Diverse recruitment to a globally structured atmospheric microbiome

Stephen D.J. Archer¹, Kevin C. Lee¹, Tancredi Caruso², Marcus H.Y. Leung³, Xinzha³, Tong³, Graham Hinchliffe¹, Teruya Maki⁴, Tina Santl-Temkiv⁵, Kimberley A. Warren-Rhodes⁶,⁷, Benito Gomez-Silva⁸, Kevin D. Hyde⁹, Celine J.N. Liu¹⁰, Antonio Alcami¹¹, Dina M. Al Mailem¹², Jonathan G. Araya¹³, S. Craig Cary¹⁴, Don A. Cowan¹⁵, Jessica Dempsey¹⁶, Claudia Etchebehere¹⁷, Batdelger Gantsetseg¹⁸, Sean Hartery¹⁹, Mike Harvey²⁰, Kazuichi Hayakawa²¹, Ian Hogg²², Mutsoe Inoue²¹, Mayada K. Kansour¹², Timothy Lawrence¹, Charles K. Lee¹⁴, Matthias Leopold²³, Christopher P. McKay⁶, Seiya Nagao²¹, Yan Hong Poh¹⁰, Jean-Baptiste Ramond²⁴, Alberto Rastrojo¹¹, Toshio Sekiguchi²¹, Joo Huang Sim¹⁰, William Stahn¹⁶, Henry J. Sun²⁵, Ning Tang²¹, Bryan Vandenbrink²², Craig Walther¹⁶, Patrick K.H. Lee³, Stephen B. Pointing⁵,²¹,²⁶*

¹ Department of Applied Ecology, Auckland University of Technology, Auckland, New Zealand
² School of Biology and Environmental Science, University College Dublin, Dublin, Ireland
³ School of Energy and Environment, City University of Hong Kong, Hong Kong, China
⁴ Department of Life Sciences, Kindai University, Osaka, Japan
⁵ Department of Biology, Aarhus University, Aarhus, Denmark
⁶ NASA Ames Research Center, Mountain View, California, USA
⁷ SETI Institute, Mountain View, California, USA
⁸ Departamento Biomédico, CeBiB, Universidad de Antofagasta, Antofagasta, Chile
⁹ Center of Excellence in Fungal Diversity, Mae Fah Luang University, Chiang Rai, Thailand
¹⁰ Yale-NUS College, National University of Singapore, Singapore
¹¹ Université Paris-Sud, Université Paris-Saclay, Gif-Sur-Yvette, France
¹² Aberystwyth University, Aberystwyth, Wales, UK
¹³ Instituto Politécnico de Tomar, Tomar, Portugal
¹⁴ The University of Queensland, Brisbane, Australia
¹⁵ Chulalongkorn University, Faculty of Medicine, Bangkok, Thailand
¹⁶ The Open University, Milton Keynes, UK
¹⁷ Leiden University, Leiden, The Netherlands
¹⁸ University of California, Berkeley, USA
¹⁹ University of Bern, Bern, Switzerland
²⁰ University of Oxford, Oxford, UK
²¹ University of Technology Sydney, Sydney, Australia
²² Leibniz Institute for Natural Product Research and Infection Biology, Braunschweig, Germany
²³ University of California, San Diego, USA
²⁴ University of Colorado Boulder, USA
²⁵ University of California, Santa Cruz, USA
²⁶ National University of Singapore, Singapore
Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas (CSIC), Universidad Autónoma de Madrid, Madrid, Spain

Department of Biological Sciences, Kuwait University, Kuwait City, Kuwait.

Instituto Antofagasta, Universidad de Antofagasta, Antofagasta, Chile

School of Science, University of Waikato, Hamilton, New Zealand

Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa

East Asian Observatory, Hilo, Hawaii, USA

Biological Research Institute Clemente Estable, Ministry of Education, Montevideo, Uruguay

Institute of Meteorology, Hydrology and Environment, Ulan Bator, Tuv, Mongolia

School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand

National Institute of Water and Atmospheric Research, Wellington, New Zealand

Institute of Nature and Environmental Technology, Kanazawa University, Kanazawa, Japan

Canadian High Arctic Research Station, Cambridge Bay, Nunavut, Canada

UWA School of Agriculture and Environment, University of Western Australia, Perth, Australia

Departamento de Genética Molecular y Microbiología, Pontificia Universidad Católica de Chile, Santiago, Chile

Desert Research Institute, Las Vegas, Nevada, USA

Department of Biological Sciences, National University of Singapore, Singapore

*e-mail: S.B.P. stephen.pointing@yale-nus.edu.sg
Summary

Atmospheric transport is critical to dispersal of microorganisms between habitats and this underpins resilience in terrestrial and marine ecosystems globally. Conventional dogma that this is a neutral process involving ubiquitous distribution in air has been challenged by recent advances. However, the lack of standardized methods and analytical frameworks have impeded synthesis and global perspective. A key unresolved question is whether microorganisms assemble to form a taxonomically distinct, geographically variable and functionally adapted atmospheric microbiome. Here we characterised global-scale patterns of microbial taxonomic and functional diversity in air within and above the atmospheric boundary layer and in underlying soils. Bacterial and fungal assemblages in air were taxonomically structured and deviated substantially from purely stochastic assembly. Fungi dominated above tropical, temperate and continental biomes whilst bacteria did so above oceans and drylands. At high altitudes bacterial diversity declined but fungal diversity was greatest. Source-tracking indicated a complex recruitment process involving local soils plus globally distributed inputs from drylands and the phyllosphere. Assemblages displayed stress-response and metabolic traits relevant to survival in air, and taxonomic and functional diversity were correlated with macroclimate and atmospheric variables. Our findings highlight a structured global atmospheric microbiome that is central to understanding regional and global ecosystem connectivity.

Main Text

Microorganisms occupy central roles in terrestrial and marine ecosystems globally. Movement of viable cells and propagules between habitats occurs largely through the troposphere, which is the atmospheric layer closest to Earth. This is critical to recruitment and turnover that drive ecological resilience of these systems, as well as influencing...
dispersal of pathogens and invasive taxa. There is also a growing awareness that microorganisms suspended in the atmosphere are potentially capable of in situ metabolic and biophysical activity that can influence climatic processes. However, despite the central importance of the atmosphere to these ecological outcomes, assessments of microbial diversity in air at broad geographic scales remain limited. As a result, there is little understanding of how variable the overall microbial composition of the atmosphere may be on a global scale, the extent to which it may be decoupled from underlying local surface communities, or the importance of environmental or biotic factors in shaping diversity. The unique role of the atmosphere as a transport medium for microorganisms has also obscured the question of whether it supports a functionally adapted microbiome with the potential for metabolic transformations and cell proliferation.

Previous research suggests that the conventional assumption of random and ubiquitous microbial transport in the atmosphere may no longer be valid. Diversity estimates for air within the boundary layer at near-ground level at various locales have indicated varied bacterial and fungal communities that were correlated with local abiotic variables such as temperature and humidity or land use. Several studies have related variation in diversity to different history of sampled air masses and this suggests combined influence of the different sources and conditions to which microorganisms are exposed during transit. Estimates made indirectly from ground-deposited desert dust or precipitation have yielded valuable insight on long-range dispersal across inter-continental scales although they reflect deposition and differ somewhat to direct estimates from air. Estimating diversity in the free troposphere at higher altitudes above the boundary layer is challenging and scarce data indicates a more restricted microbial occurrence. There is currently no consensus for diversity or abundance estimates due to the different experimental approaches, lack of ecologically relevant scaling and taxonomic resolution, and the confounding effect of
contamination during diversity estimation for an ultra-low biomass system such as air\textsuperscript{23,24}. Adaptive traits have generally been inferred from taxonomy, although laboratory estimates of metabolic activity by atmospheric bacterial isolates\textsuperscript{25}, and recovery of RNA from air and cloud water\textsuperscript{26,27}, indicate that atmospheric microorganisms are potentially active in situ.

Here we report taxonomic and functional diversity in a large globally sourced original dataset ($n = 596$) for air within the atmospheric boundary layer that delineates the majority of physical interactions with the Earth’s surface\textsuperscript{28} (near-ground air), as well as aircraft sampling of free tropospheric air at higher altitudes above the boundary layer (high-altitude air) (Fig. 1a,c; Supplementary Information, Field sampling). We combined this with concurrent sampling of underlying surface soils and sediments to allow direct air-surface connectivity comparisons. The approach employed a combination of high-throughput techniques to estimate and triangulate taxonomic diversity: shotgun metagenomics for inter-domain phylum-level comparisons and targeted amplicon sequencing for finer scale taxonomic resolution of abundant groups. This was combined with an unprecedented effort to mitigate against the occurrence of putative contaminant taxa that have plagued low-biomass microbiological studies\textsuperscript{23,24} (Supplementary Information, Decontamination of environmental sequence data). A statistical classification approach was used to identify potential recruitment sources of airborne microorganisms. This was complemented by a targeted functional analysis of our metagenomes ($n = 120$) focused on metabolic strategies and stress responses relevant to environmental conditions in the atmosphere. We combined these biotic data with a novel geospatial modelling approach to enable correlation with environmental variables to which microorganisms were exposed during transit in the atmosphere, as well as local climatic variables. We report compelling evidence for a taxonomically distinct, altitudinally and geographically variable, and functionally relevant atmospheric microbiome that is influenced by a complex suite of biotic and abiotic drivers. We demonstrate that
diverse local and global recruitment sources are important and that dryland soils are a particularly strong influence on global diversity patterns in air.

An overview of inter-domain diversity from our metagenomic libraries indicated that bacteria and fungi were the most abundant microorganisms in air with very low and patchy occurrence of archaea and protists. We therefore focused further abundance and diversity estimation effort on these two groups. On a global scale consistent patterns were observed where total biomass and estimated abundance for bacteria and fungi were substantially lower in near-ground air within the atmospheric boundary layer than in underlying soil (Fig. 1b,d,h; Extended Data Fig. 1). High-altitude air above the boundary layer indicated a further steep decline from values observed for near-ground level air. Fungi were more abundant than bacteria in air at all locations except those above oceans and drylands (Supplementary Information, Inventory of taxa). We view the abundance of microorganisms in air within the atmospheric boundary layer as analogous to that which occurs in microbially active zones of other biomes, such as the photic zone of the oceans or the topsoil of terrestrial habitats. Our data allowed a globally triangulated estimate of $1.3 \times 10^{22}$ bacterial cells and $6.1 \times 10^{23}$ fungal cells in this region of the global atmosphere.

