npg

# ORIGINAL ARTICLE

# Spatial patterns of microbial diversity and activity in an aged creosote-contaminated site

Shinjini Mukherjee[1], Heli Juottonen[1,4], Pauli Siivonen[1], Cosme Lloret Quesada[1],
Pirjo Tuomi[2], Pertti Pulkkinen[3] and Kim Yrjälä[1]

[1]Department of Biosciences, MEM Group, University of Helsinki, Helsinki, Finland; [2]Golder Associates Oy,
Helsinki, Finland and [3]Finnish Forest Research Institute (METLA), Läyliäinen, Finland

Restoration of polluted sites via *in situ* bioremediation relies heavily on the indigenous microbes
and their activities. Spatial heterogeneity of microbial populations, contaminants and soil chemical
parameters on such sites is a major hurdle in optimizing and implementing an appropriate
bioremediation regime. We performed a grid-based sampling of an aged creosote-contaminated site
followed by geostatistical modelling to illustrate the spatial patterns of microbial diversity and
activity and to relate these patterns to the distribution of pollutants. Spatial distribution of bacterial
groups unveiled patterns of niche differentiation regulated by patchy distribution of pollutants and
an east-to-west pH gradient at the studied site. Proteobacteria clearly dominated in the hot spots of
creosote pollution, whereas the abundance of Actinobacteria, TM7 and Planctomycetes was
considerably reduced from the hot spots. The pH preferences of proteobacterial groups dominating
in pollution could be recognized by examining the order and family-level responses. Acidobacterial
classes came across as generalists in hydrocarbon pollution whose spatial distribution seemed to
be regulated solely by the pH gradient. Although the community evenness decreased in the heavily
polluted zones, basal respiration and fluorescein diacetate hydrolysis rates were higher, indicating
the adaptation of specific indigenous microbial populations to hydrocarbon pollution. Combining
the information from the kriged maps of microbial and soil chemistry data provided a
comprehensive understanding of the long-term impacts of creosote pollution on the subsurface
microbial communities. This study also highlighted the prospect of interpreting taxa-specific spatial
patterns and applying them as indicators or proxies for monitoring polluted sites.
*The ISME Journal* (2014) **8**, 2131–2142; doi:10.1038/ismej.2014.151; published online 8 August 2014

## Introduction

The pivotal role of soil microbes in maintaining
overall ecosystem functioning has become very
evident in recent years (Falkowski *et al.*, 2008; Van
Der Heijden *et al.*, 2008). Constant exposure of soil
to anthropogenic disturbances such as pollution
might lead to compromised soil health and
ecosystem functioning by altering the biodiversity
(Chapin *et al.*, 2000; Loreau *et al.*, 2001; Ager *et al.*,
2010). Restoration of polluted ecosystems requires
an in-depth understanding of microbial responses
to contaminants, both in terms of community
composition and function. A majority of the studies
dealing with these questions involve experimental
amendment of contaminants followed by assess-
ment of microbial responses measured over periods

ranging from days to few years. Depending on the
level of pollution, however, the time required to
observe ecologically relevant responses towards
pollution might rather be in the order of few decades
(Lindstrom *et al.*, 1999; Davis *et al.*, 2004; Fahy *et
al.*, 2005; Oorts *et al.*, 2006). In addition, in the
restoration of polluted sites, aged contamination is
one of the most important concerns. In this context,
aged polluted sites with gradients of contaminants
can provide excellent models to study long-term
responses of microbial communities and their
activity towards different contamination levels
(Ramsey *et al.*, 2005).

The present study site is part of an old wood
preservation facility located in southeastern Finland
where creosote had been used to impregnate rail-
road sleepers between 1947 and 1958, resulting in
high concentrations of creosote in the soil. Creosote
is a dense, nonaqueous phase liquid that has led to
the contamination of soil and groundwater at many
wood preservation sites worldwide. It is a by-product
of coal tar distillation, composed of ∼85% polycyclic
aromatic hydrocarbons (PAHs), 10% phenolic
compounds and 5% *N*-, *S*- and *O*- heterocyclics
(Mueller *et al.*, 1989). Typically, 20–40% of creosote

Correspondence: S Mukherjee, Department of Biosciences, Faculty
of Biological and Environmental Sciences, University of Helsinki,
P.O. Box 56, Helsinki FI-00014, Finland.
E-mail: shinjini.mukherjee@helsinki.fi
[4]Current address: Department of Genetics and Ecology, Uppsala
University, Uppsala, Sweden.

2132

is composed of 16 (US Environmental Protection Agency (EPA) defined) priority PAH pollutants (Kulik et al., 2006).

Spatial heterogeneity of pollutants as well as other soil chemical parameters is a major concern while pre-evaluating the strategies of bioremediation or evaluating the success of remediation of polluted sites (Sturman et al., 1995; Gerhardt et al., 2009). Such heterogeneity might result from accidental spills or production/treatment facilities that create hot spots of pollutants. Subsequently, the spatial variation of pollutants and other soil chemical parameters may regulate the spatial patterns of microbial communities (Törneman et al., 2008; Bengtsson et al., 2010). Association between spatial patterns of environmental factors and the distribution of microbial taxa as well as their activities can be elucidated via geostatistical modelling (Philippot et al., 2009; Bru et al., 2010; Enwall et al., 2010; Wessén et al., 2011). Furthermore, the choice of taxonomic level to address the microbial community responses is a crucial issue. Some recent studies have highlighted the ecologically coherent strategies displayed by higher taxonomic levels in bacteria (Fierer et al., 2007; Philippot et al., 2009), indicating that it is not always necessary to dissect down to the species level to observe an ecologically relevant response. It has also been suggested that 'even if the species is a fundamental category of biological organization, it is important to understand ecology at every ecologically meaningful level of organization' (Philippot et al., 2010).

In this study, we analysed the spatial patterns of bacterial diversity and microbial activity in an aged creosote-contaminated site. We used an intensive grid-based sampling strategy to address the following questions: Does the spatial pattern of bacterial community composition in the polluted site reflect the distribution of pollutants? Is the spatial distribution of specific taxonomic groups regulated by different soil chemical factors based on the ecological and physiological properties of the groups? Can the taxa-specific spatial patterns be related to pollution levels and further used as indicators in monitoring a polluted site? And is there an association between spatial patterns of microbial activity and the pollutant distribution?

## Materials and methods

### Study site and sampling
The former wood treatment area ($\sim$7 Ha) investigated in this study is located in southeastern Finland (60° 55' 00"N and 27° 26' 00"E, Luumäki). The vegetation on this site was dominated by small Scots pines (Pinus sylvestris) and Silver birches (Betula pendula). The highest tree densities in the area were $\sim$0.8 trees m$^{-2}$. Ground vegetation was mainly lichens and mosses, most common

species being Cladonia sp. and Dicranum majus (Van Oss, 2012). No clear impact of pollution was observed on the vegetation. As per the previous site investigations (by Golder Associates, Helsinki, Finland) $\sim$1.3 Ha area in the central part, where the treatment facility was located, was characterized to be contaminated. Sampling for this study was performed in September 2011. A sampling grid was created for the site (Supplementary Figure S1). The central area that was anticipated to be more contaminated was sampled using a $20 \times 20$ m grid. Surrounding, less contaminated areas were investigated using a $40 \times 40$ or $20 \times 80$ m grid. In each plot, five 0.5 m deep sampling pits were excavated with a shovel. Before excavation, the uppermost few centimetres of organic soil was removed as the contamination was expected to be in the sandy soil below the organic layer, and organic material can affect hydrocarbon analyses. Top soil was collected from 0 to 0.1 m, and samples from five pits (four from the corners of the plot/square and one from centre) were combined to form composite samples representing each plot. Approximately 0.25 l of soil was collected from each pit and combined as 1 l composite sample in a gas tight plastic sampling bag (Tub-Ex Rilsan bag PA11, Tars, Denmark). The composite sample was mixed well in the bag and then divided into portions for different analyses. Samples were stored cooled for 1–2 days until they were delivered to laboratory for soil chemical and microbial analysis. The microbial activity assays were performed within 1 week of sample collection and the DNA extraction was performed $\sim$4 months after sampling from soil stored at $+4\,^\circ$C.

