

Contrasting soil bacterial community structure between the phyla Acidobacteria and Proteobacteria in tropical Southeast Asian and temperate Japanese forests

Naohiko T. Miyashita*

Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

(Received 20 October 2014, accepted 26 February 2015)

Soil bacterial community structures of six dominant phyla (Acidobacteria, Proteobacteria, Verrucomicrobia, Planctomycetes, Bacteroidetes and Actinobacteria) and unclassified bacteria detected in tropical Sarawakian and temperate Japanese forests were compared based on 16S rRNA gene sequence variation. The class composition in each phylum was similar among the studied forests; however, significant heterogeneities of class frequencies were detected. Acidobacteria and Proteobacteria were the most dominant phyla in all six forests, but differed in the level of bacterial species diversity, pattern of species occurrence and association pattern of species composition with physicochemical properties in soil. Species diversity among Acidobacteria was approximately half that among Proteobacteria, based on the number of clusters and the Chao1 index, even though a similar number of sequence reads were obtained for these two phyla. In contrast, species diversity within Planctomycetes and Bacteroidetes was nearly as high as within Acidobacteria, despite many fewer sequence reads. The density of species (the number of sequence reads per cluster) correlated negatively with species diversity, and species density within Acidobacteria was approximately twice that within Proteobacteria. Although the percentage of forest-specific species was high for all bacterial groups, sampling site-specific species varied among bacterial groups, indicating limited inter-forest migration and differential movement of bacteria in forest soil. For five of the seven bacterial groups, including Acidobacteria, soil pH appeared to strongly influence species composition, but this association was not observed for Proteobacterial species. Topology of UPGMA trees and pattern of NMDS plots among the forests differed among the bacterial groups, suggesting that each bacterial group has adapted and evolved independently in each forest.

Key words: 16S rRNA, Acidobacteria, Proteobacteria, tropical forest, unified neutral theory

INTRODUCTION

Soil bacteria representing as many as 10,000 distinct species have been found to coexist in a single soil sample (Gans et al., 2005; Tringe et al., 2005; Roesch et al., 2007). The most prominent phyla represented among soil bacteria around the world are Proteobacteria and Acidobacteria (Roesch et al., 2007; Lauber et al., 2009; Nemergut et al., 2010; Köberl et al., 2011; Russo et al., 2012; Tripathi et al., 2012). However, at the species level, one of the characteristics of soil bacteria distribution is the abundance of “rare biosphere” species (Ashby et al., 2007; Elshahed et al., 2008). In other words, a

large portion of soil bacterial species exists in low numbers, while only a small number of species are present at high frequencies (Sogin et al., 2006). Thus, the community structure of soil bacteria species may be characterized by an enormous number of individual species, most of which are present at low abundance, belonging to a few major phyla.

Microorganisms in soil are considered to play essential and primary roles in organic matter decomposition, which is an important process in the terrestrial ecosystem (Schlesinger, 1997; Parton et al., 2007). The aboveground biota supplies organic carbon and other resources required for the belowground biota to decompose organic matter, and organic breakdown products are then used as nutrients for the aboveground biota, especially plants (Grime, 2001; Wardle et al., 2004). This feedback inter-

Edited by Rimantas Kodzius

* Corresponding author. E-mail: arabis@kais.kyoto-u.ac.jp

action of the flow of nutrition and energy throughout the earth's ecosystem influences the maintenance of the entire biological community structure (Wardle et al., 2004). Additional studies on the relationship between the above- and belowground biota are expected to provide enlightening and valuable knowledge related to the function and activity of soil microbial communities (Harris, 2009).

In our previous study (Miyashita et al., 2013), soil bacterial community structure was compared among five tropical rain forests in Sarawak, Malaysia and a temperate forest in Japan, and the relationship between soil bacterial community structure and aboveground forest condition was explored. Although the aboveground environmental conditions differed in these six forests with respect to temperature, rainfall and vegetation, the representation of bacterial taxa at the levels of phylum, class, order, and family was similar in all forests. As reported for other forest soils, Acidobacteria and Proteobacteria were the predominant soil bacteria in these six forests. At the species level of 3% divergence, in addition to the abundance of "rare biosphere" species (e.g., singleton OTUs), most OTUs were forest-specific, and few species were shared among the forests. Species composition was influenced by soil pH, as reported for other forests (Fierer and Jackson, 2006; Lauber et al., 2009; Nacke et al., 2011; Singh et al., 2012). This pattern of species distribution discriminated the six forests distinctly and almost equally distantly. However, the level of species diversity (the number of OTUs and Chao1 index) was comparable among forests, even though the six-forest sample included five tropical and one temperate forest environment. These results indicated that geographic locale and the aboveground forest environment have minimal impact on the taxonomic composition of the bacterial communities, at least with respect to higher taxonomic levels and relative species diversity. However, the fact that the composition of the bacterial community in each forest is unique suggested that a process of adaptation to specific forest conditions has occurred.

In our previous study (Miyashita et al., 2013), the bacterial community structure at the species level was analyzed without considering phylum classification. In this report, soil bacterial community structure (composition and diversity) was investigated separately for each of six dominant phyla (Acidobacteria, Proteobacteria, Verrucomicrobia, Planctomycetes, Bacteroidetes and Actinobacteria) and unclassified bacteria detected in our previous study. The purpose of this study was to examine if the pattern of community structure for the total bacteria described by our previous study mentioned above applies at the phylum level. For this purpose, the following analyses were conducted: (1) an evaluation of the class composition and species diversity within each of the phyla and unclassified bacteria; (2) a comparison of soil bacte-

rial community structure among phyla; and (3) an examination of the relationship between the community structure at the phylum level and aboveground environmental condition. Four primary conclusions were supported by the present study: 1) the two dominant phyla Acidobacteria and Proteobacteria in the six forests analyzed were characterized by distinct levels of species diversity; 2) a statistically significant negative relationship was observed between species diversity and species density, independent of bacterial group and independent of forest geography or location; 3) pH, C/N ratio and Mg content appeared to have significant influence on species composition of the soil bacterial communities analyzed, with some variation across bacterial groups; and 4) the topology of UPGMA trees as well as NMDS plots provided evidence that bacterial communities are more similar within than between forests, suggesting that the bacterial communities in each forest evolved independently.

MATERIALS AND METHODS

The forest soils investigated were described in our previous study (Miyashita et al., 2013). Briefly, soil samples were collected in 2009 from five tropical forests in Sarawak, Malaysia and one temperate forest in Kyoto, Japan. The Sarawakian samples were obtained from a ridge location (LR) and valley location (LV) in the Lambir Hill National Park, and remnant forest (BR), secondary forest (BS) and burned forest (BB) in the Bakam Experimental Reserve. The Japanese sample was obtained from the Ashiu Research Forest (AS) in Kyoto, Japan. Five soil samples were obtained in each of the six forests ($n = 30$). The locations of these forests and procedures of hierarchical soil sampling, microbial DNA extraction, PCR amplification of the 16S rRNA region, pyrosequencing, initial sequence data processing (quality check, length cut-off and denoising) and physicochemical analyses of soil samples were as described (Miyashita et al., 2013). In the initial data processing, reads of phred quality score lower than 25 (default for `-s` option in `split-libraries.py` tool in QIIME, <http://www.qiime.org>) and shorter than 300 bp were eliminated. By denoising (`denoise_wrapper.py` tool in QIIME, Quince et al., 2009; Caporaso et al., 2010), PCR and pyrosequencing errors were corrected. Physicochemical properties of soil from two of the five sampling sites in each forest were determined; the results are presented in Supplementary Table S1.

Data analysis In this study, the previously investigated 16S rRNA sequence data (Miyashita et al., 2013) were reanalyzed for bacterial community structure of each of six dominant phyla (Acidobacteria, Proteobacteria, Verrucomicrobia, Planctomycetes, Bacteroidetes and Actinobacteria) and unclassified bacteria. Unclassified

bacteria is a group belonging to the domain Bacteria but not assigned to any existing phylum, as communicated by the Ribosomal Database Project (RDP, Cole et al., 2009). Taxonomic composition was determined in each sample ($n = 30$) using the Classifier tool in RDP. The compositions of classes in a phylum were evaluated between different forests using Spearman's rank correlation calculated with Excel and a web tool (<http://www.geninfo.osaka-u.ac.jp/testdocs/tomocom/spea.html>). To test heterogeneity in frequency of the classes in a phylum and unclassified bacteria among forests, chi-square tests were performed with Excel using marginal frequencies of the class (or unclassified bacteria) and location to obtain the expectation. The Classifier tool generated an assignment detail file for each of the six phyla and unclassified bacteria. The assignment detail files were used as the ID list files for the Sequence selection tool in RDP, and the corresponding original sequence reads were downloaded in FASTA format. Thus, analyzing 30 samples (six forests, five samples per forest) for seven bacterial groups (six phyla and unclassified bacteria), 210 FASTA data sets were generated, 30 per bacterial group. For each bacterial group, 16S rRNA sequences were aligned using the Aligner tool of RDP, and the 30 aligned data files were used for the Complete Linkage Clustering (an RDP tool). OTUs or species clusters were defined using the conventional criterion of 3% divergence of bacterial 16S rRNA sequences (Bond et al., 1995; Schloss and Handelsman, 2005; Huber et al., 2007). Following cluster analysis, α -diversity (Chao1 index (Chao, 1984); Shannon H' index (Shannon, 1948)) and β -diversity (modified Jaccard index; Chao et al., 2005) were estimated using RDP PIPELINE tools. Pearson's correlation between the diversity estimates (the number of clusters and Chao1 index) and the number of sequence reads was tested to examine the effect of the read number on the species diversity. To compare bacterial diversity in different forests, read number was normalized randomly to the minimum number of reads among sampling sites using the Daisychopper program (<http://www.genomics.ceh.ac.uk/GeneSwytc/Tools.html>). The re-sampled data set was used to estimate the diversity and diversity indices above. The obtained OTU table was used to compare the occurrence of OTUs among the sampling sites or forests. Pairwise distance measures between sampling sites obtained by 1 - the Jaccard index (β -diversity estimates) were used to construct a dendrogram by the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method, and were subjected to nonmetric multidimensional scaling (NMDS) analysis using a QIIME tool (nmds.py). The relationships between NMDS 1st and 2nd axis values and soil physicochemical properties at 12 (of 30) sampling sites were examined using Pearson's correlation (Fierer and Jackson, 2006), and this analysis was performed for the seven bacterial groups (six phyla

and unclassified bacteria), as well as for all bacteria (Miyashita et al., 2013).

