

Lijuan Yan, Petri Penttinen#, Anu Mikkonen, Kristina Lindström

Bacterial community changes in response to oil contamination and perennial crop cultivation

Affiliations:

Lijuan Yan, Petri Penttinen, Kristina Lindström

Department of Environmental Sciences, PO Box 65, FI-00014 University of Helsinki, Finland

Anu Mikkonen

Department of Biological and Environmental Science, PO Box 35, FI-40014 University of Jyväskylä, Finland

Corresponding author: Petri Penttinen, Department of Environmental Sciences, PO Box 65, FI-00014 University of Helsinki, Finland; +358 41 5243194; petri.penttinen@helsinki.fi

Present addresses:

Lijuan Yan

Department of Environmental Sciences, Niemenkatu 73, FI-15140 University of Helsinki, Finland

Petri Penttinen

Zhejiang Provincial Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon

Sequestration, School of Environmental and Resource Sciences, Zhejiang Agriculture & Forestry

University, Huancheng North Road 88, Lin'an 311300, China

Abstract

We investigated bacterial community dynamics in response to used motor oil contamination and perennial crop cultivation by 16S rRNA gene amplicon sequencing in a four-year field study.

Actinobacteria, *Proteobacteria*, *Chloroflexi*, *Acidobacteria* and *Gemmatimonadetes* were the major bacterial phyla, and *Rhodococcus* the most abundant genus. Initially, oil contamination decreased the overall bacterial diversity. *Actinobacteria*, *Betaproteobacteria* and *Gammaproteobacteria* were

sensitive to oil contamination, exhibiting clear succession with time. However, bacterial communities changed over time, regardless of oil contamination and crop cultivation. The abundance difference of most OTUs between oil-contaminated and non-contaminated plots remained the same in later sampling years after the initial abundance difference induced by oil spike. The abundances of three oil-favored actinobacteria (*Lysinimonas*, *Microbacteriaceae* and *Marmoricola*) and one betaproteobacterium (*Aquabacterium*) changed in different manner over time in oil-contaminated and non-contaminated soil. We propose that these taxa are potential bio-indicators for monitoring recovery from motor oil contamination in boreal soil. The effect of crop cultivation on bacterial communities became significant only after the crops achieved stable growth, likely associated with plant material decomposition by *Bacteroidetes*, *Armatimonadetes* and *Fibrobacteres*.

Keywords: soil microbiome; hydrocarbon contamination; bioremediation; perennial crop cultivation; 16S rRNA gene amplicon sequencing; temporal abundance change

Acknowledgements

We thank Lars Paulin and his colleagues at the Institute of Biotechnology, Helsinki, for 16S rRNA gene sequencing. Asko Simojoki and Frederick L. Stoddard are acknowledged for their supervision in soil and crop analysis. We thank Markku Tykkyläinen for his technical assistance in the field management and plant harvest. The authors also acknowledge Anni-Mari Pulkkinen and Kati P. Pulkkinen for their technical assistance with oil analysis. This work was supported by Legume Futures, an international research project funded by the European Union through the Framework 7 Programme (FP7) under grant agreement number 245216 (FP7-KBBE-2009-3); Magnus Ehrnrooth foundation; The Finnish Society of Sciences and Letters; and MUTKURY for oil analysis.

Introduction

Soil pollution by petroleum hydrocarbons (PHCs) originating from crude oil or refined petroleum products poses a significant threat to the environment. In particular, used motor oil that contains high concentrations of aliphatics, polycyclic aromatic hydrocarbons (PAHs) and heavy metals (e.g. lead, zinc, chromium, barium and arsenic) contributes to chronic hazards including carcinogenicity (Dominguez-Rosado et al. 2004, Vazquez-Duhalt 1989). Microorganisms can degrade PHCs and utilize them as carbon and energy sources in a natural attenuation process. Nitrogen is frequently a limiting factor in oil-contaminated soils (Wenzel 2009). Thus, the nitrogen-fixing legume-rhizobium symbiosis has potential to assist in the biodegradation of PHCs (Dominguez-Rosado and Pichtel 2004). The cultivation of oil-tolerant perennial legumes holds promise for accelerated oil degradation with additional bonus of biomass for bioenergy production. Fodder galega (*Galega orientalis* Lam.), a perennial forage legume, and smooth brome (*Bromus inermis* L.), a cool-season perennial sod-forming grass, are both persistent in boreal and nemoral zones and grow well in crop mixtures without N-fertilizer supply (Jasinskas et al. 2008, Kryževičienė et al. 2008). Fodder galega-*Neorhizobium galegae* symbiosis is both oil tolerant and possesses bioremediation potential in oil-contaminated soils (Jussila et al. 2006, Lindstrom et al. 2003, Mikkonen et al. 2011a, Suominen et al. 2000).

Degradation of PHCs is relatively slow in the boreal zone. To assess the sustainability of a perennial legume-cropping bioremediation system economically and environmentally, we established a four-year bioremediation field experiment with monocrops (*Bromus inermis*, a grass and *Galega orientalis*, a legume), their intercrops and bare fallow and used motor oil+/- treatments. Our previous studies (Yan et al. 2015, Yan et al. 2016) showed that depending on crop type 8%-27% of oil remained in the soil after 40 months. Furthermore, oil increased crop dry matter and nitrogen yield, indicating that the perennial legume-cropping bioremediation system can provide economic benefits as the biomass may be used for example to produce bioenergy (Yan et al. 2015).

The aims of this study were to monitor changes in bacterial communities in response to oil and crop treatments, to identify the bacterial taxa that were responsible for the changes in bacterial communities between oil-contaminated and control soil over time, and to assess the overall bioremediation system by linking the changes of bacterial community with environmental variables.

Materials and methods

Experimental design and sampling of soil and crop data

The field was established at the Viikki Experimental Farm, University of Helsinki, Finland (60°14'N, 25°01'E, 8 m AMSL). Monocultures of brome grass and fodder galega, their mixture and bare fallow were the main plots in four replicated blocks of 2.5×1.5 m per plot. Used motor oil treatments (oil+/-) were the sub-plot factors. Soil samples consisting of 16 subsamples were taken from the top 20 cm layer in each plot in July 2009, May 2010, May 2011 and May 2012 and stored at -20°C until the analysis. The oil concentration in each oil-spiked plot was determined as the difference of total solvent extractable material concentration between the plot and the average of 4 to 5 randomly selected control plots at each sampling time. Field establishment (e.g. oil spike, seed sowing and field management), sampling and measurements of oil and crop parameters are described in details in Supporting Information.

Amplicon library construction and sequence analysis

Genomic DNA was extracted from 0.5 g of homogenized soil samples using FastDNA SPIN kit for Soil (Qbiogene, USA) according to the manufacturer's instructions. The amplification of bacterial 16S rRNA gene (target region V1-V3) of soil DNA (1:50 diluted) was performed in two PCR steps prior to sequencing (Table S1). The amplicons were sequenced at the Institute of Biotechnology, University of Helsinki (Finland) using Illumina MiSeq (San Diego, CA, USA). The resulting forward and reverse reads were combined and analyzed using MOTHUR v. 1.35.1 (Schloss et al. 2009), according to the standard operating procedure for MiSeq data (Kozich et al. 2013),

accessed on April 28, 2015. Sequences were clustered into operational taxonomic units (OTUs) at 97% similarity and classified with Bayesian classifier based on SILVA reference database v119 at cutoff of 0.8. The non-rarefied OTU data was exported. Singleton OTUs were removed and a subsample file was generated, so that all samples were compared at an equivalent sequencing depth of 2977. The sequences of the partial 16S rRNA gene amplicons have been deposited in the NCBI Short Read Archive under accession no. SRP072116.

