Journal of Microbiology (2013) Vol. 51, No. 4, pp. 439–447
DOI 10.1007/s12275-013-2446-9

Bacterial Diversity in the Mountains of South-West China: Climate Dominates Over Soil Parameters[§]

Dharmesh Singh[1], Lingling Shi[2], and Jonathan M. Adams[1*]

[1]*Department of Biological Sciences, College of Natural Sciences, Seoul National University, Seoul 151-742, Republic of Korea*
[2]*Soil Ecology Group, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, 88 Xuefu Road, Kunming 650223, P. R. China*

(Received August 21, 2012 / Accepted March 20, 2013)

Certain patterns in soil bacterial diversity and community composition have become evident from metagenomics studies on a range of scales, from various parts of the world. For example, soil pH has generally been seen as dominating variation in bacterial diversity, above all other soil and climate parameters. It is important however to test the generality of these relationships by studying previously unsampled areas. We compared soil bacterial diversity and community composition under a wide range of climatic and edaphic conditions in mountainous Yunnan Province, SW China. Soil samples were taken from a range of primary forest types and altitudes, reflecting the great variation of forest environments in this region. From each soil sample, DNA was extracted and pyrosequenced for bacterial 16S rRNA gene identification. In contrast to other recent studies from other parts of the world, pH was a weaker predictor of bacterial community composition and diversity than exchangeable Ca^{2+} concentration, and also the more poorly defined environmental parameter of elevation. Samples from within each forest type clustered strongly, showing the distinctive pattern of their microbial communities on a regional scale. It is clear that on a regional scale in a very heterogeneous environment, additional factors beyond pH can emerge as more important in determining bacterial diversity.

Keywords: bacteria, diversity, soil, pyrosequencing, S.W. China

Introduction

The cataloguing and understanding of patterns in the diversity of life is one of the main aims of ecology (Huston, 1994; Adams, 2009). The field has been further invigorated recently by the discovery, made possible with advanced molecular genetic methods, of a vast hidden diversity of bacteria everywhere in nature, especially in soil (Torsvik *et al*., 1990, 2002; Curtis and Sloan, 2005; Gans *et al*., 2005).

Broad patterns of bacterial diversity of soils are still being explored (Fierer and Jackson, 2006; Lozupone and Knight, 2007; Lauber *et al*., 2009). Reviews of widely scattered sample points have suggested that pH is the most consistent predictor of bacterial diversity in soils at both continental scales (Fierer and Jackson, 2006; Lauber *et al*., 2009) and local scales (Aciego Pietri and Brookes, 2008; Rousk *et al*., 2010), and at a regional scale (Jesus *et al*., 2009; Griffiths *et al*., 2011; Tripathi *et al*., 2012). No soil parameter other than pH has emerged as such a strong predictor of variation in bacterial diversity, in normal (non-polluted) terrestrial environments. There is no obvious geographical trend in soil bacterial diversity with latitude–in contrast to large eukaryotes where a monotonic decrease in the species richness towards the poles is often seen (Willig *et al*., 2003; Hillebrand, 2004).

It is so far unclear whether there is any consistent trend in bacterial diversity with elevation and/or temperature in mountainous regions, with different studies reporting different results. Most groups of eukaryotes (e.g., mammals, amphibians) which have been studied show a diversity peak in lower to mid altitudes although a monotonic decline in diversity towards higher altitudes has also been reported (e.g., birds, trees) (Rahbek, 1995; Brown, 2001; Lomolino, 2001; McCain, 2005; Rahbek, 2005). Bryant *et al*. (2008) found a decline in diversity of one group of bacteria, *Acidobacteria*, with altitude in the western Rocky Mountains of the USA and a study of Mt. Fuji, Japan reported a hump-backed trend in bacterial diversity with increasing elevation (Singh *et al*., 2012). By contrast, Fierer *et al*. (2011) did not find any trend in total bacterial richness with altitude in a study of the tropical Andes.

To some extent, simple spatial distance may also be a predictor of variation in bacterial community structure and diversity, as was found by Martiny *et al*. (2011). However, this is not generally found to be the case (Chu *et al*., 2010).

Once the relationships between bacterial diversity and particular environmental factors are known, it will become possible to begin more detailed investigation of the structuring of bacterial communities in soils.

This study aimed to identify the environmental factor or factors which are most important in determining variation in bacterial community composition (BCC) and diversity in relation to environmental heterogeneity and geographic distance in mountainous south-west China (SW China), a region that so far has been little studied from a metagenomic viewpoint. With a great diversity of climate and geology amongst the various mountain systems, we hoped that the

*For correspondence. E-mail: jonadams@snu.ac.kr; Tel.: +82-2-880-4339; Fax: +82-2-876-8153
[§]Supplemental material for this article may be found at http://www.springerlink.com/content/120956.

relative importance of elevation (as a proxy for climate-related factors) and soil parameters could be explored.

Our working hypothesis was that pH would dominate variation in bacterial diversity in SW China, largely to the exclusion of other variables. This is the predominant pattern seen on the broad geographical scale (Fierer and Jackson, 2006; Lauber *et al.*, 2009) and on the local scale (Rousk *et al.*, 2010). We further hypothesized that pH would explain most variation in diversity of higher level taxa, such as phyla and sub-phyla.

Materials and Methods

Study area

The south-western part of China is topographically extremely diverse, and its forest zones range from 'paratropical' evergreen forest to 'boreal' type conifer forest. This area in general is well-known for spanning plant biodiversity hotspots (Myers *et al.*, 2000) (www.biodiversityhotspots.org). Most of the Yunnan province lies within the Mountains of the SW China biodiversity hotspot whereas the Ailao Shan region of Yunnan falls into the Indo Burma hotspot along with the Wuilang Shan of Central Yunnan (Myers *et al.*, 2000) (www.biodiversityhotspots.org). We sampled forested areas from Ailao Shan, Xishuangbanna, and Zhongdian mountain ranges from Yunnan province of SW China. Climates across Yunnan are generally moist, with a strong summer peak of rainfall associated with the northern hemisphere monsoon.

