Spatial scaling of forest soil microbial communities across a temperature gradient

Ye Deng[1,2,3,10*], Daliang Ning[3], Yujia Qin[3], Kai Xue[3,10], Liyou Wu[3], Zhili He[3,9], Huaqun Yin[3,4], Yuting Liang[3,5], Vanessa Buzzard[6], Sean T. Michaletz[6], and Jizhong Zhou[3,7,8,*]

[1]Institute for Marine Science and Technology, Shandong University, Qingdao 266237, China

[2]CAS Key Laboratory of Environmental Biotechnology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

[3]Institute for Environmental Genomics, Department of Microbiology and Plant Biology, and School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK 73019, USA

[4]School of Minerals Processing and Bioengineering, Central South University, Changsha 410083, China

[5]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

[6]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721

[7]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, 94270, USA

[8]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China

[9]School of Environmental Science and Engineering, Sun Yat-Sen University, Guangzhou, 510006, China

[10]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

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: 10.1111/1462-2920.14303

Running title: temperature altered microbial spatial scaling

Key words: Microbial community; Spatial scaling; Temperature gradient; Taxa-area relationship; *z*-value; Forest soils; Biodiversity

***Corresponding authors: Jizhong Zhou**

101 David L Boren Blvd, Norman, OK, 73071

Phone: 1-(405)325-6073

Fax: 1-(405)325-7552

E-mail: jzhou@ou.edu

Ye Deng

18 Shuangqing Rd, Haidian, Beijing, China, 100085

Phone: 86-(10)6284-0082

Fax: 86-(10) 6284-0082

E-mail: yedeng@rcees.ac.cn

Originality-Significance Statement

Taxa-area relationship (TAR) is one of the few fundamental laws in ecology and it is critical to our understanding of the distribution of global biodiversity; but it is still debatable whether it is universally applicable to all organisms from all domains of life and how it is affected by environmental changes such as climate warming. Using deep sequencing, our results revealed strong temperature-dependent continental scale TARs in the forest soil microbial communities with typical turnover rates (*z* values) less than 0.1, supporting the general claim that the TARs are universal and microbial turnover in space is lower than plants and animals. Such findings have important implications for predicting how biodiversity changes with temperature, especially under future climate warming scenarios.

Summary

Temperature is an important correlate of global patterns of biodiversity, yet the mechanisms driving these relationships are not well understood. Taxa-area relationships (TARs) have been intensively examined, but the effects of temperature on TARs, particularly for microbial communities, are largely undocumented. Here we present a continental-scale description of temperature-dependent nested TARs of microbial communities (bacteria and archaea) from soils of six forest sites spanning a temperature gradient from subalpine Colorado to tropical Panama. Our results revealed that spatial scaling rates (*z*-values) of microbial communities varied with both taxonomic resolutions and phylogenetic groups. Additionally, microbial TAR *z*-values increased with temperature ($r = 0.739$, $P < 0.05$), but were not correlated with other environmental variables tested ($P > 0.05$), indicating that microbial spatial scaling rate is temperature-dependent. Understanding how temperature affects the spatial scaling of microbial biodiversity is of fundamental importance for preservation of soil biodiversity and management of ecosystems.

Introduction

Temperature is one of the fundemental factors influencing the distribution of biodiversity on earth (Rohde, 1992; Allen *et al.*, 2002; Brown *et al.*, 2004). The metabolic theory of ecology (MTE) hypothesizes that temperature influences the rates of metabolic and biochemical processes following an Boltzmann-Arrhenius relationship, so that rates of mutations and speciation are predicted to increase exponentially with temperature (Brown, 2014). As a result, the richess of species number would be elevated with the increase of temperature directly (Brown, 2014). Indirectly, high temperature enhances primary productivity in most ecosystems that improves the population persistence and provides more opportunity for species coexistence (Hurlbert and Stegen, 2014). For microbes, several recent studies have showed that temperature influenced microbial biodiversity as quantified by species richness and/or α-diversity indexes (Garcia-Pichel *et al.*, 2013; Zhou *et al.*, 2016). However, much less is known about how patterns of spatial scaling (*i.e.* β diversity) vary across broad temperature gradients.

The taxa-area relationship (TAR), the general increase of species richness with area, is the most commonly used way to assess spatial scaling of biodiversity. It has been called a fundamental law of ecology (Lawton, 1999; Horner-Devine *et al.*, 2004; Woodcock *et al.*, 2006; Zhou *et al.*, 2008), and has became a important tool for assessing total species diversity (Harte *et al.*, 2009; Storch *et al.*, 2012), extinction rates (He and Hubbell, 2011), and the size of biodiversity hotspots (Guilhaumon *et al.*, 2008). However, despite more than two centuries of research (Lomolino, 2001; Harte *et al.*, 2009; Franzen *et al.*, 2012), TARs have been a subject of contention, particularly with respect to their universality across taxa and regions (Logue *et al.*, 2012; Storch *et al.*, 2012), their shape across geographical scales (Dengler, 2008; Guilhaumon *et al.*, 2008; Storch *et al.*, 2012), the variation in spatial scaling rates (*z*-values) (Horner-Devine *et*

al., 2004; Woodcock *et al.*, 2006; Zhou *et al.*, 2008), how they are influenced by biotic and abiotic variables (Ryberg and Chase, 2007; Kisel *et al.*, 2011; Powell *et al.*, 2013), and their causal mechanisms (Turner and Tjorve, 2005; Franzen *et al.*, 2012; Hanson *et al.*, 2012; Powell *et al.*, 2013). Various relationships have been proposed to describe TARs (Tjorve, 2003). One of the most commonly used way is the power-law,

$$S = cA^z \qquad (1)$$

where S is the number of species or taxa, A is the area (m^2), c (m^{-2}) is a normalization constant, and z is the species-area exponent that characterizes the rate of spatical scaling (Martín and Goldenfeld, 2006).

Microorganisms are the most diverse group of life on Earth and play critical roles in biogeochemistry and ecosystem functioning. However, our understandings of microbial spatial scaling patterns and the mechanisms driving them are limited (Green *et al.*, 2004; Horner-Devine *et al.*, 2004; Zhou *et al.*, 2008; Hanson *et al.*, 2012; Nemergut *et al.*, 2013). TARs for microbial communities are not as well documented as for plant and animal communities (Logue *et al.*, 2012), but their *z*-values are generally believed to be lower than that of plants and animals, ussually less than 0.1 (Zhou *et al.*, 2008). Recent studies found that microbial *z*-values could be varied with long-term fertilization (Liang *et al.*, 2015), artificially elevated carbon dioxide (Deng *et al.*, 2016) and even seasonal alternation (Liu *et al.*, 2018), however, the underlying mechanism behind these observations is largely unclear.