Conserved ecological patterns were evident on a global scale for bacterial and fungal diversity. Our diversity estimations using two commonly employed approaches, amplicon sequencing and shotgun metagenomics, were broadly congruent (Extended Data Fig. 2) and so we focused our fine scale phylogenetic interrogation on amplicon sequence data because this approach allowed better ecological representation of the targeted assemblages. Bacteria were least diverse in high-altitude air, followed by near-ground air, and soils were relatively taxa-rich although variable as expected for the extensive global environmental gradient encompassed by our study (Fig. 1e; Extended Data Fig. 1)\(^{29-31}\). Conversely, the fungi were most diverse in high-altitude air, followed by near-ground air and soils were least diverse
(Fig. 1i; Extended Data Fig. 1). Clear biogeographic patterns occurred where bacterial
diversity was reduced over low productivity habitats and fungal diversity was elevated above
locations with a well-developed phyllosphere, thus broadly reflecting patterns for soil
communities (Extended Data Fig. 1) \(^{29,32}\). We observed distinct patterns common to all
bacterial and fungal assemblages in air globally and these were robustly supported after
extensive effort to mitigate contaminant signal from the low biomass air and soil samples
(Supplementary Information, Decontamination of environmental sequence data): A robust
separation of assemblages by habitat type (high-altitude air, near-ground air, soil) and clear
biogeographic separation by location was observed (Fig. 1f,j; Extended Data Fig. 3). This
delineation was supported despite co-occurrence analysis indicating assemblages in air were
more variable than those in soil (Extended Data Fig. 4a). A pronounced distance decay
relationship for diversity was observed for near-ground and high-altitude air assemblages that
was comparable to those for soil (Extended Data Fig. 4b). However, there was no evidence
for a latitudinal gradient in richness and this mirrored observations for global soil diversity\(^{29}\),
and also reflected the inclusion of deserts, mountains, high latitude and ocean locations in our
study. We then applied null models to estimate nestedness, which is a measure of the extent to
which observed patterns of co-occurrence are non-random (Extended Data Fig. 5). Bacterial
and fungal assemblages in both near-ground and high-altitude air were significantly less
nested than null models and therefore taxonomically structured and non-randomly assembled,
although slightly less so than for soils \(^{31}\) (Fig. 1g,k). The non-random patterns were
fundamentally an outcome of taxa specificity to habitat and location, and we interpreted these
as indicative of strong filtering for taxa. At higher altitudes above the atmospheric boundary
layer where abiotic stressors are more pronounced, communities were even more structured
and this reflected more extreme environmental filtering. The pattern persisted between
hemispheres sampled at peak and low growing season and across major climatic boundaries
and land use types. Our data indicates that environmental filtering in both near-ground and high-altitude air results in structured and biogeographically predictable patterns for bacteria and fungi.

Taxonomic composition of assemblages was congruent with the ecological patterns observed (Fig. 2a,d; Extended Data Fig. 6). At broad taxonomic ranks (phylum-class) a remarkably consistent diversity occurred globally regardless of underlying biome or growing season. Our amplicon sequence variant (ASV) approach to diversity analysis revealed that at finer taxonomic scale (genus-ASV) and after extensive decontamination effort there were 13% of bacterial and 10% of fungal genera co-occurring among ≥50% of globally distributed air samples (Supplementary Information, Inventory of taxa). The only genus with ubiquitous representation in all air samples was *Sphingomonas*, a diverse group linked with emissions from the phyllosphere \(^{16,33}\). Despite conserved patterns of diversity at lower taxonomic ranks down to genus level, we found no evidence for a “core” group of specific ASV-defined taxa that represent an atmospheric microbiome. Nonetheless there was compelling evidence from taxonomic data for environmental filtering of assemblages in air. Bacteria enriched in near-ground air compared to soil were largely accounted for by classes supporting taxa with known tolerance to environmental stress (Actinobacteria, Alphaproteobacteria, Firmicutes and Gammaproteobacteria), although it cannot be ruled out that this also indicates taxa that possess adaptive traits that favour aerosolization \(^{34}\). At higher altitudes where environmental stress is exacerbated spore-forming Actinobacteria and Firmicutes were more abundant (Extended Data Fig. 6), suggesting selection towards survival as passive resting structures. Elevated abundance of gammaproteobacterial taxa at a single location (South Africa) was consistent with emissions of this group due to land-use as a farm \(^{35}\). For the fungi, elevated diversity in air relative to soils was largely due to macrofungi (Agaricomycetes) and prolific spore-formers (Dothidiomycetes). In the absence of observed mycelia in air samples we
concluded that spores accounted for much of the fungal signature in air (Extended Data Fig. 6). The Agaricomycetes were notably more abundant in tropical air than all other locations globally and this likely reflected global patterns for terrestrial fungi. In near-ground air the abundance of common fungal agricultural pathogens (Ustilaginomycetes) was elevated in temperate Northern Hemisphere locations sampled during peak growing season, as opposed to reduced abundance in Southern Hemisphere samples collected at the end of the growing season. This we interpreted as a signature of seasonality in land use on a global scale. Previous studies at individual near-ground locales have concluded that inter-seasonal variation may variously be absent, weak, pronounced for some taxa or stochastic. Elevated fungal diversity in ultra-low biomass high-altitude air was indicative of persistent fungal propagules that are tolerant to extreme UV and thermal stress. Residence time for cells in air may be extended at high altitudes and so this necessitates effective tolerance to these stressors during potentially long-distance dispersal. Overall, our combined ecological and taxonomic data provided strong evidence that contrary to long-held dogma in microbial ecology that microbial transport in air is ubiquitous and neutral to dispersal outcomes, instead atmospheric microbiomes exhibit a pronounced biogeography.

In order to further interrogate possible explanations for the observed patterns, we conducted source tracking analysis to assess the likely origin of bacteria and fungi encountered in the air. First, a co-occurrence analysis revealed that near-ground air displayed greatest taxonomic connectivity with local soil at any given location and less connectivity with soil from different locations (Fig. 2b,e). Assemblages in high-altitude air displayed markedly fewer shared taxa with underlying near-ground air or soil and were essentially decoupled from local underlying surface assemblages (Fig. 2b,e). Aerosolization of microorganisms not only occurs from soil but also from different terrestrial and aquatic surfaces, e.g. ocean surface waters, the phyllosphere and stochastic desert dust events.
We therefore employed a statistical classification algorithm to estimate recruitment to air microbiomes from the surface habitats of different climatic regions (Fig. 2c,f, Extended Data Fig. 7). A large volume of taxa had unexplained sources and this is likely due in part to stochastic aerosolisation events for microorganisms and also a consequence of the fact that it is impossible to exhaustively sample potential source microbiomes. For most locations local soil was the major explained source of bacteria and fungi in air, and bacteria were sourced in a more cosmopolitan manner than fungi (Extended Data Fig. 7). Many sampled air masses had significant transit over oceans and yet marine sources were a relatively minor contributor to observed diversity in air above terrestrial locations. This reflects that fewer microorganisms occur above the oceans than over land, and also the limited oceanic sources for comparison. Clear patterns for terrestrial sources were apparent. Dryland soils (dry deserts, polar/alpine and dry continental locations) were pronounced sources for bacteria globally and this may reflect the more readily aerosolised non-cohesive soils typical of these biomes. This expands the influence of deserts to a global scale beyond the well-defined intercontinental desert dust transit routes for microbial dispersal. For the fungi, polar and alpine soils were major sources and this is congruent with the notion that permanently cold surface substrates in these environments have been proposed to act as long-term reservoirs for inactive fungal propagules. The phyllosphere was a pervasive yet smaller contributor to bacterial diversity globally, and minor contribution to fungal sources likely reflects the lack of comparative data. This may emerge as a more significant source as the inventory of phyllosphere microbiomes increases. For high-altitude air, major sources were dry deserts and polar/alpine sources and this likely reflects in part the adaptive advantages that taxa from these habitats have in air, e.g. UV repair and desiccation tolerance. The ability to become aerosolized may vary between taxa in marine and terrestrial systems and so deterministic biotic drivers may also be relevant to recruitment from sources, as well as selective deposition.
during transit. Overall, the source tracking demonstrated that atmospheric diversity is driven by a complex recruitment process involving local soils plus globally distributed inputs from drylands and the phyllosphere.

To generate further insight into possible biotic drivers of the observed patterns in diversity we conducted a functional metagenomic analysis of selected metabolic and stress-response genes relevant to the atmospheric habitat (Fig. 3a,b; Supplementary Information, Metagenomic functional analysis). We targeted bacteria because they likely comprise any active fraction of the atmospheric microbiome. Differentially more abundant genes in air versus soil were inversely correlated with biomass and taxonomic richness, they affiliated with taxa observed as enriched in air, and all values were averaged by location. We therefore interpret the elevated abundance of genes in air as reflective of assemblage composition rather than an artefact of sampling effort. Distribution of marker genes in air broadly reflected that for underlying soil at any given location and this supported our identification of soil as a major source for atmospheric bacteria. Traits were widely distributed globally and those for stress tolerance were notably more abundant in bacterial assemblages in air above dry and polar/alpine regions (Fig. 3b), thus further supporting our hypothesis that microorganisms from these surface environments are adapted to survival in air. Both near-ground and high-altitude air communities possessed marker genes for cold shock, oxidative stress, sporulation, starvation, and UV-repair and these were elevated in several air assemblages compared with underlying soil (Fig. 3a). High abundance for stress-response genes occurred in air above oceans, providing further evidence that the low biomass and richness above marine surfaces reflected strongly filtered microbial diversity. This may arise during long-distance transport from largely terrestrial sources, as well as during recruitment of bacteria from the sea surface micro-layer. Metabolic marker genes for respiration were widespread, and notably for the \textit{ccoN} proteobacterial cytochrome oxidase that correlated with elevated proteobacteria in air.
versus soil. Markers for the metabolism and fixation of a variety of gaseous atmospheric substrates including carbon dioxide, hydrogen, methane, nitrogen and isoprene, as well as phototrophy were also prevalent in air (Fig. 3a). Elevated occurrence of the coxL gene associated with carbon monoxide metabolism was indicative of the potential for interaction with anthropogenic emissions. This limited functional interrogation provided a much-needed glimpse into the potential for an active and stress-adapted atmospheric microbiome.

Our data indicates that there is capacity for greater metabolic plasticity than the existing inventory from meta-omics and transformation of substrates by atmospheric isolates under laboratory conditions currently suggests.

We examined possible interactions between the taxonomic and functional diversity of assemblages and abiotic variables relevant to survival in air and soil (Fig. 4). These included both location-specific macroclimate and environmental variables encountered by microorganisms during transit acquired using a novel geospatial analysis (Extended Data Fig. 8). Significant correlations were revealed between both local macroclimate and transit abiotic variables and community metrics of taxonomic and functional diversity in air (Fig. 4a).

Relatively strong negative correlations for bacterial and fungal diversity, oxidative stress genes, and UV-repair genes in air with solar radiation and altitude (covariables of UV exposure) provided further evidence for UV exposure as a strong selective force on global bacterial and fungal diversity. Functional genes were most strongly correlated with mean annual precipitation, and this likely reflects niche differentiation of source communities in underlying soil at different climatic locations since we have shown they are coupled to diversity in local air. Transit variables were also influential on functional diversity and this was consistent with our other lines of evidence for environmental filtering. The strong correlation between occurrence of phototrophy genes and all abiotic variables suggested photoautotrophic bacteria may be subject to greater selective pressure than other groups. For
soil communities the correlations with macroclimate variables were relatively congruent with those observed for other global studies of soil microbial diversity (Fig. 4b) \(^{30}\), and this provided triangulation for our approach. These data highlight that although location is important, the conditions to which microorganisms are exposed during transit are a significant but previously overlooked correlate with taxonomic and functional diversity and are influential to dispersal outcomes.

