



**QUEEN'S  
UNIVERSITY  
BELFAST**

## **Airborne microbial transport limitation to isolated Antarctic soil habitats**

Archer, S. D.J., Lee, K. C., Caruso, T., Maki, T., Lee, C. K., Cary, S. C., Cowan, D. A., Maestre, F. T., & Pointing, S. B. (2019). Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nature Microbiology*.

**Published in:**  
Nature Microbiology

**Document Version:**  
Peer reviewed version

**Queen's University Belfast - Research Portal:**  
[Link to publication record in Queen's University Belfast Research Portal](#)

**Publisher rights**

© 2019 The Authors. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

**General rights**

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

1 **Airborne microbial transport limitation to isolated Antarctic soil habitats**

2 Stephen D.J. Archer<sup>1,2</sup>, Kevin C. Lee<sup>2</sup>, Tancredi Caruso<sup>3</sup>, Teruya Maki<sup>4</sup>, Charles K. Lee<sup>5</sup>, S.  
3 Craig Cary<sup>5</sup>, Don A. Cowan<sup>6</sup>, Fernando T. Maestre<sup>7</sup>, Stephen B. Pointing<sup>1,8,9</sup>

4 *<sup>1</sup> Yale-NUS College, National University of Singapore, Singapore 138527*

5 *<sup>2</sup> Institute for Applied Ecology New Zealand, Auckland University of Technology, Auckland*  
6 *1142, New Zealand*

7 *<sup>3</sup> School of Biological Sciences and Institute for Global Food Security, Queen's University*  
8 *Belfast, Belfast BT9 7BL, Northern Ireland, UK*

9 *<sup>4</sup> Department of Chemical Engineering, Kanazawa University, Kanazawa 920-1192, Japan*

10 *<sup>5</sup> International Centre for Terrestrial Antarctic Research, University of Waikato, Hamilton*  
11 *3240, New Zealand*

12 *<sup>6</sup> Centre for Microbial Ecology and Genomics, University of Pretoria, Pretoria 0002, South*  
13 *Africa*

14 *<sup>7</sup> Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de*  
15 *Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933*  
16 *Móstoles, Spain*

17 *<sup>8</sup> Department of Biological Sciences, National University of Singapore, Singapore 117558*

18 *<sup>9</sup> Institute of Nature and Environmental Technology, Kanazawa University, Kanazawa*  
19 *920-1192, Japan*

20

1           Dispersal is a critical yet poorly understood factor underlying macroecological  
2 patterns in microbial communities <sup>1</sup>. Airborne microbial transport is assumed to occupy  
3 a central role in determining dispersal outcomes <sup>2,3</sup> and extra-range dispersal has  
4 important implications for predicting ecosystem resilience and response to  
5 environmental change <sup>4</sup>. One of the most pertinent biomes in this regard is Antarctica  
6 given its geographic isolation and vulnerability to climate change and human  
7 disturbance <sup>5</sup>. Here we report microbial diversity in near-ground and high-altitude air  
8 above the largest ice-free Antarctic habitat as well as that of underlying soil microbial  
9 communities. We found that persistent airborne inputs were unable to fully explain  
10 local soil community assembly. Comparison with airborne microbial diversity from non-  
11 polar sources suggests that strong selection occurs during atmospheric transport  
12 resulting in regionally restricted airborne inputs. Highly specialized soil communities  
13 occurred where fungi displayed greater isolation than bacteria from non-polar sources.  
14 Overall, microbial communities from this isolated Antarctic ecosystem displayed limited  
15 connectivity to the non-polar microbial pool and alternative sources of recruitment are  
16 necessary to fully explain extant soil diversity. Our findings provide critical insights to  
17 forecast the potential outcomes for microbial communities of climate change-mediated  
18 shifts in air circulation to Antarctica.

19  
20           Airborne microbial transport on fine particulate matter has typically been  
21 regarded as neutral and ubiquitous due to the small size and survivability of cells <sup>6-8</sup>.  
22 Diverse microbial signatures have been recovered from airborne particulates after  
23 transport across inter-continental distances <sup>3,8,9</sup>. Large scale patterns in microbial  
24 diversity are therefore often viewed as developing largely due to deterministic niche-  
25 driven processes <sup>10,11</sup>. The assumption that airborne transport is a neutral and

1 ubiquitous process has, however, been challenged <sup>12</sup> and we postulate that  
2 deterministic dispersal drivers may help explain extant diversity patterns in isolated  
3 microbial communities. The McMurdo Dry Valleys comprise the largest ice-free habitat  
4 in Antarctica and one of the most isolated soil habitats on Earth. They are devoid of  
5 vascular plants and dominated by highly specialised soil microbial communities <sup>13</sup> that  
6 display adaptations to the extreme environmental conditions <sup>14</sup>. Some taxa, notably  
7 cyanobacteria, have been shown to display phylogenetic endemism in Antarctica at the  
8 level of rRNA gene-defined diversity <sup>15-17</sup>. Also, lichenised fungi displayed patterns in  
9 diversity that suggest that they radiated from local refugia rather than from exogenous  
10 sources outside Antarctica <sup>18</sup>. A global theoretical model for atmospheric aerosols  
11 estimated that the rate of airborne microbial exchange to Antarctica may be extremely  
12 low, with 90% of aerosols expected to be of local origin <sup>19</sup>. In contrast, empirical studies  
13 have claimed circum-polar distribution for some soil Cyanobacteria, chlorophyte algae,  
14 and Fungi <sup>20,21</sup>. Antarctica therefore presents a paradox in microbial biogeography with  
15 regard to microbial dispersal. Evidence for airborne microorganisms in Antarctica is  
16 scarce and limited to a few preliminary biodiversity estimates indicating cosmopolitan  
17 and local taxa are recoverable from the near-surface atmosphere <sup>22-24</sup>.

18         Here we use state-of-the art sampling, sequencing and statistical approaches to  
19 study the diversity of airborne and soil microbial communities in a typical Antarctic Dry  
20 Valley. We targeted Bacteria and Fungi because these domains are the most abundant  
21 microorganisms in the McMurdo Dry Valleys <sup>13</sup>. We tested the null hypotheses that air  
22 and soil microbial communities are a random sample of the regional and global  
23 microbial pools. For doing so, we acquired massive bulk-air samples and estimated  
24 microbial diversity in near-ground air and underlying soil for low and high elevation

1 sites, as well as polar air above the boundary layer for surface interactions and non-  
2 polar sources.

3 The incoming air mass to our study site in the McMurdo Dry Valleys largely  
4 transited above the Antarctic Plateau during the maximum predicted residence time for  
5 bacteria and fungi in air (15 days) <sup>19</sup>, whilst the most distant air mass had a non-polar  
6 origin above the coastal shelf of New Zealand (Fig. 1a). Transport was exclusively from  
7 the Polar Plateau and across the Trans-Antarctic Mountains during the average  
8 residence time for microorganisms in air (Fig. 1a). We thus envisage that severe  
9 selection pressure should occur during airborne transit in an air mass with freezing  
10 temperatures and high UV exposure at mean altitudes of 2769m (3 day transit) and  
11 3034m (15 day transit).

