, SRR11242340, SRR11242342, SRR11242343, SRR11242344, SRR11242345, SRR11242346, SRR11242347, SRR11242348, SRR11242349, SRR11242350, SRR11242351, SRR11242353, SRR11242354, SRR11242355, SRR11242356, SRR11300707, SRR11300708, SRR11300709, SRR11300710, SRR11300711, SRR11300712, SRR11300713, SRR11300714, SRR11300715, SRR11300716, SRR11300717, SRR11300718, SRR11304755

type	Run name
human excretion (*n=241*)	SRR059365, SRR060439, SRR061513, SRR061293, SRR059475, SRR061373, SRR061392, SRR513156, SRR061387, SRR513145, SRR059411, SRR063524, SRR346706, SRR059913, SRR061525, SRR512078, SRR513150, SRR063466, SRR513146, SRR059877, SRR346676, SRR061914, SRR513181, SRR513794, SRR059996, SRR513764, SRR060429, SRR059859, SRR061592, SRR514266, SRR059401, SRR346690, SRR063490, SRR063492, SRR061530, SRR353629, SRR061348, SRR062358, SRR353628, SRR059377, SRR062444, SRR061370, SRR514232, SRR061480, SRR062536, SRR062289, SRR061365, SRR062332, SRR062101, SRR059957, SRR061979, SRR059493, SRR059994, SRR059814, SRR061968, SRR063507, SRR061916, SRR062042, SRR062056, SRR063528, SRR059428, SRR061335, SRR062014, SRR062300, SRR059490, SRR059446, SRR061302, SRR059371, SRR059484, SRR059340, SRR059899, SRR062347, SRR061354, SRR059414, SRR059363, SRR062527, SRR059865, SRR061377, SRR061520, SRR061159, SRR061585, SRR059328, SRR061590, SRR062410, SRR062505, SRR059403, SRR062391, SRR059433, SRR059510, SRR061566, SRR062377, SRR062432, SRR061276, SRR060081, SRR062452, SRR061239, SRR061953, SRR061997, SRR060121, SRR062068, SRR063928, SRR063675, SRR061311, SRR061210, SRR061178, SRR061400, SRR059893, SRR061259, SRR061319, SRR059437, SRR061244, SRR062325, SRR060378, SRR061389, SRR059919, SRR060033, SRR060361, SRR061289, SRR063496, SRR059374, SRR059928, SRR062468, SRR061141, SRR062491, SRR060421, SRR060440, SRR060142, SRR062311, SRR060387, SRR060405, SRR061511, SRR059839, SRR062281, SRR063746, SRR061945, SRR059358, SRR061550, SRR061252, SRR061441, SRR061451, SRR062034, SRR061225, SRR060367, SRR062093, SRR063519, SRR064436, SRR18491211, ERR527046, ERR527053, ERR527063, SRR18490939, SRR18490983, SRR18490984, SRR18490986, SRR18490988, SRR18490991, SRR18490996, SRR18490997, SRR18490998, SRR18490999, SRR18491000, SRR18491001, SRR18491003, SRR18491004, SRR18491008, SRR18491028, SRR18491037, SRR18491039, SRR18491042, SRR18491043, SRR18491045, SRR18491046, SRR18491048, SRR18491051, SRR18491052, SRR18491059, SRR18491061, SRR18491067, SRR18491078, SRR18491093, SRR18491095, SRR18491100, SRR18491105, SRR18491109, SRR18491110, SRR18491112, SRR18491116, SRR18491118, SRR18491120, SRR18491121, SRR18491123, SRR18491125, SRR18491136, SRR18491157, SRR18491168, SRR18491177, SRR18491213, SRR18491217, SRR18491218, SRR18491219, SRR18491220, SRR18491221, SRR18491222, SRR18491224, SRR18491226, SRR18491231, SRR18491232, SRR18491236, SRR18491240, SRR18491241, SRR18491243, SRR18491244, SRR18491246, SRR18491250, SRR18491251, SRR18491253, SRR18491255, SRR18491259, SRR18491260, SRR18491261, SRR18491263, SRR18491272, SRR18491274, SRR18491277, SRR18491278, SRR18491286, SRR18491287, SRR18491288, SRR18491290, SRR18491292, SRR18491293, SRR18491294, SRR18491295, SRR18491306, SRR18491328, ERR527048, ERR527047, ERR526291, ERR527065, ERR527050, ERR527064
landfill-related (*n=34*)	SRR14102348, SRR14102349, SRR14102350, SRR14102351, SRR14102352, SRR14102353, SRR6301222, SRR6301223, SRR6301224, SRR11702754, SRR11702755, SRR11702756, SRR11702757, SRR11702758, SRR11702759, SRR11702760, SRR11702761, SRR11702762, SRR11702763, SRR11702764, SRR11702765, SRR11702766, SRR11702767, SRR11702768, SRR11702769, SRR11702770, SRR11702771, SRR10498317, SRR10498318, SRR10498319, SRR10498320, SRR10498321, SRR10498322, SRR10498323
ocean (*n=67*)	ERR598943, ERR598954, ERR598966, ERR598967, ERR598970, ERR598978, ERR598979, ERR598989, ERR598992, ERR598997, ERR599011, ERR599012, ERR599019, ERR599024, ERR599029, ERR599030, ERR599036, ERR599038, ERR599039, ERR599045, ERR599050, ERR599052, ERR599054, ERR599057, ERR599058, ERR599063, ERR599064, ERR599066, ERR599069, ERR599074, ERR599075, ERR599077, ERR599080, ERR599088, ERR599091, ERR599093, ERR599098, ERR599102, ERR599114, ERR599118, ERR599119, ERR599120, ERR599138, ERR599139, ERR599141, ERR599142, ERR599143, ERR599146, ERR599150, ERR599151, ERR599158, ERR599160, ERR599162, ERR599163, ERR599169, SRR2134631, SRR2134632, SRR2134633, SRR2134634, SRR2134636, SRR2134637, SRR2134639, SRR2134640, SRR2134641, SRR2134642, SRR2134643, SRR2134644
plant (*n=24*)	SRR10585495, SRR10585497, SRR10585498, SRR10585499, SRR10585507, SRR10585513, SRR10585514, SRR10585515, SRR10585516, SRR10585518, SRR10585527, SRR10585528, SRR10585529, SRR10585530, SRR10585531, SRR10585540, SRR10585541, SRR10585542, SRR10585543, SRR10585544, SRR10585558, SRR10585559, SRR10585560, SRR10585561

type	Run name
WWTPs-related (*n=122*)	ERR1414209, ERR1414277, ERR1414276, ERR1414278, ERR1414273, ERR1414275, ERR1414272, ERR1414271, ERR1414270, ERR1414211, ERR1414210, ERR1414225, ERR1414269, ERR1414267, ERR1414237, ERR1414224, ERR1414223, ERR1414268, ERR1414248, ERR1414247, ERR1414242, ERR1414253, ERR1414243, ERR1414238, ERR1414241, ERR1414251, ERR1414254, ERR1414252, ERR1414249, ERR1414221, ERR1414218, ERR1414244, ERR1414245, ERR1414222, ERR1414255, ERR1414240, ERR1414219, ERR1414239, ERR1414220, ERR1414256, ERR1414217, ERR1414250, ERR1414216, ERR1414274, ERR1414236, ERR1414232, ERR1414229, ERR1414234, ERR1414235, ERR1414215, ERR1414233, SRR14932484, ERR1414230, ERR1414226, ERR1414231, ERR1414214, ERR1414213, ERR1414228, ERR1414227, ERR1414212, ERR1414262, SRR14932479, ERR1414265, ERR1414257, SRR14932480, ERR1414261, ERR1414264, ERR1414263, ERR1414266, ERR1414260, ERR1414258, SRR14932686, SRR14932504, SRR14932506, SRR14932505, ERR1414259, SRR14932668, SRR14932684, SRR14932685, SRR14932478, SRR14932557, SRR14932565, SRR14932509, SRR14932503, SRR14932677, SRR14932702, SRR14932527, SRR14932477, SRR14932562, SRR14932508, SRR14932558, SRR14932563, SRR14932526, SRR14932493, SRR14932564, SRR14932510, SRR14932669, SRR14932683, SRR14932567, SRR14932561, SRR14932687, SRR14932507, SRR14932487, SRR14932566, SRR14932689, SRR14932482, SRR14932485, SRR14932559, SRR14932568, SRR14932486, SRR14932488, SRR14932483, SRR5997552, SRR5997541, SRR5997540, SRR5997551, SRR5423071, SRR5423072, SRR5422913, SRR5422949, SRR5422860, SRR1047949

3. Database of potential sources of airborne ARGs in Hong Kong for case study.

Type	Run name
WWTPs (*n=7*)	SRR13503351, SRR13503352, SRR13503358, SRR13503359, SRR13503360, SRR13503361, SRR13503362
transportation-related (*n= 21*)	SRR5312474, SRR5312475, SRR5312476, SRR6145059, SRR6145060, SRR6145062, SRR6145063, SRR6145071, SRR6145072, SRR6145074, SRR6145075, SRR6145077, SRR6145080, SRR6145081, SRR6145082, SRR6145083, SRR6145084, SRR6145088, SRR6145091, SRR6145096, SRR6145101
hospital-related (*n=12*)	SRR5312478, SRR5312479, SRR5312480, SRR5312481, SRR5312482, SRR5312483, SRR5312484, SRR5312485, SRR5312486, SRR5312487, SRR5312488, SRR5312489
farm-related (*n=2*)	SRR1022348, SRR1022349
marine (*n=13*)	SRR1022373, SRR1022377, SRR2134631, SRR2134632, SRR2134633, SRR2134634, SRR2134636, SRR2134637, SRR2134639, SRR2134640, SRR2134641, SRR2134642, SRR2134643
Freshwater (*n=7*)	SRR1022353, SRR1022378, SRR1047946, SRR1047948, SRR1047951, SRR1047952, SRR1047954
tap water (*n=1*)	SRR2134644
human excretion (*n=8*)	SRR062300, SRR062311, SRR062325, SRR062332, SRR062347, SRR062377, SRR062391, SRR062432

Appendix 4

This section provides the initial results of bioinformatic analyses in the thesis.

1. Annotation and relative abundance (%) calculation of global airborne ARGs.

ARG	ARG class	abundance
AAC(2')-I	aminoglycoside antibiotic	4.86E-04
AAC(2')-IE	aminoglycoside antibiotic	9.88E-06
AAC(2')-IIB	aminoglycoside antibiotic	1.01E-05
AAC(3)-I	aminoglycoside antibiotic	3.38E-03
AAC(3)-IB	aminoglycoside antibiotic	3.15E-04
AAC(3)-II	aminoglycoside antibiotic	1.68E-04
AAC(3)-IIIA	aminoglycoside antibiotic	1.05E-03
AAC(3)-IV	aminoglycoside antibiotic	1.18E-03
AAC(3)-IX	aminoglycoside antibiotic	5.22E-06
AAC(3)-VI	aminoglycoside antibiotic	1.34E-05
AAC(3)-VII	aminoglycoside antibiotic	2.54E-05
AAC(3)-VIII	aminoglycoside antibiotic	2.71E-05
AAC(3)-X	aminoglycoside antibiotic	2.25E-05
AAC(6')-31	aminoglycoside antibiotic	1.89E-04
AAC(6')-34	aminoglycoside antibiotic	4.53E-05
AAC(6')-I	aminoglycoside antibiotic	1.84E-03
AAC(6')-IAA	aminoglycoside antibiotic	7.95E-06
AAC(6')-IAK	aminoglycoside antibiotic	2.22E-05
AAC(6')-IB'	aminoglycoside antibiotic	1.78E-06
AAC(6')-IB7	aminoglycoside antibiotic	2.97E-05
AAC(6')-IB8	aminoglycoside antibiotic	1.98E-04
AAC(6')-IE	aminoglycoside antibiotic	3.42E-05
AAC(6')-II	aminoglycoside antibiotic	3.80E-04
AAD(9)	aminoglycoside antibiotic	6.42E-04
AADA	aminoglycoside antibiotic	3.49E-03
AADA13	aminoglycoside antibiotic	3.04E-05
AADA25	aminoglycoside antibiotic	7.49E-05
AADA9	aminoglycoside antibiotic	1.76E-04
AADB	aminoglycoside antibiotic	5.66E-06
AADD	aminoglycoside antibiotic	1.75E-03
AADE	aminoglycoside antibiotic	3.48E-03
AADK	aminoglycoside antibiotic	2.00E-04
ABCA	multidrug	9.45E-03
ABEM	multidrug	1.14E-04
ABES	multidrug	6.33E-04

ARG	ARG class	abundance
ACC-5	beta-lactams	5.19E-06
ACI-1	beta-lactams	1.01E-04
ACRA	multidrug	1.80E-03
ACRB	multidrug	1.97E-02
ACRD	aminoglycoside antibiotic	5.34E-04
ACRE	multidrug	1.63E-06
ACRF	multidrug	5.75E-04
ACRS	multidrug	1.07E-04
ACT	beta-lactams	5.83E-05
ADC-12	beta-lactams	5.21E-06
ADC-13	beta-lactams	8.25E-07
ADC-14	beta-lactams	4.67E-07
ADC-15	beta-lactams	1.44E-05
ADC-16	beta-lactams	1.38E-05
ADC-18	beta-lactams	3.65E-06
ADC-19	beta-lactams	1.40E-06
ADC-21	beta-lactams	5.11E-06
ADC-22	beta-lactams	8.74E-06
ADC-23	beta-lactams	1.60E-05
ADC-39	beta-lactams	1.78E-06
ADC-41	beta-lactams	5.48E-06
ADC-43	beta-lactams	6.30E-06
ADC-44	beta-lactams	1.73E-05
ADC-6	beta-lactams	6.12E-06
ADC-7	beta-lactams	5.91E-07
ADC-78	beta-lactams	2.45E-06
ADC-79	beta-lactams	2.61E-06
ADC-8	beta-lactams	2.84E-04
ADC-80	beta-lactams	2.69E-05
ADC-81	beta-lactams	8.98E-06
ADEA	tetracycline antibiotic	3.65E-04
ADEB	tetracycline antibiotic	2.48E-03
ADEC	tetracycline antibiotic	8.11E-04
ADEF	multidrug	3.20E-03
ADEH	multidrug	5.33E-04

ARG	ARG class	abundance
ADEI	multidrug	1.30E-03
ADEJ	multidrug	2.52E-03
ADEK	multidrug	7.07E-04
ADEL	multidrug	6.27E-04
ADEN	multidrug	4.65E-04
ADER	tetracycline antibiotic	4.16E-04
ADES	tetracycline antibiotic	1.40E-04
ADP-RIBOSYLATING_TRANSFERASE_ARR	rifamycin antibiotic	1.15E-02
AER-1	beta-lactams	4.51E-04
AIM-1	beta-lactams	1.07E-04
AMPC	multidrug	1.09E-04
AMRB	aminoglycoside antibiotic	3.02E-03
ANT(2")-I	aminoglycoside antibiotic	2.59E-04
ANT(3")-IH	aminoglycoside antibiotic	9.56E-05
ANT(3")-IIA	aminoglycoside antibiotic	5.00E-04
ANT(3")-IIB	aminoglycoside antibiotic	1.12E-05
ANT(3")-IIC	aminoglycoside antibiotic	2.09E-04
ANT(4')-IA	aminoglycoside antibiotic	6.96E-06
ANT(4')-IIA	aminoglycoside antibiotic	6.11E-05
ANT(9)-I	aminoglycoside antibiotic	1.30E-03
ANTIBIOTIC_RESISTANCE_RRNA_ADENINE_METHYLTRANSFERASE	aminoglycoside antibiotic	4.92E-04
APH(2")-IE	aminoglycoside antibiotic	1.68E-05
APH(2")-IF	aminoglycoside antibiotic	2.08E-04
APH(2")-II	aminoglycoside antibiotic	6.18E-05
APH(2")-III	aminoglycoside antibiotic	1.94E-05
APH(3')-I	aminoglycoside antibiotic	5.60E-03
APH(3")-I	aminoglycoside antibiotic	1.55E-03
APH(3')-IIA	aminoglycoside antibiotic	2.07E-03
APH(3')-IIB	aminoglycoside antibiotic	5.76E-04
APH(3'")-III	aminoglycoside antibiotic	1.57E-03
APH(3")-IV	aminoglycoside antibiotic	3.91E-06
APH(3')-IX	aminoglycoside antibiotic	1.11E-06
APH(3')-V	aminoglycoside antibiotic	1.05E-05
APH(3')-VI	aminoglycoside antibiotic	9.01E-05
APH(3')-VII	aminoglycoside antibiotic	3.12E-05
APH(4)-I	aminoglycoside antibiotic	5.41E-04
APH(4)-IB	aminoglycoside antibiotic	1.26E-06
APH(6)-I	aminoglycoside antibiotic	2.28E-03
APH(9)-IA	aminoglycoside antibiotic	6.14E-05
APMA	aminoglycoside antibiotic	1.36E-04
ARLS	multidrug	5.44E-04
ARNA	peptide antibiotic	4.10E-03
ARR-4	rifamycin antibiotic	9.76E-04
ARR-5	rifamycin antibiotic	3.47E-04
ARR-7	rifamycin antibiotic	2.72E-05
ARR-8	rifamycin antibiotic	1.21E-04
AXYX	multidrug	4.19E-05
AXYY	multidrug	1.38E-03
BACA	peptide antibiotic	5.64E-02
BACILLUS_CLUSTER_A_INTRINSIC_MPH	macrolide antibiotic	1.36E-04
BACILLUS_CLUSTER_A_INTRINSIC_MPH	macrolide antibiotic	1.36E-04
BACILLUS_SUBTILIS_MPRF	macrolide antibiotic	5.72E-05
BAER	multidrug	2.80E-03
BAES	multidrug	1.52E-03
BAHA	peptide antibiotic	1.12E-04
BASS	peptide antibiotic	1.39E-05
BCR-1	bicyclomycin	1.81E-05
BCRA	peptide antibiotic	4.22E-03
BCRB	peptide antibiotic	5.37E-05
BCRC	peptide antibiotic	1.80E-05
BEL	beta-lactams	9.65E-06
bifunctional_AAC/APH	aminoglycoside antibiotic	1.45E-03
BJP-1	beta-lactams	2.20E-03
BLAI	beta-lactams	2.32E-05
BLAR1	beta-lactams	8.33E-06
BLAZ	beta-lactams	2.84E-04
BLEO	glycopeptide antibiotic	1.32E-04
BLEOMYCIN_RESISTANCE_PROTEIN	glycopeptide antibiotic	9.83E-04
BLT	multidrug	2.44E-04
BMR	multidrug	3.87E-04
BPEF	multidrug	6.58E-03
BRP(MBL)	glycopeptide antibiotic	8.00E-04
BRUCELLA_SUIS_MPRF	peptide antibiotic	5.77E-05
BURKHOLDERIA_PSEUDOMALLEI_OMP38	multidrug	4.05E-04