Based upon our findings, we envisage a global system where for any ecological region microorganisms are emitted to and deposited from the atmosphere across multiple spatial and temporal scales, but that physical movement of air masses limits their occurrence largely within that region regardless of emission source. Our empirical data and modelling of maximum residence time for cells in air arriving at each location revealed little overlap in back trajectories and this further supports our hypothesis, and the scale of such regions may be influenced by regional air mass circulation \(^{48}\). We envisage soil and other surface habitats act as a global reservoir for bacterial diversity, whereas fungi were more diverse in air than soil and so the atmosphere may be viewed as a global reservoir for fungal diversity. We have also shown that drylands were a major contributor to global atmospheric microbial diversity and because desert soils are characterised by stochastic emissions that travel inter-continental distances, we assume stochasticity punctuates regional patterns and accounts for some of the unexplained sources \(^{40}\). Communities in air were taxonomically and functionally structured and show clear signs of selective influence on their assembly, despite evidence for recruitment from diverse habitats and many taxa with undetermined source. This is in contrast to a prevailing assumption that dispersal limitation plays a minor role in determining global biogeography of soil communities \(^{30}\). Our taxonomic data highlights that whilst the most abundant microbial groups in air are likely to comprise dormant propagules, bacteria also displayed diverse stress-response and metabolic genes and this may allow limited
proliferation of an active microbiome under certain conditions. Given the physicochemical
and dynamic complexity of the atmosphere and the broad range of correlations we observed
between taxonomic and functional diversity and abiotic factors, a potentially chaotic system
of interplay may emerge that influences atmospheric microbial ecology in a manner similar to
that envisaged for highly dispersed marine larvae\textsuperscript{49}. Taken together we anticipate these
findings will be valuable in future hypothesis-driven research to identify interactions
mediated by the atmospheric microbiome between different surface habitats across multiple
ecological scales, and in particular to testing models of recruitment, turnover, functionality
and resilience. Given that the atmosphere is also a sink for a large fraction of anthropogenic
emissions\textsuperscript{45}, it is timely that an accurate global inventory of microbial diversity is provided in
order to present a baseline for measuring future responses to change. Finally, the study
complements efforts to inventory global soil\textsuperscript{29\textendash}31 and oceanic microbiomes\textsuperscript{50} and expands
the scope of the pan-global microbiome.

References

1. Barberan, A., Casamayor, E. O. & Fierer, N. The microbial contribution to
macroecology. \textit{Front. Microbiol.} \textbf{5}, 1\textendash}8 (2014).

2. Cavicchioli, R. \textit{et al.} Scientists’ warning to humanity: microorganisms and climate
change. \textit{Nat. Rev. Microbiol.} \textbf{17}, 569\textendash}586 (2019).

3. Mayol, E. \textit{et al.} Long-range transport of airborne microbes over the global tropical and
subtropical ocean. \textit{Nat. Commun.} \textbf{8}, 201 (2017).

4. Archer, S. D. J. \textit{et al.} Airborne microbial transport limitation to isolated Antarctic soil
habitats. \textit{Nat. Microbiol.} \textbf{4}, 925\textendash}932 (2019).

5. Cáliz, J., Triadó-Margarit, X., Camarero, L. & Casamayor, E. O. A long-term survey
unveils strong seasonal patterns in the airborne microbiome coupled to general and
regional atmospheric circulations. Proc. Natl. Acad. Sci. U. S. A. 115, 12229–12234 (2018).

6. Womack, A. M., Bohannan, B. J. M. & Green, J. L. Biodiversity and biogeography of the atmosphere. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 3645–3653 (2010).

7. Lighthart, B. The ecology of bacteria in the alfresco atmosphere. FEMS Microbiol. Ecol. 23, 263–274 (2006).

8. Griffin, D. W. Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. Clin. Microbiol. Rev. 20, 459–477 (2007).

9. Deguillaume, L. et al. Microbiology and atmospheric processes: Chemical interactions of primary biological aerosols. Biogeosciences 5, 1073–1084 (2008).

10. Fröhlich-Nowoisky, J. et al. Biogeography in the air: Fungal diversity over land and oceans. Biogeosciences 9, 1125–1136 (2012).

11. Tignat-Perrier, R. et al. Global airborne microbial communities controlled by surrounding landscapes and wind conditions. Sci. Rep. 9, 1–11 (2019).

12. de Wit, R. & Bouvier, T. ‘Everything is everywhere, but, the environment selects’; what did Baas Becking and Beijerinck really say? Environ. Microbiol. 8, 755–758 (2006).

13. Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. H. H. Beyond biogeographic patterns: Processes shaping the microbial landscape. Nat. Rev. Microbiol. 10, 497–506 (2012).

14. Lowe, W. H. & McPeek, M. A. Is dispersal neutral? Trends Ecol. Evol. 29, 444–450 (2014).

15. Šantl-Tempik, T., Gosewinkel, U., Starnawski, P., Lever, M. & Finster, K. Aeolian dispersal of bacteria in southwest Greenland: Their sources, abundance, diversity and physiological states. FEMS Microbiol. Ecol. 94, fiy031 (2018).
16. Bowers, R. M., McLetchie, S., Knight, R. & Fierer, N. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. ISME J. 5, 601–612 (2011).

17. Tignat-Perrier, R. et al. Seasonal shift in airborne microbial communities. Sci. Total Environ. 716, 137129 (2020).

18. Bowers, R. M., McCubbin, I. B., Hallar, A. G. & Fierer, N. Seasonal variability in airborne bacterial communities at a high-elevation site. Atmos. Environ. 50, 41–49 (2012).

19. Els, N. et al. Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation. Aerobiologia. 35, 671–701 (2019).

20. Uetake, J. et al. Seasonal changes of airborne bacterial communities over Tokyo and influence of local meteorology. Front. Microbiol. 10, 1572 (2019).

21. Favet, J. et al. Microbial hitchhikers on intercontinental dust: catching a lift in Chad. ISME J. 7, 850–867 (2013).

22. Els, N., Baumann-Stanzer, K., Larose, C., Vogel, T. M. & Sattler, B. Beyond the planetary boundary layer: Bacterial and fungal vertical biogeography at Mount Sonnblick, Austria. Geo Geogr. Environ. 6, e00069 (2019).

23. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87 (2014).

24. Eisenhofer, R. et al. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol. 27, 105–117 (2019).

25. Khaled, A., Zhang, M., Amato, P., Delort, A.-M. & Ervens, B. Biodegradation by bacteria in clouds: An underestimated sink for some organics in the atmospheric multiphase system. Atmos. Chem. Phys. Discuss. 2020, 1–32 (2020).

26. Klein, A. M., Bohannan, B. J. M., Jaffe, D. A., Levin, D. A. & Green, J. L. Molecular
evidence for metabolically active bacteria in the atmosphere. *Front. Microbiol.* **7**, 1–11 (2016).

27. Amato, P. *et al.* Metatranscriptomic exploration of microbial functioning in clouds. *Sci. Rep.* **9**, 4383 (2019).

28. Stull, R. B. A boundary layer definition. in *Introduction to Boundary Layer Meteorology* 1–27 (Kluwer Academic Publishers, 1988).

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

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

31. Thompson, L. R. *et al.* A communal catalogue reveals Earth’s multiscale microbial diversity. *Nature* **551**, 457–463 (2017).

32. Fierer, N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* **15**, 579–590 (2017).

33. Lymperopoulou, D. S., Adams, R. I. & Lindow, S. E. Contribution of vegetation to the microbial composition of nearby outdoor air. *Appl. Environ. Microbiol.* **82**, 3822–3833 (2016).

34. Michaud, J. M. *et al.* Taxon-specific aerosolization of bacteria and viruses in an experimental ocean-atmosphere mesocosm. *Nat. Commun.* **9**, 2017 (2018).

35. Zhao, Y., Aarnink, A., De Jong, M. & Groot Koerkamp, P. W. G. (Peter). Airborne Microorganisms From Livestock Production Systems and Their Relation to Dust. *Crit. Rev. Environ. Sci. Technol.* **44**, 1071–1128 (2014).

36. Hawksworth, D. L. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* **105**, 1422–1432 (2001).

37. Gusareva, E. S. *et al.* Microbial communities in the tropical air ecosystem follow a
precise diel cycle. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23299–23308 (2019).

38. Bryan, N. C., Christner, B. C., Guzik, T. G., Granger, D. J. & Stewart, M. F.
Abundance and survival of microbial aerosols in the troposphere and stratosphere. *ISME J.* **13**, 2789–2799 (2019).

39. Vorholt, J. A. Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* **10**, 828–840 (2012).

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

41. Pointing, S. B. S. B. & Belnap, J. Microbial colonization and controls in dryland systems. *Nat. Rev. Microbiol.* **10**, 551–562 (2012).

42. Aalismail, N. A. *et al.* Functional metagenomic analysis of dust-associated microbiomes above the Red Sea. *Sci. Rep.* **9**, 1–12 (2019).

43. Reche, I. *et al.* Deposition rates of viruses and bacteria above the atmospheric boundary layer. *ISME J.* **12**, 1154–1162 (2018).

44. Hartery, S. *et al.* Constraining the Surface Flux of Sea Spray Particles From the Southern Ocean. *J. Geophys. Res. Atmos.* **125**, e2019JD032026 (2020).

45. Archer, S. D. J. J. & Pointing, S. B. Anthropogenic impact on the atmospheric microbiome. *Nat. Microbiol.* **5**, 229–231 (2020).

46. Matulová, M., Husárová, S., Capek, P., Sancelme, M. & Delort, A. M.
Biotransformation of various saccharides and production of exopolymeric substances by cloud-borne Bacillus sp. 3B6. *Environ. Sci. Technol.* **48**, 14238–14247 (2014).

47. Šantl-Temkiv, T., Finster, K., Hansen, B. M., Pašić, L. & Karlson, U. G. Viable methanotrophic bacteria enriched from air and rain can oxidize methane at cloud-like conditions. *Aerobiologia.* **29**, 373–384 (2013).

48. Uetake, J. *et al.* Airborne bacteria confirm the pristine nature of the Southern Ocean
boundary layer. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 13275–13282 (2020).

49. Álvarez-Noriega, M. *et al.* Global biogeography of marine dispersal potential. *Nat. Ecol. Evol.* **4**, 1196–1203 (2020).

50. Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science.* **348**, 1–10 (2015).

Acknowledgements

The research was funded by the Singapore Ministry of Education and Yale-NUS College, grant number R-607-265-331-121.

Author contributions

S.B.P. conceived the study, secured funding and led the project; S.D.J.A was field team leader and contributed to experimental design; S.D.J.A., K.C.L., T.M., S.B.P. and K.A.W-R conducted sampling; A.A., D.A.M., J.G.A., S.C.C., C.E., S.H., M.I., M.K.K., C.K.L., C.J.N.L., J.B.R., A.R., T.S., W.S., H.S., B.V. and C.W. supported field sampling; S.C.C., D.A.C., J.D., B.G., B.G-S., M.H., K.H., I.H., M.L., C.P.M., S.N. and N.T. facilitated access to remote field locations; T.L., Y.H.P. and J.H.S. provided scientific assistance for laboratory and field experiments; S.D.J.A. and G.H. collected metadata; S.D.J.A. performed sample processing and laboratory experiments; M.H.Y.L., K.C.L. and X.T. performed bioinformatics analysis; T.C. and K.C.L. performed ecological analysis and modelling; G.H. performed geospatial analysis and modelling; P.K.H.L. and S.B.P. supervised data analysis; S.D.J.A., T.C., B.G-S., K.C.L., P.K.H.L., M.H.Y.L., K.D.H., T.M., S.B.P., T.S-T., X.T. and K.A.W-R. interpreted the findings; S.B.P. wrote the paper.

Competing Interests
The authors declare no competing interests.

Additional Information

Extended Data is available for this paper.

Supplementary Information is available for this paper.