The objectives of this study are: (i) to document spatial scaling rates of forest soil microbial communities for nested TARs, and (ii) to determine if and how spatical scaling rates vary across a broad temperature gradient. To address these objectives, we used high throughput sequencing to analyze 126 soil samples from six forest sites spanning a large latitudinal

temperature gradient from Central America to North America. Our results showed that while spatial scaling of forest soil microbial diversity is influenced by taxonomic resolution and phylogeny, it is still highly temperature-dependent.

Results

Sampling sites and diversity estimation

A total of 126 soil samples were collected from six forest sites spanning Central America to North America. These six forest sites vary greatly in annual mean temperature, ranging from 2.5 to 25.7 °C (Table S1), as well as other climate and soil characteristics, such as annual mean precipitation, soil type, soil moisture, soil pH, soil total carbon and nitrogen (Table S1).

A mean of 59,378 sequence reads were obtained across all samples (Table S2). The sequence classification based on UCLUST at the 97% similarity threshold showed that an average of 11,256 OTUs were obtained from each sample. Rarefaction analysis showed that most of the dominant taxa were collected, but the large number of single copies indicated that the microbial communities could be still undersampled (Fig. S2). For this reason, we also used Chao1 (Chao, 1984) estimates of microbial diversity. Since the numbers of sequencing reads varied greatly among samples, we resampled 12,531 reads from each sample. To minimize the influence of resampling on OTU richness estimation, we used the average number of OTUs from 100 resamples (Table S2).

Taxa-area relationships (TARs) and *z*-value variations

We tested eight functional forms for TARs (Table S3), and AIC model selection indictated that TARs were best characterized by a power law form. Very strong linear relationships were observed between the $\log_{10}$-transformed OTU richness and $\log_{10}$-transformed

area based on empirical OTU numbers (Fig. 1A, r^2=0.990-0.999) as well as Chao1 estimates (Fig. 1B, r^2=0.966-0.999). Bootstrapping analysis with 1,000 random permutations showed that all observed *z*-values were significantly higher than zero (all $P<0.001$) for all communities (Table S4). Thus, these forest soil microbial communities have significant TARs. The *z*-values for the corresponding aboveground plant communities ranged from 0.10 to 0.31. The microbial TARs and plant TARs were highly correlated (r=0.777, P=0.036) (Fig. S3).

The mean *z*-value calculated for all taxonomic resolutions (97%, 98% and 99%) and all sites was 0.081±0.013 (mean±standard deviation), which is consistent with previous observations that the *z*-values for microbes are generally less than 0.1. However, *z*-values varied across sites (Table S4). The *z*-values also varied by taxonomic resolution (Fig. 2) with the mean *z*-value at the 99% cutoff being 0.093±0.010, which was higher than that for 97% (0.072±0.009) (Table S4). This result indicated the microbial *z* value was generally larger at the finer taxonomic level than the coarser taxonomic level, indicating *z*-value was taxonomic resolution-dependent. In addition, *z*-values varied among different phylogenetic groups, ranging from 0.027 to 0.192 (Fig. S4). For example, the average *z*-values across six sites for OD1 (0.127±0.044) is twice than that for Acidobacteria (0.076±0.012) (Fig. S4). Bootstrap analysis indicated that this variation among different phylogenetic groups is significantly larger than expected.

Microbial *z*-values co-varied with site mean annual temperatures

To determine whether spatial scaling of forest microbial communities varies with environmental conditions, relationships between *z*-values and environmental factors were further examined. Mean annual temperature and *z*-values were correlated at different taxonomic resolutions (Table 1). For instance, at 97% taxonomic resolution, *z*-values can be described by temperature-dependent equations, *z*(T) (Fig. 3), and all three linear and non-linear regressions

had very small differences of AIC values (<2) indicating they were equally good. None of the other environmental variables, including pH and precipitation, were significantly correlated with *z*-values (Table 1).

To further evaluate the influence of temperature on spatial scaling of forest soil microbial communities, a multivariate regression tree (MRT) analysis was performed between community structure and various environmental variables (temperature, pH, plant richness, total carbon and nitrogen). This revealed that temperature was the primary environmental variable to separate the tree at primary and secondary levels, while pH separated the tree at the third level (Fig. 4). These results further demonstrated that temperature is a primary driver of the spatial scaling of microbial diversity.

In addition, we performed correlation analysis between all environmental variables and *z*-values from individual phylogenetic groups at 97% taxonomic resolution (Table S5). As for the entire microbial community, *z*-values for the bacterial domain were only significant correlated with temperature ($P<0.05$). Classes Alphaproteobacteria and Gammaproteobacteria also showed significant correlations with temperature ($P=0.033$) as well as precipitation and NO_3-N (both $P=0.017$). Furthermore, the spatial scaling rate of the archaea domain was significantly correlated with soil moisture ($P=0.017$), while the Proteobacteria phylum and Clostridia class showed significant correlations with TN and NH_4-N ($P=0.017$ and 0.015 respectively), indicating different phylogenetic taxa had distinct variability in spatial scaling rates, and not all were driven by mean annual temperature.

Discussion

Rates of spatial scaling are critical for assessing taxon diversity, extinction rates and species hotspots (Guilhaumon *et al.*, 2008; Harte *et al.*, 2009; He and Hubbell, 2011; Storch *et al.*, 2012). However, these are difficult to estimate for microbial communities due to methodological complications associated with sampling, taxonomic lumping and/or spatial sizes (Cam *et al.*, 2002; Turner and Tjorve, 2005; Zhou *et al.*, 2008). Previous studies have suggested that microbial *z*-values may vary substantially, potentially as a result of differences in the spatial size, habitat type, and taxonomic groups under consideration (Green *et al.*, 2004; Horner-Devine *et al.*, 2004; Zhou *et al.*, 2008; Wang *et al.*, 2009; Gilbert *et al.*, 2012; Tu *et al.*, 2016). In this study, deep sequencing of 16S rRNA genes showed that the average *z*-value for forest soil plots (within 8×10^4 m^2) spanning a large latitudinal temperature gradient is 0.081 ± 0.013 at 97% OTU similarity cutoff (Table S4). This value is consistent with a previous study based on GeoChip analysis, where *z*-values across different phylogenetic/functional groups were below 0.1 in forest (Zhou *et al.*, 2008) and grassland (Liang *et al.*, 2015).