### Soil chemical analysis
Soil chemical analyses were performed using accredited methods in ALS Finland Oy (Helsinki, Finland), SGS Inspection Services Oy (Helsinki, Finland) and the Rovaniemi unit of Finnish Forest Research Institute (Rovaniemi, Finland). For the measurement of soil nutrients, soil samples were digested by the closed wet HNO$_3$-HCl digestion method in a microwave (CEM MDS-2000; CEM Corp., Matthews, NC, USA) and the extract was analysed by a iCAP 6500 DUO ICP-emission spectrometer (Thermo Scientific, Cambridge, UK). Concentrations of trace elements were determined according to SFS-ISO 11466. Total carbon (C) and organic carbon were analysed according to ISO 10694 and total nitrogen (N) was measured according to ISO 13878. Electrical conductivity was determined according to ISO 11265 and pH was measured according to ISO 10390. Hydrocarbon content was analysed by means of gas chromatography/flame ionization detector using method ISO EN 9377-2 and ISO 16703. PAHs were analysed using n-hexane/acetone (1:1) extraction and by gas chromatography/mass spectrometry according to the methods EPA 8270, EPA 8131, EPA 8091, ISO 6468 and ISO 18287.

*Soil DNA extraction, 16S rRNA gene pyrosequencing and sequence data analysis*
DNA was extracted from 500 mg of soil with NucleoSpin Soil (Macherey-Nagel, GmbH & Co. KG, Düren, Germany) according to the manufacturer's instructions. The bead-beating step was performed using Precellys24-Dual (Bertin Technologies, Montigny-le-Bretonneux, France). DNA purity and quantity were checked with Nanodrop (Nano-Drop Technologies, Wilmington, DE, USA) and agarose gel electrophoresis.

The 16S rRNA gene PCR amplification was carried out using a two-step PCR protocol. This method has an advantage of higher reproducibility and reduction of potential bias introduced by barcoded primers (Berry *et al.*, 2011). In the first step, the V1–V3 regions of the 16S rRNA gene were amplified in triplicate from each sample with the universal bacterial primers 27F (5′-ACCCTAG GTGCTCACGAC-AGAGTTTGATCMTGGCTCAG-3′) and 519R (5′-GCATTGCCTCAGAACGGT-GTATTAC CGCGGCTGCTG-3′). The primers 27F and 519R were modified by adding an 18-bp overhang sequence in front of the primer (overhang sequence is indicated as separated by '-' from the original primer sequences). The PCR reaction was performed using Phusion polymerase (Finnzymes, Espoo, Finland) with the HF buffer and cycling conditions consisted of an initial denaturation at 98 °C for 30 s, followed by 20 cycles of 98 °C for 10 s, 65 °C for 30 s and 72 °C for 10 s, and a final extension for 5 min. The triplicate PCR products from the first step were pooled and purified using Agencourt AMPure PCR purification system (Beckman Coulter, Danvers, MA, USA). The pooled and purified PCR products were diluted in a 1:10 ratio to a final concentration of $\sim 1$–5 ng $\mu l^{-1}$. These diluted products were used as a template for the second step PCR with primers with A and B adapters and specific barcode primers for each sample (10 cycles). A final purification as described above was performed before the sequencing. The size and quality of the PCR products were determined with Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and also by agarose gel electrophoresis. DNA concentrations were measured with a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). The sequencing was carried out at the Institute of Biotechnology in the University of Helsinki using the 454-GS FLX Titanium protocol (454 Life Sciences/Roche Diagnostics, Branford, CT, USA) that yields reads length of $\sim 400$ bp.

The processing of the 16S rRNA pyrosequences using mothur v.1.30.2 started with the flowgram data (Schloss *et al.*, 2009). The sequences contained in the flowgrams (SFF file) were sorted into groups according to barcode sequences. The standard operating procedure (http://www.mothur.org/wiki/Schloss_SOP) was followed for the analysis (Schloss *et al.*, 2011). A total of 112 729 good-quality sequences were recovered from 28 samples that were used for downstream analysis. A 97% identity in the 16S rRNA gene sequence was used as the operational taxonomic unit (OTU) definition. A total of 12 250 OTUs were detected from all the samples. In order to determine the changes in microbial community on higher taxonomic levels (phylum, class, order and family), OTUs were classified using Ribosomal database training set with the 'classify.otu' command in mothur that provides consensus taxonomy for every OTU.

*Sequence data availability*
Sff files for all the samples can be found in European Nucleotide Archive under the accession numbers ERR505173–ERR505200.

*Basal respiration rate and FDA hydrolysis*
Microbial basal respiration in the samples was evaluated in triplicates in the laboratory as $CO_2$ production after 24 h of incubation at constant temperature ($+ 14$ °C). Next, 20 ml of soil sample was weighed into 120 ml glass bottles and $CO_2$ evolution determined by taking 100 µl air sample from the headspace and analysing $CO_2$ concentration by gas chromatography on a Hewlett Packard 6890 (Agilent Technologies, Palo Alto, CA, USA) equipped with a thermal conductivity detector (200 °C, 15.0 ml min$^{-1}$ reference flow, 3.0 ml min$^{-1}$ makeup flow), and packed column Agilent 19006-80110. Helium was used as carrier gas (60.0 ml min$^{-1}$). The mean microbial respiration rate per plot was calculated as $CO_2$ µl per g soil per h.

Fluorescein diacetate (FDA) hydrolysis assay was used as a general indicator of soil hydrolytic activity, as it measures the activities of proteases, lipases and esterases that are capable of cleaving the fluorogenic FDA into fluorescein. FDA hydrolysis was analysed according to a protocol from Green *et al.* (2006). Mean FDA hydrolysis rate per plot was calculated as µg fluorescein per g soil per 3 h.

*Statistical analysis*
Nonmetric multidimensional scaling based on Bray–Curtis distances and NPMANOVA (nonparametric multivariate analysis of variance) were performed in PAST v 2.17c (Hammer *et al.*, 2009). NPMANOVA is the multivariate analogue of analysis of variance and is normally used for ecological taxa-in-samples data, where groups of samples are to be compared (Anderson, 2001). Redundancy analysis (RDA) to relate bacterial community patterns at different taxonomic levels to environmental factors was performed using the vegan package in R version 3.0.2 (http://www.r-project.org). Hellinger transformation of taxon relative abundance was carried out before performing RDA to minimize the effect of zeroes in the community data (Legendre and Gallagher, 2001). Significance of the RDA models and the explanatory factors were tested by 200 permutations. Correlations between the microbial and soil chemical properties

were calculated in Prism software (GraphPad Software Inc., San Diego, CA, USA).