Statistical analysis

Alpha diversity

Alpha diversity (Shannon-Wiener index) was calculated based on the OTU data of the whole bacterial community and the major classes (relative abundance of each class > 2%) in PRIMER V.7 (Clarke and Warwick 2001) and exported to SPSS version 22 (IBM Inc., Armonk, NY, USA) for statistical analysis. The diversity values checked with Normal Q-Q plots were roughly normally distributed. We used repeated measures analysis of variance (RM ANOVA) with the sampling time as the repeated factor (within-subject factor) to test the overall between- and within-subjects effects (sphericity assumed) on the diversity indices, based on a split-plot experimental design. When the effects of interactions between sampling times and other treatment factors were significant, the split-plot univariate analysis of variance (UV ANOVA) was applied to further test the between-subjects effects (oil, crop and oil × crop) on bacterial diversity yearly. Bonferroni multiple pairwise test was applied to compare the means. Differences were concluded significant at $p < 0.05$.

Multivariate data analysis of bacterial community data

The bacterial OTU data were non-normally distributed and contained high numbers of rare species that contributed lots of zeroes to the data set. Therefore, non-parametric multivariate methods based on Bray-Curtis dissimilarity were used to test the significance of the treatment effects on bacterial communities in PRIMER v.7 with the add-on package PERMANOVA+

(Anderson et al. 2008). Square root transformation of the rarefied OTU data was applied to ensure that abundant species did not dominate the following non-parametric statistical analysis. Bacterial OTUs were assigned into taxa based on the respective taxonomic affiliations from phylum to genus. Variation in bacterial communities over a range of taxonomic levels was firstly visualized by non-metric multidimensional scale ordination (nMDS). Repeated measures PERMANOVA was used to test the between- and within- subjects effects on transformed OTU data based on the split-plot experimental design (Anderson et al. 2008, Anderson 2001). When the interactions between time and other factors were significant, PERMANOVA was performed to test the treatment effects yearly. When the effect of each treatment factor was significant, pair-wise comparisons were obtained by doing an additional separate run of the PERMANOVA routine. The homogeneity of multivariate dispersions was tested using PERMDISP. PERMDISP was also performed to assure that the correlation structure among samples through time was ignored under permutation in multivariate data analysis. Canonical analysis of principal coordinates (CAP) (Anderson and Willis 2003) was used to characterize significant differences between *a priori* groups of bacterial overall community and at the phylum and class levels in response to oil contamination. OTUs that discriminated treatment groups were tested by SIMPER (similarity percentages) at different taxonomic levels, using a two-way crossed design with block and oil as the two factors. In this way, block effect was taken into account when determining the taxonomic groups that contributed to discriminating between oil-contaminated and non-contaminated plots. OTUs contributing >0.02% to the summed abundance over the four years were selected to visualize taxonomic contamination association networks using Cytoscape 3.2.1 (Shannon et al. 2003). Beta binomial analysis was performed to test the abundance difference of each OTU (untransformed data) between oil-contaminated and non-contaminated plots at each sampling time with the *ibb* package Pham et al. 2010) in R version 3.1.3 (<https://www.r-project.org/>). The OTUs with consistent contributions (Diss/SD ratio > 1 in SIMPER) and $p < 0.05$ in the beta binomial test were considered as

differentially abundant. The OTUs with statistically significant difference in abundance in at least one time point were selected for differential abundance analysis to test changes in abundances over time with DESeq2 (Love et al. 2014) using non-rarefied OTU data in R, as described at <http://www.bioconductor.org/help/workflows/rnaseqGene/#time>. A Draftsman plot was used to check the inter-correlations and distribution of soil parameters (oil concentration, pH, electrical conductivity EC, soil total C, soil total N and C:N ratio) and plant parameters (dry matter yield, crop total C, crop total N, crop C:N ratio, chlorophyll, %Ndfa and biologically fixed nitrogen yield), obtained from July 2009 and May 2012 sampling times. These environmental variables of different ranges were normalized. Marginal distance-based linear models (DISTLM) (McArdle and Anderson 2001) were first used to test the relationship between bacterial communities and each soil chemical parameter alone. Non-significant and multi-collinear environmental variables were removed based on sequential DISTLM tests, using R^2 criterion and plotted on a distance-based redundancy analysis ordination (db-RDA) for visualization. All nonparametric analyses were performed with 9999 permutations to determine the significance of the treatment effects. In all statistical analysis, differences were concluded significant at $p < 0.05$.

Results

Bacterial taxonomic composition and diversity

When analyzing bacterial communities in a four year field bioremediation experiment, altogether 20116 bacterial operational taxonomic units (OTUs) were detected. After subsampling to equal sequencing depth further analyses were done using 19053 OTUs. The OTUs were assigned to 32, 90, 198, 867 and 1815 taxa at phylum, class, order, family and genus levels, respectively. *Actinobacteria* (average relative abundance: 38.4%), *Proteobacteria* (29.3%), *Chloroflexi* (8.6%), *Acidobacteria* (8.3%) and *Gemmatimonadetes* (5.8%) were the most abundant phyla in the field soil (Supporting Information Fig. S1). Among the phylum *Actinobacteria*, *Actinobacteria* class accounted for 53.5%, *Thermoleophilia* for 32.1%, *Acidimicrobiia* for 5.4%, *MB-A2-108* for 3.4%

and unclassified for 5.6% of the total relative abundance. Among the phylum *Proteobacteria*, *Alphaproteobacteria* accounted for 63.3%, *Betaproteobacteria* for 18.1%, *Deltaproteobacteria* for 10.1% and *Gammaproteobacteria* for 8.0%. The proportions of unclassified sequences were 1.4% at the phylum level, 4.6% at the class, 14.4% at the order, 33.3% at the family and 57.8% at the genus level.

The alpha diversity of the overall community and at the class level exhibited a strong time-dependent pattern (RM ANOVA, $p < 0.05$). In the beginning of the experiment at 2009, when the oil concentration was on average 3.98 g kg^{-1} in bare fallow, 4.20 g kg^{-1} in brome grass, 4.59 g kg^{-1} in galega and 3.36 g kg^{-1} in mixture plots (Fig. S2), oil contamination decreased bacterial overall diversity (Fig. 1a). However, the effect of oil on bacterial diversity was not detected in the following years, even though the concentration of oil was approximately the same both in 2009 and 2010 (Fig. S2). The classes *Actinobacteria* and *Acidobacteria* were on average more diverse in the absence of oil, whereas *Gammaproteobacteria* were more diverse in the presence of oil (Table S2). Other classes showed divergent patterns in different years in response to oil contamination (Table S2). In 2012 bacterial diversity was significantly higher in the crop treatments than in the bare fallow (Fig. 1b).

Response of bacteria to treatments over time

Community level response

The bacterial communities changed over time, regardless of oil contamination (Fig. S3, Table S3), and the differences over time were statistically more significant than the differences between oil-contaminated and non-contaminated plots (Table 1). According to the canonical analysis of principle coordinates (CAP), bacterial communities were different based on time, oil, the combined factor of oil and time (OilTime) and crop, whereas the effect of crops was not revealed by PERMANOVA (Fig. 2, Table S3). CAP discriminated the bacterial communities between bare

fallow and the vegetated plots only in 2012 (Fig. S4). The relative abundances of the phyla *Bacteroidetes*, *Armatimonadetes*, and *Fibrobacteres* were positively associated with crop cultivation (Fig. S4).

