Horizon-Specific Bacterial Community Composition of German Grassland Soils, as Revealed by Pyrosequencing-Based Analysis of 16S rRNA Genes

Christiane Will, Andrea Thürmer, Antje Wollherr, Heiko Nacke, Nadine Herold, Marion Schrumpf, Jessica Gutknecht, Tesfaye Wubet, François Buscot and Rolf Daniel Nacke, Nadine Herold, Marion Schrumpf, Jessica

Published Ahead of Print 20 August 2010.

Updated information and services can be found at: http://aem.asm.org/content/76/20/6751

These include:

SUPPLEMENTAL MATERIAL
Supplemental material

REFERENCES
This article cites 54 articles, 20 of which can be accessed free at: http://aem.asm.org/content/76/20/6751#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more »

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Horizon-Specific Bacterial Community Composition of German Grassland Soils, as Revealed by Pyrosequencing-Based Analysis of 16S rRNA Genes*†

Christiane Will,1 Andrea Thu¨rmer,2 Antje Wollherr,2 Heiko Nacke,1 Nadine Herold,3 Marion Schrumpf,3 Jessica Gutknecht,4 Tesfaye Wubet,4 François Buscot,4 and Rolf Daniel1,2*

Department of Genomic and Applied Microbiology1 and Göttingen Genomics Laboratory,2 Institute of Microbiology and Genetics, Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany; Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, D-07745 Jena, Germany; and UFZ-Helmholtz Centre for Environmental Research, Department of Soil Ecology, Theodor-Lieser-Str. 4, D-06120 Halle/Saale, Germany4

Received 3 May 2010/Accepted 9 August 2010

The diversity of bacteria in soil is enormous, and soil bacterial communities can vary greatly in structure. Here, we employed a pyrosequencing-based analysis of the V2-V3 16S rRNA gene region to characterize the overall and horizon-specific (A and B horizons) bacterial community compositions in nine grassland soils, which covered three different land use types. The entire data set comprised 752,838 sequences, 600,544 of which could be classified below the domain level. The average number of sequences per horizon was 41,824. The dominant taxonomic groups present in all samples and horizons were the Acidobacteria, Betaproteobacteria, Actinobacteria, Gammaproteobacteria, Alphaproteobacteria, Deltaproteobacteria, Chloroflexi, Firmicutes, and Bacteroidetes. Despite these overarching dominant taxa, the abundance, diversity, and composition of bacterial communities were horizon specific. In almost all cases, the estimated bacterial diversity (H’) was higher in the A horizons than in the corresponding B horizons. In addition, the H’ was positively correlated with the organic carbon content, the total nitrogen content, and the C-to-N ratio, which decreased with soil depth. It appeared that lower land use intensity results in higher bacterial diversity. The majority of sequences affiliated with the Actinobacteria, Bacteroidetes, Cyanobacteria, Fibrobacteres, Firmicutes, Spirochaetes, Verrucomicrobia, Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria were derived from A horizons, whereas the majority of the sequences related to Acidobacteria, Chloroflexi, Gemmatimonadetes, Nitrospira, TM7, and WS3 originated from B horizons. The distribution of some bacterial phylogenetic groups and subgroups in the different horizons correlated with soil properties such as organic carbon content, total nitrogen content, or microbial biomass.

Soil is probably the most complex microbial environment on Earth with respect to species richness and community size. The microbial richness in soils exceeds that of other environments (44) and is higher by orders of magnitude than the biodiversity of plants and animals. Cultivated soil or grassland soil contains an estimated 2 × 10^9 prokaryotic cells per gram (12). Soil microbial communities are an important factor of agricultural and managed systems, as they are responsible for most nutrient transformations in soil and influence the above-ground plant diversity and productivity (53).