Meanwhile, we found that taxonomic resolution (80%-99% OTU similarity) had large impacts on estimated *z*-values for microbes (Fig. 2). Theory predicts that the number of higher taxa (*e.g.* genera) will increase with area at a slower rate than the number of lower taxa (*e.g.* species), such as an increase in the number of species cannot always be followed by an increase of the number of genera (Storch and Sizling, 2008; Sizling *et al.*, 2011). Thus, *z*-values for genera-area relationships are expected to be shallower than for species-area relationships (Storch and Sizling, 2008). Our results show that spatial scaling rates increased from 0.02-0.04 to 0.07-0.10 when microbial OTU similarity thresholds were changed from 0.80 to 0.99 (Fig. 2), indicating the microbial *z*-values in our surveyed areas were dependent on taxonomic resolution.

Even so, the largest *z* value of microbes (0.106 in CWT under OTU similarity 99%) was still less than the corresponding plant species *z*-values in the same area (0.311). The lower spatial scaling rates observed for microbial communities could be due to the unique characteristics of microorganisms, including high dispersal rates, high diversity, high functional redundancy, large population sizes, rapid asexual reproduction, high degrees of dormancy, and high levels of horizontal gene transfer (Horner-Devine *et al.*, 2004; Zhou *et al.*, 2008; Nemergut *et al.*, 2013). All of these factors could lead to lower spatial scaling rates for microbes at current spatial size we examined. For instance, the roles of natural selection or ecological drift will be less important when diversity, functional redundancy and population abundance are higher, and rapid dispersal will decrease β–diversity of the microbial communities (Nemergut *et al.*, 2013; Vellend *et al.*, 2014). Another potential reason for lower spatial scaling rates in microbial communities could be undersampling. Given the extremely high densities of microbes in soil (Dindal, 1990), current sampling and detection technologies can only sample a small proportion of the total microbial community. Thus, it is possible that some endemic and rare species were not detected, which could reduce the β–diversity between communities. In addition, for nested sampling designs, surveys of larger areas require counting more individuals, and thus, sampling design and not only biological and ecological processes may influence TARs (Rosenzweig, 1995; Cam *et al.*, 2002; Turner and Tjorve, 2005). However, the number of samples needed to adequately characterize microbial communities across area remains unclear. Challenges arising from undersampling cannot be resolved given current technologies, and these will remain outstanding challenges that need to be addresses in future research.

Since temperature drives almost all biochemical process, it has been always regarded as one of the primary factors shaping regional and global biodiversity patterns (Rohde, 1992; Allen

et al., 2002). Some large-scale surveys demonstrated that temperature is an important predictor for bacterial (Prober *et al.*, 2015) and fungal (Tedersoo *et al.*, 2012) diversities, as same as pH, precipitation and plant diversity. Our recent analyses based on these continental forest sites (Tu *et al.*, 2016; Zhou *et al.*, 2016) pointed out that α diversity (especially of soil bacteria, fungi, nitrogen fixers and diazotrophs) was better predicted by temperature than pH and other environmental factors, but this could be only observed when the number of replicates in each site was large enough (Zhou *et al.*, 2017). However, the temperature dependence of β diversity has not been well examined in previous studies. As expected based on MTE, most ecological patterns at all levels of biological organizations are temperature-dependent, and hence temperature should be a useful predictor for understanding TAR pattern as well. The spatial *z*-values are mathematically deduced to increase with temperature (Wang *et al.*, 2009), which was supported by recent studies on plants (Qian *et al.*, 2007; Wang *et al.*, 2009; Qiao *et al.*, 2012). Herein, Our results showed that continental-scale forest soil microbial *z*-values also increased with mean annual temperature. However, this trend was inconsistent with another bacterial study from Antarctic mountainside in highly oligotrophic soils (Okie *et al.*, 2015). According to a niche-based theory, higher temperatures promote wider niches both within and among microorganisms, and meanwhile increasing niche width increases α-diversity but decreases β-diversity (Okie *et al.*, 2015). Our recent analysis has already showed that α diversity of forest soil microbes increased with temperature across latitude (Tu *et al.*, 2016; Zhou *et al.*, 2016), which was coincided with both MTE and niche-width theories. But in this study, the results between β-diversity and temperature are consistent with MTE but not with the niche-width theory, which could be due to several reasons. At continental scales, higher temperatures are generally associated with higher plant diversity, and thus higher levels of substrates, nutrients, physical

attachments and microbial niche space, which could lead to higher γ-diversity of microbial communities. This hypothesis is supported by the observed positive correlation between *z*-values and plant richness (Table 1) as well as the strong correlation between microbial and plant TARs (Fig. S3). This could lead to higher spatial scaling rates of microbial communities (Hanson *et al.*, 2012; Nemergut *et al.*, 2013). For instance, strong positive correlations were observed between latitudinal diversity gradients and speciation rate (Allen *et al.*, 2006), suggesting the importance of temperature in generating species composition and structure. In addition, temperature could interact with other environmental factors such as water availability, carbon and nutrient availability, and pH to affect biodiversity indirectly. For example, higher water availability is associated with higher primary production (Huxman *et al.*, 2004; Campos *et al.*, 2013). Higher temperatures lead to higher decomposition rates and hence nutrient availability (Zhou *et al.*, 2012), which could support and maintain higher plant and microbial diversity. Finally, temperature could also lead to faster ecological drift, *i.e.*, stochastic processes associated with birth, death, colonization, extinction and speciation (Bell, 2001; Hubbell, 2001; Chave, 2004; Chase, 2010; Ofiteru *et al.*, 2010; Chase and Myers, 2011; Zhou *et al.*, 2013; Zhou *et al.*, 2014; Zhou and Ning, 2017). All together lead to microbial *z*-values increased significantly with annual mean temperature in this continental scale.

Interestingly, different microbial phylogenetic groups also exhibited distinct spatial scaling patterns, as same as observed in our previous studies (Zhou *et al.*, 2008; Liang *et al.*, 2015). Fungi and Gram-positive bacteria had lower *z*-values than other microbial taxa, which could be related to their high dispersal ability (Zhou *et al.*, 2008). While copiotrophic bacteria had relatively higher *z*-values than oligotrophic bacteria in an agricultural ecosystem (Liang *et al.*, 2015), but the underlying mechanisms remained largely unclear. In this study, we also

observed different *z*-values among phylogenetic groups and sites (Fig. S4), but those variabilities had no clear difference among oligotropic (*e.g.* Acidobacteria, 0.076±0.007) and copiotrophic phylogenetic (*e.g.* Proteobacteria, 0.070±0.006) groups. However, we did find that the *z*-values of different phylogenetic groups had significant correlations with some other environmental variables beyond temperature (Table S5). These results indicated the spatial scaling rates of different phylogenetic groups could be driven by multiple environmental factors, but more studies are needed to dissentangle the causal mechanisms underlying these patterns.