PERMDISP revealed unequal within-group dispersions of bacterial communities based on time, oil and the combined factors OilTime and block and time (Fig. S5). The average dissimilarity of bacterial community between oil-contaminated and non-contaminated plots were 8.0% at the phylum level, 11.7% at the class, 15.3% at the order, 24.9% at the family and 32.6% at the genus level over years.

The classes *Actinobacteria*, *Betaproteobacteria* and *Gammaproteobacteria* showed clear *a priori* group separation (total correct allocation rates > 60%) in response to time and oil contamination (Fig. 3). The separation between oil-contaminated and non-contaminated communities was particularly clear in 2009 (correct allocation rates > 80%).

At the phylum level, *Actinobacteria* contributed to an average proportion of 9.3%, *Acidobacteria* of 9.2%, *Proteobacteria* of 7.9%, *Bacteroidetes* of 7.7%, *Gemmatimonadetes* of 6.2% and *Chloroflexi* of 5.6% to the dissimilarity between oil-contaminated and non-contaminated plots over years (Fig. S1c). In the taxonomic contamination association networks (Fig. S6), taxa that were more abundant in oil-contaminated than in non-contaminated plots were assigned mostly to the phyla *Actinobacteria* and *Proteobacteria*. Changes in the other phyla were smaller (Fig. S7).

OTU level response

The abundances of 283 OTUs were different between oil-contaminated and non-contaminated soils in at least one time point (Fig. S6). Among *Actinobacteria*, for example OTUs of the genera *Nocardioides*, *Nocardia*, *Rhodococcus*, *Williamsia* and *Streptomyces* were more abundant in oil-contaminated than in non-contaminated soil in 2009. In addition to these OTUs, OTUs affiliated to the genera *Phycoccus*, *Mycobacterium* and *Candidatus Microthrix* and the family *Microbacteriaceae* were more abundant in oil-contaminated soil in 2010 and 2011. In 2012,

members of the genera *Lysinimonas* and *TMI46* became more abundant in oil-contaminated soil. Particularly, *Rhodococcus* OTU8 was the most abundant OTU in oil-contaminated soil one month after oil spike; over time its relative abundance in oil-contaminated soil decreased sharply. The relative abundances of the members of classes *Thermoleophilia* and *MB-A2-108* were consistently lower in the presence of oil.

Among *Proteobacteria*, for example OTUs assigned to the genera *Phenylobacterium*, *Aquabacterium*, *Arenimonas* and *Lysobacter* and the family *KCM-B-112* were more abundant in oil-contaminated than in non-contaminated soil in 2009 (Fig. S6). In 2010, a higher number of taxa assigned to the classes *Alphaproteobacteria* (e.g. genera *Roseomonas* and members of the family *Bradyrhizobiaceae*) and *Gammaproteobacteria* (e.g. family *Alteromonadaceae* and the order *PYR10d3*) were detected in oil-contaminated than in non-contaminated soil. Those OTUs remained more abundant in oil-contaminated soil until the late phase of bioremediation, except for an inconsistency in 2011. Oil addition consistently decreased the relative abundance of the genus *Pseudolabrys* and the order *SC-I-84*. The other major phyla showed mostly higher or equal average relative abundances in non-contaminated soil than in oil-contaminated soil.

To compare relative abundances of the OTUs in oil-contaminated and non-contaminated soil over time, we applied a differential abundance analysis. The temporal abundance changes of 52 OTUs were different between the treatments ($p < 0.05$). After Bonferroni correction, only three *Actinobacteria* OTUs (*Lysinimonas* OTU132, *Microbacteriaceae* OTU29 and *Marmoricola* OTU541) and one *Betaproteobacteria* OTU (*Aquabacterium* OTU188) showed significantly different temporal abundance changes between treatments ($p < 0.05$) (Fig. 4). The relative abundance of these OTUs remained rather stable and low in the absence of oil over time. In the oil-contaminated plots, *Marmoricola* OTU541 and *Aquabacterium* OTU188 were more abundant immediately following oil spike in 2009; the relative abundance of *Microbacteriaceae* OTU29 was highest in 2010 and *Lysinimonas* OTU132 showed an increasing abundance in the

presence of oil over time. For the other OTUs, for example *Pseudolabrys* OTU2, the temporal abundance changes were similar in both treatments after the initial abundance difference detected in the beginning of the experiment.

Linking bacterial community to environmental variables

According to DISTLM sequential tests, oil concentration explained 3.07% of the total variation in bacterial communities, while pH explained 2.49%, electrical conductivity 1.79% and the crop dry matter yield 2.25% ($p < 0.05$) (Fig. S8). In addition, the other environmental variables, except for soil total N, explained lower but significant ($p < 0.05$) proportions of variation (Table S4).

Discussion

We studied the effects of oil and vegetation on bacterial communities in a four-year field experiment using amplicon sequencing targeting the bacterial 16S rRNA gene. The composition and diversity of the soil bacterial communities changed over time, regardless of oil contamination or vegetation. Initially, oil contamination decreased diversity, showing the immediate community-level disturbance by hydrocarbons. Multivariate analysis revealed distinct differences in the composition of bacterial communities, particularly *Actinobacteria*, *Betaproteobacteria* and *Gammaproteobacteria*, between oil-contaminated and non-contaminated soil over time, revealing significant ecological impact of oil contamination.

The insurance hypothesis states that biodiversity buffers against disturbances in ecosystem functioning; a high number of species guarantees that at least some species will function after disturbance (Yachi and Loreau 1999). Thus, diversity may be considered as a measure of sustainability. Short term bioremediation is known to decrease microbial diversity (Liang et al. 2011, Ros et al. 2014), whereas the decrease in hydrocarbon concentration has been found positively related to bacterial biodiversity (Dell'Anno et al. 2012). The sharp decrease in oil

concentration during the summer of 2010 (Yan et al. 2015) was plausibly associated with the observed high bacterial diversity and in particular with a higher number and relative abundance of possible oil-degrading groups such as *Rhodococcus*. In 2011, most major bacterial classes showed lowest diversity and the differences between oil-contaminated and non-contaminated plots were not significant. Low precipitation and low temperature in May 2011 (Yan et al. 2015), both of which affect bacterial growth, may have masked the impact of oil contamination. Similarly, the high precipitation in May 2012 (Yan et al. 2015) might have contributed to the detected high variations in bacterial communities. The effect of oil on bacterial alpha diversity was not detected after 2009, and in 2012 the diversity was higher in the vegetated plots, emphasizing the ecological sustainability of the perennial legume-cropping bioremediation system. As indicated by separation of *a priori* groups based on year and contamination, the community in oil-contaminated soil could not be expected to return to the original state, and thus monitoring the changes as a community without a clean soil control is not applicable to assess recovery.