To analyze the bacterial community in soils, most approaches target the 16S rRNA gene by PCR amplification and subsequent analysis employing sequencing of clone libraries (10, 24), denaturing gradient gel electrophoresis (DGGE) (38), or terminal restriction fragment length polymorphism (T-RFLP) (17, 52). Most of these approaches provided limited insights into the structure of soil bacterial communities, as the survey sizes and the number of compared sampling sites were small with respect to the enormous bacterial diversity present in different soil samples. For example, the reported clone libraries vary considerably in size, but small sample sizes (500 or fewer 16S rRNA gene sequences) are usually analyzed and employed for the theoretical estimation of species richness (39). This provides snapshots of the predominant bacterial community members, but phylogenetic groups that are present in a low abundance and which may possess important ecosystem functions are not assessed (47). In addition, it has been shown that rich sampling (several thousands of clones) of complex bacterial communities is required to perform robust measurements and estimations of community diversity parameters (37). Thus, the detection bias accompanying analyses of small sample sizes can lead to invalidated assumptions. Genetic profiling techniques such as DGGE and T-RFLP have high-throughput capability. These approaches allow researchers to unravel differences in community structure but are limited for assessing diversity (23, 40). To deeply survey the diversity and the composition of the bacterial communities within different soil samples, large-scale pyrosequencing of partial 16S rRNA
analyses of 16S rRNA genes to assess the bacterial community. It has been shown that the microbial population in the shallow subsurface is an important part of the soil microbial biomass (18). It has been focused on the analysis of microbial communities and taxa that were present in a low abundance. However, all those studies based on partial gene sequences. Those studies provided comprehensive sets, which comprised 39,707 (30) to 152,359 (34) 16S rRNA genes has been employed recently. Previous pyrosequencing-based studies of soil (1, 30, 34, 43) have generated large data sets, which comprised 39,707 (30) to 152,359 (34) 16S rRNA partial gene sequences. Those studies provided comprehensive insights into the biogeography of bacterial soil communities and taxa that were present in a low abundance. However, all those studies focused on the analysis of microbial communities present in topsoil. The subsoil is also known to harbor an important part of the soil microbial biomass (18). It has been shown that the microbial population in the shallow subsurface is impacted by agricultural production to a similar extent as that in topsoil (5).

In this study, we performed large-scale pyrosequencing-based analyses of 16S rRNA genes to assess the bacterial community composition in topsoil and the corresponding subsoil of nine different grassland sites in the Hainich region (Thuringia, Germany). To provide a high level of coverage at the species level (97% genetic distance) and minimize detection bias, we exceeded the above-described numbers of analyzed 16S rRNA gene sequences (752,838 in this study). To examine the impact of land use on bacterial diversity and community composition, the selected grassland sites covered a range of three different land use types, including samples from unfertilized pastures grazed by cattle, fertilized mown pastures grazed by cattle, and fertilized meadows. In many recent studies, surveys were focused on comprehensive analyses of a single soil or a few soil samples (1, 14, 37, 43). This allowed the determination of overall bacterial species richness and community composition, but the assessment of spatial patterns and environmental factors that drive these patterns is hampered by the limited number of examined soils. To assess spatial distribution and the impact of soil edaphic factors and land use on community structure, we used triplicate samples of each land use type from different locations. In addition, composite samples derived from five soil cores after the separation of soil horizons were employed.

### MATERIALS AND METHODS

#### Study sites and soil sampling.
Sample soils were collected from nine different grassland sites of the Hainich region in Germany, which is located in the west of Thuringia near the border to Hessen (latitude, 51.2167°N/10°27′E; longitude, 10.45°/E 10°27′). The Hainich region is one of the three locations investigated within the framework of the German Biodiversity Exploratories initiative (www.biodiversity-exploratories.de). The nine sampling sites encompass the following three different land use types: fertilized meadow (plots 1 to 3), fertilized mown pasture grazed by cattle (plots 4 to 6), and unfertilized pasture grazed by cattle (plots 7 to 9) (for coordinates, see Table S1 in the supplemental material). Sampling was performed in April and May 2008. At each sampling site, five soil cores (8.3 cm in diameter) were sampled with a motor-driven soil column cylinder at each corner and in the center of the plot within a given area of 20 m by 20 m. The soil was classified using the World Reference Base of Soil Resources (27). The predominant soil type in the study plots is Stagnosol, which is characterized by a perched water table, strong mottling, and reducing conditions (27). For each soil core, we determined soil horizons according to the Guidelines for Soil Description (28). The horizons were homogenized and pooled into one composite sample per plot and horizon. Coarse roots and stones (>5 mm) were removed from the samples. In the majority of the samples, the horizons Ah (topsoil) and Btg (subsoil) were detected. In plots 2 and 3 the top horizon was a transitional horizon (ABth) that was employed instead of the missing Ah horizon. In plots 7 and 8 only a transitional horizon (ABth) between topsoil and parent rock material was present. This horizon was employed instead of the missing Btg horizon. Throughout the study the topsoil horizon and the subsoil horizon were designated horizon A and horizon B, respectively.

