



Spatiotemporal patterns of microbial composition and diversity in precipitation

Ken A. Aho, C.F. Weber, Brent C. Christner, Boris A. Vinatzer, Cindy E. Morris, R. Joyce, K. C. Failor, J.T. Werth, A.L.H. Bayless-edwards, David G. Schmale

► To cite this version:

Ken A. Aho, C.F. Weber, Brent C. Christner, Boris A. Vinatzer, Cindy E. Morris, et al.. Spatiotemporal patterns of microbial composition and diversity in precipitation. *Ecological monographs*, 2020, 90 (1), pp.e01394. 10.1002/ecm.1394 . hal-02626506

HAL Id: hal-02626506

<https://hal.inrae.fr/hal-02626506>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article type : Article

Running head: Microbe metacommunities in precipitation

Spatiotemporal patterns of microbial composition and diversity in precipitation

Aho, K.*¹, Weber, C. F.¹, Christner, B. C.², Vinatzer, B. A.³,
Morris, C. E.⁴, Joyce, R.², Failor, K.³, Werth, J. T.¹, Bayless-Edwards, A. L. H.¹,
Schmale III, D. G.³

*Corresponding author (ahoken@isu.edu)

¹ Biological Sciences, Idaho State University, Pocatello, ID 83209-8007.

² Department of Microbiology and Cell Science, Biodiversity Institute, University of Florida,
Gainesville, FL 32611.

³ School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA 24061-0331.

⁴ INRA, Plant Pathology Research Unit 407, Montfavet, France.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/ECM.1394](#)

This article is protected by copyright. All rights reserved

Aho, K. (Auteur de correspondance), Weber, Christner, B., Vinatzer, B., Morris, C. E., Joyce, Failor, K., Werth, BaylessEdwards, Schmale, D. (2020). Spatiotemporal patterns of microbial composition and diversity in precipitation. Ecological Monographs, 90 (1), e01394. , DOI : 10.1002/ecm.1394

Abstract:

Microbes in the atmosphere have broad ecological impacts, including the potential to trigger precipitation through species and strains that act as ice nucleation particles. To characterize spatiotemporal trends of microbial assemblages in precipitation we sequenced 16S (bacterial) and 18S (fungal) rRNA gene amplicon libraries collected from 72 precipitation events in three US states (Idaho, Louisiana, and Virginia) over four seasons. We considered these data from the perspective of a novel metacommunity framework. In agreement with our heuristic, we found evidence for distinct mechanisms underlying the composition and diversity of bacterial and fungal assemblages in precipitation. Specifically, we determined that: 1) bacterial operational taxonomic unit (OTU) composition of precipitation was strongly associated with macroscale drivers including season and high altitude characteristics of storms; 2) fungal OTU composition was strongly correlated with mesoscale drivers including particular spatial locations; 3) β -diversity (heterogeneity of taxa among samples) for both bacteria and fungi was largely maintained by turnover of taxa; however, 4) bacterial assemblages had higher contributions to total β -diversity from nestedness (i.e., lower richness assemblages were largely taxonomic subsets of richer assemblages), due to losses of taxa during dispersal, particularly among potential ice nucleation active bacteria; and 5) fungal assemblages had higher contributions to total β -diversity from turnover due to OTU replacement. Spatiotemporal trends in precipitation-borne metacommunities allowed delineation of a large number of statistically significant indicator taxa for particular sites and seasons, including trends for bacteria that are potentially ice nucleation active. Our findings advance understanding regarding the dispersion of aerosolized microbes via wet deposition, and the development of theory concerning potential assembly rules for bioaerosol assemblages.

Key words: microbial ecology; community ecology; bioprecipitation; precipitation; atmosphere; metabarcoding; metacommunity; biogeography; meteorology

Introduction

Of growing interest to ecologists are the properties of precipitation-borne microbes, particularly

species and strains that may trigger precipitation through the formation of ice in clouds (i.e., bioprecipitation, Sands et al. 1982, Morris et al. 2014a, Hummel et al. 2018). At temperatures warmer than -38 °C, phase transition from liquid water to ice requires the presence of ice nucleating particles (INPs) (Atkinson et al. 2016, Kanji et al. 2017, Knopf et al. 2018). Various particles can act as ice nucleators (Szyrmer and Zawadzki 1997, Després et al. 2012, Harrison et al. 2016); however, those most efficient in catalyzing ice formation at warmer temperatures (> -15 °C) are microbial or biogenic in origin (Murray et al. 2012, Morris et al. 2013, Fröhlich-Nowoisky 2015, Weber 2015, Fröhlich-Nowoisky 2016, Hill et al. 2017). Thus, knowledge concerning microbes in precipitation is potentially relevant to understanding global precipitation itself.

Microbes are aerosolized from all terrestrial and aquatic ecosystems (Wilson et al. 2015, Joung et al. 2017), including the surfaces of multicellular organisms, e.g., plants (Lindemann et al. 1982; Burrows et al. 2009a, Després et al. 2012), and fluxes of bioaerosols are often generated or affected by anthropogenic activities (Taha et al. 2006, Suski et al. 2018). Once aerosolized, some microbes are well suited for survival in atmospheric environments. For instance, dormant forms of spore-forming bacteria are commonly isolated from aerosols (Brunet et al. 2017). Still other microbial taxa are metabolically active in fog (Fuzzi et al. 1997) and super-cooled cloud droplets (Sattler et al. 2001), apparently using organic acids and alcohols as growth substrates (Amato et al. 2007, Vařilingom et al. 2010). Wet deposition (washout) serves as the primary mechanism for removing bioaerosols from higher atmospheric layers (Brunet et al. 2017, Hanlon et al., 2017). Because aerosolized microbial INPs may trigger ice formation and be-scavenged-by precipitation, they are predicted to have much shorter atmospheric residence times than other microbes (Burrows et al. 2009b, Stopelli et al. 2015). Thus, the ice nucleation-active (INA) phenotype allows broad dissemination of aerosolized microbes (Sands et al. 1982, Fitt et al. 1989) while limiting prolonged exposure to damaging stresses (e.g., desiccation, UV radiation) associated with atmospheric transport (Attard et al. 2012, Hill et al. 2017). Despite this information, however, the community ecology of these processes is poorly understood. For instance, it is largely unknown if atmospheric spatiotemporal characteristics, including site locations and season, correspond to general trends in precipitation-borne microbial taxa or taxa assemblages (Amato et al. 2017, pg. 8, but see Cáliz et al. 2018).

Altitudinal Constraints to Microbial Atmospheric Dispersal

The earth's planetary boundary layer (PBL; 0.01 - 2.5 km AGL) constitutes the lowest level of the troposphere. The PBL is directly affected by characteristics of the planet's terrestrial and marine surfaces, and varies continuously in thickness with surface roughness, and latitudinally, diurnally, and seasonally -- particularly over land -- due to radiative forcing (von Engelmann and Teixeira 2013).

Within the PBL, convective air patterns generate turbulent atmospheric mixing that is largely independent of trends at higher elevations. In the overlaying free troposphere, atmospheric mixing dramatically decreases and microbes may be carried over large horizontal distances via continental and global circulatory patterns (Prospero et al. 2005, Wilkinson et al. 2012). Although thermal inversion at the tropopause (6–16 km AGL) forms a natural upper elevational barrier to most microbes, high abundances of aerosolized organisms, including viable bacteria, have been observed in the lower stratosphere (Griffin 2004; Després et al. 2012; Bryan et al. 2019).

In general, the concentration of atmospheric microbes decreases with altitude (Bryan et al. 2019). Further, because of frictional and gravitational constraints, bioaerosols in the upper troposphere will generally be in the size range of bacteria (0.5 – 1.5 μm dia.; Mandrioli et al. 1984), with long residence times and the potential for global-scale dispersal (Brunet et al. 2017). Larger biological particles, including hyphal fragments and fungal spores (2–40 μm dia.), will generally be limited to the PBL, or even the surface boundary layer (SBL; 10-100 m AGL) which at its bottom has the properties of a plane-mixing layer and is subject to high turbulence and small-scale intermittent eddies (Raupach et al. 1996). Thus, larger microbes will generally have smaller dispersal distances (< 1 km) and residence times (Brunet et al. 2017, Wilkinson et al. 2012). As an exception, certain meteorological conditions may allow large dispersal distances (> 100 km) of microbes in the PBL (Aylor et al. 2006, Schmale and Ross 2015). Moreover, the strongest phenotypic correlate of aerosolized particle dispersal is aerodynamic diameter, for which physical size is an imprecise proxy (Reponen et al. 2001). Small aerodynamic diameters can contribute to the relatively long dispersal distances of some physically large spores, including the urediospores of rust fungi (Morris et al. 2014b).

A Metacommunity Framework for Microbes in Precipitation

A metacommunity defines a set of communities linked by the dispersal of potentially interacting species (Leibold et al. 2004). This perspective acknowledges the potential importance of multiscale processes in which *localities*, that hold *local* taxa assemblages, are connected to other local assemblages within a *region*. A number of meteorological dispersal models for aerosolized microbes have been previously applied, generally relying on Eulerian and Lagrangian vector field approaches from multivariate calculus (Yamamoto et al. 2012, Wilkinson et al. 2012, Schmale and Ross 2015). While neither acknowledging metacommunity theory nor explicitly considering precipitation-borne organisms, these models emphasize the dynamic, dispersal-driven constraints of aerosolized microbes, in which weather patterns serve as multi-scale conduits among local microbial assemblages.

We note that a current lack of metacommunity studies of bioaerosols can be linked partially to difficulties with the term community (*sensu* Ricklefs 1976, Stroud et al. 2015) in microbial systems -- because interactions among co-occurring microbial taxa are often negligible or unknown (Boon et al. 2014). This is particularly true for atmospheric microbes; whose densities are much lower than those in terrestrial and aquatic environments. Despite these issues, however, tools for the description and analysis of ecological communities and metacommunities have been extremely useful in clarifying processes and characteristics of microbial assemblages (Prosser et al. 2007, Boon et al. 2014).

Partitioning β -diversity. A primary use of metacommunity theory is the elucidation of mechanisms underlying the structure and composition of taxa at multiple spatiotemporal scales (Leibold et al. 2004). For this purpose, metrics that quantify compositional heterogeneity among local assemblages, i.e., β -diversity, are often used (e.g., Winegardner et al. 2017). Total β -diversity, denoted β_{Total} , can be attributed to two processes: turnover of assemblage taxa, β_{turn} , and assemblage nestedness, β_{nest} (Baselga 2010). β_{turn} and β_{nest} , in turn, result from antithetic mechanisms among assemblages: taxa replacement and taxa loss, respectively (Baselga 2010).

To illustrate these properties of β -diversity partitioning in precipitation-borne microbial assemblages, we consider three hypothetical scenarios in which twelve microbial taxa (e.g., families, genera, species, OTUs) distinguish the bioaerosol composition of three precipitation events (Fig. 1). In scenario A, the assemblages of precipitation events are completely nested, with lower richness

events containing compositional subsets of richer events. Because all compositional heterogeneity among events is due to nestedness, $\beta_{\text{turn}} = 0$, and $\beta_{\text{nest}} = \beta_{\text{Total}}$. In scenario B, all three events contain six microbial taxa, and (like scenario A) events have three species in common. However, at each event, the three remaining taxa are replaced by three other taxa and these replacements are distinct for each event. Because all compositional heterogeneity among events is due to turnover, $\beta_{\text{nest}} = 0$, and $\beta_{\text{turn}} = \beta_{\text{Total}}$. In scenario C, both nestedness and turnover occur, and both contribute to β_{Total} .

A graphical metacommunity heuristic. Figure 2 presents a metacommunity framework for precipitation-borne microbes based on the information given earlier. Local microbial assemblages are distinguished with hatching and colored arrows show subsidization fluxes. Atmospheric bioparticles with smaller aerodynamic diameters, including bacterial cells, can disperse to higher altitudes (DeLeon-Rodriguez et al. 2013, Bryan et al. 2019), where they generally occur at lower concentrations (Burrows et al. 2009a, Gandolfi et al. 2013, Griffin et al. 2017, pg. 155), and travel further horizontally (long narrow red arrows). Because of higher velocity air currents, the composition of high altitude bioparticles will be spatially macroscale-driven. Thus, changes in local microbial assemblages should occur over large distances (gradual color changes in hatching). Further, because of repeated washout of microbes as storms move in time and space, coupled with isolation from surface environments –preventing subsidization of new taxa, β_{nest} should be relatively high among high altitude assemblages (Fig. 1, scenario A), and assemblages will be poorly associated with particular precipitation sites. In contrast, lower atmospheric altitudes will contain higher concentrations of bioparticles, including relatively large microorganisms (e.g., fungal spores and pollen), with poorer dispersal, and stronger ties to local (mesoscale) conditions on the terrestrial surface. Within the PBL, relatively large microbial fluxes should occur over small horizontal distances (short wide green arrows) due to the higher depositional velocities of larger microorganisms (Raupach et al. 1996) and atmospheric mixing, causing strong spatial differentiation among local aerosolized assemblages (dramatic color changes in hatching). In this case, β_{turn} should be relatively high among local assemblages, as spatiotemporal changes on the earth's surface drive corresponding changes to microbial assemblages in the adjacent atmosphere (Fig 1, scenario B). Wet deposition scrubs/scavenges microbes from overlaying atmospheric layers (Hanlon et al., 2017), causing

precipitation samples (oval at bottom right) to contain microorganisms of varying sizes, potentially aerosolized at both local and distant sites, and transported by air currents from different altitudes.

Hypotheses

To clarify properties of precipitation-borne microbial metacommunities, we collected a large, culture-independent dataset describing fungal and bacterial taxa by sampling precipitation events at multiple seasons and sites. To complement this information, we assembled a dataset describing the meteorological characteristics of each precipitation event at multiple altitudes. We considered these data from the perspective of our metacommunity framework (Fig. 2) by testing two hypotheses.

H₁: Compared to fungi, precipitation-borne bacterial composition is driven by macroscale factors associated with high-altitude dispersal and subsidization.

P_{1a}: Variation in bacterial assemblages will be strongly associated with higher altitude (> 3000 mAGL) characteristics of storms (e.g., high altitude back-trajectory distance, high altitude air mass origin, etc.), and weakly associated with local characteristics of particular rainfall sites.

P_{1b}: Compared to fungi, total β -diversity of bacteria will be largely attributable to nestedness effects (taxa losses compared to the overall metacommunity), because isolation from the earth's surface will prevent rapid replacement of taxa lost during washout.

H₂: Compared to bacteria, fungal composition in precipitation is driven by mesoscale factors associated with low-altitude dispersal and subsidization.

P_{2a}: Fungal composition in precipitation will be strongly associated with particular rainfall locations and lower altitude (< 3000 mAGL) characteristics of storms.

P_{2b}: Compared to bacteria, total β -diversity of fungi will be largely attributable to turnover (taxa replacement) reflecting locally-driven spatiotemporal filtering.

Materials and Methods

Collection of precipitation samples

From May 2013 to July 2015 a total of 72 distinct precipitation events were sampled at locations associated with three participating U.S. universities: 20 m Above Ground Level (mAGL) on the roof of the Life Sciences Building at Louisiana State University (LSU) in Baton Rouge, LA, USA (30° 25' 12" N, 91° 10' 48" W; 11 m Height Above Ellipsoid (HAE)), 20 mAGL on the roof of the Gale Life Sciences Building at Idaho State University (ISU) in Pocatello, ID, USA (42° 52' 3.85" N, 112° 25' 47.23" W; 1440 m HAE), 10 mAGL on the roof of Latham Hall at Virginia Tech (VT) in Blacksburg, VA (37° 13' 29" N, 80° 25' 22" W; 627 m HAE) and at three sites at VT's Kentland Farm: site A (37° 12' 9" N, 80° 33' 51" W; 515 m HAE), site B (37° 11' 47" N, 80° 34' 41" W; 531 m HAE) and site C (37° 12' 8" N, 80° 35' 32" W; 594 m HAE) (see Table 1). Sample dates were placed into the following seasonal categories: spring (March 29-June 12), summer (June 25-Sept 6), winter (Dec 22-March 16) and autumn (Oct 12 - Dec 19) (Table 1). Sixty-six of the 72 precipitation events were rainstorms and the remaining six events were ice/snow storms. Ice/snow storms occurred at all three sites.

Bacterial 16S and fungal 18S rRNA gene amplicon libraries were sequenced from precipitation samples collected during all four seasons at LSU, and three seasons at ISU (spring, summer and winter), and VT (summer, winter and autumn) (Table 1). Relatively rare heavy-precipitation events contributed to the smaller number of samples at ISU. The 30 precipitation events sampled for 18S rRNA genes constituted a subset of the 71 events sampled for 16S except for one case (LSU-spring) in which a precipitation event was sampled for 18S but not 16S rRNA genes (Table 1). The subset of 18S analyses was due to the depletion of sample materials from the 16S rRNA gene amplification procedures, which generally preceded 18S rRNA gene sequencing. In 29 precipitation events both 18S and 16S rRNA genes were successfully amplified and sequenced.

For each sampling event, precipitation was collected in ~120 L galvanized collection cans that were lined with clean, sterile 94 x 122 cm polypropylene bags (Fisher Scientific, Pittsburgh, PA; Failor et al. 2017). A minimum volume of 3L was amalgamated from collection cans and sample collection was conducted using a total of eleven cans, ten for samples and one as a procedural control

(see Appendix S1, Fig S1). Immediately following each precipitation event, collected material was transferred to sterile 9L carboys (Appendix S1, Fig S2) or left in the collection bags and stored at 4 °C preceding concentrating by filtration, which was always completed within 48 h of precipitation collection. Precipitation was filtered (1-3 L) onto 0.2 µm, 47 mm Supor PES membrane filters or Millipore polycarbonate filters. Filters were placed into sterile 15 mL or 50 mL falcon tubes and frozen preceding further processing. All samples from Louisiana and Virginia were shipped overnight on dry ice to ISU for consolidation with ISU samples.