General sampling methodology

We sampled old growth forest reserves representing each of the main forest types in this region. Sets of samples were taken from broad-leaved temperate deciduous forest, cool climate coniferous forest, paratropical evergreen broad leaved forest, tropical seasonal rain forest over limestone, montane moist evergreen broad leaved forest and montane mossy forest, differing in temperature, rainfall, seasonality and underlying soil type (Table 1). The number of sites that could be sampled was limited by, a) availability of old growth forest and b) the limited distribution of areas accessible and flat enough to sample in this steeply mountainous region. Consequently, only one major forest reserve could be sampled to represent each biome.

Sampling took place at 8 forest reserves located within the following three main areas in Yunnan province of China: (see Fig. 1) Zhongdian (a high-plateau temperate region 27°51′N, 99°43′E), Ailao Shan mts (subtropical climate region 24°32′N, 101°01′16′E) and Xishuangbanna (tropical climate region 21°55′N, 101°16′E). For the Zhongdian region, study sites were located in a broad-leaved deciduous forest (Z1) and coniferous Forest (Z2). The broad-leaved forest is dominated by the species *Quercus semicarpifolia*, family *Fagaceae*, and the coniferous forest is dominated by the species

Table 1. Geographical and climactic information of sampled sites

Study site	Label	No. of samples	Vegetation	Elevation (m.a.s.l.)	Coordinates	MAT (°C)	PPT (mm)	EVAPN	Plant species	Label	Soil type	Veg. Richness	pH	SOM (g/kg)	TC (g/kg)	TN (g/kg)	TP (g/kg)	Ca (cmol/kg)	Mg (cmol/kg)	K (cmol/kg)
Zhong Dian	Z2	4	Coniferous forest	3350	27°47′N, 99°43′E	4	NA	NA	*Larix potaninii* var. *macrocarpa*; *Rhododendron simsii*; *Iris forrestii*	Z2	Alfisol	3.7	4.76	68	30.81	2.27	1.69	2.54	0.46	0.35
	Z1	4	Broadleaved deciduous forest	3250	27°38′N, 99°30′E	5	606.6	NA	*Quercus semicarpifolia*	Z1	Alfisol	8.0	5.46	NA	23.79	1.93	0.75	7.96	0.47	0.44
Ailao Shan	A2	4	Montane mossy forest	2900	24°32′N, 101°00′E	8	NA	NA	*Ilex delavayi*; *Rhodoleia championii*	A2	Alfisol	10.0	4.00	171	99.30	6.31	0.64	0.46	0.42	0.32
	A1	5	Montane moist evergreen broad-leaved forest	2450	24°28′N, 101°40′E	12	1931.1	1486	*Lithocarpus xylocarpa*; *Castanopsis wattii*	A1	Alfisol	16.0	5.01	116	91.25	6.56	1.11	1.04	0.47	0.37
	A3	4	Montane moist evergreen broad-leaved forest	2400	24°20′N, 101°31′E	10	NA	NA	*Lithocarpus xylocarpa*; *Castanopsis wattii*	A3	Alfisol	15.5	4.04	NA	123.5	8.87	NA	NA	NA	NA
Xishuangbanna	B2	5	Evergreen broad-leaved lower montane forest	1500	21°105′N, 101°12′E	12	NA	NA	*Castanopsis microcarpa*; *Aporusa yunnanensis*; *Schima superb Gardn et Champ*	B2	Ultisol	70.0	4.71	NA	17.81	1.54	0.27	0.68	0.07	0.20
	B1	5	Tropical seasonal rain forest over limestone	750	21°54′N, 101°16′E	20	NA	NA	*Cleistanthus sumatranus*; *Sumbaviopsis albicans*; *Lasiococca comberi Haines*	B1	Ultisol	29.0	7.02	NA	38.08	4.39	2.11	27.91	1.78	0.95
	B3	5	Tropical seasonal rainforest	650	21°41′N, 101°25′E	22	1186	1128	*Pometia pinnata Forst*; *Ardisia affinis*; *Cleidion brevipetiolatum pax et Hoffm*	B3	Ultisol	106.0	4.50	32	12.73	1.53	0.32	1.31	0.30	0.17

m.a.s.l, meters above sea level; MAT, mean annual temperature; PPT, precipitation; EVAPN, evaporation; SOM, soluble organic matter; TC, total carbon; TN, total nitrogen; TP, total phosphorus

Fig. 1. Map of Yunnan province with approximate locations for the three sampling areas, encompassing total 8 different forest types. Map of Yunnan province in China (inset). Modified from http://www.sacredland.org/yunnan-province/.

Larix potaninii var. *macrocarpa*. Within the Ailao Shan range, three forests of two vegetation types were selected for sampling: montane mossy forest (A2) located at the top of the mountains and dominated by *Ilex delavayi* and *Rhodoleia championii*, and montane moist evergreen broad-leaved forests A1 & A3 located at somewhat lower altitudes and dominated by species *Lithocarpus xylocarpum* and *Castanopsis wattii*. In the Xishuangbanna paratropical region, evergreen broad-leaved lower montane forest (B2) (Cao and Zhang, 1997), tropical seasonal rain forest over limestone (B1) and tropical seasonal rainforest (B3) were selected for sampling. Evergreen broad-leaved lower montane forest (B2) (Cao and Zhang, 1997) is located at higher altitudes of these mountains and dominated by such species as *Castanopsis microcarpa* and *Aporusa yunnanensis*. Another study site was located in a lower altitude tropical seasonal rainforest over limestone (B1), with vegetation dominated by *Cleistanthus sumatranus*, *Sumbaviopsis albicans*, and *Lasiococca comberi*. Non-limestone tropical seasonal rainforest (B3) was sampled in a flat area between two hills extending from east to west, containing permanent plots (dominated by *Pometia tomentosa* and *Terminalia myriocarpa*) dedicated to the long-term ecological research managed by the Tropical Rainforest Ecosystem Station. Basic geographical and climate information for all sites are listed in Table 1.