ARG	ARG class	abundance
BUT-1	beta-lactams	5.43E-06
CAMP-REGULATORY_PROTEIN	beta-lactams	4.64E-03
CARA	multidrug	8.75E-05
CARB	beta-lactams	4.81E-04
CARO	beta-lactams	5.60E-04
CAT_CHLORAMPHENICOL_ACETYLTRANSFERASE	phenicol antibiotic	2.06E-03
CATA	phenicol antibiotic	1.58E-04
CATB	phenicol antibiotic	1.04E-03
CATB10	phenicol antibiotic	1.21E-04
CATD	phenicol antibiotic	6.18E-05
CATP	phenicol antibiotic	3.38E-05
CATQ	phenicol antibiotic	5.49E-05
CATS	phenicol antibiotic	2.33E-05
CATU	phenicol antibiotic	6.90E-06
CATV	phenicol antibiotic	5.61E-06
CAZ	beta-lactams	1.41E-04
CEOB	multidrug	6.74E-03
CFRC	phenicol antibiotic	9.51E-05
CFXA2	beta-lactams	3.02E-04
CFXA3	beta-lactams	4.76E-05
CFXA6	beta-lactams	1.97E-04
CGB-1	beta-lactams	1.26E-06
CHLORAMPHENICOL_AND_FLORFENICOL_EXPORTER	multidrug	5.06E-04
CHLORAMPHENICOL_AND_FLORFENICOL_RESISTANCE	multidrug	1.01E-04
CHLORAMPHENICOL_EXPORTER	multidrug	2.14E-03
CHRB	MLS	1.89E-06
CHRYSEOBACTERIUM_MENINGOSEPTICUM_BLAB	multidrug	5.51E-06
CIPA	multidrug	1.76E-05
CLASS_A	beta-lactams	2.48E-03
CLASS_B	beta-lactams	1.77E-05
CLASS_C	beta-lactams	1.39E-03
CLASS_D	beta-lactams	9.42E-07
CLBA	multidrug	1.40E-04
CLBB	multidrug	5.44E-05
CLBC	multidrug	3.23E-05
CLOSTRIDIUM_PERFRINGENS_MPRF	peptide antibiotic	1.10E-05
CMEA	multidrug	6.72E-07
CMEB	multidrug	9.50E-04
CMEC	multidrug	1.02E-06
CMLA	phenicol antibiotic	1.91E-04
CMRA	phenicol antibiotic	3.19E-05
CMX	phenicol antibiotic	7.59E-05
CPAA	aminoglycoside antibiotic	3.77E-06
CPS-1	beta-lactams	9.77E-05
COB(I)ALAMIN_ADENOLSYLTRANSFERASE	multidrug	8.52E-04
CPXA	multidrug	9.23E-04
CRP	multidrug	8.51E-07
CRPP	fluoroquinolone antibiotic	5.89E-05
CTX-M	beta-lactams	2.34E-04
CYSTATHIONINE_beta-LYASE_PATB	fluoroquinolone antibiotic	1.61E-06
DFRA1	diaminopyrimidine	3.29E-04
DFRA10	diaminopyrimidine	1.91E-06
DFRA12	diaminopyrimidine	7.93E-05
DFRA13	diaminopyrimidine	1.08E-06
DFRA14	diaminopyrimidine	7.38E-05
DFRA15	diaminopyrimidine	5.93E-05
DFRA16	diaminopyrimidine	1.97E-05
DFRA17	diaminopyrimidine	3.68E-04
DFRA19	diaminopyrimidine	9.17E-06
DFRA20	diaminopyrimidine	8.26E-05
DFRA22	diaminopyrimidine	3.04E-05
DFRA25	diaminopyrimidine	4.73E-06
DFRA2D	diaminopyrimidine	1.62E-05
DFRA3	diaminopyrimidine	2.22E-03
DFRA5	diaminopyrimidine	1.17E-05
DFRA8	diaminopyrimidine	3.31E-05
DFRB1	diaminopyrimidine	8.08E-06
DFRB2	diaminopyrimidine	8.37E-05
DFRB3	diaminopyrimidine	2.23E-05
DFRB6	diaminopyrimidine	5.67E-05
DFRC	diaminopyrimidine	1.12E-03
DFRD	diaminopyrimidine	1.77E-04

ARG	ARG class	abundance
DFRE	diaminopyrimidine	2.43E-05
DFRF	diaminopyrimidine	9.31E-05
DFRG	diaminopyrimidine	5.28E-04
DFRK	diaminopyrimidine	1.23E-04
DHA	beta-lactams	1.16E-05
DIM-1	beta-lactams	9.69E-07
DNA-BINDING_PROTEIN_H-NS	multidrug	7.34E-04
DNA-BINDING_PROTEIN_H-NS	multidrug	7.34E-04
DNA-BINDING_TRANSCRIPTIONAL_REGULATOR_GADX	multidrug	4.64E-05
DNA-BINDING_TRANSCRIPTIONAL_REGULATOR_GADX	multidrug	4.64E-05
EDEQ	multidrug	5.79E-06
EFMA	multidrug	2.20E-05
EFPA	multidrug	3.81E-04
EFRA	multidrug	1.23E-03
EFRB	multidrug	3.50E-03
EMRA	fluoroquinolone antibiotic	6.07E-03
EMRB	fluoroquinolone antibiotic	7.98E-03
EMRB-QACA_FAMILY_MAJOR_FACILITATOR_TRANSPORTER	fluoroquinolone antibiotic	2.53E-03
EMRD	phenicol antibiotic	5.44E-04
EMRE	multidrug	1.05E-03
EMRK	tetracycline antibiotic	9.50E-04
EMRR	fluoroquinolone antibiotic	1.00E-03
EMRY	tetracycline antibiotic	8.99E-05
EPTA	peptide antibiotic	4.57E-04
EREA	MLS	2.81E-05
EREB	MLS	4.25E-06
ERM(31)	MLS	6.89E-05
ERM(33)	MLS	2.91E-05
ERM(35)	MLS	4.07E-05
ERM(36)	MLS	2.81E-04
ERM(37)	MLS	1.08E-04
ERM(38)	MLS	1.55E-04
ERM(39)	MLS	1.76E-04
ERM(41)	MLS	1.33E-04
ERM(42)	MLS	1.25E-05
ERM(43)	MLS	2.58E-05
ERM(44)	MLS	2.12E-06
ERM(47)	MLS	5.25E-05
ERM(TR)	MLS	3.58E-04
ERMA	MLS	7.28E-04
ERMB	MLS	2.76E-03
ERMC	MLS	5.19E-03
ERME	MLS	3.19E-05
ERMF	MLS	7.79E-04
ERMG	MLS	4.94E-04
ERMH	MLS	8.89E-07
ERMO	MLS	7.88E-05
ERMS	MLS	3.02E-05
ERMT	MLS	1.53E-03
ERMU	MLS	9.57E-07
ERMX	MLS	1.15E-03
ESCHERICHIA_COLI_LAMB	multidrug	7.39E-04
ESCHERICHIA_COLI_MDFA	multidrug	2.36E-06
ESCHERICHIA_COLI_MIPA	multidrug	3.36E-04
ESP-1	beta-lactams	1.83E-04
EVGS	multidrug	8.47E-05
FEZ-1	beta-lactams	6.58E-04
FARA	antibacterial free fatty acids	9.78E-04
FARB	antibacterial free fatty acids	1.74E-04
FLOR	phenicol antibiotic	6.97E-04
FMTC	tetracycline antibiotic	5.51E-04
FOMB	fosfomycin	8.02E-06
FONA-6	beta-lactams	1.40E-05
FOSA	fosfomycin	4.14E-04
FOSA3	fosfomycin	3.27E-05
FOSA4	fosfomycin	6.98E-06
FOSA5	fosfomycin	4.84E-05
FOSA6	fosfomycin	5.80E-05
FOSA7	fosfomycin	6.98E-05
FOSB	fosfomycin	2.63E-03
FOSC2	fosfomycin	1.13E-04
FOSK	fosfomycin	2.72E-05
FOSX	fosfomycin	7.03E-04
FOX	beta-lactams	1.13E-05

ARG	ARG class	abundance
FUSB	fusidic acid	8.45E-05
FUSC	fusidic acid	3.83E-05
FUSD	fusidic acid	6.79E-05
FUSH	fusidic acid	3.45E-06
GADW	multidrug	8.14E-05
GES	beta-lactams	2.84E-05
GOB-1	beta-lactams	1.85E-04
GOB-18	beta-lactams	2.69E-05
GOLS	multidrug	2.16E-03
HERA-1	beta-lactams	1.14E-05
HMB-1	beta-lactams	1.83E-06
HMRM	multidrug	2.31E-05
HP1181	multidrug	1.61E-06
ICR-MC	peptide antibiotic	7.07E-05
IMI	beta-lactams	1.12E-06
IMP	beta-lactams	2.88E-05
IND	beta-lactams	4.87E-06
JOHN-1	beta-lactams	6.80E-05
KAMB	aminoglycoside antibiotic	6.67E-06
KASUGAMYCIN_RESISTANCE_PROTEIN_KSGA	aminoglycoside antibiotic	3.46E-03
KDPE	aminoglycoside antibiotic	6.95E-03
KHM-1	beta-lactams	1.36E-06
KLEBSIELLA_PNEUMONIAE_OMPK35	multidrug	3.45E-06
KLEBSIELLA_PNEUMONIAE_OMPK36	multidrug	9.66E-05
KLEBSIELLA_PNEUMONIAE_OMPK37	multidrug	2.50E-04
KLUG-1	cephalosporin	1.12E-05
KPC	beta-lactams	3.72E-04
LARIBACTER_HONGKONGENSIS_AMPC_beta-LACTAMASE	multidrug	2.30E-06
LCR-1	beta-lactams	4.88E-05
LEN	beta-lactams	2.51E-04
LING	MLS	4.30E-05
LISTERIA_MONOCYTOGENES_MPRF	peptide antibiotic	5.99E-06
LLMA_23S_RIBOSOMAL_RNA_METHYLTRANSFERASE	lincosamide antibiotic	8.22E-04
LMRA	lincosamide antibiotic	1.79E-05
LMRB	MLS	6.94E-04
LMRC	multidrug	2.29E-04
LMRD	MLS	3.38E-04
LMRP	multidrug	2.55E-05
LNUA	MLS	3.31E-03
LNUB	MLS	3.11E-04
LNUC	MLS	9.52E-04
LNUD	MLS	3.40E-04
LNUE	MLS	1.78E-04
LNUF	MLS	3.16E-05
LNUG	MLS	2.08E-04
LPEA	MLS	6.85E-07
LPEB	MLS	1.51E-04
LRA	beta-lactams	2.31E-03
LRFA	fluoroquinolone antibiotic	6.83E-04
LSA	multidrug	7.35E-04
LSAC	multidrug	8.86E-05
LSAE	multidrug	5.22E-04
LUXR	multidrug	1.12E-04
MACA	MLS	2.58E-03
MACB	MLS	1.95E-02
MAJOR_FACILITATOR_SUPERFAMILY_TRANSPORTER	multidrug	6.23E-03
MARA	multidrug	4.19E-04
MARR	multidrug	3.36E-03
MCR-1	peptide antibiotic	3.91E-05
MCR-2	peptide antibiotic	8.95E-05
MCR-3	peptide antibiotic	6.29E-05
MCR-4	peptide antibiotic	1.48E-04
MCR-5	peptide antibiotic	5.22E-04
MDFA	tetracycline antibiotic	1.71E-03
MDSA	multidrug	9.21E-05
MDSB	multidrug	1.60E-03
MDSC	multidrug	7.36E-05
MDTA	aminocoumarin antibiotic	3.20E-04
MDTB	aminocoumarin antibiotic	1.99E-03
MDTC	aminocoumarin antibiotic	2.63E-03
MDTD	aminocoumarin antibiotic	6.65E-04
MDTE	multidrug	4.61E-04
MDTF	multidrug	1.17E-03
MDTG	fosfomycin	8.35E-04

ARG	ARG class	abundance
MDTH	fluoroquinolone antibiotic	4.48E-04
MDTK	fluoroquinolone antibiotic	2.82E-03
MDTL	multidrug	3.63E-04
MDTM	multidrug	1.35E-04
MDTN	multidrug	1.72E-04
MDTO	multidrug	1.78E-04
MDTP	multidrug	1.14E-04
MECA	beta-lactams	3.03E-04
MECB	beta-lactams	1.58E-05
MECD	beta-lactams	1.25E-06
MECI	beta-lactams	2.73E-04
MECR1	beta-lactams	4.30E-04
MEF(B)	MLS	2.26E-05
MEFA	MLS	8.06E-04
MEFC	MLS	1.21E-04
MEFE	MLS	7.16E-05
MEL	multidrug	4.97E-04
MEPA	tetracycline antibiotic	1.14E-03
MEPR	tetracycline antibiotic	4.18E-04
METALLO-beta-LACTAMASE	beta-lactams	8.46E-04
MEXA	multidrug	2.28E-03
MEXB	multidrug	9.12E-03
MEXC	multidrug	3.45E-03
MEXD	multidrug	4.38E-03
MEXE	multidrug	7.66E-03
MEXF	multidrug	2.60E-02
MEXG	multidrug	1.40E-04
MEXH	multidrug	2.64E-05
MEXI	multidrug	3.32E-03
MEXJ	multidrug	8.76E-04
MEXK	multidrug	7.81E-03
MEXL	multidrug	9.67E-04
MEXM	phenicol antibiotic	2.56E-05
MEXN	phenicol antibiotic	1.81E-03
MEXP	multidrug	5.38E-05
MEXQ	multidrug	1.91E-03
MEXS	multidrug	4.77E-04
MEXT	multidrug	5.90E-03
MEXV	multidrug	1.52E-04
MEXW	multidrug	5.33E-03
MEXX	multidrug	2.60E-03
MEXY	multidrug	2.15E-03
MEXZ	multidrug	4.17E-04
MFPA	fluoroquinolone antibiotic	2.84E-04
MGRB	multidrug	7.83E-05
MGTA	MLS	2.29E-04
MIR	beta-lactams	2.80E-05
MORAXELLA_CATARRHALIS_M35	multidrug	1.38E-04
MOX	beta-lactams	8.76E-05
MPHA	MLS	9.71E-05
MPHB	MLS	2.27E-04
MPHC	MLS	4.76E-04
MPHD	MLS	6.45E-04
MPHE	MLS	1.13E-04
MPHG	MLS	1.25E-04
MPHI	MLS	4.74E-06
MRX	multidrug	6.87E-05
MSBA	nitroimidazole antibiotic	1.33E-02
MSI-1	beta-lactams	1.18E-04
MSRA	multidrug	6.46E-04
MSRC	multidrug	4.24E-05
MSRE	multidrug	6.55E-04
MTRA	multidrug	2.79E-02
MTRC	multidrug	5.83E-05
MTRD	multidrug	1.29E-03
MTRE	multidrug	1.51E-04
MTRR	multidrug	1.67E-04
multidrug_ABC_TRANSPORTER	multidrug	5.25E-02
MUXA	multidrug	8.43E-04
MUXB	multidrug	5.55E-03
MUXC	multidrug	1.72E-03
MVAT	multidrug	2.20E-03
MYCINAMICIN-RESISTANCE_PROTEIN_MYRB	MLS	3.55E-05
MYRA	MLS	2.41E-04
NDM	multidrug	1.39E-05
NMCR	beta-lactams	4.66E-04
NORA	multidrug	1.05E-03
NORB	fluoroquinolone antibiotic	8.27E-04
NPS-1	beta-lactams	2.02E-04
OCH	beta-lactams	5.09E-05
OKP-A	beta-lactams	5.04E-05