RESULTS

Classes of Acidobacteria and Proteobacteria in soil from Sarawakian and Japanese forests Of the two most dominant phyla detected in the previous study (Miyashita et al., 2013, Supplementary Fig. S1), Acidobacteria (accounting for 32.3% of all the bacteria including unclassified bacteria) was variable in the class composition among the forests and sampling sites in a forest, with a total of 20 out of 26 Acidobacteria classes in the RDP database being represented (Fig. 1A), but two classes, Acidobacteria_Gp1 and Gp2, were predominant in all the forests, especially in LR, BR and AS forests, accounting for > 90% of all Acidobacteria. In these three forests, the class composition was similar among all five sampling sites. On the other hand, in the LV, BS and BB forests, four classes (Acidobacteria_Gp5, Gp3, Gp6 and Gp4) reached notable frequencies, and variation in the class composition among sampling sites was higher in these three forests than in LR, BR and AS forests. The LV forest had 19 classes, of which Acidobacteria_Gp25, Holophagae and Acidobacteria_Gp20 were uniquely detected in this forest. Fourteen classes of Acidobacteria were detected in the AS forest in Japan, all of which were also represented in at least one of the Sarawakian forests. To compare Acidobacteria class frequencies between forests, Spearman's rank correlation coefficients were calculated for all pairwise combinations of forests. The results showed statistical significance at least at the 1% level (range = 0.66 – 0.96, average = 0.80, Supplementary Table S2) for all comparisons, indicating similarity in class composition in Acidobacteria in all forests, although the significance was probably due to the presence of the two highly dominant classes Acidobacteria_Gp1 and Gp2 in all six forests.

The phylum Proteobacteria (accounting for 29.0% of all the bacteria) contained four out of six classes in the RDP database in all the sampling sites in all six forests, among which the class Alphaproteobacteria was the most dominant in all forests (Fig. 1B). The frequency of the class Alphaproteobacteria was very high in the LR, BR and AS forests, while in the LV, BS and BB forests the class Betaproteobacteria was represented at higher frequency than in the LR, BR and AS forests. The class composition was similar among the sampling sites in the LR, BR and AS forests, while there was greater variation in the frequencies of Deltaproteobacteria and Betaproteobacteria among sampling sites in the LV, BS and BB forests. Although Spearman's correlation was not statistically appropriate due to the small number of classes (four) in Proteobacteria, the correlations among the six forests were all positive (Supplementary Table S2). In

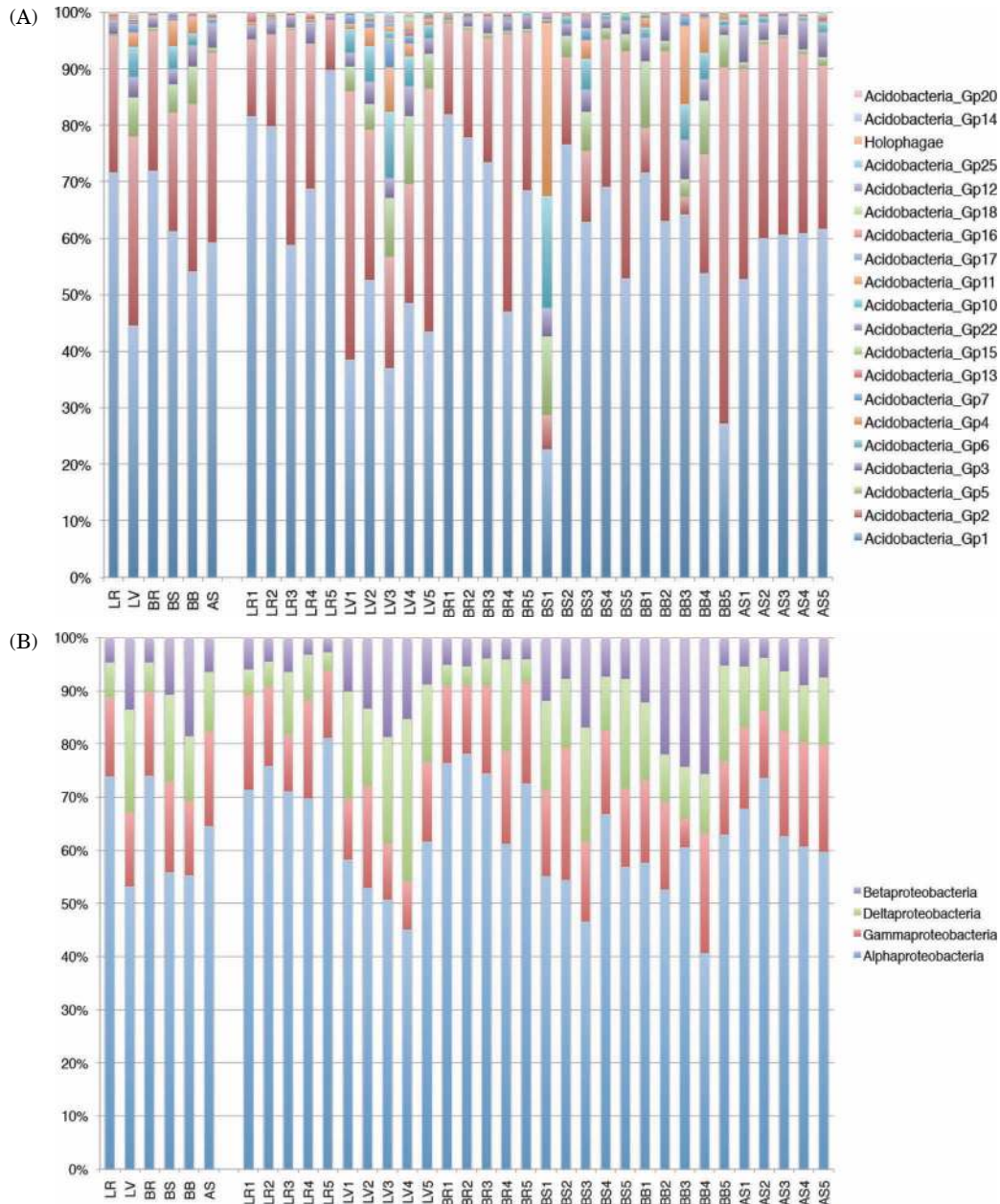


Fig. 1. Composition of classes in the phyla Acidobacteria (A) and Proteobacteria (B). Left, six columns represent the composition of classes in five pooled samples from each of six forests. Right, similar to the left panel, except that the columns represent individual data points for each of 30 sampling sites in the six forests. LR, Lambir ridge; LV, Lambir valley; BR, Bakam remnant; BS, Bakam secondary; BB, Bakam burned; AS, Ashiu.

the LV and BB forests, frequencies of Deltaproteobacteria and Betaproteobacteria were higher than in the remaining four forests (Fig. 1B). Examining the class distribution of Acidobacteria and Proteobacteria in all six forests, we found that LR, BR and AS forests grouped together and LV, BS and BB forests grouped together, because of similar class distributions in forests.

Classes of Verrucomicrobia, Planctomycetes, Bacteroidetes and Actinobacteria in soil from Sarawakian and Japanese forests For Verrucomicro-

bia, Planctomycetes, Bacteroidetes and Actinobacteria, two to four classes per phylum were represented (Supplementary Figs. S2–S5). In the phylum Verrucomicrobia (five classes listed in the RDP database), two classes, Subdivision3 and Spartobacteria, were almost equally predominant in all forests except for the BB and AS forests, where the class Opiritutae was represented at relatively high frequency (Supplementary Fig. S2). For Planctomycetes (two classes in RDP), Bacteroidetes (four classes) and Actinobacteria (five classes), a single class per phylum was predominant (Planctomycetacia, Sphin-

gobacteria and Actinobacteridae, respectively), although minor classes were detected at low but variable frequency in samples from different sampling sites and forests. Particularly in the phylum Bacteroidetes (Supplementary Fig. S4), class composition varied in different forests. For example, Bacteroidetes-incertae_sedis and Flavobacteria were lacking in LR; LV and BS contained Flavobacteria at notable frequencies; in BR, only Sphingobacteria dominated; and Bacteroidia were lacking in BB and AS (Supplementary Fig. S4).

It was noted that the number of Acidobacterial classes was much higher than for other bacterial phyla. In addition, Spearman's rank correlations for Acidobacterial classes between forests were positive and statistically significant, while similar comparisons for other classes, although positive, were mostly not significant (Supplementary Table S2). However, class frequencies in all phyla displayed significant heterogeneity by forest (1% level or lower), except for the Planctomycete class Planctomycetacia (Supplementary Table S3). Therefore, we conclude that the significant rank correlation and heterogeneity in class frequencies reported previously can be attributed mainly to Acidobacteria, instead of to all bacteria in the sampling set (Miyashita et al., 2013).

Unclassified bacteria accounted for 8.4% of the total bacteria (averaged across all forests and sites), a fraction smaller than Verrucomicrobia but larger than Planctomycetes, Bacteroidetes and Actinobacteria. These organisms comprised 6.9%–10.1% at the forest level and 4.9%–13.7% at the sampling site level (Supplementary Fig. S1). Highly significant heterogeneity in frequency of the unclassified bacteria among different forests was also detected (Supplementary Table S3).

Bacterial species diversity in the dominant phyla in forest soils in Sarawak and Japan Figure 2 shows rarefaction curves of OTUs at the 3% divergence level (species clusters) for the Acidobacteria and Proteobacteria in the six forests. These two dominant phyla clearly differed in the level of saturation: the curves for the Acidobacteria almost reached a plateau, but those for the Proteobacteria did not. Except for the unclassified bacteria, the rarefaction curves for the other four phyla did plateau, as the Acidobacteria did (Supplementary Figs. S5–S9). These results showed that the level of bacterial species diversity of the Proteobacteria and unclassified bacteria exceeded the setting of the experimental scale to estimate the exact species diversity, indicating that more sequences were needed to estimate the level of species diversity of these two bacterial groups. On the other hand, sequencing depth was sufficient to quantify bacterial species diversity for the other phyla. The unplateaued rarefaction curve for the total bacteria (Miyashita et al., 2013) was probably due to the Proteobacteria and unclassified bacteria.