In summary, understanding the mechanisms driving variation in diversity is a central challenge in ecology. Although TARs are well documented for plant and animal communities, a relatively limited number of studies have been carried out with microbial communities. Using deep sequencing to examine the forest soils with a broad climate gradient, our results demonstrate that TARs do indeed exist for forest soil microbes. Also, our results revealed that environmental temperature has a pervasive influence on microbial biodiversity patterns. In addition, it is becoming widely accepted that climate change and biodiversity are tightly coupled. While it is affected by climate change, through its associated ecosystem functioning and services, understanding biodiversity patterns and the underlying mechanisms is also critical to both climate-change mitigation and adaptation. Thus, the results presented in this study have important implications for biodiversity preservation and ecosystem management. Although predicting how biodiversity changes in response to temperature increase is difficult, the temperature-dependent TARs could provide powerful tools in projecting biodiversity change under future climate warming scenarios.

Experimental procedures

Sampling sites and climate data collection

Six forest sites were selected for this study because they provided variation in ecosystem type from subalpine to tropical forest with a broad range of average annual temperatures. Hourly temperature and precipitation data collected at weather stations at each site were used to calculate mean annual temperature and mean annual precipitation. Additional information about the experimental design can be found at http://macroeco.lternet.edu/ and also has been described previously (Zhou *et al.*, 2016).

Barro Colorado Island (BCI), Barro Colorado National Monument, is located in Panama [9.16ºN, 79.85ºW, 157 meters above the sea level (masl)]. It is a 1560 ha island that was formed by the rising waters of Lake Gatun during creation of the Panama Canal (Dietrich *et al.*, 1996). BCI has a semi-deciduous, lowland tropical moist forest with basaltic and sedimentary derived clay soils (Foster and Brokaw, 1996). Mean annual temperature was collected from the Lutz Tower probes at 42 m for 1984-2013.

The Luquillo Long Term Ecological Research (LTER) site (LUQ), El Yunque National Forest, is located in Puerto Rico (18.32ºN, 65.82ºW, 386 masl). The Luquillo array overlaps with the Luquillo Forest Dynamics Plot in tropical wet forest near the El Verde Field Station. Soils are kaolinitic oxisols (Thompson *et al.*, 2004). Air temperature data for 2001-2010 were collected from a National Atmospheric Deposition Program tower (20 m) located at the El Verde Field Station (http://luq.lternet.edu/data/luqmetadata127).

Coweeta Hydrological Laboratory LTER (CWT), Nantahala National Forest, is located in North Carolina, USA (35.05ºN, 83.43ºW, 864 masl). Soils are Evard–Cowee gravelly loam

(Bonito *et al.*, 2003). Air temperature and precipitation data were obtained for 1984-2013 from climate station 01.

H.J. Andrews Experimental Forest LTER (AND), Willamette National Forest, is located in Oregon, USA (44.23ºN, 122.15ºW, 860 masl). Plots were located in old growth coniferous (Douglas Fir) forest. Soils are Andic Haplumbrept (Whalen *et al.*, 2000). Air temperature and precipitation data for 1989-2011 were obtained from multiple sensors at the site (http://andlter.forestry.oregonstate.edu/data/abstract.aspx?dbcode=MS001).

Harvard Forest LTER (HFR) is located in Massachusetts, USA (42.54ºN, 72.18ºW, 356 masl). Harvard Forest is a second-growth deciduous hardwood forest with podzolic soils (Stout, 1952). Climate data were obtained from Fisher Tower for 2002-2013 (http://harvardforest.fas.harvard.edu/harvard-forest-weather-station).

Niwot Ridge LTER (NWT), Mountain Research Station is located in Colorado, USA (40.04ºN, 105.56ºW, 3186 masl). The forest is spruce/fir and soils are mixed Typic Humicryepts and sandy-skeletal (Williams *et al.*, 2009). Climate data were collected from the adjacent climate station C1 for 1996-2012.

Soil collection and plant survey

A nested sampling design was established at each site (Fig. S1) with 21 $1m^2$ plots in a cross pattern with 4 plots abutting a central plot and 4 plots placed 10, 50, 100 and 200 meters from the central plot in each cardinal direction (Fig. S1). Within each plot, 9 surface soil cores (0-10cm depth) were collected (Oakfield Apparatus Company model HA) in the summer of 2012 and shipped overnight on dry ice to the laboratory. Soil samples were homogenized and chemical properties (*e.g.* water content, pH, total carbon and total nitrogen) were measured as described previously (Zhou *et al.*, 2016).

Plant species-area surveys at each site were implemented using a modified "Gentry plot" methodology (Zhou *et al.*, 2016). Each Gentry plot consisted of five 100 × 2 m transects separated by a distance of 8 m, so that each Gentry plot was located within a 42 × 100 m area (see http://macroeco.lternet.edu/?q=node/13). All plant stems greater than 1 cm basal diameter that were rooted within the transects were censused and identified to species. For plants that extended outside of the transect boundaries, inclusion criteria varied by growth form; trees were censused if the center of their stem base fell within the transect bounds, lianas were censused if rooted within the transect, and hemiepiphytes were censused if any part of the aerial root of rhizome fell within the transect. Stems were tallied as separate individuals if they were not connected above-ground or below-ground within approximately 10 cm of the soil surface.

Microbial analyses

Soil microorganisms were surveyed using Illumina MiSeq technology. Soil microbial DNA extraction, 16S rRNA gene amplification, sequencing and data preprocessing have been described previously (Zhou *et al.*, 2016). Five grams of homogenized soil were used for DNA extraction. PCR was performed to amplify soil DNA using primers of 16S rRNA genes (515F and 806R), and PCR products were sequenced with Illumina MiSeq. Following quality control, all useable sequences were assigned to OTUs according to three similarity thresholds (97%, 98%, and 99%) using UCLUST (Edgar, 2010) in USEARCH v9; these analyses used both *de novo* and close reference parameters and singletons were retained. Microbial taxonomic information was assigned according to the ribosomal database project (RDP) (Wang *et al.*, 2007).

Microbial taxon richness was estimated in two ways. First, we calculated the average number of OTUs from 100 random resamplings of 12,531 sequences (the minimum number of sequences among the 126 samples). Second, we calculated the Chao1 value (Chao, 1984) as

$$S_{chao1} = S_{obs} + \frac{f_1(f_1-1)}{2(f_2+1)} \qquad (2)$$

where S_{obs} is the number of observed OTUs and f_1 and f_2 are the numbers OTUs with exactly one read and two reads, respectively, in the sample. For each sample, f_1 and f_2 were counted through each column of an OTU matrix. For each plot (Fig. S1), the OTU abundance of each replicate was summed and S_{obs}, f_1 and f_2 were calculated. Chao1 values represent the estimated taxonomic richness including both observed and unobserved species.