In each forest reserve sampled, four or five sites at least 100m apart were randomly chosen (Table 1). At each sampling site, a 2 × 2 m plot was established and five soil cores (5 cm in diameter and 10 cm in depth) were taken from the middle of four sides or four corners and the center; these samples were then placed together in a bag and mixed. Samples at all sites were taken within seven days of one another at the end of the monsoon season in September 2009. A total of 36 samples were taken from all of the 8 forest sites (Table 1). Soil samples were sieved through 2-mm-mesh sieves. Samples were stored at moderate temperatures (10–20°C) for a maximum of 2 days until they could be frozen at -80°C ready for DNA extraction. Soil samples can be stored under field conditions without refrigeration for approximately up to 2 weeks without affecting the microbial community (Lauber *et al*., 2010).

DNA extraction and pyrosequencing

DNA was extracted from each of the collected sieved soil samples using the MOBIO Power Soil DNA extraction kit (MOBIO Laboratories, USA) as directed by the manufacturer. Isolated DNA was stored at -80°C. PCR amplification used bar-coded primers targeting the V1 to V3 region of the 16S rRNA gene, with PCR conditions and primers as previously described by Hur *et al*. (2011). Briefly, PCR reactions were performed in 50 µl reactions, each containing 1 µl (20 nm) of both primers, 5 µl (PCR reaction buffer with $MgCl_2$, 10X), 1 µl (dNTP mix), 0.25 µl (Taq DNA Polymerase, 5 U/µl) (Roche Diagnostics GmbH, Germany) and 1 µl of DNA as template. We used the following PCR conditions: initial denaturation 94°C, 5 min, followed by 10 cycles (denaturation, 94°C, 30 sec; annealing, 60°C to 55°C with a touch-down program for 45 sec; elongation, 72°C, 90 sec) tailed by an additional 20 cycles (denaturation, 94°C, 30 sec; annealing, 55°C, 45 sec; elongation, 72°C, 90 sec). Pooled reactions were purified using the QIAquick PCR purification kit (Qiagen) and quantified using PicoGreen (Invitrogen) spectrofluorometrically (TBS 380, Turner Biosystems, Inc., USA). 50 ng of PCR product for each sample was combined in a single tube and sent to Chunlab Inc. (Korea) for pyrosequencing using Roche/454 GS FLX Titanium platform.

Chemical analyses

Soil pH, organic carbon content, total C, N, & P, exchangeable Mg^{2+}, K^{+}, and Ca^{2+} were analyzed according to the analytical procedures of the State Forestry Administration, P.R. China (1999), which have been described in Chan *et al*. (2006). Each sample was analyzed in duplicate. Significant differences between individual samples among the soil chemical properties were evaluated using an unassuming equal variances test Tamhane's T2, one-way ANOVA ($P \leq 0.05$, SPSS version 13.0).

Processing of pyrosequencing data and taxonomic analysis

Data were processed following the Costello analysis protocol on the Mothur platform (Schloss *et al*., 2009) with an additional step of removing the chimeric sequences (Singh *et al*., 2012). Briefly, 10 random read subsets for a read of 1,000 for each sample replicate were generated using CD-HIT (-as the number of reads affects the number of OTUs) and this was used for further analysis (http://www.mothur.org/wiki/Costello_stool_analysis) (Schloss *et al*., 2009) using EzTaxon-e alignment bacterial database as a template (Kim *et al*., 2012). To assign sequences into OTUs, we clustered sequences by the furthest neighbor method with 97% sequence similarity as the designated cut off. The assigned

OTUs were then used to calculate coverage, richness, diversity (Supplementary data Table S1) and rarefaction values for each sample. A maximum likelihood (ML) tree was inferred using FastTree2 (Price *et al.*, 2010) and was used to calculate Faith's PD value using the Mothur platform (Schloss *et al.*, 2009).

Unweighted Pairwise UniFrac distance matrix (Hamady *et al.*, 2010) and Bray-Curtis similarity matrix for the overall community (hereafter referred as *genetic matrices*), as well as for the 2 most abundant phyla (*Acidobacteria* and *Proteobacteria* which comprised around 70% of the total quality sequences obtained) were calculated for a subset of 1,000 reads using Primer v6 (Clarke and Gorley, 2006) for further statistical analysis.

Environmental variables and geographic distance

An environmental distance matrix was calculated using Euclidean distance on a dataset containing samples as rows and an environmental data as columns utilising Primer v6 (Clarke and Gorley, 2006). Environmental variables for which more than one reading was missing were excluded from the analysis (Table 1). Also C:N ratio was taken as a single variable rather than including total carbon and nitrogen as individual variables. Environmental variables were transformed where needed, and normalised before calculating the distance matrix. Geographic distance between different forest types was calculated using the Haversine formula on latitudinal and longitudinal co-ordinates and was used to make a distance matrix (i.e. spatial matrix).

Statistical analysis

Non-Metric multidimensional scaling (NMDS), analysis of similarity (ANOSIM) and Mantel tests (Relate function) were performed using Primer v6 (Clarke and Gorley, 2006). An independent matrix was calculated for each environmental variable, and then we used Mantel tests to look at the correlation between distance for each environmental variable and Bray-Curtis similarity and UniFrac distance for the whole bacterial community, as well as for the two most abundant phyla.

To disentangle the relative impact of environmental heterogeneity and spatial distance on the distribution of microbial soil communities, we used a partial Mantel test [ecodist R package; Goslee and Urban (2007)], employing genetic, spatial and environmental distance matrices calculated as specified above.

We used a multiple regression on matrices (MRM) approach to study the relative importance of each of the environmental factors in affecting community similarity (Legendre *et al.*, 1994). Before applying MRM to the dataset we checked for redundant edaphic factors using the VARCLUS procedure (Sarle, 1990) in the Hmisc R package. Mean annual temperature (MAT) and vegetation richness were highly correlated with elevation (Supplementary data Fig. S1), and thus we removed them for the MRM analysis. Similarly, exchangeable magnesium (Mg^{2+}) was highly correlated with exchangeable potassium (K^{+}) and was also removed. With the 6 environmental variables left (on the basis of VARCLUS results and BEST procedure in Primer v6), we estimated a new environmental distance matrix and performed MRM using this environmental distance matrix and genetic matrices calculated as specified above. The MRM approach allows testing the correlation between community similarity and each independent variable, keeping all other variables within the model fixed. Non-significant factors were removed sequentially and the MRM analysis was repeated until only significant factors were left in the model. P-values of two-tailed tests are reported for this analysis. Data within distance matrices is not independent and therefore significance is evaluated through permutations. Here we used 9999 permutations for Mantel tests and MRM to assess the significance level.