Table 1 summarizes the species (OTU at the 3% divergence level) diversity estimates at different sampling levels in the six forests for each bacterial phylum and unclassified bacteria. As expected from the un-plateaued rarefaction curves of the Proteobacteria and unclassified bacteria (Fig. 2 and Supplementary Fig. S10), these two bacterial groups had a much higher level of species diversity (the number of clusters and Chao1 value) than the other phyla. The species diversity of Proteobacteria and unclassified bacteria was in the range of several thousand based on the Chao1 index, while the estimated Chao1 values of other phyla were in the range of several

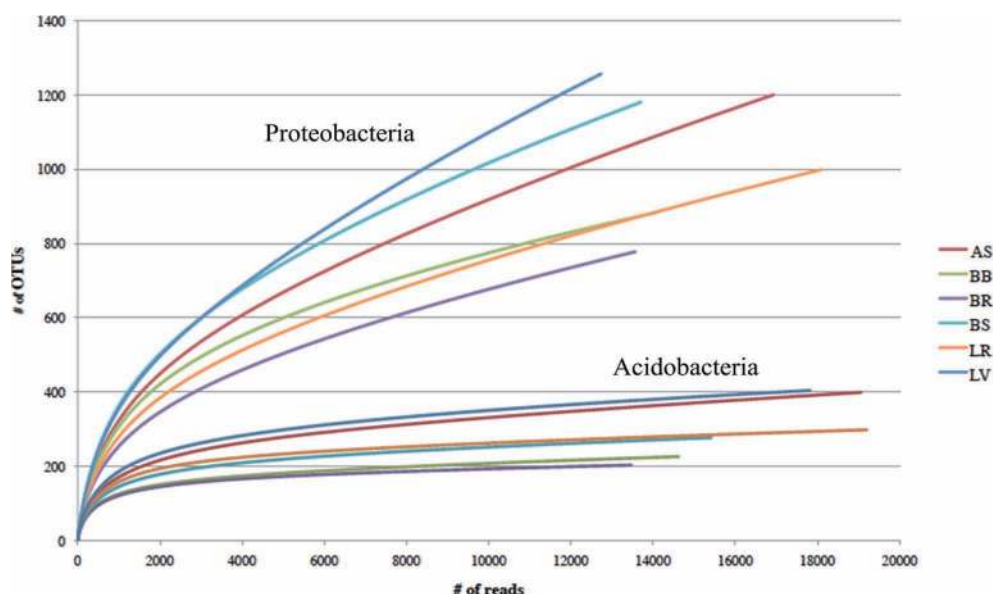


Fig. 2. Rarefaction curves of the phyla Acidobacteria (the lower six curves) and Proteobacteria (the upper six) at the 3% divergence level for each of six forests pooled over five sampling sites.

Table 1. Summary of bacterial species diversity of the six dominant phyla and unclassified bacteria

Phylum	Acidobacteria				Proteobacteria				Verrucomicrobia				Planctomycetes			
Sampling level	# of Reads	# of Clusters	ChaoI	H'	# of Reads	# of Clusters	ChaoI	H'	# of Reads	# of Clusters	ChaoI	H'	# of Reads	# of Clusters	ChaoI	H'
1) 30 sites	3320.8	247.7	311.0	4.34	2974.4	503.8	863.0	5.32	1461	140.76	192.5	3.97	534	136.9	240.3	4.39
2) 5 sites	The average of 5 sites is shown															
LR	3840.2	251	302.9	4.35	3620.4	487.4	788.4	5.12	1341.2	94	122.3	3.23	494	106.6	175.1	4.06
LV	3564.4	322	430.6	4.75	2549	551.6	1104.1	5.52	1476.8	185.6	258.7	4.52	832.2	202.2	376.4	4.87
BR	2695.2	181.6	211.6	4.05	2715.4	405.4	658.5	5.10	1271.4	106.2	149.7	3.47	567.2	131.2	196.5	4.40
BS	3083.8	224.6	279.4	4.22	2743.2	550.8	908.0	5.56	2145.6	185.8	246.3	4.36	575.8	156.4	293.4	4.52
BB	2927.8	189	229.0	3.95	2828.8	454.2	723.2	5.21	1541.6	132	185.0	3.97	462.6	122.2	186.5	4.20
AS	3813.6	318	412.4	4.70	3389.6	573.4	995.8	5.43	989.4	141	193.2	4.30	272.2	103	213.6	4.27
3) Forests																
LR	19201	405	734.1	4.66	18102	1099	2937.1	5.38	6706	185	370.3	3.38	2470	244	924.1	4.37
LV	17822	572	1381.8	4.86	12745	1408	4697.4	5.92	7384	355	1622.7	4.72	4161	483	1267.8	5.24
BR	13476	274	419.8	4.23	13577	869	2042.9	5.38	6357	199	415.3	3.68	2836	258	683.0	4.80
BS	15419	394	631.5	4.57	13716	1375	3075.1	6.06	10728	340	700.5	4.70	2879	418	1668.5	5.08
BB	14639	328	648.0	4.39	14144	1086	2599.4	5.71	7708	262	652.1	4.54	2313	308	878.0	4.98
AS	19068	516	1010.1	4.84	16948	1338	3970.7	5.70	4947	260	838.0	4.58	1361	304	2524.0	4.75
4) Total	99625	1910	3812.1	5.80	89232	5350	11891.6	6.71	43830	1130	2827.5	5.22	16020	1723	4605.8	6.34

Phylum	Bacteroidetes				Actinobacteria				Unclassified			
	# of Reads	# of Clusters	ChaoI	H'	# of Reads	# of Clusters	ChaoI	H'	# of Reads	# of Clusters	ChaoI	H'
1) 30 sites	461.7	91.3	196.8	3.69	211.5	53.6	94.4	3.50	859.6	313.7	904.2	4.99
2) 5 sites												
LR	114.6	43.8	120.9	3.35	338.6	63.6	81.7	3.76	736.2	211.4	489.1	4.57
LV	384.6	135.8	412.0	4.28	166	60.6	127.3	3.71	785.6	390.2	1244.2	5.35
BR	80.6	27.6	53.2	2.81	101.2	30.4	53.4	3.04	726.2	241	608.5	4.83
BS	374.2	96.4	182.1	3.98	65.4	36	109.3	3.23	842.4	347.8	932.5	5.25
BB	1569.6	171.6	253.7	4.00	268.2	51.2	81.8	3.28	1222.2	322.4	895.0	4.67
AS	246.4	72.8	158.7	3.74	329.8	79.6	112.8	3.96	845.2	369.2	1256.1	5.28
3) Forests												
LR	573	132	423.1	3.93	1693	111	247.9	4.01	3681	664	2202.2	5.14
LV	1923	457	1978.7	4.97	830	164	430.8	4.22	3928	1470	4801.6	6.28
BR	403	94	311.8	3.36	506	74	125.2	3.54	3631	767	2279.5	5.40
BS	1871	294	740.3	4.67	327	132	439.2	4.17	4212	1291	3775.1	6.28
BB	7848	421	945.9	4.74	1341	116	418.5	3.82	6111	1143	3680.5	5.74
AS	1232	202	700.8	4.23	1649	158	446.0	4.23	4226	1329	5402.8	6.06
4) Total	13850	1394	3680.6	5.69	6346	624	1494.4	5.35	25789	5590	15367.8	7.24

The average values over all 30 sites and over 5 sites in a forest are given in 1) 30 sites and 2) 5 sites, respectively.

hundred. Relatively higher Shannon indices (H' , a measure of the evenness of species distribution) were consistent with the presence of many Proteobacterial and unclassified bacterial species represented at very low frequency (see below).

The number of sequence reads was not positively correlated with the estimated species diversity (Table 1). The

ratio of Chao1 index to the total number of reads for the total data varied considerably among the bacterial groups (0.038 for the Acidobacteria, 0.133 for the Proteobacteria, 0.065 for the Verrucomicrobia, 0.288 for the Planctomycetes, 0.266 for the Bacteroidetes, 0.235 for the Actinobacteria and 0.596 for unclassified bacteria). Although the Acidobacteria had the highest number of sequence reads,

the level of species diversity of the Acidobacteria was approximately half that of the Proteobacteria at the sampling site and forest levels. On the other hand, despite many fewer sequence reads obtained, species diversity of Planctomycetes and Bacteroidetes was similar to that of Acidobacteria.

As detected for the total species diversity (Miyashita et al., 2013), the LV forest had higher species diversity than other forests at the sampling site and forest levels, with the exception of Planctomycetes, Actinobacteria and unclassified bacteria. Species diversity for Planctomycetes and Actinobacteria was higher at the forest level in the AS forest, and species diversity for unclassified bacteria was higher at both sampling levels in the AS forest. Species diversity for Acidobacteria and Proteobacteria was second highest in the AS forest. For the other phyla, forests having the second highest species diversity differed depending on the bacterial groups and sampling levels. For the Verrucomicrobia and Planctomycetes, the BS forest had the second highest diversity at the sampling site level, but the AS forest had higher diversity than the BS forest at the forest level. For the Bacteroidetes, the BB forest was second highest in species diversity. For the Actinobacteria, the AS forest was the highest and the BS forest was second in species diversity (Chao1 index) at the forest level. The AS forest, which is a temperate forest located on a mountain in Japan that is subject to winter snowfall, did not necessarily have lower species diversity than tropical forests in Sarawak for any of investigated bacterial groups, as might be expected from its lower level of aboveground biodiversity.

Although statistically significant correlations between the diversity estimates and the number of sequence reads over sampling sites were not detected in the previous study on the total bacteria (Miyashita et al., 2013), significant and positive correlations over sampling sites were detected at least at the 5% level in the present analysis for the individual phyla, except for the Chao1 index for the Acidobacteria and Actinobacteria (Supplementary Table S4). To eliminate the effect of read number variation on the diversity estimates among the sampling sites and forests, the diversity estimates were recalculated for the data set normalized to the minimum number of sequence reads at the sampling site level (Supplementary Table S5). The overall relationship of the species diversity among the sampling sites and forests was the same as that of non-normalized results (Table 1). In other words, the LV forest still generally had the highest species diversity for most of the bacterial groups, and the species diversity of the AS forest was comparable to those of the tropical forests in Sarawak. This result implies that read number variation among samples did not influence the relative relationship among forests and the sampling sites in a forest with respect to the species diversity for different bacterial groups.