ARG	ARG class	abundance
OKP-B	beta-lactams	2.82E-04
OKP-B-12	beta-lactams	2.26E-05
OLEB	multidrug	7.22E-05
OLEC	MLS	6.74E-03
OLED	MLS	1.40E-04
OLEI	MLS	6.65E-06
OMP36	multidrug	1.71E-04
OMPF	multidrug	1.44E-03
OMPR	multidrug	2.43E-02
OPCM	multidrug	9.66E-04
OPMB	multidrug	6.39E-04
OPMD	multidrug	3.47E-06
OPME	multidrug	1.02E-05
OPMH	triclosan	2.38E-03
OPRA	multidrug	1.74E-04
OPRC	multidrug	6.74E-04
OPRJ	multidrug	7.41E-04
OPRM	multidrug	4.77E-03
OPRN	multidrug	1.63E-03
OPRZ	multidrug	7.62E-05
OPTRA	multidrug	8.46E-04
OQXA	multidrug	6.47E-04
OQXB	multidrug	3.44E-03
OTR(B)	tetracycline antibiotic	8.16E-04
OTRA	tetracycline antibiotic	1.31E-03
OTRC	tetracycline antibiotic	9.09E-04
OXA	beta-lactams	1.98E-02
OXY	beta-lactams	3.81E-05
PATA	fluoroquinolone antibiotic	9.79E-04
PATB	fluoroquinolone antibiotic	3.95E-04
PBP-1A	multidrug	8.72E-04
PBP-1B	multidrug	7.02E-04
PBP-2X	multidrug	3.19E-04
PDC	beta-lactams	6.00E-05
PEDO-1	beta-lactams	2.01E-04
PEDO-2	beta-lactams	5.86E-05
PEDO-3	beta-lactams	8.93E-06
PENA	beta-lactams	1.18E-03
PER	beta-lactams	2.67E-04
PGPB	peptide antibiotic	9.19E-07
PMRA	fluoroquinolone antibiotic	7.83E-06
PMRF	peptide antibiotic	1.79E-03
PORIN_OMPC	multidrug	1.07E-04

ARG	ARG class	abundance
POXTA	multidrug	3.02E-04
PSE-1	multidrug	7.00E-06
PSEUDOMONAS_AERUGINOSA_EMRE	multidrug	4.51E-04
PUROMYCIN_RESISTANCE_PROTEIN	nucleoside antibiotic	3.62E-03
PVRR	multidrug	4.85E-05
QACA	fluoroquinolone antibiotic	7.96E-05
QACB	fluoroquinolone antibiotic	3.18E-04
QACG	fluoroquinolone antibiotic	9.06E-04
QACH	fluoroquinolone antibiotic	3.30E-03
QEPA	fluoroquinolone antibiotic	2.50E-03
QEPA1	fluoroquinolone antibiotic	3.03E-06
QEPA2	fluoroquinolone antibiotic	5.18E-05
QNRA	fluoroquinolone antibiotic	2.46E-06
QNRB	fluoroquinolone antibiotic	4.36E-05
QNRB15	fluoroquinolone antibiotic	5.26E-06
QNRB27	fluoroquinolone antibiotic	2.46E-06
QNRB33	fluoroquinolone antibiotic	7.56E-07
QNRB40	fluoroquinolone antibiotic	9.75E-07
QNRB62	fluoroquinolone antibiotic	1.15E-06
QNRB66	fluoroquinolone antibiotic	6.50E-06
QNRB71	fluoroquinolone antibiotic	5.85E-06
QNRB72	fluoroquinolone antibiotic	1.01E-06
QNRC	fluoroquinolone antibiotic	4.42E-06
QNRD1	fluoroquinolone antibiotic	1.52E-05
QNRD2	fluoroquinolone antibiotic	5.37E-05
QNRS	fluoroquinolone antibiotic	6.06E-04
QNRS3	fluoroquinolone antibiotic	7.41E-06
QNRS5	fluoroquinolone antibiotic	3.30E-06
QNRS6	fluoroquinolone antibiotic	8.70E-04
QNRS8	fluoroquinolone antibiotic	2.46E-05
QNRS9	fluoroquinolone antibiotic	8.94E-06
QNRVC1	fluoroquinolone antibiotic	1.39E-04
QNRVC3	fluoroquinolone antibiotic	8.21E-06
QNRVC4	fluoroquinolone antibiotic	6.47E-05
QNRVC5	fluoroquinolone antibiotic	1.19E-04
QNRVC6	fluoroquinolone antibiotic	1.74E-04
QNRVC7	fluoroquinolone antibiotic	2.46E-06
R39	beta-lactams	1.19E-05
RAMA	multidrug	3.10E-04
RBPA	rifamycin antibiotic	2.74E-03
RGT1438	rifamycin antibiotic	2.61E-04

ARG	ARG class	abundance
RHODOBACTER_SPHAEROIDES_AMPC_beta-LACTAMASE	multidrug	5.83E-05
RIFAMPIN_MONOOXYGENASE	rifamycin antibiotic	9.05E-03
RLMA(II)	MLS	4.20E-04
RM3	beta-lactams	6.17E-04
RMTD	aminoglycoside antibiotic	5.24E-06
ROB-1	beta-lactams	1.17E-05
ROSA	peptide antibiotic	1.84E-02
ROSB	peptide antibiotic	1.86E-02
RPHA	rifamycin antibiotic	3.81E-03
RPHB	rifamycin antibiotic	5.98E-03
RPOB2	rifamycin antibiotic	1.23E-01
RPSD_(RAMA_OR_SUD2)	multidrug	7.23E-04
RSA-1	beta-lactams	2.84E-06
RSA-2	beta-lactams	1.86E-06
RTG-4	multidrug	8.03E-05
RTG-5	multidrug	5.50E-06
SALA	multidrug	6.39E-05
SAT-2	nucleoside antibiotic	3.43E-05
SAT-3	nucleoside antibiotic	3.15E-05
SAT-4	nucleoside antibiotic	6.36E-04
SDEY	multidrug	4.33E-05
SDIA	multidrug	3.85E-04
SERRATIA_MARCESCENS_OMP1	multidrug	3.23E-04
SFH-1	beta-lactams	1.35E-06
SGM	aminoglycoside antibiotic	7.65E-06
SHV	beta-lactams	5.11E-04
SIM-1	beta-lactams	1.22E-05
SMB-1	beta-lactams	7.27E-06
SME-3	beta-lactams	2.55E-06
SMEB	multidrug	4.74E-04
SMEC	multidrug	1.49E-04
SMED	multidrug	6.31E-04
SMEE	multidrug	1.35E-03
SMEF	multidrug	2.80E-04
SMER	multidrug	4.18E-03
SMES	multidrug	3.43E-04
SPCN	aminoglycoside antibiotic	1.04E-05
SPG-1	beta-lactams	2.53E-04
SRMB	multidrug	3.70E-05
SRT-1	beta-lactams	1.28E-05
SRT-2	beta-lactams	2.06E-05
STAPHYLOCOCCUS_MUPA_CONFERRING_RESISTANCE_TO_MUPIROCIN	multidrug	1.54E-03
STREPTOCOCCUS_AGALACTIAE_MPRF	peptide antibiotic	2.43E-06
STREPTOMYCIN_RESISTANCE_PROTEIN	aminoglycoside antibiotic	1.05E-03
STREPTOTHRICIN_ACETYLTRANSFERASE	nucleoside antibiotic	4.25E-05
SUL1	sulfonamide antibiotic	2.44E-03
SUL2	sulfonamide antibiotic	7.40E-04
TAEA	Diterpenoids	2.95E-03
TAP	tetracycline antibiotic	2.74E-04
TCMA	tetracenomycin antibiotic	2.52E-04
TCR3	tetracycline antibiotic	2.01E-04
TEM	beta-lactams	1.06E-02
TET(33)	tetracycline antibiotic	4.27E-04
TET(38)	tetracycline antibiotic	8.60E-05
TET(40)	tetracycline antibiotic	1.72E-05
TET(42)	tetracycline antibiotic	2.38E-04
TET(45)	tetracycline antibiotic	3.61E-05
TET(59)	tetracycline antibiotic	2.44E-05
TET(K)	tetracycline antibiotic	8.91E-04
TET(W/N/W)	tetracycline antibiotic	8.44E-06
TET31	tetracycline antibiotic	2.23E-06
TET32	tetracycline antibiotic	3.00E-04
TET34	tetracycline antibiotic	7.40E-04
TET35	tetracycline antibiotic	6.53E-04
TET36	tetracycline antibiotic	6.57E-05
TET37	tetracycline antibiotic	1.81E-04
TET39	tetracycline antibiotic	4.26E-04
TET40	tetracycline antibiotic	3.93E-04
TET41	tetracycline antibiotic	1.13E-03
TET43	tetracycline antibiotic	6.72E-04
TET44	tetracycline antibiotic	5.40E-04
TETA	tetracycline antibiotic	2.67E-03
TETA(46)	tetracycline antibiotic	1.10E-04
TETA(48)	tetracycline antibiotic	1.30E-02

ARG	ARG class	abundance
TETA(60)	tetracycline antibiotic	1.18E-04
TETB	tetracycline antibiotic	3.34E-04
TETB(46)	tetracycline antibiotic	9.38E-04
TETB(48)	tetracycline antibiotic	2.36E-03
TETB(60)	tetracycline antibiotic	6.68E-04
TETC	tetracycline antibiotic	3.60E-05
TETD	tetracycline antibiotic	1.15E-05
TETE	tetracycline antibiotic	2.42E-05
TETG	tetracycline antibiotic	1.55E-04
TETH	tetracycline antibiotic	9.38E-05
TETJ	tetracycline antibiotic	1.21E-06
TETL	tetracycline antibiotic	2.68E-03
TETM	tetracycline antibiotic	2.41E-03
TETO	tetracycline antibiotic	2.98E-04
TETP	tetracycline antibiotic	1.16E-03
TETQ	tetracycline antibiotic	9.69E-04
TETR	tetracycline antibiotic	1.49E-03
TETS	tetracycline antibiotic	1.54E-04
TETT	tetracycline antibiotic	9.23E-05
TETV	tetracycline antibiotic	1.03E-03
TETW	tetracycline antibiotic	2.08E-03
TETX	tetracycline antibiotic	6.72E-04
TETY	tetracycline antibiotic	5.48E-05
TETZ	tetracycline antibiotic	6.75E-04
THIN-B	beta-lactams	1.09E-03
THIOSTREPTON_RESISTANCE_METHYLASE_TSNR	peptide antibiotic	1.98E-05
TLA-1	multidrug	5.63E-05
TLA-2	multidrug	4.28E-05
TLA-3	multidrug	1.15E-04
TLCC	multidrug	4.25E-05
TLRB_CONFERRING_TYLOSIN_RESISTANCE	multidrug	1.19E-05
TMRB	nucleoside antibiotic	1.33E-06
TOLC	multidrug	1.82E-03
TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR	multidrug	2.65E-02
TRIA	triclosan	2.12E-04
TRIB	triclosan	1.18E-04
TRIC	triclosan	2.71E-03
TRU-1	beta-lactams	2.48E-06
TRUNCATED_PUTATIVE_RESPONSE_REGULATOR_ARLR	multidrug	4.34E-03
TWO-COMPONENT_SYSTEM_RESPONSE_REGULATOR_EVGA	multidrug	7.94E-06
TYPE_A_NFXB	multidrug	6.21E-04
UGD	peptide antibiotic	1.15E-02
VANA	glycopeptide antibiotic	1.21E-03
VANB	glycopeptide antibiotic	4.48E-05
VANC	glycopeptide antibiotic	6.24E-04
VAND	glycopeptide antibiotic	1.05E-04
VANE	glycopeptide antibiotic	1.42E-04
VANG	glycopeptide antibiotic	1.78E-04
VANH	glycopeptide antibiotic	1.20E-03
VANI	glycopeptide antibiotic	3.67E-04
VANJ	glycopeptide antibiotic	1.81E-04
VANKI	glycopeptide antibiotic	1.99E-06
VANL	glycopeptide antibiotic	2.44E-05
VANM	glycopeptide antibiotic	2.07E-05
VANN	glycopeptide antibiotic	1.07E-05
VANR	glycopeptide antibiotic	1.82E-02
VANRI	glycopeptide antibiotic	1.16E-03
VANS	glycopeptide antibiotic	6.42E-03
VANSD	glycopeptide antibiotic	9.05E-06
VANT	glycopeptide antibiotic	1.37E-06
VANTG	glycopeptide antibiotic	4.46E-06
VANU	glycopeptide antibiotic	2.27E-05
VANVB	glycopeptide antibiotic	1.85E-05
VANW	glycopeptide antibiotic	1.45E-05
VANX	glycopeptide antibiotic	1.59E-03
VANXI	glycopeptide antibiotic	3.36E-05
VANY	glycopeptide antibiotic	2.96E-04
VANYG1	glycopeptide antibiotic	2.95E-05
VANZ	glycopeptide antibiotic	6.86E-05
VATA	MLS	1.01E-04
VATB	MLS	8.07E-04
VATC	MLS	1.85E-04
VATD	MLS	1.51E-04
VATE	MLS	5.07E-04
VATF	MLS	2.47E-03

2. Annotation of potential mobile ARGs, associated risk rank, and calculation of read numbers in the urban and background atmosphere globally.

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
bacA	I	peptide antibiotic	antibiotic target alteration	100	5
lnuA	I	MLS	antibiotic inactivation	90	2
ErmB	I	MLS	antibiotic target alteration	44	4
ErmC	I	MLS	antibiotic target alteration	46	1
tetM	I	tetracycline antibiotic	antibiotic target protection	43	2
norA	I	multidrug	antibiotic efflux	42	3
mdtE	I	multidrug	antibiotic efflux	35	0
mecR1	I	Beta-lactams	antibiotic target replacement	30	2
tolC	I	multidrug	antibiotic efflux	27	5
ErmT	I	MLS	antibiotic target alteration	24	0
floR	I	phenicol antibiotic	antibiotic efflux	21	4
dfrA12	I	diaminopyrimidine antibiotic	antibiotic target replacement	27	0
mecA	I	Beta-lactams	antibiotic target replacement	16	0
tetL	I	tetracycline antibiotic	antibiotic target protection	17	0
dfrA17	I	diaminopyrimidine antibiotic	antibiotic target replacement	11	1
lnuB	I	MLS	antibiotic inactivation	9	0
msrA	I	multidrug	antibiotic target protection	7	0
fosB	I	fosfomycin	antibiotic inactivation	5	0
blaZ	I	Beta-lactams	antibiotic inactivation	3	0
VEB-3	I	Beta-lactams	antibiotic inactivation	6	1
mdtL	I	multidrug	antibiotic efflux	4	1
mepA	I	tetracycline antibiotic	antibiotic efflux	2	4
vanY	I	glycopeptide antibiotic	antibiotic target alteration	0	1
mphA	I	MLS	antibiotic inactivation	0	1
tetO	II	tetracycline antibiotic	antibiotic target protection	91	8
tetW	II	tetracycline antibiotic	antibiotic target protection	81	2
penA	II	Beta-lactams	antibiotic inactivation	80	6
mdtG	II	fosfomycin	antibiotic efflux	72	3
mdtM	II	multidrug	antibiotic efflux	43	1
vatE	II	MLS	antibiotic inactivation	25	0
aadA	II	aminoglycoside antibiotic	antibiotic inactivation	34	2
emrD	II	phenicol antibiotic	antibiotic efflux	2	1
mdfA	II	tetracycline antibiotic	antibiotic efflux	1	0
macB	III	MLS	antibiotic efflux	862	369
emrB	III	fluoroquinolone antibiotic	antibiotic efflux	178	41
macA	III	MLS	antibiotic efflux	67	22
acrB	III	multidrug	antibiotic efflux	52	19
arnA	III	peptide antibiotic	antibiotic target alteration	50	15
sdiA	III	multidrug	antibiotic efflux	42	18
ampC	III	multidrug	antibiotic inactivation	21	10
vanS	III	glycopeptide antibiotic	antibiotic target alteration	20	8
MdtN	III	multidrug	antibiotic efflux	20	6
AcrA	III	multidrug	antibiotic efflux	15	4
mdtP	III	multidrug	antibiotic efflux	14	5
MdtO	III	multidrug	antibiotic efflux	13	1
ykkD	III	multidrug	antibiotic efflux	9	0
mecI	III	Beta-lactams	antibiotic target replacement	8	2
emrE	III	multidrug	antibiotic efflux	8	1
tet44	III	tetracycline antibiotic	antibiotic target protection	7	0
tet32	III	tetracycline antibiotic	antibiotic target protection	7	2
VatB	III	MLS	antibiotic inactivation	6	4
mdtK	III	fluoroquinolone antibiotic	antibiotic efflux	5	2

