Identification of Bacteria Utilizing Biphenyl, Benzoate, and Naphthalene in Long-Term Contaminated Soil

Ondrej Uhlik¹, Jiri Wald¹, Michal Strejcek¹, Lucie Musilova¹, Jakub Ridl², Miluse Hroudova², Cestmir Vlcek², Erick Cardenas³, Martina Mackova¹, Tomas Macek¹*

¹ Department of Biochemistry and Microbiology, Faculty of Food and Biochemical Technology, Institute of Chemical Technology Prague, Prague, Czech Republic, ² Department of Genomics and Bioinformatics, Institute of Molecular Genetics, Czech Academy of Sciences, Prague, Czech Republic, ³ Center for Microbial Ecology, Michigan State University, East Lansing, Michigan, United States of America

Abstract

Bacteria were identified with biodegradation of aromatic pollutants biphenyl, benzoate, and naphthalene in a long-term polychlorinated biphenyl- and polyaromatic hydrocarbon-contaminated soil. In order to avoid biases of culture-based approaches, stable isotope probing was applied in combination with sequence analysis of 16S rRNA gene pyrotags amplified from 13C-enriched DNA fractions. Special attention was paid to pyrosequencing data analysis in order to eliminate the errors caused by either generation of amplicons (random errors caused by DNA polymerase, formation of chimeric sequences) or sequencing itself. Therefore, sample DNA was amplified, sequenced, and analyzed along with the DNA of a mock community constructed out of 8 bacterial strains. This warranted that appropriate tools and parameters were chosen for sequence data processing. 13C-labeled metagenomes isolated after the incubation of soil samples with all three studied aromatics were largely dominated by Proteobacteria, namely sequences clustering with the genera Rhodanobacter, Burkholderia, Pandoraea, Dyella as well as some Rudaea- and Skermanella-related ones. Pseudomonads were mostly labeled by 13C from naphthalene and benzoate. The results of this study show that many biphenyl/benzoate-assimilating bacteria derive carbon also from naphthalene, pointing out broader biodegradation abilities of some soil microbiota. The results also demonstrate that, in addition to traditionally isolated genera of degradative bacteria, yet-to-be cultured bacteria are important players in bioremediation. Overall, the study contributes to our understanding of biodegradation processes in contaminated soil. At the same time our results show the importance of sequencing and analyzing a mock community in order to more correctly process and analyze sequence data.

Citation: Uhlik O, Wald J, Strejcek M, Musilova L, Ridl J, et al. (2012) Identification of Bacteria Utilizing Biphenyl, Benzoate, and Naphthalene in Long-Term Contaminated Soil. PLoS ONE 7(7): e40653. doi:10.1371/journal.pone.0040653

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: tomas.macek@vscht.cz

Introduction

Wide-spread use, improper handling, and disposal of certain synthetic organic chemicals have resulted in contamination of soils, waters, and sediments. Among the compounds released, biphenyl, polychlorinated biphenyls (PCBs), petroleum chemicals including polycyclic aromatic hydrocarbons (PAHs) and BTEX (benzene, toluene, ethylbenzene and xylenes), pesticides, dioxins, furans, and flame retardants are of major concern [1,2]. Their presence in the environment often poses a serious risk to both ecosystem functioning and human health. Although anthropogenic activity is a main source of these contaminants, they can as well form during natural events [3]. In addition, some biogenic sources of PAHs have been known, such as endogenic synthesis by microorganisms, phytoplankton, algae, or higher plants [4].

Although biphenyl, PCBs, and PAHs are toxic for many higher organisms, some bacteria are known to transform and/or mineralize these toxicants in the processes of bioremediation [4–6]. Under aerobic conditions, bacteria metabolize biphenyl and its monochlorinated derivatives [7], rarely also dichlorinated ones [8], in order to derive carbon and energy. Other lower chlorinated PCB congeners are usually biodegraded cometabolically [9–11] with biphenyl being the primary substrate. Degradative enzymes encoded in the bph operon transform (chloro)biphenyl to 2-hydroxyphenyl-2,4-dienoate and (chloro)benzoate. 2-hydroxy pentadienoate is further metabolized into acetylcoenzyme A, which enters the Krebs cycle. (Chloro)benzoate is catabolized to catechol, which is further degraded into central intermediates [9]. The simplest PAH, naphthalene, along with some 3- and 4-ring PAHs act as growth substrates for aerobic bacteria [12]. Higher PAHs are usually biodegraded cometabolically and their biotransformation yields no carbon or energy [13]. The upper catabolic pathway for naphthalene degradation consists of six steps with salicylic acid being the final product. Salicylic acid is further metabolized via catechol or gentisic acid by so called lower pathway of naphthalene degradation. In many cases, the same enzymes can transform not only naphthalene but also phenanthrene and anthracene [12].