Results and Discussion

Distribution of taxa and phylotypes

All the samples show a typically high diversity of bacterial phylotypes (OTUs), within the general range found for soil samples elsewhere in the world (Roesch *et al.*, 2007). In total, from 36 samples distributed across 8 forest types, we obtained 223,279 quality sequences which we were able to classify up to OTU level grouped at ≥97% similarity. Around 60% of the OTUs were represented by a single sample read. The most abundant single phylotype was an unclassified sequence belonging to the phylum *Acidobacteria*, as was the second most abundant phylotype from genus *Koribacter* represented by 6208 (approx. 2.78%) and 3941 (approx. 1.77%) sequences respectively. The third most abundant single phylotype was classified as a member of the *Alphaproteobacteria*, order *Rhizobiales* and was represented by 2607 (1.17%) sequences.

Of the classifiable sequences, 39 phyla were identified across the sample set (listed in Supplementary data Table S2). The dominant phyla in order of abundance were *Acidobacteria* (39.9%), *Alphaproteobacteria* (15.9%), *Gammaproteobacteria* (5.31%), *Actinobacteria* (5.16%), *Betaproteobacteria* (4.35%), *Bacteroidetes* (4.27%), *Deltaproteobacteria* (4.02%), *Planctomycetes* (3.83%), *Chloroflexi* (3.14%), and *Verrucomicrobia* (3.05%) etc. Less abundant phyla, but still found in all samples, belonged to *Nitrospirae* (1.89%), *Cyanobacteria* (1.73%), *Gemmatimonadetes* (1.11%), *TM7* (0.72%), and *OP10* (0.66%). Apart from these phyla, we found sequences from 31 additional phyla and additional sequences which we were not able to classify under any known phyla (denoted by Bacteria* in Supplementary data Table S2).

Across all the samples, *Acidobacteria* were the most abundant by a large margin, except for the monsoon rain forest over limestone where proteobacterial sequences (28.15%) narrowly exceeded *Acidobacteria* (24.7%). Other than this, *Cyanobacteria* were much less abundant in the monsoon rain forest over limestone. In the limestone forest there was by contrast an increased proportion of *Gemmatimonadetes* (4 times more abundant than other forest types) as well as *WS3* and *Thermobaculum* (now re-classified as a member of phylum *Chloroflexi*; (Kunisawa, 2011) which were either virtually absent or only present in a very small proportion in the other forest soil samples. Similarly, phylum *Verrucomicrobia* was found in a much higher proportion in the

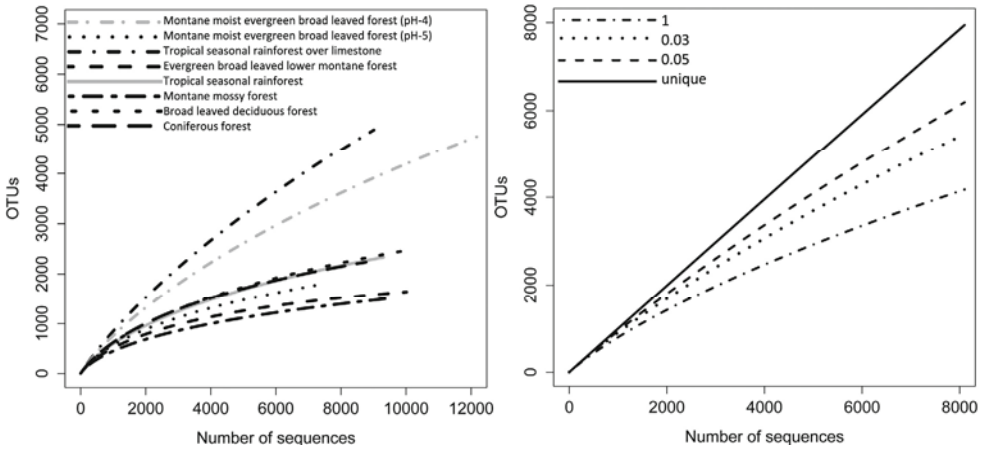

Fig. 2. Rarefaction analysis of one replicate from each forest sample calculated for the 0.3 OTU definition (A) and for one biological replicate of Tropical seasonal rainforest over limestone (L2) based on pairwise distance (B). (B) Rarefaction is shown for OTUs that contain unique sequences and OTUs with differences that do not exceed 3, 5, or 10%. Rarefaction of the other samples showed curves with similar slopes.

coniferous and broad leaved deciduous forest types in Zhongdian than at the other sites.

Phylum *Acidobacteria* was dominated in terms of abundance by classes *Acidobacteria*_Gp1, and Gp2 in all samples, except in those from the limestone forest where the acidobacterial sequences from Gp6, Gp4 and Gp5 were the most dominant. Similarly, *Alphaproteobacteria* was the most abundant taxon amongst the *Proteobacteria*, except in limestone forest where *Deltaproteobacteria* was almost as abundant as *Alphaproteobacteria*. *Sphingobacteria* and *Flavobacteria* were the most abundant in the phylum *Bacteriodetes*, and its percentage was much higher in the Zhongdian and limestone forests, accounting for around 1.14% and 4.6% respectively; even the percentage occupied by the phylum *Bacteriodetes* in these three forest types (ZD1, ZD2, and B1) was much higher compared to the remaining forest types (approximately 2.5 times higher). Phylum *Deinococcus* and one unclassified rare phyla *CHS* were represented by only a single sequence. More complete information on the relative abundances of different phyla is provided in Supplementary data Table S2.