In our study, the observed oil-favored OTUs belonged mostly to *Actinobacteria* and *Proteobacteria*, especially at the early-middle phase of bioremediation, as they were among the most abundant and diverse phyla in oil-contaminated soil. Among *Actinobacteria*, oil favored OTUs belonging to *Nocardiaceae* (including genera *Nocardia*, *Rhodococcus* and *Williamsia*), *Nocardioides* and *Streptomyces*, that are known to carry genes linked to the aerobic degradation of alkanes and a wide range of aromatic compounds, have been found in oil-contaminated environments (Akbari and Ghoshal 2015, Balachandran et al. 2012, Ferradji et al. 2014, Liao et al. 2015, Nie et al. 2014, Saito et al. 2000, Schippers et al. 2005, Song et al. 2011, Tsuboi et al. 2015). *Gammaproteobacteria* and *Betaproteobacteria* displayed clear successional changes during bioremediation and exhibited higher diversity in the early-middle stage of bioremediation. Among *Gammaproteobacteria*, *Arenimonas*, *Lysobacter*, *Alteromonadaceae* and *PYR10d3* were early-phase oil-favored taxa. *Arenimonas* and *Lysobacter* are closely related to *Pseudomonas*,

Xanthomonas and *Thermomonas* that include aerobic hydrocarbon degraders (Akbari and Ghoshal 2015, Young et al. 2007). *Alteromonadaceae* have PAH-, *n*-alkane- and toluene-degradation potential (Lamendella et al. 2014), and *PYR10d3* contains a group of uncultivated bacteria that were postulated capable of degrading pyrene or pyrene metabolites (Jones et al. 2008, Singleton et al. 2006). Among *Betaproteobacteria*, the order *Burkholderiales* were consistently more abundant in oil-contaminated than in non-contaminated soil. At least 60% of the *Burkholderiales* genomes comprise nearly all of the central ring-cleavage pathways involved in the degradation of aromatic compounds (Pérez-Pantoja et al. 2012). Thus, the competitiveness of these OTUs in oil-contaminated soil might be associated with hydrocarbon degradation by these bacteria. Similarly, the other middle-late-phase oil-responding actinobacterial taxa detected in our study included potential hydrocarbon degraders. For example, *Phycococcus* was previously noticed to carry the *nidA* gene and suggested to be involved in pyrene biodegradation (Chen et al. 2015), and *Candidatus Microthrix* was found specialized for alkane degradation (Yakimov et al. 2003).

In our experiment *Alphaproteobacteria* were more abundant in non-contaminated soil than in oil-contaminated soil, in accordance with the comparative phylogenetic study (Labbe et al. 2007), whereas other studies observed the dominance of *Alphaproteobacteria* in the presence of oil during biodegradation (Mills et al. 2003, Vinas et al. 2005). The varied soil condition, pollutant characteristics and original composition of *Alphaproteobacteria* communities among these studies are plausible explanations for the contradictory findings. In our study, *Phenylobacterium*, an alkane- and PAH-degrader found in cold environments (Giebler et al. 2013, Yang et al. 2014), was the only *Alphaproteobacteria* OTU that was consistently more abundant in the oil-contaminated soil than in non-contaminated soil. *Alphaproteobacteria* include strains with efficient PAH-degrading potential (Dunlevy et al. 2013, Giebler et al. 2013, Jones et al. 2011, Lafortune et al. 2009, Vinas et al. 2005, Yang et al. 2014, Zhang and Margesin 2014). The high oil degradation rate that occurred in summer 2010 (Yan et al. 2015) might be associated with the significant increase in

relative abundance of alphaproteobacterial *Sphingomonadaceae*, *Roseomonas* and *Bradyrhizobiaceae* in oil-contaminated soil in that summer.

The contamination association networks showed a large number of OTUs that displayed consistently higher or lower abundance in oil-contaminated plots than in the non-contaminated ones throughout the experiment. However, if the initial effect of oil spike on these OTUs was strong enough, it could have resulted in the observed consistently positive (or negative) responses of these OTUs till the end of the experiment. The differential abundance analysis was therefore performed to reveal the OTUs with time-dependent abundance differences between oil-contaminated and non-contaminated soil. Interestingly, the four OTUs affiliated to *Lysinimonas*, *Microbacteriaceae*, *Marmoricola* and *Aquabacterium*, with differential oil-specific abundance over time, were all favored by oil, revealing the high sensitivity of these bacterial taxa in response to the changing hydrocarbon composition at different phases of bioremediation. Out of the four OTUs, the relative abundance of *Aquabacterium* OTU188 was significantly different only in the first year. Interestingly, the *Aquabacterium* genus harbors *alkB* genes (Aburto-Medina et al. 2012, Giebler et al. 2013, Masuda et al. 2014), and a highly abundant *Aquabacterium* OTU positively correlated with hydrocarbon degradation rate in boreal soil contaminated by fuel oil (Mikkonen et al. 2011b). *Microbacteriaceae* OTU29 and OTU132 were more abundant in oil-contaminated soil only in the second and following years, possibly due to their increased importance in degradation of recalcitrant compounds with poor bioavailability in soil. Hence, sequences with high identity to these OTUs may be suitable indicators to monitor the ecological impact of oil contamination and the following recovery process. However, whether these indicators have general applicability must be confirmed in future studies.

Bioremediation includes complex interactions between biotic and abiotic factors. Assessment of the relationship between environmental factors and the response of the bacterial community is important to optimize remediation strategies for oil-contaminated soil. In assessing

the bioremediation we must keep in mind that not all changes are due to the contamination. The diversity and composition of bacterial community in our field soil could be interwoven with environmental factors that differ between years. In addition to oil concentration, EC, pH and dry matter yield explained low but significant proportions of variations in bacterial communities. Both EC and pH are known to shape soil bacterial communities (Johnson et al. 2003, Kuramae et al. 2010, Lauber et al. 2009). Our earlier report found them sensitive soil parameters that significantly decreased over years in our field, independent of oil addition (Yan et al. 2015).

Distinct differences in the composition and diversity of bacterial communities between vegetated plots and bare fallow were evident only after the perennial crops had established their stable growth in 2012, showing the significant impact of multi-year agricultural cultivation and soil management on soil bacterial communities. Changes in microbial community in the presence of vegetation were suggested attributable to an input of nutrients from the decomposition of vegetation cover (Hobbie 2015). We observed a strong positive correlation between *Bacteroidetes* and crop cultivation, agreeing with the previous studies (Acosta-Martinez et al. 2008, da C Jesus et al. 2009, Kuramae et al. 2010). Bacteria in the phyla *Armatimonadetes* and *Fibrobacteres* are capable of decomposing plant materials, polysaccharide-based substances, and photosynthetic biomass (Lee et al. 2014, Ransom-Jones et al. 2012), possibly explaining why these phyla were more abundant in vegetated plots than in the bare fallow. Since biological nitrogen fixation yield that correlated positively with crop dry matter yield was associated with the changes in bacterial community, the differences in bacterial community composition between legume and non-legume cropping systems would have most likely become significant during a longer period of crop cultivation. Since the unmeasured environmental factors most probably played important roles in affecting bacterial community dynamics as well as bioremediation efficiency, most of the variation in soil bacterial communities in the field remained unexplained.

In conclusion, bacterial communities changed over time, regardless of oil contamination or vegetation. After the initial abundance difference, the abundances of all but four OTUs changed in a similar manner in both oil-contaminated and non-contaminated plots. In the future whole metagenomic profiling can be incorporated to address the functional significance of the observed community differences.