12 In general terms, we found that alpha diversity metrics for air and elevated high-  
13 altitude soils were distinct from those observed in valley soils, with taxa richness being  
14 more variable in valley soils than in air and elevated soils (Fig. 1b,c). These results  
15 highlight the more heterogenous nature of valley soils as a habitat compared with bulk  
16 phase air or elevated altitude soils, where conditions are generally unfavourable to  
17 colonisation. The elevated altitude mineral soil sites may therefore be representative of  
18 near-term airborne deposition to this system. Bacterial taxa richness was similar among  
19 all samples although slightly higher in air and elevated soil (Fig. 1b). Conversely, the  
20 richness of fungal taxa was highest in valley soils, and lower in air and elevated soils  
21 (Fig. 1c). Soils also displayed greater evenness and lower richness values than air  
22 samples and this tentatively indicated that airborne fungi are under strong selective  
23 pressure. Ordination analyses of weighted UniFrac distances for bacteria and fungi  
24 supported these trends in alpha diversity (Fig. 1b,c). Valley soil bacterial communities

1 separated clearly from air and elevated soil communities (Fig. 1b). A similar although  
2 less pronounced pattern was observed for fungi (Fig. 1c).

3 Taxonomic assignment of bacteria and fungi revealed further complexity.  
4 Airborne phylum-level bacterial diversity was dominated by Proteobacteria,  
5 Bacteroidetes and Firmicutes (Fig. 1d, Online Supplementary Information Fig. S2). Phyla  
6 with high relative abundance comprised spore-formers and taxa with known UV and/or  
7 desiccation tolerant traits viewed as advantageous during atmospheric transport and  
8 survival in Antarctic soil. The airborne bacterial samples supported relatively high  
9 levels of taxa associated with marine influence <sup>8</sup> suggesting recruitment during transit  
10 over the Southern Ocean (Online Supplementary Information, Fig. S2). Terrestrial  
11 bacteria were also transported and may have benefited from islands acting as stepping  
12 stones for dispersal <sup>8</sup>. Soil communities supported greater abundance of Actinobacteria  
13 and other taxa typical from arid soils <sup>10</sup>. Near-ground air supported 3-5 fold more  
14 habitat-specific taxa than high-altitude air; samples from the former habitat were most  
15 similar to their underlying soil communities. Valley soils supported 56.4% soil-specific  
16 taxa compared with only 15.8% in elevated high-altitude soils. Valley soils shared very  
17 few taxa with the total air sample pool (4.5%) whilst different air habitats (valley,  
18 elevated and high-altitude) shared approximately half the taxa encountered in each  
19 habitat.

20 The most abundant fungal taxa in air were basidiomycetous yeasts (Fig. 1e,  
21 Online Supplementary Information Fig. S2) whereas soils were dominated by  
22 unclassified fungi also including yeasts. The yeasts are thought to be well-adapted to  
23 growth in Antarctic soil habitats <sup>25</sup>. Ascomycetes were also commonly encountered but  
24 Chytrids occurred only in valley soil and air. Valley soils supported 48.3% habitat-  
25 specific taxa while elevated soils and high-altitude air showed the lowest number of

1 habitat-specific taxa (4.9-8.7%). The different air habitats shared approximately half the  
2 taxa encountered in each habitat. Fungi are well-adapted to conditions anticipated  
3 during atmospheric transport due to the production of resistant spores and UV-  
4 protective compounds. Local recruitment may, however, be limited to asexual states  
5 since teleomorph fruiting structures are not known from Antarctic fungi.

6 We achieved near-asymptote in diversity estimation for all samples (Online  
7 Supplementary Information, Supplementary Methods). Thus we further interrogated  
8 the phylogenetic diversity of air and soil by generating distribution heatmaps by habitat  
9 for the 1,000 most abundant taxa. This analysis captured 91% of total bacterial and  
10 96% of total fungal diversity in the libraries (Fig. 2a,b). We also incorporated diversity  
11 data from air originating at the nearest non-polar land mass into this analysis. A striking  
12 pattern emerged where soil bacterial and fungal assemblages in the McMurdo Dry  
13 Valleys were only partially recruited from local air taxa. This pattern cannot be further  
14 explained by airborne recruitment from exogenously sourced aerosols. These findings  
15 support the notion of a system operating in stark contrast with the long-held  
16 assumption that microbial dispersal is ubiquitous and deterministic niche processes are  
17 the primary driver of community assembly in terrestrial surfaces <sup>1,12</sup>.

18 We therefore further interrogated this association using Ecological Network  
19 Analysis (Fig. 2c). Overall non-polar air displayed least connectivity to all other  
20 Antarctic habitats as previously observed by the weak associations and greater Bray-  
21 Curtis distances between them (Fig. 2c). Bacterial communities clustered by habitat  
22 type (Fig. 2c). This pattern is likely to be indicative of selection pressures due to local  
23 environmental filtering, which could combine a mixture of biotic and abiotic factors.  
24 Conversely, fungal communities associated by geographic distance and were thus more  
25 likely to be influenced by dispersal limitation. No significant distance-decay

1 relationships were observed for airborne or soil communities between valley and  
2 elevated sites (air bacteria  $R^2 = 0.009$ , soil bacteria  $R^2 = 0.006$ , air fungi  $R^2 = 0.016$ , soil  
3 fungi  $R^2 = 0.024$ ). These results indicate that dispersal within the Dry Valleys may be  
4 limited and likely reflects the associated steep environmental gradients present in this  
5 region. These findings provide empirical support to prevailing theoretical models of  
6 emission and transport for biological particles in the atmosphere that predicts  
7 relatively low exchange between Antarctic and non-polar air as well as reduced  
8 residence time in air for fungi compared to bacteria due to allometric considerations <sup>19</sup>.

9         We conducted additional analyses (Nestedness Analysis and Net Relatedness  
10 Index analysis) to reveal the extent to which taxonomic and phylogenetic structuring  
11 reflected the likelihood of exogenous recruitment (Fig. 3). Null models applied to  
12 nestedness metrics showed that bacterial and fungal communities overall were  
13 significantly anti-nested (NODF <30) (Fig. 3a). This general result implies that passive  
14 sampling from the regional pool alone was not sufficient to explain the structure of local  
15 Antarctic communities. Both Bacteria and Fungi, however, were significantly nested for  
16 taxonomic composition (NODFc) under the hypothesis that nestedness can be  
17 maximised when ordering sites from the most connected, to the least connected, to a  
18 global species pool. This suggests that species poor assemblages of the least connected  
19 sites are a proper subset of richer, more connected sites. Fungi were markedly more  
20 nested (NODFc = 62) than Bacteria (NODFc = 18) (Fig. 3a), suggesting a potential major  
21 role of dispersal limitation for this group.