TAR relationship

To characterize TARs, Eq. (1) was log-transformed to give

$$\log(S) = \log(c) + z \bullet \log(A) \qquad (3)$$

where A was the area under consideration and S was the total number of observed/predicted OTUs. The distances D (m) of each sampling location to the plot center were 1, 10, 50, 100 and 200 meters, yielding sampling areas of 200, 5000, 20000 and 80000 m^2, respectively (Figure S1), as given by $A = \left(\sqrt{2} * D\right)^2 = 2D^2$. Linear regression of log(S) on log(A) was performed using the function **lm** in the statistical software R. Significance of the z-value could not be tested using linear regression since the nested sampling design violated the assumption of independence; therefore, bootstrapping with replacement of log(S) was used. Estimated and observed z-values were compared with Student's t-test after 10,000 bootstraps between area and richness. Additionally, we observed that z-values had either linear or non-linear temperature-dependence. Based on these models, the change of z-values could be estimated by a known range of the mean

annual temperatures. For a given area *A*, microbial richness *S* could be roughly predicted as $S = cA^z$ where *z* is a range from above estimation.

Statistical methods

Pearson correlations were used to evaluate relationships between *z*-values and environmental variables. Analyses were conducted using the package **mmSAR** (Guilhaumon *et al.*, 2010) in R. Goodness of fit was assessed using coefficients of determination (r^2) and Akaike information criterion (AIC).

Multivariate regression trees (MRT) were used to evaluate species-environment relationships (De'Ath, 2002). MRT clusters samples by repeated splitting of data, with each split defined by a simple rule based on environmental variables. Cross validation was used to determine optimal tree steps (De'Ath, 2002). These analyses were conducted using the package **mvpart** in R.

Data Availability

The OTU tables of 16S sequences that support the findings of this study are available in the institute website, http://ieg.ou.edu/4download. The raw sequencing data have been deposited in the NCBI Sequence Read Archive under accession code PRJNA308872.

Acknowledgment

Y.D. was supported by the National Natural Science Foundation of China (Grant No. 31540071), the China National Key Research and Development Program (No. 2016YFC0500702), Strategic Priority Research Program of the Chinese Academy of Sciences (CAS) (Grant XDB15010302) and CAS 100 talent program. The experimental project was supported by the U.S. National Science Foundation MacroSystems Biology program under the contract (NSF EF-1065844), by

the Office of the Vice President for Research at the University of Oklahoma, by the Collaborative Innovation Center for Regional Environmental Quality at the Tsinghua University and the National Science Foundation of China (41430856). The authors claimed no Conflict of Interest in this study.

References

Allen, A.P., Brown, J.H., and Gillooly, J.F. (2002) Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* **297**: 1545-1548.

Allen, A.P., Gillooly, J.F., Savage, V.M., and Brown, J.H. (2006) Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc Natl Acad Sci USA* **103**: 9130-9135.

Bell, G. (2001) Ecology - Neutral macroecology. *Science* **293**: 2413-2418.

Bonito, G.M., Coleman, D.C., Haines, B.L., and Cabrera, M.L. (2003) Can nitrogen budgets explain differences in soil nitrogen mineralization rates of forest stands along an elevation gradient? *Forest Ecology and Management* **176**: 563-574.

Brown, J.H. (2014) Why are there so many species in the tropics? *Journal of Biogeography* **41**: 8-22.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., and West, G.B. (2004) Toward a metabolic theory of ecology. *Ecology* **85**: 1771-1789.

Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J., and Green, J.L. (2008) microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proc Natl Acad Sci USA* **105 Suppl 1**: 11505-11511.

Cam, E., Nichols, J.D., Hines, J.E., Sauer, J.R., Alpizar-Jara, R., and Flather, C.H. (2002) Disentangling sampling and ecological explanations underlying species-area relationships. *Ecology* **83**: 1118-1130.

Campos, G.E.P., Moran, M.S., Huete, A., Zhang, Y.G., Bresloff, C., Huxman, T.E. *et al.* (2013) Ecosystem resilience despite large-scale altered hydroclimatic conditions. *Nature* **494**: 349-352.

Chao, A. (1984) Nonparametric-Estimation of the Number of Classes in a Population. *Scand J Stat* **11**: 265-270.

Chase, J.M. (2010) Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. *Science* **328**: 1388-1391.

Chase, J.M., and Myers, J.A. (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philos Trans R Soc B-Biol Sci* **366**: 2351-2363.

Chave, J. (2004) Neutral theory and community ecology. *Ecol Lett* **7**: 241-253.

De'Ath, G. (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* **83**: 1105-1117.

Deng, Y., He, Z., Xiong, J., Yu, H., Xu, M., Hobbie, S.E. *et al.* (2016) Elevated carbon dioxide accelerates the spatial turnover of soil microbial communities. *Glob Chang Biol* **22**: 957-964.

Dengler, J. (2008) Pitfalls in Small-Scale Species-Area Sampling and Analysis. *Folia Geobot* **43**: 269-287.

Dietrich, W.E., Windsor, D.M., and Dunne, T. (1996) Geology, climate, and hydrology of Barro Colorado Island. In *The Ecology of a Tropical Moist Forest*. Leigh, E.G., Rand, J., A.S. , and Windsor, D.M. (eds). Washington, DC, USA: Smithsonian Institution Press, pp. 21-46.

Dindal, D.L. (1990) *Soil biology guide*. New York: Wiley.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.

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

Foster, R.B., and Brokaw, B.V.L. (1996) Structure and history of the vegetation of Barro Colorado Island. In *The Ecology of a Tropical Moist Forest*. Leigh, E.G., Rand, J., A.S., and Windsor, D.M. (eds). Washington, DC, USA: Smithsonian Institution Press, pp. 67-81.

Franzen, M., Schweiger, O., and Betzholtz, P.E. (2012) Species-Area Relationships Are Controlled by Species Traits. *Plos One* **7**: e37359.

Garcia-Pichel, F., Loza, V., Marusenko, Y., Mateo, P., and Potrafka, R.M. (2013) Temperature Drives the Continental-Scale Distribution of Key Microbes in Topsoil Communities. *Science* **340**: 1574-1577.

Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbrueck, L., Reeder, J., Temperton, B. *et al.* (2012) Defining seasonal marine microbial community dynamics. *ISME J* **6**: 298-308.

Green, J.L., Holmes, A.J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M. *et al.* (2004) Spatial scaling of microbial eukaryote diversity. *Nature* **432**: 747-750.

Guilhaumon, F., Mouillot, D., and Gimenez, O. (2010) mmSAR: an R-package for multimodel species-area relationship inference. *Ecography* **33**: 420-424.

Guilhaumon, F., Gimenez, O., Gaston, K.J., and Mouillot, D. (2008) Taxonomic and regional uncertainty in species-area relationships and the identification of richness hotspots. *Proc Natl Acad Sci USA* **105**: 15458-15463.