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
ErmA	III	MLS	antibiotic target alteration	5	0
ErmX	III	MLS	antibiotic target alteration	5	1
tetA	III	tetracycline antibiotic	antibiotic efflux	3	1
mefA	III	MLS	antibiotic efflux	2	1
tetQ	III	tetracycline antibiotic	antibiotic target protection	2	2
mdtF	III	multidrug	antibiotic efflux	2	2
tetB	III	tetracycline antibiotic	antibiotic efflux	2	0
vanG	III	glycopeptide antibiotic	antibiotic target alteration	1	1
ErmF	III	MLS	antibiotic target alteration	1	1
ErmG	III	MLS	antibiotic target alteration	1	0
CfxA2	III	Beta-lactams	antibiotic inactivation	0	1
PBP-2X	III	multidrug	antibiotic target replacement	0	1
bcrA	IV	peptide antibiotic	antibiotic efflux	37	73
MexT	IV	multidrug	antibiotic efflux	51	31
MexF	IV	multidrug	antibiotic efflux	20	16
mdtC	IV	aminocoumarin antibiotic	antibiotic efflux	13	19
mexW	IV	multidrug	antibiotic efflux	94	23
MexD	IV	multidrug	antibiotic efflux	87	19
emrA	IV	fluoroquinolone antibiotic	antibiotic efflux	81	18
OprN	IV	multidrug	antibiotic efflux	67	17
mdtB	IV	aminocoumarin antibiotic	antibiotic efflux	64	22
tetR	IV	tetracycline antibiotic	antibiotic target protection	52	10
ceoB	IV	multidrug	antibiotic efflux	50	9
AmrB	IV	aminoglycoside antibiotic	antibiotic efflux	50	4
tetC	IV	tetracycline antibiotic	antibiotic target protection	48	14
OpcM	IV	multidrug	antibiotic efflux	46	2
sul1	IV	sulfonamide antibiotic	antibiotic target replacement	45	5
oprA	IV	multidrug	antibiotic efflux	38	9
mdtA	IV	aminocoumarin antibiotic	antibiotic efflux	30	9
rosA	IV	peptide antibiotic	antibiotic efflux	28	8
mexY	IV	multidrug	antibiotic efflux	25	16
MexB	IV	multidrug	antibiotic efflux	24	7
tetX	IV	tetracycline antibiotic	antibiotic inactivation	22	5
RosB	IV	peptide antibiotic	antibiotic efflux	21	8
mdtD	IV	aminocoumarin antibiotic	antibiotic efflux	21	6
adeB	IV	tetracycline antibiotic	antibiotic efflux	19	4
mexH	IV	multidrug	antibiotic efflux	16	2
smeF	IV	multidrug	antibiotic efflux	16	6
OprJ	IV	multidrug	antibiotic efflux	15	3
lmrB	IV	MLS	antibiotic efflux	13	6
tetT	IV	tetracycline antibiotic	antibiotic target protection	13	5
MexE	IV	multidrug	antibiotic efflux	13	6
adeA	IV	tetracycline antibiotic	antibiotic efflux	12	0
aadK	IV	aminoglycoside antibiotic	antibiotic inactivation	12	0
cmeB	IV	multidrug	antibiotic efflux	10	4
vgaE	IV	multidrug	antibiotic target protection	9	6
carA	IV	multidrug	antibiotic target protection	9	3
mexI	IV	multidrug	antibiotic efflux	8	0
MexC	IV	multidrug	antibiotic efflux	8	5
LmrA	IV	lincosamide antibiotic	antibiotic target alteration	8	1
vanD	IV	glycopeptide antibiotic	antibiotic target alteration	7	0
tcr3	IV	tetracycline antibiotic	antibiotic efflux	7	0
mtrR	IV	multidrug	antibiotic efflux	7	0
PmrA	IV	fluoroquinolone antibiotic	antibiotic efflux	7	0
opmD	IV	multidrug	antibiotic efflux	7	1
srmB	IV	multidrug	antibiotic target protection	6	2
adeJ	IV	multidrug	antibiotic efflux	6	1

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
tetY	IV	tetracycline antibiotic	antibiotic inactivation	6	1
vanE	IV	glycopeptide antibiotic	antibiotic target alteration	5	0
smeE	IV	multidrug	antibiotic efflux	5	0
smeC	IV	multidrug	antibiotic efflux	5	1
sul2	IV	sulfonamide antibiotic	antibiotic target replacement	4	2
otrA	IV	tetracycline antibiotic	antibiotic target protection	4	1
PBP-1B	IV	multidrug	antibiotic target replacement	4	0
bleO	IV	glycopeptide antibiotic	antibiotic inactivation	4	0
qepA	IV	fluoroquinolone antibiotic	antibiotic efflux	4	0
oprM	IV	multidrug	antibiotic efflux	4	3
emrK	IV	tetracycline antibiotic	antibiotic efflux	4	1
smeB	IV	multidrug	antibiotic efflux	4	1
vgaA	IV	multidrug	antibiotic target protection	3	0
qacG	IV	fluoroquinolone antibiotic	antibiotic efflux	3	0
tet34	IV	tetracycline antibiotic	antibiotic target protection	3	0
tet39	IV	tetracycline antibiotic	antibiotic target protection	3	0
adeK	IV	multidrug	antibiotic efflux	3	0
abeS	IV	multidrug	antibiotic efflux	3	1
qacA	IV	fluoroquinolone antibiotic	antibiotic efflux	3	0
mtrE	IV	multidrug	antibiotic efflux	2	0
tetH	IV	tetracycline antibiotic	antibiotic target protection	2	1
FosX	IV	fosfomycin	antibiotic inactivation	2	0
PBP-1A	IV	multidrug	antibiotic target replacement	2	1
SRT-2	IV	Beta-lactams	antibiotic inactivation	2	3
marR	IV	multidrug	antibiotic efflux; reduced permeability to antibiotic	2	0
smeD	IV	multidrug	antibiotic efflux	2	0
tet36	IV	tetracycline antibiotic	antibiotic target protection	2	0
tetD	IV	tetracycline antibiotic	antibiotic target protection	1	0
adeC	IV	tetracycline antibiotic	antibiotic efflux	1	0
tet35	IV	tetracycline antibiotic	antibiotic target protection	1	2
lmrP	IV	multidrug	antibiotic efflux	1	0
vanA	IV	glycopeptide antibiotic	antibiotic target alteration	1	1
msrC	IV	multidrug	antibiotic target protection	1	0
bcrC	IV	peptide antibiotic	antibiotic target alteration	1	1
tet31	IV	tetracycline antibiotic	antibiotic target protection	1	1
tetZ	IV	tetracycline antibiotic	antibiotic inactivation	1	0
tetV	IV	tetracycline antibiotic	antibiotic efflux	1	0
Erm(39)	IV	MLS	antibiotic target alteration	1	0
MexA	IV	multidrug	antibiotic efflux	1	2
vanC	IV	glycopeptide antibiotic	antibiotic target alteration	1	0
acrF	IV	multidrug	antibiotic efflux	1	1
tet41	IV	tetracycline antibiotic	antibiotic target protection	0	1
fusH	IV	fusidic acid	antibiotic inactivation	0	1
oleB	Unassessed	multidrug	antibiotic target protection	34	5
dfrA20	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	6	1
cmrA	Unassessed	phenicol antibiotic	antibiotic efflux	3	1
vgaD	Unassessed	multidrug	antibiotic target protection	3	0
tet37	Unassessed	tetracycline antibiotic	antibiotic inactivation	2	1
vgaB	Unassessed	multidrug	antibiotic target protection	2	0
vanM	Unassessed	glycopeptide antibiotic	antibiotic target alteration	2	0
tet43	Unassessed	tetracycline antibiotic	antibiotic target protection	1	0
blaI	Unassessed	Beta-lactams	antibiotic inactivation	1	0
mexG	Unassessed	multidrug	antibiotic efflux	0	1
msbA	Unassessed	nitroimidazole antibiotic	antibiotic efflux	292	91
adeL	Unassessed	multidrug	antibiotic efflux	286	71
evgS	Unassessed	multidrug	antibiotic efflux	254	96

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
PvrR	Unassessed	multidrug	resistance by absence	249	58
mtrA	Unassessed	multidrug	antibiotic efflux	194	36
patA	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	188	48
abcA	Unassessed	multidrug	antibiotic efflux	186	111
tetA(48)	Unassessed	tetracycline antibiotic	antibiotic efflux	181	46
golS	Unassessed	multidrug	antibiotic efflux	168	35
cpxA	Unassessed	multidrug	antibiotic efflux	165	47
NmcR	Unassessed	Beta-lactams	antibiotic inactivation	155	45
olec	Unassessed	MLS	antibiotic efflux	140	43
smeS	Unassessed	multidrug	antibiotic efflux	128	28
baeS	Unassessed	multidrug	antibiotic efflux	86	26
MuxB	Unassessed	multidrug	antibiotic efflux	75	20
kdpE	Unassessed	aminoglycoside antibiotic	antibiotic efflux	71	17
adeR	Unassessed	tetracycline antibiotic	antibiotic efflux	66	9
smeR	Unassessed	multidrug	antibiotic efflux	61	15
mexK	Unassessed	multidrug	antibiotic efflux	61	15
TaeA	Unassessed	Diterpenoids	antibiotic efflux	57	13
baeR	Unassessed	multidrug	antibiotic efflux	57	19
arlS	Unassessed	multidrug	antibiotic efflux	55	22
mexS	Unassessed	multidrug	antibiotic efflux	55	20
ChrB	Unassessed	MLS	antibiotic target alteration	47	8
dfrE	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	46	9
optrA	Unassessed	multidrug	antibiotic target protection	43	21
tetA(46)	Unassessed	tetracycline antibiotic	antibiotic efflux	42	10
tetA(60)	Unassessed	tetracycline antibiotic	antibiotic efflux	40	9
bcr-1	Unassessed	bicyclomycin	antibiotic efflux	40	13
MuxC	Unassessed	multidrug	antibiotic efflux	38	6
lmrC	Unassessed	multidrug	antibiotic target protection	35	11
otr(B)	Unassessed	tetracycline antibiotic	antibiotic efflux	34	14
patB	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	30	10
mexN	Unassessed	phenicol antibiotic	antibiotic efflux	29	7
OpmH	Unassessed	triclosan	antibiotic efflux	29	8
mexJ	Unassessed	multidrug	antibiotic efflux	29	8
qacH	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	28	10
lrfA	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	27	6
efrB	Unassessed	multidrug	antibiotic efflux	25	8
lmrD	Unassessed	MLS	antibiotic efflux	23	6
rphB	Unassessed	rifamycin antibiotic	antibiotic inactivation	23	6
cmx	Unassessed	phenicol antibiotic	antibiotic efflux	23	0
vanRI	Unassessed	glycopeptide antibiotic	antibiotic target alteration	23	7
adeF	Unassessed	multidrug	antibiotic efflux	22	1
cob(I)	Unassessed	multidrug	antibiotic efflux	22	1
efpA	Unassessed	multidrug	antibiotic efflux	21	5
tetB(46)	Unassessed	tetracycline antibiotic	antibiotic efflux	21	4
RlmA(II)	Unassessed	MLS	antibiotic target alteration	19	2
vantg	Unassessed	glycopeptide antibiotic	antibiotic target alteration	19	9
mexV	Unassessed	multidrug	antibiotic efflux	19	1
adeH	Unassessed	multidrug	antibiotic efflux	18	4
efrA	Unassessed	multidrug	antibiotic efflux	17	3
TriA	Unassessed	triclosan	antibiotic efflux	17	3
dfrA3	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	17	4
mexZ	Unassessed	multidrug	antibiotic efflux	16	2
adeS	Unassessed	tetracycline antibiotic	antibiotic efflux	14	6
AxyY	Unassessed	multidrug	antibiotic efflux	13	5
major_facilitator_superfamily_transporter	Unassessed	multidrug	antibiotic efflux	16	5

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
tetB(60)	Unassessed	tetracycline antibiotic	antibiotic efflux	13	8
tcmA	Unassessed	tetracenomycin antibiotic	antibiotic efflux	13	2
ermB	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	12	8
hp1181	Unassessed	multidrug	antibiotic efflux	12	4
opmE	Unassessed	multidrug	antibiotic efflux	12	2
rphA	Unassessed	rifamycin antibiotic	antibiotic inactivation	12	1
YojI	Unassessed	peptide antibiotic	antibiotic efflux	12	3
MCR-5	Unassessed	peptide antibiotic	antibiotic target alteration	12	0
two-component_system_response_regulator_EvgA	Unassessed	multidrug	antibiotic efflux	10	1
AAC(6')-IE	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	9	0
OpmB	Unassessed	multidrug	antibiotic efflux	9	3
rpsD_(ramA_or_sud2)	Unassessed	multidrug	antibiotic target alteration	8	3
vanSD	Unassessed	glycopeptide antibiotic	antibiotic target alteration	8	4
bcrB	Unassessed	peptide antibiotic	antibiotic efflux	8	0
mdsb	Unassessed	multidrug	antibiotic efflux	8	2
hmrM	Unassessed	multidrug	antibiotic efflux	7	4
LpeB	Unassessed	MLS	antibiotic efflux	7	2
poxtA	Unassessed	multidrug	antibiotic target protection	7	4
msrE	Unassessed	multidrug	antibiotic target protection	7	1
mdsC	Unassessed	multidrug	antibiotic efflux	7	1
MCR-3	Unassessed	peptide antibiotic	antibiotic target alteration	7	4
salA	Unassessed	multidrug	antibiotic target protection	6	2
farA	Unassessed	antibacterial free fatty acids	antibiotic efflux	6	2
dfrG	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	6	0
cfrC	Unassessed	phenicol antibiotic	antibiotic target alteration	6	1
EdeQ	Unassessed	multidrug	antibiotic inactivation	6	0
otrC	Unassessed	tetracycline antibiotic	antibiotic efflux	5	0
MuxA	Unassessed	multidrug	antibiotic efflux	5	3
MCR-2	Unassessed	peptide antibiotic	antibiotic target alteration	5	0
tet(K)	Unassessed	tetracycline antibiotic	antibiotic efflux	5	0
myrA	Unassessed	MLS	antibiotic target alteration	5	1
TriC	Unassessed	triclosan	antibiotic efflux	5	2
farB	Unassessed	antibacterial free fatty acids	antibiotic efflux	4	1
Bmr	Unassessed	multidrug	antibiotic efflux	4	0
LsaE	Unassessed	multidrug	antibiotic target protection	4	0
apmA	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	4	1
oqxB	Unassessed	multidrug	antibiotic efflux	4	2
ramA	Unassessed	multidrug	antibiotic efflux; reduced permeability to antibiotic	4	0
acrD	Unassessed	aminoglycoside antibiotic	antibiotic efflux	4	1
oqxA	Unassessed	multidrug	antibiotic efflux	4	2
BJP-1	Unassessed	Beta-lactams	antibiotic inactivation	3	0
cpaA	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	3	0
lnuD	Unassessed	MLS	antibiotic inactivation	3	0
MvaT	Unassessed	multidrug	antibiotic efflux	3	2
vanI	Unassessed	glycopeptide antibiotic	antibiotic target alteration	3	0
TLA-1	Unassessed	multidrug	antibiotic inactivation	3	1
mphD	Unassessed	MLS	antibiotic inactivation	3	1
tet(33)	Unassessed	tetracycline antibiotic	antibiotic efflux	3	0
vgaC	Unassessed	multidrug	antibiotic target protection	3	2
efmA	Unassessed	multidrug	antibiotic efflux	3	0
AIM-1	Unassessed	Beta-lactams	antibiotic inactivation	3	0
oleI	Unassessed	MLS	antibiotic inactivation	3	0