22         The Net Relatedness Index (NRI) added phylogenetic support to the findings of  
23 our Network Analysis and Nestedness Analysis, by further demonstrating that local  
24 Antarctic communities were not a random sample of the overall species pool. Antarctic  
25 bacteria displayed greater and highly significant phylogenetic clustering compared to



1 non-polar samples, which were almost randomly structured and in some case over-  
2 dispersed (Fig. 3b). Although the pattern itself does not prove any specific process, the  
3 results clearly indicate that Antarctic bacterial communities both in soil and air must  
4 have been selected non-randomly, which is consistent with both the taxonomic and  
5 phylogenetic observations for our air and soil communities. This result is congruent  
6 with observations of biodiversity for other soils in the Dry Valleys region <sup>26</sup>. The Fungi  
7 were always significantly clustered but bacterial communities were always much more  
8 clustered than fungi at any given airborne or soil location. The fungal data should,  
9 however, be interpreted with care given current drawbacks with phylogenetic  
10 reconstruction based on ITS and despite our efforts to correct for them. Nonetheless  
11 even a cautious interpretation of the data suggests a limited extent of input from fungal  
12 taxa not present in local reservoirs. This interpretation concurs with our other lines of  
13 evidence presented here and with studies on fungal dispersal from other biomes <sup>27</sup>.

14         Contrary to the view that “everything is everywhere” in terms of airborne  
15 microbial transport, our data indicates that the aerosphere is a strongly selective  
16 habitat that limits dispersal, although the extent may vary between taxonomic groups  
17 and spatial scales. We conclude that inter-continental microbial connectivity to the  
18 McMurdo Dry Valleys of East Antarctica is limited, and this supports the hypothesis that  
19 the Hadley Cell circulation acts as a dispersal barrier to the poles even during the  
20 austral summer when the Polar Vortex is annually at its weakest. The Antarctic  
21 continent supports other smaller ice-free soil regions and whilst we are unable to  
22 directly extrapolate our data to these, it is reasonable to expect similar patterns given  
23 what is known of air circulation to the continent; that is, that other Antarctic ice-free  
24 areas may also be somewhat decoupled from global microbial reservoirs. An exception  
25 may be the peninsula in West Antarctica due to its proximity to the South American

1 continent. Comparison of our soil biodiversity estimates with those for other Dry  
2 Valleys locations suggest there is a common core diversity throughout the Antarctic Dry  
3 Valleys <sup>26</sup>. Hence, we expect our data to be broadly applicable to this region. The low  
4 level of airborne immigration from exogenous sources may represent an inherently low  
5 flux for Antarctica and this may help to explain the unique microbial composition of  
6 Antarctic soils compared to others globally <sup>28</sup>.

7         We have presented multiple lines of evidence to refute the null hypothesis that  
8 local air and soil microorganisms are a random sample of phylogenetic diversity in the  
9 regional/global pools. Sources of recruitment other than persistent airborne transport  
10 are therefore necessary to fully explain the extant Antarctic soil microbial diversity  
11 patterns observed. One potential explanation is stochastic storm events where  
12 particulate matter supporting biological propagules is thought to be transported on  
13 local scales within the Dry Valleys <sup>29</sup>, although we did not encounter any such events  
14 during our sampling expedition. A further source may be local dispersal from  
15 geothermal refugia as they are important reservoirs for radiative dispersal of animal,  
16 plant and lichen taxa <sup>5</sup>. The periodicity from which dispersal from such reservoirs  
17 occurs is, however, unknown. An additional reservoir may be the moisture-sufficient  
18 soil around lakes where microbial mats are known to persist over inter-annual periods  
19 <sup>26</sup>. Local refugia may be important in facilitating resilience at the landscape scale where  
20 severe local extinction pressure occurs due to stochasticity and steep environmental  
21 gradients for abiotic variables.

22         Challenges remain in deciphering the relationship of diversity patterns to  
23 biomass and ecosystem function <sup>30</sup>, but the revelation that airborne connectivity is  
24 largely localised rather than being an inter-continental scale process emphasises the  
25 conservation value of the McMurdo Dry Valleys as a unique ecosystem. This is

1 particularly pertinent in light of a predicted increase in stochasticity for atmospheric air  
2 circulation as a result of climate change, which may led to an increased flux of foreign  
3 and invasive taxa into Antarctic ecosystems. Such an increased flux acting in concert  
4 with warmer temperatures could profoundly alter the unique biota inhabiting the  
5 Antarctic Dry Valleys, one of the of the last pristine ecosystems on Earth.

## 6 7 **Methods**

### 8 **Sample Recovery**

9 We employed a high-volume liquid impinger apparatus (Coriolis  $\mu$ , Bertin  
10 Technologies, France) and developed a novel collection protocol optimised for low-  
11 temperature environments. This involved collection of samples directly into *RNAlater*  
12 nucleic acid preservative solution (Invitrogen, Carlsbad, CA). Evaporation was  
13 compensated for using a peristaltic pump set to between 0.5 and 1.3mL per minute with  
14 a mixture of phosphate-buffered saline (PBS) and 20% v/v *RNAlater*. For each sampling  
15 interval a random collection cone was also assembled into the machine but not  
16 activated, and these were used as the negative controls. All collection cones were  
17 soaked in 1.5% sodium hypochlorite (NaClO) then washed with 70% ethanol and three  
18 washes of Milli-Q H<sub>2</sub>O before being filled with filtered *RNAlater*. All sampling equipment  
19 was disassembled between locations and cleaned with NaClO, ethanol and Milli-Q H<sub>2</sub>O.  
20 Samples in *RNAlater* were stored at 4°C during transit from Antarctica and until  
21 processed.

22 Air mass at near-surface (1.5m above ground) and corresponding soil sampling  
23 (top 10mm surface soil after removing pebbles and rocks) was conducted from 11<sup>th</sup> –  
24 23<sup>rd</sup> January 2017 from eight locations throughout the Wright Valley floor (77.518633  
25 S, 161.768783 E) and high elevation locations at the valley ridge and a high elevation

1 inter-valley feature known as Bull Pass (7.47085 S, 161.77345 E). Air mass above the  
2 boundary layer for surface influence was also recovered by mounting the apparatus in a  
3 helicopter with an external sampling port (flightpath: 2,000m A.M.S.L., 77.440836 S,  
4 162.657553 E to 77.524583 S, 161.690917 E). Overall, we retrieved 33 massive bulk air  
5 samples plus underlying soil samples from 18 locations (high altitude Antarctic air  
6 samples and non-polar air did not have accompanying soil samples). Massive bulk-  
7 phase air volumes (72,000L per sample) were collected for each discreet sampling  
8 location. Biotic data were retrieved for 30 air samples with a total sampled volume of  
9 2,160,000L, and from all soil samples.

