Living with detergents: pyrosequencing-based assessment of bacterial community structures in soils subjected for decades to contamination by detergents

Rafael A. Rojas-Herrera · Alejandra S. Ramos-Castillo · Héctor Estrada-Medina ·
César De los Santos-Briones · Miguel A. Keb-Llanes · Roberto C. Barrientos-Medina ·
Yuri J. Peña-Ramírez · Aileen O'Connor-Sánchez

Abstract Although detergents are major environmental pollutants for soil and water, their impact in bacterial community structures has not been addressed. We compared the bacterial community structures as well as several edaphic parameters between severely detergent-contaminated soils and non-contaminated soils in the State of Yucatán, México. The results indicate that sodium concentration, salinity, and electrical conductivity were significantly higher in contaminated samples, and that this correlates with different bacterial community structures. The most important differences were that (i) samples with detergent presented a lower species richness; (ii) *Proteobacteria* was the most abundant phylum in soils with detergent, while *Actinobacteria* and *Acidobacteria* were the most abundant phyla in soils without detergent; (iii) *Rhodomicrobium*, *Hydrogenophaga,* and *Thiobacillus* were the most abundant genera in soils with detergent, while *Acidobacteria* dominated soils without detergent. With the continual increase of the human population without access to a proper disposal of waste waters, these modifications may contribute to bringing about changes in the ecological parameters of the region.

Keywords Soil bacterial community · 16S rRNA pyrosequencing · Detergent contamination

R. A. Rojas-Herrera
Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, Mérida, Yucatán, Mexico

A. S. Ramos-Castillo · C. De los Santos-Briones ·
M. A. Keb-Llanes · A. O'Connor-Sánchez (✉)
Centro de Investigación Científica de Yucatán, Unidad de Biotecnología, Mérida Yucatán, Mexico
e-mail: aileen@cicy.mx

H. Estrada-Medina
Facultad de Medicina Veterinaria y Zootecnia, Departamento de Manejo y Conservación de Recursos Naturales, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, Merida Yucatán, México

R. C. Barrientos-Medina
Facultad de Medicina Veterinaria y Zootecnia, Departamento de Ecología, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, Merida Yucatán, México

Y. J. Peña-Ramírez
Departamento de Ciencias de la Sustentabilidad. El Colegio de la Frontera Sur - Unidad Campeche, Campeche, México

Introduction

The structure of a microbial community is determined by a complex combination of factors, such as the overall habitat characteristics and historical modifications, the physical structure of the substrates, and the changes in current environmental parameters (Jeffries et al. 2011). Several of these factors, e.g. salinity (Lozupone and Knight 2007), substrate type (Jeffries et al. 2011), soil management (Wessén et al. 2010), and oil pollution (Kimes et al. 2013), have been extensively analyzed as determinants for bacterial taxa distribution. In contrast, to our knowledge, the influence of detergents on soil bacterial community structures has not been investigated.

In the State of Yucatán, Mexico, in most rural houses, the sewage produced daily by laundry and general washing has for decades been discharged directly onto the soil of the backyards. This wastewater contains not only a mixture of various detergents, bought mainly in powdered form, but occasionally also bleach, soap, hot water, and even food scraps.

Most powder detergents are composed of surfactants (20–35 %); builders-sequestrants (polyphosphates, zeolites, chelating agents) (15–45 %); alkaline components (silicates,

sodium carbonate, etc.) (10–30 %); bleaching agents (perborates) and bleach activators (0–30 %); carboxymethyl-cellulose, enzymes, foam regulators, optical brighteners, and perfumes (2–7 %); and sodium sulfate (10–30 %) (Zoller 1998). It is to be expected that they exert a selective pressure when present in soils, bringing about modifications in the structure of the edaphic bacterial communities.

The gene encoding the small-subunit rRNA serves as a prominent tool for the phylogenetic analysis and classification of bacteria, owing to its high degree of conservation and its fundamental function in living organisms (Vos et al. 2012). Pyrosequencing of this gene has proved to be a cost-effective method for the characterization of bacterial communities and, although it may have a moderate bias (Kumar et al. 2011), it is widely used to get a culture-independent general view about the phylogenetic profile of bacterial communities.

In the present work, bacterial community structures of three samples from detergent-contaminated soils and three from non-contaminated soils were assessed by 16S rRNA tag-pyrosequencing. The results may contribute to a better understanding of how selective pressure imposed by this anthropogenic contamination can modify soil bacterial communities.

Materials and methods

Environmental samples Six soil samples were collected to a depth of 5–15 cm in three house backyards located in the town of Hunucmá, Yucatán, Mexico (21°1'5.30" N / 89°52'31.21" W) (Fig. 1), in February 2012. Three of these soil samples (SDet1, SDet2, and SDet3) were taken from points where sewage produced by laundry and general washing has been discharged for at least 20 years, and three other samples (S1, S2, and S3) were taken in the same backyards, but at points located at least 20 m away from any wastewater discharges. Samples were placed in sterile plastic bags, sealed, and transported on ice to the laboratory, where they were stored at 4 °C for up to 2 days before further processing.

Soils characterization Soil samples were submitted to the Laboratorio de análisis de suelos, plantas y agua (LASPA) (Mérida, Yucatán, Mexico) for physical and chemical analysis. There, samples were dried to a constant weight in an oven at 105 °C, and sieved to 2 mm. All the analyses were carried out following standard procedures: particle size analysis with a Bouyoucus densimeter (Gee and Bauder 1986), water content using gravimetry (Gardner 1986), REDOX potentiometry (Patrick et al. 1996), pH potentiometry (Thomas 1996), phosphorus by Olsen method (Kuo 1996), total nitrogen by Kjeldahl method (Bremmer 1996), potassium using flame photometry (Helmke and Sparks 1996), organic matter by colorimetric determination (Nelson and Sommers 1996), salinity and electric conductivity potentiometry (Rhoades 1996).