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
acrE	Unassessed	multidrug	antibiotic efflux	2	1
AAC(6')-Ib'	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	2	0
dfrK	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	2	0
vanL	Unassessed	glycopeptide antibiotic	antibiotic target alteration	2	1
pmrF	Unassessed	peptide antibiotic	antibiotic target alteration	2	0
tet(42)	Unassessed	tetracycline antibiotic	antibiotic efflux	2	0
kasugamycin_resistance_protein_ksgA	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	2	1
RbpA	Unassessed	rifamycin antibiotic	antibiotic target protection	2	0
marA	Unassessed	multidrug	antibiotic efflux	2	1
rgt1438	Unassessed	rifamycin antibiotic	antibiotic inactivation	2	0
mexL	Unassessed	multidrug	antibiotic efflux	2	1
MSI-1	Unassessed	Beta-lactams	antibiotic inactivation	2	0
abeM	Unassessed	multidrug	antibiotic efflux	2	1
SPG-1	Unassessed	Beta-lactams	antibiotic inactivation	2	0
APH(9)-Ia	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	2	0
vanKI	Unassessed	glycopeptide antibiotic	antibiotic target alteration	2	0
TriB	Unassessed	triclosan	antibiotic efflux	1	0
tap	Unassessed	tetracycline antibiotic	antibiotic efflux	1	0
mexQ	Unassessed	multidrug	antibiotic efflux	1	0
VatI	Unassessed	MLS	antibiotic inactivation	1	0
blt	Unassessed	multidrug	antibiotic efflux	1	1
FosC2	Unassessed	fosfomycin	antibiotic inactivation	1	0
clbC	Unassessed	multidrug	antibiotic target alteration	1	0
catB10	Unassessed	phenicol antibiotic	antibiotic inactivation	1	0
MCR-4	Unassessed	peptide antibiotic	antibiotic target alteration	1	0
mel	Unassessed	multidrug	antibiotic target protection	1	0
lnuE	Unassessed	MLS	antibiotic inactivation	1	0
mphI	Unassessed	MLS	antibiotic inactivation	1	0
lnuG	Unassessed	MLS	antibiotic inactivation	1	0
SAT-4	Unassessed	nucleoside antibiotic	antibiotic inactivation	1	1
mepR	Unassessed	tetracycline antibiotic	antibiotic efflux	1	0
mtrD	Unassessed	multidrug	antibiotic efflux	1	0
mdsA	Unassessed	multidrug	antibiotic efflux	1	0
GOB-18	Unassessed	Beta-lactams	antibiotic inactivation	1	0
CRP	Unassessed	multidrug	antibiotic efflux	1	5
QepA2	Unassessed	fluoroquinolone antibiotic	antibiotic efflux	1	0
tet(W/N/W)	Unassessed	tetracycline antibiotic	antibiotic target protection	1	0
adeI	Unassessed	multidrug	antibiotic efflux	1	3
CfxA6	Unassessed	Beta-lactams	antibiotic inactivation	1	0
tet(40)	Unassessed	tetracycline antibiotic	antibiotic efflux	1	0
BUT-1	Unassessed	Beta-lactams	antibiotic inactivation	1	0
APH(3')-IIa	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	1	1
Erm(38)	Unassessed	MLS	antibiotic target alteration	1	0
lsaC	Unassessed	multidrug	antibiotic target protection	1	0
FosA3	Unassessed	fosfomycin	antibiotic inactivation	1	0
Mrx	Unassessed	multidrug	antibiotic efflux	0	2
vatF	Unassessed	MLS	antibiotic inactivation	0	1
dfrF	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	0	1
acrS	Unassessed	multidrug	antibiotic efflux	0	2
CarO	Unassessed	Beta-lactams	antibiotic inactivation	0	1
LCR-1	Unassessed	Beta-lactams	antibiotic inactivation	0	2
nps-1	Unassessed	Beta-lactams	antibiotic inactivation	0	1
emrY	Unassessed	tetracycline antibiotic	antibiotic efflux	0	1
gadW	Unassessed	multidrug	antibiotic efflux	0	2
lnuC	Unassessed	MLS	antibiotic inactivation	0	1
ANT(3")-IIc	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	0	2

subclass	risk rank	ARG Class	resistance mechanism	read number	
				urban	background
QnrVC5	Unassessed	fluoroquinolone antibiotic	antibiotic target protection	0	2
mphG	Unassessed	MLS	antibiotic inactivation	0	1
QnrS5	Unassessed	fluoroquinolone antibiotic	antibiotic target protection	0	2
QnrVC1	Unassessed	fluoroquinolone antibiotic	antibiotic target protection	0	2
mecB	Unassessed	Beta-lactams	antibiotic target replacement	0	1
ANT(3")-IIa	Unassessed	aminoglycoside antibiotic	antibiotic inactivation	0	1
dfrD	Unassessed	diaminopyrimidine antibiotic	antibiotic target replacement	0	1
ADC-43	Unassessed	Beta-lactams	antibiotic inactivation	0	1
tetB(48)	Unassessed	tetracycline antibiotic	antibiotic efflux	0	1

References

Agler, M. T., et al. (2016). Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLOS Biology,* 14(1), e1002352.

Albuquerque, L. and M. S. da Costa (2014). The Family Gaiellaceae. The Prokaryotes: Actinobacteria. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt and F. Thompson. Berlin, Heidelberg, Springer Berlin Heidelberg: 357-360.

Allen, H. K., et al. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol,* 8(4), 251-259.

Aller, J. Y., et al. (2005). The sea surface microlayer as a source of viral and bacterial enrichment in marine aerosols. *Journal of Aerosol Science,* 36(5), 801-812.

Alneberg, J., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat Methods,* 11(11), 1144-1146.

Amann, R. I., et al. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews,* 59(1), 143-169.

Amato, P., et al. (2007). An important oceanic source of micro-organisms for cloud water at the Puy de Dôme (France). *Atmospheric Environment,* 41(37), 8253-8263.

Andersson, A. F., et al. (2008). Comparative Analysis of Human Gut Microbiota by Barcoded Pyrosequencing. *PLOS ONE,* 3(7), e2836.

Andreeva, I., et al. (2002). Biogenic component of atmospheric aerosol in the south of west siberia, Chem. *Sust. Dev,* 10, 523-537.

Andreeva, I. S., et al. (2001). Seasonal variations in the microorganisms concentration in the biogenic component of atmospheric aerosol in the South of Western Siberia. *Journal of Aerosol Science,* 32, 369-376.

Andreote, F. D., et al. (2009). Culture-Independent Assessment of Rhizobiales-Related Alphaproteobacteria and the Diversity of Methylobacterium in the Rhizosphere and Rhizoplane of Transgenic Eucalyptus. *Microbial Ecology,* 57(1), 82-93.

Ansari, T. U., et al. (2015). Model simulations of fungal spore distribution over the Indian region. *Atmospheric Environment,* 122, 552-560.

Antipov, D., et al. (2019). Plasmid detection and assembly in genomic and metagenomic data sets. *Genome Res,* 29(6), 961-968.

Arango-Argoty, G., et al. (2018). DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome,* 6(1), 23.

Årdal, C., et al. (2020). Antibiotic development — economic, regulatory and societal challenges. *Nature Reviews Microbiology,* 18(5), 267-274.

Ariya, P. A. and M. Amyot (2004). New Directions: The role of bioaerosols in atmospheric chemistry and physics. *Atmospheric Environment,* 38(8), 1231-1232.

Arndt, D., et al. (2012). METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res,* 40(Web Server issue), W88-95.

Ashley, S. L., et al. (2020). Lung and gut microbiota are altered by hyperoxia and contribute to oxygen-induced lung injury in mice. *Science Translational Medicine,* 12(556), eaau9959.

Asnicar, F., et al. (2020). Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nature Communications,* 11(1), 2500.

Bahram, M., et al. (2018). Structure and function of the global topsoil microbiome. *Nature,* 560(7717), 233-237.

Bahram, M., et al. (2016). Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *Isme j,* 10(4), 885-896.

Bakutis, B., et al. (2004). Analyses of airborne contamination with bacteria, endotoxins and dust in livestock barns and poultry houses. *Acta Veterinaria Brno,* 73(2), 283-289.

Baltz, R. (2005). Antibiotic discovery from actinomycetes: will a renaissance follow the decline and fall? *Sim News,* 55, 186-196.

Baltz, R. H. (2008). Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol,* 8(5), 557-563.

Banerjee, S., et al. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology,* 16(9), 567-576.

Baquero, F., et al. (2008). Antibiotics and antibiotic resistance in water environments. *Current Opinion in Biotechnology,* 19(3), 260-265.

Bär, M., et al. (2002). Modelling the survival of bacteria in drylands: the advantage of being dormant. *Proc Biol Sci,* 269(1494), 937-942.

Barabási, A.-L. (2009). Scale-free networks: a decade and beyond. *Science,* 325(5939), 412-413.

Barabási, A.-L. and R. Albert (1999). Emergence of scaling in random networks. *Science,* 286(5439), 509-512.

Barberán, A., et al. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal,* 6(2), 343-351.

Barberán, A., et al. (2015). Continental-scale distributions of dust-associated bacteria and fungi. *Proceedings of the National Academy of Sciences,* 112(18), 5756-5761.

Bauer, H., et al. (2002a). Bacteria and fungi in aerosols generated by two different types of wastewater treatment plants. *Water Res,* 36(16), 3965-3970.

Bauer, H., et al. (2002b). The contribution of bacteria and fungal spores to the organic carbon content of cloud water, precipitation and aerosols. *Atmospheric Research,* 64(1), 109-119.

Belnap, J. and D. A. Gillette (1998). Vulnerability of desert biological soil crusts to wind erosion: the influences of crust development, soil texture, and disturbance. *Journal of Arid Environments,* 39(2), 133-142.

Bengtsson, B., et al. (2017). High occurrence of mecC-MRSA in wild hedgehogs (Erinaceus europaeus) in Sweden. *Veterinary Microbiology,* 207, 103-107.

Berendonk, T. U., et al. (2015). Tackling antibiotic resistance: the environmental framework. *Nature Reviews Microbiology,* 13(5), 310-317.

Berglen, T. F., et al. (2004). A global model of the coupled sulfur/oxidant chemistry in the troposphere: The sulfur cycle. *Journal of Geophysical Research: Atmospheres,* 109(D19).

Berry, D. and S. Widder (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology,* 5.

Bertolini, V., et al. (2013). Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy. *Appl Microbiol Biotechnol,* 97(14), 6561-6570.

Bhullar, K., et al. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLOS ONE,* 7(4), e34953.

Biswas, R., et al. (2021). Overview on the role of heavy metals tolerance on developing antibiotic resistance in both Gram-negative and Gram-positive bacteria. *Archives of Microbiology,* 203(6), 2761-2770.

Blais-Lecours, P., et al. (2015). Non-culturable bioaerosols in indoor settings: Impact on health and molecular approaches for detection. *Atmospheric Environment,* 110, 45-53.

Blanchard, D. C. (1975). Bubble scavenging and the water-to-air transfer of organic material in the sea. Applied Chemistry at Protein Interfaces, AMERICAN CHEMICAL SOCIETY. 145: 360-387.

Blanchard, D. C., et al. (1981). Bubble scavenging of bacteria in freshwater quickly produces bacterial enrichment in airborne jet drops1. *Limnology and Oceanography,* 26(5), 961-964.

Bolotin, E. and R. Hershberg (2017). Horizontally acquired genes are often shared between closely related bacterial species. *Front Microbiol,* 8, 1536.

Bolyen, E., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol,* 37(8), 852-857.

Boreson, J., et al. (2004). Correlating bioaerosol load with PM2.5 and PM10cf concentrations: a comparison between natural desert and urban-fringe aerosols. *Atmospheric Environment,* 38(35), 6029-6041.

Bovallius, A., et al. (1980). Long-range transmission of bacteria. *Ann N Y Acad Sci,* 353, 186-200.

Bowers, R. M., et al. (2013). Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere. *Environmental Science & Technology,* 47(21), 12097-12106.

Bowers, R. M., et al. (2009). Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. *Appl Environ Microbiol,* 75(15), 5121-5130.

Bowers, R. M., et al. (2012). Seasonal variability in airborne bacterial communities at a high-elevation site. *Atmospheric Environment,* 50, 41-49.

Bowers, R. M., et al. (2011a). Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *Isme j,* 5(4), 601-612.

Bowers, R. M., et al. (2011b). Sources of bacteria in outdoor air across cities in the midwestern United States. *Appl Environ Microbiol,* 77(18), 6350-6356.

Bradford, M. A., et al. (2008). Thermal adaptation of soil microbial respiration to elevated temperature. *Ecol Lett,* 11(12), 1316-1327.

Breza-Boruta, B. (2016). The assessment of airborne bacterial and fungal contamination emitted by a municipal landfill site in Northern Poland. *Atmospheric Pollution Research,* 7(6), 1043-1052.

Brodie, E. L., et al. (2007). Urban aerosols harbor diverse and dynamic bacterial populations. *Proceedings of the National Academy of Sciences,* 104(1), 299-304.

Brown, J. K. M. and M. S. Hovmøller (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science,* 297(5581), 537-541.

Brown, M. G. and D. L. Balkwill (2009). Antibiotic resistance in bacteria isolated from the deep terrestrial subsurface. *Microb Ecol,* 57(3), 484-493.

Buchfink, B., et al. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat Methods,* 12(1), 59-60.

Burrows, S. M., et al. (2009b). Bacteria in the global atmosphere – Part 2: Modeling of emissions and transport between different ecosystems. *Atmos. Chem. Phys.,* 9(23), 9281-9297.

Burrows, S. M., et al. (2009a). Bacteria in the global atmosphere – Part 1: Review and synthesis of literature data for different ecosystems. *Atmos. Chem. Phys.,* 9(23), 9263-9280.

Burrows, S. M., et al. (2013). Estimating bacteria emissions from inversion of atmospheric transport: sensitivity to modelled particle characteristics. *Atmos. Chem. Phys.,* 13(11), 5473-5488.

Bush, K. and G. A. Jacoby (2010). Updated functional classification of beta-lactamases. *Antimicrob Agents Chemother,* 54(3), 969-976.

Cáliz, J., et al. (2018). A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences,* 115(48), 12229-12234.

Cao, C., et al. (2014). Inhalable microorganisms in Beijing's PM2. 5 and PM10 pollutants during a severe smog event. *Environmental Science & Technology,* 48(3), 1499-1507.

Caporaso, J. G., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal,* 6(8), 1621-1624.

Casabianca, S., et al. (2013). Quantification of the Toxic Dinoflagellate Ostreopsis spp. by qPCR Assay in Marine Aerosol. *Environmental Science & Technology,* 47(8), 3788-3795.

Chaumeil, P. A., et al. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics,* 36(6), 1925-1927.

Chen, B., et al. (2013). Metagenomic profiles of antibiotic resistance genes (ARGs) between human impacted estuary and deep ocean sediments. *Environ Sci Technol,* 47(22), 12753-12760.

Chen, B., et al. (2016). Metagenomic analysis revealing antibiotic resistance genes (ARGs) and their genetic compartments in the Tibetan environment. *Environ Sci Technol,* 50(13), 6670-6679.

Chen, L., et al. (2005). VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res,* 33(suppl_1), D325-D328.

Chen, P.-S. and C.-S. Li (2005). Bioaerosol characterization by flow cytometry with fluorochrome. *Journal of Environmental Monitoring,* 7(10), 950-959.

Chen, P., et al. (2022). Antibiotic resistance genes in bioaerosols: Emerging, non-ignorable and pernicious pollutants. *Journal of Cleaner Production,* 348, 131094.

Chen, R., et al. (2017). Fine particulate air pollution and daily mortality. A nationwide analysis in 272 Chinese cities. *Am J Respir Crit Care Med,* 196(1), 73-81.

Chen, S., et al. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics,* 34(17), i884-i890.

Cheng, Y. S. (2003). Aerosol Deposition in the Extrathoracic Region. *Aerosol Sci Technol,* 37(8), 659-671.

Chmielowiec-Korzeniowska, A., et al. (2020). Staphylococcus aureus carriage state in healthy adult population and phenotypic and genotypic properties of isolated strains. *Postepy Dermatol Alergol,* 37(2), 184-189.

Chng, K. R., et al. (2020). Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. *Nature Medicine,* 26(6), 941-951.

Cho, B. C. and C. Y. Hwang (2011). Prokaryotic abundance and 16S rRNA gene sequences detected in marine aerosols on the East Sea (Korea). *FEMS Microbiol Ecol,* 76(2), 327-341.

Cho, S.-H., et al. (2005). Aerodynamic characteristics and respiratory deposition of fungal fragments. *Atmospheric Environment,* 39(30), 5454-5465.

Chow, J. C., et al. (1993). The dri thermal/optical reflectance carbon analysis system: description, evaluation and applications in U.S. Air quality studies. *Atmospheric Environment. Part A. General Topics,* 27(8), 1185-1201.

Christie, G. and P. Setlow (2020). Bacillus spore germination: Knowns, unknowns and what we need to learn. *Cellular Signalling,* 74, 109729.

Christner, B. C., et al. (2008). Ubiquity of biological ice nucleators in snowfall. *Science,* 319(5867), 1214.

Clauß, M. (2015). Particle size distribution of airborne microorganisms in the environment-a review. *Landbauforsch Appl Agric Forestry Res,* 65(2), 77-100.

Cohen, A. J., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet,* 389(10082), 1907-1918.

Comte, J., et al. (2016). Co-occurrence patterns in aquatic bacterial communities across changing permafrost landscapes. *Biogeosciences,* 13(1), 175-190.

Condit, R., et al. (2002). Beta-diversity in tropical forest trees. *Science,* 295(5555), 666-669.

Cottee-Jones, et al. (2012). Perspective: The keystone species concept: a critical appraisal. *Frontiers of Biogeography,* 4(3), 117-127.

Crawford, I., et al. (2016). Observations of fluorescent aerosol–cloud interactions in the free troposphere at the High-Altitude Research Station Jungfraujoch. *Atmos. Chem. Phys.*, 16(4), 2273-2284.

Crook, B., et al. (1991). Airborne dust, ammonia, microorganisms, and antigens in pig confinement houses and the respiratory health of exposed farm workers. *Am Ind Hyg Assoc J,* 52(7), 271-279.

Cruz, P. and M. P. Buttner (2016). Analysis of Bioaerosol Samples. Manual of Environmental Microbiology: 3.2.3-1-3.2.3-9.

Cuadrat, R. R. C., et al. (2020). Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. *Gigascience,* 9(5).

Curtis, T. P., et al. (2002). Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci U S A,* 99(16), 10494-10499.

Czekalski, N., et al. (2015). Does human activity impact the natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss lakes. *Environment International,* 81, 45-55.

D'Costa, V. M., et al. (2006). Sampling the antibiotic resistome. *Science,* 311(5759), 374-377.

D'Costa, V. M., et al. (2011). Antibiotic resistance is ancient. *Nature,* 477(7365), 457-461.

Danko, D., et al. (2021). A global metagenomic map of urban microbiomes and antimicrobial resistance. *Cell,* 184(13), 3376-3393.e3317.

Dantas, G., et al. (2008). Bacteria subsisting on antibiotics. *Science,* 320(5872), 100-103.