#### Determination of microbial biomass.
To determine microbial biomass, we performed phospholipid fatty acid analysis (PLFA) on soil samples from the A and B horizons of the sampling sites. The composite samples were kept frozen at −80°C after sampling and freeze-dried prior to PLFA extractions. PLFA extractions were performed by using a modified Bligh and Dyer (4) method. Briefly, 2 g...
of freeze-dried sample was extracted twice in a chloroform-methanol-citrate buffer (1:2:0.8), followed by overnight phase separation. Fatty acids in the organic phase were then separated by using a silica-bonded phase column (silica-based solid-phase extraction [SPE-SI] Bond Elut, 3 ml, 500 mg: Varian Inc., Darmstadt, Germany) to remove glycolipids and neutral lipids. The polar lipids were then converted to fatty acid methyl esters by mild alkaline methanolysis. Methyl-esterified fatty acids were analyzed by using a Hewlett-Packard 6890 gas chromatograph equipped with a DB-5MS column (60-m length; Agilent Technologies, Böblingen, Germany) and interfaced with an Agilent 5973 mass selective detector. Peak areas of each lipid were converted to n mole/g soil using internal standards (19:0 nonadecanoic methyl ester). The total n mole/g dry soil (sum of all lipids present, 20 or fewer carbons in length) was used as an index of microbial biomass (19, 25).

**RESULTS AND DISCUSSION**

General characteristics of the soil samples and the pyrosequencing-derived data set. In this study, we assessed and compared the compositions of soil bacterial communities present in the A and B horizons of nine different grassland sites of the Hainich region in Germany by a pyrosequencing-based analysis of the 16S rRNA gene sequences. The grassland sites covered a range of the following three different land use types: fertilized meadow (plots 1 to 3), fertilized mown pasture grazed by cattle (plots 4 to 6), and unfertilized pasture grazed by cattle (plots 7 to 9). The soil type of all samples was Stagnosol, except for plot 1, which was a Vertic Cambisol. In addition, further analysis of the Stagnosols revealed that plot 6 was a Luvis Stagnosol, whereas the other plots were Vertic Stagnosols (see Table S1 in the supplemental material).

The two analyzed soil horizons showed significant differences with respect to edaphic properties such as soil pH, OC content, N content, C-to-N ratio, and water content (Table 1). The pH in the A horizons ranged from 6.03 to 7.20, and the pH in the B horizons ranged from 6.30 to 7.40. In general, the pH value of the B horizon was higher than that of the corresponding A horizon, whereas the water content, the amounts of OC and N, and the C-to-N ratio showed 1.1- to 2.0-fold, 2.0- to 10.9-fold, 1.8- to 7.3-fold, and 1.2- to 1.5-fold decreases with depth, respectively.

The microbial biomass in the B horizons of all samples was lower by 48.7 to 96.3% than that in the corresponding A horizons (Table 1). A decrease in the total microbial biomass with soil depth was previously reported (5, 9, 16, 18). The total microbial biomass was positively correlated with the concentration of OC (r = 0.88; P < 0.01). This supports the assumption reported previously by Blume et al. (5), that carbon availability is closely associated with microbial biomass. In addition, significant correlations of microbial biomass with the concentration of N (r = 0.84; P < 0.01) and the C-to-N ratio (r = 0.89; P < 0.01) were detected.