Metagenomic DNA extraction Metagenomic DNA was extracted from soil samples using a PowerMax DNA isolation kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's protocol. Following elution, DNA samples were concentrated by ethanol precipitation, resuspended in 10 mM Tris-HCl (pH 8.0) and quantified with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Massively parallel bTEFAP Purified metagenomic DNA was submitted to the Research and Testing Laboratory (RTL) (Lubbock, TX, USA) for tag-pyrosequencing. Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) was performed as described previously using Gray28F (5'-TTTG ATCNTGGCTCAG-3') and Gray519R (5'-GTNTTACNGC GGCKGCTG-3') primers (Dowd et al. 2008).

HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA, USA) was used for PCR under the following conditions: 94 °C for 3 min followed by 32 cycles of 94 °C for 30 s; 60 °C for 40 s and 72 °C for 1 min; and a final elongation step at 72 °C for 5 min. A secondary PCR was performed for FLX (Roche, Nutley, NJ, USA) amplicon sequencing under the same condition by using designed special fusion primers with different tag sequences (Dowd et al. 2008). Tag-encoded FLX amplicon pyrosequencing analyses used a Roche 454 FLX instrument with Titanium reagents; Titanium procedures were based on RTL protocols (www.researchandtesting.com).

Data analysis and bacteria identification Following sequencing, all failed sequence reads, low quality sequence ends, tags, and primers were removed, and sequence collections were depleted of any non-bacterial ribosome sequences and chimeras using B2C2 software (Gontcharova et al. 2010), as described previously (Dowd et al. 2008; Ishak et al. 2011). To determine the identity of bacteria in the remaining sequences, these were denoised, assembled into clusters, and queried using a distributed BLASTn.NET algorithm (Dowd et al. 2005) against a database of high-quality, 16S bacterial sequences derived from NCBI. Database sequences were characterized as high-quality based upon criteria similar to those used by RDP (Cole et al. 2009). Using a .NET and C# analysis pipeline, the resulting BLASTn outputs were compiled and validated using taxonomic distance methods, and data reduction analysis was performed as described previously (Ishak et al. 2011). Based upon the above BLASTn-derived sequence identity (percent of total-length query sequence which aligns with a given database sequence) and validation using taxonomic distance methods, the bacteria were classified at the appropriate taxonomic levels based upon the following criteria: sequences with identity scores (relative to known or well-characterized 16S sequences) greater than 97 % identity (<3 % divergence) were resolved at the species level, between 95 % and 97 % at the genus level, between

90 % and 95 % at the family level, between 85 % and 90 % at the order level, 80 and 85 % at the class level, and 77 to 80 % at the phylum level.

Sequencing reads were aligned and clustered following the Ribosomal Database Project (RDP-Release 10) pyrosequencing pipeline (http://pyro.cme.msu.edu/). Shannon, Chao 1, and evenness indices, as well as rarefaction curves, were obtained using the RDP tools.

Canonical correspondence analysis To analyze the possible correlation between the community compositions and the environmental factors, a canonical correspondence analysis was performed. An ordination table was constructed based on the abundance of shared genera among samples and values of the environmental variables. A permutation test of 5,000 repetitions was performed to check the strength of the correlation matrix. All analyses were run on the PAST package (Hammer et al. 2001).

Accession numbers All the 16S rRNA gene sequences were deposited in the NCBI Sequence Read Archive (SRA) under the accession numbers: SRS472403 (S1), SRS472404 (S2), SRS466159 (S3), SRS472396 (SDet1), SRS472401 (SDet2), SRS472402 (SDet3).

Results and discussion

Soil characterization As can be seen in Table 1, SDet soil samples were coarser than S samples. This can be due to the scattering properties that detergents have, as the stream of wastewater breaks up soil aggregates and washes out the clay fraction. Possibly something similar happens with the humus, because the color in all SDet samples was lighter than in S ones (visual observations), and they had a lower content of organic carbon (Table 1).

Soils subjected to wastewater discharges showed higher water contents than S soils, a logical result because the discharges are composed mainly of water and they are dumped at least once a day; in fact, SDet samples were muds.

SDet samples had lower REDOX potential values than S samples. This is probably also related to the water dumping, because the frequent soaking of SDet soils creates anoxic conditions. Although it is worth mentioning that this kind of soil has a good permeability, which allows most of the wastewater and the compounds it contains to filter into deeper soil profiles; the soil remains moist due to the frequent water discharges.

The pH values of all samples ranged from 7.7 to 8.9; there were no relevant differences between S and SDet samples in this parameter. These pH values are very similar to others previously reported for soils of the same region of Yucatán (Aguila 2007), suggesting that these soils have a good buffering capacity, minimizing the effect of the wastewater on their pH.

There was a higher phosphorus concentration in SDet samples than in S samples, with exception of S2. This was also an expected result because detergents contain high amounts of phosphates.

Nitrogen showed the same range of values in both sets of samples. These were relatively high values, but similar to

Table 1 Physical and chemical characterization of the soils. Abbreviations: S, soil samples without detergent; SDet, soil samples with detergents; EC, electrical conductivity; m ± s.d., mean plus or minus standard deviation

Parameter	Soil samples							
	Without detergents				With detergents			
	S1	S2	S3	mS±s.d.	SDet1	SDet2	SDet3	mSDet±s.d.
Sand (%)	28.4	24.4	38.4	30.4±7.2	46.4	32.4	54.4	44.4±11.1
Silt (%)	26.7	44.7	52.7	41.3±13.3	30.7	52.7	38.7	40.7±11.1
Clay (%)	44.8	30.8	8.8	28.2±18.1	22.8	14.8	6.8	14.8±8
Water content (%)	6.8	13.0	3.1	7.7±4.9	9.0	42.5	26.6	26.0±16.7
REDOX potential (mv)	78.3	81.2	83.6	81.0±2.6	73.6	76.4	64.5	71.5±6.2
pH (Water) 1:2	8.0	7.9	8.1	8.0±0.0	8.3	7.7	8.9	8.3±0.6
P (mg/kg)	3.8	41.4	19.5	21.5±18.8	42.8	40.2	44.4	42.5±2.1
N (%)	0.69	0.76	0.32	0.59±0.23	0.66	0.82	0.42	0.63±0.2
Na (cmol/kg)	2.3	3.8	1.5	2.5±1.2	16.4	21.1	11.7	16.4±4.71
Salinity (%)	0.43	0.38	0.29	0.36±0.07	1.19	2.26	1.03	1.49±0.66
EC (1:5) mS/cm	0.36	0.38	0.28	0.34±0.05	0.82	2.68	1.00	1.5±1.02
Organic carbon (%)	11.0	13.9	13.7	12.9±1.6	8.9	10.0	6.7	8.5±1.7