Sampling procedures were designed to ensure that 16S and 18S rRNA gene sequences were not obtained from contaminated sampling equipment, sampling processes, or downstream molecular procedures. Procedural controls were created by lining one galvanized steel can with a sterile polypropylene bag (just as was done for precipitation sample collection) and covering the can with an ethanol-cleaned lid (Appendix S1, Fig S1). One L of autoclaved nanopure water was poured into the control bag and this water was filtered in the same manner as the precipitation samples. All filters containing procedural controls were put through the same DNA extraction and gene amplification procedures to assess spurious amplification of 16S and 18S rRNA genes due to procedural and/or reagent contamination.

DNA extraction from precipitation samples

Microbial wet lab procedures, including choices of primers, followed guidelines of the Earth Microbiome Project (Thompson et al. 2017). In preparation for DNA extraction, filters were cut using ethanol cleaned and flamed scissors. All filter cutting was performed in a laminar flow hood and filters were only handled with sterile tweezers. Each filter was cut into the smallest possible pieces (< 5mm in length) just above a bead beating tube from the FastDNA SPIN Kit for Soil (MP Biomedicals, USA). Once in the bead beating tube, the manufacturer's protocol for the FastDNA SPIN Kit for Soil was followed, starting with the addition of MT and phosphate buffer solutions. The manufacturer's protocol was followed with the exception of the following steps: in step 4, a mini bead beater was used to homogenize extracts for 70 seconds, step 5 was performed for 8 minutes, a 15 mL tube was always used for step 7, and 100 µL of DNase/pyrogen-free water was used to elute DNA in step 16. The resulting extracts were further purified with the aid of a silica spin filter, following steps

14-22 of the protocol for the MoBio Power Soil Kit (MoBio Laboratories, Carlsbad, CA). In these steps, 160 μ L of solution C4 was added to the DNA extraction. The manufacturer's protocol was followed for other steps, with the exception of adding 25 μ L of solution C6 and incubating for 1 minute at room temperature prior to centrifugation. All DNA extracts were stored at -20 °C until PCR amplification of the 16S and 18S rRNA genes was attempted.

Bacterial 16S rRNA gene fragment amplification

The V4 regions of bacterial 16S rRNA gene fragments were PCR-amplified in triplicate from each DNA extract for sequencing. PCR amplification was carried out using barcoded primers following the methods of Caporaso et al. (2012). The 25 μ L PCR mix contained the following: 1X 5-PRIME HotMaster Mix, 0.2 μ M 515F, 0.2 μ M 806R and nuclease free water (13 μ L). Thermal-cycling was carried out using 1-6 μ L of DNA template in an Eppendorf PRO S Master Cycler under the following conditions: initial denaturation at 94 °C for 3 minutes, 32 cycles of denaturation at 94 °C for 1 minute, annealing at 50 °C for 1 minute and extension at 72 °C for 1:45, followed by a final extension at 72 °C for 10 minutes. We resolved electrophoretically separated PCR products on a 1% TAE gel stained with ethidium bromide to confirm successful amplifications in each reaction. Triplicate reactions were pooled before purification and concentration using the Qiagen MinElute PCR Cleanup kit (Qiagen, Valencia, California, USA). A 35% guanidine-HCl wash step was performed to ensure that large primer-dimers were removed from the samples. Amplicons were stored at -20 °C preceding sequencing at Idaho State University's Molecular Research Core Facility (Pocatello, ID, USA) on the Illumina MiSeq Platform using the V2 500 bp kit.

Fungal 18S rRNA gene fragment amplification

Fungal 18S rRNA gene fragments from the V9 region were PCR-amplified in triplicate from each DNA extract. Steps for 16S amplification were also followed for 18S amplification with two exceptions. First, the 25 μ L PCR mix contained 1X 5-PRIME HotMaster Mix, 0.2 μ M 1391F, 0.2 μ M EukBr and nuclease free water (13 μ L). Second, thermal cycling used the following conditions: initial denaturation at 94 °C for 3 minutes, 32 cycles of denaturation at 94 °C for 0:45, annealing at 57 °C for 1 minute and extension at 72 °C for 1:30, followed by a final extension at 72 °C for 10

minutes.

Sequence data analysis

Sequence data were analyzed using the mothur (ver. 1.42.3) software package (Schloss et al. 2009). Contigs were assembled and parsed on the basis of unique barcodes attached to the 806R primer (Caporaso et al. 2012). Sequences that did not contain exact matches to the primer and barcodes utilized in the PCR amplification were discarded. Sequences were filtered for quality with a 50-base sliding window. Sequences not obtaining an average quality score of 25, containing ambiguous bases, homopolymers (> 7 bases), having lengths > 259 bases, or deemed chimeric using the UCHIME algorithm (Edgar et al. 2011) were eliminated.

All 16S rRNA gene sequences were aligned to the SILVA bacterial 16S rRNA gene reference (release 132) (Pruesse et al. 2007) truncated to contain the amplified V4 hypervariable region. Sequences were taxonomically classified using the classify.seqs command with arguments set to default values, with the exception of the minimum consensus confidence value being set at 80%. All 16S rRNA gene sequences that were classified as eukaryotic, archaea, mitochondria, or chloroplast were removed from analyzed datasets. Further, pseudoreplicate 16S rRNA gene samples with fewer than 40,000 total reads, following removal of non-informative OTUs, were eliminated from downstream analyses, and thus were not averaged with other pseudoreplicates to obtain a single independent 16S observation per storm (see Appendix S1: Table S2). The 16S rRNA gene sequence library size was normalized by randomly sampling a total of 44,621 sequences from each library, which was the number of sequences in the smallest library.

The sequences of 18S rRNA gene amplicons for each sample were aligned in mothur to the SILVA 18S rRNA gene reference (release 132), truncated to contain the V9 hypervariable region, with minimum consensus confidence set at 80%. Non-fungal sequences were eliminated from the 18S dataset. Analogous to methods for bacteria, pseudoreplicate fungal samples with fewer than 20,000 reads were eliminated. Average read totals for events were approximately 40% lower for 18S than for 16S rRNA gene sequence libraries. The 18S sequence library size was normalized by randomly sampling a total of 28,586 sequences from each library; the number of sequences in the smallest library.

Both 16S and 18S rRNA gene sequences were clustered into operational taxonomic units (OTUs) based on a sequence dissimilarity of ≤ 0.03 , using the Needleman-Wunsch algorithm (Kunin et al. 2010, Smith and Peay 2014). Following the quality control procedures described above, the 16S rRNA gene dataset contained 34,107 OTUs describing 71 independent precipitation event samples, and the 18S rRNA gene dataset contained 16,597 OTUs describing 30 independent precipitation event samples. Both 16S and 18S rRNA gene sequence libraries were deposited into MG-RAST metagenomics RAST (Rapid Annotation using Subsystem Technology) server (<https://metagenomics.anl.gov>). Accession numbers for sequences are given in Appendix S1 (§ S1.6.1).

Collection of Meteorological Data

Storm structure classifications were based on cloud top data retrieved from The National Weather Service (NWS) archive of Geostationary Operational Environmental Satellite-East (GOES-East) infrared satellite imagery, and Next Generation Radar (NEXRAD). Level II and III radar reflectivities were provided by the National Climatic Data Center's (NCDC) website (<https://www.ncdc.noaa.gov/nexradinv/>). Troposphere temperature profiles were retrieved from The NWS radiosonde data archive of stations located in Lake Charles, Louisiana (KLC) Slidell Muni, Louisiana (KLI), Pocatello, Idaho (KSFX), and Roanoke Virginia (KFCX).

Seventy-two-hour backward trajectories of air masses present over Baton Rouge, LA, Blacksburg, VA, and Pocatello, ID at times of collection were calculated using the NOAA Air Resources Laboratory Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model. Altitudes of 100, 1000, 5000, 6000, 8000 and 9000 mAGL were chosen for trajectory examination to represent vertical sections of ecological importance in our metacommunity framework (Fig. 2). The 9000 m AGL upper altitudinal bound was chosen due to its proximity to the homogeneous freezing layer ($\sim -38^{\circ}\text{C}$), where ice nucleation by heterogeneous mechanisms is no longer necessary for the freezing of water molecules.

Classification of convective vs. stratiform precipitation was based on the methods of Biggerstaff and Listemaa (2000), and Anagnostou (2004), in addition to the inspection of tropospheric stability indices from National Weather Service soundings (NWS 2018)

<https://www.weather.gov/lmk/indices>), and infrared and visible satellite imagery obtained from GOES-East (NOAA GOES-East 2018). Stratiform precipitation is defined specifically as precipitation collected from stratus and nimbostratus-like cloud systems independent of trailing stratiform regions from convective storm formations (Biggerstaff and Listemaa 2000, Houze 2014). Precipitation events were also assigned to six morphological classes: Air mass, Mesoscale Convective System (MCS), Multicell, Nimbostratus, and Squall.

Precipitation event trajectories were classified into five storm origin categories based on a conventional scheme for North American air masses (AMS 2018). Seventy-two-hour trajectories, which originated over marine surfaces, were classified as maritime (m), whereas those that originated over terrestrial surfaces were classified as continental (c). Likewise, those events that originated at latitudes above the subtropical regions of the Northern Hemisphere ($> 23.5^{\circ}\text{N}$) were classified as Polar (P), and those originating in the subtropical and tropical regions of the Northern Hemisphere ($< 23.5^{\circ}\text{N}$) were classified as tropical (T). Thus, six possible air mass combination classes were considered: Pacific Maritime Polar (mP1), Atlantic Maritime Polar (mP2), Pacific Maritime Tropical (mT1), Atlantic Maritime Tropical (mT2), Continental Polar (cP), and Continental Tropical (cT).

Statistical Analysis

General constraints. Meteorological data, and univariate summaries of microbial data, were largely analyzed using conventional univariate statistical approaches including general linear models (e.g. ANOVA, ANCOVA) or robust analogues, and tabular methods. Microbial compositional summaries and hypothesis tests required multivariate approaches. To address missing cells, a single ISU autumn and VT summer observation were imputed for ANOVA-type analyses considering site and season interactions, using the random forest algorithm (Stekhoven and Bühlmann 2011). Imputed points were assumed to be missing completely at random (MCAR). Lack of balance in site and season levels prompted the use of type II (marginal) sums of squares in ANOVA-type analyses, or the emulation of type II sums of squares by changing the order of predictors in linear models and reporting summaries from the last specified predictor. Family-wise type I error (FWE) was controlled, in pairwise comparisons following omnibus ANOVA-type analyses, using Tukey's (1949) HSD method. In multiple pairwise tests subsequent to non-ANOVAs, Holm's (1979) sequential Bonferroni procedure

was used to control FWE (see Aho 2013, Ch. 10). For indicator species analyses (see below) large families of tests for individual taxa were required. In this case, p -values were adjusted for false discovery rate (FDR) within taxonomic hierarchies (i.e., phylum, class, order, family, genus) using the method of Benjamini and Hochberg (1995). Control of FDR is more appropriate than control of FWE in exploratory applications involving a large number of null hypothesis tests (cf. Hulsman et al. 2013). Throughout the paper we use the significance level $\alpha = 0.05$.

Diversity. We measured bacterial and fungal OTU α -diversity in precipitation events using the Shannon-Wiener index (Shannon 1948). We partitioned total β -diversity (β_{Total}) into nestedness (β_{nest}) and spatial turnover (β_{turn}) components using the incidence-based pairwise measures given in Baselega (2010), i.e., the means of the β_{sor} , β_{nes} , and β_{sim} dissimilarity matrices, respectively. In the context of our work, β_{nest} represents variation in bacterial and fungal local assemblages due to subsetting of richer assemblages (Fig 1; scenario A), whereas β_{turn} represents local assemblage heterogeneity due to taxa replacement in precipitation events over spatiotemporal gradients (Fig 1; scenario B). By definition, β_{Total} , β_{nest} , and β_{turn} are in the interval $[0, 1]$, with larger proportions indicating greater assemblage heterogeneity, and $\beta_{\text{Total}} = \beta_{\text{nest}} + \beta_{\text{turn}}$. Bootstrapping, based on random sampling of rain events with replacement, was used to derive standard errors for β_{Total} , β_{nest} , and β_{turn} and to perform null hypothesis tests of equal partitioning components. Controls of microbial assemblage nestedness were examined using the matrix temperature approach of Rodríguez-Gironés, and Santamaría (2006).

For our study, OTUs were defined using rRNA gene sequences of the V4 and V9 regions for bacteria and fungi, respectively. However, inherent physical differences and putative differences in evolutionary history exist for these gene regions. These discrepancies limit comparability of bacterial and fungal diversity using rRNA gene sequences. For example, in a simpler application involving only eukaryotes, Tragin et al. (2017) found that V4 and V9 regions of 18S rRNA genes in phytoplankton samples had similar levels of α -diversity, but the V9 region identified 20% more OTUs. To address these issues, we constrained fungal and bacterial comparisons in four ways. First, we formally compared bacteria and fungi diversity using only β -partitioning measures (i.e., trends in overall heterogeneity of assemblages due to OTU turnover and nestedness), as reflected in our

hypotheses. Second, to address richness biases in the 16S V4 regions and 18S V9 rRNA gene regions, we chose β -partitioning metrics shown to be independent of the richness of OTU assemblages (Baselga and Leprieur 2015). Third, we corrected for inherent differences in the heterogeneity of 16S and 18S rRNA genes by dividing fungal and bacterial β_{nest} and β_{turn} measures by their respective measures of β_{Total} . Thus, in bacterial/fungal comparisons, both β_{nest} and β_{turn} were scaled to be in [0, 1], and represented the proportion of variability in total β -diversity due to nesting and turnover, respectively. Fourth, formal comparisons of fungi and bacteria β -partitioning measures were based only on the 29 rain events for which both 16S and 18S rRNA gene sequences were obtained. The complete set of 72 rain events were a mix of paired (spatiotemporally dependent) and non-paired events (Table 1). Bootstrapping significance tests were constrained so that bacterial and fungal data were paired by precipitation event preceding sampling with replacement. Corresponding p -values were based on one-tailed (reflecting hypothesized directions) percentile confidence intervals (Aho 2013) for the true difference in β -partitioned values.

PERMANOVA. Permutational Multivariate Analysis of Variance (PERMANOVA; Anderson 2001) was used to consider the effect of macroscale (large spatial scale and seasonal), and mesoscale (site-specific) drivers of bacterial and fungal precipitation-borne assemblages (Fig. 2). An unresolved problem in next-generation sequencing (NGS) of PCR rRNA gene amplicons is the occurrence of microorganisms whose genomes contain multiple copies of the rRNA gene, causing particular OTUs to have inflated read abundances (Herrera 2009). To address this issue, we conducted PERMANOVAs using both raw read abundances of OTUs and binary-transformed (presence/absence) OTU data. PERMANOVA p -values were based on 10000 permutations.

To decrease the propensity for single precipitation events or OTUs to influence multivariate analyses, we applied a Wisconsin double transformation (Bray and Curtis 1957) to both 16S and 18S OTU by rain event data matrices. Because of computational constraints, random forest imputations of single Idaho and VT missing cells (and concomitant PERMANOVAs) were based on only the 1500 most abundant 16S and 18S OTUs. Bray-Curtis dissimilarity (Bray and Curtis 1957) was used as the underlying resemblance matrices in PERMANOVA and ordinations (see below). Bray-Curtis

dissimilarity effectively represents and distinguishes multivariate assemblages given sparse (few non-zero entries) datasets which may contain many samples with nothing in common (Aho et al. 2008).

Ordination. The relation of precipitation events in OTU space was depicted using Nonmetric Multidimensional Scaling ordinations (NMDS; Kruskal and Wish 1978). These analyses were based on non-imputed datasets with all recorded OTUs. The final NMDS projections were based on the best (lowest stress) solutions resulting from twenty randomized starting configurations. To enhance interpretability, NMDS axes were centered and rotated to principal components. Vector (quantitative data) and factor (categorical data) fitting (Oksanen et al. 2016) was used to examine correlations between raw meteorological data and NMDS scatter. Permutational approaches were used to test null hypotheses of no association in these procedures using 1000 iterations.

Indicator Species Analysis. Indicator species analysis (ISA; Dufrene and Legendre 1997) using non-imputed data was used to test for association of taxa designations of bacterial and fungal OTUs to season and site levels. All unclassified taxa (i.e., “unclassified”, “uncultured”, “unknown”, “Incertae Sedis”, “unnamed”) were eliminated from ISA procedures. Specific taxa were required to occur at two or more precipitation events for inclusion in analyses. *P*-values for tests of the null hypothesis that associations between species and categorical levels were no better than random were based on permutation procedures using 1000 iterations.

Software. All analyses were performed within the R statistical environment (R Core Team 2017) with heavy reliance on the packages *asbio* (Aho 2018) and *car* (Fox and Weisburg 2011) for implementation of conventional general linear models; *vegan* (Oksanen et al. 2016) for most compositional analyses of assemblages including PERMANOVA, ordinations, and matrix temperature calculations; *betapart* (Baselga and Orme 2012) for beta-diversity partitioning analyses; *labdsv* (Roberts 2016) for ISA; and *missForest* (Stekhoven 2013) for imputation.

Results

Meteorological Differences among Sites and Seasons

As expected (Fig. 2), higher altitude atmospheric layers traveled further in 72 h back-trajectories (slope $\approx 0.01 \text{ m m}^{-1} \text{ hr}^{-1}$, $F_{1,420} = 391.3$, $p < 2.2 \times 10^{-16}$). Significant differences in distance traveled occurred by site when holding season and elevation constant ($F_{2,420} = 11.2$, $p = 0.0002$), with Idaho storms traveling the furthest and LSU storms arising more locally. These differences were increasingly apparent at higher altitudes (interaction effect: $F_{2,420} = p = 0.0007$). Strong seasonal effects were also evident ($F_{3,420} = 41.58$, $p < 2.2 \times 10^{-16}$), with shorter summer trajectories and longer trajectories in the winter. No trajectory distance interactions occurred between site and season.