The pyrosequencing-based analysis of the V2-V3 region of the 16S rRNA gene was employed for assessments of bacterial community compositions from the A and B horizons of the nine sampling sites. Short pyrosequencing reads assess the microbial diversity almost as reliably as near-full-length sequences when appropriate primers are chosen. Primers derived from V2-V3 region of the 16S rRNA gene were shown previously to be suitable for this purpose (36). In addition, this region is the most effective region for universal genus identification (7, 42). Across all 18 samples, we recovered 752,838 quality sequences with a read length of ≥200 bp. The average read length was 262 bp. The number of sequences per sample ranged from 25,851 to 61,366, with an average of 41,824 (see Table S2 in the supplemental material). We were able to classify 600,544 (79.77%) of the quality sequences below the domain level. The percentage of classified 16S rRNA gene sequences was in the range of those of other pyrosequencing-based studies (35), but the average number of sequences per sample and the total number of analyzed sequences exceeded those of other previously reported studies of pyrosequencing-based determinations of soil bacterial community composition (34, 43).

**Nucleotide sequence accession number.** The 16S rRNA gene sequences derived from pyrosequencing have been deposited in the NCBI Sequence Read Archive under accession number SRA012068.1.
ational taxonomic units (OTUs) at sequence divergences of 3% (species level) and 20% (phylum level). The rarefaction analysis of bacterial communities derived from the A and B horizons of the nine sampling sites is depicted in Fig. 1. At a 20% genetic distance, almost all curves showed saturation, indicating that the surveying effort covered almost the full extent of taxonomic diversity at this genetic distance. In addition, a comparison of rarefaction analyses with the number of OTUs estimated by the Chao1 richness estimator revealed that 83 to 100% of the estimated taxonomic richness was covered by the sequencing effort. At a 3% genetic distance, the observed richness was 63 to 80% of that predicted by the Chao1 richness estimator (Table 2). Thus, we did not survey the full extent of taxonomic diversity at the species level. Taking into account that at genetic distances below 5%, rarefaction analyses underestimate the bacterial richness whereas Chao1 estimators overestimate it (43), a substantial fraction of the bacterial diversity at the species level was assessed by the surveying effort. It is important that pyrosequencing provides an unprecedented sampling depth compared to that of traditional Sanger sequencing of 16S rRNA genes (51), but the intrinsic error of pyrosequencing could result in the overestimation of rare phylotypes, since each pyrosequencing read is treated as a unique identifier of a community member and correction by assembly and sequencing depth, which is typically applied during genome projects, is not feasible (26, 32). To ensure per-base error rates lower than that of conventional Sanger sequencing, we used quality filtering of the pyrosequencing-derived data set, such as the removal of reads with atypical lengths (26). In addition, to minimize the overestimation of rare phylotypes, clustering and diversity estimates were performed only at genetic divergences of ≥3% (32).

Acosta-Martínez et al. (1) postulated previously that in managed soils, the maximum number of OTUs is less than 3,400 at a genetic distance of 3%. This is in contrast to our results, as up to 4,781 and 6,231 OTUs were predicted for fertilized meadows (plot 3, A horizon) and fertilized mown pasture grazed by cattle (plot 5, A horizon), respectively (Table 2). The differences in the results might be explained by the different surveying efforts. Several studies (13, 37, 43, 46, 56) showed that the number of analyzed sequences per sample has an effect on the predicted number of OTUs. For example, Roesch et al. (43) previously plotted the number of observed OTUs against the sequencing effort using the bacterial community present in a
Canadian forest soil sample. The employment of the whole data set (53,632 sequences) revealed 5,500 OTUs at a genetic distance of 3%, whereas the reduction of the same data set to 30,000 sequences yielded 3,500 OTUs. In general, fewer sequences result in lower curve progression and a lower number of predicted OTUs. In addition, the comparison of richness estimates between different surveys might be hampered by the differences in sequence conservation and sequence length of the analyzed 16S rRNA gene regions. Recently, Engelbrektson et al. (15) showed that amplicon length and differences in the analyzed 16S rRNA gene regions markedly influence estimates of richness and evenness.

The Shannon index of diversity ($H'$) was determined for all samples (Table 2). At a genetic distance of 3%, it ranged from 5.65 to 7.16 in the A horizons and from 5.01 to 6.72 in the B horizons. The predicted diversity in the topsoil exceeded that of the corresponding subsoil, except for plot 1 (Table 2). To our knowledge, no other study assessing bacterial diversity along a soil profile was conducted with a comparable surveying effort. However, a significant decrease of bacterial diversity with soil depth was also recorded by a community analysis employing terminal restriction fragment length polymorphism (33) and phospholipid fatty acid analysis (18) of soil profiles derived from Californian grassland and soil samples of the Sedgwick Reserve (California), respectively.