In order to understand natural biodegradative processes in contaminated environments, it is crucial to identify bacteria involved in pollutants metabolism. Most of our knowledge on biodegradative processes in soil has still been based on highly
biased results of cultivation studies. Stable isotope probing (SIP) [14], in contrast, enables researchers to link metabolic activity and phylogeny without previous cultivation of microbes. Therefore, SIP is considered one of the leading molecular tools for investigating the diversity of bacteria potentially responsible for ecologically relevant processes. Applying this method first helped describe the utilization of one-carbon compounds [e.g. 14–16] and has since been used to detect microbial communities active in the utilization of a wide variety of compounds, including xenobiotics [reviewed by 17,18]. Huge progress in application of SIP has been made as a result of its combination with high-throughput sequencing [19]. For the purposes of bacterial identification, the most informative approach appears to be pyrosequencing of 16 S rRNA tag-encoded amplicons. Recently, combination of SIP with sequence analysis of 16 S rRNA gene pyrotag has been used to probe populations associated with rapid response to soil rewetting [20], nitrification [21], nitrogen incorporation in petroleum-contaminated arctic soils [22], earthworm-mediated shaping of communities [23], toluene degradation at a tar-oil-contaminated aquifer [24], or biphenyl-oxidation in a tidal mudflat [25]. At the same time, bottlenecks of this approach have been realized which are associated especially with the noise introduced during amplification and pyrosequencing leading to an overestimation of diversity [26]. Such a realization has resulted in the introduction of several denoising algorithms or improved operational taxonomic unit (OTU) clustering [27–30] in order to increase data analysis exactness.

The aim of the presented study was to identify bacterial populations in contaminated soil that derive carbon from biphenyl, naphthalene, and benzoate. In order to reach this goal, we applied DNA-based stable isotope probing with subsequent pyrosequencing of 16 S rRNA gene tag-encoded amplicons. Sequence data were analyzed along with a mock community comprised of known strains in order to determine appropriate parameters for sequence data processing. Results of the study help unveil which bacteria are potentially involved in biodegradation of aromatic pollutants. In addition, the experimental design with labeled substrates as well as key intermediates of degradation pathways (biphenyl and benzoate, respectively) helps clarify the flow of carbon in the contaminated soil. The experimental design also has another advantage that it can detect abilities of strains to biodegrade multiple contaminants.

Results

Stable Isotope Probing of DNA

Microcosms constructed with contaminated soil and amended with 13C-labeled biphenyl, benzoate, or naphthalene were destructively harvested after 4 and 14 days of incubation. Majority of unlabeled DNA occurred at buoyant density of around 1.6 g.mL⁻¹. Labeled DNA was localized in fractions with buoyant densities 1.623–1.673 g.mL⁻¹ as is indicated by comparing the distribution of 16 S rRNA genes in gradient fractions of unlabeled control and 13C-DNA. These fractions were compiled and further analyzed as 13C-DNA (“heavy” DNA). One of the major issues associated with stable isotope probing is to control for background 12C-DNA contamination in the heavy-DNA-containing fractions.

To cope with this phenomenon, fractions of unlabeled control with buoyant densities of 1.623–1.673 g.mL⁻¹ were compiled and further analyzed along with the samples. The sequences from all samples that were ≥99% identical to those of unlabeled control were subtracted from analysis as they were considered background 12C-DNA contamination.

Distributions of 16 S rRNA genes in gradient fractions indicate that in case of biphenyl significant labeling was achieved on day 14, whereas on day 4 DNA was not labeled enough to be distinguished from the unlabeled control (the curves have almost the same behavior). Incubation time of 4 days, however, was sufficient for cells to derive carbon from 13C-labeled benzoate and naphthalene (Figure S1).

Identification of Bacteria

The majority of 16 S rRNA gene sequences retrieved from the contaminated soil clustered with Proteobacteria and Acidobacteria – together these two phyla represented two thirds of the sequences (Figure 1). Sequences of the phylum Proteobacteria were also the predominant ones retrieved from 13C-DNA. 13C-DNA obtained after a 4-day incubation with 13C-biphenyl contained less than 25% of valid sequences (Table 1). This was caused by subtracting the majority of sequences that were also detected in the unlabeled control. Such a small number of sequences was expected based on data obtained from qPCR (Figure S1), which suggested there was not enough DNA labeling on day 4. Analyzing 13C-DNA isolated after 14-day incubation of the soil with biphenyl showed that carbon from biphenyl has been mainly utilized by the genera Rhodanobacter and Burkholderia (38% of all sequences) (Table 2). The major OTU identified as Rhodanobacter was most likely associated with R. sphaerophylii as determined by RDP Seqmatch. In addition to Rhodanobacter, other Gammaproteobacteria from the family Xanthomonadaceae have been detected, including Dyella and Rudnea-related populations. Betaproteobacteria were represented in addition to Burkholderia by Pandorea. Further bacteria included in the top OTUs belonged to Alphaproteobacteria (Table 2). Many of the sequences clustering with major OTUs were not classified at the level of genus, suggesting a possible detection of yet-to-be-described bacterial species.

The majority of sequences detected after 13C-benzoate labeling clustered with pseudomonads. This OTU was also detected in 13C-biphenyl labeled DNA on day 14 but was very scarce (0.3% sequences). The second most abundant cluster of sequences after both 4- and 14-day incubation was identical to the Rhodanobacter cluster which dominated carbon acquisition from 13C-biphenyl. Similarly, other populations were detected to derive carbon from benzoate as well as biphenyl, including those associated with Burkholderia, Pandorea, or Rudnea-like and Skermanella-like Proteobacteria, and some Acidobacteria (Table 2, Table S1). In addition, carbon from 13C-benzoate was acquired by populations of Azitolobacter (3% and 6% sequences on day 4 and 14, respectively) and Gram-positive bacilli (2% and 7% sequences on day 4 and 14, respectively) (Table 2).