Fisher's exact tests identified significant differences ($p \leq 0.03$) in cloud air mass origins (i.e., cP, cT, mP1, mP2, mT1, mT2) between sites at all altitudes but one (8000m) when controlling for FWE using Holm's (1979) procedure. These differences were largely due to the maritime/tropical origin of most LSU and VT precipitation events, compared to Idaho events, which were mostly polar/continental. Idaho maritime events were never Atlantic, whereas VT and LSU maritime events were frequently Atlantic in origin (Fig. 3). Significantly ($p < 0.025$) distinct seasonal characteristic for air mass origin were evident at all altitudes above 5000 after controlling for FWE. This finding was largely due to precipitation events from LSU (the best-sampled site) being mostly maritime, except for summer storms which were primarily continental (Fig. 3). Weather morphology classes (Airmass, MCS, Multicell, Nimbostratus, Squall) differed significantly by season ($p = 0.008$), with nimbostratus events occurring primarily in the winter and fall, but not site ($p = 0.389$; Fig. 3). Similar patterns were apparent for stratiform/convective classes with significant differences occurring with season ($p = 0.003$), due to the absence of stratiform storms during summer months, but a lack of strong differences with site ($p = 0.102$).

Microbial Diversity Varies with Season, and Particularly Sites

Site, season, and season by site interactions were not significantly associated with richness and Shannon diversity for either the bacterial or fungal datasets (Table 2). LSU had significantly higher bacterial β_{turn} and β_{Total} than ISU, and LSU had significantly higher β_{turn} and β_{Total} than ISU in the fungal dataset (Table 2; see Appendix S1: Tables S86 and S87 for bootstrap confidence intervals). Thus, for both 16S and 18S rRNA gene data, LSU events had higher levels of compositional heterogeneity, particularly compared to ISU, and this heterogeneity was due to relatively high taxa

turnover in the bacterial dataset and relatively high taxa nestedness in the fungal dataset.

Fungal and Bacterial β -Diversities are Distinct-- β_{turn} higher for Fungi, β_{nest} higher for Bacteria.

Fungal and bacterial β -diversity partitioning components differed. Spatiotemporal turnover was higher for fungal assemblages in precipitation, causing $\beta_{turn}/\beta_{Total}$ outcomes for fungi to be significantly higher than for bacteria ($p < 0.0001$; percentile bootstrap confidence intervals; $R = 1000$; Fig. 4). Assemblage nesting was higher for bacteria, causing bacterial $\beta_{nest}/\beta_{Total}$ values to be significantly higher than for fungi ($p < 0.0001$; Fig. 4).

Bacterial Composition Varies Subequally with Site and Season. Fungal Composition Varies with Site

PERMANOVA analyses of raw and binary transformed data were in strong agreement (Appendix S1: Tables S24, S25, S56, S57). Thus, we present only analyses of raw reads here (Table 3). For 16S rRNA gene library reads, OTU compositional differences had identical levels of statistical significance for site ($F_{2,61} = 6.32$, $p = 0.0001$) and season ($F_{3,61} = 2.7$, $p = 0.0001$). For 18S rRNA gene reads, however, much stronger compositional differences were apparent among sites ($F_{2,20} = 3.1$, $p = 0.0002$) compared to seasons ($F_{3,20} = 1.46$, $p = 0.06$). Interactions between site and season were absent in both bacterial and fungal datasets, indicating that seasonal assemblage fluctuations were somewhat consistent across sites (Table 3). In pairwise comparisons of bacterial OTU composition, all sites and seasons were significantly different after controlling FWE (Table 3). In pairwise comparisons of fungal OTU composition, all sites were significantly different (Table 3). Seasons were not compared in the fungal dataset due to a lack of omnibus significance. Notwithstanding the effects of smaller sample sizes and inherent inconsistencies in 16S and 18S rRNA gene heterogeneity, our PERMANOVA results suggest that compositional variations were smaller for fungal assemblages, compared to bacterial assemblages, particularly with respect to seasons.

Ordinations Demonstrate Spatiotemporal Variation in Assemblages. Patterns of variation in microbial metacommunity composition detected in PERMANOVA are reflected in NMDS ordinations of 16S and 18S rRNA gene data with respect to both site (Fig. 5) and season (Fig. 6). Also evident, with vector and factor fitting, were the longer distances traveled by Idaho events over 72 hour

back-trajectories, and the continental and Atlantic sources of LSU and VT storms, compared to the Pacific maritime and polar sources of Idaho storms (Fig. 5a, 6a). Correlations of the NMDS projections to meteorological variables are given in Table 4. Significant 16S rRNA gene correlates included season ($p = 0.025$; $R^2 = 0.11$), but not site ($p = 0.154$, $R^2 = 0.13$). Correlations of the projection to storm travel distances strengthened as elevation increased. Specifically, for elevations of 100, 1000, 5000, 8000, and 9000m, R^2 values were: 0.01, 0.06, 0.15, 0.23, 0.22, 0.19, respectively, and p -values were: 0.64, 0.10, 0.02, 0.001, <0.001, and 0.001, respectively (Table 4). These results demonstrate that strong relationships exist between meteorological measures and bacterial OTU composition, particularly storm characteristics at higher altitudes.

In the 18S rRNA gene ordination (Figs. 5b, 6b), fewer significant relationships occurred between abiotic variables and the projection of precipitation events; and the ordination solution was inferentially weaker (stress = 0.159), compared to the bacterial ordination (stress = 0.08). Despite these differences, we present two-dimensional solutions for both datasets to facilitate their comparison. The fungal projection was strongly associated with site ($p = 0.002$; $R^2 = 0.28$; Table 4) but not season ($p = 0.13$; $R^2 = 0.17$; Table 4), reversing trends observed for bacteria (Table 4). Other significant correlates of the fungal ordination were meteorological characteristics of moderate altitudes, i.e., air mass categories at 1000m ($p = 0.028$; $r = 0.25$), and 6000m ($p = 0.003$; $r = 0.35$), and precipitation type, i.e., rain or ice ($p = 0.01$, $R^2 = 0.15$). Thus, the NMDS ordination of bacterial OTUs was more strongly correlated with variation in large spatiotemporal scale measures, e.g., season and higher altitude back trajectory distance, whereas the NMDS ordination of fungal OTUs was more strongly correlated with smaller scale spatiotemporal measures, including particular site locations, and moderate altitude air mass categories (Table 4).

Spatiotemporal Dynamics of Taxonomic Composition

The average consensus confidence level for alignment of 16S rRNA gene sequences to taxa was greater than 98.3% for all taxonomic levels, and equaled 100% for levels above order (Appendix S1: Table S28). For 18S rRNA gene sequences, consensus confidence was greater than 98.9% for all taxonomic levels, and equaled 100% for levels above order. For 16S rRNA gene data, unclassified taxa comprised 4%, 5%, 11%, 13%, and 31% of reads across samples for phylum, class, order, family

and genus level designations, respectively. Taxonomic designations based on 18S rRNA gene sequences included a large proportion of unclassifiable OTUs for levels below order. Specifically, unclassified taxa constituted 0%, 9%, 32%, 60%, and 71% of reads for phylum, class, order, family and genus level designations, respectively. As a result, we do not present ISA results for fungal taxonomic levels below order. Complete results for fungal families and genera taxa, however, are given in Appendix S1. A large number of bacterial and fungal taxa were statistically significant indicators of sites and/or seasons (Table 5). Thus, only the highest average abundance bacterial and fungal taxa are summarized graphically (Figs 7-10), and only these taxa are considered in the text below. Complete ISA results and taxonomic summaries are given in Appendix S1.

Bacteria. Six named phyla and one unclassified group comprised 98.4% of the total number of 16S rRNA gene sequence reads. In particular, Proteobacteria, Cyanobacteria, Bacteroidetes, Actinobacteria, Firmicutes, Acidobacteria and “unclassified” made up 58%, 14%, 8.1%, 7.6%, 6.1%, 1.6% and 3.6% of read totals, respectively (Figs. 7, 8). Proteobacterial subdivisions were dominant in lower taxonomic hierarchies with Gammaproteobacteria, Alphaproteobacteria and Deltaproteobacteria making up 26.1%, 25.9% and 4.9% of class read totals, respectively, and, Pseudomonadales (Gammaproteobacteria), Betaproteobacteriales (Betaproteobacteria), Rhizobiales (Alphaproteobacteria), and Sphingomonadales (Alphaproteobacteria), and Burkholderiales (Betaproteobacteria) making up 9.3%, 8.9%, 8.7%, and 8.5%, respectively, of order read totals.

A large percentage of bacterial taxa were significant study site indicators (18% of phyla, 20%, of classes, 18% of orders, 22% of families, and 21% of genera), although much lower percentages of significant taxa remained after controlling for FDR (8% of phyla, 0%, of classes, 3% of orders, 6% of families, and 5% of genera) (Table 5). With respect to bacterial indicators of seasons, 18% of phyla, 21%, of classes, 17% of orders, 17% of families, and 14% of genera were significant indicators, and 0% of phyla, 1%, of classes, 2% of orders, 3% of families, and 2% of genera were statistically significant after controlling for FDR (Table 5).

ISU had the highest number of significant bacterial site indicators across all taxonomic levels (Appendix S1: Table S45). Ignoring FDR, ISU significant indicators included Actinobacteria (phylum), Actinobacteria (class), Cytophagales and Micrococcales (order), Burkholderiaceae,

Sphingobacteriaceae, and Microbacteriaceae (family) and *Hymenobacter* (genus) (Fig. 7). FDR-adjusted significant ISU indicators included *Massilia* (genus). Ignoring FDR, significant LSU indicators include Archangiaceae and Chroococcidiopsaceae (family), and *Methylobacterium* (genus). FDR-adjusted significant LSU indicators included Nostocales (order) and *Cloacibacterium* (genus). Ignoring FDR, significant VT indicators included Acidobacteriia (class) and Acidobacteriales (family). Significant FDR-adjusted VT indicators included X1174.901.12 (genus).

The season with the largest number of significant bacterial indicators was winter, followed by summer, spring and autumn (Appendix S1: Table S39). Indicators of autumn included the Pseudomonadales (order), and its subtaxa Pseudomonadaceae (family), and *Pseudomonas* (genus) (Fig. 8). Indicators of spring included Alphaproteobacteria (class), and its subtaxa: Sphingomonadales (order) Sphingomonadaceae (family) and *Sphingomonas* (genus). The latter three remained statistically significant after controlling FDR. Indicators of summer included three taxa that remained significant after controlling FDR: Deltaproteobacteria (class), and its subtaxa Myxococcales (order), Cystobacteraceae (family), and *Melittangium* (genus), along with the genus *Cloacibacterium*. Indicators of winter included two taxa that remained significant after controlling FDR: Hymenobacteraceae (Family) and *Hymenobacter* (genus), along with Clostridia, Thermoleophilia, and Verrucomicrobiae (class), Cytophagales (order), and *Pedobacter* (genus).

Fungi. A relatively large number of fungal taxa (20% of phyla, 15%, of classes, and 29% of orders) were significant site indicators, and 6% of orders remained statistically significant site indicators after controlling for FDR (Table 5). As with bacterial results, ISU had the highest number of indicative taxa, with 15 significant ISA orders, before controlling for FDR, compared to LSU (five orders) and VT (seven orders) (Appendix S1: Table S76). Fungal taxa were poor predictors of seasons. No phyla, only 12% of classes, and 4% of orders were statistically significant ISA season indicators (Table 5, Fig 10), and only 2% of genera were significant seasonal indicators after controlling for FDR (Table 5). The most distinct season levels were summer and spring, which had eight and seven significant order indicators, respectively, before controlling for FDR (Appendix S1: Table S71).

Fungal indicators for ISU included Cystobasidiomycetes, Arthoniomycetes, and Tremellomycetes (class) and Capnodiales and Filobasidiales (order). The latter was significant after

controlling FDR (Fig. 9). LSU indicators included Basidiomycota (phylum). VT indicators included Auriculariales (order). Ignoring FDR, spring was indicated by Arthoniomycetes and Taphrinomycetes (class), and summer was indicated by Ustilaginomycetes (class) and the order Ustilaginales (Fig. 10).

Discussion

Bioaerosols constitute a vital component of the biosphere. However, fundamental ecological components of these systems, including microbial composition of precipitation, and the spatiotemporal dynamics and drivers of these assemblages, remain poorly understood (Delort and Amato 2017). The limited number of existing studies of precipitation-borne microbes have: 1) generally not considered taxa other than bacteria, or distinguished INPs, 2) concentrated on nearby or identical sites, and/or few points in time, and 3) have not effectively utilized existing analytical and theoretical frameworks for community and metacommunity ecology (Bigg and Miles 1964, Bigg 1973, Constantinidou et al. 1990, Amato et al. 2007, Christner et al. 2008, Kourtev et al. 2011, Huffman et al. 2013, Joly et al. 2013, Tobo et al. 2013, Šantl-Temkiv et al. 2015, DeMott et al. 2016). To address this deficiency, we conducted comparative analyses using what is, to date, the most extensive dataset of large-scale spatiotemporal trends in microbial assemblages distributed in precipitation. These data were considered using a metacommunity heuristic (Fig 2) that predicts bioaerosol dispersal patterns are largely a function of altitude, which in turn serves as an environmental filter based on organism size. The metacommunity framework prompted the evaluation of two hypotheses.

Hypothesis 1: Precipitation-borne bacterial assemblages are correlated with macroscale factors

Bacteria assemblages in precipitation were more strongly associated with macroscale spatiotemporal processes than fungi (H_1). This trend was demonstrated through both the relatively strong associations of bacterial assemblages to non-local high-altitude meteorological variables (P_{1a}), and the higher levels of nestedness in bacterial assemblages, compared to fungi (P_{1b}).

Prediction 1a. Assemblages of bacterial OTUs were strongly associated with spatially distant storm characteristics, compared to fungal OTUs. For instance, strong correlations occurred between, high-

elevation air trajectory distances and the bacterial OTU NMDS ordination, although this effect was absent in the fungal projection (Table 4).

Strong bacterial macroscale associations were also evident from the perspective of large-scale *temporal* (seasonal) drivers. Season explained more variation in the bacterial OTU ordination than site in several multivariate analyses, and this trend was reversed for fungal OTUs (Figs. 4, 5; Tables 3, 4). This suggests that macroscale seasonal forces are more important than short-interval temporal subsidies in determining the composition of precipitation-borne bacterial assemblages. This finding was reiterated in taxonomic analyses, as a large number of significant bacterial indicators occurred for season compared to fungi (Table 5, Figs. 7 and 9). The distinctiveness of ISU bacterial OTU composition (Table 3, Figs. 5, 8) is likely due to the large spatial distances of ISU to LSU and VT, and ISU's unique meteorology (e.g., Pacific maritime and polar storm sources), compared to LSU and VT, which are more similar meteorologically (Fig. 3). Seasonal trends in the composition of aerosolized bacteria have been previously reported (e.g., Brodie et al. 2007, Bowers et al. 2013, Striluk et al. 2016), and a small set of studies have observed temporal variations in the bacterial composition of clouds or precipitation samples taken from the same or nearby locations (Kourtev et al. 2011, Vařtilingom et al. 2012, Huffman et al. 2013, řantl-Temkiv et al. 2015).

Although all four seasons had significantly different bacterial assemblages in pairwise comparisons (Table 3), the uniqueness of winter is evident through both the small area of its confidence ellipse in the bacterial ordination (Fig. 6), and in the relatively high number of significant ISA winter indicator taxa (Fig. 7). Along with the distinct terrestrial surface conditions of these seasons, a potential driver for this phenomenon is the relatively unique (non-ecotonal) conditions of winter storms. The morphology of winter storms (which were limited to nimbostratus and MCS events) were highly distinct compared to summer, autumnal and spring storms (Fig. 3).

Prediction 1b. Compared to fungi, the β -diversity of bacterial assemblages at particular sites had a larger β_{nest} component (due to species losses in local assemblages compared to the overall metacommunity species pool) and a smaller β_{turn} component (species replacement in time and space) (Fig. 4). The average bacterial OTU richness of individual precipitation events was only $1,911.4/34,107 = 5.6\%$ of the total richness of the entire 16S rRNA gene dataset, whereas the

corresponding fungal OTU richness was much higher: $1377.8/16,597 = 8.3\%$. This discrepancy may occur because fungal OTUs lost from washout are more readily replaced by other OTUs from local subsidization events.

Richness “hotspots” may contribute to high levels of β_{nest} (Wang et al. 2018). In our study, the richness sample variance was more than five times greater for bacteria than fungi ($1,134,132/205,963 = 5.5$) for paired events. The two most bacterially diverse precipitation events occurred at ISU (richness = 3865 OTUs) and LSU (richness = 4253 OTUs). These events were linked to very different meteorological processes (e.g., 72 hr. back trajectory storm origins at 9000m were mP1 and mT1 for ISU and LSU, respectively), suggesting that a variety of aerosolization conditions may result in high levels of bacterial diversity in precipitation.

Nestedness of bacterial assemblages in precipitation may be reinforced by the predicted shorter residence time of INA taxa in the troposphere (Burrows et al. 2009b, Stopelli et al. 2015). The gammaproteobacterial genus *Pseudomonas* contains the vast majority of documented INA microbial strains (Joly et al. 2013), and was assigned in our sequence analyses to 203 OTUs (see Methods). When considered alone, these 203 OTUs resulted in much higher levels of precipitation event nestedness ($\beta_{\text{nest}} = 0.16$; $\beta_{\text{nest}}/\beta_{\text{Total}} = 0.32$) than when considering all bacterial OTUs ($\beta_{\text{nest}} = 0.12$; $\beta_{\text{nest}}/\beta_{\text{Total}} = 0.16$). These results suggest that bacterial assemblages in precipitation events with few *Pseudomonas* OTUs are compositional subsets of events with many *Pseudomonas* OTUs--potentially from washout of particular INA strains within storm systems.