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

Correspondence and requests for materials should be addressed to S.B.P., e-mail:

stephen.pointing@yale-nus.edu.sg
**Figure Legends**

**Fig. 1 | The global atmospheric microbiome is taxonomically structured.** a, Back trajectories for near-ground air (blue lines) and high-altitude air (red lines), plus soil sampling locations (green boxes) (total independent samples $n = 596$). b, Global distribution for microbial biomass in air and soil. c, Mean transit altitudes for sampled air. d, h, Global bacterial (d) and fungal (h) abundance distribution. e, i, Global taxonomic richness of bacterial (e) and fungal (i) ASV. f, j, Community dissimilarity (Jaccard Index) for bacterial (f) and fungal (j) air and soil assemblages by location. g, k, Modelled nestedness estimates for bacterial (g) and fungal (k) assemblages across phylogenetic ranks. Networks constructed for each habitat and location are shown in Extended Data Fig. 5. HA air, high-altitude air; NG air, near-ground air. Locations: 1, Canada; 2, Mongolia; 3, Spain; 4, Japan; 5, California, USA; 6, Kuwait; 7, Hilo, Hawaii, USA; 8, Mauna Kea, Hawaii, USA; 9, Singapore; 10, Coral Sea; 11, Namibia; 12, Chile; 13, South Africa; 14, Australia; 15, Uruguay; 16, New Zealand; 17, Southern Ocean; 18, Antarctica. Meta-data for each location are shown in Supplementary Information, Field sampling.
Fig. 2 | Atmospheric assemblages reflect diverse recruitment sources. a,d, Taxonomic composition of ASV-defined bacterial and fungal assemblages in air and underlying soil among globally distributed climatic locations (bacteria $n = 529$; fungi $n = 444$). Asterisks indicate locations where aircraft sampling of high-altitude air was undertaken. b,e, Shared ASV-defined bacterial (b) and fungal (e) taxa among habitat types within and between locations. Locations are shown in rank order of explained source for each climate. HA air, high-altitude air; NG air, near-ground air. c,f, Source contribution to observed bacterial (c) and fungal (f) diversity in air, averaged for each source biome in order to mitigate sample size effects. Locations are numbered as shown in Fig. 1.
Fig. 3 | The atmospheric microbiome displays functional traits relevant to survival and metabolism in air. 

**a.** Functional metagenomics profiling of targeted stress-response and metabolic genes by habitat type (n = 120). HA air, high-altitude air; NG air, near-ground air.  

**b.** Summary for distribution of stress-response and metabolic genes by climatic region, with all locations globally pooled by climate (n = 120). Oxid. stress, oxidative stress; Trace gas met., trace gas metabolism.
Fig. 4 | Taxonomic and functional diversity is correlated with multiple macroclimate and atmospheric transit variables. a, Correlation of location-specific macroclimate factors and atmospheric variables encountered during transit for air assemblages. b, Correlation of location-specific macroclimate factors with local soil assemblages. Blue circles denote positive correlations and red circles denote negative correlations. Circle colour intensity and size denote magnitude of correlation. MAT, mean annual temperature; MAP, mean annual precipitation; RH, relative humidity; UV, ultraviolet radiation. Abundance, qPCR estimated gene copy number; Richness, Chao1 estimation from rRNA gene diversity.
Extended Data Figure Legends

Extended Data Fig. 1 | Global patterns in alpha diversity for bacteria and fungi in air and soil. 

a, Total biomass as recoverable environmental DNA (n = 596). b-e, Bacterial abundance and alpha diversity metrics (n = 529); f-i, Fungal abundance and alpha diversity metrics (n = 444) estimated from amplicon sequence data. Abbreviations: MK, Mauna Kea; NZ, New Zealand; SA, South Africa; SO, Southern Ocean; HA air, high-altitude air; NG air, near-ground air.
Extended Data Fig. 2 | Comparison of diversity estimation using shotgun metagenomics and amplicon sequencing of air and soil. a, Relative abundance of twenty most abundant bacterial phyla estimated from metagenomes (n = 120); b, Relative abundance of twenty most abundant bacterial phyla estimated from amplicon sequencing (n = 529); c, Relative abundance of fungal phyla estimated from metagenomes (n = 120); d, Relative abundance of fungal phyla estimated from amplicon sequencing (n = 444). Diversity in metagenomes was estimated based upon classified reads, unclassified reads comprised on average 77.9% of sample reads (SD 13.1%). HA air, high-altitude air; NG air, near-ground air; NA, no taxonomy assigned. Locations are numbered as shown in Fig. 1.
Extended Data Fig. 3 | Global patterns in beta diversity for bacteria and fungi in air and soil. a-b, Jaccard Index for global bacterial (a) and fungal (b) assemblages, decomposed using PCoA. Clustering by habitat type (green, soil; blue, NG air; red, HA air) was also conserved when analysis was reiterated using only locations where concurrent sampling for all habitat types occurred in order to mitigate against potential sample-size effects. c-d, Hellinger Distance for global bacterial (a) and fungal (b) assemblages, decomposed as weighted community diversity-abundance using tSNE. HA air, high-altitude air; NG air, near-ground air. Separation by location is evident with the exception of a cluster for the regionally proximal New Zealand, Southern Ocean and Antarctic locations (Bacteria n = 529; Fungi n = 444).
Extended Data Fig. 4 | Global distribution of bacteria and fungi. 

a, Co-occurrence of bacterial (n = 529) and fungal (n = 444) ASV in globally distributed air and soil. 

b, Global-scale distance decay plots for bacteria and fungi in air and soil. Shaded area surrounding line of best fit indicates 95% confidence intervals. HA air, high-altitude air; NG air, near-ground air.
Extended Data Fig. 5 | Null model networks for bacterial and fungal assemblages in air and soil. a, High-altitude air. b, Near-ground air. c, Soil. For nestedness estimates from phylum to family level all taxa were used, and for genus and ASV level the 1,000 most prevalent taxa. The null models were based upon statistical mechanics reconstruction of the taxa by location bipartite network. Maximum-likelihood was used to estimate the probability distribution that maximised the entropy function of the null network conditional on the constraint of the observed degree sequence, which was enforced as an average vector (Canonical ensemble).
Extended Data Fig. 6 | Taxonomic composition of bacterial and fungal assemblages in air and soil. 
a, Relative abundance of bacterial classes in soil (n = 79), near-ground air (NG air) (n = 437) and high-altitude air (HA air) (n = 13). 
b, Relative abundance of fungal classes in soil (n = 70), near-ground air (NG air, n = 363) and high-altitude air (HA air, n = 11).
Extended Data Fig. 7 | Source contribution to observed diversity in air. a, Estimated contribution of aquatic, phyllosphere and soil sources to bacterial diversity. b, Estimated contribution of aquatic, phyllosphere and soil sources to fungal diversity. Location 3 (Spain) was impacted by a minor intrusion of Sahara Desert atmospheric dust during sampling. HA air, high-altitude air; NG air, near-ground air. Green shading indicates source locations sampled in this study, grey boxes indicates data from other studies employed in the meta-analysis. Coloured boxes indicate climate as shown in Fig. 2.
Extended Data Fig. 8 | Atmospheric variables encountered by microorganisms during transit to each sampling location. a, Temperature. b, Solar flux. c, Relative humidity. d, Transit over land/ocean. e, Transit altitude. f, Transit velocity. Data was obtained from the NOAA HYSPLIT-model and long-range trajectories and abiotic variables were estimated using the GDAS database (https://ready.arl.noaa.gov/HYSPLIT.php).
Methods

Global sampling

The sampling campaign retrieved 596 air and soil samples from 18 locations spanning all major climatic regions and continents plus two oceans. Several dryland (desert) locations were incorporated because they comprise the most abundant terrestrial biome on earth and are also the single largest natural source of particulate emissions to the atmosphere. The Southern Hemisphere was sampled during April-May 2019 and the Northern Hemisphere during June-July 2019 (Supplementary Information, Field sampling). The two circumnavigations encompassed every major climatic biome and included high and low growing seasons. Oceanic and remote land samples were retrieved during independent voyages (May 2017 – June 2018) using the same sampling methodology and samples from the two previously interrogated locations were re-sequenced for this study. Bulk phase boundary layer air was sampled at 1.5m above the surface (near-ground air, n = 501) using tripod-mounted air samplers and also above the boundary layer for surface interactions at 2,000 m above local surface level using aircraft mounted-air samplers (high-altitude air, n = 11). Concurrent sampling of underlying soil was conducted within a 25 m radius of air sampling devices (soil, n = 84). Ship-board sampling was conducted at 25m above the ocean surface to avoid sea-spray contamination. Logistical challenges limited high-altitude air sampling to six locations although these were nonetheless able to capture a broad geographic and climatic range for both hemispheres.

The use of high-throughput DNA sequencing for samples from low biomass habitats such as air raises the issue of confounding signal due to contaminants that are otherwise indistinguishable in higher biomass samples. We employed an experimental design for sample recovery and quality filtering of sequence data that embraced recommended best practice for minimising contaminant signal (Supplementary Information, Decontamination...
of environmental sequence data). Recovery of bulk phase air was achieved using three Coriolis μ high-volume impingement devices (Bertin Instruments, France) operated concurrently. This device has been shown to perform well against other samplers. All equipment was transported between locations in sterile containers and bags. Each device was dis-assembled and contact surfaces soaked for one hour with 1.5% v/v sodium hypochlorite (NaClO) followed by three washes of Milli-Q H₂O prior to and after each sampling in order to minimise contamination from cells or nucleic acids. All apparatus and work surfaces used during sampling and sample processing were also cleaned in this way prior to use. All operators wore surface sterilised nitrile gloves during field collections. Randomised collection cones were assembled into the devices without activating the air pump, and these were used as the negative sampling controls at each location. Additional control samples for potential human contamination were provided via swabs from the inside of anonymised used nitrile gloves (human operator controls).

Samplers were located 3m apart from each other at each sampling location and all inlets were aligned facing prevalent local wind direction. Bulk air was recovered at 300 L/min⁻¹ and particulates recovered after cyclonic deceleration into a sterile phosphate-buffered saline (PBS) impingement medium in each collection cone. Samplers were only approached from downwind during operation. Each device was used to collect discreet 18 m³ air samples as this volume has been shown to result in recoverable environmental DNA. Samples were recovered hourly between 10:00 – 16:00 hrs daily, and then processed immediately by syringe filtration onto a 25 mm polycarbonate filter with 0.2 µm pore size and preserved in 0.5 mL of DNA/RNA Shield (Zymo Research, USA) at ambient temperature during transit and then frozen at -20° C until processed for DNA extraction in the laboratory.
At each location undisturbed surface soil or sediment samples (upper 2cm soil captured in sterile 50 mL falcon tubes) were also collected from the base of each Coriolis µ device as well as 25 m away in 120° intervals from the point of sampling. In recognition of inherent soil heterogeneity each sample comprised five subsamples that were mixed and then resampled to yield each sample for analysis. For each sample 0.5 g was preserved for DNA extraction in 0.5 mL of DNA/RNA Shield (Zymo Research, USA) at ambient temperature during transit and then frozen at -20° C until processed for DNA extraction in the laboratory. The remaining sample fraction was archived. It was recognised that soil is not the primary reservoir for terrestrial fungal diversity but in the absence of a practical means to globally sample the diversity of other fungal substrates we accepted this limitation to the study.

DNA extractions from samples were performed in randomised sample batches each with discreet laboratory controls to assess potential laboratory or reagent contamination. Environmental DNA was recovered from filtered air and soil samples using a CTAB-based manual extraction protocol optimised for low biomass samples. DNA yield was quantified using the Qubit 2.0 Fluorometer (Invitrogen, USA) and samples were then stored at -20 °C until processed.