Geostatistical analysis was performed in ArcGIS 10.1 (ArcGIS Desktop, ESRI, Redlands, CA, USA). Semivariogram modelling was performed for quantifying the scale of heterogeneity (patch size) and the spatially dependent predictability of the different microbiological and soil chemical properties studied (Goovaerts, 1998, 1999). A suitable model was chosen for each parameter by comparing 11 models, and the model was selected based on the best fit to the empirical semivariogram as well as cross-validation. Cross-validation uses all the data to estimate the trend and autocorrelation models by removing each data location one at a time and predicting the associated data value. The predicted and actual values at the location of the omitted point are compared; this procedure is then repeated for a second point, and so on. For all points, cross-validation compares the measured and predicted values. Spherical, pentaspherical or stable models produced the best fit for most of the parameters in this study (Supplementary Table S2). Ordinary kriging was used to create prediction maps for various microbiological and soil chemical parameters. Ordinary kriging assumes the model: $Z(s) = \mu + \varepsilon(s)$; where $Z(s)$ is the variable of interest, $\mu$ is the deterministic surface (the trend) and $\varepsilon(s)$ is a spatially autocorrelated error. The symbol $s$ indicates the location; containing the spatial x (longitude) and y (latitude) coordinates.

## Results

### Soil chemistry

Spatial distribution of PAHs at the former wood preservation site showed several hot spots, surrounded by less contaminated soil areas (Figure 1). The PAH concentrations in the hot spots ranged up to 714 p.p.m. Among the 16 PAHs analysed, the highest concentrations were measured for 3 to 5 ring
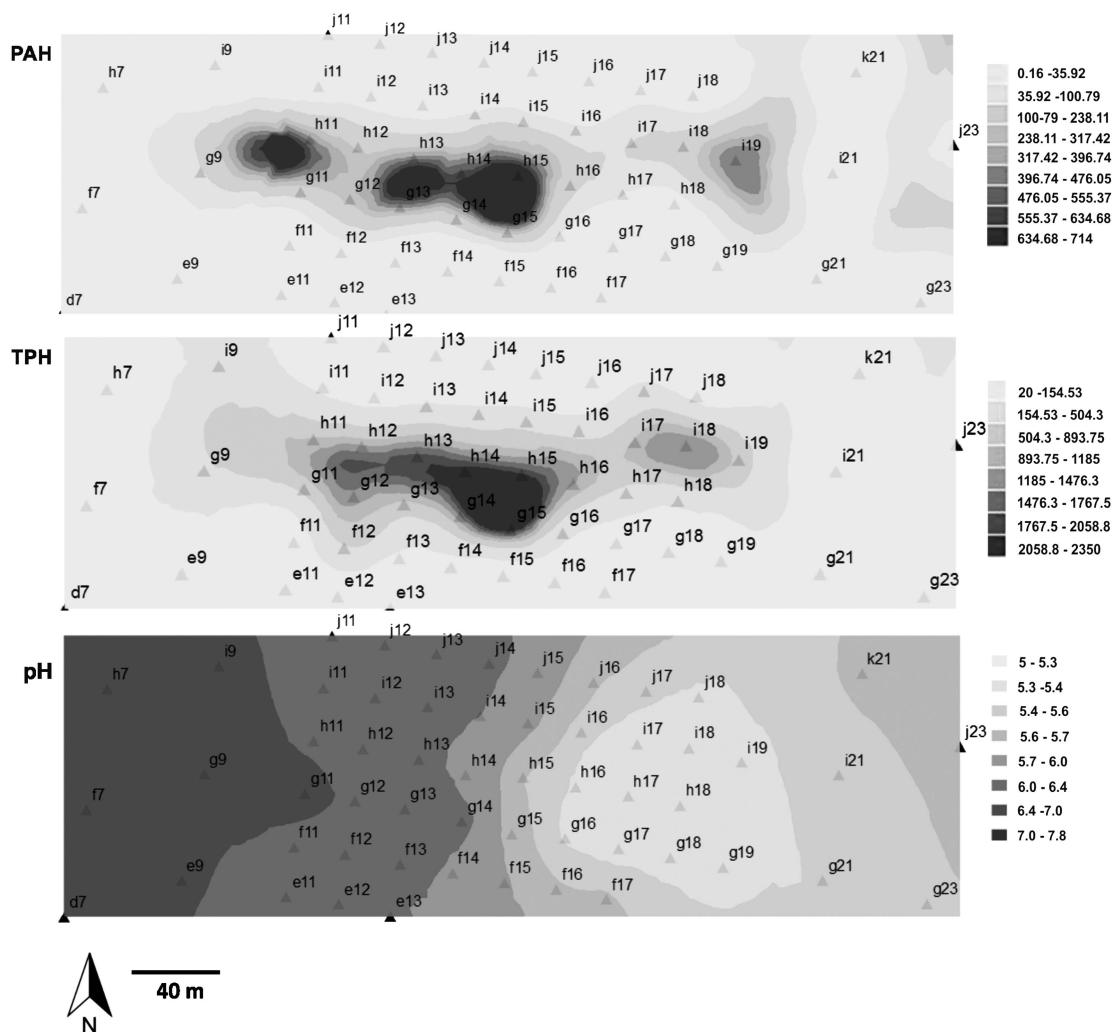


**Figure 1** Kriged maps of distribution of PAHs (p.p.m.), TPHs (p.p.m.) and pH on the creosote-polluted site. The grey triangles represent the lower left corner of each sampling plot marked by letters. The colour codes are given next to each map. A full colour version of this figure is available at *The ISME Journal* online.

compounds such as pyrene (4 ring), fluoranthene (4), benz(a)anthracene (4), chrysene(4), benzo(b)fluoranthene (4 + 1), anthracene (3), benzo(k)fluoranthene (4) and benzo[a]pyrene (5). Low-molecular-weight PAHs were found in extremely low concentrations in the soil samples indicating their degradation, evaporation or transportation from the top soil. Total hydrocarbon content (C10–C40; referred to as TPH in the following text and figures) was also high with the hot spots overlapping with the PAH hot spots (Figure 1). In addition to patchy distribution of pollutants, the site displayed a pH gradient (5 to 7.8) with pH decreasing towards the east. Soil texture was classified as sandy with mostly medium and very fine sand particles (Supplementary Table S1).

*The α- and β-diversity patterns in the polluted site*
Nonmetric multidimensional scaling revealed that the major separation of communities was based on the level of pollution (Figure 2). Bacterial communities in clean soil and low concentration of pollution clearly differed from those in moderate and high levels of pollution. Based on NPMANOVA, communities in soils with 0–10 p.p.m. PAH were different from those of 80–200 p.p.m. soils ($P = 0.01$) and 400–750 p.p.m. PAH soils ($P = 0.0003$). Soils with 15–50 p.p.m. PAH showed a significant difference with 80–200 p.p.m. PAH ($P = 0.01$) and 400–750 PAH ($P = 0.003$) soil bacterial communities. At pollution levels above 200 p.p.m., there was no clear separation between the communities based on PAH level. It should however be noted that TPH concentrations were highest in some of the samples with moderate levels of PAH pollution.