Variation in the number of 16S rRNA sequences per OTU Based on the concerted evolution and conserved copy number variation of 16S rRNA cistrons in genomes of bacterial species (Liao, 1999; Klappenbach et al., 2001; Miyashita et al., 2013), the number of 16S rRNA sequence reads per OTU was regarded as representative of the number of individual bacterial species in the forest soils investigated (Miyashita et al., 2013). Here, for each phylum and unclassified bacteria, the average and maximum numbers of 16S rRNA sequence reads per OTU were estimated (Table 2). There was considerable variation in the estimated average number of reads per OTU among the bacterial groups. At the sampling site and forest levels, the Acidobacteria and Verrucomicrobia had higher average values, and the Proteobacteria and unclassified bacteria had lower average values. These results indicate that a bacterial group with a higher species diversity would have a lower average number of reads per species (OTU). In other words, as the number of species (diversity) increased in a bacterial group, individual species would be represented at a lower frequency (i.e., their density would decrease). This negative relationship between species diversity and density was confirmed by a significantly positive Pearson's correlation between the Chao1 and Shannon indices over the six phyla and unclassified bacteria for the total data ($r = 0.93$, $df = 5$, $P < 0.01$), since a higher Shannon index indicates a larger number of species of low frequencies. However, the maximum number of reads was highest for Acidobacteria and Verrucomicrobia, and lowest for Actinobacteria (Table 2). It seemed that the maximum number of reads was not necessarily related to species diversity.

A negative relationship between diversity and density of bacterial species was also observed among the investigated forests for some bacterial groups. In the LV and AS forests, species diversity was generally higher, and the average and maximum number of sequence reads per OTU was generally lower (Tables 1 and 2). However, over all six forests, Pearson's correlation between the Chao1 and Shannon indices was significantly positive only for Acidobacteria ($r = 0.88$, $df = 4$, $P < 0.05$) and the unclassified bacteria ($r = 0.84$, $df = 4$, $P < 0.05$), while the correlation coefficients for other phyla were positive but not significant ($r = 0.55$ for the Proteobacteria, 0.68 for the Verrucomicrobia, 0.10 for the Planctomycetes, 0.79 for the Bacteroidetes, and 0.79 for the Actinobacteria). Bacteroidetes were an exception, having lower average and maximum sequence reads in the LV and AS forests, as well as the second highest species diversity in the BB forest at the sampling site and forest levels. The BB forest had a very high average and maximum number of sequence reads, probably indicating that the fire-burned aboveground condition in the BB forest selectively modulated diversity and/or density of Bacteroidetes species.

To clarify the difference in the distribution of the read

Table 2. Average number and maximum number of 16S rRNA sequence reads per OTU at the 3% divergence level for the six phyla and unclassified bacteria at different sampling levels

		Acido bacteria	Proteo bacteria	Verruco microbia	Plancto mycetes	Bacteroidetes	Actinobacteria	Unclassified
Average number per OTU	30 sites	13.4	5.9	10.4	3.9	5.1	3.9	2.7
Average over five sites	LR	14.8	7.5	14.1	4.6	2.6	5.3	3.5
	LV	11.2	4.5	8.1	3.8	2.7	2.7	2.1
	BR	14.8	6.6	12	3.9	2.8	3.3	3
	BS	13.9	4.9	11	3.7	3.9	1.8	2.4
	BB	15.2	6.2	11	3.8	8.7	5	3.8
	AS	12.1	5.9	7	2.6	3.4	4.1	2.3
Forest	LR	47.4	16.5	36.2	10.1	4.3	15.3	5.5
	LV	31.2	9.1	20.8	8.6	4.2	5.1	2.7
	BR	49.2	15.6	31.9	11	4.3	6.8	4.7
	BS	39.1	10	31.6	6.9	6.4	2.5	3.3
	BB	44.6	13	29.4	7.5	18.6	11.6	5.3
	AS	36.9	12.7	19	4.5	6.1	10.4	3.2
Total		52.2	16.7	38.8	9.3	9.9	10.2	4.6
Maximum number of reads	30 sites	913	285	446	68	482	113	297
Average over five sites	LR	403.4	209.4	253.2	42	13.8	27.4	64.2
	LV	224.4	84	90.2	30.8	31	13.8	54.4
	BR	275.8	124.2	148	32.8	16.8	14.2	42
	BS	310	91.4	170.6	38.8	27.6	9.4	38.4
	BB	401.6	114.4	131.2	25.4	208	41.6	141.2
	AS	216.2	140.6	64.2	13.8	20	21.6	34.2
Forest	LR	1362	1118	1238	187	51	93	185
	LV	889	380	397	114	94	66	162
	BR	891	527	616	142	84	71	140
	BS	1368	381	617	124	73	31	142
	BB	1327	500	334	77	663	117	297
	AS	869	599	297	48	70	88	113
Total		4014	2449	2598	423	724	126	424

numbers per OTU among the investigated phyla, rank plots of the relative proportion of reads per OTU in the total number of reads were compared among the seven bacterial groups and the total bacteria (Fig. 3). A comparable rank plot for total bacteria from our previous analysis is also shown (Fig. 3, black dashed curve; Miyashita et al., 2013). Within each bacterial group, the percentage of reads per OTU decreased rapidly as rank increased. Rank plots indicated that in Acidobacteria, Verrucomicrobia and Bacteroidetes, predominant OTUs were present, while in the other groups they were absent (see below). The rank plot of unclassified bacteria was generally lower than those of other bacterial groups, which was consistent with the low density of OTUs in this group (Table 2).

Pattern of OTU occurrence among forests in different phyla and unclassified bacteria In addition to the predominance of singleton OTUs in forest soils for the total bacteria, we showed that approximately 30% of the OTUs were uniquely represented in one sampling site, and that > 50% were uniquely represented in one forest (Miyashita et al., 2013). Here, this analysis was extended as follows (Table 3). Acidobacteria and Verrucomicrobia, which had low species diversity (Table 1), had lower percentages of singleton OTUs (31.3% and 42.8% for the pooled data, respectively) than that for the total bacteria (48.2%, Miyashita et al., 2013). On the other hand, the other bacterial groups, including Proteobacteria, had higher percentages of singletons than the total bacteria. Proteobacteria, with high species diversity, had a higher percentage of singleton OTUs than Acido-

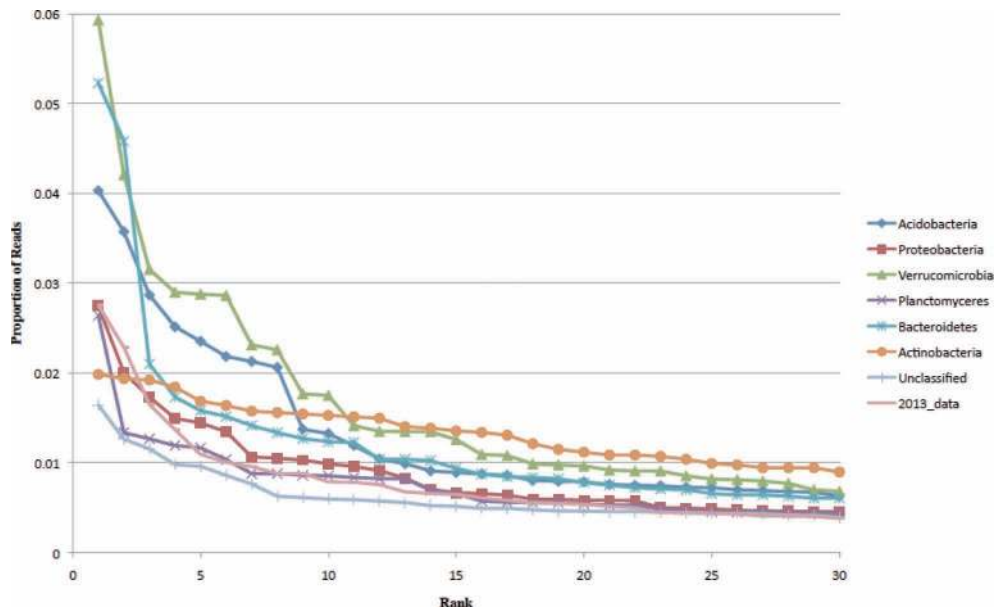


Fig. 3. Rank plot against the proportion of sequence reads in the total number of reads for the top 30 ranked OTUs for the six phyla and unclassified bacteria. The plot for the total bacteria analyzed in Miyashita et al. (2013) was included.

bacteria, with low species diversity. However, Pearson's correlation between the Chao1 index and the proportion of singleton OTUs was positive but not significant over the seven bacterial groups for the pooled data ($r = 0.56$, $df = 5$, $P > 0.05$). In most bacterial groups, the percentage of singletons was highest in the LV and AS forests, which had higher species diversity (Table 1). There was a positive relationship between species diversity and the percentage of singleton OTUs with respect to both the bacterial groups and forests. The Bacteroidetes was exceptional, in that fewer singleton OTUs were detected in the BB forest, where species diversity and the average number of reads per OTU were higher; this confirmed the above-noted unusual species diversity and density of Bacteroidetes in the BB forest.

The percentage of sampling site-specific OTUs in a forest varied by bacterial group (Table 3), being 10.3% and 17.6% on average for Acidobacteria and Verrucomicrobia, respectively, approximately 30% on average for Proteobacteria and for the total bacteria (Miyashita et al., 2013), and higher for Bacteroidetes and unclassified bacteria. However, the percentage of forest-specific OTUs (approximately 80%) was much higher for all the bacterial groups than that of sampling site-specific OTUs in a forest. The percentage of forest-specific OTUs was less variable among the bacterial groups.

The percentage of OTUs shared in multiple sampling sites in a forest also varied among bacterial groups (Table 3), and was negatively related to the percentage of site-specific OTUs. Acidobacteria and Verrucomicrobia had relatively higher percentages (on average 28.1% and 23.8%, respectively) of shared OTUs among five sites in

one forest, while the other bacterial groups, especially Bacteroidetes and unclassified bacteria, had lower percentages (on average 7.2% and 3.8%, respectively) of OTUs detected among five sampling sites. The number of shared OTUs among all 30 sampling sites or among the six forests was very low and fewer than 1% of OTUs were detected in all six forests or all Sarawakian forests. Although the AS forest in Japan was geographically distant from the five Sarawakian forests, over 20% of OTUs of all the bacterial groups detected in the AS forest (1 - the percentage of the AS forest-specific clusters) were shared with at least one of the Sarawakian forests. Actually, the corresponding percentages for the five Sarawakian forests (1 - the percentage of forest-specific clusters) were almost at the same level for all the bacterial groups, implying that the six investigated forests are equally separated in terms of the percentage of shared OTUs, despite the large difference in the geographical distance between them (see below). The percentage of OTUs in any one forest shared with the other five forests (1 - the percentage of forest-specific clusters) was higher than that shared among sampling sites in a single forest for most bacterial groups.