10 At each station and time interval air flow rates of 300L/min were employed for  
11 4hrs into 15ml RNA*later*. This approach overcomes limitations from earlier studies  
12 where low volume pumps have necessitated long sampling durations with uncertain  
13 microbial survival and recovery and impaction techniques that are known to bias  
14 against certain phyla <sup>31,32</sup>. Non-polar air samples were collected from New Zealand's  
15 North Island (36.916153 S, 174.645760 E) during the same austral summer season and  
16 using the same method. We selected this place because our HYSPLIT back trajectory  
17 analysis (see below) indicated that this was the nearest non-polar land mass from  
18 which air mass arriving at the Dry Valleys location was derived. These samples were  
19 used to make broad diversity comparisons with possible exogenous sources although  
20 we acknowledge that additional variability is likely among non-polar aerosols given  
21 inherent uncertainties over their trajectory to the Antarctic. Each location was sampled  
22 at three discreet time intervals during the austral summer field season.

23 Sampling was conducted within local weather parameters as follows: Relative  
24 humidity 20-58%, Temperature -3.3-6.9°C, wind speed 0.6-9.5m/S, wind direction E-  
25 ESE (Kestrel 3500 Weather Meter, Nielson-Kellerman Co, Minnesota, USA); Total near-

1 ground air particulate matter 2939-6558 $\mu\text{g}/\text{m}^3$  PM<sub>2.5-10</sub> (Aerotrak, TSI Incorporated,  
2 Minnesota, USA). Back-trajectories of air mass arriving at each sampling interval were  
3 generated using the National Oceanic and Atmospheric Administration (NOAA)  
4 HYSPLIT-WEB model (<https://ready.arl.noaa.gov/HYSPLIT.php>). Separate calculations  
5 were made at 3d and 15d because they represent the average minimum and maximum  
6 residence times for microorganisms in air transported to the McMurdo Dry Valleys <sup>19</sup>.  
7 HYSPLIT back trajectories were calculated using the GDAS database and the model  
8 vertical velocity option. Three day back trajectories travelled between 598 and 2581 km  
9 with an average of 1350 km at an average altitude of 2769 m and a maximum altitude of  
10 5174 m above mean sea level (AMSL) Fifteen day back trajectories travelled between  
11 4673 and 11216 km with an average of 6886 km at an average altitude of 3034 m and a  
12 maximum altitude of 8211 m AMSL.

### 13 **Environmental DNA sequencing and bioinformatics**

14 Airborne microbial samples from the RNA<sub>later</sub> preservation solution were  
15 filtered onto a 25mm 0.2 $\mu\text{m}$  polycarbonate filter and stored frozen until processed.  
16 Total DNA was directly extracted using a CTAB protocol <sup>33</sup>. DNA was extracted from  
17 three 0.75g  $\pm$ 0.025 of soil using the same CTAB protocol. DNA yield for these ultra-low  
18 biomass samples was quantified using the Qubit 2.0 Fluorometer (Invitrogen) in the  
19 range 1.06-8.44ng. Samples were then stored at -20°C until processed. We used DNA  
20 yield as an indirect estimate for biomass. We were unable to successfully apply direct  
21 cell/fluorescent particle counting due to the extremely low cell numbers in Antarctic  
22 air, although we acknowledge the validity of this approach in studies of higher biomass  
23 aerial habitats <sup>8</sup>. We did not apply real-time quantitative PCR as a further indirect  
24 estimate of biomass since the approach has significant limitations that preclude

1 meaningful estimates of biomass from environmental samples with domain-specific  
2 PCR primers <sup>34,35</sup>.

3         Illumina MiSeq libraries were prepared as per manufacturer's protocol  
4 (Metagenomic Sequencing Library Preparation Part # 15044223 Rev. B; Illumina, San  
5 Diego, CA, USA) and as previously described with PhiX positive controls <sup>36</sup>. We targeted  
6 Bacteria and Fungi since these domains are the most abundant microorganisms in the  
7 McMurdo Dry Valleys <sup>26</sup>. PCR was conducted with primer sets targeting the V3-V4  
8 regions of bacterial and archaeal 16S rRNA gene: PCR1 forward (5' TCGTCGGCAG  
9 CGTCAGATGT GTATAAGAGA CAGCCTACGG GNGGCWGCAG 3') and PCR1 reverse (5'  
10 GTCTCGTGGG CTCGGAGATG TGTATAAGAG ACAGGACTAC HVGGGTATCT AATCC 3') and  
11 the internal transcribed spacer region of fungal 18S and 5.8S rRNA genes: ITS1 forward  
12 (5' CTTGGTCATTTAGAGGAAGTAA 3') and ITS2 reverse (5' GCTGCGTTCTTCATCGATGC  
13 3'). These primers for Bacteria and Fungi are widely accepted to capture the broadest  
14 estimates of diversity <sup>10,37,38</sup> and were used according to recommended workflows for  
15 the Earth Microbiome Project (<http://www.earthmicrobiome.org>). Total sequence  
16 library sizes were 3,994,561 for bacteria and 2,437,256 for fungi before filtering and  
17 total counts in the processed dataset were 1,333,553 bacterial and 2,220,883 fungal  
18 sequences. A total of 3,636 bacterial and 5,525 fungal taxa were identified from these.  
19 All sequence data generated by this study has been submitted to the NCBI Sequence  
20 Read Archive under BioProject PRJEB27416 with accession numbers ERS3573837 to  
21 ERS3573946.

22         Sequencing data for 16S rRNA gene amplicons was processed based on the  
23 DADA2 v1.8 <sup>39</sup> pipeline. Primers sequences were removed using cutadapt <sup>40</sup> to remove  
24 forward (CCTACGGGNGGCWGCAG) and reverse (GACTACHVGGGTATCTAATCC). The  
25 reads were uniformly trimmed to 280 bp (forward) and 250 bp (reverse) and then

1 filtered by removing reads exceeding maximum expected error of 2 for forward reads  
2 and 5 for reverse reads or reads containing ambiguity N symbol. The reads were used to  
3 train the error model and then dereplicated to acquire unique sequences, which were  
4 used to infer sequence variants with the trained error model. The forward and reverse  
5 reads were merged and chimeric sequences were removed. For bacteria we used  
6 amplicon sequence variants (ASV) to assign operational taxonomic units (OTUs) since  
7 this has been shown as the most robust method currently available for bacterial 16S  
8 rRNA gene-defined taxa identification <sup>41</sup>. OTUs were given taxonomic assignment using  
9 DADA2 with SILVA nr v132 database <sup>42</sup> to provide species level assignment based on  
10 exact match between ASVs and known reference sequences. For fungal ITS1 amplicon  
11 data, the sequences were processed using USEARCH v9.0.2132 <sup>43</sup>. The forward and  
12 reverse paired-end sequences were merged and filtered to remove reads >1 maximum  
13 expected error per sequence. Additionally, anomalous sequences (<200 or >500 bp in  
14 length or exceeding 20 homopolymers) were also removed. After dereplication and  
15 removal of singletons, the reads were clustered at 97% identity threshold to obtain  
16 representative sequences as OTUs <sup>37</sup>. Unfiltered reads were mapped onto these OTUs to  
17 produce an abundance table of the occurrence of these OTUs within the communities.  
18 The representative sequences were given taxonomic assignment using USEARCH  
19 SINTAX classifier and RDP Warcup training set v2 (rdp\_its\_v2) <sup>44</sup>.