Environmental and climate metadata and modelling

Local climate metadata for each sampling location were retrieved from public databases: Mean annual precipitation and mean annual temperature, Köppen-Geiger climate classifications, growing season that defined the time period when photoautotrophy can occur. Other local variables were recorded using hand held devices: Temperature (°C), relative humidity (%), wind speed (m/S) and wind direction (Kestrel Meters, USA); Particulate matter (PM 2.5 and PM 10.0) (HoldPeak, China).
Back trajectories and metadata relevant to the *in situ* conditions that airborne microorganisms were exposed to during their transit towards each sampling location was modelled from National Oceanic and Atmospheric Administration (NOAA) databases. Fourteen day back trajectories of air masses for each sampling time were generated because this is the estimated maximum residence time for microorganisms in the troposphere. Data was obtained from the NOAA HYSPLIT-model and long range trajectories were estimated using the GDAS database (https://ready.arl.noaa.gov/HYSPLIT.php), Data were processed using ArcGIS Pro, version 2.6 (https://www.esri.com). The following variables were calculated at intervals along each tropospheric transport path from the HYSPLIT models: altitude (m above ground level, AGL), wind speed (m/S), direction, temperature (°C), Relative humidity (%), solar irradiance (W/m²), precipitation events, and transit duration over land or ocean surface. Points were plotted on WGS 1984 Web Mercator coordinate system with date line wrapping and climate mapping at 1km resolution.

**Biomass estimation**

DNA yield per m³ air or per gram of soil were used as a proxy for total biomass. Due to uncertainties surrounding currently available technology for the direct estimation of cell numbers using real-time fluorescence measurements, we opted for a well-established indirect approach using real-time quantitative PCR (qPCR) for the most abundant microbial groups (bacteria and fungi). Estimates were also triangulated with read abundance from our metagenomes with the awareness that metagenomic databases currently yield high numbers of unassigned reads across all domains. Primers used for bacteria targeted the 16S rRNA gene V3-V4 hypervariable region (Fwd 341 5′- CCTACGGGNGGCWGCAG-3′ and Rev 785 5′- GACTACHVGGGTATCTAATCC -3′) with LightCycler 480 SYBR Green I Master mix (Roche Holding, Switzerland). Primers for fungi targeted the 18S rRNA gene
(FungiQuant-F 5′-GGRAAACTCACCAGGTCCAG-3′ and FungiQuant-R 5′-
GSWCTATCCCACCKACGA-3′) with TaqMan probe FungiQuant-PrbLNA 6FAM-5′-
TGGTGCATGGCGT-3′-BBQ. TaqMan Fast Advanced Master Mix was used for qPCR
reactions with the following conditions: denaturing step: 95 °C for 20 s; cycling step: 35
cycles of 95 °C for 1 s and 60 °C for 20 s. A qPCR standard for each target sequence was
developed to estimate gene copy number using pooled samples. These were amplified using
TaqMan Fast Advanced Master Mix as described above but without fluorescent markers
(Applied biosystems, USA) and quantified using a Bioanalyzer (Agilent Technologies, USA).
Serial dilutions of the template were used to generate standard curves. Although estimates of
cell abundance using qPCR are flawed due to uncertainties over gene copy number among
diverse phyla, and the issue of multicellular fungi and taxonomic unit assignation, these
were not sources of systematic bias in our study since they applied to all samples equally.

**Amplicon sequencing**

Targeted high-throughput amplicon sequencing was employed to gain insight into taxonomic
diversity of bacteria and fungi because they comprised the majority of microbial reads in our
metagenomic libraries and have been estimated as the most abundant microorganisms in
aerosols. Amplicon sequence libraries were prepared using Illumina MiSeq v3 600 cycle
chemistry as per manufacturer’s protocol with PhiX positive spike-in controls. Template
DNA in samples was normalised to 2.5ng/ul prior to two-step PCR amplification for the
bacterial 16S rRNA gene V3-V4 hypervariable region (Fwd 341 5′-
CCTACGGGNGGCWGCAG-3′ and Rev 785 5′- GACTACHVGGGTATCTAATCC -3′)
13,14, and fungal ITS1 region (Fwd ITS1 5′- CTTGGTCATTAGAGGAGTAA -3′ and Rev
ITS2 5′-GCTGCGTCTTCATCGATGC -3′) 19,20. The amplicon sequence libraries were
first processed with cutadapt v2.7 21 to remove primer sequences. Amplicon sequence
variants (ASVs) were generated for 16S rRNA amplicons (truncLen=c(230,220), maxN=0, maxEE=c(2.5), truncQ=2) and ITS amplicons (minLen=50, maxN=0, maxEE=c(5,8), truncQ=2) dada2 v1.14. Pseudo-pooling was used in ASV calling to increase sensitivity and accuracy in alpha diversity estimation. Taxonomic classification was conducted in dada2 with SILVA v138 and UNITE v7.2 as references. After the decontamination steps, true samples with > 1,000 reads (16S rRNA n = 529, ITS n = 444) were used for all subsequent analyses. All the raw high-throughput sequencing data were deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession PRJEB42754.

Diversity estimation occurred only after a stringent decontamination protocol to mitigate putative contamination in sequence libraries from our ultra-low biomass air and soil samples (Supplementary Information, decontamination of environmental sequence data). This comprised subtractive filtering steps at the level of ASV as follows: 1) Removal of ASV with suspicious frequency and/or prevalence. 2) Removal of all ASV encountered in any of the sampling, human operator or laboratory controls; 3) A highly stringent subtractive filtering at genus level based on a meta-analysis of putative contaminants from other studies of low biomass samples regardless of whether or not they were also encountered in our sampling, laboratory or operator control samples. Suspiciously frequent and/or prevalent ASV were identified using the [isContaminant] function of the R package decontam and removed if they met the stringent statistical threshold for frequency (0.1) or prevalence (0.5). The prevalence test was used as a further check on the step for removal of ASV from controls. Putative human-associated contamination was addressed by removing all ASV affiliating with the following bacterial genera: Bacteroides, Bifidobacterium, Corynebacterium, Cutibacterium/Propionibacterium, Escherichia, Faecalibacterium, Haemophilus, Klebsiella, Lactobacillus, Listeria, Moraxella, Neisseria, Porphyromonas, Prevotella, Shigella, Salmonella, Staphylococcus, Streptococcus and Veillonella. We also removed ASV
affiliating with the human-associated fungal genera and fungi *Candida* and *Malassezia*. Other ASV affiliating with potential reagent contaminant genera were manually inspected and removed if they showed strong inverse correlation to DNA concentration in samples, with particular focus on the following genera: *Acidovorax, Acinetobacter, Afipia, Aquabacterium, Bacillus, Bradyrhizobium, Burkholderia, Clostridium, Dysgonomonas, Enterobacter, Kocuria, Methylobacterium, Pseudomonas, Ralstonia, Renibacterium, Rhizobacter, Romboutsia, Sphingobium, Sphingomonas, Tardiphaga, Turicibacter, Variovorax*. This latter approach reflected that many potential reagent contaminant genera include a number of environmental taxa that are reported as abundant and functionally important in the atmosphere (e.g. 16,27) and the air is recognised as a potential source of putative contaminants in studies of other habitats. Overall, the multi-step decontamination process identified 1,079 bacterial ASV and 229 fungal AV as suspected contaminants.

**Shotgun metagenomics**

Independent replicates were pooled by sampling day and device to yield 120 pooled samples and 3 pooled controls for metagenomics sequencing. Libraries were prepared using a low-input preparation protocol where required 28 and using the Nextera XT library kit and sequenced (2 × 150 bp paired-end) on an Illumina NextSeq 500 (Illumina, USA). Kneaddata (v0.7.4, default settings, https://github.com/biobakery/kneaddata) was used to remove low-quality reads and human DNA using the human genome hG37 as reference from raw fastq files.

Similar to the steps adopted for amplicon sequencing, filtered metagenomics reads were further processed in a multi-step fashion to systematically identify and remove potential contaminating nearest taxonomic units (NTU) (Supplementary Information, decontamination of environmental sequence data). Filtered reads from the controls were co-assembled into
contigs using the “assembly” module of MetaWRAP (v1.2.1). Reads in the samples that mapped to the contigs constructed (≥ 1,000 bp) in the controls were removed using Kneaddata. Next, taxonomic classification for NTU was performed using Kraken (v2.0.9-beta) based on the PlusPFP database (Dec 2nd, 2020 update) and species-level NTU classification was optimized using Bracken (v2.6.0). Fungal species were identified using FindFungi (v0.23.3). Species-level information from Kraken2 and FindFungi were processed to identify potential contaminating taxa using the same decontam settings applied to our ASV data. The decontam algorithm identified 23 fungal and 280 bacteria suspected contaminant NTU. Last, a genus level subtractive filtering for NTU commonly associated with contamination in other low biomass systems (i.e. as listed above for amplicon-based contaminant removal) was performed. The genus-level subtractive filtering identified an additional 7 fungal and 244 bacterial suspected contaminant NTU. All bacterial contaminants were subsequently removed using the “extract_kraken_reads.py” command (option --exclude and --include-children) from KrakenTools (v2.0.8-beta, https://github.com/jenniferlu717/KrakenTools). Reads cleared of bacterial contaminants were subjected to another round of contaminants removal using Kneaddata to discard reads that mapped to representative genomes of the fungal contaminants. Genus-level subtractive filtering was not applied to archaea or protists, and viruses were poorly represented in our metagenome libraries although methodological limitations may have reduced their detection. After all the decontamination steps, a total of 1,498,558,646 high-quality paired-end reads were retained across the entire dataset of 120 metagenomes, averaging 12,487,988 reads per sample. All the raw metagenomics sequencing data were deposited in the NCBI SRA under BioProject accession PRJNA694999.

Taxonomic profiles (phylum and species) of the high-quality reads were generated with Kraken2 and Bracken. In addition, FindFungi was used to identify fungal phyla. Functional potentials of the metagenomes were queried using HUMAnN (v3.0.0.alpha.3), generating a
total of 1,855,295 UniRef90 features, which were subsequently converted to 10,440 protein families (Pfams). Pfams corresponding to genes encoding carbon fixation, cold shock response, nitrogen cycle, oxidative stress, phototrophy, respiration, sporulation, starvation, trace gas metabolism, and UV repair proteins (Supplementary Information, Metagenomic functional analysis) were examined to understand factors shaping the abundance of metabolic and stress response potentials.