Das, B., et al. (2022). HT-ARGfinder: A comprehensive pipeline for identifying horizontally transferred antibiotic resistance genes and directionality in metagenomic sequencing data. *Frontiers in Environmental Science,* 10.

Davey, H. M. (2003). Flow cytometric techniques for the detection of microorganisms. Advanced Flow Cytometry: Applications in Biological Research. R. C. Sobti and A. Krishan. Dordrecht, Springer Netherlands: 91-97.

Davies, J. and D. Davies (2010). Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev,* 74(3), 417-433.

De Varigny, H. (1883). Les Organismes vivants de l'Atmosphère. *Nature,* 28(708), 76-77.

DeLeon-Rodriguez, N., et al. (2013). Microbiome of the upper troposphere: Species composition and prevalence, effects of tropical storms, and atmospheric implications. *Proceedings of the National Academy of Sciences,* 110(7), 2575-2580.

Deng, W., et al. (2016a). Distribution of bacteria in inhalable particles and its implications for health risks in kindergarten children in Hong Kong. *Atmospheric Environment,* 128, 268-275.

Deng, Y., et al. (2012). Molecular ecological network analyses. *BMC Bioinformatics,* 13(1), 113.

Després, V., et al. (2012). Primary biological aerosol particles in the atmosphere: a review. *Tellus B: Chemical and Physical Meteorology,* 64(1), 15598.

Després, V. R., et al. (2007). Characterization of primary biogenic aerosol particles in urban, rural, and high-alpine air by DNA sequence and restriction fragment analysis of ribosomal RNA genes. *Biogeosciences,* 4(6), 1127-1141.

Dickson, R. P. and G. B. Huffnagle (2015). The lung microbiome: new principles for respiratory bacteriology in health and disease. *PLoS Pathog,* 11(7), e1004923.

Dijkshoorn, L., et al. (2007). An increasing threat in hospitals: multidrug-resistant Acinetobacter baumannii. *Nat Rev Microbiol,* 5(12), 939-951.

Drautz-Moses, D. I., et al. (2022). Vertical stratification of the air microbiome in the lower troposphere. *Proceedings of the National Academy of Sciences,* 119(7), e2117293119.

Du, W., et al. (2017). Simultaneous measurements of particle number size distributions at ground level and 260 m on a meteorological tower in urban Beijing, China. *Atmos. Chem. Phys.,* 17(11), 6797-6811.

Dubuis, M.-E., et al. (2017). Bioaerosols concentrations in working areas in biomethanization facilities. *Journal of the Air & Waste Management Association,* 67(11), 1258-1271.

Dutil, S., et al. (2008). Measurement of airborne bacteria and endotoxin generated during dental cleaning. *Journal of occupational and environmental hygiene,* 6(2), 121-130.

Eakins, B. W. and G. F. Sharman Volumes of the World's Oceans from ETOPO1, NOAA National Geoghysical Data Center.(2010).

Echeverria-Palencia, C. M., et al. (2017). Disparate antibiotic resistance gene quantities revealed across 4 major cities in California: a survey in drinking water, air, and soil at 24 public parks. *ACS Omega,* 2(5), 2255-2263.

Elbert, W., et al. (2007). Contribution of fungi to primary biogenic aerosols in the atmosphere: wet and dry discharged spores, carbohydrates, and inorganic ions. *Atmos. Chem. Phys.*, 7(17), 4569-4588.

Eldridge, D. J. and J. F. Leys (2003). Exploring some relationships between biological soil crusts, soil aggregation and wind erosion. *Journal of Arid Environments,* 53(4), 457-466.

Enright, M. C., et al. (2002). The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *Proc Natl Acad Sci U S A,* 99(11), 7687-7692.

Fahlgren, C., et al. (2015). Seawater mesocosm experiments in the Arctic uncover differential transfer of marine bacteria to aerosols. *Environmental Microbiology Reports,* 7(3), 460-470.

Fajardo, A., et al. (2008). The neglected intrinsic resistome of bacterial pathogens. *PLOS ONE,* 3(2), e1619.

Fan, Y., et al. (2014). Heavy metal and antibiotic resistance in bacteria isolated from the environment of swine farms. *Journal of the Chemical Society of Pakistan,* 36(2).

Fang, Z., et al. (2007). Culturable airborne bacteria in outdoor environments in Beijing,China. *Microb Ecol,* 54(3), 487-496.

Fauset, S., et al. (2015). Hyperdominance in Amazonian forest carbon cycling. *Nature Communications*, 6(1), 6857.

Faust, K. and J. Raes (2012). Microbial interactions: from networks to models. *Nature Reviews Microbiology,* 10(8), 538-550.

Fernstrom, A. and M. Goldblatt (2013). Aerobiology and its role in the transmission of infectious diseases. *J Pathog,* 2013, 493960.

Fierer, N. and R. B. Jackson (2006). The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A,* 103(3), 626-631.

Fierer, N., et al. (2008). Short-Term Temporal Variability in Airborne Bacterial and Fungal Populations. *Appl Environ Microbiol*, 74(1), 200-207.

Fleming, A. (1929). On the antibacterial action of cultures of a Penicillium, with special reference to their use in the isolation of B. influenzæ. *The Bristish Journal of Experiemntal Pathology,* 10(3), 226-236.

Fouladi, F., et al. (2020). Air pollution exposure is associated with the gut microbiome as revealed by shotgun metagenomic sequencing. *Environment International,* 138, 105604.

Franchitti, E., et al. (2020). Methods for Bioaerosol Characterization: Limits and Perspectives for Human Health Risk Assessment in Organic Waste Treatment. *Atmosphere,* 11(5), 452.

Franzetti, A., et al. (2011). Seasonal variability of bacteria in fine and coarse urban air particulate matter. *Appl Microbiol Biotechnol,* 90(2), 745-753.

Friedl, M. A., et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment,* 83(1), 287-302.

Fröhlich-Nowoisky, J., et al. (2016). Bioaerosols in the Earth system: Climate, health, and ecosystem interactions. *Atmospheric Research,* 182, 346-376.

Fu, L., et al. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics,* 28(23), 3150-3152.

Fulton, J. D. (1966). Microorganisms of the upper atmosphere. *Applied Microbiology,* 14(2), 237-240.

Gabey, A. M., et al. (2013). Observations of fluorescent and biological aerosol at a high-altitude site in central France. *Atmos. Chem. Phys.,* 13(15), 7415-7428.

Gadd, G. M. (2010). Metals, minerals and microbes: geomicrobiology and bioremediation. *Microbiology (Reading),* 156(Pt 3), 609-643.

Gandolfi, I., et al. (2015). Spatio-temporal variability of airborne bacterial communities and their correlation with particulate matter chemical composition across two urban areas. *Appl Microbiol Biotechnol,* 99(11), 4867-4877.

Gandolfi, I., et al. (2011). Antibiotic resistance in bacteria associated with coarse atmospheric particulate matter in an urban area. *J Appl Microbiol,* 110(6), 1612-1620.

Gao, J.-F., et al. (2017a). Airborne bacterial communities of PM2.5 in Beijing-Tianjin-Hebei megalopolis, China as revealed by Illumina MiSeq Sequencing: a case study. *Aerosol and Air Quality Research,* 17(3), 788-798.

Gao, M., et al. (2015). Seasonal size distribution of airborne culturable bacteria and fungi and preliminary estimation of their deposition in human lungs during non-haze and haze days. *Atmospheric Environment,* 118, 203-210.

Gao, M., et al. (2017b). Size-related bacterial diversity and tetracycline resistance gene abundance in the air of concentrated poultry feeding operations. *Environmental Pollution,* 220, 1342-1348.

García-Álvarez, L., et al. (2011). Meticillin-resistant Staphylococcus aureus with a novel mecA homologue in human and bovine populations in the UK and Denmark: a descriptive study. *The Lancet Infectious Diseases,* 11(8), 595-603.

Gaston, K. J., et al. (2000). Abundance–occupancy relationships. *Journal of Applied Ecology,* 37(s1), 39-59.

Gat, D., et al. (2017). Origin-dependent aariations in the atmospheric microbiome community in eastern Mediterranean dust storms. *Environmental Science & Technology,* 51(12), 6709-6718.

Geng, H., et al. (2016). Changes in the structure of the microbial community associated with Nannochloropsis salina following treatments with antibiotics and bioactive compounds. *Front Microbiol,* 7, 1155.

Genitsaris, S., et al. (2017). Variability of airborne bacteria in an urban Mediterranean area (Thessaloniki, Greece). *Atmospheric Environment,* 157, 101-110.

Ghosh, B., et al. (2022). Bioaerosol and its impact on human health. Airborne Particulate Matter: Source, Chemistry and Health. S. Sonwani and A. Shukla. Singapore, Springer Nature Singapore: 167-193.

Gibbs, S. G., et al. (2006). Isolation of antibiotic-resistant bacteria from the air plume downwind of a swine confined or concentrated animal feeding operation. *Environ Health Perspect,* 114(7), 1032-1037.

Gilbert, J. A., et al. (2012). Defining seasonal marine microbial community dynamics. *The ISME Journal,* 6(2), 298-308.

Gilbert, Y., et al. (2010). Airborne bacteria and antibiotic resistance genes in hospital rooms. *Aerobiologia,* 26(3), 185-194.

Gillings, M. R., et al. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *Isme j,* 9(6), 1269-1279.

Glaeser, S. P. and P. Kämpfer (2014). The Family Sphingomonadaceae. The Prokaryotes: Alphaproteobacteria and Betaproteobacteria. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt and F. Thompson. Berlin, Heidelberg, Springer Berlin Heidelberg: 641-707.

Glassing, A., et al. (2016). Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathogens,* 8(1), 24.

Glickman, M. E. and D. A. v. Dyk (2007). Basic Bayesian Methods. Totowa, NJ, Humana Press.

Gokul, J. K., et al. (2016). Taxon interactions control the distributions of cryoconite bacteria colonizing a High Arctic ice cap. *Mol Ecol,* 25(15), 3752-3767.

Gong, Z., et al. (2018). Optical trapping-Raman spectroscopy (OT-RS) with embedded microscopy imaging for concurrent characterization and monitoring of physical and chemical properties of single particles. *Analytica Chimica Acta,* 1020, 86-94.

Gou, H., et al. (2016). Assessment of microbial communities in PM1 and PM10 of Urumqi during winter. *Environ Pollut,* 214, 202-210.

Gou, H., et al. (2017). Assessment of microbial communities in TSP and PM10 of Shihezi during spring. *Chinese Journal of Environmental Engineering,* 11(4), 2343-2349.

Graham, B., et al. (2003). Composition and diurnal variability of the natural Amazonian aerosol. *Journal of Geophysical Research: Atmospheres,* 108(D24).

Graham, E. B., et al. (2017). Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. *Environ Microbiol,* 19(4), 1552-1567.

Griffin, D. W. (2007). Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin Microbiol Rev,* 20(3), 459-477, table of contents.

Griffin, D. W., et al. (2018). 2.3. Global-scale atmospheric dispersion of microorganisms.

Griffin, D. W., et al. (2006). Airborne microorganisms in the African desert dust corridor over the mid-Atlantic ridge, Ocean Drilling Program, Leg 209. *Aerobiologia,* 22(3), 211-226.

Gruber, S., et al. (1998). Vertical distribution of biological aerosol particles above the North Sea. *Journal of Aerosol Science,* 29, S771-S772.

Gunde-Cimerman, N., et al. (2018). Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol Rev,* 42(3), 353-375.

Gusareva, E. S., et al. (2019). Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proceedings of the National Academy of Sciences,* 116(46), 23299-23308.

Haiko, J. and B. Westerlund-Wikström (2013). The role of the bacterial flagellum in adhesion and virulence. *Biology (Basel),* 2(4), 1242-1267.

Hall, B. G. and M. Barlow (2004). Evolution of the serine β-lactamases: past, present and future. *Drug Resistance Updates,* 7(2), 111-123.

Hallar, A. G., et al. (2011). Atmospheric bioaerosols transported via dust storms in the western United States. *Geophysical Research Letters,* 38(17).

Hamady, M., et al. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods,* 5(3), 235-237.

Hannah Ritchie and M. Roser (2019). Outdoor Air Pollution. *Our World in Data*.

Hantsch, L., et al. (2013). Species richness and species identity effects on occurrence of foliar fungal pathogens in a tree diversity experiment. *Ecosphere,* 4(7), art81.

Hara, K. and D. Zhang (2012). Bacterial abundance and viability in long-range transported dust. *Atmospheric Environment,* 47, 20-25.

Harrison, R. M., et al. (2005). Climate factors influencing bacterial count in background air samples. *Int J Biometeorol,* 49(3), 167-178.

He, T., et al. (2021). Intracellular and Extracellular Antibiotic Resistance Genes in Airborne PM2.5 for Respiratory Exposure in Urban Areas. *Environmental Science & Technology Letters,* 8(2), 128-134.

Heald, C. L. and D. V. Spracklen (2009). Atmospheric budget of primary biological aerosol particles from fungal spores. *Geophysical Research Letters,* 36(9).

Hendriksen, R. S., et al. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications,* 10(1), 1124.

Herlemann, D. P. R., et al. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME Journal,* 5(10), 1571-1579.

Hernando-Amado, S., et al. (2019). Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nature Microbiology,* 4(9), 1432-1442.

Hill, R., et al. (2016). Temporal and spatial influences incur reconfiguration of Arctic heathland soil bacterial community structure. *Environ Microbiol,* 18(6), 1942-1953.

Hill, T. C. J., et al. (2014). Measurement of ice nucleation-active bacteria on plants and in precipitation by quantitative PCR. *Appl Environ Microbiol,* 80(4), 1256-1267.

Hoose, C., et al. (2010). How important is biological ice nucleation in clouds on a global scale? *Environmental Research Letters,* 5(2), 024009.

Horneck, G., et al. (1994). Long-term survival of bacterial spores in space. *Advances in Space Research,* 14(10), 41-45.

Horneck, G., et al. (2010). Space microbiology. *Microbiol Mol Biol Rev,* 74(1), 121-156.

Horrevorts, A. M., et al. (1990). Ecology of Pseudomonas aeruginosa in patients with cystic fibrosis. *J Med Microbiol,* 31(2), 119-124.

Hospodsky, D., et al. (2012). Human occupancy as a source of indoor airborne bacteria. *PLOS ONE,* 7(4), e34867.

Hu, J., et al. (2018). Metagenomic profiling of ARGs in airborne particulate matters during a severe smog event. *Sci Total Environ,* 615, 1332-1340.

Huffman, J. A., et al. (2013). High concentrations of biological aerosol particles and ice nuclei during and after rain. *Atmos. Chem. Phys.,* 13(13), 6151-6164.

Hultin, K. A. H., et al. (2011). Aerosol and bacterial emissions from Baltic Seawater. *Atmospheric Research,* 99(1), 1-14.

Hummel, M., et al. (2015). Regional-scale simulations of fungal spore aerosols using an emission parameterization adapted to local measurements of fluorescent biological aerosol particles. *Atmos. Chem. Phys.,* 15(11), 6127-6146.

Humphries, M. D. and K. Gurney (2008). Network 'Small-World-Ness': a quantitative method for determining canonical network equivalence. *PLOS ONE,* 3(4), e0002051.

Hyatt, D., et al. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics,* 11(1), 119.

Imshenetsky, A. A., et al. (1978). Upper boundary of the biosphere. *Appl Environ Microbiol,* 35(1), 1-5.

Innocente, E., et al. (2017). Influence of seasonality, air mass origin and particulate matter chemical composition on airborne bacterial community structure in the Po Valley, Italy. *Sci Total Environ,* 593-594, 677-687.

Jackson, R. B., et al. (1997). A global budget for fine root biomass, surface area, and nutrient contents. *Proceedings of the National Academy of Sciences,* 94(14), 7362-7366.

Jaenicke, R. (2005). Abundance of cellular material and proteins in the atmosphere. *Science,* 308(5718), 73-73.

Jalili, D., et al. (2021). Assessment of Airborne Bacterial and Fungal Communities in Shahrekord Hospitals. *Journal of Environmental and Public Health,* 2021, 8864051.

Jari Oksanen, et al. (2018). vegan: Community Ecology Package. R package, Version 2.0-4. Available at http://CRAN.R-project.org/package=vegan.

Ji, Y., et al. (2016). Structure and function of methanogenic microbial communities in sediments of Amazonian lakes with different water types. *Environ Microbiol,* 18(12), 5082-5100.

Jiang, W., et al. (2015). Optimized DNA extraction and metagenomic sequencing of airborne microbial communities. *Nature Protocols,* 10(5), 768-779.

Jiang, Y., et al. (2017). Plant cultivars imprint the rhizosphere bacterial community composition and association networks. *Soil Biology and Biochemistry*, 109, 145-155.

Jin, L., et al. (2016). Airborne particulate matter pollution in urban China: a chemical mixture perspective from sources to impacts. *National Science Review*, 4(4), 593-610.

Jones, A. M. and R. M. Harrison (2004). The effects of meteorological factors on atmospheric bioaerosol concentrations—a review. *Science of The Total Environment*, 326(1), 151-180.

Junker, B. H. (2008). Networks in Biology. Analysis of Biological Networks: 1-14.