those previously reported for these kinds of soil (Aguila 2007; Ruíz-Garvia 2008). It was expected that the nitrogen would not be higher in the SDet samples, as the detergents do not contain substantial amounts of nitrogen, but the results also suggest that the wastewater does not influence the level of nitrogen normally present in soils.

The sodium content was about one order of magnitude higher in SDet samples than in S samples, and salinity and electrical conductivity, parameters closely related to each other, presented similar patterns. These values are very high for soils, but they are understandable because detergents contain high concentrations of sodium phosphates and sodium sulfates; thus, after so many years of contamination by detergents, sodium should be the major cation accounting for salinity in the SDet samples.

As wastewater discharges were mainly from laundry, it was to be expected that besides the detergents they contained some organic matter. Nevertheless, as shown in Table 1, SDet samples showed even lower levels of organic carbon than S samples. As we discussed above, this might mean that the stream of water with detergents washes out the organic matter.

General analysis of the pyrosequencing-derived dataset A total of 54,210 rRNA quality sequences with an average read length of 550 bp were generated through the 454 tag-sequencing.

At the phylum level, all samples showed similar numbers of OTUs, ranging from 154 to 187, with the exception of SDet3, which had 224. Nevertheless, at species level, S samples showed a larger number of OTUs, ranging from 1,043 to 1,296 OTUs, than SDet samples, which contained from 707 to 981 OTUs (Table 2).

The Shannon diversity index values (H') at phylum level suggest that bacterial diversity was slightly lower in SDet samples than in S samples; nevertheless, at species level the values were similar in all samples, ranging from 5.17 to 6.62. All these H' values are in the middle range normally reported for soils (e.g., Gołębiewski et al. 2014; Miyashita et al. 2013; Sun et al. 2014), suggesting that neither set of samples have an outstanding feature with respect to bacterial diversity.

Chao 1 richness estimates and rarefaction curves of all samples (Fig. 2) suggest that most of the estimated diversity contained within these communities was captured by our sequencing efforts. At phylum level, Chao 1 had very similar values among both SDet and S sets of samples, ranging from 154 to 187 OTUs, except for sample SDet3, where 225 OTUs were predicted. Nevertheless, at species level, SDet samples presented lower Chao 1, ranging from 921 to 1,012 OTUs, than S samples, which went from 1,067 to 1,326 OTUs. SDet samples also presented lower evenness values than S samples, meaning that the relative abundances of their taxa have more heterogeneous values. It is not common to find in soils Chao 1 values at species level inferior to 1,000 (see, for example, Yun et al. 2014; Jung et al. 2014; Will et al. 2010), which might mean that possibly the kind of soils we analyzed have a general low species richness. To our knowledge, this is the first study concerning the bacterial communities in the Leptosols of this region of Yucatán, thus further studies would be necessary to determine whether this low richness is indeed typical of the area. Rarefaction curves of all samples (Fig. 2) showed a leveling off, indicating that the number of analyzed reads was representative of the communities, both at a phylum and species level. Consistent with Chao 1, rarefaction curves at species level predicted lower species richness in SDet

Table 2 General analysis of the pyrosequencing-derived datasets. The number of OTUs, Shannon diversity, Chao I, and evenness were analyzed at 20 % (phylum level) and 3 % (species level) sequence dissimilarity for each soil sample. Abbreviations: S, soil samples without detergent; SDet, soil samples with detergents

Sample	#Reads	Phylum level				Species level			
		#OTUs	Shannon (H')	Chao I	Evenness	#OTUs	Shannon (H')	Chao I	Evenness
S1	7,911	154	3.74	154	0.74	1,140	6.29	1,153	0.89
S2	5,653	187	4.28	187	0.82	1,296	5.17	1,326	0.92
S3	6,450	183	4.15	183	0.80	1,043	6.04	1,067	0.87
SDet1	12,562	185	3.68	186	0.67	981	5.47	1,012	0.79
SDet2	11,295	175	3.61	175	0.74	707	6.62	727	0.79
SDet3	10,339	224	4.03	225	0.59	921	5.62	921	0.83

samples than in S samples. This lower species richness and diversity in SDet samples suggests that the diversity of niches in the soils subjected to detergents contamination is less than in S soils.

Bacterial community structures All the generated reads were classified as belonging to the domain Bacteria. The reads of all the six samples together clustered in a total of 32 phyla (data not shown), out of which 14 had a relative abundance >1 % in at least one sample. These major phyla were present in all samples, and together accounted for about 97 % of the total detected phyla for each sample (Fig. 3a). The dominant phylum in both sets of samples was *Proteobacteria*. Nevertheless, this was much more abundant in SDet samples, ranging from 61 to 72 %, than in S samples, where it ran from 30 to 52 %. According to Spain et al. (2009), who reviewed the abundances of *Proteobacteria* in soils considering only studies based on 16S rRNA gene sequences analysis, soils normally have a relative abundance of *Proteobacteria* of about 40 %± 8 %. Thus, the relative abundances that were found in SDet samples can be considered high. A possible explanation for this is that the other phyla normally present in S soils are more sensitive to the conditions generated by laundry sewage, thus their relative abundances are lower. Together, *Proteobacteria*, *Actinobacteria* and *Acidobacteria* accounted for about three quarters of the total phyla found in all samples. Nevertheless, *Actinobacteria* and *Acidobacteria*, in contrast to *Proteobacteria*, were less abundant in SDet samples than in most S samples (Fig. 3a). The minor phyla (with relative abundance <1 %) did not follow a distinctive pattern between S and SDet samples (data not shown).