The nestedness-prone patterns of *Pseudomonas* OTUs are graphically depicted in Fig. 11. Rows (precipitation events) and columns (OTUs) in Fig. 11 are arranged to maximize event/OTU co-occurrences (black cells) in the upper left-hand part of the matrices (see Fig S34 in the Supplemental Material for a drastically simplified representation). Under *perfect* nestedness (see Fig. 1, scenario A), it would be possible to arrange the matrix so that *all* black cells were in the region above the designated red curve (Atmar and Patterson 1993, Rodríguez-Gironés, and Santamaría 2006). The matrix temperature index, denoted T , can be used to measure nestedness of multivariate observations, e.g., the microbial composition of precipitation events (Atmar and Patterson 1993). T decreases with nestedness with $T = 0$ indicating perfect nestedness. Matrix temperature was calculated for precipitation events using all bacterial OTUs (Fig 11a) and using only the 203 OTUs designated as

Pseudomonas (Fig. 11b). We found that the matrix temperature of events when only considering *Pseudomonas* ($T = 4.81$) was significantly lower than matrix temperature of events when considering all OTUs ($T = 7$), based on bootstrap tests ($p = 0.008$, 1000 iterations, Appendix S1: pg. 97).

To further assess the contributions of individual taxa to bacterial assemblage nestedness, we derived β_{nest} and β_{turn} dissimilarity matrices (Baselga 2010) based on the 1698 bacterial genera identified, and generated PCoA ordinations from those matrices (NMDS projections failed to converge for the β_{turn} matrix). The gene library reads for *Pseudomonas* were more strongly correlated to the PCoA projection of the β_{nest} dissimilarity matrix than for any other genera ($R^2 = 0.73$, $p = 0.001$; Supplementary Materials; Table S89). For the remaining bacterial genera, $R^2 \leq 0.38$ (Table S89). Further, *Pseudomonas* had one of the *weakest* correlations to the PCoA projection of β_{turn} dissimilarities ($R^2 = 0.02$, $p = 0.553$). These results support the premise that patterns of relatively high levels of nestedness and low levels of turnover within precipitation-borne aerosolized bacterial assemblages are driven in large part by selective washout of *Pseudomonas* OTUs as storm systems develop and age. These conclusions are supported by Stopelli et al. (2015) who found that “older” clouds (those less recently derived from a storm-generating source) have markedly lower proportions of INA taxa, due to ongoing precipitation events as storms move in space and time.

Hypothesis 2: Precipitation-borne fungal assemblages are correlated with mesoscale processes

In agreement with our heuristic, fungal assemblages were mostly associated with mesoscale spatiotemporal drivers (H_2). This outcome was evident in the strong associations of fungal OTU composition to particular spatial locations and moderate altitude characteristics of storms (P_{2a}), and the higher turnover (species replacement) contributions to β -diversity in fungi, compared to bacteria (P_{2b}).

Prediction 2a. Fungal assemblages were strongly associated with factors depicting relatively small changes in space and time. For example, the fungal NMDS projection was more strongly correlated with lower altitude air mass origin classes than the bacterial NMDS ordination (Figs. 5, 6; Table 4). Both PERMANOVA (Table 3) and factor-fitting (Table 4) identified stronger fungal OTU compositional associations for sites compared to continental-scale season effects. Further, ISA

identified few fungal seasonal indicators, and large number of significant site indicators (Table 5). This outcome is likely due to a number of physical factors including the fact that fungal spores and free-living aerosolized fungal organisms are generally much larger (2-56 μm diam.) than prokaryotic cells (0.2-10 μm diam.) (Amato et al. 2017, Griffin et al. 2017). Long-distance atmospheric dispersal of fungal spores is generally considered to be less than 10 km (Golan and Pringle 2017). Thus, the fungal composition of precipitation is more likely to be influenced by proximate aerosolization events, including agricultural soil disturbance, and spore dispersal as a part of sexual or asexual reproduction (Wilkinson et al. 2011).

A factor unconsidered by our metacommunity framework is the propensity of precipitation to scrub larger aerosols ($> 1.5\mu\text{m}$) more efficiently than smaller aerosols (Gregory 1961, Jung et al. 2011), although sub-micrometer bioaerosols can be effectively scavenged by precipitation (Pratt et al. 2009). Size-dependent scrubbing undoubtedly contributes to the composition of metacommunities in precipitation. Indeed, this factor emphasizes the desirability of INA phenotypes for taxa that may otherwise be washout-resistant and prone to prolonged separation from the earth's surface.

Prediction 2b. β_{turn} was the largest contributor to β_{Total} for both bacteria and fungi (Table 2, Fig. 4). Despite this general pattern, however, nestedness and turnover components of β_{Total} differed for bacterial and fungal assemblages. Compared to bacteria, fungal β -diversity at particular sites had higher contributions from turnover (taxa replacement) over time, and lower contributions from nestedness (Fig. 4). This outcome reiterates the importance of mesoscale drivers for fungal assemblages in precipitation, with taxa replacement corresponding to locality-specific temporal variations and environmental filtering (cf. Douda et al. 2018). Macroscale effects in airborne bacteria species sorting and opposing mesoscale mechanisms for fungal assemblages likely contribute to pronounced spatiotemporal variations in bacterial and fungal metacommunities conjoined in precipitation (e.g., Myers et al. 2012).

Taxonomic Composition of Sequence Libraries

Bacteria. Our findings agree with previous studies of aerosolized microbes that have reported dominance of Actinobacteria, Cyanobacteria, Firmicutes, Bacteroidetes, and Proteobacteria,

particularly Alphaproteobacteria and Gammaproteobacteria (Lindeman et al. 1986, Bowers et al. 2013, Striluk et al. 2016, many others). These findings are also in general agreement with the limited number of studies that have examined the microbes in clouds and precipitation (Vařtilingom et al. 2012, Huffman et al. 2013, řantl-Temkiv et al. 2015, Amato et al. 2017). Reflecting our work, cloud studies from puy de Dôme, France report high proportions of culturable bacteria in cloud water samples related to the genera *Pseudomonas* and *Sphingomonas*, along with notable levels of *Methylobacterium*, *Massilia*, *Bacillus*, *Streptomyces*, *Curtobacterium*, and *Arthrobacter* (Vařtilingom et al. 2012).

Among numerous bacterial indicators for ISU was the soil-associated actinobacterial genus *Arthrobacter* (Fig. 7). High abundance of this taxon may be due to aerosolization from agricultural lands surrounding ISU. LSU was strongly indicated by the genus *Cloacibacterium*, a wastewater-associated, gram-negative, non-motile, facultative anaerobe from the family Flavobacteriaceae. VT was strongly indicated by the *Edaphobacter*, a gram-negative obligatively aerobic genus in Acidobacteria.

Strongly significant spring indicator genera included *Massilia* and *Sphingomonas* (Fig. 8). *Massilia* is a major group of aerobic flagellated rods that include rhizosphere and root colonizing bacteria, and has been previously detected in aerosols (Blatny et al. 2011, Fahlgren 2011), including cloud-top samples (Vařtilingom et al. 2012) and the stratosphere (Bryan 2017). *Sphingomonas* was previously found to be dominant in precipitation sampling events in winter (December and November) and spring (April, March, and June) at Puy de Dôme, France (Vařtilingom et al. 2012).

INP concentrations in precipitation were previously reported to be highest during late autumn months and lowest in the summer at our LSU site (Christner et al. 2008; Fig. 1B). Our current work, based on a preponderance of LSU samples, supports this finding. The genus *Pseudomonas* was a significant ISA indicator for autumn in our study, and was at lowest abundances during the summer (Fig. 8). This outcome, however, was driven in part by two unusual events (both LSU autumnal storms), identifiable as isolated points along the far-right side of the first axis of bacterial NMDS ordination (Fig 5a, 6a). Read totals ($\bar{x} \pm se$) for *Pseudomonas* at these events ($43,867 \pm 370$) were much higher than at other events ($2,202 \pm 361$). Indicator values (see Dufrene and Legendre 1998)

for *Pseudomonas* remained higher for august than other seasons with removal of the two events; *Pseudomonas* *p*-values for august, however, increased from 0.01 to 0.07.

Notably, *Pseudomonas* was the most common bacterial genus cultured during repeated sampling of clouds from puy de Dôme, and *P. syringae* was found to be the most commonly isolated species from this genus (Vaïtilingom et al. 2012). *P. syringae* is the most effective naturally-occurring INP known (Maki et al. 1974), and frequently occurs in precipitation events globally (Sands et al. 1982, Amato et al. 2007, Monteil et al. 2014, Du et al. 2017).

Fungi. The dominance of Ascomycota, and particularly Basidiomycota, in our samples (Figs. 9, 10) agrees with previous work on aerosolized fungi, including investigations of cloud and precipitation microbiology (Amato et al. 2007, Vaïtilingom et al. 2012, Huffman et al. 2013). Although non-significant, a weak trend of higher Linnaean fungal diversity during august and summer months (Supplemental Material, Table S53) was evident and is supported by a number of aerosolized microbial studies (e.g., Burch and Levetin 2002).

As noted repeatedly, trends in fungal indicators were site (not season) dependent. Among abundant, strongly significant taxa were ISU fungal indicators including the order Capnodiales, an epiphytic Dothideomycetes, and the class Tremellomycetes, a saprophytic Basidiomycetes. VT was strongly indicated by Microbotryales, a Basidiomycetes that includes plant pathogens (Fig 9).

The high fraction of unclassified fungal taxa in our data is potentially tied to our use of the 18S rRNA gene V9 region over ITS approaches (Lepère et al 2019). While this approach may hamper taxonomic analyses at levels below order, it did not affect OTU-based inferences with respect to our research hypotheses, which were largely concerned with β -partitioning.

Conclusions

Our work clarifies the mechanisms and constraints of bioaerosols in the Earth-atmosphere system by compiling and analyzing the most extensive spatiotemporal dataset of microbial assemblages in precipitation to date. We considered these data using a novel metacommunity framework that predicts dispersal patterns as a function of altitude, based on reported general limitations to distributions of airborne microorganisms. In accordance with our heuristic, we found that bacterial

composition of precipitation was more strongly associated with macroscale-driven storm characteristics compared to fungi assemblages, which were associated with mesoscale drivers, including particular spatial locations and lower altitude characteristics of storms. These patterns were demonstrated using a number of approaches including PERMANOVA, ordination, and β -partitioning. In the later analyses, bacteria had higher contributions to total β -diversity from nestedness due to non-random species losses. Hypothetically, this may be due to the inefficiency of taxa subsidizations to high altitudes following washout events. Notably, bacterial nestedness appeared to be reinforced by general trends within the genus *Pseudomonas*, which contains the highest number of documented INA species and strains. Conversely, total β -diversity of fungi had higher contributions from turnover over time --due to species replacement, hypothetically driven by local environmental filtering.

Acknowledgments

This research was made possible by grants from the Division of Environmental Biology of the National Science Foundation (NSF DEB 1241069, NSF DEB 1241161/1643288). We thank J.S. Lockhart, Q.L. Washburn, S. Zorio, C. Baxter, H. Lavendar, J. Farrar, A. Gilliland, G. Hunt, and J. Tucher for technical and intellectual support, and ISU MRCF for microbial sequencing and seed grants.

Literature Cited

- Aho, K. 2013. Foundational and applied statistics for biologists using R. CRC Press, Boca Raton FL, USA.
- Aho, K. 2018. asbio: A collection of statistical tools for biologists. R package version 1.5. <https://cran.r-project.org/web/packages/asbio/index.html> (Accessed 12/21/2018)
- Aho, K., D. W. Roberts, and T. Weaver. 2008. Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods. *Journal of Vegetation Science* 19: 549-562.
- Amato, P., M. Parazols, M. Sancelme, G. Mailhot, P. Laj, and A. M. Delort. 2007. An important oceanic source of micro-organisms for cloud water at the Puy de Dôme (France). *Atmospheric Environment* 41:8253-8263.

- Amato, P., E. Brisbois, M. Draghi, C. Duchaine, J. Frolich-Nowoisky, J. A. Huffman, G. Mainelis, and M. Thibaudon. 2017. Main biological aerosols, specificities, abundance and diversity. Pages 3-11 *in* Delort, A-M, and P. Amato, editors. Microbiology of Aerosols. Wiley, Hoboken NJ, USA.
- Amato, P., M. Joly, L. Besaury, A. Oudart, N. Taib, A. I. Moné, L. Deguillaume, A.-M. Delort, and D. Debroas. 2017. Active microorganisms thrive among extremely diverse communities in cloud water. PloS one 12: e0182869.
- AMS (American Meteorological Society). 2018. Airmass Classification. http://glossary.ametsoc.org/wiki/Airmass_classification (Accessed 12/18/2018).
- Anagnostou, E. N. 2004. A convective/stratiform precipitation classification algorithm for volume scanning weather radar observations. Meteorological Applications 11: 291-300.
- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26: 32-46.
- Atkinson, J.D., B.J. Murray, and D. O'Sullivan. 2016. Rate of homogenous nucleation of ice in supercooled water. The Journal of Physical Chemistry A 120:6513-6520.
- Atmar, W., and B. D. Patterson. 1993. The measure of order and disorder in the distribution of species in fragmented habitat. Oecologia 96: 373-382.
- Attard, E., H. Yang, A.-M. Delort, P. Amato, U. Pöschl, C. Glaux, T. Koop, and C. E. Morris. 2012. Effects of atmospheric conditions on ice nucleation activity of *Pseudomonas*. Atmospheric Chemistry and Physics 12: 10667-10677.
- Aylor, D. E., M. T. Boehm, and E. J. Shields. 2006. Quantifying aerial concentrations of maize pollen in the atmospheric surface layer using remote-piloted airplanes and Lagrangian stochastic modeling. Journal of Applied Meteorology and Climatology 45: 1003-1015.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19: 134-143.
- Baselga, A., and C. D. L. Orme. 2012. betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution 3: 808-812.
- Baselga, A., and F. Leprieur. 2015. Comparing methods to separate components of beta diversity. Methods in Ecology and Evolution 6: 1069-1079.

- Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57: 289–300.
- Bigg, E. K. 1973. Ice nucleus concentrations in remote areas. *Journal of the Atmospheric Sciences*, 30: 1153-1157.
- Bigg, E. K., and G. T. Miles. 1964. The results of large-scale measurements of natural ice nuclei. *Journal of the Atmospheric Sciences* 21: 396-403.
- Biggerstaff, M. I., and S. A. Listemaa. 2000. An improved scheme for convective/stratiform echo classification using radar reflectivity. *Journal of Applied Meteorology* 39: 2129-2150.
- Blatny, J. M., J. Ho, G. Skogan, E. M. Fykse, T. Aarskaug, and V. Waagen. 2011. Airborne *Legionella* bacteria from pulp waste treatment plant: aerosol particles characterized as aggregates and their potential hazard. *Aerobiologia* 27: 147-162.
- Boon, E., C. J. Meehan, C. Whidden, D. H. J. Wong, M. G. Langille, and R. G. Beiko. 2014. Interactions in the microbiome: communities of organisms and communities of genes. *FEMS Microbiology Reviews*, 38: 90-118.
- Bowers, R. M., S. McLetchie, R. Knight, and N. Fierer. 2011. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *The ISME Journal* 5: 601–612.
- Bowers, R. M., N. Clements, J. B. Emerson, C. Wiedinmyer, M. P. Hannigan, and N. Fierer. 2013. Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere. *Environmental Science and Technology*, 47: 12097-1210
- Bray, J. R. and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* 27: 325-349.
- Brodie, E. L., T. Z. DeSantis, J. P. M. Parker, I. X. Zubietta, Y. M. Piceno, and G. L. Andersen. 2007. Urban aerosols harbor diverse and dynamic bacterial populations. *Proceedings of the National Academy of Sciences* 104: 299-304.
- Brunet, Y., N. Wéry, and A. Galès, A. 2017. Short-Scale Transport of Bioaerosols. Pages 137-154 *in* Delort, A-M and P. Amato, editors. *Microbiology of Aerosols*. Wiley, Hoboken NJ, USA.
- Bryan, N.C. 2017. Microbial distributions and survival in the troposphere and stratosphere.

PhD Thesis, Department of Biological Sciences, Louisiana State University.

Bryan, N.C., B.C. Christner, T.G. Guzik, D.J. Granger, and M.F. Stewart. 2019. Abundance and survival of microbial aerosols in the troposphere and stratosphere. *ISME Journal*, doi: 10.1038/s41396-019-0474-0.

Burch, M., and E. Levetin. 2002. Effects of meteorological conditions on spore plumes. *International Journal of Biometeorology* 46: 107-117.

Burrows, S. M., W. Elbert, M. G. Lawrence, and U. Pöschl. 2009a. Bacteria in the global atmosphere—Part 1: Review and synthesis of literature data for different ecosystems. *Atmospheric Chemistry and Physics* 9: 9263-9280.

Burrows, S. M., T. Butler, P. Jöckel, H. Tost, A. Kerkweg, U. Pöschl, and M. G. Lawrence. 2009b. Bacteria in the global atmosphere—Part 2: Modeling of emissions and transport between different ecosystems. *Atmospheric Chemistry and Physics* 9: 9281-9297.

Cáliz, J., X. Triadó-Margarit, L. Camarero, and E. O. Casamayor. 2018. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceeding of the National. Academy of Science*. 115: 12229-12234.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J. A. Gilbert, G. Smith, and R. Knight. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal* 6: 1621-1624.

Christner, B. C., R. Cai, C. E. Morris, K. S. McCarter, C. M. Foreman, M. L. Skidmore, S. M. Montross, and D. Sands. 2008. Geographic, seasonal, and precipitation chemistry influence on the abundance and activity of biological ice nucleators in rain and snow. *Proceedings of the National Academy of Sciences* 105: 18854-18859.

Constantinidou, H., S. S. Hirano, L. S. Baker, and C. D. Upper. 1990. Atmospheric dispersal of ice nucleation-active bacteria: The role of rain. *Phytopathology* 80: 934-937.