Kallmeyer, J., et al. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *Proceedings of the National Academy of Sciences*, 109(40), 16213-16216.

Kan, H., et al. (2012). Ambient air pollution, climate change, and population health in China. *Environ Int*, 42, 10-19.

Kang, D. D., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 7, e7359.

Kang, K., et al. (2018). The environmental exposures and inner- and intercity traffic flows of the metro system may contribute to the skin microbiome and resistome. *Cell Rep*, 24(5), 1190-1202.e1195.

Kellogg, C. A. and D. W. Griffin (2006). Aerobiology and the global transport of desert dust. *Trends Ecol Evol*, 21(11), 638-644.

Kembel, S. W., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26(11), 1463-1464.

Kim, D., et al. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res*, 26(12), 1721-1729.

Klein, A. M., et al. (2016). Molecular Evidence for Metabolically Active Bacteria in the Atmosphere. *Frontiers in Microbiology*, 7.

Klein, E. Y., et al. (2018). Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc Natl Acad Sci U S A*, 115(15), E3463-e3470.

Knights, D., et al. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nat Methods,* 8(9), 761-763.

Kourtev, P. S., et al. (2011). Atmospheric cloud water contains a diverse bacterial community. *Atmospheric Environment,* 45(30), 5399-5405.

Krawczyk, P. S., et al. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res,* 46(6), e35.

Kruse, H. and H. Sørum (1994). Transfer of multiple drug resistance plasmids between bacteria of diverse origins in natural microenvironments. *Appl Environ Microbiol,* 60(11), 4015-4021.

Landrigan, P. J. (2017). Air pollution and health. *The Lancet Public Health,* 2(1), e4-e5.

Lang-Yona, N., et al. (2014). Marine aerosol as a possible source for endotoxins in coastal areas. *Science of The Total Environment,* 499, 311-318.

Langfelder, P. and S. Horvath (2012). Fast R functions for Robust correlations and Hierarchical clustering. *J Stat Softw,* 46(11).

Langmead, B. and S. L. Salzberg (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9(4), 357-359.

Laxminarayan, R., et al. (2013). Antibiotic resistance—the need for global solutions. *The Lancet Infectious Diseases,* 13(12), 1057-1098.

Layeghifard, M., et al. (2017). Disentangling interactions in the microbiome: a network perspective. *Trends Microbiol,* 25(3), 217-228.

Leck, C. and E. K. Bigg (2005). Biogenic particles in the surface microlayer and overlaying atmosphere in the central Arctic Ocean during summer. *Tellus B: Chemical and Physical Meteorology,* 57(4), 305-316.

Lee, S. H., et al. (2010). Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. *Sci Total Environ,* 408(6), 1349-1357.

Lengke, M. and G. Southam (2006). Bioaccumulation of gold by sulfate-reducing bacteria cultured in the presence of gold(I)-thiosulfate complex. *Geochimica et Cosmochimica Acta,* 70(14), 3646-3661.

Li, B., et al. (2015). Profile and fate of bacterial pathogens in sewage treatment plants revealed by high-throughput metagenomic approach. *Environmental Science & Technology,* 49(17), 10492-10502.

Li, D., et al. (2016b). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods,* 102, 3-11.

Li, D., et al. (2020). Surf zone microbiological water quality following emergency beach nourishment using sediments from a catastrophic debris flow. *Water Res,* 176, 115733.

Li, D., et al. (2021a). Limited bacterial removal in full-scale stormwater biofilters as evidenced by community sequencing analysis. *Environmental Science & Technology,* 55(13), 9199-9208.

Li, J., et al. (2018). Global survey of antibiotic resistance genes in Air. *Environmental Science & Technology,* 52(19), 10975-10984.

Li, J. and K. Osada (2007). Preferential settling of elongated mineral dust particles in the atmosphere. *Geophysical Research Letters,* 34(17).

Li, J., et al. (2016a). Bioaerosol emissions and detection of airborne antibiotic resistance genes from a wastewater treatment plant. *Atmospheric Environment,* 124, 404-412.

Li, P., et al. (2021b). Characterization, factors, and UV reduction of airborne bacteria in a rural wastewater treatment station. *Science of The Total Environment,* 751, 141811.

Li, X., et al. (2019). Air pollution: a global problem needs local fixes. *Nature,* 570, 437-439.

Liang, J., et al. (2020b). Identification and quantification of bacterial genomes carrying antibiotic resistance genes and virulence factor genes for aquatic microbiological risk assessment. *Water Res,* 168, 115160.

Liang, Z., et al. (2020a). Pollution profiles of antibiotic resistance genes associated with airborne opportunistic pathogens from typical area, Pearl River Estuary and their exposure risk to human. *Environ Int,* 143, 105934.

Lighthart, B. (1997). The ecology of bacteria in the alfresco atmosphere. *FEMS Microbiology Ecology,* 23(4), 263-274.

Lindow, S. E. and M. T. Brandl (2003). Microbiology of the phyllosphere. *Appl Environ Microbiol,* 69(4), 1875-1883.

Lindsay, J. A. and M. T. Holden (2004). Staphylococcus aureus: superbug, super genome? *Trends Microbiol,* 12(8), 378-385.

Lindsley, W. G., et al. (2006). A two-stage cyclone using microcentrifuge tubes for personal bioaerosol sampling. *Journal of Environmental Monitoring,* 8(11), 1136-1142.

Ling, A. L., et al. (2013). Tetracycline resistance and Class 1 integron genes associated with indoor and outdoor aerosols. *Environ Sci Technol,* 47(9), 4046-4052.

Lis, D. O., et al. (2009). Methicillin resistance of airborne coagulase-negative staphylococci in homes of persons having contact with a hospital environment. *Am J Infect Control,* 37(3), 177-182.

Liu, B. and M. Pop (2009). ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res,* 37(Database issue), D443-447.

Locey, K. J. and J. T. Lennon (2016). Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences,* 113(21), 5970-5975.

Lomolino, M. V. and J. H. Brown (2009). The reticulating phylogeny of island biogeography theory. *Q Rev Biol,* 84(4), 357-390.

Lu, J., et al. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science,* 3, e104.

Lu, R., et al. (2018). Bacterial community structure in atmospheric particulate matters of different sizes during the haze days in Xi'an, China. *Sci Total Environ,* 637-638, 244-252.

Lupatini, M., et al. (2014). Network topology reveals high connectance levels and few key microbial genera within soils. *Frontiers in Environmental Science,* 2.

Ma, B., et al. (2016). Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *The ISME Journal,* 10(8), 1891-1901.

Ma, B., et al. (2020a). Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome,* 8(1), 82.

Ma, M., et al. (2020b). Characterization of bacterial communities during persistent fog and haze events in the Qingdao coastal region. *Frontiers of Environmental Science & Engineering,* 15(3), 42.

Machado, E. C., et al. (2023). Antibiotic resistance profile of wastewater treatment plants in Brazil reveals different patterns of resistance and multi resistant bacteria in final effluents. *Science of The Total Environment*, 857, 159376.

Madelin, T. M. (1994). Fungal aerosols: A review. *Journal of Aerosol Science*, 25(8), 1405-1412.

Maki, T., et al. (2011). Characterization of halotolerant and oligotrophic bacterial communities in Asian desert dust (KOSA) bioaerosol accumulated in layers of snow on Mount Tateyama, Central Japan. *Aerobiologia*, 27(4), 277-290.

Maki, T., et al. (2022). Long-range transport of airborne bacteria over East Asia: Asian dust events carry potentially nontuberculous Mycobacterium populations. *Environment International*, 168, 107471.

Manaka, A., et al. (2017). Comparison of 16S ribosomal RNA gene sequence analysis and conventional culture in the environmental survey of a hospital. *J Pharm Health Care Sci*, 3, 8.

Manara, S., et al. (2018). Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of Staphylococcus aureus strains in a paediatric hospital. *Genome Med*, 10(1), 82.

Marshall, B. M., et al. (2009). Commensals: underappreciated reservoir of antibiotic resistance. *Microbe Magazine*, 4, 231-238.

Martínez, J. L. (2008). Antibiotics and antibiotic resistance genes in natural environments. *Science*, 321(5887), 365-367.

Martinez, J. L. and F. Baquero (2000). Mutation frequencies and antibiotic resistance. *Antimicrobial Agents and Chemotherapy*, 44(7), 1771-1777.

Martinez, J. L., et al. (2009). A global view of antibiotic resistance. *FEMS Microbiol Rev*, 33(1), 44-65.

Martiny, J. B. H., et al. (2011). Drivers of bacterial β-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences*, 108(19), 7850-7854.

Marvasi, M., et al. (2017). Fitness of antibiotic-resistant bacteria in the environment: a laboratory activity. *J Microbiol Biol Educ*, 18(1).

Matthias-Maser, S., et al. (1999). The size distribution of marine atmospheric aerosol with regard to primary biological aerosol particles over the South Atlantic Ocean. *Atmospheric Environment*, 33(21), 3569-3575.

Matthias-Maser, S., et al. (2000a). Seasonal variation of primary biological aerosol particles in the remote continental region of Lake Baikal/Siberia. *Atmospheric Environment,* 34(22), 3805-3811.

Matthias-Maser, S., et al. (1995). Seasonal variation of primary biological aerosol particles. *Journal of Aerosol Science*, 26, S545-S546.

Matthias-Maser, S., et al. (2000b). Primary biological aerosol particles at the high alpine site of Jungfraujoch/Switzerland. *Journal of Aerosol Science,* 31, 955-956.

May, M. (2021). Tomorrow's biggest microbial threats. *Nature Medicine,* 27(3), 358-359.

Mayol, E., et al. (2017). Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nat Commun,* 8(1), 201.

McCluskey, C. S., et al. (2019). Numerical representations of marine ice-nucleating particles in remote marine environments evaluated against observations. *Geophysical Research Letters,* 46(13), 7838-7847.

McEachran, A. D., et al. (2015). Antibiotics, bacteria, and antibiotic resistance genes: aerial transport from cattle feed yards via particulate matter. *Environ Health Perspect,* 123(4), 337-343.

Menzel, U. and M. U. Menzel (2016). Package 'RMThreshold'.

Meola, M., et al. (2015). Bacterial Composition and Survival on Sahara Dust Particles Transported to the European Alps. *Frontiers in Microbiology,* 6.

Mériaux, A., et al. (2006). Bioaerosols in Peat Moss Processing Plants. *Journal of occupational and environmental hygiene*, 3(8), 408-417.

Mestrovic, T., et al. (2022). The burden of bacterial antimicrobial resistance in the WHO European region in 2019: a cross-country systematic analysis. *The Lancet Public Health,* 7(11), e897-e913.

Miao, J., et al. (2017). 16SPIP: a comprehensive analysis pipeline for rapid pathogen detection in clinical samples based on 16S metagenomic sequencing. *BMC Bioinformatics,* 18(16), 568.

Miller, F. J., et al. (1988). Influence of breathing mode and activity level on the regional deposition of inhaled particles and implications for regulatory standards. *The Annals of Occupational Hygiene,* 32(inhaled_particles_VI), 3-10.

Mittapally, S., et al. (2018). Metal ions as antibacterial agents. *Journal of Drug Delivery and Therapeutics*, 8, 411-419.

Morawska, L., et al. (2022). The physics of respiratory particle generation, fate in the air, and inhalation. *Nature Reviews Physics*.

Morlon, H., et al. (2011). Spatial patterns of phylogenetic diversity. *Ecology Letters*, 14(2), 141-149.

Morris, C. E. and L. L. Kinkel (2002). Fifty years of phyllosphere microbiology: significant contributions to research in related fields. *Phyllosphere microbiology*, 365-375.

Morris, C. E., et al. (2013). Movement of Bioaerosols in the Atmosphere and the Consequences for Climate and Microbial Evolution. Aerosol Science: 393-415.

Mulani, M. S., et al. (2019). Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: a review. *Frontiers in Microbiology*, 10.

Murata, K. and D. Zhang (2014). Transport of bacterial cells toward the Pacific in Northern Hemisphere westerly winds. *Atmospheric Environment*, 87, 138-145.

Murata, K. and D. Zhang (2016). Concentration of bacterial aerosols in response to synoptic weather and land-sea breeze at a seaside site downwind of the Asian continent. *Journal of Geophysical Research (Atmospheres)*, 121, 11,636-611,647.

Murray, C. J. L., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, 399(10325), 629-655.

Nazaries, L., et al. (2013). Methane, microbes and models: fundamental understanding of the soil methane cycle for future predictions. *Environ Microbiol*, 15(9), 2395-2417.

Neff, J. C., et al. (2008). Increasing eolian dust deposition in the western United States linked to human activity. *Nature Geoscience*, 1(3), 189-195.

Negron, A., et al. (2020). Using flow cytometry and light-induced fluorescence to characterize the variability and characteristics of bioaerosols in springtime in Metro Atlanta, Georgia. *Atmos. Chem. Phys.*, 20(3), 1817-1838.

Newman, D. J. and G. M. Cragg (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products,* 79(3), 629-661.

Ning, D., et al. (2019). A general framework for quantitatively assessing ecological stochasticity. *Proceedings of the National Academy of Sciences,* 116(34), 16892-16898.

Ning, D. and M. D. Ning (2021). Package 'NST'.

O'Dowd, C., et al. (2015). Connecting marine productivity to sea-spray via nanoscale biological processes: Phytoplankton Dance or Death Disco? *Scientific Reports,* 5(1), 14883.

Ochman, H., et al. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature,* 405(6784), 299-304.

Oh, M., et al. (2018). MetaCompare: a computational pipeline for prioritizing environmental resistome risk. *FEMS Microbiol Ecol,* 94(7).

Paine, R. T. (1966). Food web complexity and species diversity. *The American Naturalist,* 100, 65-75.

Pal, C., et al. (2016). The structure and diversity of human, animal and environmental resistomes. *Microbiome,* 4(1), 54.

Park, Y. and J. S. Bader (2011). Resolving the structure of interactomes with hierarchical agglomerative clustering. *BMC Bioinformatics,* 12(1), S44.

Parte, A., et al. (2012). Volume 5: The Actinobacteria. Bergey's Manual of Systematic Bacteriology. M. Goodfellow, P. Kämpfer, H.-J. Busse et al., Springer New York, NY. 5: 599.

Patrick R. Amestoy, et al. (2020). igraph: Network Analysis and Visualization. R package, Version 1.3.5. Available at https://cran.r-project.org/web/packages/igraph/index.html.

Peccia, J. and M. Hernandez (2006). Incorporating polymerase chain reaction-based identification, population characterization, and quantification of microorganisms into aerosol science: A review. *Atmospheric Environment,* 40(21), 3941-3961.

Peel, M. C., et al. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.,* 11(5), 1633-1644.

Petroselli, C., et al. (2021). Characterization of long-range transported bioaerosols in the Central Mediterranean. *Science of The Total Environment,* 763, 143010.

Poirel, L., et al. (2010). Diversity, epidemiology, and genetics of class D beta-lactamases. *Antimicrob Agents Chemother,* 54(1), 24-38.

Polymenakou, P. N., et al. (2008). Particle size distribution of airborne microorganisms and pathogens during an intense African dust event in the eastern Mediterranean. *Environ Health Perspect,* 116(3), 292-296.

Pöschl, U. (2005). Atmospheric aerosols: composition, transformation, climate and health effects. *Angewandte Chemie International Edition,* 44(46), 7520-7540.

Pöschl, U. and M. Shiraiwa (2015). Multiphase chemistry at the atmosphere–biosphere interface influencing climate and public health in the anthropocene. *Chemical Reviews,* 115(10), 4440-4475.

Pósfai, M., et al. (2003). Aerosol bacteria over the Southern Ocean during ACE-1. *Atmospheric Research,* 66(4), 231-240.

Pratt, K. A., et al. (2009). In situ detection of biological particles in cloud ice-crystals. *Nature Geoscience,* 2(6), 398-401.

Protection, C. f. H. (2017). Hong Kong Strategy and action plan on antimicrobial resistance (2017–2022).

Proulx, S. R., et al. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution,* 20(6), 345-353.

Pruden, A., et al. (2006). Antibiotic resistance genes as emerging contaminants: studies in northern Colorado. *Environmental Science & Technology,* 40(23), 7445-7450.

Pruesse, E., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res,* 35(21), 7188-7196.

Qin, N., et al. (2020). Longitudinal survey of microbiome associated with particulate matter in a megacity. *Genome Biol,* 21(1), 55.

Quinn, R. (2013). Rethinking antibiotic research and development: World War II and the Penicillin collaborative. *American Journal of Public Health,* 103(3), 426-434.

Reshef, D. N., et al. (2011). Detecting novel associations in large data sets. *Science,* 334(6062), 1518-1524.

Rice, E. W., et al. (2020). Determining hosts of antibiotic resistance genes: a review of methodological advances. *Environmental Science & Technology Letters,* 7(5), 282-291.

Rice, L. B. (2008). Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J Infect Dis,* 197(8), 1079-1081.

Riera-Ruiz, C., et al. (2014). First report of bacterial panicle blight of rice caused by Burkholderia gladioli in Ecuador. *Plant Dis,* 98(11), 1577.

Riesenfeld, C. S., et al. (2004). Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol,* 6(9), 981-989.

Rinttilä, T., et al. (2004). Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *J Appl Microbiol,* 97(6), 1166-1177.