20 The resulting OTUs were then processed as previously described <sup>45</sup>. The R  
21 packages phyloseq <sup>46</sup>, DESeq2 <sup>47</sup> and ggplot2 <sup>48</sup> were used for downstream analysis and  
22 visualisation including ordination and alpha/beta diversity calculations. Despite  
23 inherent bias due to underlying differences in substrate biomass influencing species  
24 richness estimates with any cross-habitat biogeographic analysis <sup>49</sup>, we are confident  
25 that comparable yet inherently low biomass in all our air and ultra-oligotrophic mineral

1 soil samples minimised such influence and this was reflected in our diversity estimates.  
2 An exception was that fungi in non-polar air were markedly more taxon-rich than in  
3 Antarctic samples. Therefore we used guild analysis (FUNGuild<sup>50</sup>) to establish the  
4 predominantly phyllosphere origin of fungi in non-polar samples which are absent in  
5 Antarctica as well as the latitudinal gradient in fungal diversity<sup>51</sup> (lack of database  
6 depth for FUNGuild limited its value to identifying ecological guilds for non-polar fungi  
7 only). For heatmap visualisations the 1,000 most abundant OTUs in each data set were  
8 selected, and these captured 91% bacterial and 96% fungal sequences. All other  
9 analysis used the entire sequence library data. Our heatmap analysis therefore had high  
10 confidence since the unsampled 'tail' comprised only extremely rare sequence variants  
11 at very low/singleton abundance.

## 12 **Statistical Treatments**

13 We used multiple statistical approaches to test the null hypothesis that local soil  
14 and air sample communities were a random sample of the of the regional pool.  
15 Rejection of the hypothesis (i.e., non-random patterns) yields observational evidence  
16 for the alternative hypothesis that local communities are a non-random selection from  
17 the regional pool. Specifically, we expected local soil communities to be selected against  
18 the extreme conditions found at soils in the McMurdo Dry Valleys soil and thus to  
19 display clustering. At the same time, we expected air samples to be less structured or  
20 even random and thus better reflect the regional pool, although some structuring due to  
21 local influences from atmospheric stressors such as low temperatures and UV exposure  
22 were also be expected.

23 We employed approaches that utilised both taxonomic identity and phylogenetic  
24 structure of the communities. Ecological Network Analysis is a commonly employed  
25 tool to infer biotic interactions within and between communities by visualising links



1 between species nodes. Potential relationships pertinent to our system include  
2 connectivity, clustering and nestedness which are informative to interpreting the  
3 biogeographic patterns of species occurrence <sup>52</sup>. We performed Ecological Network  
4 Analysis on our samples using the R package phyloseq <sup>46</sup> with maximum Bray-Curtis  
5 distance of 0.2 for Bacteria and 0.4 for Fungi to establish connection between nodes  
6 (representing communities). The nodes were positioned using the Fruchterman-  
7 Reingold method <sup>53</sup>.

8         Nestedness is a widespread biogeographical pattern that emerges when species  
9 composition of small assemblages is a nested subset of larger regional assemblages. We  
10 quantified this pattern using the metrics NODF <sup>54</sup> and its compositional (NODFc) and  
11 incidence (NODFr) version. This metric is currently considered one of the most effective  
12 and statistically robust, especially in relation to the null models that are used to test  
13 whether observed metrics are smaller or larger than expected by chance <sup>52,55</sup>. The  
14 metric ranges from 0, that is no nestedness or perfect antinestedness, to 100 (perfect  
15 nestedness). In a perfectly nested assemblage, species poor communities are just a  
16 subset of species richer communities, or less frequent species always occur in subsets of  
17 sites where most widespread species also occur, or a combination of both. We applied  
18 the NODF metrics to the matrix of the investigated sites both for Bacteria and Fungi. We  
19 simulated null models using a random swap algorithm (R project, vegan package,  
20 function “oecosimu” <sup>56</sup>) with fixed row and column sums. This combination to create a  
21 random matrix is the most conservative in terms of Type I and II errors and is  
22 particularly recommended when species co-occurrence is critical to the tested  
23 hypothesis <sup>52</sup>. We calculated Standardised Effect Sizes and tested for significance of  
24 effects with 999 permutations and at  $P < 0.05$ . We calculated NODF and null models by a  
25 specific order of sites in the matrix that reflected our main null hypothesis. This was

1 that local communities are just a passive sampling from regional species pools.  
2 Specifically, we used two complementary orderings of the sites to test our hypothesis.  
3 First, we created a connectivity gradient, which assumed New Zealand is the site more  
4 connected to the global species pool while soils in the Antarctic Dry Valleys are the least  
5 connected (the exact order was: non-polar>high-altitude air>elevated air>valley  
6 air>elevated soil>valley soil). In the second test of the hypothesis local, more isolated  
7 soils were assumed to select more than connected soil and air habitats. In this case, sites  
8 were ordered as follows: Valley Soil>Elevated Soil>Valley Air>Elevated Air>high-  
9 altitude air>non-polar air.

10 Community phylogenetic metrics were calculated using the R <sup>57</sup> packages picante  
11 <sup>58</sup>, ape <sup>59</sup>, phylobase <sup>60</sup>, adephylo <sup>61</sup>, and phytools <sup>62</sup>. Phylogenetic trees for community  
12 phylogenetic structure analysis were constructed for all OTUs with FastTree v2.1.9 on  
13 multiple alignment of sequences produced by MUSCLE v3.8.31. For Bacteria an  
14 approximately Maximum-Likelihood approach was used whilst for Fungi an alignment-  
15 free distance approach with Neighbour-Joining method was employed in order generate  
16 a ITS-based phylogenetic tree for community metrics <sup>63</sup>. The distance approach was  
17 used for Fungi as the hypervariability of ITS1 loci hinders multiple sequence alignment  
18 required for most phylogenetic analyses. To validate this approach, we compared our  
19 tree topology with the most recent whole genome phylogenies for the Fungi <sup>64,65</sup>. This  
20 approach was robust at higher taxonomic levels (there is no consensus for fungal  
21 phylogenies using multiple loci or whole genomes below Order rank) and has been used  
22 successfully with other eukaryotic taxa as a workflow for Net Relatedness analysis <sup>66</sup>.  
23 Although we acknowledge limitations to this approach, the advantages of using ITS loci  
24 for taxonomic identification vastly outweighed its shortcomings, and we also  
25 triangulated data from this test with additional analytical approaches, so overall our

1 inclusion of this test is justified and interpreted conservatively. Mean phylogenetic  
2 distance (MPD) <sup>67</sup> was calculated to measure phylogenetic distance between ASVs and  
3 OTUs in each sample. A null model algorithm based on independent swap (999  
4 randomisation) was used to test the extent of phylogenetically clustering (positive  
5 values) or overdispersion (negative values) <sup>68</sup>. Results for NRI were expressed as effects  
6 size,  $(MPD - MPD_{null}) / SD(MPD_{null})$ . Distance Decay of phylogenetic versus geographic  
7 distance for bacteria and fungi was estimated using the R package Vegan <sup>56</sup>.