DeLeon-Rodriguez, N., T. L. Lathem, R. L. M. Rodriguez, J. M. Barazesh, B. E. Anderson, A. J. Beyersdorf, L. D. Ziemba, M. Bergin, A. Nenes, and K. T. Konstantinidis. 2013. Microbiome of the upper troposphere: species composition and prevalence, effects of tropical storms, and atmospheric implications. *Proceedings of the National Academy of Sciences*, 110: 2575-2580.

- DeMott, P. J., T. C. J. Hill, C. S. McCluskey, K. A. Prather, D. B. Collins, R. C. Sullivan, M. J. Ruppel, R. H. Mason, V. E. Irish, T. Lee, C. Y. Hwang, T. S. Rhee, J. R. Snider, G. R. McMeeking, S. Dhaniyala, E. R. Lewis, J. J. B. Wentzell, J. Abbatt, C. Lee, C. M. Sultana, A. P. Ault, J. L. Axson, M. Diaz Martinez, I. Venero, G. Santos-Figueroa, M. D. Stokes, G. B. Deane, O. L. Mayol-Bracero, V. H. Grassian, T. H. Bertram, A. K. Bertram, B. F. Moffett, and G. D. Franc. 2016. Sea spray aerosol as a unique source of ice nucleating particles. *Proceedings of the National Academy of Sciences* 113: 5797-5803.
- Després, V., J. A. Huffman, S. M. Burrows, C. Hoose, A. Safatov, G. Buryak, J. Fröhlich-Nowoisky, W. Elbert, M., U. Pöschl, and R. Jaenicke. 2012. Primary biological aerosol particles in the atmosphere: a review. *Tellus B: Chemical and Physical Meteorology* 64: 15598.
- Douda, J., A. Havrdová, J. Doudová., and B. Mandák. 2018. Legacy of post-glacial colonization affects β -diversity: Insights into local community assembly processes. *Journal of Biogeography* doi: 10.1111/jbi.13235
- Dufrene, M., and P. Legendre, P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* 67: 345-366.
- Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince, and R. Knight. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27: 2194-2200.
- Fahlgren, C., G. Bratbak, R. A. Sandaa, R. Thyraug, and U. L. Zweifel. 2011. Diversity of airborne bacteria in samples collected using different devices for aerosol collection. *Aerobiologia* 27: 107-120.
- Failor, K., D. G. Schmale, B. A. Vinatzer, and C. L. Monteil. 2017. Ice nucleation active bacteria in precipitation are unexpectedly diverse and nucleate ice by employing different mechanisms. *The ISME Journal* 11: 1331–1344.
- Fitt, B. D., H. A. McCartney, and P. J. Walklate. 1989. The role of rain in dispersal of pathogen inoculum. *Annual Review of Phytopathology* 27: 241-270.
- Fox J., and S. Weisberg. 2011. *An R companion to applied regression*, second edition. Sage, Thousand Oaks CA, USA.

- Fröhlich-Nowoisky, J., T. C. Hill, B. G. Pummer, P. Yordanova, G. D. Franc, and U. Pöschl. (2015). Ice nucleation activity in the widespread soil fungus *Mortierella alpina*. *Biogeosciences* 12: 1057-1071.
- Fröhlich-Nowoisky, J., C. J. Kampf, B. Weber, J. A. Huffman, C. Pöhlker, M. O. Andreae, Naama Lang-Yona, S. M. Burrows, S. S. Gunthee, W. Elbert, H. Sua, P. Hoor, E. Thines, T. Hoffmann, V. R. Després, and U. Pöschl. 2016. Bioaerosols in the Earth system: Climate, health, and ecosystem interactions. *Atmospheric Research* 182: 346-376.
- Fuzzi, G., P. Mandrioli, and A. Perfetto. 1997. Fog droplets: An atmospheric source of secondary biological aerosol particles. *Atmospheric Environment* 31:287–290.
- Garcia, E. B., R. Hanlon, M. R. Makris, C. W. Powers, C. Jimenez-Sanchez, O. Karatum, L. C. Marr, D. C. Sands, and D. G. Schmale III. 2019. Microbial diversity of individual raindrops collected from simulated and natural precipitation events. *Atmospheric Environment* 209: 102-111.
- Gandolfi, I., V. Bertolini, R. Ambrosini, G. Bestetti, and A. Franzetti. 2013. Unravelling the bacterial diversity in the atmosphere. *Applied Microbiology and Biotechnology* 97: 4727-4736.
- Golan J. J., and A. Pringle. 2017. Long-distance dispersal of fungi. *Microbiology Spectrum* doi: 10.1128/microbiolspec.FUNK-0047-2016.
- Gregory, P. H. 1961. *The microbiology of the atmosphere*. Leonard Hill Books Ltd., London
- Griffin, D. W. 2004. Terrestrial microorganisms at an altitude of 20,000 m in Earth's atmosphere. *Aerobiologia* 20:135-140.
- Griffin, D. W., C. Gonzalez-Martin, C. Hoose, and D. J. Smith. 2017. Global-scale atmospheric dispersion of microorganisms. Page 155-194 *in* Delort, A-M and P. Amato editors. *Microbiology of Aerosols*. Wiley, Hoboken NJ, USA.
- Hanlon, R., C. Powers, K. Failor, C. L. Monteil, B. A. Vinatzer, and D. G. Schmale. 2017. Microbial ice nucleators scavenged from the atmosphere during simulated rain events. *Atmospheric Environment* 163: 182-189.
- Harrison, A. D., T. F. Whale, M. A. Carpenter, M. A. Holden, L. Neve, D. O'Sullivan, J. V. Temprado, and B. J. Murray. 2016. Not all feldspars are equal: a survey of ice nucleating

properties across the feldspar group of minerals. *Atmospheric Chemistry and Physics*, 16: 10927-10940.

Hill, T. C. J., P. J. DeMott, F. Conen, and O. Möhler. 2017. Impacts of bioaerosols on atmospheric ice nucleation processes. Page 195-219 in Delort, A-M and P. Amato editors. *Microbiology of Aerosols*.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6: 65–70.

Houze Jr. R. A. 2014. *Cloud dynamics*, 2nd ed. Elsevier, Oxford, UK.

Huffman, J. A., A. J. Prenni, P. J. DeMott, C. Pöhlker, R. H. Mason, N. H. Robinson, J. Fröhlich-Nowoisky, Y. Tobo, V. R. Després, E. Garcia, D. J. Gochis, E. Harris, I. Müller-Germann, C. Ruzene, B. Schmer, B. Sinha, D. A. Day, M. O. Andreae, J. L. Jimenez, M. Gallagher, S. M. Kreidenweis, A. K. Bertram, and U. Pöschl. 2013. High concentrations of biological aerosol particles and ice nuclei during and after rain. *Atmospheric Chemistry and Physics* 13: 6151.

Hulsman, M., L. Geris, and J. T. Reinder. 2103. Computational analysis of high-throughput material screens. Page 103 in de Boer, J. and C. A. van Blitterswijk, editors. *Materiomics: high-throughput screening of biomaterial properties* Cambridge University Press, Cambridge, UK.

Hummel, M., C. Hoose, B. Pummer, C. Schaupp, J. Fröhlich-Nowoisky, and O. Möhler. 2018. Simulating the influence of primary biological aerosol particles on clouds by heterogeneous ice nucleation. *Atmospheric Chemistry and Physics*, 18: 20.

Joly, M., E. Attard, M. Sancelme L. Deguillaume, C. Guilbaud, C. E. Morris, P. Amato, and A-M Delort. 2013. Ice nucleation activity of bacteria isolated from cloud water. *Atmospheric Environment*, 70: 392–400.

Joung, Y. S., Z. Ge, C. R. Buie. 2017. Bioaerosol generation by raindrops on soil. *Nature Communications* 8: 14668.

Jung, C. H., S. Bae, Y. P. Kim. 2011. Approximated solution on the properties of the scavenging gap during precipitation using harmonic mean method. *Atmospheric Research* 99: 496-504.

Kang, K. S., G. T. Veeder, P. J. Mirrasoul, T. Kaneko, and I. W. Cottrell. 1982. Agar-like

polysaccharide produced by a *Pseudomonas* species: Production and basic properties. *Applied and Environmental Microbiology* 43: 1086-1091.

Kanji, Z. A., L. A. Ladino, H. Wex, Y. Boose, M. Burkert-Kohn, D. J. Cziczo, and M. Krämer, 2017. Overview of ice nucleating particles. *Meteorological Monographs*, 58, 1-1.

Knopf, D. A., P. A. Alpert, and B. Wang. 2018. The role of organic aerosol in atmospheric ice nucleation: a review. *ACS Earth and Space Chemistry* 2: 168-202.

Kourtev, P. S., K. A. Hill, P. B. Shepson, and A. Konopka. 2011. Atmospheric cloud water contains a diverse bacterial community. *Atmospheric Environment*, 45: 5399-5405.

Kruskal, J. B., and M. Wish. 1978. *Multidimensional scaling*. Sage Publications, Beverly Hills, CA, USA.

Kunin, V., A. Engelbrektson, H. Ochman, and P. Hugenholtz. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology* 12: 118-123.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7: 601-613.

Lepère, C., I. Domaizon, J-F Humbert, L. Jardillier, M. Hugoni, and D. Debroas. 2019. Diversity, spatial distribution and activity of fungi in freshwater ecosystems. *PeerJ* 7: e6247.

Lindemann, J., H. A. Constantinidou, W. R. Barchet, and C. D. Upper. 1982. Plants as sources of airborne bacteria, including ice nucleation-active bacteria. *Applied and Environmental Microbiology*, 44: 1059-1063.

Lindemann, J., and C. D. Upper. 1985. Aerial dispersal of epiphytic bacteria over bean plants. *Applied and Environmental Microbiology* 50: 1229-1232.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs and H. Wagner. 2016. *vegan: community ecology package*. R package version 2.3-5. <https://cran.r-project.org/web/packages/vegan/index.html> (Accessed 12/21/2018)

Maki, L. R., E. L. Galyan, M. M. Chang-Chien, and D. R. Caldwell. 1974. Ice nucleation

induced by *Pseudomonas syringae*. Applied microbiology 28: 456-459.

- Mandrioli, P., M. G. Negrini, G. Cesari, and G. Morgan. 1984. Evidence for long range transport of biological and anthropogenic aerosol particles in the atmosphere. Grana 23: 43-53.
- Monteil, C. L., F. Lafolie, J. Laurent, J. C. Clement, R. Simler, Y. Travi, and C. E. Morris. 2014. Soil water flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine headwaters. Environmental Microbiology 16, 2038-2052.
- Morris, C. E., D. C. Sands, C. Glaux, J. Samsatly, S. Asaad, A. R. Moukahel, F. L. T. Goncalves, and E. K. Bigg. 2013. Urediospores of rust fungi are ice nucleation active at $> -10^{\circ}$ C and harbor ice nucleation active bacteria. Atmospheric Chemistry and Physics 13: 4223-4233.
- Morris, C. E., F. Conen, J. Alex Huffman, V. Phillips, U. Pöschl, and D. C. Sands. 2014a. Bioprecipitation: a feedback cycle linking Earth history, ecosystem dynamics and land use through biological ice nucleators in the atmosphere. Global change biology 20: 341-351.
- Morris, C. E., C. Leyronas, and P. C. Nicot. 2014b. Movement of bioaerosols in the atmosphere and the consequences for climate and microbial evolution. Page 393-416 in Colbeck, I, and M. Lazaridis editors. Aerosol Science: Technology and Applications. John Wiley & Sons, Hoboken NJ, USA.
- Murray, B. J., D. O'Sullivan, J. D. Atkinson, and M. E. Webb. 2012. Ice nucleation by particles immersed in supercooled cloud droplets. Chemical Society Reviews 41: 6519-6554.
- Myers, J. A., J. M. Chase, I. Jiménez, P. M. Jørgensen, A. Araujo-Murakami, N. Paniagua-Zambrana, and R. Seidel, R. 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecology letters 16: 151-157.
- NOAA GOES-East. 2018. <https://www.star.nesdis.noaa.gov/GOES/index.php> (Accessed 7/3/2018)
- NWS. 2018. <https://www.weather.gov/lmk/indices> (Accessed 7/3/2018)
- Pratt, K. A., P. J. DeMott, J. R. French, Z. Wang, D. L. Westphal, A. J. Heymsfield, C. H. Twohy, A. J. Prenni, and K. A. Prather. 2009. In situ detection of biological particles in cloud ice-crystals. Nature Geoscience 2: 398-401.
- Prospero, J. M., E. Blades, G. Mathison, R. Naidu. 2005. Interhemispheric transport of

viable fungi and bacteria from Africa to the Caribbean with soil dust. *Aerobiologia* 21: 1-19.

Prosser, J. I., B. J. Bohannon, T. P. Curtis, R. J. Ellis, M. K. Firestone, R. P. Freckleton, J. L. Green, G. E. Green, K. Killham, J. J. Lennon, and A. M. Osborn 2007. The role of ecological theory in microbial ecology. *Nature Reviews Microbiology* 5: 384.

Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* 35: 7188-7196.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. (Accessed 12/20/2018)

Raupach, M., Finnigan, J. J., & Brunet, Y. (1996). Coherent eddies and turbulence in vegetation canopies: the mixing-layer analogy. Page 351-382 in Garratt, J. R., and P. A. Taylor, editors. *Boundary-layer meteorology 25th anniversary volume, 1970–1995*. Springer, Dordrecht, Netherlands.

Reponen, T., S. A. Grinshpun, K. L. Conwell, J. Wiest, M. Anderson. 2001. Aerodynamic versus physical size of spores: Measurement and implication for respiratory deposition. *Grana* 40: 119-125.

Ricklefs, R. E., 1976. *The Economy of Nature: A Textbook in Basic Ecology*. Chiron Press, New York

Roberts, D. 2016. labdsv: ordination and multivariate analysis for ecology. R package version 1.8-0.

Rodríguez-Gironés, M. A., and L. Santamaría. 2006. A new algorithm to calculate the nestedness temperature of presence–absence matrices. *Journal of Biogeography* 33: 924-935.

Sands, D. C., V. E. Langhans, A. L. Scharen, G. de Smet. 1982. The association between bacteria and rain and possible resultant meteorological implications. *Quarterly Journal of the Hungarian Meteorological Service* 86:148–152.

Šantl-Temkiv, T., M. Sahyounb, K. Finster, S. Hartmann, S. Augustin-Bauditz, F. Stratmann, H.

- Wex, T. Clauss, N. Woetmann Nielsen, J. Havskov Sørensen, U. Smith Korsholm, L. Y. Wick, and U. Gosewinkel Karlson. 2015. Characterization of airborne ice-nucleation-active bacteria and bacterial fragments. *Atmospheric Environment* 109: 105-117.
- Sattler, B., H. Puxbaum, and R. Psenner. 2001. Bacterial growth in supercooled cloud droplets. *Geophysical Research Letters* 28: 239-242.
- Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75: 7537-7541.
- Schmale III, D. G., and S. D. Ross. 2015. Highways in the sky: scales of atmospheric transport of plant pathogens. *Annual review of phytopathology*, 53: 591-611.
- Shannon, C. E. 1948 A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423 and 623–656.
- Singh, A. K., P. K. Pindi, S. Dube, V. R. Sundareswaran, and S. Shivaji. 2009. Importance of trmE for growth of the psychrophile *Pseudomonas syringae* at low temperatures. *Applied and Environmental Microbiology* 75: 4419-4426.
- Smith, D. P., and K. G. Peay. 2014. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS One* 9: e90234.
- Stekhoven, D. J., and P. Bühlmann. 2011. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28: 112-118.
- Stekhoven, D. J. 2013. MissForest: non-parametric missing value imputation for mixed-type data. R package version 1.4.
<https://cran.r-project.org/web/packages/missForest/index.html> (Accessed 12/21/2018)
- Stopelli, E., F. Conen, C. E. Morris, E. Herrmann, N. Bukowiecki, and C. Alewell. 2015. Ice nucleation active particles are efficiently lost from precipitating clouds. *Scientific Reports* 5:16433.
- Striluk, M. L., K. Aho, and C. F. Weber, C. F. 2016. The effect of season and terrestrial biome

on the abundance of plant growth-promoting bacteria in the lower atmosphere. *Aerobiologia* 33: 137-149.

Stroud, J. T., M. R. Bush, M. C. Ladd, R. J. Nowicki, A. A. Shantz, and J. Sweatman. 2015. Is a community still a community? Reviewing definitions of key terms in community ecology. *Ecology and evolution* 5: 4757-4765.

Suski, K. J., T. C. Hill, E. J. Levin, A. Miller, P. J. DeMott, and S. M. Kreidenweis. 2018. Agricultural harvesting emissions of ice-nucleating particles. *Atmospheric Chemistry and Physics* 18: 13755-13771.

Szyrmer, W., and I. Zawadzki. 1997. Biogenic and anthropogenic sources of ice-forming nuclei: a review. *Bulletin of the American Meteorological Society*, 78: 209-228.

Taha, M. P. M., G. H. Drew, P. J. Longhurst, R. Smith, and S. J. Pollard. 2006. Bioaerosol releases from compost facilities: Evaluating passive and active source terms at a green waste facility for improved risk assessments. *Atmospheric Environment*, 40: 1159-1169.

Thompson, L. R., J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons, G. Ackermann, J. A. Navas-Molina, S. Janssen, E. Kopylova, Y. Vázquez-Baeza¹, A. González, J. T. Morton, S. Mirarab, Z. Z. Xu, L. Jiang, M. F. Haroon, J. Kanbar, Q. Zhu, S. J. Song, T. Kosciolk, N. A. Bokulich, J. Lefler, C. J. Brislawn, G. Humphrey, S. M. Owens, J. Hampton-Marcell, D. Berg-Lyons, V. McKenzie, N. Fierer, J. A. Fuhrman, A. Clauset, R. L. Stevens, A. Shade, K. S. Pollard, K. D. Goodwin, J. K. Jansson, J. A. Gilbert, R. Knight, and the Earth Microbiome Project Consortium. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, 551: 457–463

Tobo, Y., A. J. Prenni, P. J. DeMott, J. A. Huffman, C. S. McCluskey, G. Tian, C. Pöhlker, U. Pöschl, and S. M. Kreidenweis. 2013. Biological aerosol particles as a key determinant of ice nuclei populations in a forest ecosystem. *Journal of Geophysical Research: Atmospheres* 118: 100-110.