Relative divergence of soil bacteria communities at the phylum level The relationship among soil bacterial communities at the 30 sampling sites in this study was evaluated by constructing UPGMA trees (Fig. 4 and Supplementary Figs. S11–S15) or NMDS plots (Fig. 5 and Supplementary Figs. S16–S20), which were based on the modified Jaccard index. UPGMA trees revealed well-separated clusters corresponding to each forest, consis-

Table 3. Summary of OTU occurrence among the sampling sites and forests for each phylum and unclassified bacteria

	forest	Acido -bacteria	Proteo -bacteria	Verruco -microbia	Plancto -mycetes	Bacteroidetes	Actino -bacteria	Unclassified
Singleton OTUs (%)	LR	96 (23.7)	572 (52.0)	77 (41.6)	138 (56.6)	87 (65.9)	37 (33.3)	454 (68.4)
	LV	193 (33.7)	839 (59.6)	167 (47.0)	252 (52.2)	331 (72.4)	98 (59.8)	1086 (73.9)
	BR	54 (19.7)	414 (47.6)	75 (37.7)	101 (39.1)	66 (70.2)	32 (43.2)	513 (66.9)
	BS	95 (24.1)	685 (49.8)	104 (30.6)	245 (58.6)	169 (57.5)	96 (72.7)	869 (67.3)
	BB	80 (24.4)	489 (45.0)	79 (30.2)	151 (49.0)	162 (38.5)	55 (47.4)	723 (63.3)
	AS	154 (29.8)	740 (55.3)	102 (39.2)	221 (72.7)	134 (66.3)	72 (45.6)	997 (75.0)
	Pooled	598 (31.3)	2714 (50.7)	484 (42.8)	892 (51.8)	803 (57.6)	315 (50.5)	3676 (65.8)
Site-specific OTUs in each forest (%) (average over five sites)	LR	23.2 (9.2)	127.6 (26.2)	17.8 (18.9)	29.4 (27.6)	19 (43.4)	8.4 (13.2)	103 (48.7)
	LV	44.4 (13.8)	180.8 (32.8)	36.2 (19.5)	56.2 (27.8)	73.2 (53.9)	21.4 (35.3)	243.4 (62.4)
	BR	14.2 (7.8)	93.6 (23.1)	17.2 (16.2)	22.4 (17.1)	15.4 (55.8)	8.2 (27.0)	115.2 (47.8)
	BS	27 (12.0)	164.6 (29.9)	25.8 (13.9)	57 (36.4)	43 (44.6)	21.2 (58.9)	210.2 (60.4)
	BB	20.2 (10.7)	121.4 (26.7)	21.6 (16.4)	36 (29.5)	44.6 (26.0)	12.2 (23.8)	173.8 (53.9)
	AS	36 (11.3)	163.4 (28.5)	22.2 (15.7)	45.8 (44.5)	28.8 (39.6)	15.4 (19.3)	216 (58.5)
Forest-specific OTUs (%)	LR	254 (62.7)	615 (56.0)	99 (53.5)	182 (74.6)	96 (72.7)	78 (70.3)	458 (69.0)
	LV	418 (73.1)	856 (50.8)	209 (58.9)	373 (77.2)	365 (79.9)	124 (75.6)	1051 (71.5)
	Lambir	684 (72.8)	1520 (64.8)	314 (60.2)	559 (79.1)	464 (80.3)	203 (75.7)	1529 (73.9)
	BR	113 (41.2)	398 (45.8)	75 (37.7)	189 (73.3)	65 (69.1)	44 (59.5)	490 (63.9)
	BS	211 (53.6)	788 (57.3)	170 (50)	321 (76.8)	205 (69.7)	87 (65.9)	883 (68.4)
	BB	172 (52.4)	636 (58.6)	131 (50)	209 (67.9)	344 (81.7)	86 (74.1)	855 (74.8)
	Bakam	554 (69.4)	1996 (71.0)	424 (66.1)	762 (84.3)	642 (85.3)	227 (76.4)	2359 (80.9)
	AS	378 (73.3)	874 (65.3)	155 (59.6)	229 (75.3)	159 (78.7)	113 (71.5)	1053 (79.2)
	Pooled	1546 (80.9)	4167 (77.9)	839 (74.2)	1503 (87.2)	1234 (88.5)	532 (85.3)	4790 (85.7)
Shared among 30 sites (%)		1 (0.05)	6 (0.11)	1 (0.09)	0	0	0	0
Shared among five sites in each forest	LR	122 (30.1)	201 (18.3)	50 (27.0)	42 (17.2)	8 (6.1)	31 (27.9)	55 (8.3)
	LV	149 (26.0)	187 (13.3)	97 (27.3)	56 (11.6)	23 (5.0)	17 (10.4)	35 (2.4)
	BR	97 (35.4)	171 (19.7)	52 (26.1)	50 (19.4)	4 (4.3)	8 (10.8)	46 (6.0)
	BS	94 (23.9)	157 (11.4)	73 (21.5)	40 (9.6)	21 (7.4)	3 (2.3)	21 (1.6)
	BB	61 (18.6)	124 (11.4)	43 (16.4)	12 (3.9)	36 (8.6)	13 (11.2)	9 (0.8)
	AS	179 (34.7)	250 (18.6)	64 (24.6)	32 (10.5)	24 (11.9)	37 (23.4)	50 (3.8)
	Average	117 (28.1)	181.7 (15.5)	63.2 (23.8)	38.7 (12.0)	19.3 (7.2)	18.2 (14.3)	36 (3.8)
Shared among six forests (%)		6 (0.3)	21 (0.4)	4 (0.4)	0	0	0	1 (0.02)
Shared among 5 Sarawak forests		9 (0.6)	26 (0.6)	7 (0.7)	1 (0.07)	1 (0.08)	0	2 (0.04)
Shared between AS (Japan) & Sarawak		138 (26.7)	464 (34.7)	105 (40.4)	75 (24.7)	43 (21.3)	45 (28.5)	276 (20.8)

The number of OTUs is shown, with the percentage in parenthesis.

tent with a lower percentage of site-specific OTUs than that of forest-specific OTUs and a higher percentage of shared OTUs among sampling sites within one forest than that shared among forests (Table 3). External branches were longer in UPGMA trees for Planctomy-

cetes, Bacteroidetes, Actinobacteria and unclassified bacteria than for Acidobacteria, Proteobacteria and Verrucomicrobia, as expected from a lower percentage of shared OTUs in the less dominant phyla (Table 3), indicating that OTU composition among sampling sites for

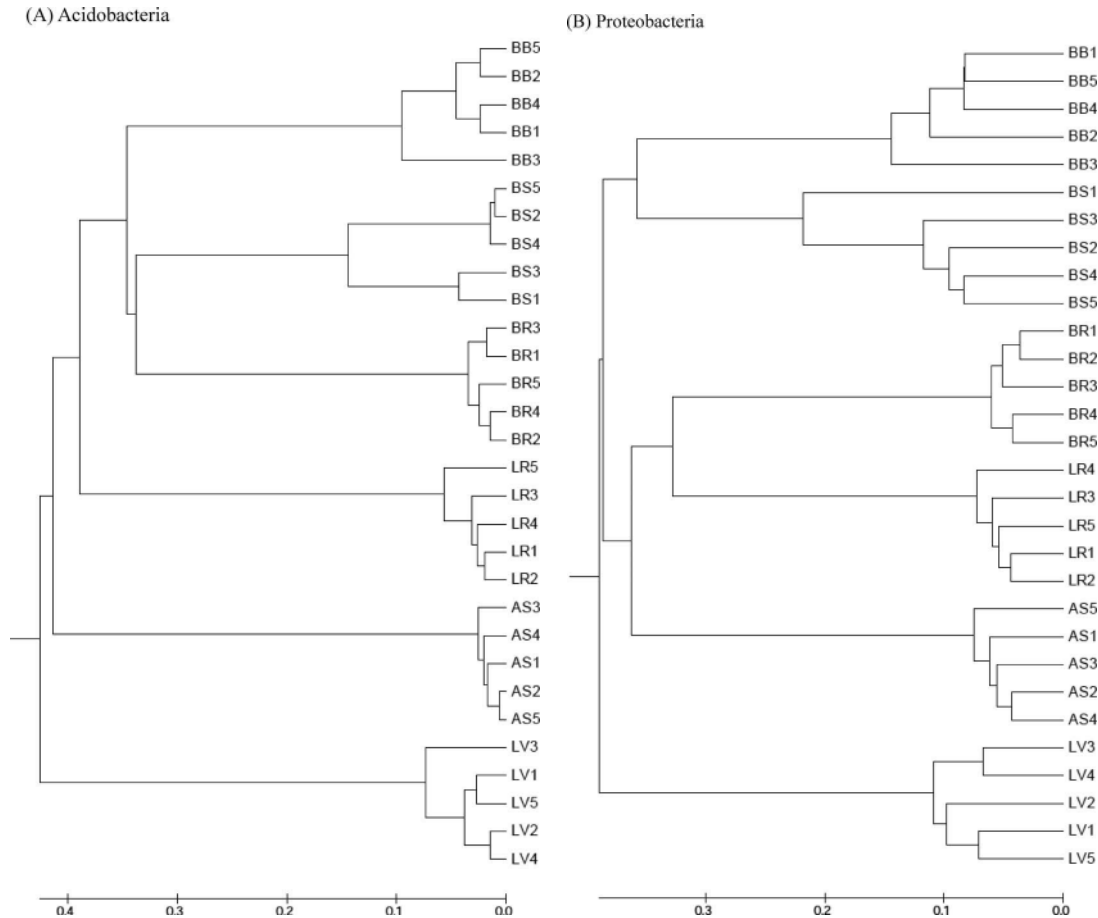


Fig. 4. UPGMA dendrograms for the 30 sampling sites in six forests, constructed using the Jaccard index calculated from data at the 3% divergence level. (A) Acidobacteria; (B) Proteobacteria.

the former bacterial groups was more divergent than for the latter three dominant phyla. The topology of UPGMA trees differed among the seven bacterial groups. For Acidobacteria and Proteobacteria, the LV forest was located at the outer edge of the dendrogram, although for the other bacterial groups the LV forest was more internal. For Acidobacteria, the BR, BS and BB forests formed a cluster, while for Proteobacteria, BR did not cluster with BS or BB. The topologies of the UPGMA trees for Verrucomicrobia and total bacteria were identical (Miyashita et al., 2013), while those for other bacterial groups were distinct from that of total bacteria. For Planctomycetes, the AS forest was isolated from the Sarawakian forests, but for the remaining five bacterial phyla and for unclassified bacteria, the UPGMA trees did not reveal separation between Sarawakian forests and the AS forest, despite the latter's greater geographic isolation.