The number of reads per sample ranged from 1499 to a maximum of 10366 sequences in a single replicate but even at this depth of sequencing, we had not surveyed the full extent of bacterial diversity/richness as evidenced by the lack of asymptotes in the rarefaction curves generated after assigning OTUs at the ≥97% similarity level of taxonomic resolution (Fig. 2).

Variability in bacterial richness and diversity

To avoid bias due to differences in the number of reads, we sub-sampled the bacterial communities for each of the samples and calculated phylotype richness (OTUs) and diversity (Faith's PD) per 1,000 reads. Bacterial OTU richness was positively correlated with pH, (Pearson correlation, R^2=0.21; $P \leq 0.004$; Supplementary data Fig. S2) and linearly correlated with elevation (Pearson correlation, R^2= 0.13; $P \leq 0.02$). Our results corroborate the correlations observed earlier between bacterial diversity and soil pH (in the interval of 4–7), as can be seen in previous studies; for example, Lauber *et al.* (2009) and Rousk *et al.* (2010). This confirms that the pattern is robust across different spatial scales and soil types. Another factor, elevation, which was shown to be linearly correlated here has been found to behave in a contradictory fashion in different studies, for example: (Bryant *et al.*, 2008)-(linear), (Fierer *et al.*, 2011)-(no trend), (Singh *et al.*, 2012)-(hump-backed trend) and (Wang *et al.*, 2012)-(hollow). Elevation is discussed in more detail in the next section.

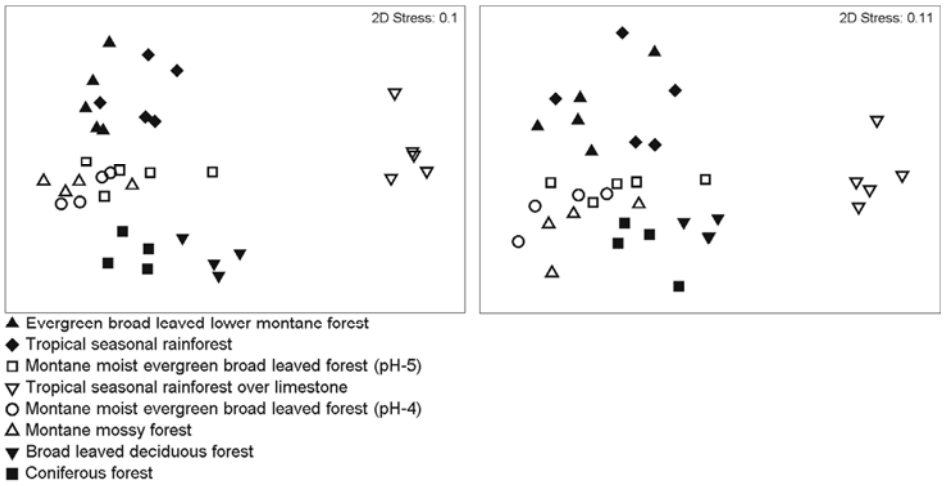

Fig. 3. Non Metric Multi Dimensional Scaling plots of forest samples. (A) Bacterial community similarity (Bray- Curtis Index; square root transformed) (B) Unweighted Pairwise UniFrac community distance.

Table 2. **Mantel test (Spearman rank correlation coefficient) between species community similarity (Bray-Curtis index) and environmental variables estimated using RELATE function in Primer v6**

Phylum	Correlation								
	Soil pH	Elevation (m.a.s.l.)	C:N Ratio	MAT (°C)	TP (g/kg)	Ca (cmol/kg)	K (cmol/kg)	Mg (cmol/kg)	Veg. Richness
Full community	0.659	0.545	0.423	0.492	0.503	0.725	0.792	0.688	0.513
Acidobacteria	0.617	0.603	0.5	0.4	0.536	0.636	0.844	0.746	0.434
Proteobacteria	0.654	0.407	0.315	0.435	0.413	0.739	0.692	0.603	0217

* All correlations were significant at $P \leq 0.01$

Environmental heterogeneity explains a majority of the variation in bacterial community composition

Visualization of the genetic matrices on NMDS (Fig. 3) was associated with large significant R values for all pairwise comparisons between the 8 different forest types (ANOSIM, 0.231–0.975, $P \leq 0.0001$, Supplementary data Table S3) and from significant differences between the limestone forest and other habitats (Fig. 3, Supplementary data Table S3). This result suggests that communities were either structured by environmental variables, geographic (spatial) distance or by a combination of both (Ramette and Tiedje, 2007). Furthermore, there was a high correlation between community similarity and environmental variables for the overall community and for the two most abundant phyla (Tables 2 and 3). This strongly suggests that environmental heterogeneity plays an important role in the BCC and diversity, rather than one edaphic factor being responsible for most of the variability found in the BCC and diversity among forest sites. These correlations (rho values as high as 0.8 or 0.9) initially seem unrealistic; however overlaying a NMDS plot of the Euclidean distances from the normalized environmental variables (Supplementary data Fig. S3) on the NMDS plots in Fig. 3, show that the actual position of the sites in relation to each other in terms of abundance and environmental datasets is very similar, and it should be no surprise that the correlations from the Mantel test are so high. Apart from this, the abundances do split very strongly into the various sites; samples from the same site are similar in abundance as well as in environmental variables.

Partial Mantel tests revealed that the similarity in the total bacterial community as well as in the 2 most abundant phyla was highly correlated with the environment (ρ=-8099, P= 0.0001) (Table 4), but not with geographical distance (Martiny *et al.*, 2011). This indicates that the geographic distance played an insignificant role in explaining community composition compared to that of the environment (Tables 2, 3, and 4). Our results here agree with studies by Horner-Devine *et al.* (2004) and Van der Gucht *et al.* (2007), who found that taxa-area relationship for bacteria in salt marsh sediments and shallow lakes water respectively, were best explained by environmental heterogeneity and not by geographic distance, and that environmental heterogeneity often increases with area.