At species level, there were 17 major OTUs (abundances >3 % in at least one sample), which together accounted for about 35 % of the species diversity in every sample (Fig. 3b). In S samples, the most abundant species was *Acidobacterium* sp., ranging from 11 to 26 %, and although it was also present in SDet samples, its abundance was much lower, ranging from 0.2 to 4.4 %, suggesting that the contaminated soil conditions affected it in a negative manner. *Acidobacterium* is a genus belonging to the phylum *Acidobacteria*, which is one of the most abundant bacterial groups in soils and sediments, but is very poorly known because its members are very difficult to culture (Ward et al. 2009). The three most abundant species in SDet samples were *Rhodomicrobium* sp., ranging from 2.8 to 13.5 %, *Hydrogenophaga* sp., ranging from 0.9 to 9.3 %, and *Thiobacillus* sp., ranging from 5.4 to 8.6 %. These species were detected in S soils as well, but with much lower abundances, suggesting that in contrast to *Acidobacterium* sp., they are relatively tolerant to SDet conditions. The rest of the major species did not follow a specific pattern when comparing S and SDet samples, with exception of *Phenylobacterium falsum* and *Thauera* sp., which were detected exclusively in SDet samples. It is interesting that *Thauera* has been reported as a genus isolated from wastewater treatment plants (Liu et al. 2013a), and it is known for its versatile metabolism and its ability to handle transitions from oxic to anoxic respiration (Liu et al. 2013b). It may be logical to find it in SDet soils, because these environments experience frequent changes from oxic conditions, when the soil turns relatively dry, to anoxic conditions, when the soil is soaked. *Rhodomicrobium* is a genus containing budding photoheterotrophic purple species (Duchow and Douglas 1949; Whittenbury and Dow 1977) that belongs to the order *Rhizobiales*, class *Alphaproteobacteria*; *Hydrogenophaga* is a genus of yellow-pigmented hydrogen-oxidizing bacteria (Willems et al. 1989; Yoon et al. 2008) belonging to the order *Burkholderiales*, class *Betaproteobacteria*; and *Thiobacillus* is a very heterogeneous genus of small, rod-shaped, autotrophic bacteria (Kelly and Wood 2000) belonging to the order *Hydrogenophilales*, class *Betaproteobacteria*. Thus, the three most abundant species in SDet samples belong to *Proteobacteria*, the most abundant phylum in the SDet samples.

Jaccard similarity coefficient For a general comparison of the degree of similarity of all the samples with each other, a heat map based on the Jaccard similarity index and a dendogram

Fig. 2 Rarefaction curves at
20 % (**a**) and 3 % (**b**)
dissimilarity. Taxa accumulation
curves as a function of the number
of sequences using resampling of
bacterial 16S rRNA gene
sequences. *Curves* were
calculated using the RDP tools.
Detergent-contaminated soils
(SDet); Soils without detergent
contamination (S)
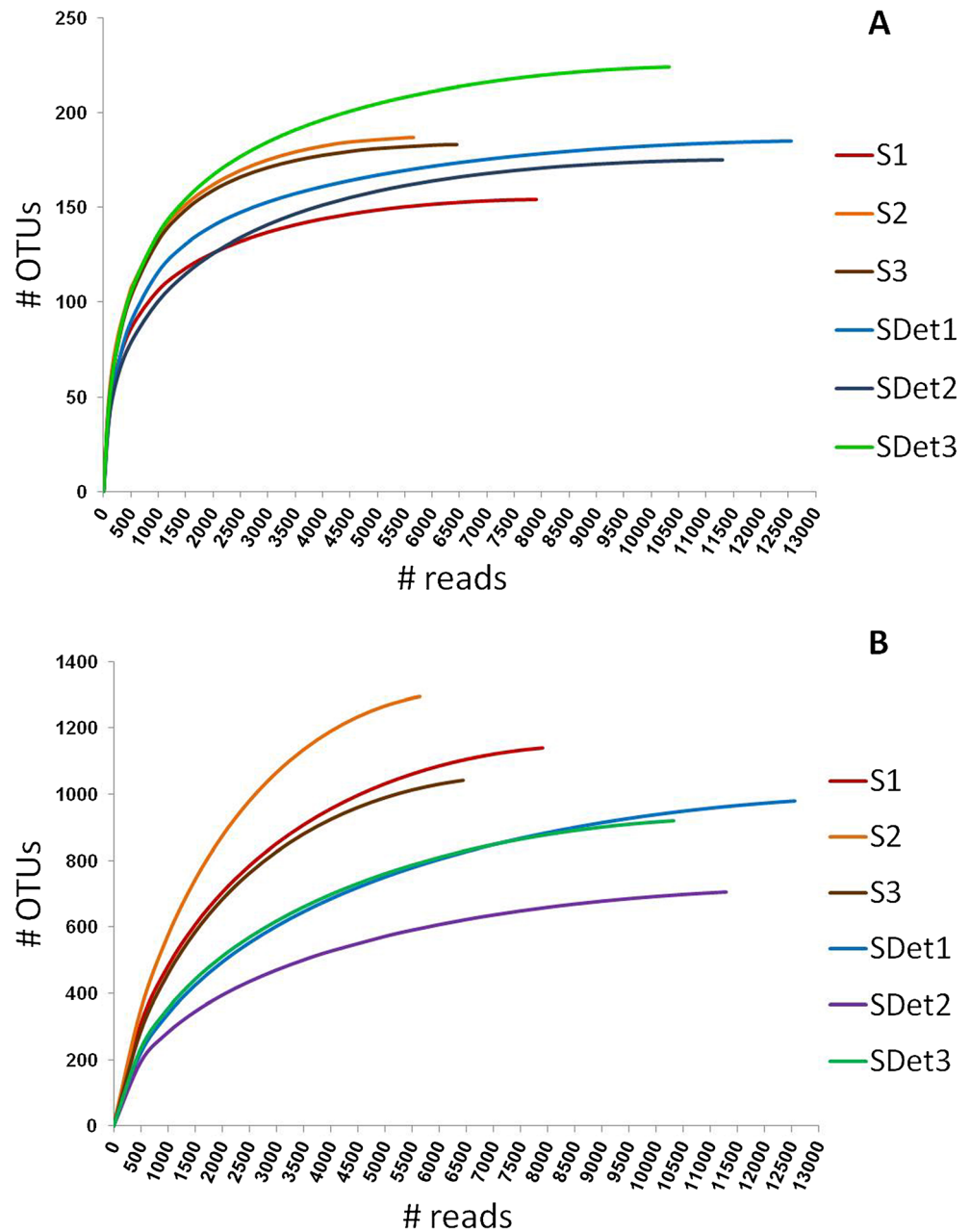
were constructed. As can be seen in Fig. 4, samples SDet2 and SDet3 proved to be the most similar to each other. Samples S1, S2, and S3 also showed similarity. Nevertheless, SDet1 unexpectedly appeared to be more similar to S samples than to SDet ones. The soils in this region are quite homogeneous (Aguila 2007; Ruíz-Garvia 2008); nevertheless, as the samples were not generated under controlled laboratory conditions, but came from a natural environment, there may be parameters that were not considered and may have some influence on the bacterial communities, such as the proximity of some plant species or general animal and human activity, thus it is not surprising that one sample does not fit in the general pattern. With the exception of this sample, the heat