8

## 9 **Acknowledgements**

10 Field and logistical support was provided by Antarctica New Zealand and the United  
11 States Antarctic Program. The authors thank Craig Cary (University of Waikato) for  
12 facilitating access to the McMurdo Dry Valleys. The research was funded by a grant from  
13 the New Zealand Ministry of Business, Innovation & Employment (UOWX1401) and the  
14 Yale-NUS College Start-Up Fund. F.T.M. is supported by the European Research Council  
15 (BIODESERT project, ERC Grant agreement n° 647038).

16

## 17 **Author contributions**

18 S.D.J.A. and S.B.P. conceived the study; S.D.J.A and C.K.L. conducted fieldwork; T.M.  
19 developed and validated the helicopter sampling method; S.D.J.A. performed laboratory  
20 experiments; S.D.J.A., K.C.L., T.C., and S.B.P. performed data analysis and interpretation;  
21 D.A.C., F.T.M. and S.B.P. critically assessed and interpreted the findings; S.B.P. wrote the  
22 manuscript.

23

## 1 **Materials & Correspondence**

2 Correspondence and requests for materials should be addressed to S.B.P. Email:

3 [stephen.pointing@yale-nus.edu.sg](mailto:stephen.pointing@yale-nus.edu.sg).

4

## 5 **Data Availability**

6 All sequence data generated by this study has been submitted to the EMBL European

7 Nucleotide Archive (ENL) under BioProject PRJEB27416 with accession numbers

8 ERS3573837 to ERS3573946. PLEASE CHECK PUBLIC AVAILABILITY

9

## 10 **References**

- 11 1. Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. H. Beyond  
12 biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev.*  
13 *Microbiol.* **10**, 497–506 (2012).
- 14 2. Burrows, S. M., Elbert, W., Lawrence, M. G. & Pöschl, U. Bacteria in the global  
15 atmosphere – Part 1: Review and synthesis of literature data for different  
16 ecosystems. *Atmos. Chem. Phys.* **9**, 9263–9280 (2009).
- 17 3. Kellogg, C. A. & Griffin, D. W. Aerobiology and the global transport of desert dust.  
18 *Trends Ecol. Evol.* **21**, 638–644 (2006).
- 19 4. Wilson, J. R. U., Dormontt, E. E., Prentis, P. J., Lowe, A. J. & Richardson, D. M.  
20 Something in the way you move: dispersal pathways affect invasion success.  
21 *Trends Ecol. Evol.* **24**, 136–44 (2009).
- 22 5. Chown, S. L. *et al.* The changing form of Antarctic biodiversity. *Nature* **522**, 431  
23 (2015).
- 24 6. de Wit, R. & Bouvier, T. ‘Everything is everywhere, but, the environment selects’;

- 1        what did Baas Becking and Beijerinck really say? *Environ. Microbiol.* **8**, 755–758  
2        (2006).
- 3    7.    Finlay, B. J. & Clarke, K. J. Ubiquitous dispersal of microbial species. *Nature* **400**,  
4        828–828 (1999).
- 5    8.    Mayol, E. *et al.* Long-range transport of airborne microbes over the global tropical  
6        and subtropical ocean. *Nat. Commun.* **8**, 201 (2017).
- 7    9.    Favet, J. *et al.* Microbial hitchhikers on intercontinental dust: catching a lift in  
8        Chad. *ISME J.* **7**, 850–67 (2013).
- 9    10.   Delgado-Baquerizo, M. *et al.* A global atlas of the dominant bacteria found in soil.  
10       *Science (80-. )*. **359**, 320–325 (2018).
- 11   11.   Wang, J. *et al.* Phylogenetic beta diversity in bacterial assemblages across  
12        ecosystems: deterministic versus stochastic processes. *ISME J.* **7**, 1310–1321  
13        (2013).
- 14   12.   Lowe, W. H. & McPeck, M. A. Is dispersal neutral? *Trends Ecol. Evol.* **29**, 444–50  
15        (2014).
- 16   13.   Pointing, S. B. *et al.* Highly specialized microbial diversity in hyper-arid polar  
17        desert. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19964–19969 (2009).
- 18   14.   Chan, Y., Van Nostrand, J. D., Zhou, J., Pointing, S. B. & Farrell, R. L. Functional  
19        ecology of an Antarctic Dry Valley. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8990–5  
20        (2013).
- 21   15.   Bahl, J. *et al.* Ancient origins determine global biogeography of hot and cold desert  
22        cyanobacteria. *Nat. Commun.* **2**, 163 (2011).
- 23   16.   Jungblut, A., Lovejoy, C. & Vincent, W. Global distribution of cyanobacterial  
24        ecotypes in the cold biosphere. *ISME J.* **4**, 191–202 (2010).
- 25   17.   Vyverman, W. *et al.* Evidence for widespread endemism among Antarctic micro-