In the NMDS plots, sampling sites from one forest clustered together, consistent with lower divergence within a forest than between forests. The relative position and distance between forests differed for different bacterial groups, as noted in the UPGMA trees. For Acidobacteria

and Proteobacteria, the LR, LV, BR and AS forests were well separated, while the BS and BB forests were close together. In all bacterial groups, the LR forest was associated with negative NMDS1 axis values and the LV forest with mostly positive NMDS1 axis values, indicating a large difference in OTU composition between these two geographically proximate forests. The three Bakam forests, BR, BS and BB, were positioned close to each other in NMDS plots for Acidobacteria, Proteobacteria, Verrucomicrobia and Planctomycetes, but were relatively far from each other for other bacterial groups, with values for the BR forest restricted to the 3rd quadrant. Although the AS forest was an outer branch of the UPGMA tree only for Planctomycetes, it was furthest from all other clusters in the NMDS plots for Acidobacteria, Verrucomicrobia and Planctomycetes. In summary, the UPGMA trees and NMDS plots provided evidence that bacterial communities from different sampling sites in a single forest are more similar to each other than those in different forests. However, no clear relationship between bacterial community (OTU composition) and geographical location was detected. Therefore, the bacterial communities of each bacterial group most likely evolved independently

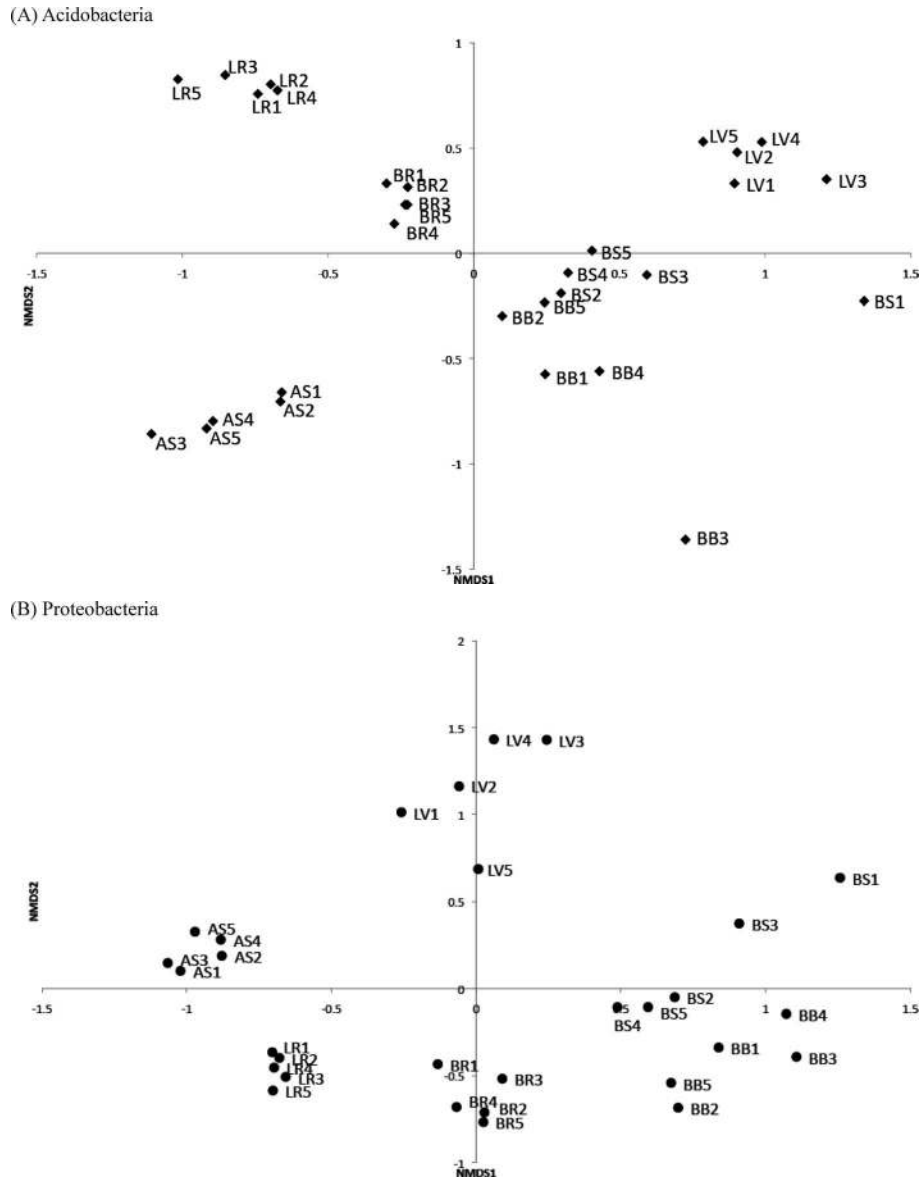


Fig. 5. NMDS plots for the 30 sampling sites in six forests, constructed using the Jaccard index calculated from data at the 3% divergence level. (A) Acidobacteria; (B) Proteobacteria.

in the forests investigated here.

Influence of physicochemical properties of forest soil on bacterial communities at the phylum level

As done before (Fierer and Jackson, 2006), to evaluate the effect of environmental factors on bacterial communities, Pearson's correlation coefficients between the NMDS 1st and 2nd axis values and 14 physicochemical properties of the investigated forest soils (Supplementary Table 1) were examined for each bacterial group (Table 4). For the total bacteria (2013_total in Table 4), only pH and C/N ratio were significantly associated with the NMDS 1st axis value, which had the primary importance. However, the pattern of association between NMDS axis values and

physicochemical properties was complex for each of the bacterial groups. With respect to the association with the NMDS 1st axis value, Acidobacteria and Verrucomicrobia showed significant correlation for pH (positive), C/N ratio (negative) and Mg content (positive), although silt (%) was also significant for Verrucomicrobia. The other bacterial groups with positive and significant association with pH, Planctmycetes, Bacteroidetes and unclassified bacteria, also had a similar pattern of association with pH (positive), C/N ratio (negative) and Mg content (positive), although the correlation coefficients were not always significant. Since correlation coefficients among the physicochemical properties Mg content, C/N ratio and pH were all significant for the present data set (Supplemen-

Table 4. Correlation coefficient between NMDS 1st and 2nd axis values and physicochemical properties of forest soil for 12 sampling sites in the six forests

Phylum	N M D S	Physicochemical property													
		Ca	Mg	K	Na	C	N	C/N ratio	Clay	Silt	Fine Sand	Coarse Sand	Sand total	pH	H ₂ O
Acidobacteria	1	0.28	0.62*	0.36	0.28	-0.48	-0.44	-0.81**	-0.15	0.42	0.17	-0.24	-0.07	0.88**	-0.34
	2	-0.30	0.00	-0.24	0.37	-0.61	-0.61*	-0.26	-0.54	-0.36	0.15	0.66 *	0.52	-0.11	-0.64*
Proteobacteria	1	-0.36	-0.06	-0.28	-0.13	-0.65*	-0.64*	-0.33	-0.53	-0.04	0.62*	0.08	0.37	0.45	-0.52
	2	0.77**	0.87**	0.81**	0.49	0.07	0.12	-0.80*	0.44	0.68*	-0.48	-0.48	-0.57	0.72**	0.12
Verrucomicrobia	1	0.48	0.71**	0.57	0.15	-0.13	-0.09	-0.73**	0.19	0.68*	-0.04	-0.57	-0.40	0.90**	0.03
	2	0.86**	0.61*	0.85**	0.22	0.80**	0.83**	-0.21	0.90**	0.71**	-0.81**	-0.71**	-0.90**	0.16	0.76**
Planctomycetes	1	0.28	0.47	0.29	0.06	-0.24	-0.22	-0.55	-0.04	0.38	0.26	-0.40	-0.13	0.83**	-0.13
	2	0.76**	0.35	0.64*	0.21	0.83**	0.85**	0.06	0.84**	0.42	-0.73**	-0.54	-0.74**	0.00	0.72**
Bacteroidetes	1	0.42	0.51	0.38	0.07	-0.05	-0.02	-0.59*	0.23	0.52	-0.14	-0.45	-0.37	0.73**	0.02
	2	0.33	-0.05	0.22	0.11	0.60*	0.59*	0.28	0.64*	0.07	-0.77**	-0.09	-0.46	-0.39	0.49
Actinobacteria	1	0.20	0.25	0.16	0.04	-0.13	-0.11	-0.44	0.15	0.33	-0.15	-0.23	-0.23	0.47	-0.07
	2	0.19	-0.04	0.16	-0.03	0.47	0.47	0.10	0.51	0.19	-0.66*	-0.11	-0.42	-0.45	0.41
Unclassified	1	0.30	0.53	0.30	0.21	-0.36	-0.33	-0.67*	-0.05	0.35	0.12	-0.25	-0.10	0.83**	-0.26
	2	-0.19	-0.41	-0.46	-0.22	0.07	0.05	0.43	-0.09	-0.53	0.07	0.36	0.28	-0.37	-0.10
2013_total	1	0.13	0.38	0.20	0.03	-0.38	-0.35	-0.59*	-0.14	0.38	0.27	-0.31	-0.06	0.78**	-0.23
	2	-0.89**	-0.75**	-0.93**	-0.31	-0.62*	-0.66*	0.40	-0.85**	-0.82**	0.77**	0.76**	0.91**	-0.43	-0.66*

Degrees of freedom =10, significance value at the 5% level = 0.58 and at the 1% level = 0.71.

tary Table S6), it was not possible to distinguish the effect of any one of these properties from the others on the OTU composition of these bacterial groups.