References

- Aburto-Medina A, Adetutu EM, Aloor S, Weber J, Patil SS, Sheppard PJ, Ball AS, Juhasz AL (2012) Comparison of indigenous and exogenous microbial populations during slurry phase biodegradation of long-term hydrocarbon-contaminated soil. *Biodegradation* 23:813-822
- Acosta-Martinez V, Dowd S, Sun Y, Allen V (2008) Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biol Biochem* 40:2762-2770
- Akbari A, Ghoshal S (2015) Effects of diurnal temperature variation on microbial community and petroleum hydrocarbon biodegradation in contaminated soils from a sub-Arctic site. *Environ Microbiol*
- Anderson M, Gorley RN, Clarke RK (2008) *Permanova for Primer: Guide to Software and Statistical Methods*
- Anderson MJ, Willis TJ (2003) Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84:511-525
- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32-46
- Andreoni V, Bernasconi S, Colombo M, Van Beilen JB, Cavalca L (2000) Detection of genes for alkane and naphthalene catabolism in *Rhodococcus* sp. strain 1BN. *Environ Microbiol* 2:572-577
- Balachandran C, Duraipandiyar V, Balakrishna K, Ignacimuthu S (2012) Petroleum and polycyclic aromatic hydrocarbons (PAHs) degradation and naphthalene metabolism in *Streptomyces* sp.(ERI-CPDA-1) isolated from oil contaminated soil. *Bioresour Technol* 112:83-90
- Chaudhary P, Sharma R, Singh SB, Nain L (2011) Bioremediation of PAH by *Streptomyces* sp. *Bull Environ Contam Toxicol* 86:268-271
- Chen S, Peng J, Duan G (2015) Enrichment of functional microbes and genes during pyrene degradation in two different soils. *Journal of Soils and Sediments*:1-10
- Clarke K, Warwick R (2001) *An approach to statistical analysis and interpretation. Change in Marine Communities* 2

- da C Jesus E, Marsh TL, Tiedje JM, de S Moreira, Fatima M (2009) Changes in land use alter the structure of bacterial communities in Western Amazon soils. *The ISME journal* 3:1004-1011
- Dell'Anno A, Beolchini F, Rocchetti L, Luna GM, Danovaro R (2012) High bacterial biodiversity increases degradation performance of hydrocarbons during bioremediation of contaminated harbor marine sediments. *Environmental Pollution* 167:85-92
- Dominguez-Rosado E, Pichtel J (2004) Phytoremediation of soil contaminated with used motor oil: II. Greenhouse studies. *Environ Eng Sci* 21:169-180
- Dominguez-Rosado E, Pichtel J, Coughlin M (2004) Phytoremediation of soil contaminated with used motor oil: I. Enhanced microbial activities from laboratory and growth chamber studies. *Environ Eng Sci* 21:157-168
- Dunlevy SR, Singleton DR, Aitken MD (2013) Biostimulation reveals functional redundancy of anthracene-degrading bacteria in polycyclic aromatic hydrocarbon-contaminated soil. *Environ Eng Sci* 30:697-705
- Ferradji FZ, Mnif S, Badis A, Rebbani S, Fodil D, Eddouaouda K, Sayadi S (2014) Naphthalene and crude oil degradation by biosurfactant producing *Streptomyces* spp. isolated from Mitidja plain soil (North of Algeria). *Int Biodeterior Biodegrad* 86:300-308
- Giebler J, Wick LY, Chatzinotas A, Harms H (2013) Alkane-degrading bacteria at the soil-litter interface: comparing isolates with T-RFLP-based community profiles. *FEMS Microbiol Ecol* 86:45-58. doi: 10.1111/1574-6941.12097 [doi]
- Hobbie SE (2015) Plant species effects on nutrient cycling: revisiting litter feedbacks. *Trends in ecology & evolution*
- Jasinskas A, Zaltauskas A, Kryzeviciene A (2008) The investigation of growing and using of tall perennial grasses as energy crops. *Biomass Bioenergy* 32:981-987. doi: <http://dx.doi.org/10.1016/j.biombioe.2008.01.025>
- Johnson M, Lee K, Scow K (2003) DNA fingerprinting reveals links among agricultural crops, soil properties, and the composition of soil microbial communities. *Geoderma* 114:279-303
- Jones MD, Crandell DW, Singleton DR, Aitken MD (2011) Stable-isotope probing of the polycyclic aromatic hydrocarbon-degrading bacterial guild in a contaminated soil. *Environ Microbiol* 13:2623-2632
- Jones MD, Singleton DR, Carstensen DP, Powell SN, Swanson JS, Pfaender FK, Aitken MD (2008) Effect of incubation conditions on the enrichment of pyrene-degrading bacteria identified by stable-isotope probing in an aged, PAH-contaminated soil. *Microb Ecol* 56:341-349
- Jussila MM, Jurgens G, Lindström K, Suominen L (2006) Genetic diversity of culturable bacteria in oil-contaminated rhizosphere of *Galega orientalis*. *Environmental pollution* 139:244-257
- Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq

Illumina sequencing platform. *Appl Environ Microbiol* 79:5112-5120. doi: 10.1128/AEM.01043-13 [doi]

Kryževičienė A, Jasinskas A, Gulbinas A (2008) Perennial grasses as a source of bioenergy in Lithuania. *Agron.Res* 6:229-239

Kuramae EE, Gamper HA, Yergeau E, Piceno YM, Brodie EL, DeSantis TZ, Andersen GL, van Veen JA, Kowalchuk GA (2010) Microbial secondary succession in a chronosequence of chalk grasslands. *The ISME journal* 4:711-715

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. doi: FEM250 [pii]

Lafortune I, Juteau P, Déziel E, Lépine F, Beaudet R, Villemur R (2009) Bacterial diversity of a consortium degrading high-molecular-weight polycyclic aromatic hydrocarbons in a two-liquid phase biosystem. *Microb Ecol* 57:455-468

Lamendella R, Strutt S, Borglin S, Chakraborty R, Tas N, Mason OU, Hultman J, Prestat E, Hazen TC, Jansson JK (2014) Assessment of the Deepwater Horizon oil spill impact on Gulf coast microbial communities. *Frontiers in microbiology* 5

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. doi: 10.1128/AEM.00335-09 [doi]

Lee KC, Dunfield PF, Stott MB (2014) The Phylum Armatimonadetes. In: *The Prokaryotes*. Springer, pp 447-458

Liang Y, Van Nostrand JD, Deng Y, He Z, Wu L, Zhang X, Li G, Zhou J (2011) Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. *The ISME journal* 5:403-413

Liao J, Wang J, Huang Y (2015) Bacterial Community Features Are Shaped by Geographic Location, Physicochemical Properties, and Oil Contamination of Soil in Main Oil Fields of China. *Microb Ecol*:1-10

Lindstrom K, Jussila MM, Hintsa H, Kaksonen A, Mokolke L, Makelainen K, Pitkajarvi J, Suominen L (2003) Potential of the Galega-Rhizobium galegae system for bioremediation of oil-contaminated soil. *Food Technology and Biotechnology* 41:11-16

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550

Masuda H, Shiwa Y, Yoshikawa H, Zylstra GJ (2014) Draft Genome Sequence of the Versatile Alkane-Degrading Bacterium *Aquabacterium* sp. Strain NJ1. *Genome Announc* 2:10.1128/genomeA.01271-14. doi: 10.1128/genomeA.01271-14 [doi]

McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82:290-297

- Mikkonen A, Kondo E, Lappi K, Wallenius K, Lindström K, Hartikainen H, Suominen L (2011a) Contaminant and plant-derived changes in soil chemical and microbiological indicators during fuel oil rhizoremediation with *Galega orientalis*. *Geoderma* 160:336-346
- Mikkonen A, Lappi K, Wallenius K, Lindström K, Suominen L (2011b) Ecological inference on bacterial succession using curve-based community fingerprint data analysis, demonstrated with rhizoremediation experiment. *FEMS Microbiol Ecol* 78:604-616
- Mills DK, Fitzgerald K, Litchfield CD, Gillevet PM (2003) A comparison of DNA profiling techniques for monitoring nutrient impact on microbial community composition during bioremediation of petroleum-contaminated soils. *J Microbiol Methods* 54:57-74
- Nie Y, Chi C, Fang H, Liang J, Lu S, Lai G, Tang Y, Wu X (2014) Diverse alkane hydroxylase genes in microorganisms and environments. *Scientific reports* 4
- Pérez-Pantoja D, Donoso R, Agulló L, Córdova M, Seeger M, Pieper DH, González B (2012) Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. *Environ Microbiol* 14:1091-1117
- Pham TV, Piersma SR, Warmoes M, Jimenez CR (2010) On the beta-binomial model for analysis of spectral count data in label-free tandem mass spectrometry-based proteomics. *Bioinformatics* 26:363-369. doi: 10.1093/bioinformatics/btp677 [doi]
- Ransom-Jones E, Jones DL, McCarthy AJ, McDonald JE (2012) The Fibrobacteres: an important phylum of cellulose-degrading bacteria. *Microb Ecol* 63:267-281
- Ros M, Rodríguez I, García C, Hernández MT (2014) Bacterial community in semiarid hydrocarbon contaminated soils treated by aeration and organic amendments. *Int Biodeterior Biodegrad* 94:200-206
- Saito A, Iwabuchi T, Harayama S (2000) A novel phenanthrene dioxygenase from *Nocardioides* sp. Strain KP7: expression in *Escherichia coli*. *J Bacteriol* 182:2134-2141
- Schippers A, Schumann P, Sproer C (2005) *Nocardioides oleivorans* sp. nov., a novel crude-oil-degrading bacterium. *Int J Syst Evol Microbiol* 55:1501-1504. doi: 10.1093/ijsem/55.4.1501 [pii]
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537-7541. doi: 10.1128/AEM.01541-09 [doi]
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13:2498-2504
- Shen F, Lin J, Huang C, Ho Y, Arun A, Young L, Young C (2009) Molecular detection and phylogenetic analysis of the catechol 1, 2-dioxygenase gene from *Gordonia* spp. *Syst Appl Microbiol* 32:291-300

- Singleton DR, Sangaiah R, Gold A, Ball LM, Aitken MD (2006) Identification and quantification of uncultivated Proteobacteria associated with pyrene degradation in a bioreactor treating PAH-contaminated soil. *Environ Microbiol* 8:1736-1745
- Song X, Xu Y, Li G, Zhang Y, Huang T, Hu Z (2011) Isolation, characterization of *Rhodococcus* sp. P14 capable of degrading high-molecular-weight polycyclic aromatic hydrocarbons and aliphatic hydrocarbons. *Mar Pollut Bull* 62:2122-2128
- Suominen L, Jussila M, Mäkeläinen K, Romantschuk M, Lindström K (2000) Evaluation of the *Galega-Rhizobium* galegae system for the bioremediation of oil-contaminated soil. *Environmental pollution* 107:239-244
- Tsuboi S, Yamamura S, Nakajima-Kambe T, Iwasaki K (2015) Diversity of alkane hydroxylase genes on the rhizoplane of grasses planted in petroleum-contaminated soils. *SpringerPlus* 4:1-10
- Vazquez-Duhalt R (1989) Environmental impact of used motor oil. *Sci Total Environ* 79:1-23
- Vinas M, Sabate J, Espuny MJ, Solanas AM (2005) Bacterial community dynamics and polycyclic aromatic hydrocarbon degradation during bioremediation of heavily creosote-contaminated soil. *Appl Environ Microbiol* 71:7008-7018. doi: 71/11/7008 [pii]
- Wenzel WW (2009) Rhizosphere processes and management in plant-assisted bioremediation (phytoremediation) of soils. *Plant Soil* 321:385-408
- Yachi S, Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc Natl Acad Sci USA* 96:1463-1468
- Yakimov MM, Lunsdorf H, Golyshin PN (2003) *Thermoleophilum album* and *Thermoleophilum minutum* are culturable representatives of group 2 of the Rubrobacteridae (Actinobacteria). *Int J Syst Evol Microbiol* 53:377-380. doi: 10.1099/ijs.0.02425-0 [doi]
- Yan L, Sinkko H, Penttinen P, Lindström K (2016) Characterization of successional changes in bacterial community composition during bioremediation of used motor oil-contaminated soil in a boreal climate. *Sci Total Environ* 542:817-825
- Yan L, Penttinen P, Simojoki A, Stoddard FL, Lindström K (2015) Perennial crop growth in oil-contaminated soil in a boreal climate. *Sci Total Environ* 532:752-761
- Yang S, Wen X, Zhao L, Shi Y, Jin H (2014) Crude oil treatment leads to shift of bacterial communities in soils from the deep active layer and upper permafrost along the China-Russia Crude Oil Pipeline route. *PLoS One* 9:e96552. doi: 10.1371/journal.pone.0096552 [doi]
- Young CC, Kampfer P, Ho MJ, Busse HJ, Huber BE, Arun AB, Shen FT, Lai WA, Rekha PD (2007) *Arenimonas malthae* sp. nov., a gammaproteobacterium isolated from an oil-contaminated site. *Int J Syst Evol Microbiol* 57:2790-2793. doi: 57/12/2790 [pii]
- Zhang D, Margesin R (2014) Characterization of culturable heterotrophic bacteria in hydrocarbon-contaminated soil from an alpine former military site. *World Journal of Microbiology and Biotechnology* 30:1717-1724

Table 1. Average Bray-Curtis similarity of bacterial OTU-based community profiles between and within groups in oil-contaminated (Oil+) and control (Oil-) plots at different sampling times (Time).

Comparison between/within groups	Average Bray-Curtis similarity (%)				
	Time	July 2009	May 2010	May 2011	May 2012
Oil-	July 2009	<i>46.1</i>			
	May 2010	45.5	<i>45.6</i>		
	May 2011	43.9***	44.5**	<i>45.6</i>	
	May 2012	43.5***	44.2	43.5**	<i>43.3</i>
Oil+	July 2009	<i>45.1</i>			
	May 2010	42.8***	<i>44.0</i>		
	May 2011	41.6***	42.9***	<i>45.0</i>	
	May 2012	41.4***	42.8***	42.9***	<i>43.7</i>
Between Oil+ and Oil-		44.3*	43.2*	44.5 <i>a</i>	42.4*

“*” $P < 0.05$, “**” $p < 0.01$, “***” $p < 0.001$ and “*a*” $p = 0.053$, produced from PERMANOVA pair-wise test with 9999 permutations, based on a split-plot design. Similarity values of within-group comparisons are in italics.

Figure captions

Fig. 1 Effects of a) oil and b) perennial crops on soil bacterial diversity (Shannon-Wiener index) over time. Abbreviations: “*Oil+*” oil-contaminated plots, “*Oil-*” control plots, “*P(Time)*” *p*-value of sampling times, produced from repeated measures ANOVA based on a split-plot design, “*P(Oil treatment)*” *p*-value of oil treatment and “*P(Crop treatment)*” *p*-value of crop treatment, produced from univariate ANOVA based on a split-plot design.

Fig. 2 Canonical ordination for the Bray-Curtis distance-based discriminant analysis of square root transformed bacterial community data based on *a priori* groups of a) sampling time; b) oil treatment; c) the combined factor group of oil treatment and sampling times; and d) crop treatment. The proportion of samples that were correctly allocated to their own group is indicated next to the corresponding group. Abbreviations: “+” oil-contaminated plots, “-” control plots, “*m*” the number of eigenvectors (PCO axes) chosen in the subset matrix for the canonical analysis, “*prop.G*” the proportion of variation in the data cloud that is described by the resemblance matrix explained by the first *m* PCO axes, “*P*” *p*-value, tested based on 9999 permutations, “*%correct*” the total percentage of samples that were correctly allocated to their own groups using the first *m* PCO axes for the model, “ δ_1^2 ” and “ δ_2^2 ” eigenvalues (squared canonical correlation) of the first and second canonical axes, respectively.