Statistical treatments and ecological modelling

Statistical analysis: General processing of the community data including the calculation of relative abundance and estimates of alpha diversity were conducted using the R package phyloseq\(^33\) and visualised using ggplot2\(^34\). Calculation of geographic distances were performed using R package geosphere\(^35\) function distGeo with WGS84 ellipsoid. Source tracking was conducted using FEAST\(^36\) with data from other studies (processed using dada2 following the same parameters as this study) as additional sources/sinks\(^37-43\) and NCBI BioProject PRJEB42801. For correlation analysis between abiotic and biotic variables the Pearson correlation coefficient for multiple pairwise combinations were calculated using the R package Corrplot\(^44\). In order to visualise patterns of community dissimilarity, two methods were used. Hellinger distances were ordinated with t-distributed stochastic neighbour embedding (tSNE) using R package Rtsne\(^45\). We also calculated Jaccard sample pair-wise distances based on the 10,000 most abundant and frequent ASVs using the R package vegan\(^46\). A preliminary analysis based on all reads and ASVs showed qualitatively similar patterns but higher noise (i.e. the amount of variance accounted for by the major ordination axes was relatively low due to a very high number of ASVs found only at one or two locations). We decomposed the Jaccard matrix with Principal Coordinate Analysis (PCoA) which provided a quantification of the variance accounted by each ordination axis\(^47\).
Network null models: A statistical mechanics approach was employed for network construction \(^{48}\), and defined our networks as bipartite matrices with two layers: location and taxa. Analyses were performed at multiple taxonomic ranks: Phylum, Class, Order, Family, Genus, and ASV. In order to fully test our hypothesis, we employed degree sequence constraints to enforce that for each taxon, the total number of locations in which the taxon was found was a constraint, and for each location the total number of taxa found in that location, disregarding location or taxa identity, was also a constraint. We used maximum-likelihood models \(^{49,50}\), to estimate the probability distribution that maximised entropy for the canonical ensembles. We sampled the resulting probability distribution to obtain 999 null matrices representing an unbiased sample of the canonical ensemble of our location by taxa matrices using the MatLab routine Max&Sam \(^{51}\), and imported the null matrices in R for downstream analyses. We used the R packages bipartite \(^{52}\) and vegan \(^{46}\) to calculate the nestedness metric of NODF and Jaccard dissimilarity on the observed and null matrices. We then used the standard definition of effect size \(^{53}\) to quantify the difference between observed metrics and the null distribution of the metrics. Since the distribution of the 999 null metrics were approximately normal, an effect size larger than 2 standard errors corresponded to taxonomic composition that diverged more than expected under purely random assembly with an approximate P-value < 0.05. We calculated Z-scores for nestedness to indicate the number of standard deviations a given data point lay from the mean using the commonly employed NODF metric metric \(^{54}\), and also using Jaccard dissimilarity as this estimates resemblance on average better than raw forms of the indices such as NODF, particularly in the face of confounding effects of spatial scale and conspecific aggregation \(^{55}\).

**Methods References**

1. Pointing, S. B. S. B. & Belnap, J. Disturbance to desert soil ecosystems contributes to
dust-mediated impacts at regional scales. *Biodivers. Conserv.* **23**, 1659–1667 (2014).

2. Archer, S. D. J. *et al.* Air mass source determines airborne microbial diversity at the ocean–atmosphere interface of the Great Barrier Reef marine ecosystem. *ISME J.* **14**, 871–876 (2020).

3. Archer, S. D. J. *et al.* Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nat. Microbiol.* **4**, 925–932 (2019).

4. Stull, R. B. A boundary layer definition. in *Introduction to Boundary Layer Meteorology* 1–27 (Kluwer Academic Publishers, 1988).

5. Eisenhofer, R. *et al.* Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol.* **27**, 105–117 (2019).

6. Dybwad, M., Skogan, G. & Blatny, J. M. Comparative testing and evaluation of nine different air samplers: End-to-end sampling efficiencies as specific performance measurements for bioaerosol applications. *Aerosol Sci. Technol.* **48**, 282–295 (2014).

7. University of East Anglia Climatic Research Unit; Harris, I.C.; Jones, P.D.; Osborn, T. No Title. *CRU TS4.04: Climatic Research Unit (CRU) Time-Series (TS) version 4.04 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901-Dec. 2019)* (2020).

8. Beck, H. E. *et al.* Present and future köppen-geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 1–12 (2018).

9. Fischer, G. *World Food and Agriculture to 2030/50: How do climate change and bioenergy alter the long-term outlook for food, agriculture and resource availability?* in *Expert Meeting on How to Feed the World in 2050* 1–49 (2009). doi:10.1111/j.1365-2230.2009.03481.x.

10. Burrows, S. M. *et al.* Bacteria in the global atmosphere – Part 2: Modeling of emissions and transport between different ecosystems. *Atmos. Chem. Phys.* **9**, 9281–
11. Huffman, J. A. et al. Real-time sensing of bioaerosols: Review and current perspectives. *Aerosol Sci. Technol.* **54**, 465–495 (2020).

12. Hospodsky, D., Yamamoto, N. & Peccia, J. Accuracy, precision, and method detection limits of quantitative PCR for airborne bacteria and fungi. *Appl. Environ. Microbiol.* **76**, 7004–7012 (2010).

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

14. Klindworth, A. et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1 (2013).

15. Liu, C. M. et al. FungiQuant: a broad-coverage fungal quantitative real-time PCR assay. *BMC Microbiol.* **12**, 255 (2012).

16. Gusareva, E. S. et al. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23299–23308 (2019).

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

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

19. Gardes, M. & Bruns, T. D. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol. Ecol.* **2**, 113–118 (1993).

20. White, T., Burns, T., Lee, S. & Taylor, J. No Title. in *PCR protocols: a guide to methods and applications* (eds. Innis, M., Gelfand, D., Sninsky, J. & White, T.) 315–322 (Academic Press, 1990).
21. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).

22. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W. & A. J. DADA2: High resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

23. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).

24. Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* **47**, D259–D264 (2019).

25. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).

26. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).

27. Vaïtilingom, M. *et al.* Potential impact of microbial activity on the oxidant capacity and organic carbon budget in clouds. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 559–564 (2013).

28. Rinke, C. *et al.* Validation of picogram- and femtogram-input DNA libraries for microscale metagenomics. *PeerJ* **4**, e2486 (2016).

29. Uritskiy, G. V, DiRuggiero, J. & Taylor, J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).

30. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).

31. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species
abundance in metagenomics data. PeerJ Comput. Sci. 3, e104 (2017).

32. Franzosa, E. A. et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat. Methods 15, 962–968 (2018).

33. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8, e61217 (2013).

34. Wickham, H. ggplot2. (Springer New York, 2009). doi:10.1007/978-0-387-98141-3.

35. Hijmans, R. J. R package: Geosphere: spherical trigonometry. (2019).

36. Shenhav, L. et al. FEAST: fast expectation-maximization for microbial source tracking. Nat. Methods 16, 627–632 (2019).

37. Ueki, T., Fujie, M., Romaidi & Satoh, N. Symbiotic bacteria associated with ascidian vanadium accumulation identified by 16S rRNA amplicon sequencing. Mar. Genomics 43, 33–42 (2019).

38. Purahong, W. et al. Plant Microbiome and Its Link to Plant Health: Host Species, Organs and Pseudomonas syringae pv. actinidiae Infection Shaping Bacterial Phyllosphere Communities of Kiwifruit Plants. Front. Plant Sci. 9, 1563 (2018).

39. Hermans, S. M., Buckley, H. L., Case, B. S. & Lear, G. Connecting through space and time: catchment-scale distributions of bacteria in soil, stream water and sediment. Environ. Microbiol. 22, 1000–1010 (2020).

40. Yang, J. et al. Habitat filtering shapes the differential structure of microbial communities in the Xilingol grassland. Sci. Rep. 9, 19326 (2019).

41. Wang, G. et al. Soil microbial legacy drives crop diversity advantage: Linking ecological plant–soil feedback with agricultural intercropping. J. Appl. Ecol. (2021) doi:https://doi.org/10.1111/1365-2664.13802.

42. Carini, P. et al. Effects of Spatial Variability and Relic DNA Removal on the Detection of Temporal Dynamics in Soil Microbial Communities. MBio 11,
Szoboszlay, M., Näther, A., Mullins, E. & Tebbe, C. C. Annual replication is essential in evaluating the response of the soil microbiome to the genetic modification of maize in different biogeographical regions. *PLoS One* **14**, e0222737 (2019).

Wei, T. & Simco, V. R package: Corrplot: Visualisation of a correlation matrix. (2017).

Krijthe, J. E. Rtsne: t-distributed stochastic neighbor embedding using a Barnes-Hut implementation. (2015).

Oksanen, J. *et al.* R package: Vegan. (2017).

Legendre, P., Borcard, D. & Peres-Neto, P. R. Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecol. Monogr.* **75**, 435–450 (2005).

Cimini, G. *et al.* The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71 (2019).

Squartini, T. & Garlaschelli, D. *Maximum-entropy networks: Pattern detection, network reconstruction and graph combinatorics*. (Springer, 2017).

Squartini, T. & Garlaschelli, D. Analytical maximum-likelihood method to detect patterns in real networks. *New J. Phys.* **13**, 83001 (2011).

Rossana. MATLAB package: MAX&SAM. (2021).

Dormann, C. F., Frund, J., Bluthgen, N. & Gruber, B. Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *Open Ecol. J.* **2**, 7–24 (2009).

Gotelli, N. J. & Ellison, A. M. *A primer of ecological statistics*. (Oxford University Press, 2012).

Almeida-Neto, M., Guimarães, P., Guimarães, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept
and measurement. *Oikos* **117**, 1227–1239 (2008).

55. Keil, P. Z-scores unite pairwise indices of ecological similarity and association for binary data. *Ecosphere* **10**, e02933 (2019).
Diverse recruitment to a globally structured atmospheric microbiome

Supplementary Information
Decontamination of environmental sequence data

Fig. S1. Summary of steps taken to minimise contamination during environmental DNA recovery and sequencing from low biomass air and soil samples.
Fig. S2 | Outcome of robust decontamination effort for amplicon sequencing of ultra-low biomass environmental samples. As per recommended best practice\(^1\), we report the read depth and taxonomic composition of taxa removed during extensive decontamination effort that comprised screening and subtractive filtering of taxa encountered in any of the sampling, laboratory or human operator controls; removal of taxa with suspicious frequency and/or prevalence; and a stringent genus-level filtering of putative contaminants identified in a meta-analysis of other studies on low biomass systems. a, Bacteria read depth (controls: \( n = 35 \); samples: \( n = 529 \), 67 samples had <1,000 reads after filtering). b, Fungi read depth (controls: \( n = 35 \); samples: \( n = 444 \), 152 samples had <1,000 reads after filtering). Boxplots indicate median (line) and interquartile range (boxes). Red symbols denote human operator controls. c, Taxonomic profile of putative contaminant bacterial ASV, and d, fungal ASV removed from soil, near-ground air (NG air) and high-altitude air (HA air) samples. Relative abundance indicates occurrence as a percentage of all contaminant ASV and not the occurrence of legitimate sample ASV. Locations: 1, Canada; 2, Mongolia; 3, Spain; 4, Japan; 5, California, USA; 6, Kuwait; 7, Hilo, Hawaii, USA; 8, Mauna Kea, Hawaii, USA; 9, Singapore; 10, Coral Sea; 11, Namibia; 12, Chile; 13, South Africa; 14, Australia; 15, Uruguay; 16, New Zealand; 17, Southern Ocean; 18, Antarctica. Location metadata are shown in Supplementary Information.
Fig. S3 | Outcome of robust decontamination effort for shotgun metagenomics sequencing of ultra-low biomass environmental samples. As per recommended best practice, we report the read depth and taxonomic composition of taxa removed during extensive decontamination effort that comprised screening and subtractive filtering of taxa encountered in any of the sampling, laboratory or human operator controls; removal of taxa with suspicious frequency and/or prevalence; and a stringent genus-level filtering of putative contaminants identified in a meta-analysis of other studies on low biomass systems. a, Metagenomes read depth for controls (n = 3, pooled by type from 35 independent control samples) and samples (n = 120). b, Taxonomic profile of putative contaminant bacterial NTU, and c, fungal NTU removed from soil, near-ground air (NG air) and high-altitude air (HA air) samples. No fungal contaminants were recovered from soil at location 18. Relative abundance indicates occurrence as a percentage of all contaminant NTU and not the occurrence of legitimate sample NTU. Locations: 1, Canada; 2, Mongolia; 3, Spain; 4, Japan; 5, California, USA; 6, Kuwait; 7, Hilo, Hawaii, USA; 8, Mauna Kea, Hawaii, USA; 9, Singapore; 10, Coral Sea; 11, Namibia; 12, Chile; 13, South Africa; 14, Australia; 15, Uruguay; 16, New Zealand; 17, Southern Ocean; 18, Antarctica. Location metadata are shown in Supplementary Information.
Table S1. Summary of decontamination effort reported for other environmental sequencing studies of low biomass air and soil microbial communities. A tick denotes the step was reported, NR denotes not reported.