Bacterial diversity (Shannon index) correlated negatively with PAH concentration (Spearman's $r = -0.45$, $P = 0.008$; Supplementary Table S3). The major individual PAHs showing negative correlation with diversity were fluoranthrene, pyrene,
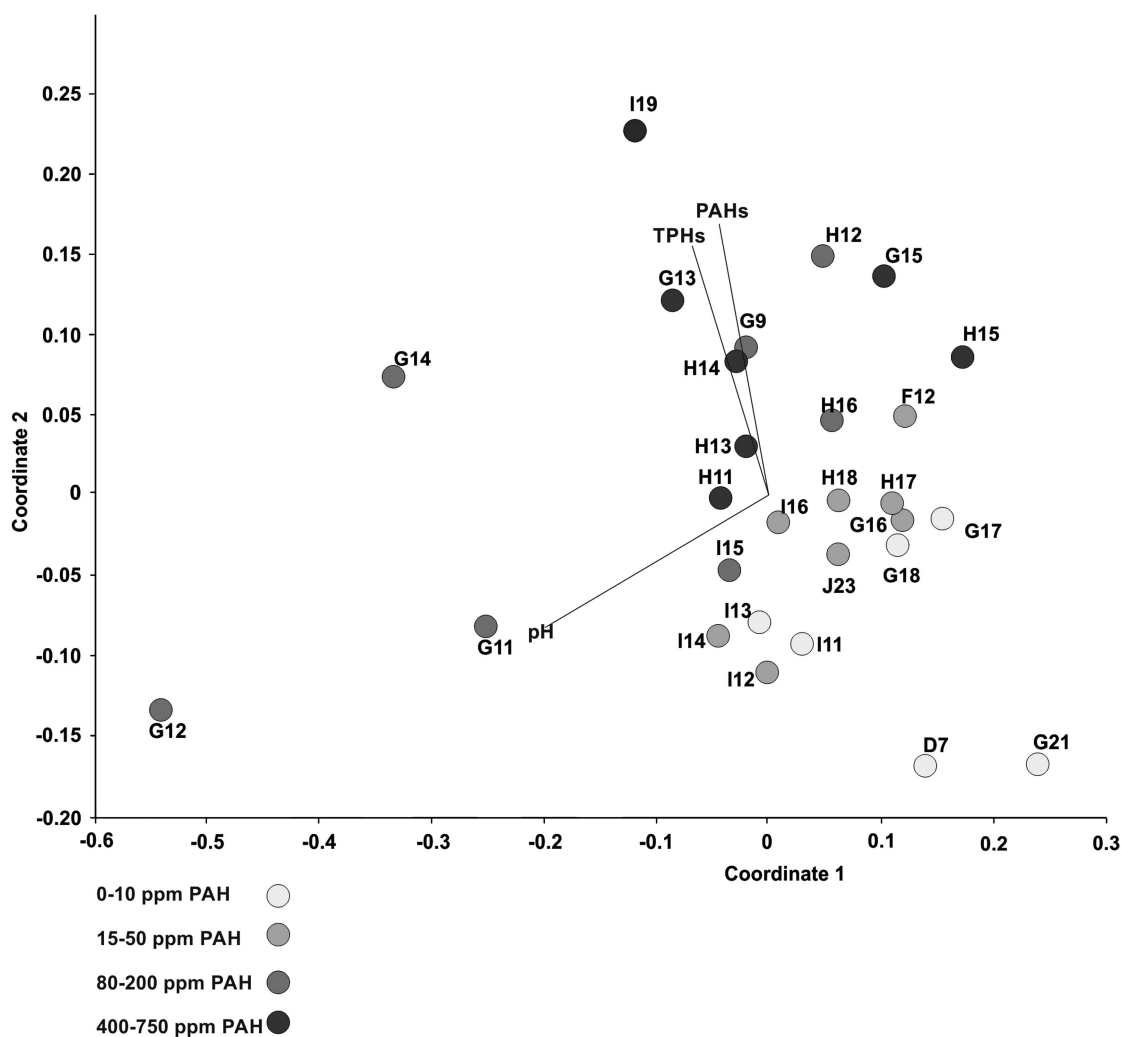


**Figure 2** Nonmetric multidimensional scaling plot of bacterial 16S rRNA gene pyrosequencing data (OTU cutoff 97%). Letters accompanying the symbols represent the sampling plot and the colour codes denote different levels of PAH pollution. Soil chemical parameters PAHs, TPHs and pH were included in the analysis and the length of the arrow represents the strength of correlation. A full colour version of this figure is available at *The ISME Journal* online.

benzo(*a*)anthracene, chrysene and benzo(*b*)fluoran-threne. Species evenness was similarly lower at higher pollution levels. In addition to total PAHs, TPH concentrations and six different PAHs were also correlated with evenness (Supplementary Table S3).

*Spatial patterns of bacterial community composition related to patchy distribution of pollution and pH gradient*
To further relate the bacterial community structure to spatial variation of pollutants and soil chemistry, we carried out geostatistical analysis. Kriged maps based on semi-variogram modelling indicated differences in spatial distribution at phylum level (Figure 3). Proteobacteria (37–75%), Acidobacteria (15–31%), Bacteroidetes (2–15%) and Actinobacteria (1–11%) were among the most abundant phyla in the subsurface bacterial communities, whereas Planctomycetes, Armatimonadetes, Verrucomicrobia, Gemmatimonadetes and TM7 were relatively lower in abundance. Proteobacteria displayed a patchy distribution; the hot spots of proteobacterial abundance appeared to coincide with the hot spots of pollutants and the range of spatial autocorrelation was similar for spatial patterns of Proteobacteria and pollutant distribution (Supplementary Table S2). On the other hand, hot spots of Actinobacteria, TM7,

Armatimonadetes and Planctomycetes overlapped with the cold spots of pollution. Acidobacteria, Verrucomicrobia and Bacteriodetes displayed a distribution pattern that seemed to be related to the pH gradient. To confirm these spatial community patterns and the environmental factors explaining them, we carried out RDA. The RDA model including PAH, TPH, pH and carbon to nitrogen ratio (C/N) explained 32% of community variation at the phylum level, and PAH ($P = 0.05$), TPH ($P = 0.01$) and pH ($P = 0.04$) were the factors best explaining the variation (Figure 4a). A pollution gradient separated Proteobacteria, abundant in the pollution hot spots, from phyla less abundant in the hot spots such as Actinobacteria. On the other hand, Acidobacteria and Bacteriodetes were clearly separated along a pH gradient.

As phylum is a relatively high-level taxonomic unit that often encompasses metabolically diverse group of organisms, we further decided to analyse the response to the environmental gradients at class, order and family level. At class level, the environmental factors explained 40% the community variation, with PAH, TPH and pH as the best explanatory factors ($P < 0.01$; Figure 4b). Out of the four Proteobacterial classes observed in this study, the abundances of Beta- and Gammaproteobacteria
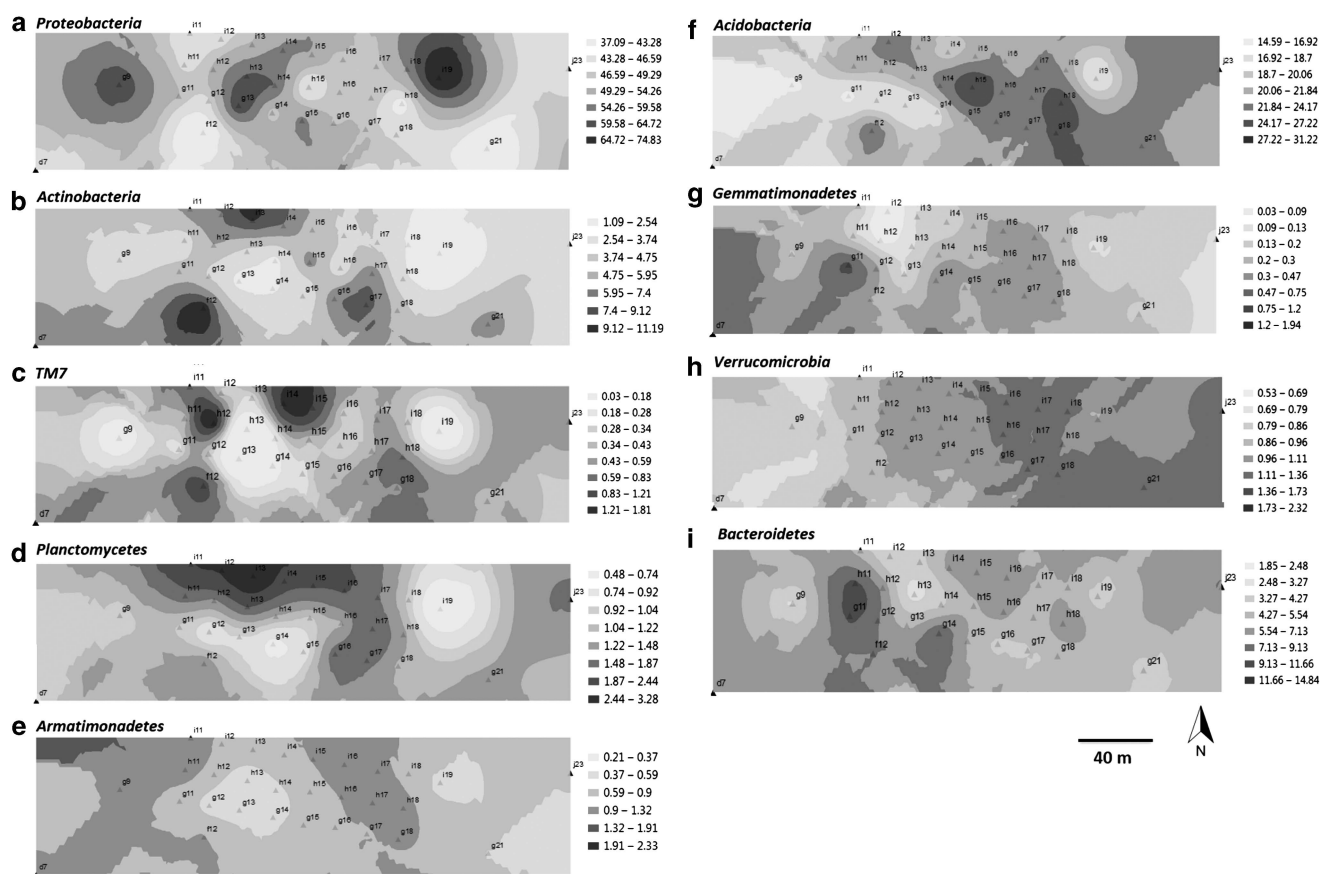