We performed a MRM with 6 out of 9 environmental variables (see methods and Supplementary data Fig. S1) to look at the relative importance of environmental variables in explaining community similarity. For the whole community (both UniFrac and Bray-Curtis matrices), environmental distance was able to predict around 75% of the total variability (Table 5). Similar results were found at the sub-phyla level: environmental distance explained 77% and 82% of the variation in Acidobacteria for Bray-Curtis and UniFrac respectively, and over 65% for both indices in Proteobacteria (Table 5). Given that our analysis likely captures only a subset of the relevant environmental factors that determine soil bacterial community structure, it appears likely that species sorting – rather than dispersal lag – is dominant in microbial biogeography within this region (Van der Gucht *et al.*, 2007).

As expected from other studies (Lauber *et al.*, 2009; Rousk *et al.*, 2010), pH was an important variable in explaining the bacterial community composition. However, it explained a smaller amount of variation than other factors such as elevation or exchangeable calcium. The pH of our samples covers quite a wide range, from strongly acidic to neutral: ranging from 3.89 to 7.13, and it is therefore surprising that it was not such a dominant predictor of BCC. The role of soil pH in shaping the bacterial community has already been widely discussed (Lauber *et al.*, 2009; Rousk *et al.*, 2010).

In this study, soil C:N ratio was a better predictor of BCC than pH, contrary to what has been found in other studies (Lauber *et al.*, 2009). Soil pH has generally been found to be correlated with C:N ratio (Fierer and Jackson, 2006; Aciego Pietri and Brookes, 2008), which in this study explained a high proportion of the variability in the overall community similarity. Some studies have suggested an important role for C:N ratio in BCC. For example, Wardle (1992), suggested that although soil pH is probably as important as soil C and N concentrations in influencing the size of microbial biomass, the C:N ratio has been identified as a key property

Table 3. **Mantel test (Spearman rank correlation coefficient) between species community similarity (Bray-Curtis index) and environmental variables estimated using RELATE function in Primer v6**

Phylum	Correlation								
	Soil pH	Elevation (m.a.s.l.)	C:N Ratio	MAT (°C)	TP (g/kg)	Ca (cmol/kg)	K (cmol/kg)	Mg (cmol/kg)	Veg. Richness
Full community	0.787	0.812	0.699	0.819	0.683	0.761	0.744	0.721	0.785
Acidobacteria	0.698	0.461	0.401	0.429	0.471	0.742	0.741	0.643	0.287
Proteobacteria	0.646	0.338	0.338	0.371	0.368	0.727	0.608	0.559	0.173

* All correlations were significant at $P \leq 0.01$

Table 4. Correlation (Spearman ρ) between bacterial community and either geographic distance or environmental distance for all pairwise samples
Significant correlations are denoted in bold (based on 9999 permutations, $P \leq 0.0001$). 40% and 30% denote the approximate number of sequences belonging to the respective phyla out of the total quality sequences.

Correlation between:				
Bacterial community similarity (Bray-Curtis) and:	Controlling for:	Whole bacterial community	Acidobacteria (40%)	Proteobacteria (30%)
Geographic distance	Environment	-0.08	0.10	-0.06
Environmental distance	Geography	**-0.81**	**-0.85**	**-0.70**
Bacterial community distance (UniFrac) and:				
Geographic distance	Environment	-0.03	-0.20	-0.04
Environmental distance	Geography	**0.72**	**0.81**	**0.65**

Significant correlations are denoted in bold (based on 9999 permutations, $P \leq 0.0001$). 40% and 30% denote the approximate number of sequences belonging to the respective phyla out of the total quality sequences.

responsible for the difference in soil bacterial community diversity and structure among the different types of organic litter. Plant litter, which is the origin of a substantial proportion of the organic material in the soil 'B' layer, can alter the decomposition rate, especially during the earlier phases after it is incorporated into the soil, through its C:N ratio. A high C:N ratio can decrease the rate of decomposition and nutrient release, and consequently alter the microbial community (Kumar and Goh, 2000; Wang *et al*., 2004). Variation in plant litter C:N ratio might affect C:N ratio in the soil, resulting in distinctive bacterial communities.

Apart from soil pH and C:N ratio, the importance of variables such as exchangeable Ca^{2+} and K^{+} in explaining variation in BCC has not been well documented. To our knowledge, the present study is the first to report exchangeable calcium and potassium as major variables explaining bacterial community structure. In particular, calcium was very important in explaining the community similarity at both the bacterial domain and the subphyla levels. Many factors including soil pH are controlled by the role of soil calcium as a base, competing with H^{+} and Al^{3+} cations for exchange sites on soil particle surfaces, especially in the 2:1 layer type clays and within soil organic matter (Reich *et al*., 2005). Neutral conditions in soil are produced partly as a consequence of bicarbonate equilibrium. Calcium and its carbonates preserve the bicarbonate equilibrium by restraining the decrease of pH by carbonate dissolution and the increase by $CaCO_3$ precipitation. Cycles of calcium and inorganic carbon are intertwined, and the inorganic carbon cycle is linked with the organic carbon, through its labile forms such as CO_2 and HCO_3^{-} (Zavarzin, 2002). A recent study (Sridevi *et al*., 2012), which investigated the major differences in microbial diversity between calcium treated and reference watershed soils at Hubbard Brook Experimental forest (HBEF), NH, USA, reported that a Ca-amended watershed had significantly different soil chemical properties as well as different relative abundances of c. 300 bacterial taxa in the two soil horizons. They concluded that the increase in soil Ca, along with changes in other inter-related soil parameters (e.g. pH and Al), may in combination have been responsible for the observed changes in microbial populations in the Ca-amended watershed soil. Hence, the finding that calcium explains a significant portion of the BCC is not unexpected.

Elevation emerged as the most significant variable explaining the BCC in the MRM results. This reinforces our findings reported earlier for the Mt. Fuji (Singh *et al*., 2012), where elevation also emerged as the strongest variable explaining the trends in bacterial richness and diversity.