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

Harte, J., Smith, A.B., and Storch, D. (2009) Biodiversity scales from plots to biomes with a universal species-area curve. *Ecol Lett* **12**: 789-797.

He, F.L., and Hubbell, S.P. (2011) Species-area relationships always overestimate extinction rates from habitat loss. *Nature* **473**: 368-371.

Horner-Devine, M.C., Lage, M., Hughes, J.B., and Bohannan, B.J.M. (2004) A taxa-area relationship for bacteria. *Nature* **432**: 750-753.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton, N.J.: Princeton University Press.

Hurlbert, A.H., and Stegen, J.C. (2014) When should species richness be energy limited, and how would we know? *Ecol Lett* **17**: 401-413.

Huxman, T.E., Smith, M.D., Fay, P.A., Knapp, A.K., Shaw, M.R., Loik, M.E. *et al.* (2004) Convergence across biomes to a common rain-use efficiency. *Nature* **429**: 651-654.

Kisel, Y., McInnes, L., Toomey, N.H., and Orme, C.D.L. (2011) How diversification rates and diversity limits combine to create large-scale species-area relationships. *Philos Trans R Soc B-Biol Sci* **366**: 2514-2525.

Lauber, C.L., Ramirez, K.S., Aanderud, Z., Lennon, J., and Fierer, N. (2013) Temporal variability in soil microbial communities across land-use types. *ISME J* **7**: 1641-1650.

Lawton, J.H. (1999) Are there general laws in ecology? *Oikos* **84**: 177-192.

Liang, Y., Wu, L., Clark, I.M., Xue, K., Yang, Y., Van Nostrand, J.D. *et al.* (2015) Over 150 years of long-term fertilization alters spatial scaling of microbial biodiversity. *mBio* **6**: e00240-00215.

Liu, S., Wang, H., Deng, Y., Tian, P., and Wang, Q. (2018) Forest conversion induces seasonal variation in microbial beta-diversity. *Environ Microbiol* **20**: 111-123.

Logue, J.B., Langenheder, S., Andersson, A.F., Bertilsson, S., Drakare, S., Lanzen, A., and Lindstrom, E.S. (2012) Freshwater bacterioplankton richness in oligotrophic lakes depends on nutrient availability rather than on species-area relationships. *ISME J* **6**: 1127-1136.

Lomolino, M.V. (2001) The species-area relationship: new challenges for an old pattern. *Prog Phys Geogr* **25**: 1-21.

Martín, H.G., and Goldenfeld, N. (2006) On the origin and robustness of power-law species-area relationships in ecology. *Proc Natl Acad Sci USA* **103**: 10310-10315.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F. *et al.* (2013) Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* **77**: 342-356.

Ofiteru, I.D., Lunn, M., Curtis, T.P., Wells, G.F., Criddle, C.S., Francis, C.A., and Sloan, W.T. (2010) Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci USA* **107**: 15345-15350.

Okie, J.G., Van Horn, D.J., Storch, D., Barrett, J.E., Gooseff, M.N., Kopsova, L., and Takacs-Vesbach, C.D. (2015) Niche and metabolic principles explain patterns of diversity and distribution: theory and a case study with soil bacterial communities. *Proc Biol Sci* **282**: 20142630.

Powell, K.I., Chase, J.M., and Knight, T.M. (2013) Invasive Plants Have Scale-Dependent Effects on Diversity by Altering Species-Area Relationships. *Science* **339**: 316-318.

Prober, S.M., Leff, J.W., Bates, S.T., Borer, E.T., Firn, J., Harpole, W.S. *et al.* (2015) Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecol Lett* **18**: 85-95.

Qian, H., Fridley, J.D., and Palmer, M.W. (2007) The latitudinal gradient of species-area relationships for vascular plants of North America. *Am Nat* **170**: 690-701.

Qiao, X.J., Tang, Z.Y., Shen, Z.H., and Fang, J.Y. (2012) What causes geographical variation in the species-area relationships? A test from forests in China. *Ecography* **35**: 1110-1116.

Rohde, K. (1992) Latitudinal Gradients in Species-Diversity - the Search for the Primary Cause. *Oikos* **65**: 514-527.

Rosenzweig, M.L. (1995) *Species diversity in space and time*. Cambridge ; New York: Cambridge University Press.

Ryberg, W.A., and Chase, J.M. (2007) Predator-dependent species-area relationships. *Am Nat* **170**: 636-642.

Sharp, C.E., Brady, A.L., Sharp, G.H., Grasby, S.E., Stott, M.B., and Dunfield, P.F. (2014) Humboldt's spa: microbial diversity is controlled by temperature in geothermal environments. *ISME J* **8**: 1166-1174.

Sizling, A.L., Kunin, W.E., Sizlingova, E., Reif, J., and Storch, D. (2011) Between Geometry and Biology: The Problem of Universality of the Species-Area Relationship. *Am Nat* **178**: 602-611.

Storch, D., and Sizling, A.L. (2008) The Concept of Taxon Invariance in Ecology: Do Diversity Patterns Vary with Changes in Taxonomic Resolution? *Folia Geobot* **43**: 329-344.

Storch, D., Keil, P., and Jetz, W. (2012) Universal species-area and endemics-area relationships at continental scales. *Nature* **488**: 78-81.

Stout, B. (1952) Species distributions and soils in the Harvard Forest. *Harvard Forest Bulletin* **24**: 1-29.

Tedersoo, L., Bahram, M., Toots, M., Diedhiou, A.G., Henkel, T.W., Kjoller, R. *et al.* (2012) Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. *Mol Ecol* **21**: 4160-4170.

Thompson, J., Brokaw, N., Zimmerman, J.K., Waide, R.B., E.M. Everham, I., and Schaefer, D.A. (2004) Luquillo Forest Dynamics Plot. In *Tropical Forest Diversity and Dynamism: Results from a Long-Term Tropical Forest Network*. Losos, E., and E.G. Leigh, J. (eds). Chicago, USA: University of Chicago Press, pp. 540-550.

Tjorve, E. (2003) Shapes and functions of species-area curves: a review of possible models. *J Biogeogr* **30**: 827-835.

Tu, Q.C., Deng, Y., Yan, Q.Y., Shen, L.N., Lin, L., He, Z.L. *et al.* (2016) Biogeographic patterns of soil diazotrophic communities across six forests in the North America. *Mol Ecol* **25**: 2937-2948.

Turner, W.R., and Tjorve, E. (2005) Scale-dependence in species-area relationships. *Ecography* **28**: 721-730.

Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J. *et al.* (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**: 1420-1430.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**: 5261-5267.