The bacterial diversity at a genetic distance of 3% was strongly related to the content of OC and N as well as to the C-to-N ratio. Positive correlations between the $H'$ and the OC content ($r = 0.60; P < 0.01$), the N content ($r = 0.58; P < 0.05$), and the C-to-N ratio ($r = 0.65; P < 0.01$) were observed. Similar correlations were detected by analyzing soil samples from South American grasslands, in which the $H'$ correlated positively with the microbial biomass C and N ($r = 0.53$ to $0.58$; $P = 0.02$ to $0.03$) (3). Interestingly, the mean $H'$ was lower in fertilized meadows (plots 1 to 3), with intermediate values in fertilized mown meadows (plots 4 to 6) and the highest values in unfertilized pastures (plots 7 to 9), over both horizons (Table 2). Thus, a higher bacterial diversity in samples from unfertilized plots, which represent the lowest land use intensity in this study, is indicated. Nevertheless, within the same land use types, strong variations in diversity were observed.

### Distribution of taxa and phylotypes across all samples.

The 600,544 classifiable sequences were affiliated with 23 phyla across the entire data set. The dominant phyla across all samples were *Acidobacteria*, *Betaproteobacteria*, *Actinobacteria*, *Gammaproteobacteria*, *Bacteroidetes* and *Deltaproteobacteria*. *Chloroflexi*, *Firmicutes*, and *Bacteroidetes* represented approximately 75% of all classified sequences, corresponding sequences represented 0.74 and 2.86% of all classified sequences in each respective horizon. The most abundant phylotype within one individual sample (plot 3, B horizon) was a member of the *Actinobacteria* (Acetobacteraceae) and a member of acidobacterial subgroup 4, respectively. The corresponding sequences represented 0.74 and 2.86% of all classified sequences in each respective horizon. The most abundant phylotype within one individual sample (plot 3, B horizon) was a member of *Actinobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria*, which represented approximately 75% of all classified sequences, corresponded roughly to those found by other studies (21, 34). Thus, despite the different surveying efforts and sampling sites used in the different studies, it is indicated that a variety of soils contain the same dominant bacterial groups.

### Table 2. Species richness estimates obtained at genetic distances of 3% and 20%.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Plot</th>
<th>Shannon index ($H'$)*</th>
<th>Rarefaction (no. of OTUs)</th>
<th>Chao1b (no. of OTUs)</th>
<th>Coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3% 20%</td>
<td>3% 20%</td>
<td>3% 20%</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>5.92 2.67</td>
<td>1,629 103</td>
<td>2,335 124</td>
<td>69.8 83.1</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>6.11 4.03</td>
<td>2,730 224</td>
<td>4,084 236</td>
<td>66.9 94.8</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>6.76 4.27</td>
<td>3,307 262</td>
<td>4,781 272</td>
<td>69.2 96.2</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>6.30 2.93</td>
<td>2,805 57</td>
<td>4,395 59</td>
<td>63.8 96.6</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>7.07 4.49</td>
<td>3,957 335</td>
<td>6,231 366</td>
<td>63.2 91.6</td>
</tr>
<tr>
<td>A</td>
<td>6</td>
<td>6.15 2.91</td>
<td>2,344 57</td>
<td>3,551 57</td>
<td>66.0 100</td>
</tr>
<tr>
<td>A</td>
<td>7</td>
<td>7.16 4.50</td>
<td>4,329 385</td>
<td>6,487 407</td>
<td>66.7 94.7</td>
</tr>
<tr>
<td>A</td>
<td>8</td>
<td>5.65 2.51</td>
<td>1,516 63</td>
<td>1,924 68</td>
<td>78.8 93.3</td>
</tr>
<tr>
<td>A</td>
<td>9</td>
<td>7.05 4.41</td>
<td>4,056 381</td>
<td>6,232 438</td>
<td>65.1 87.0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>6.72 4.34</td>
<td>3,528 340</td>
<td>5,168 360</td>
<td>68.3 94.5</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>5.01 2.64</td>
<td>1,022 84</td>
<td>1,399 94</td>
<td>73.0 89.6</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>5.14 2.50</td>
<td>1,122 76</td>
<td>1,509 82</td>
<td>74.4 92.6</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>5.57 2.72</td>
<td>1,388 55</td>
<td>1,745 55</td>
<td>79.5 99.4</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>6.15 3.98</td>
<td>2,450 237</td>
<td>3,635 252</td>
<td>67.4 94.0</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>5.64 2.83</td>
<td>1,741 69</td>
<td>2,420 73</td>
<td>72.0 94.8</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>6.51 4.15</td>
<td>2,392 267</td>
<td>3,293 293</td>
<td>72.6 91.2</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>5.57 2.88</td>
<td>1,923 54</td>
<td>2,854 57</td>
<td>67.4 94.7</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>6.09 4.10</td>
<td>2,402 258</td>
<td>3,606 280</td>
<td>66.6 92.1</td>
</tr>
</tbody>
</table>