Sequences detected in 13C-DNA isolated from 13C-naphthalene-amended microcosms were also largely predominated by pseudomonads. Their predominance is especially obvious on day 4. On day 14, the peak representing heavy DNA was much smaller compared to day 4 (Figure S1) and also many sequences detected in heavy DNA on day 14 were classified as 13C-background (Table 1). The major cluster of pseudomonads corresponds to the major ones detected in benzoate-amended microcosms, i.e. identical populations of pseudomonads seem to acquire carbon from both, benzoate and naphthalene. Many other sequences detected in 13C-DNA after naphthalene labeling were identical to those detected after biphenyl and benzoate labeling, including Rhodanobacter, Burkholderia, Pandorea, Dyella as well as many OTUs not classified at the level of genus. These populations thus seem to play a key role in intrinsic bioremediation of a wider variety of aromatic contaminants in the soil.
Mock Community Analysis

Pyrosequencing of 16S rRNA genes of 8 strains included in our mock community was performed in order to (i) analyze the accuracy and efficiency of the amplification and sequencing strategy and (ii) determine appropriate parameters for sequence data processing. A mock community was constructed with strains from phyla Proteobacteria, Actinobacteria and Firmicutes (Table 3). Members of these phyla are most commonly associated with the degradation of aromatics. From over 40,000 reads obtained by pyrosequencing, a random subset of 1,000 and 10,000 sequences was chosen which was further used as a testing set for determining the appropriate analysis procedures. Details of the analysis performed in mothur software package, version 1.25 [31] are described in Material and methods section.

The number of OTUs (defined at 3% distance) detected in the subset differed depending on the program used to detect and remove putative chimeric sequences. Using UCHIME [32] for checking chimeras, the processing resulted in 19 OTUs (defined at 3% distance) whereas using Perseus [29], the final number of OTUs was reduced to 13. The same trend was obvious when a subset of 10,000 reads of mock community were analyzed – using UCHIME and Perseus, the total number of OTUs was 64 and 41. After OTUs represented by 3 or less sequences were subtracted, the numbers of OTUs were reduced to 14 and 8, respectively. Applying Perseus, therefore, resulted in the expected number of OTUs (i.e. 8) after the exclusion of tripletons, doubletons, and singletons. After normalizing sequence data, the spurious OTUs were restricted to singletons only (Table S1).

Other algorithms for chimera checking, such as Bellerophon [33] or B2C2 [34], which is a stand-alone software not implemented in mothur, had a worse performance than both Perseus and UCHIME (data not shown). Therefore, sequence data were processed according to commands described in Table S2 with the use of Perseus for identifying chimeric sequences.

| Sample      | Sequences after processing | Valid sequences | Normalized sequences |
|-------------|---------------------------|----------------|---------------------|
| Total community | 31388                     | 31388          | 244                 |
| Bp4         | 1479                      | 343            | 343                 |
| Bp14        | 23189                     | 22610          | 329                 |
| Bz4         | 5314                      | 4750           | 327                 |
| Bz14        | 25029                     | 21485          | 305                 |
| Np4         | 20721                     | 19625          | 332                 |
| Np14        | 8029                      | 4685           | 336                 |

Table 1. Number of sequences obtained after sequence processing, after subtracting sequences detected also in control DNA (valid sequences), and after normalizing.
Table 2. Top OTUs detected in $^{13}$C-DNA after incubation of soil with $^{13}$C-biphenyl, $^{13}$C-benzoate, and $^{13}$C-naphthalene.

| % of seq* | Identification | Closest type strain(s)* | score* | OTU co-occurence* |
|-----------|----------------|-------------------------|--------|-------------------|
| biphenyl, 14 days | | | | |
| 28 | Rhodanobacter | R. spathiphylli B39; AM087226 | 1.000 | Bz4, Bz14, Np4, Np14 |
| 10 | Burkholderia | B. xenovorans LB400; U86373 | 1.000 | Bz4, Bz14, Np4, Np14 |
| 6 | unclassified Acetobacteraceae | Acidicaldus organivorans Y008; AY140238 | 0.686 | Bz4, Bz14, Np14 |
| 5 | unclassified Rhodospirillaceae | Skermanella xinjiangensis 10-1-101; EU586202 | 0.750 | Bz4, Bz14, Np4, Np14 |
| 4 | Dyella | D. ginsengisoli Gsoil 3046; AB245367 | 1.000 | Bz14, Np4 |
| 3 | unclassified Xanthomonadaceae | Rudaea cellulolytica K53-4; EU741687 | 0.815 | Bz4, Bz14 |
| 2 | unclassified Bradyrhizobiaceae | Bradyrhizobium spp. (more type strains) | 1.000 | Bz4, Bz14 |

benzoate, 4 days | | | | |
| 25 | Pseudomonas | P. umsongensis Ps 3-10; AF468450 | 1.000 | Bz14, Bp4, Bp14, Np4, Np14 |
| 9 | Rhodanobacter | R. spathiphylli B39; AM087226 | 1.000 | Bz14, Bp4, Bp14, Np4, Np14 |
| 6 | Gemmatimonas | G. aurantiaca T-27; AB072735 | 0.570 | Np14 |
| 3 | unclassified Bacteria | Desmopora activa IMMB L-1269; AM940019 | 0.486 | Bz4, Bz14 |
| 3 | Sphingomonas | S. maloi IFO 10550-T; Y09638 | 1.000 | Np14 |