**Figure 3** Kriged maps of distribution of relative abundances (%) of nine major bacterial phyla (marked as **a–i**) observed in this study. The colour codes are given next to each map. A full colour version of this figure is available at *The ISME Journal* online.
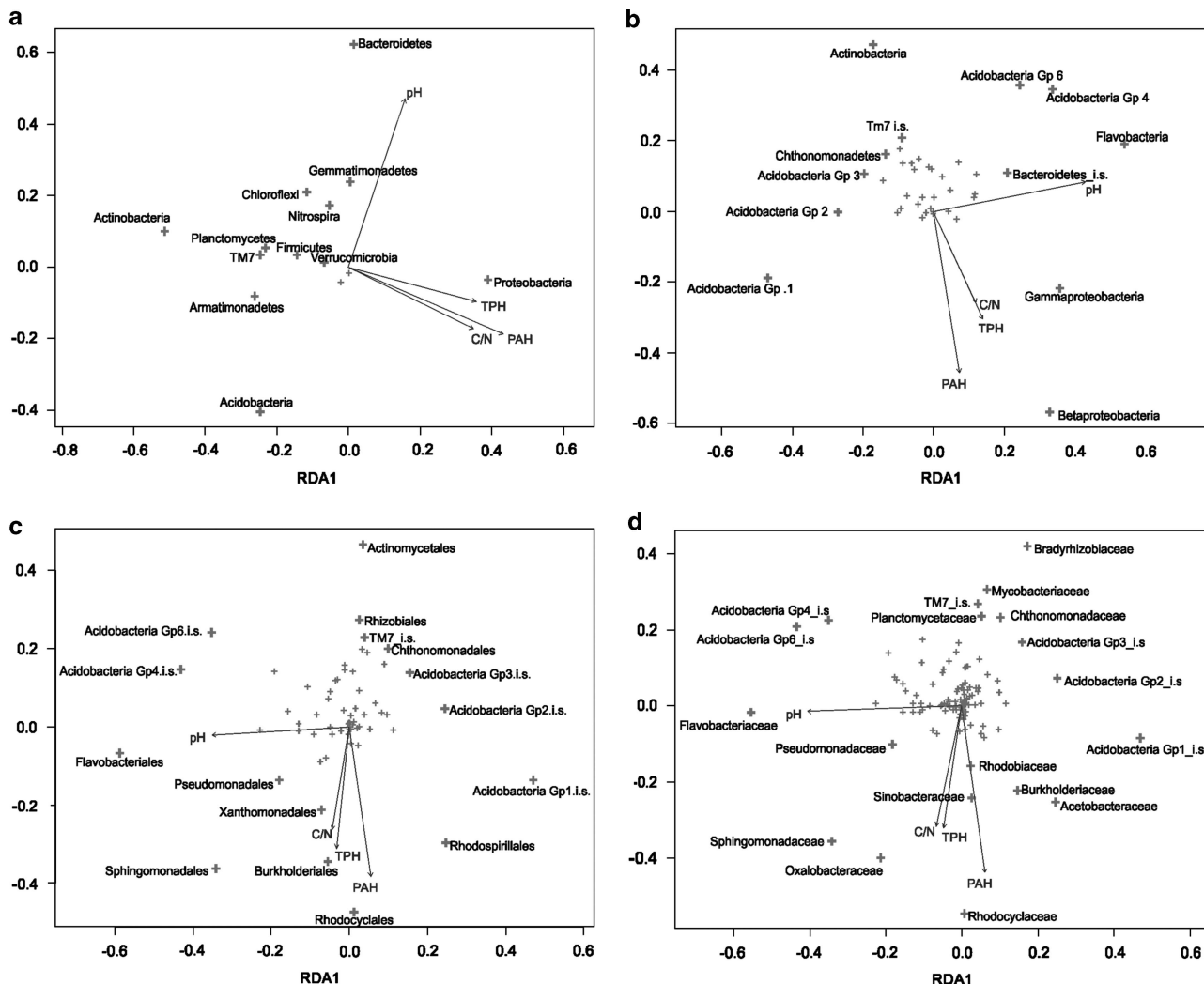
**Figure 4** Redundancy analysis (RDA) at (**a**) phylum, (**b**) class, (**c**) order and (**d**) family level. Taxa are represented by red crosses and only those taxa that showed strong responses to environmental gradients have been labelled. A full colour version of this figure is available at *The ISME Journal* online.

could be explained by high pollution. Although the spatial patterns of Betaproteobacteria with small-scale heterogeneity and patchiness matched the patterns of pollutant distribution, Gammaproteobacteria displayed large-scale heterogeneity reflecting gradients (Supplementary Figure S2) that could be explained by positive correlation of Gammaproteobacteria with pH. The responses of Alphaproteobacteria could not be explained by the environmental gradients studied, and Deltaproteobacteria decreased in abundance in the hot spots of pollution. As noticed from the kriged maps (Supplementary Figure S3) and RDA (Figure 4c), Acidobacteria Gp1, Gp2 and Gp3 were more common in plots with low pH, whereas Acidobacteria Gp 4 and Gp 6 were favoured by higher pH. The ranges of spatial autocorrelation displayed by Acidobacteria Gp1 and Gp6 were higher as compared with other bacterial taxa and this could be because of their dependence on the pH gradient (Supplementary Table S2).

At order level, the alphaproteobacterial orders of Sphingomonadales and Rhodospirillales could be related to high pollution levels, whereas Rhizobiales followed the opposite trend. Betaproteobacterial order Rhodocyclales was also favoured in high PAH levels. Apart from the pollution level, the distribution of Sphingomonadales and Rhodospirillales was also explained by pH (Figure 4c). At the family level, a distinct response to pH could be noticed in the bacterial groups that at higher taxonomic level were favoured by pollution (Figure 4d). Sphingomonadaceae, Oxalobacteraceae and Pseudomonadaceae appeared to be common at high pH, whereas Burkholderiaceae and Acetobacteraceae were common at low pH.