* A higher number indicates more diversity. 
* Nonparametric richness estimator based on the distribution of singletons and doubletons. 
* The estimates were calculated by employing the tools Aligner, Complete Linkage Clustering, and Rarefaction of the RDP pyrosequencing pipeline (11). The results from the rarefaction analyses are also depicted in Fig. 1.
FIG. 2. Distribution of phylogenetic groups in the A and B horizons derived from the different grassland sampling sites. Plot numbers are given below the graph. A description of the plots is given in Table 1. A and B indicate the different horizons. Shown are the percentages of the classified sequences. Phylogenetic groups accounting for ≥0.25% of the classified sequences are summarized in the artificial group “others.”

the above-mentioned member of acidobacterial subgroup 4, representing 8.36% of the sequences from that soil.

Distribution and abundance of the predominant phylum Acidobacteria. Members of the phylum Acidobacteria were predominant across all samples. This finding is in accordance with findings of other studies of the composition of soil-derived bacterial communities from a variety of environments, such as pristine forest, grassland, and agricultural soils (29). Here, members of the Acidobacteria form a significant fraction (12.68 to 49.86%) of the 16S rRNA gene sequences in both horizons from all land use types (see Tables S3 and S4 in the supplemental material). Correspondingly, members of this phylum have been reported to constitute an average of 20% and a maximum of approximately 50% of bacterial communities derived from various soils (13). Thus, based on their abundance and their presence in various soil types, the Acidobacteria appear to play an important role in the ecosystem function of soils.