benzoate, 14 days | | | | |
| 28 | Pseudomonas | P. umsongensis Ps 3-10; AF468450 | 1.000 | Bz4, Bp4, Bp14, Np4, Np14 |
| 10 | Rhodanobacter | R. spathiphylli B39; AM087226 | 1.000 | Bz4, Bp14, Np4, Np14 |
| 7 | Sporosarcina | Bacillus spp. (more type strains) | 1.000 | Bz4 |
| % of seq* | Identificationa | Closest type strain(s)b | scorec | OTU co-occurence*d |
|-----------|-----------------|------------------------|--------|------------------|
| 6         | Azotobacter     | *Pseudomonas azotifigens* GH33b; AB189452 | 0.943 Bz4 |
| 4         | unclassified Xanthomonadaceae | *Rudaee cellulosyltica KS3-4* EU741687 | 0.815 Bz4, Bp14 |
| 3         | Burkholderia    | *B. xenovorans* LB400; U86373 B | 1.000 Bz4, Bp14, Np4, Np14 |
| 2         | Pandoraea       | *P. apista* LM6 16407; AF139173 | 1.000 Bz4, Bp4, Bp14, Np4 |
| 2         | unclassified Acetobacteraceae | *Inquilinus limosus* AU476; AY043374 | 0.700 Bz4, Bp14 |
| 2         | unclassified Rhodospirillaceae | *Skermanella xinjiangensis* 10-1-101; EU586202 | 0.750 Bz4, Bp14, Np4, Np14 |
| naphthalene, 4 days | 81 | *Pseudomonas* | *P. umsongensis* Ps 3–10; AF468450 | 1.000 Np14, Bp4, Bp14, Bz4, Bz14 |
| 8         | Rhodanobacter   | *R. spathiphily* B39; AM087226 | 1.000 Np14, Bp14, Bz4, Bz14 |
| 2         | unclassified Rhodospirillaceae | *Oceanibaculum pacificum* MC2UP-L3; FJ463255 | 0.703 Bz4 |
| 2         | Acidobacteria Gp6 | *Ruminobacter amylophilus* DSM 1361; Y15992 | 0.510 Bp4, Bz4, Bz14 |
| 2         | Dongia         | *D. mobilis* LM22; FAJ45532 | 0.844 Bp4, Bz4, Bz14 |
| 1         | unclassified Proteobacteria | *Caloramator fervidus* RT4.B1; L09187 | 0.541 Np4 |
| 1         | unclassified Rhodospirillaceae | *Skermanella xinjiangensis* 10-1-101; EU586202 | 0.750 Np14, Bp14, Bz4, Bz14 |
| naphthalene, 14 days | 40 | *Pseudomonas* | *P. umsongensis* Ps 3–10; AF468450 | 1.000 Np4, Bp4, Bp14, Bz4, Bz14 |
| 15        | Rhodanobacter   | *R. spathiphily* B39; AM087226 | 1.000 Np4, Bp14, Bz4, Bz14 |
| 8         | unclassified Acetobacteraceae | *Roseomonas stagni* HS-69; AB369258 | 0.734 – |
| 7         | unclassified Rhodospirillaceae | *Skermanella xinjiangensis* 10-1-101; EU586202 | 0.750 Np4, Bp14, Bz4, Bz14 |
| 5         | Aeromicrobium   | *A. marinum* T2; AY166703 | 1.000 – |

*Note: Identification includes the species name and strain accession numbers.*
Discussion

In order to better understand, optimize and/or monitor bioremediation processes, a fundamental goal has been to identify bacteria involved in biodegradation of pollutants. The aim of this study was to investigate which bacterial populations participate in biphenyl, benzoate, and naphthalene biodegradation in real long term contaminated soil. This soil represents a unique ecosystem since it was gradually contaminated for 30 years and it has been deposited at the present location for 15 years [35]. Such a long time should have been sufficient for the establishment of stable microbial populations. Cultivation-based techniques highly underestimate diversity as only 1% of microbes are routinely deposited at the present location for 15 years [36,37]. Linking contaminant transformation to phylogenetic identity of active microorganisms without cultivation has thus become the main challenge. SIP is one of the available methods of functional molecular microbial ecology which can meet this challenge [38]. In this study, SIP was combined with pyrosequencing of 16 S rRNA gene amplicons. Such an approach should warrant diversity analysis in sufficient depth.

The results of stable isotope probing experiments show that biotransformation of aromatic pollutants in this soil is mainly mediated by Proteobacteria. There were populations detected that derived carbon from both labeled biphenyl and benzoate, including Rhodanobacter, Burkholderia, Pandoraea, Dyella, as well as Rudana- and Skermanella-related bacteria. As a sufficient amount of labeled DNA was not achieved on day 4 compared to unlabeled control, the question remains whether 13C from biphenyl was derived either directly by biphenyl degradation or indirectly via cross-feeding (possibly with benzoate as a substrate). If the latter was the case, one would expect similar distribution of sequences in 13C-DNA isolated after the incubation with both labeled biphenyl and benzoate. However, the richest clusters in 13C-DNA after labeling with benzoate were associated with pseudomonads (Table 2) whereas in 13C-DNA after labeling with biphenyl, pseudomonads were much poorer in abundance (Table S1). Therefore, it seems to be more likely that the populations detected are capable of transforming biphenyl via benzoate into compounds of intermediary metabolism. This may as well be supported by previous reports showing that members of these genera isolated in pure cultures are capable of degrading biphenyl via benzoate to Krebs cycle intermediates. These include some of the model PCB degrading strains Burkholderia xenovorans LB400 [39,40] or Pandoraea profenens (formerly Comamonas testosteroni) B-356 [41,42]. Recently also the closest type strain to the Dyella cluster, D. ginsengioli, has been described to degrade biphenyl [43]. The question remained whether such populations can be metabolically active directly in the soil, which has been confirmed by our results. In previous SIP experiments, however, only Burkholderia members were found to be actively deriving carbon from biphenyl [44] and benzoate [45,46] in soils. All other dominant populations acquiring carbon from biphenyl and benzoate directly in this soil are reported here for the first time. In addition, based on cultivation studies Rhodanobacter members were previously associated only with the degradation of PAHs [47], chlorobenzoates [48], or certain pesticides [49]. This paper, to the best of our knowledge, is the first one to report bacteria clustering with Rhodanobacter spp. to derive carbon from biphenyl. However, there have been studies published detecting Rhodanobacter spp. in total community DNA isolated from PCB-contaminated soils [50,51].