*Microbial activity*
Microbial activity in soil was assessed by two different methods: basal respiration assay and FDA hydrolysis assay. Soil basal respiration is a measure

2138

of the total biological activity and provides information on the biodegradability potential of hydrocarbons in soils and has been used as a relevant parameter during bioremediation treatments (Margesin *et al.*, 2003; Tuomi *et al.*, 2004). The rate of FDA hydrolysis in soil describes a broad range of enzyme activities as FDA is hydrolyzed by a number of different enzymes such as proteases, lipases and esterases (Margesin *et al.*, 2003).

The basal respiration rate varied from 0.1 to 1.1 µl $CO_2$ per g soil per h and the FDA hydrolysis rate was in the range of 19.6–140.5 µg fluorescein per g soil per 3 h (Figure 5). The basal respiration rate was positively correlated with TPHs and total PAHs, whereas the FDA hydrolysis rate had a positive correlation with five different PAHs, indicating that the microbial activity was higher in the polluted plots (Supplementary Table S3). Total organic carbon, nitrogen, C/N, Cu and Pb concentration showed positive correlations with both measures of microbial activity. In addition, both measures of microbial activity showed a highly significant correlation with each other ($r = 0.75$, $P < 0.0001$).

## Discussion

Spatial analysis of bacterial communities at different taxonomic levels and the dependence of the communities on specific spatial patterns of soil chemical parameters enabled us to infer possible ecological roles of bacterial taxa in an aged contaminated site. This integrated approach revealed two major spatial patterns, patchy distribution of pollutants and a pH gradient, that were clearly related to the spatial distribution of bacterial groups. Together with the spatial patterns of relative abundance of bacteria, the spatial patterns of microbial activity provided a holistic view of the impact of anthropogenic disturbance on the subsurface soil microbiology of this site.

Proteobacteria, which consists of myriad of well-known degraders, came across as the only phylum that dominated in the zones with high pollution, whereas the relative abundance of Actinobacteria was dramatically reduced in these zones. The latter result contradicts previous findings, where PAH pollution has been observed to increase the relative abundance of Actinobaceria
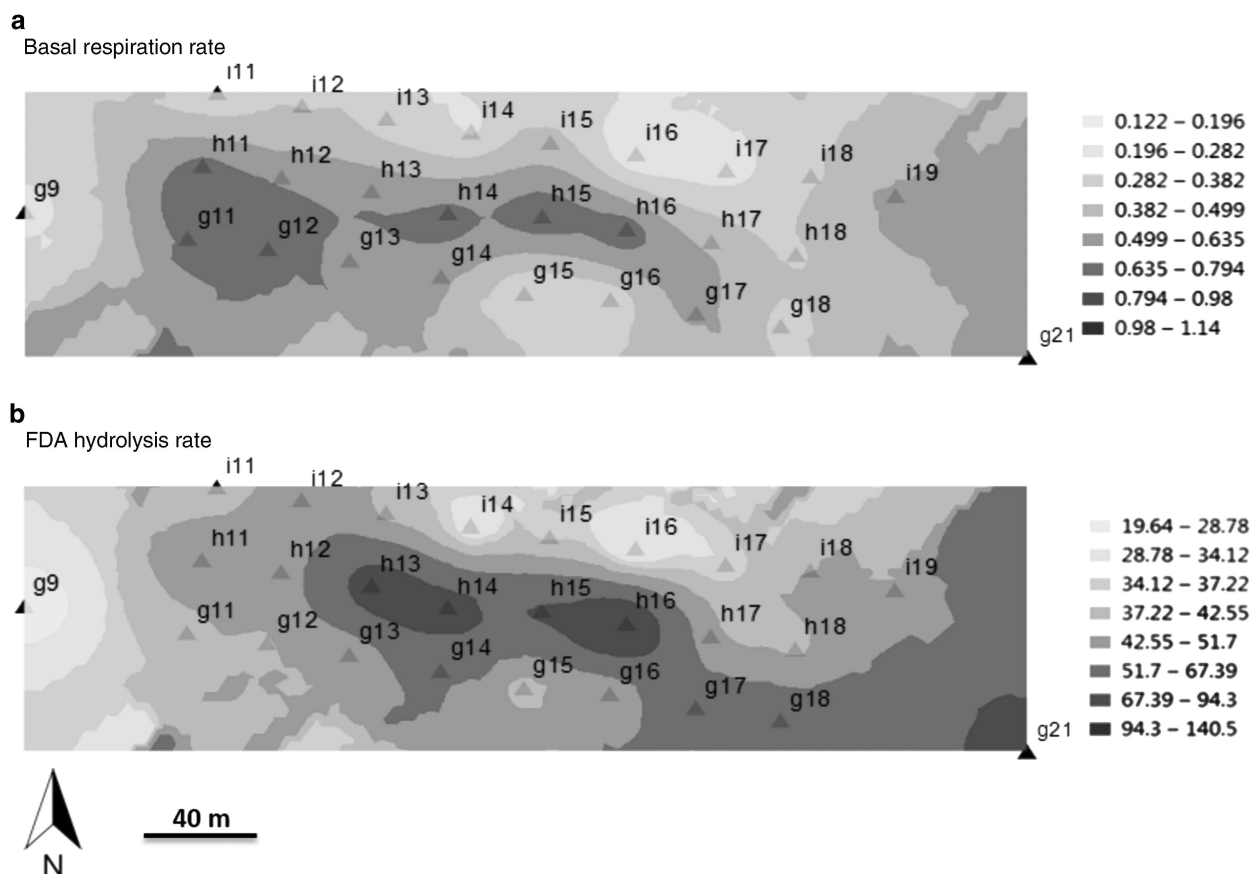


**Figure 5** Kriged maps of microbial activity on the polluted site. (**a**) Basal respiration rate given as $CO_2$ µl per g soil per h and (**b**) FDA hydrolysis rate (µg fluorescein per g soil per 3 h). The colour codes are given next to each map. A full colour version of this figure is available at *The ISME Journal* online.

(Johnsen et al., 2007; Antizar-Ladislao et al., 2008). Actinobacteria includes some of the most robust known PAH and TPH degraders (Kanaly and Harayama, 2000; Arenskötter et al., 2004; Larkin et al., 2005). A study on a similar creosote-contaminated site in Sweden, however, reported a decline in actinobacterial populations (assessed by phospholipid fatty acid analysis) from the hot spots of pollution (Törneman et al., 2008). Several investigations on hydrocarbon-contaminated cold sites have also demonstrated the dominance of Proteobacteria over Actinobacteria (Whyte et al., 2002; Margesin et al., 2003; Labbe et al., 2007; Ringelberg et al., 2008). Our findings suggest that among both groups of robust degraders, Actinobacteria were outcompeted by Proteobacteria in this aged creosote-polluted site. It is noteworthy that the proteobacterial responses to pollution varied at lower levels of taxonomic resolution. This can be expected as the physiological diversity within the phylum Proteobacteria is well known. Alphaproteobacteria, which also comprises orders displaying a diverse niche specialization (Ettema and Andersson, 2009), displayed patterns that could not be explained by any environmental parameter measured. Distinct response to contamination by Alphaproteobacteria could be noticed only upon further analysis at order and family level.