map and the dendogram show that S samples are closer to each other than to SDet ones, and that the latter group together.

Canonical correspondence analysis The relationship among microbial community arrangement of contaminated and non-contaminated samples and environmental variables was analyzed by canonical correspondence analysis (CCA), which permits a straight analysis of community profiles with respect to specific environmental variables by constraining ordination axes to be linear combinations of environmental variables (Ter-Braak 1986).

As can be seen in Fig. 5, axes 1 and 2 explained 71.2 % of the total variance. Axis 1 correlated with a positive gradient of

Fig. 3 Relative abundances at 20 % (**a**) and 3 % (**b**) dissimilarity. Bacterial community composition in detergent-contaminated soils (SDet) and non-contaminated soils (S). The abundance is presented in terms of the percentage of the total bacterial sequences in each sample. Only taxa >1 % were considered
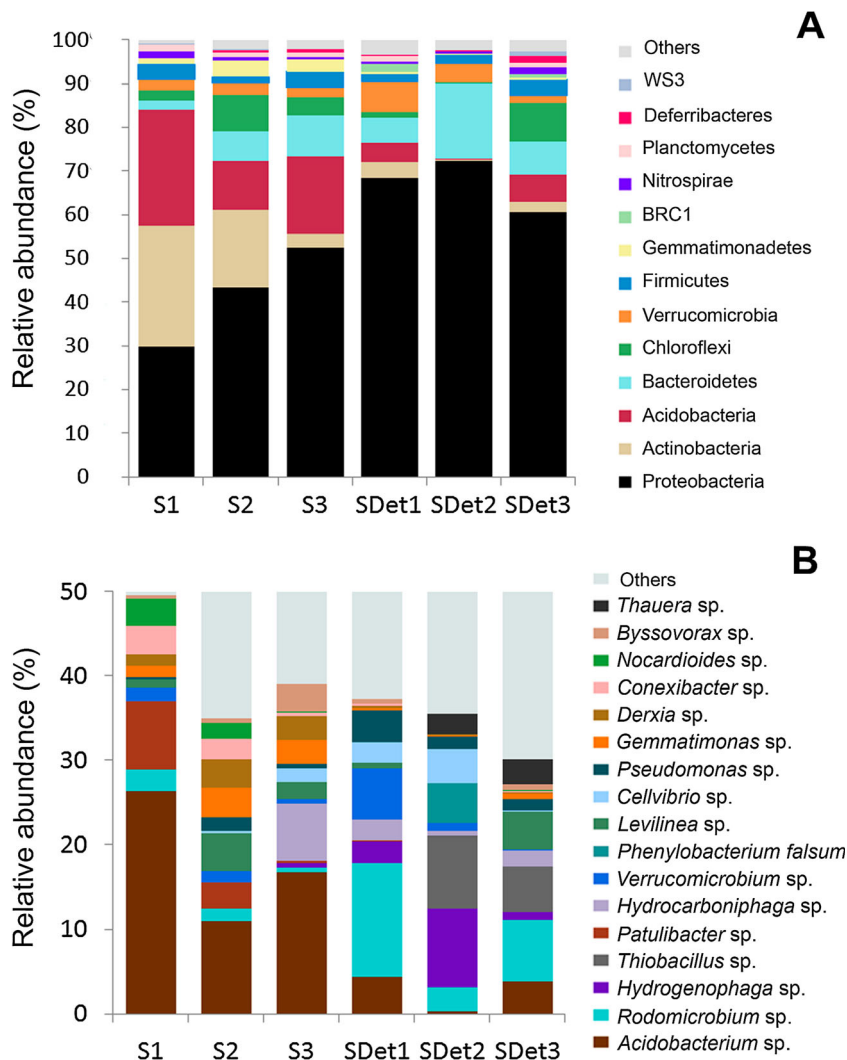
clay, REDOX, and organic carbon, and with a negative gradient of WC, EC, salinity, Na, and P, while Axis 2 correlated with a positive gradient of N and a negative gradient of pH. S samples presented a tendency to be located in the range of positive values of Axis 1, with higher values of clay, REDOX, and organic carbon, while SDet samples had a tendency to fall in the negative values of Axis 1, presenting higher values of WC, EC, salinity, Na, and P.