Pseudomonads, which have been reported to be the most prevailing group of bacteria to degrade complex organic compounds [52], largely dominated carbon acquisition from labeled benzoate and naphthalene in this study. A much lower number of sequences associated with this genus were detected to have derived carbon from biphenyl. Many strains belonging to pseudomonads had been reported to degrade biphenyl and PCBs [53–56], PAHs [12], or other aromatic contaminants. Pseudomonas spp. were also detected in other SIP experiments to utilize (i) naphthalene in PAH-contaminated soil [57–59], aquifer [60], and an aerobic bioreactor used to treat contaminated soil [61]; (ii) benzoate in soil [46]; and (iii) biphenyl in PCB-contaminated river sediments [62] or soil [63]. Major OTUs detected in 13C-DNAs were also largely detected in the total community DNA (Figure 1). In addition, the OTU associated with Rhodanobacter spathiphylli was the most abundant taxon of all (Table S1). Therefore, these populations do not seem to be just opportunistic growers taking advantage of external substrates they were provided with during SIP but are rather regular abundant members in this contaminated soil adapted for the biodegradation of the pollutants.

Table 2. Cont.

| % of seq* | Identificationb | Closest type strain(s)c | scored | OTU co-occurencee |
|----------|----------------|------------------------|--------|------------------|
| 4        | unclassified Proteobacteria | Caloramator fervidus RT4.81; L09187 | 0.541 | Np4 |
| 4        | unclassified Bacteria | Caldiferroica yamamurae YMD0722; AB308475 | 0.551 | Bp14, Bz4 |
| 2        | Verrucomicrobiab Subdv. 3 | uncultured Verrucomicrobiom DEV008; AJ401115 | 0.611 | -- |
| 2        | Propionibacterium | P. granulosum DSM 20700; AJ003057 | 1.000 | Np4 |
| 2        | Gemmatimonas | G. aurantiaca T-27; AB072735 | 0.570 | Bz4 |

*Identification was performed by mothur-implemented RDP reference files [78] and the closest type strain was determined by RDP Seqmatch with the representative sequence of each OTU [76]. The entire dataset is in Table S1.

*bRelative abundance of sequences.

cIdentification of OTU based on identification of the representative at the level of genus as determined by RDP classifier (using 50% threshold).

dScore represents S_ab score – the number of (unique) 7-base oligomers shared between the sequence data and a given RDP sequence divided by the lowest number of unique oligos in either of the two sequences.

eRefers to samples where the same OTU was detected.

doi:10.1371/journal.pone.0040653.t002
Other metabolically active Proteobacteria included *Azotobacter*, *Sphingomonas*, and unclassified Rhodospirillaceae, whose sequences were retrieved from 13C-DNA isolated after incubation with 13C-labeled benzoate only, or benzoate and naphthalene, or all three studied compounds, respectively. A relatively large proportion of sequences retrieved from 13C-DNA were not classified at the level of genus. These can be either sequences of bacteria with no cultured and studied representatives or chimeric sequences that were not removed by Perseus. The fact the same OTUs are often present in more samples (Table 2) suggests the former to be much more likely.

In addition to Proteobacteria, bacteria lineages of Gemmatimonadetes, Acidobacteria, and Gram-positive bacilli were detected to derive carbon from some of the labeled aromatics. Ability to degrade several organic pollutants have already been demonstrated for genera *Azotobacter*, *Sphingomonas*, *Bacillus*, and *Gemmatimonas* [64–67], with *Bacillus* and *Sphingomonas* being detected by SIP experiments as a benzoate [46] and biphenyl and pyrene [63,68,69] utilizers, respectively.

Contaminated soil from the locality of Lhenice has been subjected to phyto/rhizoremediation studies previously, including isolation of biphenyl (and potentially PCB) degrading bacteria. Only isolates of the genus *Achromobacter*, however, were identified in this study to be metabolically active in biphenyl biotransformation (Table S1, OTU 64) and were previously obtained in pure culture [70]. Although Ionescu et al. [71] isolated a *Burkholderia* member, the identification was based on biochemical characteristics; 16 S rRNA sequencing showed this strain to cluster with the genus *Achromobacter* [35]. In addition, a time-course stable isotope probing experiment was conducted with the soil from this locality detecting *Hydrogenophaga* and *Panibacillus* as dominant biphenyl-metabolizing genera. Yet the microcosms were constructed by converting the soil into slurry and enriching biphenyl utilizers prior to stable isotope probing [30]. These factors can create favorable conditions for different microbial populations, which we avoided in this study. This is in agreement with results of another study [25], in which *Panibacillus* sequences were also found to dominate 13C-biphenyl-labeled DNA in tidal mudflats.