Fig. 3 Bray-Curtis distance-based canonical ordination of bacterial sub-communities with total correct allocation over 60% based on *a priori* groups of the combined factor of oil treatment and sampling times (OilTime). The relative abundance of each class is shown in brackets. Observations with over 60% of correct allocation to their own group are circled with the proportions indicated close to the circles. Abbreviations: “+” oil-contaminated plots, “-” control plots, “*P*” *p*-value, tested based on 9999 permutations, “*Total correct allocation*” the percentage of samples that were correctly allocated to their own group using the first *m* PCO axes for the model, “ δ_1^2 ” and

“ δ^2 ” eigenvalues (squared canonical correlation) of the first and second canonical axes, respectively.

Fig. 4 Temporal abundance changes of representative OTUs. a-c) The three types of significantly different temporal abundance changes between treatments ($p < 0.05$). a) OTU188, 100% identity to the genus *Aquabacterium*, b) OTU29, 100% identity to family *Microbacteriaceae*, and c) OTU132, 100% identity to the genus *Lysinimonas*. d) OTU2, 100% identity to genus *Pseudolabrys*, as an example of similar temporal abundance changes in both treatments after the initial abundance difference. Symbols show the average relative abundances in oil-contaminated and non-contaminated soil and error bars denote the standard error.