On the other hand, Proteobacteria and Actinobacteria did not have a significant association between pH and the NMDS 1st axis value (Table 4). Actinobacteria showed no significant association with any of the 14 physicochemical properties examined. Proteobacteria had a distinct pattern of association between the NMDS 1st axis value and physicochemical properties of forest soil; significant correlations were detected for C, N and fine sand percentages. Since the correlation coefficients among C, N and fine sand percentages were also all significant (Supplementary Table S6), it was not possible to discriminate the effect of any one property from the others on the OTU composition of the Proteobacteria.

With respect to the NMDS 2nd axis values, Verrucomicrobia and Planctomycetes showed significant correlations for 11 and eight out of 14 properties, respectively, and the direction of all the correlation coefficients was the same, except for C/N ratio and pH, for which significance was not detected for either phylum. This result indicated that these two phyla were influenced more similarly by soil physicochemical properties among the bacterial groups affected by pH noted above. However, the pattern of association of the NMDS 2nd axis values with physicochemical properties of forest soil was too complex among bacterial groups to support any clear conclusions.

DISCUSSION

Bacterial class composition in the six dominant phyla in tropical forests in Sarawak and a temperate forest in Japan It was known that Acidobacteria, especially in soils, is ubiquitous and diverse, reflecting many characteristics in physiology and geochemistry (Barns et al., 1999; Cole et al., 2005). The class composition in Acidobacteria in the investigated forest soils was very complex, comprising a total of 20 classes, and was variable across sampling sites and forests. In the LV, BS and BB forests, the class composition of Acidobacteria was more variable among the sampling sites than that in the other three forests, LR, BR and AS (Fig. 1A), thus delineating two groups of forests.

On the other hand, for Proteobacteria, four classes out of six in the RDP database were detected in soils of the studied forests; the other two classes of the Proteobacteria, Epsilonproteobacteria and Zetaproteobacteria, which are associated with animal guts and deep-sea environments (Inagaki et al., 2003; Emerson et al., 2007; Dorer et al., 2009; Yarza et al., 2010), were not detected. Proteobacteria was less variable in class composition among the sampling sites and forests than Acidobacteria (Fig. 1B). Classes Alphaproteobacteria and Betaproteobacteria, whose frequencies were negatively correlated, separated the same two groups of studied forests as for Acidobacteria (Fig. 5B). In the other four less dominant phyla, although one or two dominant classes reached a high fre-

quency, the class composition varied among the sampling sites and forests, especially in the Bacteroidetes.

Spearman's rank correlations for classes in Acidobacteria between forests were all significantly positive and no drastic difference in class composition was detected in the other phyla. It might be concluded that for the six phyla examined the class composition was generally similar among the soils in the tropical forests in Sarawak and a temperate forest in Japan, which differ in their above-ground environmental conditions, as detected for the total bacteria (Miyashita et al., 2013). However, highly significant heterogeneity of class frequency among the six forests suggested that frequencies of the classes in a phylum were influenced by unknown forest-specific factors, such as differences in unnoticed contemporary aboveground environmental conditions and evolutionary history in belowground biota.

Bacterial community structure at the phylum level in tropical forests in Sarawak and a temperate forest in Japan The two dominant phyla Acidobacteria and Proteobacteria differed in their level of species diversity in the investigated forests (Table 1), as indicated by the saturation levels in the rarefaction curves (Fig. 2). Despite the similar number of sequence reads obtained for the two phyla, species diversity of Proteobacteria was approximately 3-fold higher than that of Acidobacteria for the total sample. For unclassified bacteria, species diversity was comparable to that of Proteobacteria, while the number of sequence reads per cluster was much lower. The number of sequence reads of Planctomycetes and Bacteroidetes was much smaller than that of Acidobacteria, but the species diversity estimates of the two phyla were almost at the same level as Acidobacteria. Species diversity is independent of the number of sequence reads obtained, and may be a characteristic associated with a particular phylum.

Acidobacteria and Proteobacteria also differed in the evenness of species distribution. Shannon indices of the Proteobacteria and Acidobacteria (Table 1) indicated that Proteobacterial species were present in lower abundance than Acidobacterial species (Table 2, Fig. 3). On the whole, a negative relationship between species diversity and density was observed, implying that as species diversity (the number of species) in a phylum increases, individual species of the phylum tend to be present in lower abundance. This result may be relevant to the unified neutral theory of biodiversity (Hubbell, 2001), which characterized the relationship between species diversity and relative abundance in biological communities. However, additional studies involving a greater diversity of bacteria at the phylum level and more geographically distinct forest environments are necessary. Only after such studies are complete will it be possible to evaluate the general applicability of the theory with regard to commu-

nities of soil bacteria.

Furthermore, this negative relationship between species diversity and density seemed to be present among the investigated forests. The forests with higher species diversity had lower species density for the Acidobacteria and unclassified bacteria, for which a statistically significant correlation between the Chao1 and Shannon indices was detected. To confirm the negative relationship at the forest level for the other phyla, the number of forests examined should be increased in future study.

The above results also suggested that both biological and environmental factors influence the balance between species diversity and density in soil bacterial communities. In other words, life-history characteristics inherent to a particular bacterial phylum/group, including rate of cell division, motility, and birth-death rate, may play a role in determining species diversity and density within that bacterial group. At the same time, forest-specific quality and availability of environmental resources could also influence diversity and density to a similar extent, and could do so in a positive or negative manner (Hughes Martiny et al., 2006).

Pattern of OTU occurrence at the phylum level and propensity and capacity for bacterial movement in forest soil Singleton OTUs were abundant in all the bacterial groups in the investigated forests in Sarawak and Japan, but there was variation in the percentage of singleton OTUs among the bacterial groups. Proteobacteria with high species diversity had a higher percentage of singleton OTUs than Acidobacteria with low species diversity. Pearson's correlation between the species diversity and the proportion of singleton OTUs was positive over the bacterial groups, but not statistically significant. This non-significant correlation suggested that the percentage of singleton OTUs was not the main cause of the negative correlation between species diversity and density. The rank plot of Proteobacteria was kept lower than that of Acidobacteria. These results suggested that there is variation in the proportion of "rare biosphere" species among the phyla, although the proportion of "rare biosphere" species in soil is indeed high (Ashby et al., 2007; Elshahed et al., 2008).

In this study, sampling site-specific OTUs were represented at lower and more variable frequency than forest-specific OTUs, which were represented at a similar and higher frequency in all bacterial groups (Table 3); this indicated greater similarity in bacterial community structure within a forest than between forests, suggesting local adaptation to a specific forest soil environment. Correspondingly, the percentage of shared OTUs among sampling sites in a forest was higher than that among all the investigated forests, which was almost nil, indicating that ubiquitous OTUs were rare, and varied among the forests and bacterial groups. These results suggested that

movement of bacterial species varies according to the local forest environment and the bacterial group. Bacterial movement through soil is strongly influenced by physicochemical properties of the soil, especially soil texture, pH, water content and ionic strength (Abu-Ashour et al., 1994; Suni and Romantschuk, 2004). The variation in the percentage of shared OTUs among sampling sites in a forest for a particular bacterial group may be related to differences in the physicochemical properties of the forest soil. The explanation for the difference in the percentage of sampling site-specific OTUs between the bacterial groups could be associated with the mobility of bacteria, for example flagella-driven motility. Although it has not been proved that such mobility occurs generally in natural soils (Stotzky, 1986), the difference in the number, shape and movement of flagella that exists among bacterial groups may account for the variation in motility among bacterial groups in a forest, especially in water-saturated conditions (Harshey, 2003).

On the other hand, the percentage of OTUs in a particular forest shared with any of the other five forests (1- the percentage of forest-specific clusters) was higher than that shared among sampling sites in a forest for most of the bacterial groups. This result seemed to be contradictory and suggested that occasional long-distance dispersal between pairs of investigated forests occurred more frequently than dispersal within a forest. Some form of occasional long-distance dispersal may not be inconsistent with the fact that geographical distance between forests had no effect on the percentage of OTUs shared by one forest with the other five forests; the six forests in Sarawak and Japan were equally separated in terms of their OTU composition. This idea has been formalized into a hypothesis stating that “everything is everywhere, but the environment selects” (de Wit and Bouvier, 2006; O'Malley, 2008). In the present study, the extremely low frequency of ubiquitous OTUs (< 1%) argues against this hypothesis. Finally, variation among bacterial groups in the percentages of OTUs in a forest shared with the other five forests suggested that the mechanisms (or the efficiency of those mechanisms) for long-distance dispersal vary from one bacterial group to another. It should be stated here that a discussion on the dispersal of bacteria between forests based on an analysis of relatively short sequence reads (ca. 500 bp) and OTUs with 97% identity, as employed in this study, may be too speculative, since the exact determination of completely identical species and recent descendants from a single common bacterial strain would be necessary to investigate long-distance dispersal of bacterial species.

Relative divergence of bacterial communities among Sarawakian and Japanese forest soils and influence of physicochemical properties of forest soil on OTU composition at the phylum level The

UPGMA trees and NMDS plots indicated that for all the bacterial groups the sampling sites within a forest were relatively closely grouped, compared to the divergence between investigated forests, consistent with the close proximity of the sampling sites in a forest and bacterial movement in soil discussed above. These results indicated that the spatial distribution of bacteria was not random, suggesting the influence of contemporary differences in forest environments. On the other hand, the relative relationship among the investigated forests (topology of the UPGMA trees and scattering pattern in the NMDS plots) differed among the bacterial groups, and none of the trees or plots were the same. There was no clear pattern of association with either geographical origin or aboveground environmental conditions. In other words, forests of the three regions (two locations in Sarawak and one in Japan) were intermixed, and the AS forest in Sarawak in terms of bacterial species composition. Therefore, it could be concluded that the soil bacterial communities examined here evolved independently from each other, in a manner independent of relative geographical distance and aboveground environmental conditions.