It was previously highlighted that pyrosequencing of amplicons can overestimate microbial diversity and the datasets need qualified processing before any conclusions are made [26]. This can be performed, for instance, through denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions [28], through denoising flowgrams before these are converted to sequences [29,30], or through improved OTU clustering [27]. Many of these studies also highlight the need to remove chimeric sequences in order to accurately estimate the diversity of a sample. In our study, we applied the mock community analysis-based approach for optimization of data processing in mothur software package (Table S2). Our data processing was based on standardized operating procedure defined for the needs of Human Microbiome Project [72], which uses UCHIME to identify chimeric sequences. Based on our mock community processing Perseus performed even better. Therefore, it was applied when processing sequence data.

Overall, our data show that biotransformation of selected aromatics was mainly performed by the same populations of Proteobacteria. Stable isotope probing combined with ultra-deep sequencing is shown here to be a helpful tool to monitor natural attenuation and investigate bioremediation potential. Results obtained within this study bring insight into the identity of bacteria actively metabolizing contaminants in soil. It points out broader biodegradation abilities of some soil microbiota, which can be further employed in bioremediation of sites with mixed organic contaminants. Many of the metabolically active populations are as-yet uncultured and unclassified bacteria. Therefore, this study shows the importance of cultivation-independent techniques to investigate biodegradation processes. At the same time our results show the importance of sequencing and analyzing a mock community in order to more correctly process and analyze sequence data.

### Materials and Methods

#### SIP Microcosms

The source of soil samples was a dumpsite of contaminated soil in Lhenice, south Bohemia, Czech Republic [73]. The soil originated from a tarmacadum-producing plant in Mileskova, South Bohemia, where it was gradually contaminated from 1960s till 1990. In 1996, the soil was transferred and deposited at the present location. The main contaminants are PCBs. Their total amounts have decreased over the years by natural attenuation, from the firstly determined concentration of 470 µg PCBs/g soil in 1999 [53] to about 100 µg PCBs/g soil in 2008. The sum of PCBs was based solely on indicator congeners so the actual numbers of all congeners are even higher. In addition to PCBs, the soil contains biphenyl, PAHs, pesticides (DDT, traces of hexachlorobenzene and lindane), and heavy metals (Table 4).

The soil samples for the experiment were collected from the depth of about 0.5 m in the late summer of 2008. After sieving and homogenization, a subsample of 60 g was fine-sieved and transferred on a slide of filter-paper in a box with a perforated bottom. The box was closed and inserted into another box with water on the bottom so that the soil was not in contact with the liquid. In such arrangement, the soil was left to moisturize for 3 days in order to increase bacterial activity. SIP microcosms were set in 100 mL serum bottles (Sigma-Aldrich, USA) in triplicates, each containing 2.5 g soil. 13C-labeled biphenyl and naphthalene (Sigma-Aldrich, USA) were dissolved in acetone (concentration 50 mg/mL), and 10 µL of the solutions were applied onto the inner walls of empty serum bottles. The soil was added after acetone was completely evaporated leaving behind crystals of the substrates. Subsequently, the soil was moistened with 500 µL of mineral salt solution [50]. Bottles were sealed with crimp seals with silicone septa to maintain stable atmosphere with continuous evaporation of the substrates at the beginning of the incubation. 13C-labeled benzoate (Sigma-Aldrich, USA) was applied directly in mineral salt solution (500 µL, concentration 1 mg/mL) onto the soil. Microcosms were destructively harvested by freezing (−80°C) after 4 and 14 days of incubation at 20°C. In order to minimize a possible influence of soil heterogeneity, triplicate SIP microcosms

### Table 3. Bacterial strains used for the preparation of the mock community.

| Bacterium                        | BioProject Accession | Reference |
|---------------------------------|----------------------|-----------|
| *Achromobacter xylosoxidans* A8 | PRJNA59899           | [79]      |
| *Pseudomonas putida* JB         | to be released       | [35]      |
| *Rhodobacter capsulatus* SB 1003| PRJNA47509           | [80]      |
| *Agrobacterium tumefaciens* C58 | PRJNAS7865           | [81]      |
| *Arthrobacter chlorophenolicus* A6| PRJNAS9869          | [82]      |
| *Bacillus pumilus* SAFR-032    | PRJNAS9017           | [83]      |
| *Micrococcus luteus* NCTC 2665 | PRJNAS9033           | [84]      |
| *Rhodococcus jostii* RHA1      | PRJNAS8325           | [85]      |

doi:10.1371/journal.pone.0040653.t003
were set for each substrate and each time point. After the incubation period, triplicate samples were pooled for downstream processes to recover $^{13}$C-DNA from all three replicates at the same time.