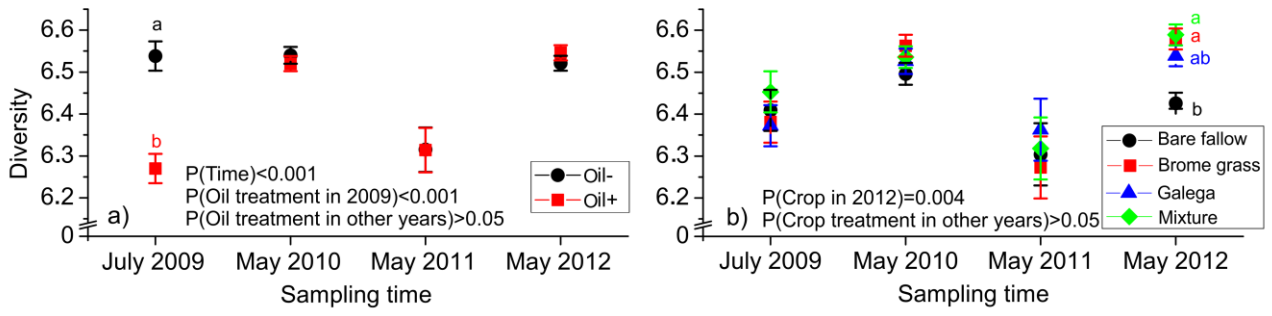


Fig. 1

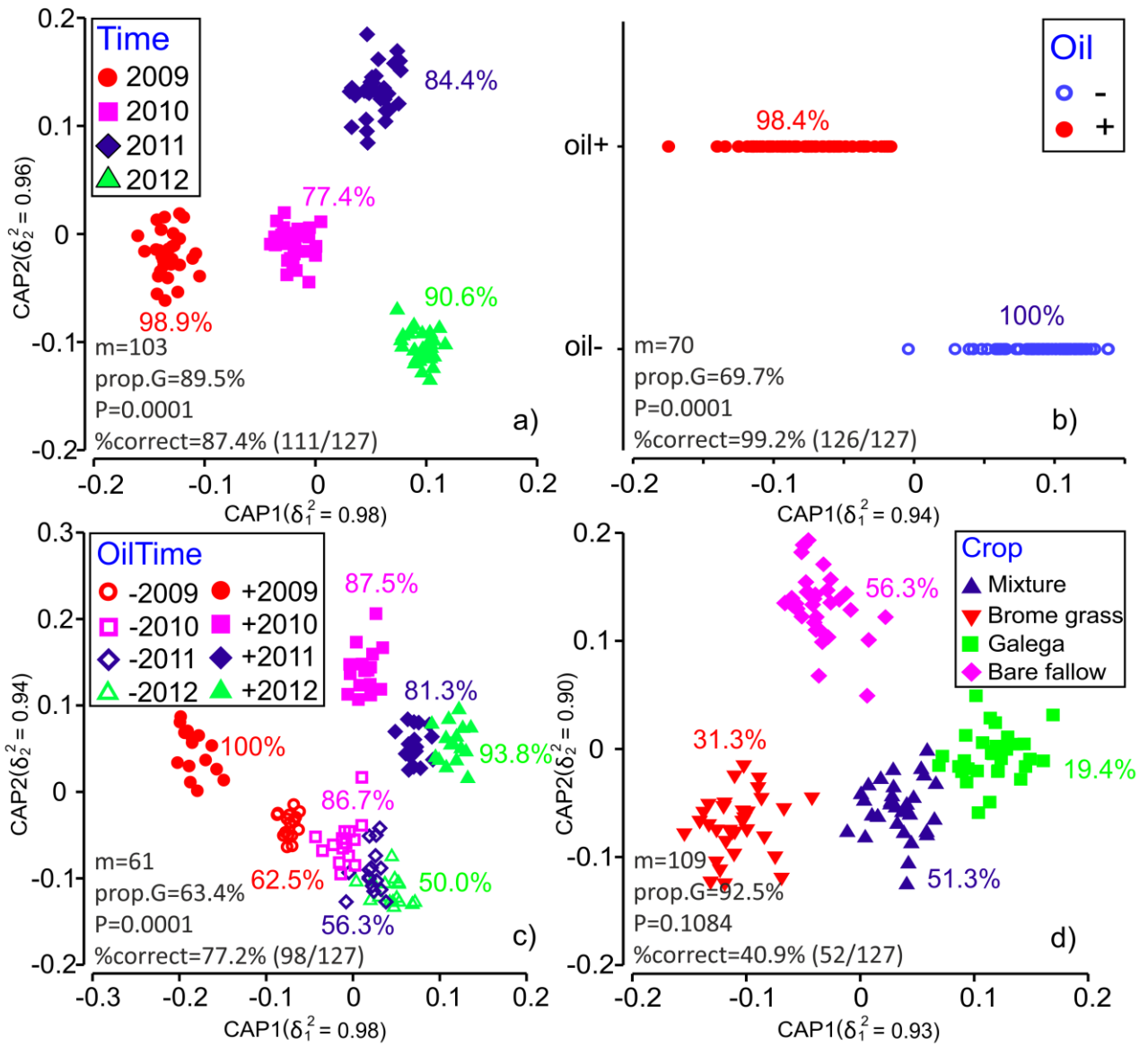


Fig. 2

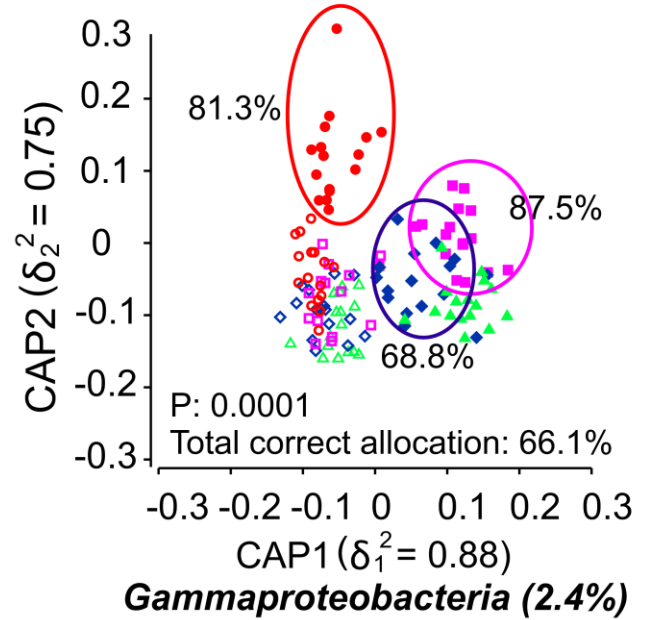
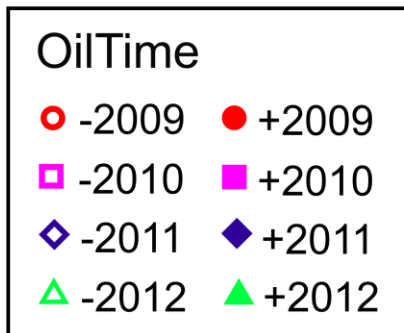
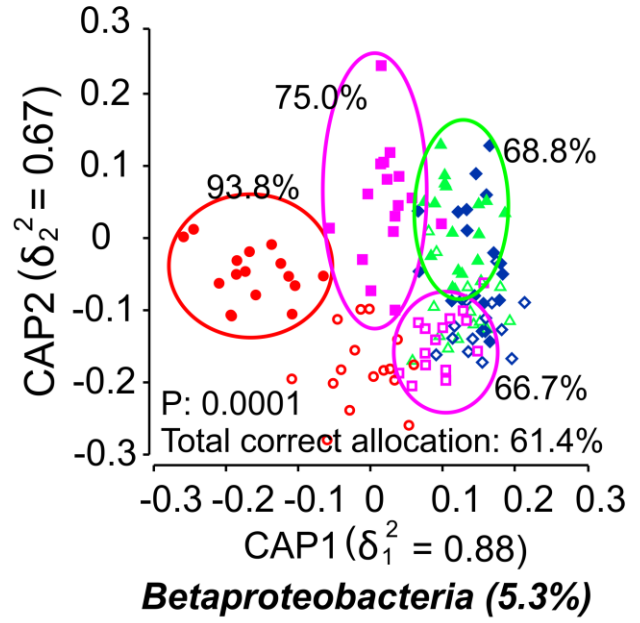
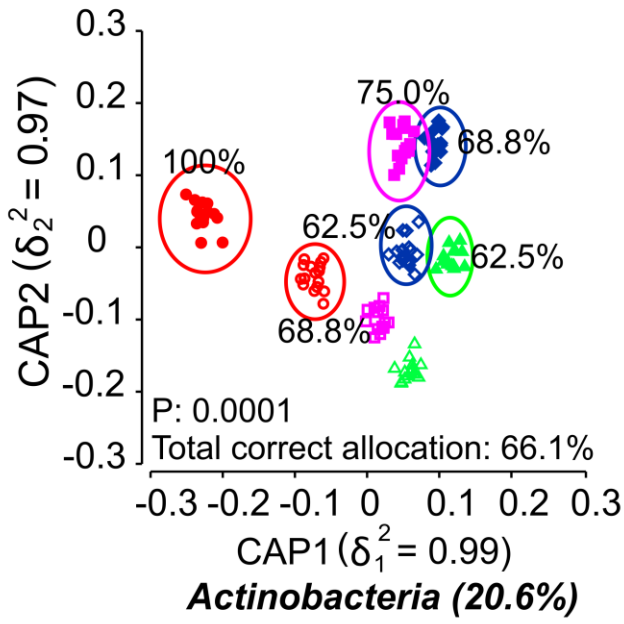


Fig. 3

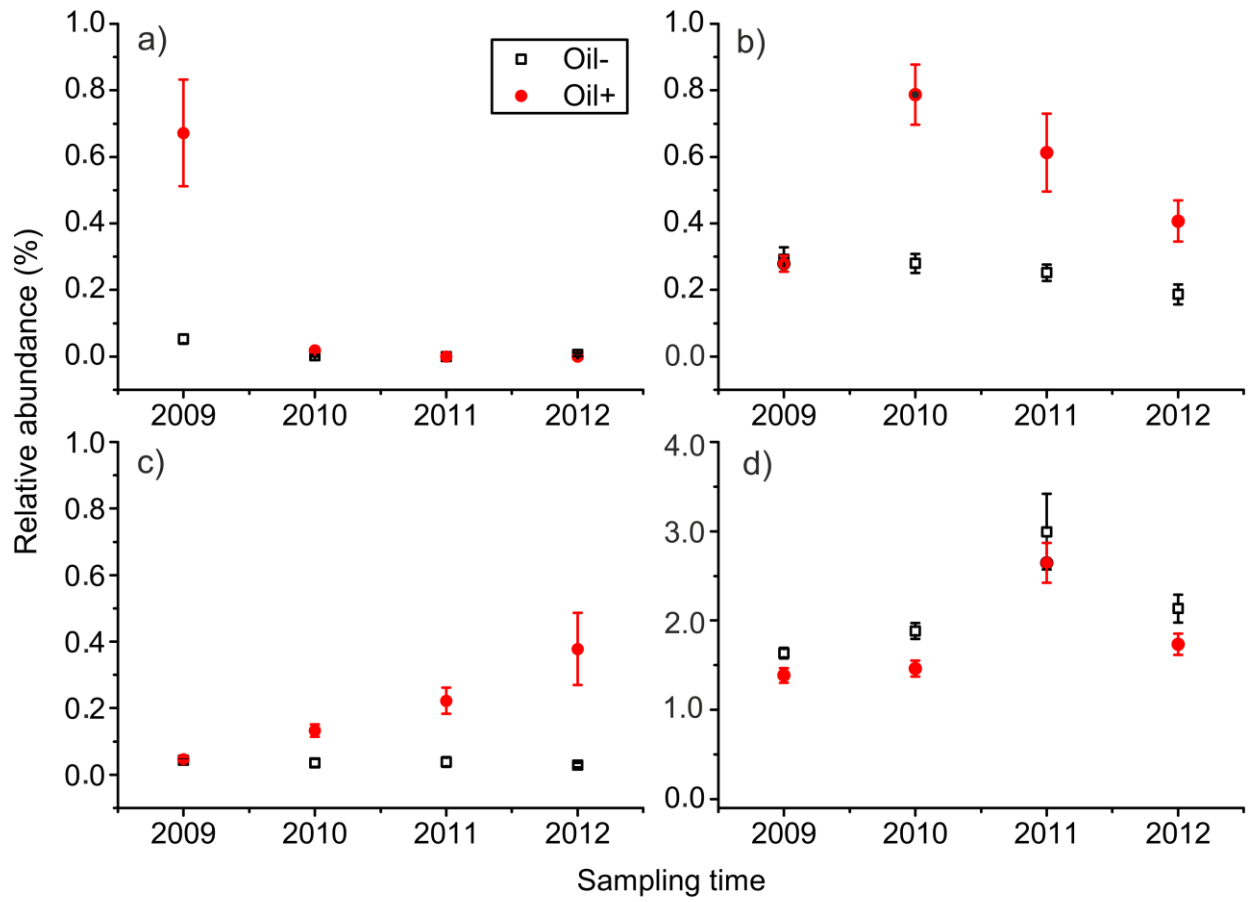


Fig. 4