The presence of a pH gradient at the Luumäki site played an important role in shaping the bacterial communities. This could be anticipated as pH has been shown to be a master environmental variable in regulating the bacterial community structure in soils (Fierer and Jackson, 2006; Lauber et al., 2009). Furthermore, the simultaneous presence of patches of toxic creosotes and a gradient of pH in the nutrient-poor sandy soil revealed some interesting patterns of niche differentiation. Proteobacterial families that dominated in high pollution levels in this site also responded to the pH gradient. Burkholderiaceae and Acetobacteraceae prevailed in lower pH, whereas Pseudomonadaeceae and Sphingomonadaceae were shown to prefer higher pH. Several genera belonging to Acetobacteraceae have been linked with hydrocarbon degradation at moderately low pH levels (Hamamura et al., 2005; Röling et al., 2006). Sphingomonadaceae, Burkholderiaceae and Pseudomonadaceae are recognized for their extraordinary catabolic flexibility and are among the most common hydrocarbon degraders in a variety of soils (Leys et al., 2004; Stolz, 2009; Silby et al., 2011; Pérez-Pantoja et al., 2012); however, a thorough knowledge on pH optima of these groups is still lacking. Although these results provide clues about the pH preferences of these bacterial groups in the studied site, it remains to be shown whether they successfully degrade PAH or TPH compounds under these conditions. Acidobacteria have widely been accepted as oligotrophs, thriving in low pH and being among the most abundant groups of soil bacteria (Fierer et al., 2007).

The contrasting tendencies of pH preference displayed by different classes of Acidobacteria in this site were in agreement with some recent studies (Rousk et al., 2010; George et al., 2011; Foesel et al., 2014). Unlike most other phyla, Acidobacteria seemed to be unaffected by contamination levels that suggests a role as generalists in hydrocarbon-contaminated soils (Mukherjee et al., 2013).

The reduced bacterial diversity in the most polluted plots compared with the clean plots in this aged polluted site was a clear indication of the long-term negative impact of creosote pollution. The decrease in evenness of the bacterial community with pollution was explained by the clear dominance of Proteobacteria over all the other bacterial groups. In contrast to the bacterial diversity, the total microbial activity (basal respiration and FDA hydrolysis rate) was found to be increased in areas with higher pollution. In a study of the impact of hydrocarbons on soil enzyme activities, the presence of diesel oil resulted in significantly increased FDA hydrolysis compared with those in uncontaminated controls (Margesin et al., 2003). Together with decreased bacterial diversity and evenness in pollution, this could be an indication of adaptation or enrichment of specific indigenous soil microorganisms, most likely the members of Proteobacteria, to the contamination and their utilization of hydrocarbons as a major carbon source. It is also worth mentioning here that the microbial activity assays used in this study measure the activity of total soil biota including fungi that often play an important role in the biodegradation of PAHs (Cerniglia, 1993; Harms et al., 2011). Although some studies have shown the potential role of fungi in biodegradation of PAHs in creosote-contaminated soils (Atagana et al., 2006; Byss et al., 2008), Törneman et al. (2008) reported a decline in fungal population in a similar aged creosote-polluted site in Sweden. They attributed this decline to the toxic effects exerted by PAHs. It would therefore be interesting to investigate the spatial distribution of fungi in the Luumäki site to be able to better explain the relationships between creosote concentration and microbial activity. Furthermore, bioavailability of PAHs is known to be inversely related to the soil PAH contact time, that is, ageing (Alexander, 2000; Ling et al., 2010), that might also be an important factor effecting the microbial metabolism of PAHs in the studied site. Activity assays specifically indicating the biodegradation potential of the indigenous microbial communities in the studied site could probably provide a deeper insight into the adaptation and responses of soil biota to aged creosote pollution.

Upscaling of bioremediation continues to be challenging both from the perspective of suitable sampling designs for site characterization and for the suitable scale of remedial treatments. Spatial heterogeneity on polluted sites is a major hurdle in this upscaling. Comprehensive information

of spatial distribution of indigenous microbial populations could play a crucial role in overcoming this hurdle and in optimizing and implementing an appropriate bioremediation regime. Although geostatistical approaches are often employed in characterizing the extent and distribution of pollutants on-site, such an approach is very rarely used for the characterization of microbial populations. Our study brings forward the importance of integrating the spatial analysis of microbial and soil chemical parameters for an extensive characterization of polluted site. A combined effect of patchiness of contaminants and a pH gradient explained the patterns of niche differentiation displayed by the bacterial communities in this site. In addition, the spatial patterns exhibited by various microbial taxa in the polluted and clean soils could be used as proxies during further monitoring of bioremediation. This approach is a step forward towards opening the microbial 'black box' and incorporating the information in ecosystem engineering.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

## Disclaimer

We confirm that the material presented in this manuscript is original research, has not been published previously and has not been submitted for publication elsewhere.

## References

Ager D, Evans S, Li H, Lilley AK, Van Der Gast CJ. (2010). Anthropogenic disturbance affects the structure of bacterial communities. *Environ Microbiol* **12**: 670–678.

Anderson MJ. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**: 32–46.

Alexander M. (2000). Aging, bioavailability, and overestimation of risk from environmental pollutants. *Environ Sci Technol* **34**: 4259–4265.

Antizar-Ladislao B, Spanova K, Beck AJ, Russell NJ. (2008). Microbial community structure changes during bioremediation of PAHs in an aged coal-tar contaminated soil by in-vessel composting. *Int Biodeterior Biodegrad* **61**: 357–364.

Arenskötter M, Bröker D, Steinbüchel A. (2004). Biology of the metabolically diverse genus Gordonia. *Appl Environ Microbiol* **70**: 3195–3204.

Atagana HI, Haynes R, Wallis F. (2006). Fungal bioremediation of creosote-contaminated soil: a laboratory scale bioremediation study using indigenous soil fungi. *Water Air Soil Pollut* **172**: 201–219.

Bengtsson G, Törneman N, Yang X. (2010). Spatial uncoupling of biodegradation, soil respiration, and PAH concentration in a creosote contaminated soil. *Environ Pol* **158**: 2865–2871.

Berry D, Ben Mahfoudh K, Wagner M, Loy A. (2011). Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl Environ Microbiol* **77**: 7846–7849.

Bru D, Ramette A, Saby N, Dequiedt S, Ranjard L, Jolivet C et al. (2010). Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale. *ISME J* **5**: 532–542.

Byss M, Elhottová D, Tříska J, Baldrian P. (2008). Fungal bioremediation of the creosote-contaminated soil: influence of *Pleurotus ostreatus* and *Irpex lacteus* on polycyclic aromatic hydrocarbons removal and soil microbial community composition in the laboratory-scale study. *Chemosphere* **73**: 1518–1523.

Chapin FS III, Zavaleta ES, Eviner VT, Naylor RL, Vitousek PM, Reynolds HL et al. (2000). Consequences of changing biodiversity. *Nature* **405**: 234–242.

Cerniglia CE. (1993). Biodegradation of polycyclic aromatic hydrocarbons. *Curr Opin Biotechnol* **4**: 331–338.

Davis MRH, Zhao F, McGrath SP. (2004). Pollution-induced community tolerance of soil microbes in response to a zinc gradient. *Environ Toxicol Chem* **23**: 2665–2672.

Enwall K, Throbäck IN, Stenberg M, Söderström M, Hallin S. (2010). Soil resources influence spatial patterns of denitrifying communities at scales compatible with land management. *Appl Environ Microbiol* **76**: 2243–2250.

Ettema TJ, Andersson SG. (2009). The α-proteobacteria: the Darwin finches of the bacterial world. *Biol Lett* **5**: 429–432.

Fahy A, Lethbridge G, Earle R, Ball AS, Timmis KN, McGenity TJ. (2005). Effects of long-term benzene pollution on bacterial diversity and community structure in groundwater. *Environ Microbiol* **7**: 1192–1199.

Falkowski PG, Fenchel T, Delong EF. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–1039.

Fierer N, Bradford MA, Jackson RB. (2007). Toward an ecological classification of soil bacteria. *Ecology* **88**: 1354–1364.