Tukey, J. 1949. Comparing individual means in the analysis of variance. *Biometrics* 5: 99–114.

Vätilingom, M., P. Amato, M. Sancelme, P. Laj, M. Leriche, and A. M. Delort. 2010.

Contribution of microbial activity to carbon chemistry in clouds. *Applied and Environmental Microbiology* 76: 23-29.

Vařilingom, M., E. Attard, N. Gaiani, M. Sancelme, L. Deguillaume, A. I. Flossmann, P. Amato, and A.-M. Delort. 2012. Long-term features of cloud microbiology at the puy de Dôme (France). *Atmospheric Environment* 56: 88-100.

von Engel, A., and J. Teixeira. 2013. A planetary boundary layer height climatology derived from ECMWF reanalysis data. *Journal of Climate* 26: 6575-6590.

Weber, C.F. 2015. *Polytrichum commune* spores nucleate ice and associated microorganisms increase the temperature of ice nucleation activity onset. *Aerobiologia* 32: 353-361.

Wilkinson, D. M., S. Koumoutsaris, E. A. Mitchell, and I. Bey. 2012. Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography* 39: 89-97.

Wilson, T. W., L. A. Ladino, P. A. Alpert, M. N. Breckels, I. H. Mason, G. McFiggans, L. A. Miller, J. J. Nájera, E. Polishchuk, S. Rae, C. L. Schiller, M. Si, J. V. Temprado, T. F. Whale, J. P. S. Wong, O. Wurl, J. D. Yakobi-Hancock, Jonathan P. D. Abbatt, J. Y. Aller, A. K. Bertram, D. A. Knopf, and B. J. Murray. 2015. A marine biogenic source of atmospheric ice-nucleating particles. *Nature* 525: 234-238.

Winegardner, A. K., P. Legendre, B. E. Beisner, and I Gregory-Eaves, I. 2017. Diatom diversity patterns over the past c. 150 years across the conterminous United States of America: Identifying mechanisms behind beta diversity. *Global ecology and biogeography* 26: 1303-1315.

Yamamoto, N., K. Bibby, J. Qian, D. Hospodsky, H. Rismani-Yazdi, W. W. Nazaroff, and J. Peccia. 2012. Particle-size distributions and seasonal diversity of allergenic and pathogenic fungi in outdoor air. *The ISME Journal* 6: 1801.

DATA AVAILABILITY

Data are available from the Northwest Knowledge Network: <https://doi.org/10.7923/6t5b-y124>

Tables

Table 1. Sample numbers for 72 precipitation events: 71 events sampled for 16S rRNA gene libraries and 30 events sampled for 18S rRNA gene libraries. The first number in each cell defines the number of 16S samples and the second number gives the number of 18S samples.

Site	Spring	Summer	Winter	Autumn	Total
ISU	2;2	1;1	2;1	0;0	5;4
LSU	5;2	15;7	12;5	15;6	47;20
VT ^a	4;1	0;0	10;4	5;1	19;6
Total	11;5	16;8	24;10	20;7	71;30

^aVT includes VT-KF-A, VT-KF-B, and VT-KF-C

Table 2. OTU richness and diversity summaries for 16S ($n = 71$) and 18S ($n = 30$) precipitation event data by site and season. Entries are means \pm SEM. Rows within site and season sections provide comparisons of levels in site and season. Entries within row with the same letter within the same section are not significantly different after controlling for FWE among site or season levels using Tukey's HSD or Bonferroni adjustment of confidence intervals (for β -partitioning tests). Rows within sections without letters indicate that site or season effects were not significant in the corresponding omnibus test. Formal β -partitioning comparisons of 16S and 18S were made using paired precipitation samples ($n = 29$) in which both 16S and 18S rRNA amplifications could be made and are shown in Fig. 4.

	Site			Season			
	ISU	LSU	VT	Autumn	Spring	Summer	Winter
16S							
Rich.	2185 \pm 531	1849 \pm 135	2001 \pm 176	1844 \pm 209	1760 \pm 338	1755 \pm 148	2141 \pm 192
SW-div.	5.3 \pm 0.3	4.5 \pm 0.2	4.9 \pm 0.3	4.5 \pm 0.3	4.3 \pm 0.4	4.6 \pm 0.1	5 \pm 0.2
β_{turn}	0.46 \pm 0.07 B	0.58 \pm 0.07 A	0.56 \pm 0.1 AB	0.6 \pm 0.02	0.62 \pm 0.03	0.58 \pm 0.03	0.6 \pm 0.02
β_{nest}	0.21 \pm 0.01	0.14 \pm 0.02	0.11 \pm 0.02	0.15 \pm 0.03	0.15 \pm 0.02	0.1 \pm 0.04	0.11 \pm 0.03
β_{Total}	0.67 \pm 0.02 B	0.72 \pm 0.03 A	0.68 \pm 0.02 AB	0.74 \pm 0.02	0.77 \pm 0.02	0.67 \pm 0.03	0.71 \pm 0.02
18S							
Rich.	1215 \pm 25.5	1442 \pm 110.8	1243.5 \pm 165.8	1387 \pm 262	1374 \pm 159	1424 \pm 151	1319 \pm 110
SW-div.	4 \pm 0.2	3.8 \pm 0.2	3.9 \pm 0.3	3.8 \pm 0.4	3.9 \pm 0.3	3.9 \pm 0.2	3.8 \pm 0.2
β_{turn}	0.64 \pm 0.12	0.7 \pm 0.0	0.67 \pm 0.12	0.71 \pm 0.02	0.74 \pm 0.02	0.66 \pm 0.02	0.74 \pm 0.02
β_{nest}	0.01 \pm 0.02 B	0.06 \pm 0.01 A	0.05 \pm 0.02 AB	0.08 \pm 0.06	0.03 \pm 0.02	0.07 \pm 0.06	0.04 \pm 0.06
β_{Total}	0.65 \pm 0.07 B	0.76 \pm 0.02 A	0.73 \pm 0.07 AB	0.79 \pm 0.09	0.78 \pm 0.02	0.73 \pm 0.09	0.78 \pm 0.09

Table 3. PERMANOVA results for 16S and 18S rRNA gene raw reads. Both omnibus (left side of table) and pairwise comparisons (right side of table) are shown. For omnibus tests, significant results are bolded. For pairwise tests, levels sharing a lettered exponent are not significantly different after controlling FWE using Holm's sequential Bonferroni procedure. Two imputed data points were required to address empty cells in PERMANOVAs (see Methods).

Omnibus PERMANOVA					Pairwise comparisons
	Df	<i>F</i>	<i>R</i> ²	<i>p</i> -val	
16S					
Site	2	6.32	0.14	0.0001	ISU ^A , LSU ^B , VT ^C Autumn ^A , Spring ^B , Summer ^C , Winter ^D
Season	3	2.65	0.09	0.0001	
Site × Season	6	1.18	0.08	0.1049	
Error	61				
18S					
Site	2	3.10	0.17	0.0002	ISU ^A , LSU ^B , VT ^C Autumn, Spring, Summer, Winter
Season	3	1.46	0.12	0.0572	
Site × Season	6	0.91	0.15	0.6932	
Error	20				

Table 4. Results of vector (quantitative predictors) and factor fits (categorical predictors) of NMDS ordinations of 16S (Figs 5a, 6a) and 18S (Figs 5b, 6b) datasets. Statistically significant results are bolded.

	16S		18S	
	R^2	p -value	R^2	p -value
Vectors				
dist. traveled 100m	0.012	0.636	0.095	0.257
dist. traveled 1000m	0.062	0.103	0.111	0.224
dist. traveled 5000m	0.154	0.016	0.035	0.627
dist. traveled 6000m	0.234	0.001	0.084	0.318
dist. traveled 8000m	0.218	< 0.001	0.109	0.219
dist. traveled 9000m	0.193	0.001	0.115	0.204
Cloud top temp.	0.018	0.546	0.008	0.915
Factors				
Site	0.130	0.154	0.278	0.002
Season	0.109	0.025	0.170	0.125
Air mass 100m	0.201	0.030	0.191	0.063
Air mass 1000m	0.039	0.590	0.251	0.028
Air mass 5000m	0.025	0.849	0.183	0.217
Air mass 6000m	0.071	0.277	0.345	0.003
Air mass 8000m	0.061	0.301	0.217	0.114
Air mass 9000m	0.072	0.181	0.245	0.059
Weather morphology	0.142	0.075	0.166	0.511
Convective or Stratiform	0.031	0.105	0.053	0.206
Precip. type (rain or ice)	0.021	0.175	0.152	0.007

Table 5. Summary of indicator species analyses (ISA). Shown for bacterial and fungal reads, for both site and season, are the number of taxa considered in ISA (those occurring at ≥ 2 events) and the numbers and proportions of significant ($\alpha = 0.05$) indicator taxa. Significant results with and without adjustment for FDR are given.

	Phylum	Class	Order	Family	Genus
Site					
Bacteria					
No. taxa considered	39	91	238	417	1153
No. significant non-FDR-adjusted taxa	7	18	43	91	247
No. significant FDR-adjusted taxa	3	0	7	27	55
Prop. significant non-FDR-adjusted taxa	0.18	0.20	0.18	0.22	0.21
Prop. significant FDR-adjusted taxa	0.08	0.00	0.03	0.06	0.05
Fungi					
No. taxa considered	5	26	65		
No. significant non-FDR-adjusted taxa	1	4	19		
No. significant FDR-adjusted taxa	0	0	4		
Prop. significant non-FDR-adjusted taxa	0.20	0.15	0.29		
Prop. significant FDR-adjusted taxa	0.00	0.00	0.06		
Season					
Bacteria					
No. taxa considered	39	91	238	417	1153
No. significant non-FDR-adjusted taxa	7	19	40	69	160
No. significant FDR-adjusted taxa	0	1	5	12	22
Prop. significant non-FDR-adjusted taxa	0.18	0.21	0.17	0.17	0.14
Prop. significant FDR-adjusted taxa	0.00	0.01	0.02	0.03	0.02
Fungi					
No. taxa considered	5	26	65		
No. significant non-FDR-adjusted taxa	0	3	9		
No. significant FDR-adjusted taxa	0	0	1		
Prop. significant non-FDR-adjusted taxa	0.00	0.12	0.14		
Prop. significant FDR-adjusted taxa	0.00	0.00	0.02		

Figure captions

Figure 1. Scenarios in which total β -diversity, is driven by (A) nestedness, (B) turnover, and (C) a combination of nestedness and turnover. Figure follows Baselga (2010). Measures of β_{turn} , β_{nest} and β_{Total} in the figure are computed using the incidence-based pairwise metrics proposed by Baselga (2010) and utilized by Baselga and Orme (2012).

Figure 2. A metacommunity framework for the origin, mixing, transport and deposition of aerosolized and precipitation-borne microorganisms. Local microbial assemblages are distinguished with hatching, and colored arrows show subsidization fluxes of microbial assemblages. PBL = Planetary Boundary Layer, SBL = Surface Boundary Layer.

Figure 3. Radar charts showing, for 72h back trajectories, proportions of occurrences of precipitation-event air mass origins at two altitudes, and weather morphologies, by site and season. For origins: mP1 = Pacific maritime polar, mT1 = Pacific maritime tropical, mT2 = Atlantic maritime tropical, cP = continental polar and cT = continental tropical. Note that an additional origin class occurs at 1000m.

Figure 4. Comparisons of bacterial and fungal beta-partitioning ratios for paired precipitation samples ($n = 29$) for which both 16S and 18S rRNA gene data were available. Bars are means and error bars are bootstrap standard errors ($R = 500$). Triple asterisks indicate significant differences between 16S and 18S ratios based on percentile bootstrap confidence intervals ($R = 1000$) at $\alpha = 0.001$.

Figure 5. Two-dimensional NMDS ordinations of 72 precipitation events; (a) 71 precipitation events based on bacterial (16S) OTUs (stress = 0.08) and (b) 30 precipitation events based on (18S) fungal OTUs (stress = 0.159). The figures above are intended to emphasize patterns among and within sites. Also shown are examples of significant results for vector fitting of quantitative data (blue arrows) and factor fitting of categorical data (grayscale text). Vector fits depict relationships between distances traveled of precipitation events at four altitudes (5000, 6000, 8000, and 9000m) and the NMDS projection. Longer arrows indicate stronger correlations and arrows point in the direction of most rapid increase in distance traveled. Factor fits are depicted as bivariate centroids of (in this case) levels of air mass types of precipitation events at 100m elevation: mP1 = Pacific maritime polar, mT2 = Atlantic maritime tropical, cP = continental polar and cT = continental tropical. Ellipses shown with dashed lines are 95% confidence envelopes for the true factor level bivariate centroids.

Figure 6. Identical NMDS projections to Fig. 5, but with seasons emphasized. See Fig. 5 caption for explanation of plot components.

Figure 7. Most abundant taxa in 16S rRNA gene libraries across sites. Specifically, the bars show the six most abundant (high read) phyla, and the 15 most abundant classes, orders, families and genera. Also shown is an “other” category, comprising all other named taxa, and “unclassified” and “uncultured” IDs. Asterisks (*) denote a significant ISA taxon and are accompanied by bold font entries in the legend. Double asterisks (**) denote a significant ISA taxon after controlling for FDR and are accompanied by bold text and underlined entries in the legend.

Figure 8. Most abundant taxa in 16S rRNA gene libraries across seasons. Specifically, the bars show the six most abundant (high read) phyla, and the 15 most abundant classes, orders, families and genera. See Fig. 7 caption for additional details.

Figure 9. Most abundant taxa in 18S rRNA gene libraries across sites. Specifically, the bars show the four most abundant phyla, and the 15 most abundant classes and orders. Also shown is an “other” category, comprising all other named taxa. See Fig. 7 caption for additional details.

Figure 10. Most abundant taxa in 18S rRNA gene libraries across seasons. Specifically, the bars show the four most abundant phyla, and the 15 most abundant classes and orders. Also shown is an “other” category, comprising all other named taxa. See Fig. 7 caption for additional details.

Figure 11. Matrix temperature (nestedness) plots (Atmar and Patterson 1993) summarizing: (a) all 34107 16S OTUs, and (b) 203 *Pseudomonas* OTUs at 71 precipitation events. Blackened cells in matrices indicate the occurrence of an OTU at an event. Matrix rows (precipitation events) and columns (OTUs) are arranged to maximize nestedness by maximizing the number of occurrences above the red curve. Events with fewer *Pseudomonas* OTUs were generally compositional subsets of the richer events, resulting in relatively high levels of event nestedness compared to OTUs from all genera.