Elevation in itself is not an environmental variable, but one that is related to a range of variables that affect the ecosystem such as temperature, precipitation, and vegetation richness (see VARCLUS results, Supplementary data Fig. S1) and O_2 availability. This may explain why in this study, elevation plays such an important role in predicting BCC. The highest OTU richness is found at 750 m and then decreases with increasing elevation (Supplementary data Fig. S2). This general relationship may be related in some way to temperature itself. Earlier studies have shown that seasonal changes in temperature and vegetation, led to the replacement of dominant soil microbial groups in a wheat field (Smit *et al*., 2001) and in grasslands (Lipson and Schmidt, 2004). When MAT is plotted instead of elevation, richness also follows this trend – with greater OTU richness in warmer climate

Table 5. Results of multiple regression on matrices analysis for the total bacterial community and the two most abundant phyla

Environmental variables	Whole community		Acidobacteria		Proteobacteria	
	Bray-Curtis R^2=0.76	UniFrac R^2=0.74	Bray-Curtis R^2=0.77	UniFrac R^2=0.82	Bray-Curtis R^2=0.68	UniFrac R^2=0.65
Sqr (C:N ratio)	3.62*	-	-	-	3.70*	-
Sqr (Elevation)	-0.32***	0.002***	-0.37***	0.002***	-0.26***	0.001**
pH	-1.64*	0.03***	-2.20*	0.05***	-2.31**	0.02**
Ln (Ca^{2+})	-2.02***	0.03***	-1.24*	0.03***	-2.62***	0.03***
Sqr (K^{+})	-9.27**		-26.02***	0.14**	-	-
P	-	-	-	-	-	-

Bray-Curtis here represents the transformed and normalized community similarity and UniFrac denotes non-transformed community distance (1-similarity). R^2 is the variation that is explained by the 6 final environmental variable's partial regression coefficients, selected after VARCLUS (see materials and methods, Supplementary data Fig. S1). Total phosphorus content of soil was not found to explain any significant amount of variation at any level. Significance level was shown with ***$P \leq 0.0001$; **$P \leq 0.001$; and *$P \leq 0.01$.

sites (Supplementary data Fig. S4). We can say that elevation trends in BCC are a probable consequence of many different factors which simply mirror elevation (such as temperature, precipitation, vegetation species richness etc.,) and thus play a potential role in shaping the bacterial biodiversity.

Overlap in bacterial community composition between different forest types

All of the forest types sampled show at least some species overlap with the other sites. The number of shared OTUs ranged approximately from 8.3% to 12.9% (at the 97% sequence similarity level) out of a subset of 1,000 reads per sample in pairwise comparisons of all forest types excluding tropical seasonal rainforest over limestone (B2). On average, limestone forest showed less overlap (Fig. 3), with only 3.7% of overlapping phylotypes in comparison with other forest types (Supplementary data Table S3: ANOSIM results with large significant R values for all pairwise comparisons between the 8 different forest types). The amount of turnover between adjacent samples within a forest type (i.e., replicate samples) is comparatively higher than the overall mean between different forest sites. Within a forest type, as successive new samples are added there is a steep rise in diversity with no sign of an asymptote being reached (Supplementary data Fig. S5).

Conclusions

Overall our results suggest that while pH is important, it is not always the dominant factor predicting variation in soil bacterial diversity and community composition on a regional scale. It appears that variation in diversity and community composition is instead a more complex result of multiple factors that vary by region, ecological system and scale. This study reveals the potentially important role of elevation, C:N ratio, exchangeable Ca^{2+} and K^{+} ions in controlling the diversity and community structure of soil bacteria.

Acknowledgements

Singh. D is supported by the Korean Government Scholarship Program, Ministry of Education, Science, and Technology, South Korea.

References

Aciego Pietri, J.C. and Brookes, P.C. 2008. Relationships between soil pH and microbial properties in a UK arable soil. *Soil Biol. Biochem.* **40**, 1856–1861.

Adams, J.M. 2009. Species richness : patterns in the diversity of life. Springer in association with Praxis, Chichester.

Brown, J.H. 2001. Mammals on mountainsides: elevational patterns of diversity. *Global Ecol. Biogeogr.* **10**, 101–109.

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**, 11505–11511.

Cao, M. and Zhang, J.H. 1997. Tree species diversity of tropical forest vegetation in Xishuangbanna, SW China. *Biodivers. Conserv.* **6**, 995–1006.

Chan, O.C., Yang, X.D., Fu, Y., Feng, Z.L., Sha, L.Q., Casper, P., and Zou, X.M. 2006. 16S rRNA gene analyses of bacterial community structures in the soils of evergreen broad-leaved forests in south-west China. *FEMS Microbiol. Ecol.* **58**, 247–259.

Chu, H., Fierer, N., Lauber, C.L., Caporaso, J.G., Knight, R., and Grogan, P. 2010. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ. Microbiol.* **12**, 2998–3006.

Clarke, K.R. and Gorley, R.N. 2006. Primer v6: User Manual/ Tutorials. Primer-E Ltd: Plymouth, UK.

Curtis, T.P. and Sloan, W.T. 2005. Exploring microbial diversity - A vast below. *Science* **309**, 1331–1333.

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

Fierer, N., McCain, C.M., Meir, P., Zimmermann, M., Rapp, J.M., Silaman, M.R., and Knight, R. 2011. Microbes do not follow the elevational diversity patterns of plants and animals. *Ecology* **92**, 797–804.

Gans, J., Wolinsky, M., and Dunbar, J. 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**, 1387–1390.

Goslee, S.C. and Urban, D.L. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22**, 1–19.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., and Whiteley, A.S. 2011. The bacterial biogeography of British soils. *Environ. Microbiol.* **13**, 1642–1654.

Hamady, M., Lozupone, C., and Knight, R. 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* **4**, 17–27.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. *Am. Nat.* **163**, 192–211.

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.

Hur, M., Kim, Y., Song, H.R., Kim, J.M., Choi, Y.I., and Yi, H. 2011. Effect of genetically modified poplars on soil microbial communities during the phytoremediation of waste mine tailings. *Appl. Environ. Microbiol.* **77**, 7611–7619.

Huston, M.A. 1994. Biological diversity : the coexistence of species on changing landscapes. Cambridge University Press, Cambridge, UK.