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

Foesel BU, Nägele V, Naether A, Wüst PK, Weinert J, Bonkowski M et al. (2014). Determinants of Acidobacteria activity inferred from the relative abundances of 16S rRNA transcripts in German grassland and forest soils. *Environ Microbiol* **16**: 658–675.

George IF, Hartmann M, Liles MR, Agathos SN. (2011). Recovery of as-yet-uncultured soil Acidobacteria on dilute solid media. *Appl Environ Microbiol* **77**: 8184–8188.

Gerhardt KE, Huang XD, Glick BR, Greenberg BM. (2009). Phytoremediation and rhizoremediation of organic soil contaminants: potential and challenges. *Plant Sci* **176**: 20–30.

Goovaerts P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma* **89**: 1–45.

Goovaerts P. (1998). Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. *Biol Fertility Soils* **27**: 315–334.

Green V, Stott D, Diack M. (2006). Assay for fluorescein diacetate hydrolytic activity: optimization for soil samples. *Soil Biol Biochem* **38**: 693–701.

Hamamura N, Olson SH, Ward DM, Inskeep WP. (2005). Diversity and functional analysis of bacterial communities associated with natural hydrocarbon seeps in acidic soils at Rainbow Springs, Yellowstone National Park. *Appl Environ Microbiol* **71**: 5943–5950.

Hammer Ø, Harper D, Ryan P. (2009). PAST: paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol Electronica* **4**: 9.

Harms H, Schlosser D, Wick LY. (2011). Untapped potential: exploiting fungi in bioremediation of hazardous chemicals. *Nat Rev Microbiol* **9**: 177–192.

Johnsen AR, Schmidt S, Hybholt TK, Henriksen S, Jacobsen CS, Andersen O. (2007). Strong impact on the polycyclic aromatic hydrocarbon (PAH)-degrading community of a PAH-polluted soil but marginal effect on PAH degradation when priming with bioremediated soil dominated by mycobacteria. *Appl Environ Microbiol* **73**: 1474–1480.

Kanaly RA, Harayama S. (2000). Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. *J Bacteriol* **182**: 2059–2067.

Kulik N, Goi A, Trapido M, Tuhkanen T. (2006). Degradation of polycyclic aromatic hydrocarbons by combined chemical pre-oxidation and bioremediation in creosote contaminated soil. *J Environ Manage* **78**: 382–391.

Labbe D, Margesin R, Schinner F, Whyte LG, Greer CW. (2007). Comparative phylogenetic analysis of microbial communities in pristine and hydrocarbon-contaminated Alpine soils. *FEMS Microbiol Ecol* **59**: 466–475.

Larkin MJ, Kulakov LA, Allen CC. (2005). Biodegradation and Rhodococcus–masters of catabolic versatility. *Curr Opin Biotechnol* **16**: 282–290.

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

Legendre P, Gallagher ED. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**: 271–280.

Leys NM, Ryngaert A, Bastiaens L, Verstraete W, Top EM, Springael D. (2004). Occurrence and phylogenetic diversity of Sphingomonas strains in soils contaminated with polycyclic aromatic hydrocarbons. *Appl Environ Microbiol* **70**: 1944–1955.

Lindstrom JE, Barry RP, Braddock JF. (1999). Long-term effects on microbial communities after a subarctic oil spill. *Soil Biol Biochem* **31**: 1677–1689.

Ling W, Zeng Y, Gao Y, Dang H, Zhu X. (2010). Availability of polycyclic aromatic hydrocarbons in aging soils. *J Soils Sediments* **10**: 799–807.

Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime J, Hector A *et al.* (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294**: 804–808.

Margesin R, Labbe D, Schinner F, Greer C, Whyte L. (2003). Characterization of hydrocarbon-degrading microbial populations in contaminated and pristine alpine soils. *Appl Environ Microbiol* **69**: 3085–3092.

Mueller JG, Chapman PJ, Pritchard PH. (1989). Creosote-contaminated sites. Their potential for bioremediation. *Environ Sci Technol* **23**: 1197–1201.

Mukherjee S, Heinonen M, Dequvire M, Sipilä T, Pulkkinen P, Yrjälä K. (2013). Secondary succession of bacterial communities and co-occurrence of phylotypes in oil-polluted Populus rhizosphere. *Soil Biol Biochem* **58**: 188–197.

Oorts K, Bronckaers H, Smolders E. (2006). Discrepancy of the microbial response to elevated copper between freshly spiked and long-term contaminated soils. *Environ Toxicol Chem* **25**: 845–853.

Pérez-Pantoja D, Donoso R, Agulló L, Córdova M, Seeger M, Pieper DH *et al.* (2012). Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. *Environ Microbiol* **14**: 1091–1117.

Philippot L, Andersson SG, Battin TJ, Prosser JI, Schimel JP, Whitman WB *et al.* (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Microbiol* **8**: 523–529.

Philippot L, Bru D, Saby N, Čuhel J, Arrouays D, Šimek M *et al.* (2009). Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree. *Environ Microbiol* **11**: 3096–3104.

Ramsey PW, Rillig MC, Feris KP, Gordon NS, Moore JN, Holben WE *et al.* (2005). Relationship between communities and processes; new insights from a field study of a contaminated ecosystem. *Ecol Lett* **8**: 1201–1210.

Ringelberg D, Richmond M, Foley K, Reynolds C. (2008). Utility of lipid biomarkers in support of bioremediation efforts at army sites. *J Microbiol Methods* **74**: 17–25.

Röling WFM, Ortega-Lucach S, Larter SR, Head IM. (2006). Acidophilic microbial communities associated with a natural, biodegraded hydrocarbon seepage. *J Appl Microbiol* **101**: 290–299.

Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG *et al.* (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* **4**: 1340–1351.

Schloss PD, Gevers D, Westcott SL. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* **6**: e27310.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Silby MW, Winstanley C, Godfrey SAC, Levy SB, Jackson RW. (2011). Pseudomonas genomes: diverse and adaptable. *FEMS Microbiol Rev* **35**: 652–680.

Stolz A. (2009). Molecular characteristics of xenobiotic-degrading sphingomonads. *Appl Microbiol Biotechnol* **81**: 793–811.

Sturman PJ, Stewart P, Cunningham A, Bouwer E, Wolfram J. (1995). Engineering scale-up of in situ bioremediation processes: a review. *J Contam Hydrol* **19**: 171–203.

Törneman N, Yang X, Bååth E, Bengtsson G. (2008). Spatial covariation of microbial community composition and

polycyclic aromatic hydrocarbon concentration in a creosote-polluted soil. *Environ Toxicol Chem* **27**: 1039–1046.

Tuomi PM, Salminen JM, Jørgensen KS. (2004). The abundance of nahAc genes correlates with the 14C-naphthalene mineralization potential in petroleum hydrocarbon-contaminated oxic soil layers. *FEMS Microbiol Ecol* **51**: 99–107.

Van Der Heijden MGA, Bardgett RD, Van Straalen NM. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* **11**: 296–310.

Van Oss L. (2012). Gebiedsplan van voormalig productiebos Somerharju. BSc thesis. Hogeschool InHolland, 97 pp.

Wessén E, Söderström M, Stenberg M, Bru D, Hellman M, Welsh A *et al.* (2011). Spatial distribution of ammonia-oxidizing bacteria and archaea across a 44-hectare farm related to ecosystem functioning. *ISME J* **5**: 1213–1225.

Whyte L, Schultz A, Beilen J, Luz A, Pellizari V, Labbé D *et al.* (2002). Prevalence of alkane monooxygenase genes in Arctic and Antarctic hydrocarbon-contaminated and pristine soils. *FEMS Microbiol Ecol* **41**: 141–150.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)