Figures

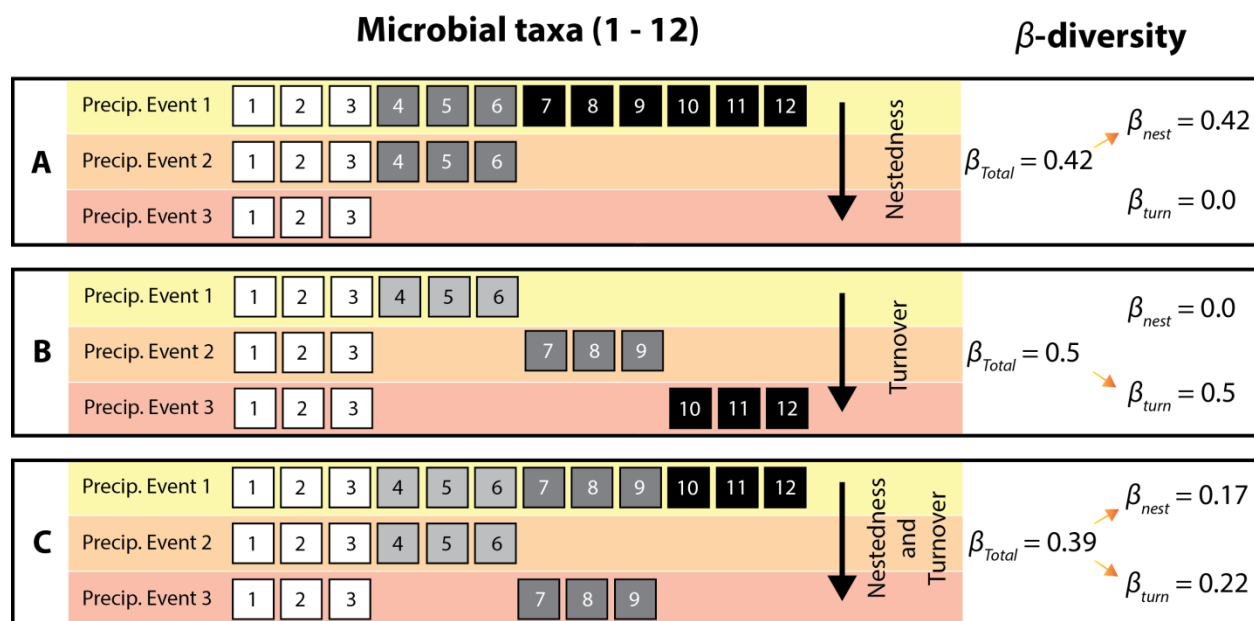


Figure 1

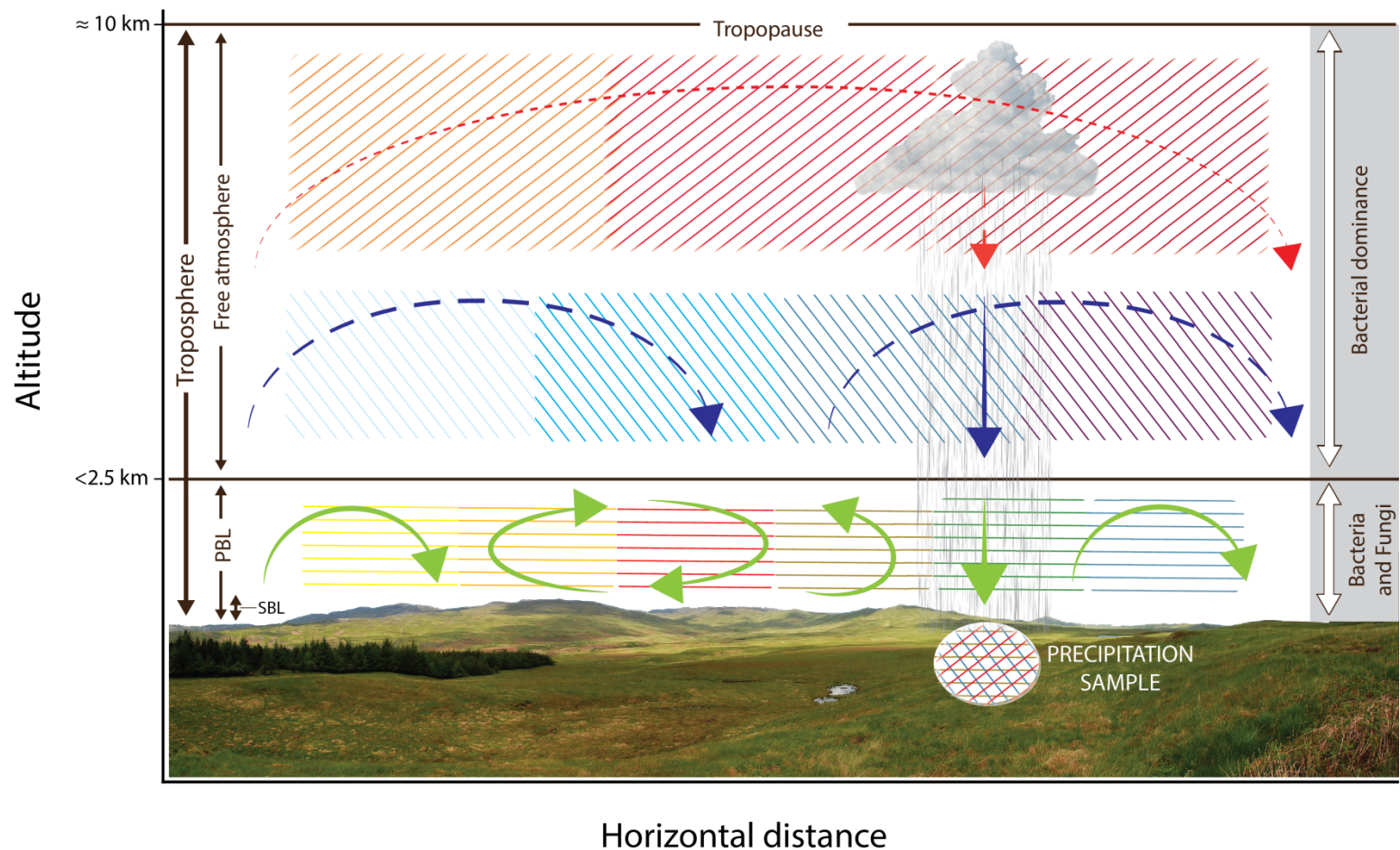


Figure 2

This article is protected by copyright. All rights reserved

Aho, K. (Auteur de correspondance), Weber, Christner, B., Vinatzer, B., Morris, C. E., Joyce, Failor, K., Werth, BaylessEdwards, Schmale, D. (2020). Spatiotemporal patterns of microbial composition and diversity in precipitation. *Ecological Monographs*, 90 (1), e01394. , DOI : 10.1002/ecm.1394