Wang, Z., Brown, J.H., Tang, Z., and Fang, J. (2009) Temperature dependence, spatial scale, and tree species diversity in eastern Asia and North America. *Proc Natl Acad Sci USA* **106**: 13388-13392.

Whalen, J.K., Bottomley, P.J., and Myrold, D.D. (2000) Carbon and nitrogen mineralization from light- and heavy-fraction additions to soil. *Soil Biol Biochem* **32**: 1345-1352.

Williams, M.W., Seibold, C., and Chowanski, K. (2009) Storage and release of solutes from a subalpine seasonal snowpack: soil and stream water response, Niwot Ridge, Colorado. *Biogeochemistry* **95**: 77-94.

Woodcock, S., Curtis, T.P., Head, I.M., Lunn, M., and Sloan, W.T. (2006) Taxa-area relationships for microbes: the unsampled and the unseen. *Ecol Lett* **9**: 805-812.

Zhou, J., and Ning, D. (2017) Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol Mol Biol Rev* **81**.

Zhou, J., Liu, W., Deng, Y., Jiang, Y.H., Xue, K., He, Z. *et al.* (2013) Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. *mBio* **4**: e00584-00512.

Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J.D. *et al.* (2014) Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc Natl Acad Sci U S A* **111**: E836-845.

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D. *et al.* (2016) Temperature mediates continental-scale diversity of microbes in forest soils. *Nat Commun* **7**: 12083.

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D. *et al.* (2017) Correspondence: Reply to 'Analytical flaws in a continental-scale forest soil microbial diversity study'. *Nat Commun* **8**: 15583.

Zhou, J.Z., Kang, S., Schadt, C.W., and Garten, C.T. (2008) Spatial scaling of functional gene diversity across various microbial taxa. *Proc Natl Acad Sci USA* **105**: 7768-7773.

Zhou, J.Z., Xue, K., Xie, J.P., Deng, Y., Wu, L.Y., Cheng, X.H. *et al.* (2012) Microbial mediation of carbon-cycle feedbacks to climate warming. *Nat Clim Change* **2**: 106-110.

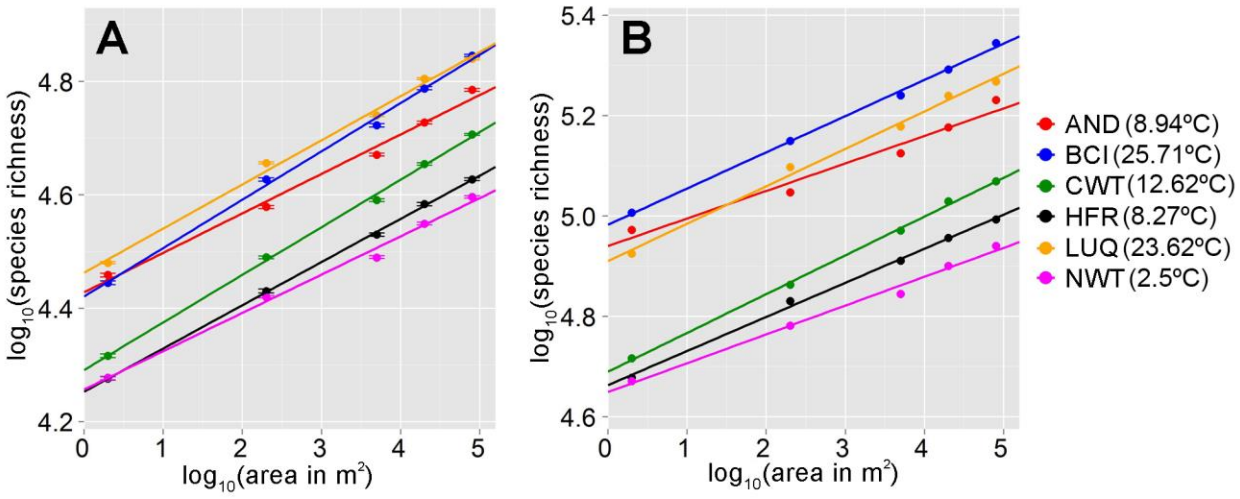
Figure legends:

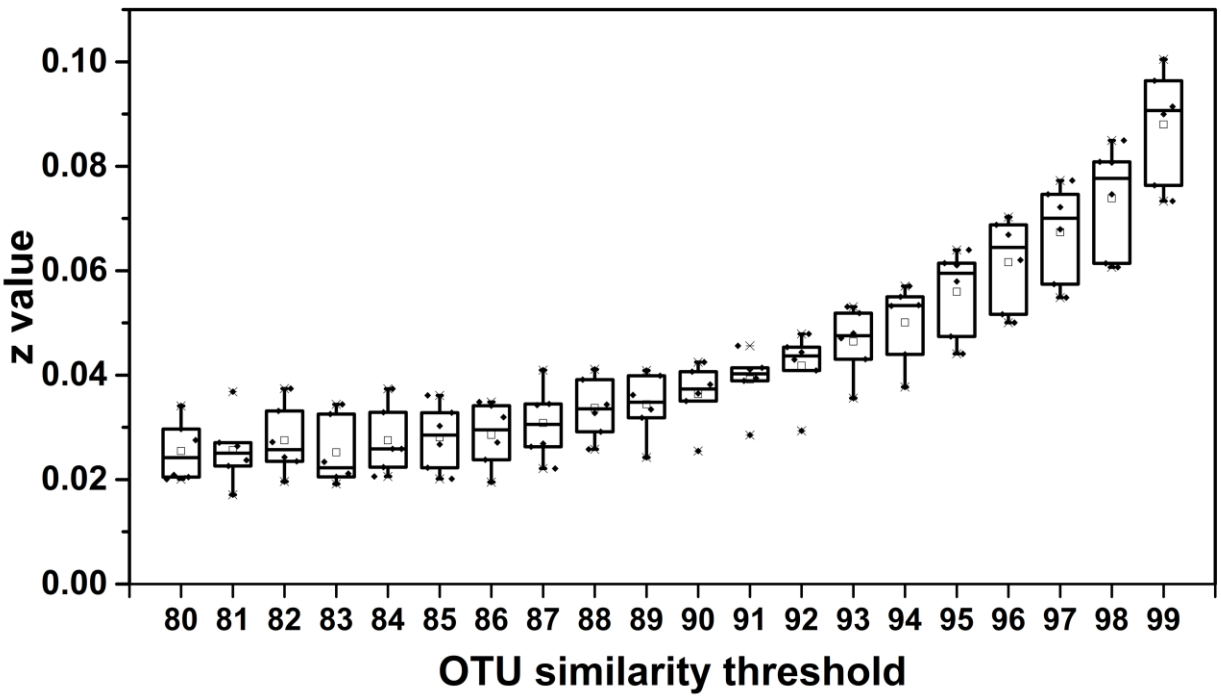
Fig. 1. Species-area relationships for six forests spanning a latitudinal temperature gradient from Central America to North America. (A) Species richness was estimated as observed 16S OTU numbers with 97% similarity. Error bars represent the maximum and minimum values from 100 times of sequence resampling. (B) Species richness was estimated as the Chao1 index. The site mean annual temperatures have been marked in the legend