Robert J. Hijmans, et al. (2015). geosphere: Spherical Trigonometry. R package, Version 1.5-10. Available at https://cran.r-project.org/package=geosphere.

Rogerson, A. and A. Detwiler (1999). Abundance of airborne heterotrophic protists in ground level air of South Dakota. *Atmospheric Research,* 51(1), 35-44.

Rognes, T., et al. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ,* 4, e2584.

Romano, S., et al. (2020). Airborne bacteria structure and chemical composition relationships in winter and spring PM10 samples over southeastern Italy. *Science of The Total Environment,* 730, 138899.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *J Stat Softw,* 48(2), 1 - 36.

Roszak, D. B. and R. R. Colwell (1987). Survival strategies of bacteria in the natural environment. *Microbiol Rev,* 51(3), 365-379.

Rypdal, K., et al. (2009). Climate and air quality-driven scenarios of ozone and aerosol precursor abatement. *Environmental Science & Policy*.

Šantl-Temkiv, T., et al. (2018). Aeolian dispersal of bacteria in southwest Greenland: their sources, abundance, diversity and physiological states. *FEMS Microbiology Ecology,* 94(4).

Sattler, B., et al. (2001). Bacterial growth in supercooled cloud droplets. *Geophysical Research Letters*, 28(2), 239-242.

Saunders, A. M., et al. (2016). The activated sludge ecosystem contains a core community of abundant organisms. *Isme j*, 10(1), 11-20.

Schloss, P. D., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*, 75(23), 7537-7541.

Schmieder, R. and R. Edwards (2011). Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. *PLOS ONE*, 6(3), e17288.

Schulze, E.-D. (1989). Air pollution and forest decline in a spruce (picea abies) forest. *Science*, 244(4906), 776-783.

Segata, N., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol*, 12(6), R60.

Seifried, J. S., et al. (2015). Spatial distribution of marine airborne bacterial communities. *Microbiologyopen*, 4(3), 475-490.

Sesartic, A., et al. (2012). Bacteria in the ECHAM5-HAM global climate model. *Atmos. Chem. Phys.*, 12(18), 8645-8661.

Shade, A. and N. Stopnisek (2019). Abundance-occupancy distributions to prioritize plant core microbiome membership. *Current Opinion in Microbiology*, 49, 50-58.

Shaffer, B. T. and B. Lighthart (1997). Survey of culturable airborne bacteria at four diverse locations in oregon: urban, rural, forest, and coastal. *Microb Ecol*, 34(3), 167-177.

Shamarina, D., et al. (2017). Communicating the promise, risks, and ethics of large-scale, open space microbiome and metagenome research. *Microbiome*, 5(1), 132.

Sharoni, S., et al. (2015). Infection of phytoplankton by aerosolized marine viruses. *Proceedings of the National Academy of Sciences*, 112(21), 6643-6647.

Shen, F., et al. (2019). Culturability, metabolic activity and composition of ambient bacterial aerosols in a surrogate lung fluid. *Sci Total Environ*, 690, 76-84.

Silver, S. and T. Phung le (2005). A bacterial view of the periodic table: genes and proteins for toxic inorganic ions. *J Ind Microbiol Biotechnol,* 32(11-12), 587-605.

Smets, W., et al. (2016). Airborne bacteria in the atmosphere: Presence, purpose, and potential. *Atmospheric Environment,* 139, 214-221.

Smith, D. J. (2013). Microbes in the upper atmosphere and unique opportunities for astrobiology research. *Astrobiology,* 13(10), 981-990.

Smith, D. J., et al. (2012). Free tropospheric transport of microorganisms from Asia to North America. *Microb Ecol,* 64(4), 973-985.

Sommer, M. O. A., et al. (2009). Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science,* 325(5944), 1128-1131.

Song, W., et al. (2019). MetaCHIP: community-level horizontal gene transfer identification through the combination of best-match and phylogenetic approaches. *Microbiome,* 7(1), 36.

Stamenković, O., et al. (2019). Anthropogenic pressure explains variations in the biodiversity of pond communities along environmental gradients: a case study in south-eastern Serbia. *Hydrobiologia,* 838(1), 65-83.

Stern, R. A., et al. (2021). The Microbiome of Size-Fractionated Airborne Particles from the Sahara Region. *Environ Sci Technol,* 55(3), 1487-1496.

Stoica, P. and Y. Selen (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine,* 21(4), 36-47.

Stone, W., et al. (2016). Microbial metabolism in bentonite clay: Saturation, desiccation and relative humidity. *Applied Clay Science,* 129, 54-64.

Stopnisek, N. and A. Shade (2021). Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. *The ISME Journal,* 15(9), 2708-2722.

Streets, D. G. and S. T. Waldhoff (2000). Present and future emissions of air pollutants in China:: SO2, NOx, and CO. *Atmospheric Environment,* 34(3), 363-374.

Sturm, R. (2012). Modeling the deposition of bioaerosols with variable size and shape in the human respiratory tract – A review. *Journal of Advanced Research,* 3(4), 295-304.

Sugden, R., et al. (2016). Combatting antimicrobial resistance globally. *Nature Microbiology,* 1(10), 16187.

Sun, R., et al. (2022). Bacterial concentrations and water turbulence influence the importance of conjugation versus phage-mediated antibiotic resistance gene transfer in suspended growth systems. *ACS Environmental Au,* 2(2), 156-165.

Sun, X., et al. (2020). Exploring the disparity of inhalable bacterial communities and antibiotic resistance genes between hazy days and non-hazy days in a cold megacity in Northeast China. *J Hazard Mater,* 398, 122984.

Sunagawa, S., et al. (2015). Structure and function of the global ocean microbiome. *Science,* 348(6237), 1261359.

Székely, A. J. and S. Langenheder (2014). The importance of species sorting differs between habitat generalists and specialists in bacterial communities. *FEMS Microbiology Ecology,* 87(1), 102-112.

Tacconelli, E., et al. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis,* 18(3), 318-327.

Takahashi, S., et al. (2014). Development of a prokaryotic universal primer for simultaneous analysis of bacteria and archaea using next-generation sequencing. *PLOS ONE,* 9(8), e105592.

Tanaka, D., et al. (2015). Seasonal variations in airborne bacterial community structures at a suburban site of central Japan over a 1-year time period using PCR-DGGE method. *Aerobiologia,* 31(2), 143-157.

Tesson, S. V. M., et al. (2016). Airborne microalgae: insights, opportunities, and challenges. *Appl Environ Microbiol,* 82(7), 1978-1991.

Thompson, L. R., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature,* 551(7681), 457-463.

Tie, X., et al. (2006). Chemical characterization of air pollution in Eastern China and the Eastern United States. *Atmospheric Environment,* 40(14), 2607-2625.

Tignat-Perrier, R., et al. (2019). Global airborne microbial communities controlled by surrounding landscapes and wind conditions. *Scientific Reports,* 9(1), 14441.

Toth, M., et al. (2010). An antibiotic-resistance enzyme from a deep-sea bacterium. *Journal of the American Chemical Society,* 132(2), 816-823.

Tripathi, V. and E. Cytryn (2017). Impact of anthropogenic activities on the dissemination of antibiotic resistance across ecological boundaries. *Essays Biochem,* 61(1), 11-21.

Uetake, J., et al. (2020). Airborne bacteria confirm the pristine nature of the Southern Ocean boundary layer. *Proc Natl Acad Sci U S A,* 117(24), 13275-13282.

Uritskiy, G. V., et al. (2018). MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome,* 6(1), 158.

Veal, D. A., et al. (2000). Fluorescence staining and flow cytometry for monitoring microbial cells. *Journal of Immunological Methods,* 243(1), 191-210.

Veron, F. (2015). Ocean Spray. *Annual Review of Fluid Mechanics,* 47(1), 507-538.

Vick-Majors, T. J., et al. (2014). Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *The ISME Journal,* 8(4), 778-789.

Vikesland, P. J., et al. (2017). Toward a comprehensive strategy to mitigate dissemination of environmental sources of antibiotic resistance. *Environmental Science & Technology,* 51(22), 13061-13069.

Vives-Rego, J., et al. (2000). Current and future applications of flow cytometry in aquatic microbiology. *FEMS Microbiology Reviews,* 24(4), 429-448.

Vorholt, J. A. (2012). Microbial life in the phyllosphere. *Nature Reviews Microbiology,* 10(12), 828-840.

Wang, F., et al. (2016). Effect of exposure to staphylococcus aureus, particulate matter, and their combination on the neurobehavioral function of mice. *Environmental Toxicology and Pharmacology,* 47, 175-181.

Wang, H., et al. (2017). Combined use of network inference tools identifies ecologically meaningful bacterial associations in a paddy soil. *Soil Biology and Biochemistry,* 105, 227-235.

Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.

Weon, H.-Y., et al. (2008). Methylobacterium iners sp. nov. and Methylobacterium aerolatum sp. nov., isolated from air samples in Korea. *International Journal of Systematic and Evolutionary Microbiology*, 58(1), 93-96.

West, J. J., et al. (2016). What we breathe impacts our health: improving understanding of the link between air pollution and health. *Environmental Science & Technology*, 50(10), 4895-4904.

Whitman, W. B., et al. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A*, 95(12), 6578-6583.

WHO (2021). "Ambient (outdoor) air quality and health.". from https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.

Wilson, T. W., et al. (2015). A marine biogenic source of atmospheric ice-nucleating particles. *Nature*, 525(7568), 234-238.

Womack, A. M., et al. (2015). Characterization of active and total fungal communities in the atmosphere over the Amazon rainforest. *Biogeosciences*, 12(21), 6337-6349.

Womack, A. M., et al. (2010). Biodiversity and biogeography of the atmosphere. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1558), 3645-3653.

Wood, D. E., et al. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol*, 20(1), 257.

Wright, G. D. (2007). The antibiotic resistome: the nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, 5(3), 175-186.

Wu, D., et al. (2017). Antibiotic resistance genes and associated microbial community conditions in aging landfill systems. *Environmental Science & Technology*, 51(21), 12859-12867.

Wu, D., et al. (2022). Inhalable antibiotic resistomes emitted from hospitals: metagenomic insights into bacterial hosts, clinical relevance, and environmental risks. *Microbiome*, 10(1), 19.

Wu, L., et al. (2019). Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nature Microbiology*, 4(7), 1183-1195.

Wu, Y. W., et al. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics,* 32(4), 605-607.

Wunderlin, T., et al. (2016). Physical isolation of endospores from environmental samples by targeted lysis of vegetative cells. *J Vis Exp,* (107), e53411.

Xia, X., et al. (2015). Bacterial communities in marine aerosols revealed by 454 pyrosequencing of the 16S rRNA gene. *Journal of the Atmospheric Sciences,* 72(8), 2997-3008.

Xiao, N., et al. (2022). Disentangling direct from indirect relationships in association networks. *Proceedings of the National Academy of Sciences,* 119(2), e2109995119.

Xie, J., et al. (2019). Bacteria and antibiotic resistance genes (ARGs) in PM2.5 from China: implications for human exposure. *Environmental Science & Technology,* 53(2), 963-972.

Xie, J., et al. (2018). Seasonal disparities in airborne bacteria and associated antibiotic resistance genes in PM2.5 between urban and rural sites. *Environmental Science & Technology Letters,* 5(2), 74-79.

Xie, J., et al. (2022). Inhalable antibiotic resistome from wastewater treatment plants to urban areas: bacterial hosts, dissemination risks, and source contributions. *Environmental Science & Technology,* 56(11), 7040-7051.

Xu, C., et al. (2017). Bacterial characterization in ambient submicron particles during severe haze episodes at Ji'nan, China. *Sci Total Environ,* 580, 188-196.

Xu, C., et al. (2019). Profile of inhalable bacteria in PM2.5 at Mt. Tai, China: Abundance, community, and influence of air mass trajectories. *Ecotoxicology and Environmental Safety,* 168, 110-119.

Yadav, R. K. P., et al. (2005). Bacterial colonization of the phyllosphere of Mediterranean perennial species as influenced by leaf structural and chemical features. *Microbial Ecology,* 50(2), 185-196.

Yamaguchi, N., et al. (2014). Long-range Transportation of Bacterial Cells by Asian Dust. *Genes and Environment,* 36(3), 145-151.

Yan, D., et al. (2018). Structural variation in the bacterial community associated with airborne particulate matter in Beijing, China, during hazy and nonhazy days. *Appl Environ Microbiol,* 84(9).

Yang, K., et al. (2019). Airborne bacteria in a wastewater treatment plant: Emission characterization, source analysis and health risk assessment. *Water Res,* 149, 596-606.

Yee, L. D., et al. (2020). Natural and anthropogenically influenced isoprene oxidation in southeastern United States and central Amazon. *Environ Sci Technol,* 54(10), 5980-5991.

Yoo, K., et al. (2019). Bacillus-dominant airborne bacterial communities identified during Asian dust events. *Microbial Ecology,* 78(3), 677-687.

Yu, Z., et al. (2021). Synergistic effect of sulfidated nano zerovalent iron and persulfate on inactivating antibiotic resistant bacteria and antibiotic resistance genes. *Water Res,* 198, 117141.

Yuan, H., et al. (2017). Cell concentration, viability and culture composition of airborne bacteria during a dust event in Beijing. *J Environ Sci (China),* 55, 33-40.

Yuan, M. M., et al. (2021). Climate warming enhances microbial network complexity and stability. *Nature Climate Change,* 11(4), 343-348.

Yue, Y., et al. (2018). Size-resolved endotoxin and oxidative potential of ambient particles in Beijing and Zürich. *Environ Sci Technol,* 52(12), 6816-6824.

Zhang, A.-N., et al. (2021). An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nature Communications,* 12(1), 4765.

Zhang, Q., et al. (2020b). Cyanobacterial blooms contribute to the diversity of antibiotic-resistance genes in aquatic ecosystems. *Communications Biology,* 3(1), 737.

Zhang, X.-X., et al. (2009a). Antibiotic resistance genes in water environment. *Appl Microbiol Biotechnol,* 82(3), 397-414.

Zhang, Y. L. and F. Cao (2015). Fine particulate matter (PM 2.5) in China at a city level. *Sci Rep,* 5, 14884.

Zhang, Z. and J. Zhang (2009b). A big world inside small-world networks. *PLOS ONE,* 4(5), e5686.

Zhang, Z., et al. (2022). Assessment of global health risk of antibiotic resistance genes. *Nature Communications,* 13(1), 1553.

Zhao, D., et al. (2016). Network analysis reveals seasonal variation of co-occurrence correlations between Cyanobacteria and other bacterioplankton. *Science of The Total Environment,* 573, 817-825.

Zhao, Y., et al. (2020). Monitoring antibiotic resistomes and bacterial microbiomes in the aerosols from fine, hazy, and dusty weather in Tianjin, China using a developed high-volume tandem liquid impinging sampler. *Sci Total Environ,* 731, 139242.

Zhen, Q., et al. (2017). Meteorological factors had more impact on airborne bacterial communities than air pollutants. *Sci Total Environ,* 601-602, 703-712.

Zhou, J., et al. (2018). Defining the sizes of airborne particles that mediate influenza transmission in ferrets. *Proceedings of the National Academy of Sciences,* 115(10), E2386-E2392.

Zhu, G., et al. (2021). Air pollution could drive global dissemination of antibiotic resistance genes. *The ISME Journal,* 15(1), 270-281.

Zhu, Y.-G., et al. (2017). Continental-scale pollution of estuaries with antibiotic resistance genes. *Nature Microbiology,* 2(4), 16270.

Ziemba, L. D., et al. (2016). Airborne observations of bioaerosol over the Southeast United States using a Wideband Integrated Bioaerosol Sensor. *Journal of Geophysical Research: Atmospheres,* 121(14), 8506-8524.

Zweifel, U. L., et al. (2012). High bacterial 16S rRNA gene diversity above the atmospheric boundary layer. *Aerobiologia,* 28(4), 481-498.

Publications from the Current PhD Study

1. **Zhao, J.**, Jin, L., Wu, D., Xie, J. W., Li, J., Fu, X. W., Cong, Z. Y., Fu, P. Q., Zhang, Y., Luo, X. S., Feng, X. B., Zhang, G., Tiedje, J. M., & Li, X.D. (2022). Global airborne bacterial community—interactions with earth's microbiomes and anthropogenic activities. *Proceedings of the National Academy of Sciences, USA,* 119(42), e2204465119.

2. Wu, D., **Zhao, J.**, Su, Y. L., Yang, M. J., Dolfing, J., Graham, D. W., Yang, K., & Xie, B. (2023). Explaining the resistomes in a megacity's water supply catchment: roles of microbial assembly-dominant taxa, niched environments and pathogenic bacteria. *Water Research*, 228, 119359.

3. Wu, D., Jin, L., Xie, J. W., Liu, H., **Zhao, J.**, Ye, D., & Li, X. D., (2022). Inhalable antibiotic resistomes emitted from hospitals: metagenomic insights into bacterial hosts, clinical relevance, and environmental risks. *Microbiome*, 10(1), 19.

4. Wu, D., Su, Y. L., Wang, P. L., **Zhao, J.**, Xie, J. W., & Xie, B. (2022). Uncover landfilled antimicrobial resistance: a critical review of antibiotics flux, resistome dynamics and risk assessment. *National Science Open*, 1(2097-1168), 20220012.