The phylum Acidobacteria is divided into 26 subgroups, but only little is known with respect to the physiological and metabolic capabilities of the different subgroups (2). We detected 18 and 22 of these subgroups in the A and B horizons, respectively. Most abundant in both horizons were subgroups 6, 4, 16, and 7 (see Tables S4 and S5 in the supplemental material). In the A horizons, these subgroups were represented by 47, 21, 13, and 7% of all acidobacterial sequences, respectively, and in the B horizons, these subgroups were represented by 26, 38, 5, and 14% of all acidobacterial sequences, respectively (see Table S5 in the supplemental material). These results are in contrast to a previous study by Hansel et al. (21) of samples from a continuous watershed soil profile (Oak Ridge), which is the only other report of acidobacterial diversity with respect to soil horizon. In the A horizon, those researchers detected primarily subgroups 3 (21%), 4 (29%), and 6 (29%), whereas in our samples, these subgroups were represented by 5, 21, and 47%, respectively. In the B horizon, Hansel et al. (21) detected primarily subgroups 1 (32%) and 2 (61%), which were represented by less than 1% of all acidobacterial sequences derived from our soil samples. The predominant subgroups in the B horizons from the Hainich region were subgroups 4, 5, and 7. The major differences in the occurrences of acidobacterial subgroups in the B horizon might be due to the dissimilar pH values of the samples used in both studies. The pH in our subsoil samples ranged from 6.30 to 7.40 (Table 1) whereas the pH of the soil samples studied by Hansel et al. (21) was 4.5. It was reported previously that the abundance of the phylum Acidobacteria correlates with the soil pH (22, 30). Lauber et al. (34) showed previously that acidobacterial subgroups 1 and 2 were most abundant in acidic soils and decreased with the increase of the pH. Here, no significant correlations of changes in the abundance of the dominant acidobacterial subgroups and other phylogenetic groups with pH were observed. A reason for this finding might be that the sampling effort in most of the other studies was much less than that of this study. Another possibility is that almost all the pH values of our samples were near neutral. Correspondingly, a relatively small pH range was covered by our soil samples (Table 1), so there is simply a lower pH range from which to determine correlations. Nevertheless, we observed negative correlations between the abundant acidobacterial subgroup 4 and the OC content (r = −0.84; P < 0.01), N content (r = −0.83; P < 0.01), or C-to-N ratio (r = −0.77; P < 0.05) in the A horizons (see Fig. S1 in the supplemental material). In the B horizons, subgroup 4 also correlated negatively with the C-to-N ratio (r = −0.70; P < 0.05), whereas the relative abundance of subgroup 6 showed a positive correlation with the C-to-N ratio (r = 0.70; P < 0.05). Thus, the subgroup distribution varied with respect to the soil profile (horizon) and soil properties and provided some insights into the conditions that are required by the different subgroups. For example, low nutrient/OC conditions (B horizons) appear to favor subgroups 4 and 7, whereas higher nutrient/OC conditions (A horizons) favor subgroup 16. Interestingly, for members of subdivision 6, a high tolerance to nutrient/OC availability was indicated, as they constituted a substantial fraction in the A horizon and the B horizon (8.87 and 9.73% of all classified sequences, respectively).
The compositions of the bacterial community and the distributions of the phyla varied between A and B soil horizons. The most frequently present phyla in the A horizons were Acidobacteria (13 to 23%), Betaproteobacteria (14 to 23%), Gammaproteobacteria (10 to 26%), Actinobacteria (5 to 17%), and Alphaproteobacteria (9 to 14%). The most abundant phyla in the B horizons were Acidobacteria (28 to 50%), Betaproteobacteria (10 to 18%), Actinobacteria (4 to 15%), Chloroflexi (3 to 12%), and Alphaproteobacteria (5 to 10%) (Fig. 2).

We analyzed the respective abundances of the 15 most represented phyla in the A and B horizons. For almost all phyla and land use types, a significant ($P < 0.00001$) difference in distribution between the two horizons was apparent (see Table S6 in the supplemental material). The distribution of selected phyla in the two horizons is shown in Fig. 3 (see also Fig. S2 in the supplemental material). The majority of sequences affiliated with the Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, Spirochaetes, Verrucomicrobia, Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria were derived from A horizons, whereas the majority of the sequences related to Acidobacteria, Chloroflexi, Gemmatimonadetes, Nitrospira, TM7, and WS3 originated from B horizons. In many other reports, a pH gradient was identified as a major factor for changes in soil community structure, but as mentioned above, this trend was not observed in this study. Therefore, other factors appear to control the distribution of the phyla along the soil profile. As stated above, the total biomass decreased with soil depth (Table 1). A significant correlation between the total microbial biomass and the occurrence of several phyla was recorded. The relative abundance of the Acidobacteria, Chloroflexi, and Nitrospira, which increased with depth, correlated negatively with total biomass ($r = -0.53$ to $-0.79; P < 0.05$). The relative abundances of the Actinobacteria, Bacteroidetes, Verrucomicrobia, Alphaproteobacteria, and Gammaproteobacteria, which decreased with depth, showed a positive correlation with biomass ($r = 0.69$ to 0.84; $P < 0.01$).

A positive correlation of some of the dominant acidobacterial groups (see above) and the alaphaproteobacterial order Rhizobiales with the concentration of OC ($r = 0.77; P < 0.05$), the concentration of N ($r = 0.73; P < 0.05$), and the C-to-N ratio ($r = 0.87; P < 0.01$) was detectable in the B horizon (see Fig. S3 in the supplemental material). A statistically significant positive correlation of the Rhizobiales with the C-to-N ratio ($r = 0.80; P < 0.01$) was also detected in the A horizon. Taking into account that the Rhizobiales include the genera Rhizobium and Bradyrhizobium, which comprise members that are able to fix nitrogen and are associated with roots of legumes, a positive correlation was expected (49). In addition, land use and management regimens seem to have an impact on the Rhizobiales, as the fertilized plots cluster and the plots with cattle cluster (data not shown). An impact of fertilization on the structure and diversity of rhizobial populations was observed previously in other studies (6, 41). For example, rhizobial populations differed between cultivated and uncultivated Mexican soils. In addition, the affinity of host cultivars for different members of the Rhizobiales influenced the composition of rhizobial populations (54).