$^{13}$C-DNA Isolation

DNA was extracted with a PowerMax Soil DNA Isolation Kit (MoBio Laboratories Inc., USA) using the standard protocol. After final elution, DNA was concentrated by ethanol precipitation with glycogen (Roche, Germany) as described previously [50]. Pellets were dissolved in 50 μL of molecular biology water (Sigma-Aldrich, USA), and samples were diluted to a concentration of 100 ng/μL. In 2 mL centrifugation cuvettes, 8 μL of DNA solution (800 ng DNA) was mixed with cesium trichloroacetate (Amersham, UK) adjusted to a density of 1.6 g/mL. Iodopicn centrifugation was performed on a Discovery 90 Ultracentrifuge using TFT-80.2 Fixed-Angle Ultracentrifuge (Beckman Coulter, USA) and Harvard Pump 11 Plus Single Syringe (Harvard Apparatus, USA), each gradient was fractionated 70 hours in 2 mL cuvettes. Using a Beckman Fraction Recovery System (Beckman Coulter, USA) to remove residual primer-dimers according to manufacturer’s instructions. Amplicons were unidirectionally sequenced from the forward primer using GS FLX Titanium chemistry followed by Pyrosequencing. The amplicon pool was purified with a QIAquick PCR purification kit (Qiagen, Germany) and mixed together with other barcoded samples for pyrosequencing. The amplicon pool was purified using AMPure XP Beads (Agencourt, Beckman Coulter, USA) to remove residual primer-dimers according to manufacturer’s instructions. Amplicons were unidirectionally sequenced from the forward primer using GS FLX Titanium chemistry followed by amplicon analysis of signal processing (Roche).

Mock Community as a Control for Pyrosequencing

A mock community was constructed by mixing genomic DNA from eight bacterial strains (Table 3). After the strains were grown overnight in liquid Luria Bertani medium (Oxoid, UK), their genomic DNA was isolated with a PureLink Genomic DNA Mini Kit (Invitrogen, USA). The number of 16 S rRNA gene copies per ng of each genomic DNA was determined by real-time qPCR as described above. Genomic DNA samples of the selected strains were mixed so that the numbers of 16 S rRNA genes from each strain were of the same order of magnitude. Only the amount of Rhabdovirus jostii RHA1 16 S rRNA genes was an order of magnitude lower.

Table 4. Analysis of inorganic nutrients and contaminants available in the soil.

| Parameter                        | Unit    | Result | Standard error |
|----------------------------------|---------|--------|----------------|
| inorganic carbon                 | % dry matter  | 1.46   | 11%            |
| inorganic nitrogen               |         | 0.12   | 55%            |
| inorganic sulfur                 |         | <0.1   | -              |
| sum of PCBs 28, 52, 101, 118, 138, 153, 180 | mg/kg dry matter | 96.74  | 40%            |
| sum of carcinogenic PAHs         |         | 0.33   | 30%            |
| sum of non-carcinogenic PAHs     |         | 0.30   | 30%            |
| Fe                               |         | 51.9   | 20%            |
| Ni                               |         | 88.0   | 20%            |
| Zn                               |         |        |                |

Results shown are averages from 5 independently measured samples (performed commercially).

aBased on method CZ_SOP_D06_07_121.
bBased on method US EPA 2082.
cBased on methods EPA 8270, EPA 8131, EPA 8091.
dBased on method US EPA 2007.

doi:10.1371/journal.pone.0040653.t004

Bacteria Utilizing Aromatics in Contaminated Soil
Analysis of 16 S rRNA Pyrotags

Raw pyrosequencing data (*.sff files) were processed within mothur software package, version 1.25 [31]. Except some minor modifications, processing was based on standardized operating procedure defined for the needs of Human Microbiome Project [72]. The whole procedure is described in Table S2 and was set this way based on mock community (Table 3) analysis. Sequences in 13C-DNA which were more than 99% identical to those detected in control set (DNA from fractions of unlabeled control corresponding to those where 13C-DNA was detected in samples after probing) were considered contamination and were subtracted from analyses. Sequences with less than 3% sequence similarity were considered the same operational taxonomic units (OTUs). These were classified by mothur-implemented RDP reference files [78]. The closest type strain was determined by RDP Seqmatch with the representative sequence of each OTU [76].

Nucleotide Sequence Accession Numbers
The nucleotide sequences have been submitted to the European Nucleotide Archive under the accession number ERP001002.

Supporting Information

Figure S1 Quantitative PCR-based detection of 13C-DNA in density gradient fractions (average from two independent measurements).

References

1. Chaudhry Q, Blom-Zandstra M, Gupta S, Joner EJ (2005) Utilising the synergy between plants and rhizosphere microorganisms to enhance breakdown of organic pollutants in the environment. Environ Sci Pollut Res Int 12: 34–48.
2. McGuinness M, Dowling D (2009) Plant-associated bacterial degradation of toxic organic compounds in soil. Int J Environ Res Public Health 6: 2226–2247.
3. Manoli E, Samara C (1999) Polyacrylic aromatic hydrocarbons in natural waters: sources, occurrence and analysis. Trends in Analytical Chemistry 18: 417–420.
4. Mueggenburg MP, Sistrom JJ, Whiteley AM (2010) Environmental bacterial community structure and function in the rhizosphere of hydrocarbon-degrading plants. Appl Environ Microbiol 76: 7537–7541.

Table S1 Clusters (defined at 3% distance) of sequences detected in total community and 13C-DNA after 4- and 14-day incubation of soil with 13C-labeled substrates (biphenyl – Bp, benzoate – Bz, and naphthalene – Np). Numbers of sequences in OTUs are normalized. Identification was performed by mothur-implemented RDP reference files [78].