There was difference in the pattern of association between physicochemical properties of forest soil and OTU composition among the bacterial groups. Although pH is known to be an important distinguishing factor among total bacterial communities (Fierer and Jackson, 2006; Jesus et al., 2009; Lauber et al., 2009), Proteobacteria did not, here, show a statistically significant association with pH, as observed before (Lauber et al., 2009), whereas five out of the seven bacterial groups, including Acidobacteria, were associated with pH (Table 4). The bacterial groups associated positively with pH also tended to show strong associations with C/N ratio and Mg content. On the other hand, Proteobacteria was associated with C and N content (%) and fine sand percentage. Thus, the bacterial groups investigated could be delineated with respect to the relationship with pH. In this study, Actinobacteria did not show a significant association with any of 14 physicochemical properties examined. This result contrasted with some previous studies that showed an association between the community structure of Actinobacteria and pH (Lauber et al., 2009; Tripathi et al., 2012). Remarkably, Na content did not show any significant association with NMDS 1st and 2nd axis values for any bacterial group, despite the larger variance (0.33) in %Na than in pH (0.06) in this data set (Supplementary Table S1). In contrast, other studies have suggested that salinity has a significant influence on the global distribution of bacteria (Lozupone and Knight, 2007; Tamames et al., 2010). To resolve these discrepancies, we have begun to collect samples and analyze bacterial community structures in forest soils that are widely distributed in Southeast Asia. In addition, since bacterial

communities seem to evolve independently of above-ground environmental conditions, as shown here, to understand more fully the relationship between above- and belowground biota, it will be necessary to analyze community structures for microbes other than bacteria, such as fungi, which play an important and primary role in the degradation of organic matter (van der Heijden et al., 1998; Wardle et al., 2004; Lauber et al., 2008).

The author expresses his thanks to Drs. K. R. Takahashi and H. Iwanaga for their comments. He also thanks Dr. Benli Chai at RDP for advice and help in using analysis tools. This work was partly supported by a grant (D-0901) from the Ministry of the Environment of Japan. This paper is contribution #612 from the Laboratory of Plant Genetics, Kyoto University, Japan.

REFERENCES

- Abu-Ashour, J., Joy, D. M., Lee, H., Whiteley, H. R., and Zelin, S. (1994) Transport of microorganisms through soil. *Water Air Soil Poll.* **75**, 141–158.
- Ashby, M. N., Rine, J., Mongodin, E. F., Nelson, K. E., and Dimster-Denk, D. (2007) Serial analysis of rRNA genes and unexpected dominance of rare members of microbial communities. *Appl. Environ. Microb.* **73**, 4532–4542.
- Barns, S. M., Takala, S. L., and Kuske, C. R. (1999) Wide distribution and diversity of members of the bacterial kingdom *Acidobacteria* in the environment. *Appl. Environ. Microb.* **65**, 1731–1737.
- Bond, P. L., Hugenholz, P., Keller, J., and Blackall, L. L. (1995) Bacterial community structure of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. *Appl. Environ. Microb.* **61**, 1910–1916.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336.
- Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* **11**, 265–270.
- Chao, A., Chazdon, R. L., Colwell, R. K., and Shen, T. J. (2005) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.* **8**, 148–159.
- Cole, J. R., Chai, B., Farris, R. J., Wang, Q., Kulam, S. A., McGarrell, D. M., Garrity, G. M., and Tiedje, J. M. (2005) The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res.* **33**, D294–D296.
- Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., McGarrell, D. M., Marsh, T., Garrity, G. M., et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141–D145.
- de Wit, R., and Bouvier, T. (2006) *‘Everything is everywhere, but, the environment selects’*; what did Baas Becking and Beijerinck really say? *Environ. Microbiol.* **8**, 755–758.
- Dorer, M. S., Talarico, S., and Salama, N. R. (2009) *Helicobacter pylori*’s unconventional role in health and disease. *PLoS Pathog.* **5**, e1000544.
- Elshahed, M. S., Youssef, N. H., Spain, A. M., Sheik, C., Najar, F. Z., Sukharnikov, L. O., Roe, B. A., Davis, J. P., Schloss, P. D., Bailey, V. L., et al. (2008) Novelty and uniqueness patterns of rare members of the soil biosphere. *Appl. Environ. Microb.* **74**, 5422–5428.
- Emerson, D., Rentz, J. A., Lilburn, T. G., Davis, R. E., Aldrich, H., Chan, C., and Moyer, C. L. (2007) A novel lineage of Proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. *PLoS One* **2**, e667.
- Fierer, N., and Jackson, R. B. (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. USA* **103**, 626–631.
- Gans, J., Wolinsky, M., and Dunbar, J. (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**, 1387–1390.
- Grime, J. P. (2001) *Plant Strategies, Vegetation Processes and Ecosystem Properties*. Wiley, Chichester, UK.
- Harris, J. (2009) Soil microbial communities and restoration ecology, Facilitators or followers? *Science* **325**, 573–574.
- Harshey, R. M. (2003) Bacterial motility on a surface: Many ways to a common goal. *Annu. Rev. Microbiol.* **57**, 249–273.
- Hubbell, S. P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, New Jersey, USA.
- Huber, J. A., Welch, D. B. M., Morrison, H. G., Huse, S. M., Neal, P. R., Butterfield, D. A., and Sogin, M. L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318**, 97–100.
- Hughes Martiny, J. B., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., et al. (2006) Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112.
- Inagaki, F., Takai, K., Kobayashi, H., Nealson, K. H., and Horikoshi, K. (2003) *Sulfurimonas autotrophica* gen. nov., sp. nov., a novel sulfur-oxidizing ϵ -proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *Int. J. Syst. Evol. Microbiol.* **53**, 1801–1805.
- Jesus, E. D. C., Marsh, T. L., Tiedje, J. M., and Moreira, F. M. D. S. (2009) Changes in land use alter the structure of bacterial communities in Western Amazon soils. *ISME J.* **3**, 1004–1011.
- Klappenbach, J. A., Saxman, P. R., Cole, J. R., and Schmidt, T. M. (2001) rrndb: the ribosomal RNA operon copy number database. *Nucleic Acid Res.* **29**, 181–184.
- Köberl, M., Müller, H., Ramadan, E. M., and Berg, G. (2011) Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. *PLoS One* **6**, e24452.
- Lauber, C. L., Strickland, M. S., Bradford, M. A., and Fierer, N. (2008) The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol. Biochem.* **40**, 2407–2415.
- 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. Microb.* **75**, 5111–5120.
- Liao, D. (1999) Concerted evolution: Molecular mechanism and biological implications. *Am. J. Hum. Genet.* **64**, 24–30.
- Lozupone, C. A., and Knight, R. (2007) Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. USA* **104**, 11436–11440.
- Miyashita, N. T., Iwanaga, H., Charles, S., Diway, B., and Chong, L. (2013) Soil bacterial community structure in five tropical forests in Malaysia and one temperate forest in Japan revealed by pyrosequencing analyses of 16S rRNA

- gene sequence variation. *Genes Genet. Syst.* **88**, 93–103.
- Nacke, H., Thürmer, A., Wollherr, A., Will, C., Hodac, L., Herold, N., Schönig, I., Schrupf, M., and Daniel, R. (2011) Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. *PLoS One* **6**, e17000.
- Nemergut, D. R., Cleveland, C. C., Wieder, W. R., Washenberger, C. L., and Townsend, A. R. (2010) Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. *Soil Biol. Biochem.* **42**, 2153–2160.
- O'Malley, M. A. (2008) *Everything is everywhere: but the environment selects*: ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Phil. Biol. & Biomed. Sci.* **39**, 314–325.
- Parton, W., Silver, W. L., Burke, I. C., Grassens, L., Harmon, M. E., Currie, W. S., King, J. V., Adair, E. C., Brandt, L. A., Hart, S. C., and Fasth, B. (2007) Global-scale similarities in nitrogen release patterns during long-term decomposition. *Science* **315**, 361–364.
- Quince, C., Lanzén, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F., and Sloan, W. T. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* **6**, 639–641.
- Roesch, L. F. W., 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.
- Russo, S. E., Legge, R., Weber, K. A., Brodie, E. L., Goldfarb, K. C., Benson, A. K., and Tan, S. (2012) Bacterial community structure of contrasting soils underlying Bornean rain forests: Inferences from microarray and next-generation sequencing methods. *Soil Biol. Biochem.* **55**, 48–59.
- Schlesinger, W. H. (1997) *Biogeochemistry: an Analysis of Global Change*. Academic Press, San Diego.
- Schloss, P. D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microb.* **71**, 1501–1506.
- Shannon, C. E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423.
- Singh, D., Takahashi, K., Kim, M., Chun, J., and Adams, J. M. (2012) A hump-backed trend in bacterial diversity with elevation of Mount Fuji, Japan. *Microb. Ecol.* **63**, 429–437.
- Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., and Herndle, G. J. (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA* **103**, 12115–12120.
- Stotzky, G. (1986) Influence of soil mineral colloids on metabolic processes, growth, adhesion, and ecology of microbes and viruses. In: *Interaction of Soil Minerals with Natural Organics and Microbes*. Special Publ. No. 17. (eds: P. M. Huang and M. Schnitzer), p. 305–428. Soil Science Society of America, Madison, Wisconsin, USA.
- Suni, S., and Romantschuk, M. (2004) Mobilisation of bacteria in soil by electro-osmosis. *FEMS Microbiol. Ecol.* **49**, 51–57.
- Tamames, J., Abellán, J. J., Pignatelli, M., Camacho, A., and Moya, A. (2010) Environmental distribution of prokaryotic taxa. *BMC Microbiol.* **10**, 85.
- Tringe, S. G., von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W., Podar, M., Short, J. M., Mathur, E. J., Detter, J. C., Bork, P., Hugenholtz, P., and Rubin, E. M. (2005) Comparative metagenomics of microbial communities. *Science* **308**, 554–557.
- Tripathi, B. M., Kim, M., Singh, D., Lee-Cruz, L., Lai-Hoe, A., Ainuddin, A. N., Go, R., Rahim, R. A., Husni, M. H. A., Chun, J., et al. (2012) Tropical soil bacterial communities in Malaysia: pH dominates in the equatorial tropics too. *Microb. Ecol.* **64**, 474–484.
- van der Heijden, M. G. A., Klironomos, J. N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A., and Sanders, I. R. (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**, 69–72.
- Wardle, D. A., Bardgett, R. D., Klironomos, J. N., Setälä, H., van der Putten, W. H., and Wall, D. H. (2004) Ecological linkages between aboveground and belowground biota. *Science* **304**, 1629–1633.
- Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K. H., Glöckner, F. O., and Rosselló-Móra, R. (2010) Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses. *Syst. Appl. Microbiol.* **33**, 291–299.