| Habitat               | Publication | Sampling controls | Reagent controls | Human operator controls | Control sequences removed | Suspectively frequent / prevalent taxa removed | Subtractive filtering of common contaminant taxa | Contaminants reported |
|-----------------------|-------------|------------------|------------------|------------------------|--------------------------|---------------------------------------------|-----------------------------------------------|-----------------------|
| Air and soil          | This study  | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Archer et al., 2020 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Signoret-Porier et al., 2020 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Uetake et al., 2020 | NR              | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air and soil          | Archer et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Els et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Signoret-Porier et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Tignat-Perrier et al., 2020 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Archer et al., 2020 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Tignat-Perrier et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Guseva et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Els et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Signoret-Porier et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Tignat-Perrier et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Guseva et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Cloud water           | Amato et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Rain & snow           | Caliz et al., 2019 | NR               | NR               | NR                     | NR                       | NR                                          | NR                                             | NR                    |
| Air                   | Maki et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Santl-Temkiv et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Smith et al., 2018 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Global soils          | Delgado-Baquerizo et al., 2018 | NR              | NR               | NR                     | NR                       | NR                                          | NR                                             | NR                    |
| Global soils          | Els et al., 2019 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Mayol et al., 2017 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Cloud water           | Amato et al., 2017 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Maki et al., 2017 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Global soils          | Thompson et al., 2017 | NR              | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Desert soil           | Lee et al., 2016 | NR               | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Desert soil           | Grothem et al., 2016 | NR              | NR               | NR                     | NR                       | NR                                          | NR                                             | NR                    |
| Desert dust           | Barberan et al., 2014 | NR              | NR               | NR                     | NR                       | NR                                          | NR                                             | NR                    |
| Air                   | Bottos et al., 2013 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Woe et al., 2013 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Desert dust           | Favet et al., 2013 | NR              | NR               | NR                     | NR                       | NR                                          | NR                                             | NR                    |
| Air (aeroplane)       | Delaunay-Rodriguez et al., 2013 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
| Air                   | Bowers et al., 2011 | ✓                | ✓                | ✓                      | ✓                        | ✓                                           | ✓                                             | ✓                     |
Field sampling

Table S2. Globally distributed sampling locations included in the study. Locations are ordered by latitude from north to south. Full metadata for each sampling location and event are available upon request at http://atmospheric-microbiome.com/.

| Location No | GPS (decimal degrees) | Atmospheric cell<sup>1</sup> | Climate<sup>+</sup> | Altitude (m AMSL) | Country/Ocean | Surface cover |
|-------------|------------------------|-----------------------------|------------------|------------------|---------------|---------------|
| 1           | 69.131, -105.057       | N Polar                     | Polar (EF)       | 50               | Canada        | Arctic tundra |
| 2           | 44.573, 105.648        | N Ferrel                    | Dry (BWk)        | 1,235            | Mongolia      | Steppe        |
| 3           | 40.825, -3.961         | N Ferrel                    | Continental (Dsby)| 1,830            | Spain         | Grassland, Sierra de Guadarrama National Park<sup>*</sup> |
| 4           | 37.308, 137.232        | N Ferrel                    | Temperate (Cfa)  | 6                | Japan         | Coastal forest |
| 5           | 35.142, -116.104       | N Ferrel                    | Dry (BWk)        | 300              | USA, California | Mojave Desert |
| 6           | 28.951, 48.192         | N Hadley                    | Dry (BWh)        | 0                | Kuwait        | Kuwait Desert  |
| 7           | 19.703, -155.090       | N Hadley                    | Tropical (Af)    | 120              | USA, Hawaii (Hilo) | Coastal forest |
| 8           | 19.823, -155.478       | N Hadley                    | Polar (ET)       | 4,200            | USA, Hawaii (Mauna Kea) | Mountain |
| 9           | 1.306, 103.772         | Equatorial                  | Tropical (Af)    | 19               | Singapore     | Tropical forest, suburban |
| 10          | -20.827, 153.067       | S Hadley                    | Tropical ocean   | 10               | Coral Sea     | Ocean          |
| 11          | -23.603, 15.038        | S Hadley                    | Dry (BWh)        | 380              | Namibia       | Namib Desert   |
| 12          | -24.105, -70.016       | S Hadley                    | Dry (BWh)        | 90               | Chile         | Atacama Desert |
| 13          | -25.753, -28.258       | S Hadley                    | Temperate (Cwb)  | 1,380            | South Africa  | Livestock and arable farmland |
| 14          | -32.898, 116.906       | S Ferrel                    | Temperate (Csa)  | 320              | Australia     | Arable farmland |
| 15          | -34.354, -57.235       | S Ferrel                    | Temperate (Cfa)  | 10               | Uruguay       | Rural grassland |
| 16          | -36.916, 174.646       | S Ferrel                    | Temperate (Cfb)  | 61               | New Zealand   | Suburban       |
| 17          | -60.288, 171.424       | S Polar                     | Polar ocean      | 10               | Southern Ocean | Ocean          |
| 18          | -75.520, 163.949       | S Polar                     | Polar (EF)       | 98               | Antarctica    | Polar desert   |

<sup>1</sup> Atmospheric cells describe the prevailing air movement in latitude-defined cells: Polar Cells occur at 60° and higher latitudes, Ferrel Cells at mid-latitudes between 30-60°, and Hadley Cells at latitudes of 30° and below.

<sup>+</sup> Climate codes follow the Köppen climate classification, major delineations were: A, tropical; B, dry; C, temperate; D, continental, E, polar.<sup>28</sup>

<sup>*</sup>This location was affected by a Sahara Desert atmospheric dust intrusion during sampling.
Inventory of taxa

Table S3. List of bacterial genera prevalent in ≥ 50% of soil and/or air samples.

| Prevalent in soil only | Prevalent in both | Prevalent in air only |
|------------------------|-------------------|----------------------|
| Acidovorax             | Abditibacterium   | Acidibacter          |
| Acinetobacter          | Actinoplanes      | Amaricoccus          |
| Ammoniphilus           | Adhaeribacter,    | Archangium           |
|                        | Allorhizobium-    |                      |
|                        | Neorhizobium-     |                      |
|                        | Pararhizobium-    |                      |
|                        | Rhizobium,        |                      |
|                        | Altererythrobacter|                      |
|                        | Arthrobacter,     |                      |
|                        | Bacillus,         |                      |
| Aquabacterium          | Bdello vibrio,    |                      |
|                        | Belnapia,         |                      |
|                        | Blastocella,      |                      |
| Atopostipes            | Blastococcus,     |                      |
|                        | Bryobacter,       |                      |
|                        | Candidatus        |                      |
|                        | Alysiopsphaera,   |                      |
|                        | Canditanus        |                      |
|                        | Udaeobacter,      |                      |
|                        | Cellulomonas,     |                      |
|                        | Chthoniobacter,   |                      |
|                        | Conexibacter,     |                      |
|                        | Devosia,          |                      |
|                        | Ellin6055,        |                      |
|                        | Flaviaesturariibacter, |                |
|                        | Flavisolibacter,  |                      |
|                        | Flavobacterium,   |                      |
|                        | Friedmanniella,   |                      |
|                        | Gaiella,          |                      |
|                        | Gemmatimonas,     |                      |
|                        | Geodermatophilus, |                      |
|                        | Haliangium,       |                      |
|                        | Hymenobacter,     |                      |
|                        | Lautropia,        |                      |
|                        | Luteitalea,       |                      |
|                        | Lyso bacter,      |                      |
|                        | Marmoricola,      |                      |
|                        | Massilia,         |                      |
|                        | Methylobacterium- |                      |
|                        | Methylorubrum,    |                      |
|                        | Micromonospora,   |                      |
|                        | Microvirga,       |                      |
|                        | Modestobacter,    |                      |
|                        | Mycobacterium,    |                      |
|                        | Nocardioides,     |                      |
|                        | Noviherbaspirillum,|                      |
|                        | Novosphingobium,  |                      |
|                        | Oligoflexus,      |                      |
|                        | Pedobacter,       |                      |
|                        | Peredibacter,     |                      |
|                        | Pirellula,        |                      |
|                        | Pontibacter,      |                      |
|                        | Pseudomonas,      |                      |
|                        | Pseudonocardia,   |                      |
|                        | Psychroglaciecola,|                      |
|                        | Quadrisphaera,    |                      |
|                        | Ramlibacter, RB41,|                      |
|                        | Rhizobacter,      |                      |
|                        | Rhodocytophaga,   |                      |
|                        | Roseisolibacter,  |                      |
|                        | Roseomonas,       |                      |
|                        | Rubellimicrobium, |                      |
|                        | Rubrobacter,      |                      |
|                        | Segetibacter,     |                      |
|                        | Skermanella,      |                      |
|                        | Solirubrobacter,  |                      |
|                        | Sphingomonas,     |                      |
|                        | Spirosoma,        |                      |
|                        | Sternotrophobacter,|                      |
|                        | Steroidobacter,   |                      |
|                        | Streptomyces,     |                      |
|                        | Subgroup_10,      |                      |
|                        | Sumerlaea,        |                      |
|                        | Truepera,         |                      |
|                        | Variovorax,       |                      |
|                        | YC-ZSS-LKJ147     |                      |
| Deinococcus            | Ellin517          |                      |
| Domibacillus           | Ellin6067         |                      |
| Georgiaea              | Ferruginibacter   |                      |
| Kineococcus            | Flavibacter,      |                      |
| Kocuria                | Flavibacter,      |                      |
| Limnobacter            | Flavibacter,      |                      |
| Longimicrobium         | Flavibacter,      |                      |
| Lateimonas             | Geodermatophilus, |                      |
| Lysinibacillus         | Haliangium,       |                      |
| Microbacterium         | Hymenobacter,     |                      |
| Oceanobacillus         | Lautropia,        |                      |
| Ornithininicrobium     | Luteitalea,       |                      |
| Paenibacillus          | Lyso bacter,      |                      |
| Paenichlostridium      | Marmoricola,      |                      |
| Pantoea                | Massilia,         |                      |
| Paracoccus             | Methylorubrum,    |                      |
| Planococcus            | Micromonospora,   |                      |
| Planomicrobium         | Microvirga,       |                      |
| Pseudarthrobacter      | Modestobacter,    |                      |
| Rufibacter             | Mycobacterium,    |                      |
| Salinicoccus           | Nocardioides,     |                      |
| Salinimicrobium        | Noviherbaspirillum,|                      |
| Soliacillus            | Novosphingobium,  |                      |
| Sporosarcina           | Oligoflexus,      |                      |
| Tumebacillus           | Pedobacter,       |                      |
| Turricibacter          | Peredibacter,     |                      |
| UCG-005                | Pirellula,        |                      |