The bacterial genera that showed a higher correlation with the positive values of Axis 1 and, therefore, with S soils, were *Patulibacter, Nocardioides*, *Conexibacter, Acidobacterium, Derxia, Gemmatimonas,* and *Byssovorax,* while the genera that showed a higher correlation with the negative values of Axis 1 and, therefore, with SDet soils, were *Phenylobacterium, Thiobacillus, Hydrogenophaga, Thauera, Cellvibrio, Pseudomonas,* and *Rhodomicrobium.* The genera *Verrucomicrobium, Hydrocarboniphaga*, and *Levilinea* did not seem to be clearly related with any kind of soil. On the

other hand, the analysis showed that N and pH affect the bacterial community structures only to a very minor extent.

All these relationships found by canonical correspondence analysis are in agreement with the previous observations, suggesting that there is indeed a clear difference in the edaphic parameters between S samples and SDet samples and that those differences correlate with different bacterial communities.

Conclusions

The results obtained suggest that the soils subjected to wastewater discharges evinced several differences with respect to the soils that have not received the discharges. It cannot be said for sure that S soils were totally free of detergent contamination because possibly some minimal amounts of

Fig. 4 Jaccard similarity coefficient. Heat map based on the Jaccard similarity index comparing the most abundant taxa in detergent-contaminated soils (SDet) and in non-contaminated soils (S) as reveled by 16S rRNA gene pyrosequencing analysis
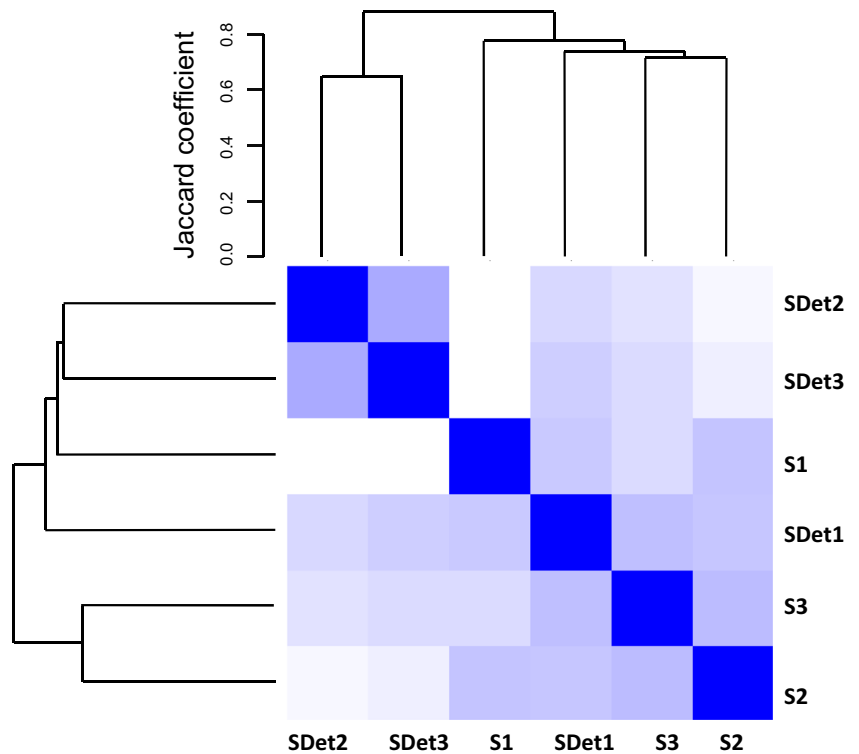
wastewater were spread around the backyard by human or animal activity, but there is no doubt that both kinds of soil are clearly different. Besides the higher water content, which was obviously due to the water discharge itself, the main edaphic
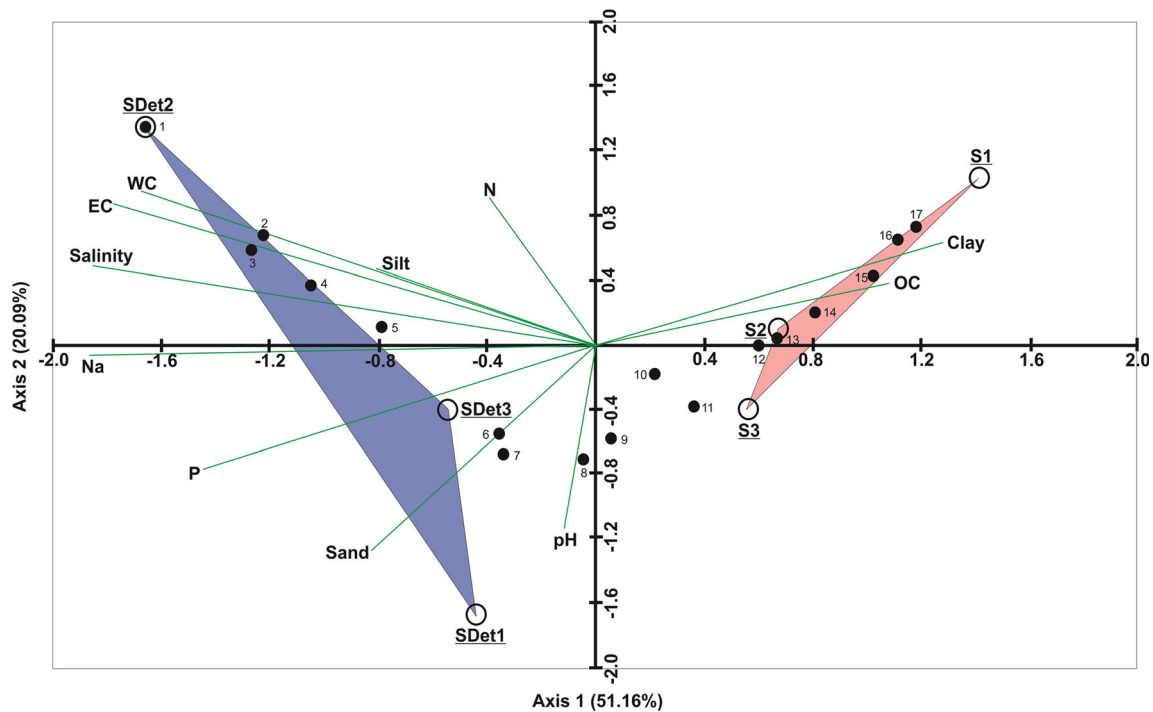