- 1 organisms. *Polar Sci.* **4**, 103–113 (2010).
- 2 18. Fraser, C. I., Terauds, A., Smellie, J., Convey, P. & Chown, S. L. Geothermal activity  
3 helps life survive glacial cycles. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5634–9 (2014).
- 4 19. Burrows, S. M. *et al.* Bacteria in the global atmosphere – Part 2: Modeling of  
5 emissions and transport between different ecosystems. *Atmos. Chem. Phys.* **9**,  
6 9281–9297 (2009).
- 7 20. Kleinteich, J. *et al.* Pole-to-Pole Connections: Similarities between Arctic and  
8 Antarctic Microbiomes and Their Vulnerability to Environmental Change. *Front.*  
9 *Ecol. Evol.* **5**, 137 (2017).
- 10 21. Cox, F., Newsham, K. K., Bol, R., Dungait, J. A. J. & Robinson, C. H. Not poles apart:  
11 Antarctic soil fungal communities show similarities to those of the distant Arctic.  
12 *Ecol. Lett.* **19**, 528–536 (2016).
- 13 22. Kobayashi, F. *et al.* Atmospheric bioaerosols originating from Adélie penguins  
14 (*Pygoscelis adeliae*): Ecological observations of airborne bacteria at Hukuro Cove,  
15 Langhovde, Antarctica. *Polar Sci.* **10**, 71–78 (2016).
- 16 23. Pearce, D. A., Hughes, K. A., Lachlan-Cope, T., Harangozo, S. A. & Jones, A. E.  
17 Biodiversity of air-borne microorganisms at Halley station, Antarctica.  
18 *Extremophiles* **14**, 145–159 (2010).
- 19 24. Bottos, E. M. E. M., Woo, A. C. A. C., Zawar-Reza, P., Pointing, S. B. S. B. & Cary, S. C.  
20 S. C. Airborne Bacterial Populations Above Desert Soils of the McMurdo Dry  
21 Valleys, Antarctica. *Microb. Ecol.* **67**, 120–128 (2013).
- 22 25. Rao, S. *et al.* Low-diversity fungal assemblage in an Antarctic Dry Valleys soil.  
23 *Polar Biol.* **35**, 567–574 (2011).
- 24 26. Cowan, D. A., Makhalanyane, T. P., Dennis, P. G. & Hopkins, D. W. Microbial ecology  
25 and biogeochemistry of continental Antarctic soils. *Front. Microbiol.* **5**, 154

- 1 (2014).
- 2 27. Brown, S. P. & Jumpponen, A. Phylogenetic diversity analyses reveal disparity  
3 between fungal and bacterial communities during microbial primary succession.  
4 *Soil Biol. Biochem.* **89**, 52–60 (2015).
- 5 28. Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities  
6 and their functional attributes. *Proc. Natl. Acad. Sci. USA* **109**, 21390–21395  
7 (2012).
- 8 29. Atkins, C. B. & Dunbar, G. B. Aeolian sediment flux from sea ice into Southern  
9 McMurdo Sound, Antarctica. *Glob. Planet. Change* **69**, 133–141 (2009).
- 10 30. Pointing, S. B., Fierer, N., Smith, G. J. D., Steinberg, P. D. & Wiedmann, M.  
11 Quantifying human impact on Earth’s microbiome. *Nat. Microbiol.* **1**, 16145  
12 (2016).
- 13 31. Dybwad, M., Skogan, G. & Blatny, J. M. Comparative testing and evaluation of nine  
14 different air samplers: End-to-end sampling efficiencies as specific performance  
15 measurements for bioaerosol applications. *Aerosol Sci. Technol.* **48**, 282–295  
16 (2014).
- 17 32. Šantl-Temkiv, T. *et al.* High-Flow-Rate Impinger for the Study of Concentration,  
18 Viability, Metabolic Activity, and Ice-Nucleation Activity of Airborne Bacteria.  
19 *Environ. Sci. Technol.* **51**, 11224–11234 (2017).
- 20 33. Archer, S. D. J., McDonald, I. R., Herbold, C. W., Lee, C. K. & Cary, C. S. Benthic  
21 microbial communities of coastal terrestrial and ice shelf Antarctic meltwater  
22 ponds. *Frontiers in Microbiology* **6**, 485 (2015).
- 23 34. Fierer, N., Jackson, J. A., Vilgalys, R. & Jackson, R. B. Assessment of Soil Microbial  
24 Community Structure by Use of Taxon-Specific Quantitative PCR Assays. *Appl.*  
25 *Environ. Microbiol.* **71**, 4117–4120 (2005).

- 1 35. Props, R. *et al.* Absolute quantification of microbial taxon abundances. *ISME J.* **11**,  
2 584–587 (2017).
- 3 36. Warren-Rhodes, K. *et al.* Subsurface microbial habitats in an extreme desert Mars-  
4 analogue environment. *bioRxiv* 269605 (2018). doi:10.1101/269605
- 5 37. Schoch, C. L. *et al.* Nuclear ribosomal internal transcribed spacer (ITS) region as a  
6 universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6241–6  
7 (2012).
- 8 38. Thompson, L. R. *et al.* A communal catalogue reveals Earth’s multiscale microbial  
9 diversity. *Nature* **551**, 457–463 (2017).
- 10 39. Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W. & A, A. J. DADA2: High  
11 resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–  
12 583 (2016).
- 13 40. Martin, M. Cutadapt removes adapter sequences from high-throughput  
14 sequencing reads. *EMBnet.journal* **17**, 10 (2011).
- 15 41. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should  
16 replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**,  
17 2639–2643 (2017).
- 18 42. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data  
19 processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
- 20 43. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST.  
21 *Bioinformatics* **26**, 2460–2461 (2010).
- 22 44. Deshpande, V. *et al.* Fungal identification using a Bayesian classifier and the  
23 Warcup training set of internal transcribed spacer sequences. *Mycologia* **108**, 1–5  
24 (2016).
- 25 45. Maki, T. *et al.* Variations in the structure of airborne bacterial communities in



- 1 Tsogt-Ovoo of Gobi desert area during dust events. *Air Qual. Atmos. Heal.* **10**, 249–  
2 260 (2017).
- 3 46. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive  
4 Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).
- 5 47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and  
6 dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- 7 48. Wickham, H. *ggplot2*. (Springer New York, 2009). doi:10.1007/978-0-387-98141-  
8 3
- 9 49. Pointing, S. B. *et al.* Biogeography of photoautotrophs in the high polar biome.  
10 *Front. Plant Sci. Funct. Plant Ecol.* **6**, 692 (2015).
- 11 50. Nguyen, N. H. *et al.* FUNGuild: An open annotation tool for parsing fungal  
12 community datasets by ecological guild. *Fungal Ecol.* **20**, 241–248 (2016).
- 13 51. Tedersoo, L. *et al.* Global diversity and geography of soil fungi. *Science (80-. )*. **346**,  
14 (2014).
- 15 52. Ulrich, W., Almeida-Neto, M. & Gotelli, N. J. A consumer's guide to nestedness  
16 analysis. *Oikos* **118**, 3–17 (2009).
- 17 53. Fruchterman, T. M. J. & Reingold, E. M. *Graph Drawing by Force-directed*  
18 *Placement.* **21**, (1991).
- 19 54. Almeida-Neto, M., Guimarães, P., Guimarães, P. R., Loyola, R. D. & Ulrich, W. A  
20 consistent metric for nestedness analysis in ecological systems: reconciling  
21 concept and measurement. *Oikos* **117**, 1227–1239 (2008).
- 22 55. Ulrich, W. *et al.* A comprehensive framework for the study of species co-  
23 occurrences, nestedness and turnover. *Oikos* **126**, 1607–1616 (2017).
- 24 56. Oksanen, J. *et al.* R package: Vegan. (2017).
- 25 57. Swenson, N. G. *Functional and Phylogenetic Ecology in R*. (Springer New York,