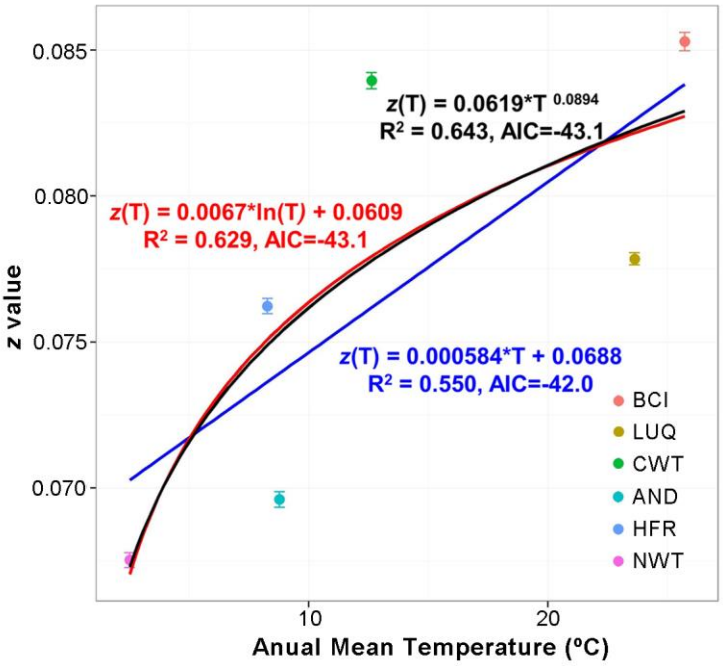
Fig. 2. Spatial scaling rates of taxa-area relationships (*z*-values) for microbial communities from six forest soils at a series of taxonomic resolution (from 80% to 99%). The *z*-value for each site was measured by using Chao1 estimated richness.

Fig. 3. Linear and non-linear regressions on microbial community *z*-values on mean annual temperature (97% similarity level). The error bars represent the maximum and minimum *z*-values obtained from 100 resample for each site. The regression curve in blue is linear regression, while the one in red and in black are logarithmic and exponential regressions respectively.

Fig. 4. Multivariate Regression Tree (MRT) for microbial community classification. Mean annual temperature, pH, plant richness, total carbon and total nitrogen were used for MRT modeling, but only significant splitting steps were plotted. The significant steps were determined by cross-validation with the least relative errors (De'Ath, 2002). *N* is the number of samples can be groups in this branch. The abbreviation "Temp" in the branches represented the mean annual temperature. The bar chart for each branch shows the relative abundances of major classes in those samples.







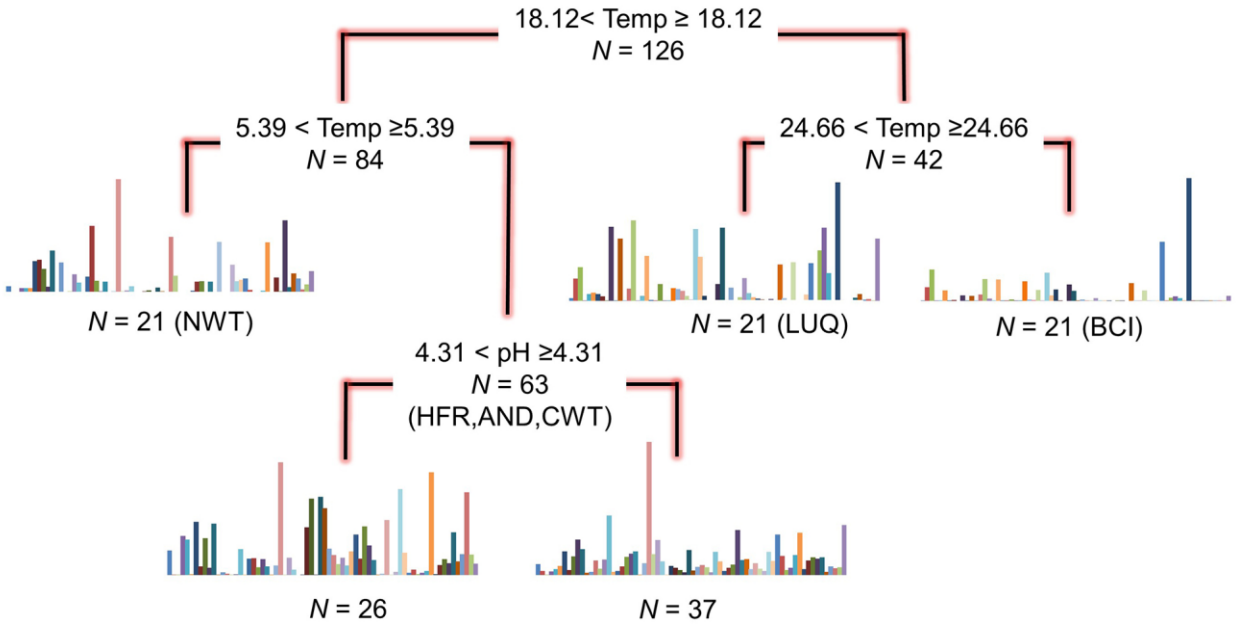


Table 1. Correlations of spatial scaling (*z*) of microbial communities as a whole and environmental conditions at different taxonomic resolutions

Environmental conditions	97%		98%		99%	
	r	*P*	*r*	*P*	*r*	*P*
Latitude	-0.680	0.069	-0.659	0.077	-0.564	0.122
Elevation	-0.680	0.069	-0.682	0.068	-0.658	0.078
Annual mean temperature	0.739	0.047*	0.732	0.049*	0.647	0.083
Mean annual precipitation	0.630	0.090	0.645	0.083	0.579	0.114
Soil moisture	0.336	0.257	0.357	0.244	0.335	0.258
pH	0.227	0.332	0.215	0.341	0.099	0.426
Total carbon	-0.684	0.067	-0.682	0.068	-0.630	0.090
Total nitrogen	-0.265	0.306	-0.276	0.298	-0.249	0.317
NH4.N	0.454	0.183	0.425	0.201	0.309	0.276
NO3.N	0.364	0.239	0.364	0.239	0.280	0.296
Plant richness	0.718	0.054	0.684	0.067	0.588	0.110
Plant biomass	-0.424	0.201	-0.406	0.212	-0.435	0.194

* $P < 0.05$