Table S2 Commands used for pyrosequencing data processing in mothur software package, version 1.25 [31,72]. The table also summarizes the number of valid and unique sequences from a 10,000-read subset of mock community that passed the selected criteria.

Acknowledgments

The authors acknowledge Mary Beth Leigh, Mary-Catherine Lecuis, Qong Wang, Patrick Schloss and Sarah Westcott for the advice on experimental setup and data analysis.

Author Contributions

Conceived and designed the experiments: OU MM TM. Performed the experiments: JW OU JR MH. Analyzed the data: OU MS LM. Contributed reagents/materials/analysis tools: CV EC. Wrote the paper: OU MS LM MM TM.
Bacteria Utilizing Aromatics in Contaminated Soil

33. Haber T, Faulkner G, Hugenholtz P (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics 20: 2317–2319.

34. Gontcharova V, Yoon E, Wolkod RT, Hollister EB, Gentry TJ, et al. (2010) Black Box Chimera Check (BBCC): a Windows-based software for batch dechimerization of bacterial 16 S rRNA gene datasets. The Open Microbiology Journal 4: 47–52.

35. Macková M, Prouzoš P, Stursa P, Ryšalová E, Uličný O, et al. (2009) Phyto/chloroformex remediation studies using long-term PCB-contaminated soil. Environ Sci Pollut Res 16: 817–829.

36. Lozoue CA, Knight R (2008) Species diversity and the measurement of microbial diversity. FEMS Microbiol Rev 32: 557–578.

37. Zhang L, Xu Z (2008) Assessing bacterial diversity in soil. J Soil Sediment 8: 379–381.

38. Gutiérrez-Zamora ML, Manfied M (2010) An appraisal of methods for linking environmental processes to specific microbial taxa. Reviews in Environmental Science and Bioengineering 3: 153–187.

39. Bopp LH (1986) Degradation of highly chlorinated PCBs by Pseudomonas strain LB400. J Indust Microbiol 1: 23–29.

40. Chain PS, Denef VJ, Konstantinidis KT, Vergez LM, Agullo L, et al. (2006) Bf2072. J Biol Chem 281: 8480–8488. A new thermophile, 9.73 Mbp genome shaped for versatility. Proc Natl Acad Sci USA 103: 15280–15287.

41. Vézina J, Barriault D, Sylvèstre M (2008) Diversity of the C-terminus portion of the biphenyl dioxygenase large subunit. J Mol Microbiol Biotechnol 15: 139–145.

42. Hurtubise Y, Barriault D, Powlowski J, Sylvestre M (1995) Purification and characterization of the Comamonas testosteroni B-556 biphenyl dioxygenase component. J Bacteriol 177: 6610–6618.

43. Li A, Qu Y, Zhou J, Ma F (2009) Characterization of a newly isolated biphenyl-degrading bacterium, Dyella ginsengisoli sp. nov. J Microbiol Biotechnol 19: 139–146.

44. Tillmann S, Stromple C, Timmis KN, Abraham WR (2005) Stable isotope probing of the polycyclic aromatic hydrocarbon-degrading bacterial guild in a PCB-contaminated soil. Appl Environ Microbiol 71: 1836–1845.

45. Kinzer P, Tishinaka B, Fetzner S, Langens F (1996) Degradation of benzoate via benzoyl-coenzyme A and glutamate by Bacillus stearothermophilus. Environ Microbiol 8: 523–532.

46. Luo W, D’Angelo EM, Coyne MS (2008) Organic carbon effects on aerobic bacterial communities in aerobic bioreactors treating polycyclic aromatic hydrocarbon-contaminated soil. Biodegradation 22: 1061–1073.

47. Zhang L, Sekiguichi Y, Hanada S, Hugenholtz P, Kim H, et al. (2003) Comamonas aktinomycetemcomitans gen. nov., sp. nov., a gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Comamonadota.Intl. Journal of systematic and evolutionary microbiology 53: 1153–1163.

48. Singleton DR, Saagia R, Gold A, Ball LM, et al. (2005) Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a biocorrosor treatating contaminated soil. Appl Environ Microbiol 71: 1209–1212.

49. Sul WJ, Park J, Quensn JF, III, Rodrigues Jr., Seliger L, et al. (2009) DNA-stable isotope probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from polychlorinated biphenyl-contaminated river sediment. Appl Environ Microbiol 75: 5501–5506.

50. Leigh MB, Pellizari VH, Uličný O, Sutka R, Rodrigues J, et al. (2007) Bipheryl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). Environ Sci Pollut Res 14: 430–436.

51. PloS ONE | www.plosone.org 10 July 2012 | Volume 7 | Issue 7 | e40653

52. E. coli, works in collaboration with other research groups in the area of biodegradation in contaminated sediment. Proc Natl Acad Sci USA 100: 13591–13596.

53. Singleton DR, Powell SN, Saagia R, Gold A, Ball LM, et al. (2005) Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a biocorrosor treatating contaminated soil. Appl Environ Microbiol 71: 1209–1212.

54. Sul WJ, Park J, Quensn JF, III, Rodrigues Jr., Seliger L, et al. (2009) DNA-stable isotope probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from polychlorinated biphenyl-contaminated river sediment. Appl Environ Microbiol 75: 5501–5506.

55. Leigh MB, Pellizari VH, Uličný O, Sutka R, Rodrigues J, et al. (2007) Bipheryl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). Environ Sci Pollut Res 14: 430–436.