Jesus, E.D., Marsh, T.L., Tiedje, J.M., and Moreira, F.M.D. 2009. Changes in land use alter the structure of bacterial communities in Western Amazon soils. *ISME J.* **3**, 1004–1011.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., and Chun, J. 2012. Introducing EzTaxon-e: A prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* **62**, 716–721.

Kumar, K. and Goh, K.M. 2000. Crop residues and management practices: Effects on soil quality, soil nitrogen dynamics, crop yield, and nitrogen recovery. *Adv. Agron.* **68**, 197–319.

Kunisawa, T. 2011. The phylogenetic placement of the non-phototrophic, Gram-positive thermophile '*Thermobaculum terrenum*' and branching orders within the phylum '*Chloroflexi*' inferred from gene order comparisons. *Int. J. Syst. Evol. Microbiol.* **61**, 1944–1953.

Lauber, C.L., Hamady, M., Knight, R., and Fierer, N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* **75**, 5111–5120.

Lauber, C.L., Zhou, N., Gordon, J.I., Knight, R., and Fierer, N. 2010. Effect of storage conditions on the assessment of bacterial

community structure in soil and human-associated samples. *FEMS Microbiol. Lett.* **307**, 80–86.

Legendre, P., Lapointe, F.J., and Casgrain, P. 1994. Modeling brain evolution from behavior: A permutational regression approach. *Evolution* **48**, 1487–1499.

Lipson, D.A. and Schmidt, S.K. 2004. Seasonal changes in an alpine soil bacterial community in the Colorado Rocky Mountains. *Appl. Environ. Microbiol.* **70**, 2867–2879.

Lomolino, M.V. 2001. Elevational gradients of species-density: historical and prospective views. *Global Ecol. Biogeogr.* **10**, 3–13.

Lozupone, C.A. and Knight, R. 2007. Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. USA* **104**, 11436–11440.

Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., and Horner-Devine, M.C. 2011. Drivers of bacterial beta-diversity depend on spatial scale. *Proc. Natl. Acad. Sci. USA* **108**, 7850–7854.

McCain, C.M. 2005. Elevational gradients in diversity of small mammals. *Ecology* **86**, 366–372.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.

Price, M.N., Dehal, P.S., and Arkin, A.P. 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.

Rahbek, C. 1995. The Elevational Gradient of Species Richness: A Uniform Pattern? *Ecography* **18**, 200–205.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. *Ecol. Lett.* **8**, 224–239.

Ramette, A. and Tiedje, J.M. 2007. Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc. Natl. Acad. Sci. USA* **104**, 2761–2766.

Reich, P.B., Oleksyn, J., Modrzynski, J., Mrozinski, P., Hobbie, S.E., Eissenstat, D.M., Chorover, J., Chadwick, O.A., Hale, C.M., and Tjoelker, M.G. 2005. Linking litter calcium, earthworms and soil properties: A common garden test with 14 tree species. *Ecol. Lett.* **8**, 811–818.

Roesch, L.F., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., Daroub, S.H., Camargo, F.A.O., Farmerie, W.G., and Triplett, E.W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1**, 283–290.

Rousk, J., Baath, E., Brookes, P.C., Lauber, C.L., Lozupone, C., Caporaso, J.G., Knight, R., and Fierer, N. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* **4**, 1340–1351.

Sarle, W.S. 1990. The VARCLUS Procedure. SAS/STAT User's Guide. SAS Institute, Inc., Cary, NC, USA.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., and *et al.* 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541.

Singh, D., Takahashi, K., Kim, M., Chun, J., and Adams, J.M. 2012. A hump-backed trend in bacterial diversity with elevation on mount fuji, Japan. *Microb. Ecol.* **63**, 429–437.

Smit, E., Leeflang, P., Gommans, S., van den Broek, J., van Mil, S., and Wernars, K. 2001. Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Appl. Environ. Microbiol.* **67**, 2284–2291.

Sridevi, G., Minocha, R., Turlapati, S.A., Goldfarb, K.C., Brodie, E.L., Tisa, L.S., and Minocha, S.C. 2012. Soil bacterial communities of a calcium-supplemented and a reference watershed at the Hubbard Brook Experimental Forest (HBEF), New Hampshire, USA. *FEMS Microbiol. Ecol.* **79**, 728–740.

State Forestry Administration, P.R.C.e. 1999. Forest Soil Analysis Methods LY/T, pp. 1210–1275. Standard Press of China, Beijing, China.

Torsvik, V., Goksoyr, J., and Daae, F.L. 1990. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**, 782–787.

Torsvik, V., Ovreas, L., and Thingstad, T.F. 2002. Prokaryotic diversity - Magnitude, dynamics, and controlling factors. *Science* **296**, 1064–1066.

Tripathi, B.M., Kim, M., Singh, D., Lee-Cruz, L., Lai-Hoe, A., Ainuddin, A.N., Go, R., Rahim, R.A., Husni, M.H., Chun, J., and Adams, J.M. 2012. Tropical soil bacterial communities in Malaysia: pH dominates in the equatorial tropics too. *Microb. Ecol.* **64**, 474–484.

Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., Jeppesen, E., Conde-Porcuna, J.M., Schwenk, K., Zwart, G., and *et al.* 2007. The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proc. Natl. Acad. Sci. USA* **104**, 20404–20409.

Wang, W.J., Baldocka, J.A., Dalala, R.C., and Moody, P.W. 2004. Decomposition dynamics of plant materials in relation to nitrogen availability and biochemistry determined by NMR and wet-chemical analysis. *Soil Biol. Biochem.* **36**, 2045–2058.

Wang, J.J., Soininen, J., He, J.Z., and Shen, J. 2012. Phylogenetic clustering increases with elevation for microbes. *Environ. Microbiol. Rep.* **4**, 217–226.

Wardle, D.A. 1992. A comparative-assessment of factors which influence microbial biomass carbon and nitrogen levels in soil. *Biol. Rev.* **67**, 321–358.

Willig, M.R., Kaufman, D.M., and Stevens, R.D. 2003. Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annu. Rev. Ecol. Syst.* **34**, 273–309.

Zavarzin, G.A. 2002. Microbial geochemical calcium cycle. *Microbiology* **71**, 1–17.