Fig. 5 Canonical correspondence analysis (CCA) showing correlations between the main bacterial genera and the soil variables. Only genera with a relative abundance >3 % in at least one sample were considered for the analysis. Green lines indicate the magnitude of measured variables associated with bacterial community structures. Circles represent the different soil samples. Genera are represented by solid dark circles: 1, *Phenylobacterium*; 2, *Thiobacillus*; 3, *Hydrogenophaga*; 4, *Thauera*; 5, *Cellvibrio;* 6, *Pseudomonas*; 7, *Rhodomicrobium*; 8, *Verrucomicrobium*; 9, *Hydrocarboniphaga*; 10, *Levilinea*; 11, *Byssovorax*; 12, *Gemmatimonas*; 13, *Derxia*; 14, *Acidobacterium*; 15, *Conexibacter*; 16, *Nocardioides*; 17, *Patulibacter*

parameters that had notoriously higher values in SDet samples than in S samples were sodium, salinity, and electrical conductivity, all of them tightly related with each other. This suggests that the sodium salts contained in the detergents are only partially lixiviated and an important part remains in the contaminated soils.

These edaphic differences showed to be related to different bacterial community structures. In fact, some of these changes in populations could have arisen from direct changes originating from the detergents, like salt concentrations or conductivity. The most prominent and general differences were that (i) SDet samples presented a lower species richness; (ii) *Proteobacteria* was the most abundant phylum in SDet soils, while *Actinobacteria* and *Acidobacteria* were the most abundant phyla in S soils; (iii) *Rhodomicrobium*, *Hydrogenophaga*, and *Thiobacillus* were the most abundant genera in SDet samples, while *Acidobacteria* dominated S samples.

All the results of this work together suggest that the wastewater contamination has modified several important soil characteristics and has caused modifications in the bacterial community structures, imposing a selective pressure that reduced the species richness; decreased the proliferation of some taxa, like the phyla *Actinobacteria* and *Acidobacteria*, and the genus *Acidobacteria*; and allowed the existence of some genera, like *Phenylobacterium* and *Thauera*, which do not thrive in non-contaminated soils.

With the continual increase of the human populations living in rural areas in Yucatán, these edaphic and bacterial modifications may contribute to bringing about changes in the ecological parameters of the region. They may have an impact, for example, in the relation between bacteria and plants, or between bacteria and microbial eukaryotes, causing ecological imbalances. Further work would be necessary to test this, but the present work provides the first insight to start wider studies.

In further work, it would also be interesting to investigate how the wastewater discharges modify the underground water characteristics and its bacterial communities. Because the ground in this region is mainly formed by a highly permeable karstic rock, it is to be expected that most of the waste water, with all its contaminants, percolate into the water lens, and as this lens is the only source of fresh drinkable water in the region, it is important to know the modifications it is undergoing.

On the other hand, as many important industrial biocatalyst reactions are developed in the presence of surfactants, we believe that the metagenome of the detergent-contaminated soils may be a good source of genes encoding enzymes able to perform well in those reactions, thus possibly providing a way to take some advantage of the contamination.

References

Aguila AE (2007) Soil fertility in calcareous tropical soils from Yucatan, Mexico, and Villa Clara, Cuba, affected by land use and soil moisture effects. Ph. D. dissertation, Georg August University, Göttingen, Germany

Bremmer JM (1996) Nitrogen-Total. In: Sparks DL (ed) Methods of soil analysis: part 3. chemical methods. agronomy monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 1085–1123

Cole JR, Wang Q, Cárdenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM (2009) The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141–D145

Dowd S, Zaragoza J, Rodríguez JR, Oliver MJ, Payton PR (2005) Windows. NET Network distributed basic local alignment search toolkit (W.ND-BLAST). BMC Bioinforma 6:93

Dowd SE, Callaway TR, Wolcott RD, Sun Y, McKeehan T, Hagevoort RG, Edrington TS (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiol 8:125. doi:10.1186/1471-2180-8-125

Duchow E, Douglas HC (1949) *Rhodomicrobium vannielii*, a new photoheterotrophic bacterium. J Bacteriol 58(4):409–416

Gardner WH (1986) Water content. In: Klute A (ed) Methods of soil analysis: part 1. Physical and Mineralogical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 493–541

Gee GW, Bauder JW (1986) Particle-size analysis. In: Klute A (ed) Methods of soil analysis: part 1. Physical and Mineralogical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 383–409

Gołębiewski M, Deja-Sikora E, Cichosz M, Tretyn A, Wróbel B (2014) 16S rDNA Pyrosequencing analysis of bacterial community in heavy metals polluted soils. Microb Ecol 67:635–647. doi:10.1007/s00248-013-0344-7

Gontcharova V, Youn E, Wolcott RD, Hollister EB, Gentry TJ, Dowd SE (2010) Black box chimera check (B2C2): a Windows-based software for batch depletion of chimeras from bacterial 16S rRNA gene datasets. Open Microbiol J 4:47–52

Hammer O, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Paleontogia Eletronica 4(1):1–9

Helmke PA, Sparks DL (1996) Lithium, sodium, potassium, rubidium, and cesium. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 551–574

Ishak HD, Plowes R, Sen R, Kellner K, Meyer E, Estrada DA, Dowd SE, Mueller UG (2011) Bacterial diversity in *Solenopsis invicta* and *Solenopsis geminata* ant colonies characterized by 16S amplicon 454 pyrosequencing. Microb Ecol 61:821–831. doi:10.1007/s00248-010-9793-4

Jeffries TC, Seymour JR, Gilbert JA, Dinsdale EA, Newton K, Leterme SSC, Roudnew B, Smith RJ, Seuront L, Mitchell JG (2011)