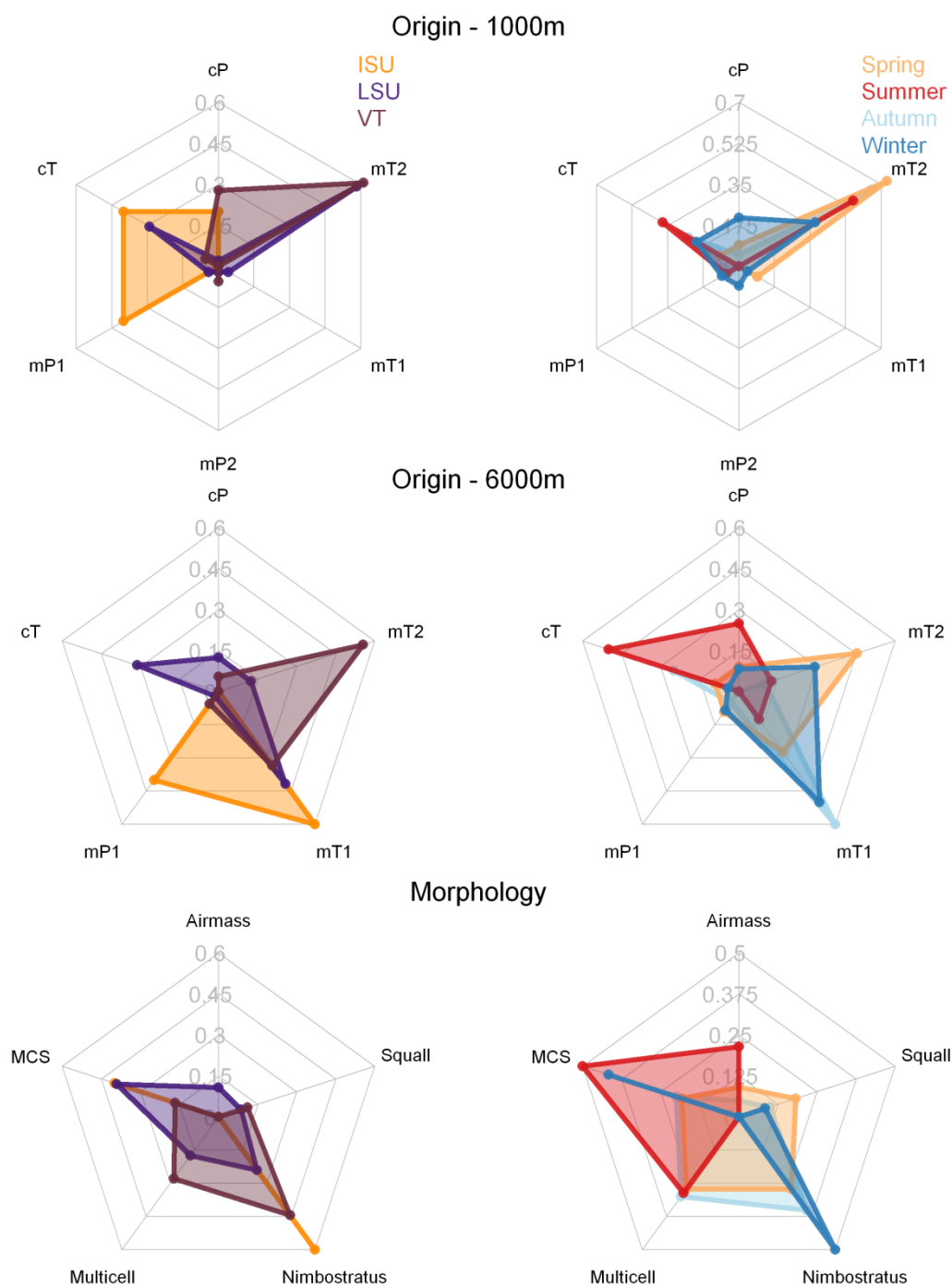


Figure 3

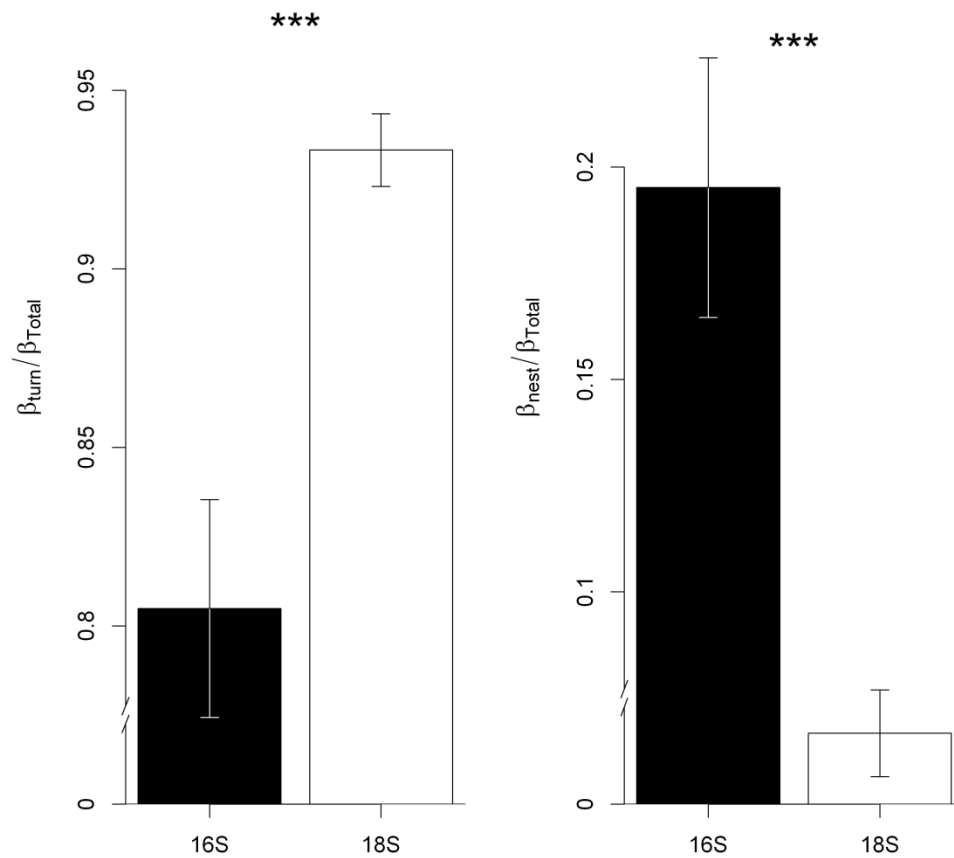


Figure 4

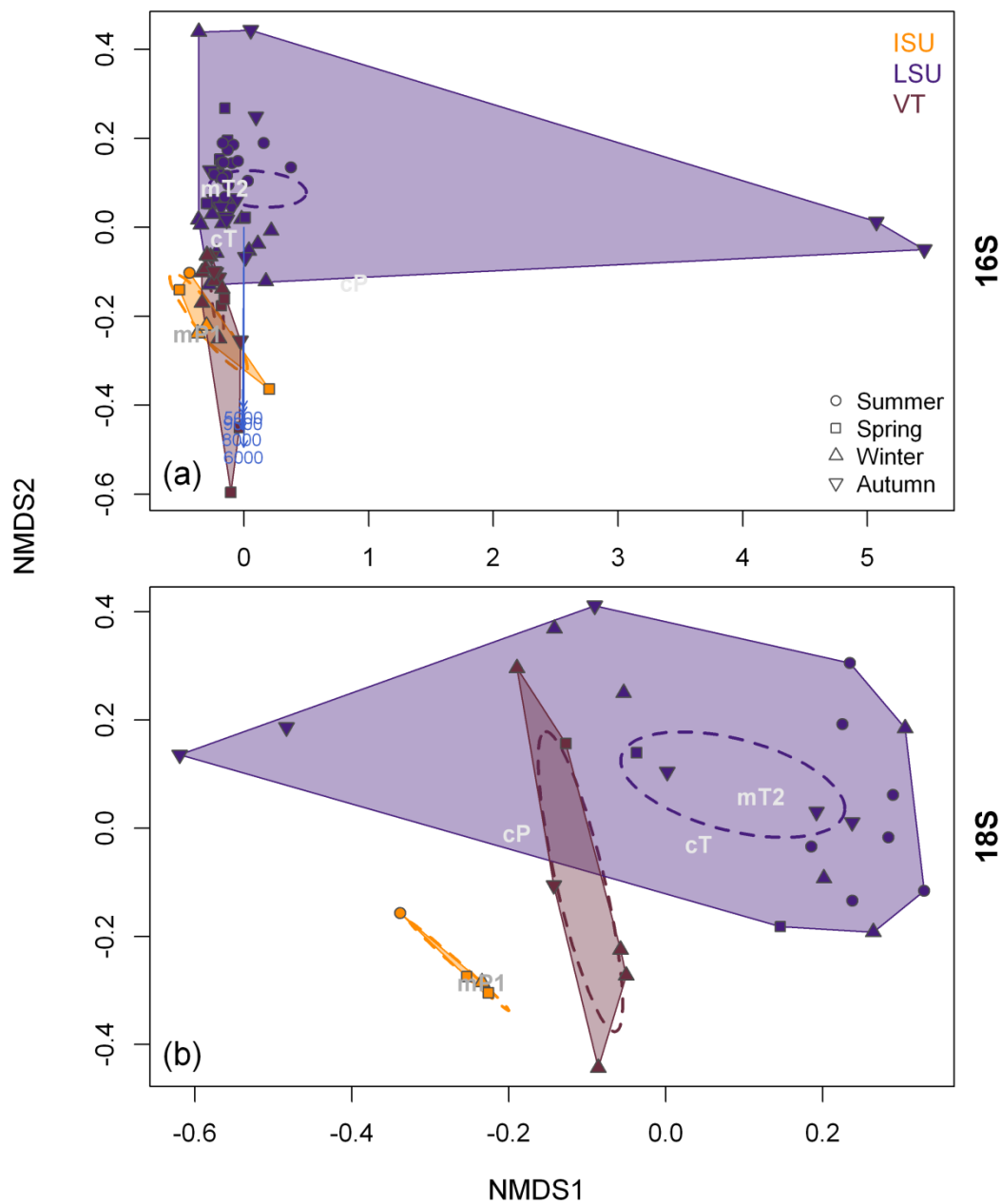


Figure 5

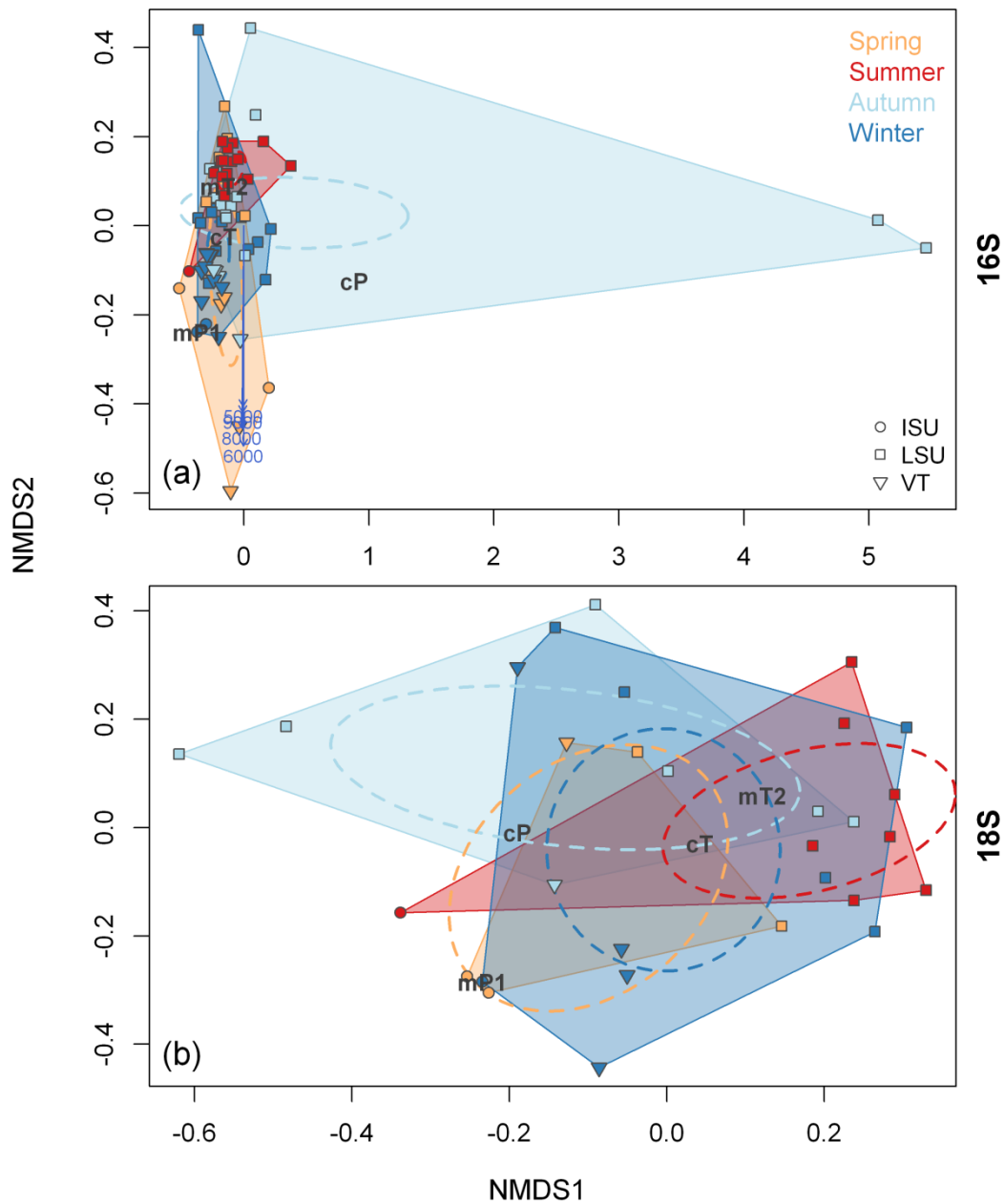


Figure 6

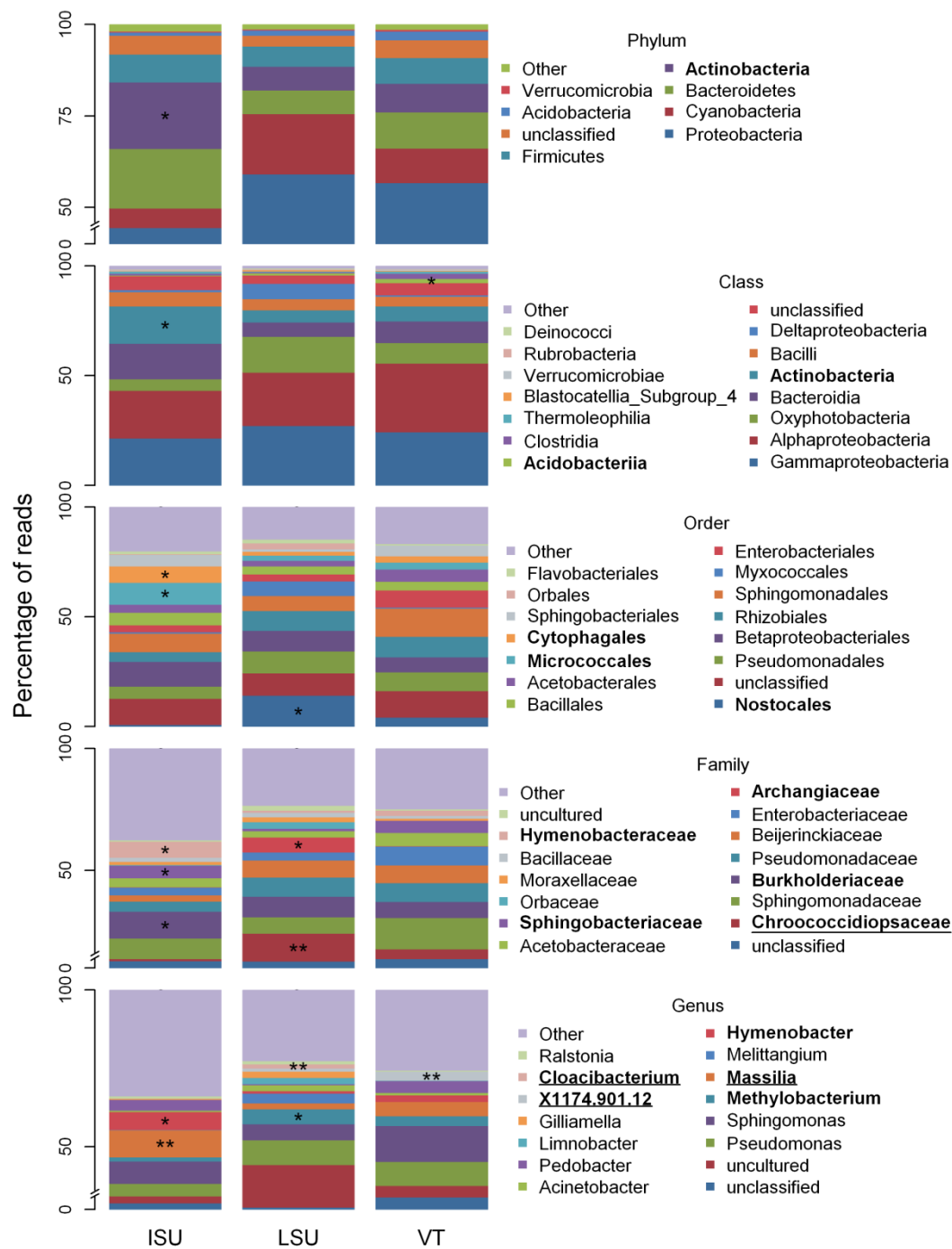


Figure 7

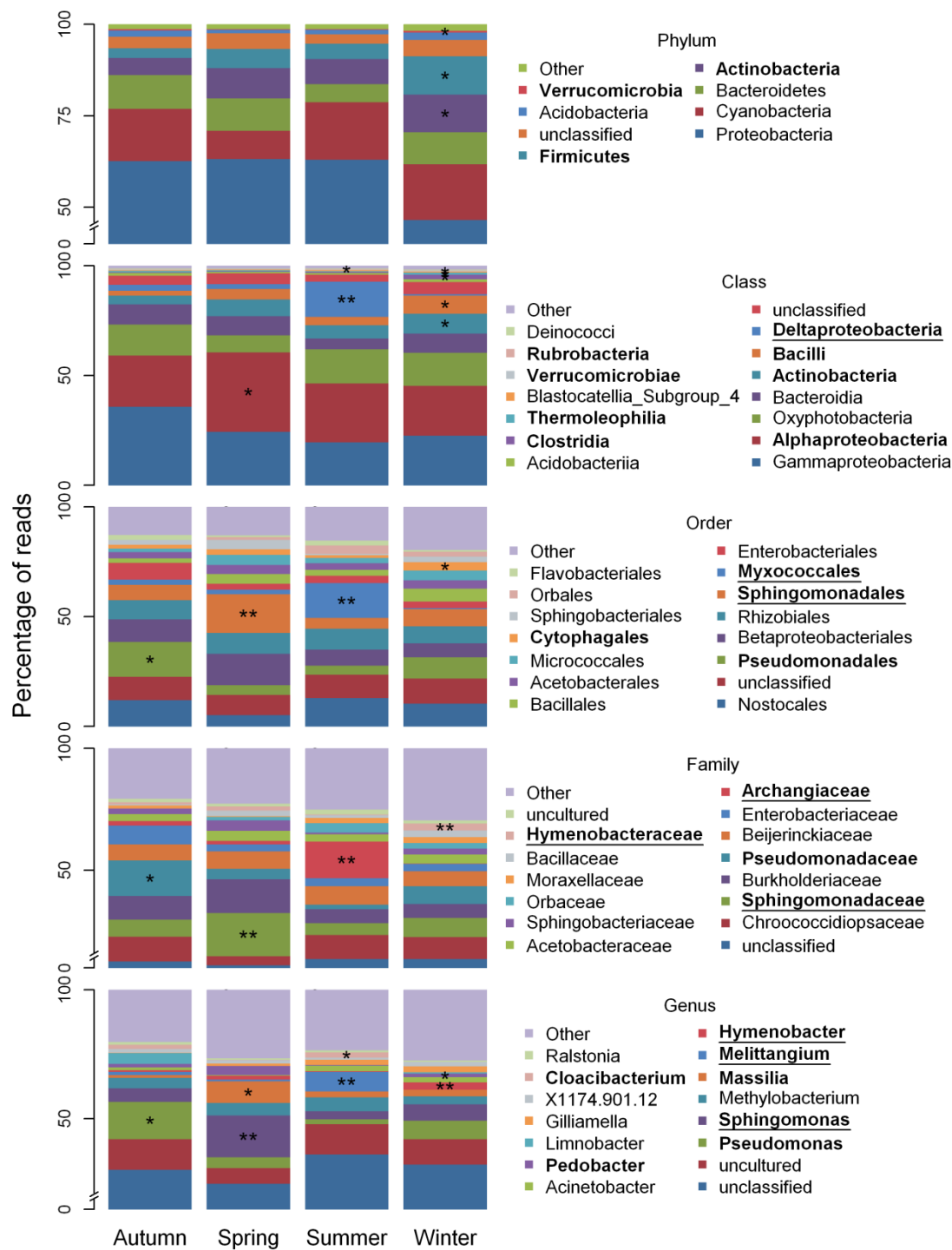


Figure 8

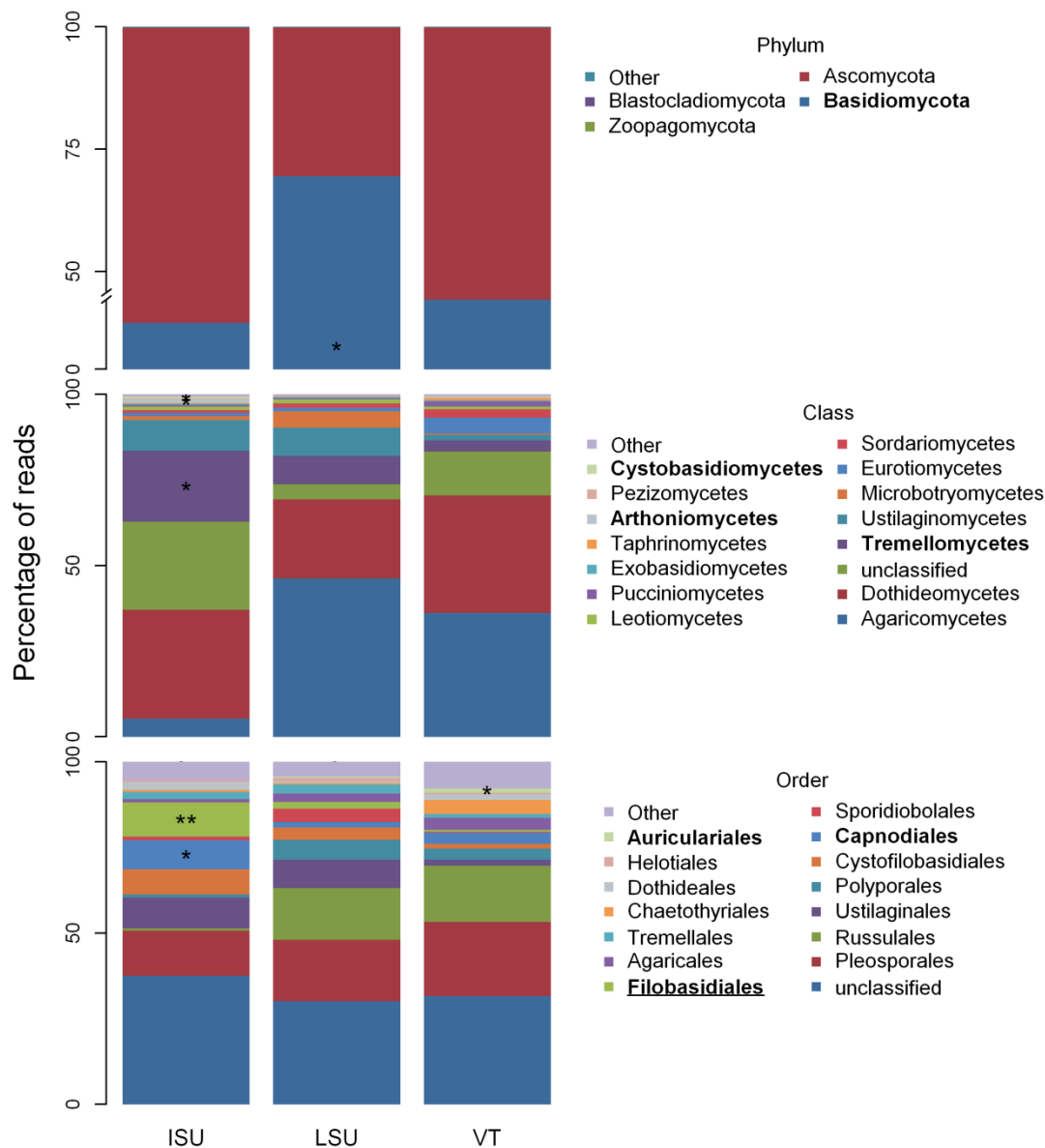


Figure 9

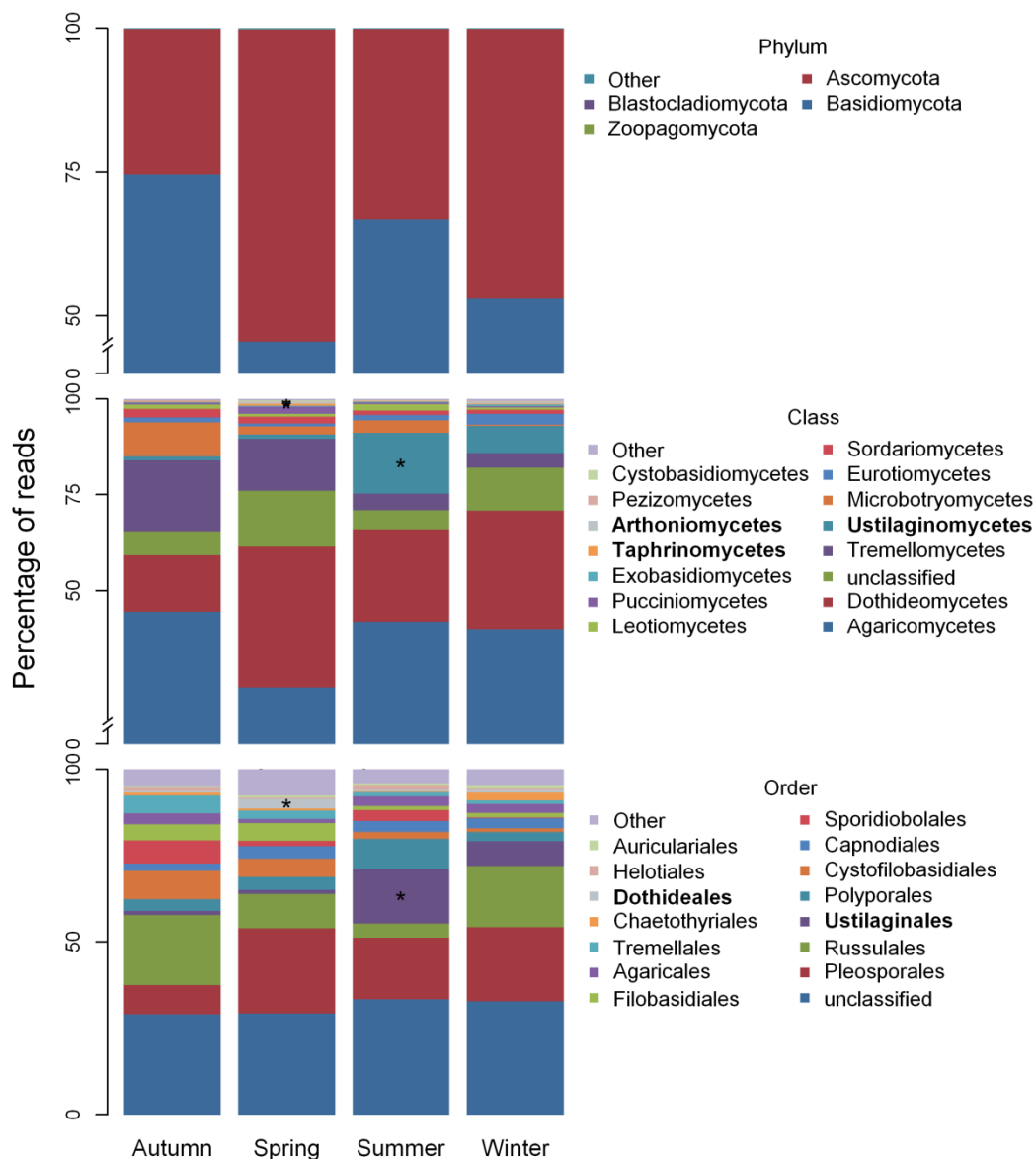


Figure 10

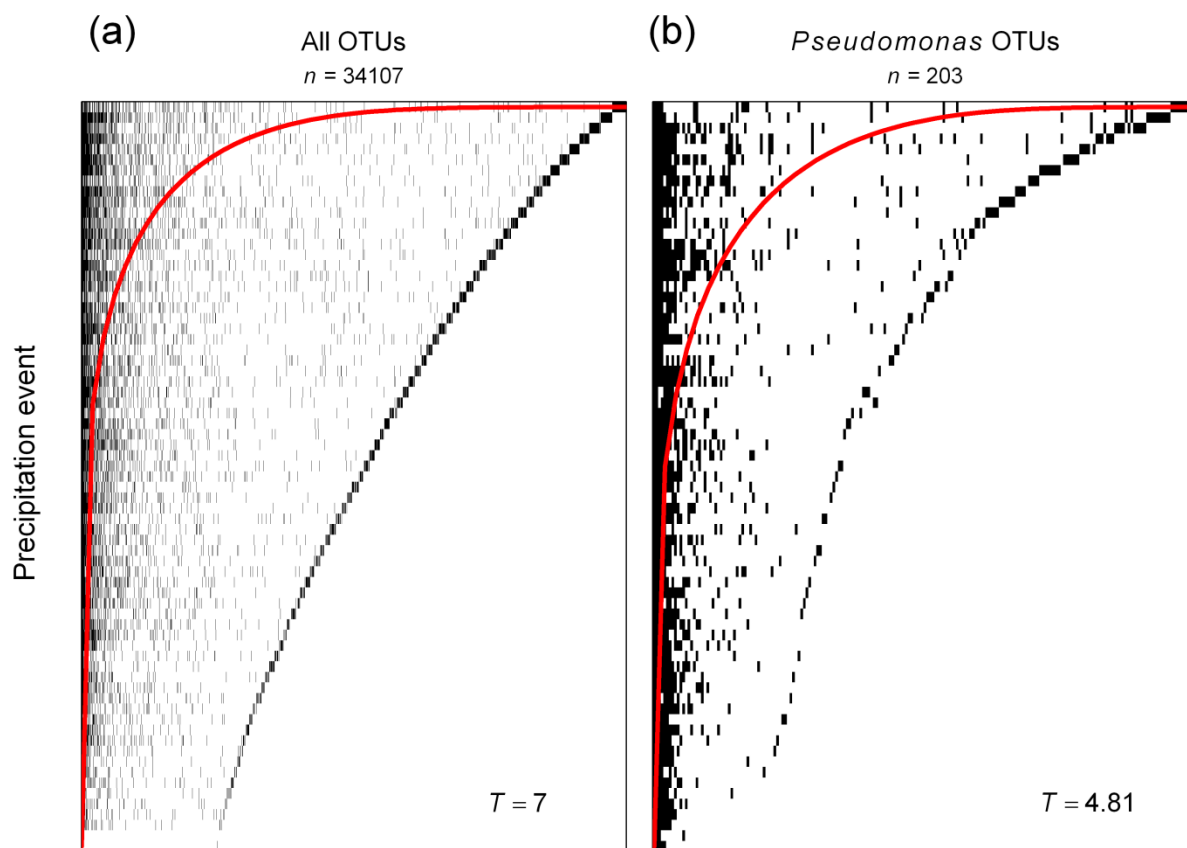


Figure 11