- 1           2014). doi:10.1007/978-1-4614-9542-0
- 2   58.   Kembel, S. W. *et al.* Picante: R tools for integrating phylogenies and ecology.
- 3           *Bioinformatics* **26**, 1463–1464 (2010).
- 4   59.   Paradis, E. Ape package. (2018). Available at:
- 5           <https://www.rdocumentation.org/packages/ape/versions/5.2>. (Accessed: 24th
- 6           October 2018)
- 7   60.   Michonneau, F. Phylobase package. Available at:
- 8           <https://www.rdocumentation.org/packages/phylobase/versions/0.8.4>.
- 9           (Accessed: 24th October 2018)
- 10   61.   Dray, S. Adephylo package. Available at:
- 11           <https://www.rdocumentation.org/packages/adephylo/versions/1.1-11>.
- 12           (Accessed: 24th October 2018)
- 13   62.   Zhang, J. Phylotools package. Available at:
- 14           <https://www.rdocumentation.org/packages/phylotools/versions/0.2.2>.
- 15           (Accessed: 24th October 2018)
- 16   63.   Höhl, M., Rigoutsos, I. & Ragan, M. A. Pattern-based phylogenetic distance
- 17           estimation and tree reconstruction. *Evol. Bioinform. Online* **2**, 359–75 (2007).
- 18   64.   Choi, J. & Kim, S.-H. A genome Tree of Life for the Fungi kingdom. *Proc. Natl. Acad.*
- 19           *Sci. U. S. A.* **114**, 9391–9396 (2017).
- 20   65.   Ebersberger, I. *et al.* A Consistent Phylogenetic Backbone for the Fungi. *Mol. Biol.*
- 21           *Evol.* **29**, 1319–1334 (2012).
- 22   66.   Boyle, E. E. & Adamowicz, S. J. Community phylogenetics: Assessing tree
- 23           reconstruction methods and the utility of DNA barcodes. *PLoS One* **10**, (2015).
- 24   67.   Kembel, S. W. & Hubbell, S. P. The phylogenetic structure of a neotropical forest
- 25           tree community. *Ecology* **87**, 86–99 (2006).

1 68. Webb, C. O., Ackerly, D. D. & Kembel, S. W. Phylocom: software for the analysis of  
2 phylogenetic community structure and trait evolution. *Bioinformatics* **24**, 2098–  
3 2100 (2008).

4 69. Horner-Devine, M. C. & Bohannan, B. J. M. Phlogenetic lcustering and  
5 overdispersion in bacterial communities. *Ecology* **87**, S100–S108

6

7

1 **Display Items (2 Figures, 1 table)**

2 **Fig. 1. REVISE FORMAT OF LEGEND AND IMAGE Antarctic air and soil habitats**

3 **support distinct bacterial and fungal communities. a)** Route predictions for average  
4 minimum and maximum modelled residence time for microorganisms in air based on  
5 HYSPLIT back trajectory analyses. Back trajectories indicate distance travelled for  
6 sampled air mass at 3 d (598 - 2581 km distance, average altitude of 2769 m, maximum  
7 altitude of 5174 m A.M.S.L.) and 15 d (4673 - 11216 km distance average altitude  
8 3034m, maximum altitude 6886 km). **b) and c)** Alpha diversity estimates (Chao1  
9 richness and Pielou's relative evenness) and visualisation of community dissimilarity  
10 using Principal Co-ordinate Analysis of weighted UniFrac distance by habitat for **b)**  
11 **Bacteria and c) Fungi.** Boxplot whiskers represent 1.5 times the interquartile range  
12 from the first to the third quartiles or the maximum/minimum data point within the  
13 range. **d) and e)** Distribution and relative abundance of **d) Bacteria and e) Fungi** in  
14 Antarctic air and soil. Each stack bar represents data from three pooled replicates for  
15 each substrate location. Diversity is shown at phylum level as this is the highest  
16 taxonomic rank at which between-substrate differences are noticeable. Venn diagrams  
17 show amplicon sequence variants (ASVs) and operational taxonomic units (OTUs) count  
18 and percentage occurrence within and between each habitat. The high altitude samples,  
19 i.e., those without underlying soil, do not have corresponding soil samples. Sampling  
20 locations: valley (soil and 1.5m above ground), elevated (soil and 1.5m above ground at  
21 higher altitude locations at Bull Pass and valley ridges), 2000m (helicopter samples).  
22 Interactive graphics identifying taxonomic composition to lower taxonomic ranks  
23 within each sample are presented in the Online Supplementary Information (Fig. S2).  
24 Comparison with non-polar samples is given in the Online Supplementary Information  
25 Fig. S2.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

**Fig. 2. REVISE FORMAT OF LEGEND AND IMAGE Comparison of bacterial and fungal diversity from Antarctic and non-polar sources. a)** Distribution and relative abundance for the 1,000 most abundant bacterial amplicon sequence variants (ASVs) and fungal operational taxonomic units (OTUs). **b)** Rarefaction curves are shown for Bacteria and Fungi for each Antarctic habitat to illustrate sampling depth to near-asymptote. **c)** Co-occurrence associations derived from Ecological Network Analysis. We enforced maximum Bray-Curtis distance of 0.2 for Bacteria and 0.4 for Fungi to establish connection between nodes (representing communities). The nodes were positioned using the Fruchterman-Reingold method. Sampling locations: valley (soil and 1.5m above ground), elevated (soil and 1.5m above ground at higher altitude locations at Bull Pass and valley ridges), 2000m (helicopter samples) and New Zealand (non-polar). Full taxonomic comparison for all polar and non-polar samples is given in the Online Supplementary Information Fig. S2.

**Fig. 3. CAN WE CONVERT THE NRI TO A TABLE TOO? PLEASE CAN YOU PROVIDE THE DAT IN TALE FORM AND CHECK LEGENDS BELOW**

**Phylogenetic structuring of local and global pools for bacterial and fungal diversity. a)** Nestedness estimates made using the NODF model (where 0 = no nestedness, 100 = perfect nestedness). Fungi were more nested (NODFc = 62) than Bacteria (NODFc = 18). Bacteria and Fungi were significantly nested for taxa composition under the hypothesis that nestedness can be maximised by ordering sites from the most connected to the least connected. Least connected sites are demonstrated as a proper subset of richer, more connected sites. **b)** Net Relatedness Index analysis of phylogenetic structure within each sample type for Bacteria (b) and Fungi (f). Error

1 bars show the standard error of the mean for all samples in a given substrate type.  
2 Values for highly dispersed non-polar bacterial communities associated with forest soil  
3 are given for comparison (bn) and indicated by an asterisk <sup>69</sup>. Sampling locations: valley  
4 (soil and 1.5m above ground), elevated (soil and 1.5m above ground at higher altitude  
5 locations at Bull Pass and valley ridges), 2000m (helicopter samples) and New Zealand  
6 (non-polar).

7