Substrate type determines metagenomic profiles from diverse chemical habitats. PLoS One 6(9):e25173. doi:10.1371/journal.pone.0025173

Jung J, Choi S, Hong H, Sung JS, Park W (2014) Effect of red clay on diesel bioremediation and soil bacterial community. Microb Ecol. doi:10.1007/s00248-014-0420-7

Kelly DP, Wood AP (2000) Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. Int J Syst Evol Microbiol 50:511–516

Kimes NE, Callaghan AV, Aktas DF, Smith WL, Sunner J, Golding BT, Drozdowska M, Hazen TC, Suflita JM, Morris PJ (2013) Metagenomic analysis and metabolite profiling of deep-sea sediments from the Gulf of Mexico following the deepwater horizon oil spill. Front Microbiol 4(50):1–17

Kumar PS, Brooker MR, Dowd SE, Camerlengo T (2011) Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. PLoS One 6(6):e20956. doi:10.1371/journal.pone.0020956

Kuo S (1996) Phosphorous. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 869–920

Liu B, Frostegård Å, Shapleigh JP (2013a) Draft genome sequences of five strains in the genus Thauera. Genome Announc 1(1):e00052–12. doi:10.1128/genomeA.00052-12

Liu B, Mao Y, Bergaust L, Bakken LR, Frostegård Å (2013b) Strains in the genus *Thauera* exhibit remarkably different denitrification regulatory phenotypes. Environ Microbiol. doi:10.1111/1462-2920.12142

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci U S A 104(27):11436–11440

Miyashita NT, Iwanaga H, Charles S, Diway B, Sabang J, 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

Nelson DW, Sommers LE (1996) Total carbon, organic carbon, and organic matter. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 961–1010

Patrick WH Jr, Gambrell RP, Faulkner SP (1996) REDOX measurements of soils. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 1085–1123

Rhoades JD (1996) Salinity: electrical conductivity and total dissolved solids. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 417–436

Ruíz-Garvia C (2008) Production potential and ecosystem quality of secondary forest recovered from agriculture – tools for land use decisions. Ph. D. dissertation, Georg August University, Göttingen, Germany

Spain AM, Krumholz LR, Elshahed MS (2009) Abundance, composition, diversity and novelty of soil Proteobacteria. ISME J 3:992–1000

Sun B, Wang F, Jiang Y, Li Y, Dong Z, Li Z, Zhang X-X (2014) A long-term field experiment of soil transplantation demonstrating the role of contemporary geographic separation in shaping soil microbial community structure. Ecol Evol 4(7):1073–1087. doi:10.1002/ece3.1006

Ter-Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67(5):1167–1179

Thomas GW (1996) Soil pH and soil acidity. In: Sparks DL (ed) Methods of soil analysis: part 3. Chemical Methods. Agronomy Monograph. American Society of Agronomy-Soil Science Society of America, Madison, pp 475–490

Vos M, Quince C, Pijl AS, de Hollander M, Kowalchuk GA (2012) A comparison of rpoB and 16S rRNA as markers in pyrosequencing studies of bacterial diversity. PLoS One 7(2):e30600. doi:10.1371/journal.pone.0030600

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M, Xie G, Haft DH, Sait M, Badger J, Barabote RD, Bradley B, Brettin TS, Brinkac LM, Bruce D, Creasy T, Daugherty SC, Davidsen TM, DeBoy RT, Detter JC, Dodson RJ, Durkin AS, Ganapathy A, Gwinn-Giglio M, Han CS, Khouri H, Kiss H, Kothari SP, Madupu R, Nelson KE, Nelson WC, Paulsen I, Penn K, Ren Q, Rosovitz MJ, Selengut JD, Shrivastava S, Sullivan SA, Tapia R, Thompson LS, Watkins KL, Yang Q, Ch Y, Zafar N, Zhou L, Kuske CR (2009) Three genomes from the phylum acidobacteria provide Insight into the lifestyles of these microorganisms in soils. Appl Environ Microb 75(7):2046–2056. doi:10.1128/AEM.02294-08

Wessén E, Hallin S, Philippot L (2010) Differential responses of bacterial and archaeal groups at high taxonomical ranks to soil management. Soil Biol Biochem 42:1759–1765

Whittenbury R, Dow CS (1977) Morphogenesis and differentiation in rhodomicrobium vannielii and other budding and prosthecate bacteria. Bact Rev 41(2):754–808

Will C, Thürmer A, Antje Wollherr A, Nacke H, Herold N, Schrumpf M, Gutknecht J, Wubet T, Buscot F, Daniel R (2010) Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. Appl Environ Microb 76(20):6751–6759

Willems A, Busse J, Goor M, Pot B, Falsen E, Jantzen E, Hoste B, Gillis M, Kersters K, Auling G, DeLey J (1989) Hydrogenophaga, a new genus of hydrogen-oxidizing bacteria that includes *Hydrogenophaga flava* comb. nov. (formerly *Pseudomonas flava*, *Hydrogenophaga palleronii* (formerly *Pseudomonas palleronii*), *Hydrogenophaga pseudoflava* (formerly *Pseudomonas pseudoflava* and "*Pseudomonas carboxydoflava*"), and *Hydrogenophaga taeniospiralis* (formerly *Pseudomonas taeniospiralis*). Int J Syst Bact 39(3):319–333

Yoon K-S, Tsukada N, Sakai Y, Ishii M, Igarashi Y, Nishihara H (2008) Isolation and characterization of a new facultatively autotrophic hydrogen-oxidizing betaproteobacterium, *Hydrogenophaga* sp. AH-24. FEMS Microbiol Lett 278:94–100. doi:10.1111/j.1574-6968.2007.00983.x

Yun J, Ju Y, Deng Y, Zhang H (2014) bacterial community structure in two permafrost wetlands on the Tibetan plateau and Sanjiang plain. China Microb Ecol. doi:10.1007/s00248-014-0415-4

Zoller U (1998) The case of persistent ("hard") nonionic surfactants in the environment. Toxicol Environ Chem 66(1–4):145–157