Striking is the distribution of the phyla Fibrobacteres and Nitrospira, which occurred almost exclusively in the A or the B horizon, respectively. Members of the Fibrobacteres are part of the microbial community in the first stomach of ruminant animals and degrade plant-based cellulose (31). Taking into account that members of the Fibrobacteres hardly occurred in grassland samples without the presence of cattle (i.e., plots 1 to 3) and almost exclusively in the topsoil, it can be assumed that members of this phylum were introduced into the samples by cattle. Members of the Nitrospira are found in interspace soils and rarely in the rhizosphere (13). In the latter environment, heterotrophic root-associated microorganisms suppress the growth of autotrophic Nitrospira. This might explain why in our samples, members of the Nitrospira occurred in the subsoil, with less rooting than in the topsoil. In addition, the concentrations of OC and N decreased with soil depth (Table 1). Thus, chemolithoautotrophic organisms adapted to darkness, like Nitrospira, have a selective advantage in subsoil samples.

**Conclusions.** Although we recovered an average of 41,824 sequences per sample, we did not survey the full extent of bacterial richness at the species level within an individual soil or horizon. Thus, an increase in surveying efforts would probably result in the identification of more bacterial taxa, which are present in a low abundance. In most cases, the B horizons showed a lower estimated bacterial diversity than the corresponding A horizons (Table 2). Correspondingly, a greater coverage of the bacterial community in the B horizons can be achieved by using the same surveying effort. In addition, the identification of bacterial taxa at the finest level of taxonomic resolution is currently not feasible by applying large-scale pyrosequencing. However, the advancement of the technology will result in an increase of the read length, and this limitation will become less relevant in the near future.

To provide a robust assessment of the impact of land use, soil factors, or soil depth on bacterial diversity, distribution,
and community composition, we used triplicate samples of each land use type from different locations. Overall, the abundance, composition, and diversity of the bacterial communities were strongly depth dependent. The Shannon index of diversity along with the nutrient content (N and OC), water content, and biomass decreased with depth (Tables 1 and 2). Based on the sharp decrease of the OC content (up to 10.9-fold) in the B horizons compared to the corresponding A horizons (Table 1) and other surveys (18, 33, 57), the concentration of OC appears to be the major driver for the diversity and structure of bacterial communities along the soil profile at near-neutral pH values. Nevertheless, we observed a variability of bacterial communities within an individual land use type, and exceptions to the above-mentioned general results were found; i.e., a slightly higher Shannon index was recorded for the B horizon of plot 1. Thus, it is advisable to survey as many soil samples as possible for the identification of general patterns and comparison of the results with those of other soil surveys. One caveat of the latter, however, is the limited comparability of different surveys, as sampling strategy, survey effort, number and type of soil factors measured, and approaches used to analyze the sequence data vary considerably (37). To take full advantage of the increasing number of data sets on soil bacterial communities, minimal requirements for sampling and the set of analyzed soil factors as well as rules for sequence analysis and phylogenetic assignment should be defined.

ACKNOWLEDGMENTS

We thank Rudolf-Josef Fischer (Department of Medical Informatics and Biomathematics, Westphalian Wilhelms University, Münster, Germany) for supporting the statistical analyses. We thank the local implementation team of the Hainich exploratory and the BEO (Biodiversity Exploratories Office) for organizing the coordinated soil sampling campaign.

The work has been funded by the DFG Priority Program 1374, Infrastructure-Biodiversity-Exploratories (DA 374/4-1).

Field work permits were given by the responsible state environmental offices of Thuringen.

REFERENCES

30. Morales, S., T. Cosart, J. Johnson, and W. Holben. 2009. Extensive phylogenetic analysis of a soil bacterial community illustrates extreme taxon evenness and the effects of ampiclon length, degree of coverage, and DNA