4

5
Table S4. List of fungal genera prevalent in ≥ 50% of soil and/or air samples.

| Prevalent in soil only | Prevalent in both | Prevalent in air only |
|------------------------|-------------------|-----------------------|
| Acremonium             | Alternaria        | Exophiala             |
| Agaricus               | Aspergillus       | Knufia                |
| Botrytis               | Bipolaris         | Mortierella           |
| Ceriporia              | Chaetomium        | Oedocephalum          |
| Cladosporium           | Coprinellus       | Powellomyces          |
| Dioszegia              | Coprinopsis       | Rhizophlyctis         |
| Ganoderma              | Curvularia        | Saitozyma             |
| Gymnopus               | Didymella         | Stagonosporopsis      |
| Mycosphaerella         | Filobasidium      | Westerdykella         |
| Neoascochyta           | Fomitopsis        |                       |
| Nigrospora             | Fusarium          |                       |
| Paradendryphiella      | Gibberella        |                       |
| Paraphaeosphaeria      | Naganishia        |                       |
| Phellinus              | Neocamarosporium  |                       |
| Phlebiopsis            | Papiliotrema      |                       |
| Psathyrella            | Penicillium       |                       |
| Resinicium             | Peniophora        |                       |
| Sarocladium            | Periconia         |                       |
| Selenophoma            | Phaeococcomyces   |                       |
| Spegazzinia            | Phaeosphaeria     |                       |
| Torula                 | Phanerochaete     |                       |
| Toxicocladosporium     | Phlebia           |                       |
| Trametes               | Preussia          |                       |
| Tranzscheliella        | Pseudopithomyces  |                       |
| Tulostoma              | Vishniacozyma     |                       |
| Udeniomyces            |                   |                       |
| Ustilago               |                   |                       |
| Wallemia               |                   |                       |
Table S5. Fungi:bacteria abundance ratios in globally distributed air and soil. Estimates were based upon quantitative PCR of rRNA genes (B, bacteria \( n = 529 \); F, fungi \( n = 444 \)).

| Habitat type | F<B (rank order) | F>B (rank order) |
|--------------|------------------|------------------|
| High-altitude air | Polar desert, Antarctica | Arable, Australia |
| | Desert, Namibia | Desert, Namibia |
| | Desert, Chile | Desert, Chile |
| | Desert, USA | Desert, USA |
| Near-ground air | Coral Sea | Suburban, New Zealand |
| | Southern Ocean | Desert, USA |
| | Desert, Chile | Tropical forest-suburban, Singapore* |
| | Polar desert, Antarctica | Desert, Kuwait |
| | Mountain summit, Hawaii | Forest, Japan* |
| | Grassland, Spain* | Arable, Australia |
| | | Desert, Namibia |
| | | Arctic tundra |
| | | Grassland, Spain |
| | | Forest, Hawaii* |
| | | Grassland, Uruguay* |
| | | Arable, South Africa |
| | | Steppe, Mongolia |
| Soil | Arctic tundra | Arable, Australia |
| | Desert, Namibia | Desert, Chile |
| | Steppe, Mongolia | Mountain summit, Hawaii |
| | Polar desert, Antarctica | Desert, Kuwait |
| | | Grassland, Spain |
| | | Grassland, Uruguay |
| | | Desert, USA |
| | | Arable, South Africa |

* Spain location was surrounded by drylands over which recent air transit occurred, and was also affected by a Sahara Desert atmospheric dust intrusion during sampling.
+ denotes locations where metagenomes also indicated F>B. It should be noted that estimation of relative abundance between domains using metagenome reads has many uncertainties when using existing bioinformatics approaches.
Metagenomic functional analysis

Table S6. Functional genes targeted in the metagenomic inquiry of air and soil. A suite of respiratory genes was used as a general marker of potential for metabolically active taxa, and targeted metabolic and stress response genes were selected based upon substrates and stressors encountered in the atmospheric habitat. No hits were recorded for *ina* genes and this likely reflects low homology between taxa.

| Trait                              | Protein                                                                 | Gene    |
|------------------------------------|-------------------------------------------------------------------------|---------|
| Respiration                        | ATP synthase                                                           | *atpA*  |
|                                    | Cytochrome cbb3 oxidase                                                | *ccoN*  |
|                                    | Cytochrome aa3 oxidase                                                 | *coxA*  |
|                                    | Cytochrome bd oxidase                                                  | *cydA*  |
|                                    | Cytochrome bo3 oxidase                                                 | *cyoA*  |
|                                    | NADH-ubiquinone oxidoreductase subunit F                               | *nuoF*  |
|                                    | Succinate dehydrogenase/fumarate reductase                            | *sdhA*  |
|                                    |                                                                        | *fdrA*  |
| Carbon fixation                    | Ribulose 1,5-bisphosphate carboxylase/oxygenase                        | *rbcL*  |
| Nitrogen fixation                  | Nitrogenase                                                            | *nifH*  |
| Phototrophy                        | Photosystem I reaction centre protein                                  | *psaA*  |
|                                    | Photosystem II reaction centre protein                                  | *psbA*  |
|                                    | Microbial rhodopsin                                                    | *RHO*   |
| Atmospheric trace gas metabolism   | Aerobic carbon monoxide dehydrogenase                                  | *coxL*  |
| (carbon monoxide, hydrogen, methane, isoprene) | Hydrogenase                                                           | *Fe*    |
|                                    | Hydrogenase                                                            | *FeFe*  |
|                                    | Isoprene oxidation gene                                                | *isoA/mmoX* |
|                                    | Soluble methane monooxygenase                                          | *mmoX*  |
|                                    | Particulate methane monooxygenase                                      | *pmoA*  |
| cold-shock                         | Cold shock protein A                                                   | *cspA*  |
|                                    | Cold shock protein B                                                   | *cspB*  |
|                                    | Cold shock protein G                                                  | *cspG*  |
|                                    | Cold shock protein I                                                  | *cspI*  |
| Oxidative stress                   | Protein-methionine sulfoxide reductase                                 | *msrP*  |
|                                    | Protein-methionine sulfoxide reductase                                 | *msrQ*  |
| Sporulation                        | Sporulation protein 0A                                                 | *spo0A* |
| Starvation/stationary phase        | DNA-protection during starvation protein                               | *dps*   |
|                                    | Outer membrane protein                                                | *slp*   |
| UV response/repair                 | Deoxyribodipyrimidine photo-lyase                                      | *phrB*  |
|                                    | UvrABC system protein A                                               | *uvrA*  |
|                                    | UvrABC system protein B                                               | *uvrB*  |
|                                    | UvrABC system protein C                                               | *uvrC*  |
| Ice nucleation                     | Ice nucleation protein A                                              | *inaA*  |
|                                    | Ice nucleation protein X                                              | *inaX*  |
|                                    | Ice nucleation protein W                                              | *inaW*  |
|                                    | Ice nucleation protein Z                                              | *inaZ*  |
Supplementary Information References

1. Eisenhofer, R. et al. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol. 27, 105–117 (2019).
2. Archer, S. D. J. et al. Air mass source determines airborne microbial diversity at the ocean–atmosphere interface of the Great Barrier Reef marine ecosystem. ISME J. 14, 871–876 (2020).
3. Tignat-Perrrier, R. et al. Seasonal shift in airborne microbial communities. Sci. Total Environ. 716, 137129 (2020).
4. Uetake, J. et al. Airborne bacteria confirm the pristine nature of the Southern Ocean boundary layer. Proc. Natl. Acad. Sci. U. S. A. 117, 13275–13282 (2020).
5. Archer, S. D. J. et al. Airborne microbial transport limitation to isolated Antarctic soil habitats. Nat. Microbiol. 4, 925–932 (2019).
6. Els, N. et al. Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation. Aerobiologia. 35, 671–701 (2019).
7. Tignat-Perrrier, R. et al. Global airborne microbial communities controlled by surrounding landscapes and wind conditions. Sci. Rep. 9, 1–11 (2019).
8. Gusareva, E. S. et al. Microbial communities in the tropical air ecosystem follow a precise diel cycle. Proc. Natl. Acad. Sci. U. S. A. 116, 23299–23308 (2019).
9. Amato, P. et al. Metatranscriptomic exploration of microbial functioning in clouds. Sci. Rep. 9, 4383 (2019).
10. Cäliz, J., Triadó-Margarit, X., Camarero, L. & Casamayor, E. O. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. Proc. Natl. Acad. Sci. U. S. A. 115, 12229–12234 (2018).
11. Maki, T. et al. Aeolian Dispersal of Bacteria Associated With Desert Dust and Anthropogenic Particles Over Continental and Oceanic Surfaces. J. Geophys. Res. Atmos. 124, 5579–5588 (2019).
12. Štekl-Temkiv, T., Gosewinkel, U., Starnawski, P., Lever, M. & Finster, K. Aeolian dispersal of bacteria in southwest Greenland: Their sources, abundance, diversity and physiological states. FEMS Microbiol. Ecol. 94, fiy031 (2018).
13. Smith, D. J. et al. Airborne Bacteria in Earth’s Lower Stratosphere Resemble Taxa Detected in the Troposphere: Results From a New NASA Aircraft Bioaerosol Collector (ABC). Front. Microbiol. 9, 1752 (2018).
14. Delgado-Baquerizo, M. et al. A global atlas of the dominant bacteria found in soil. Science. 359, 320–325 (2018).
15. Bahram, M. et al. Structure and function of the global topsoil microbiome. Nature 560, 233–237 (2018).
16. Mayol, E. et al. Long-range transport of airborne microbes over the global tropical and subtropical ocean. Nat. Commun. 8, 201 (2017).
17. Amato, P. et al. Active microorganisms thrive among extremely diverse communities in cloud water. PLoS One 12, e0182869 (2017).
18. Maki, T. et al. Variations in airborne bacterial communities at high altitudes over the Noto Peninsula (Japan) in response to Asian dust events. Atmos. Chem. Phys. 17, 11877–11897 (2017).
19. Thompson, L. R. et al. A communal catalogue reveals Earth’s multiscale microbial diversity. Nature 551, 457–463 (2017).
20. Lee, K. C. et al. Niche Filtering of Bacteria in Soil and Rock Habitats of the Colorado Plateau Desert, Utah, USA. Front. Microbiol. 7, 1489 (2016).
21. Van Goethem, W. M., Makhalanyane, T. P., Valverde, A., Cary, S. C. & Cowan, D. A. Characterization of bacterial communities in lithobions and soil niches from Victoria Valley, Antarctica. FEMS Microbiol. Ecol. 92, fiw051 (2016).
22. Barberán, A., Henley, J., Fierer, N. & Casamayor, E. O. Structure, inter-annual recurrence, and global-scale connectivity of airborne microbial communities. Sci. Total Environ. 487, 187–195 (2014).
23. Bottos, E. M. E. M., Woo, A. C. A. C., Zawar-Reza, P., Pointing, S. B. S. B. & Cary, S. C. S. C. Airborne Bacterial Populations Above Desert Soils of the McMurdo Dry Valleys, Antarctica. Microb. Ecol. 67, 120–128 (2013).
24. Woo, A. C. et al. Temporal variation in airborne microbial populations and microbially-derived allergens in a tropical urban landscape. Atmos. Environ. 74, 291–300 (2013).
25. Favet, J. et al. Microbial hitchhikers on intercontinental dust: catching a lift in Chad. ISME J. 7, 850–867 (2013).
26. DeLeon-Rodriguez, N. et al. Microbiome of the upper troposphere: Species composition and prevalence, effects of tropical storms, and atmospheric implications. Proc. Natl. Acad. Sci. U. S. A. 110, 2575–2580 (2013).
27. Bowers, R. M., McLeitchie, S., Knight, R. & Fierer, N. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. ISME J. 5, 601–612 (2011).
28. Peel, M. C. & Finlayson, B. L. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633–1644 (2007).