Appl Biochem Biotechnol (2008) 151:638–652
DOI 10.1007/s12010-008-8275-3

Functional Gene Abundances (*nahAc*, *alkB*, *xylE*) in the Assessment of the Efficacy of Bioremediation

Jani M. Salminen · Pirjo M. Tuomi · Kirsten S. Jørgensen

Received: 21 December 2007 / Accepted: 1 May 2008 /
Published online: 1 July 2008

Abstract In this study, we compared the mineralization rates of three selected ^{14}C-labeled hydrocarbon compounds, octacosane, toluene, and naphthalene, with the presence of the corresponding functional genes (*alkB*, *xylE*, *nahAc*) in a large number of soil samples representing different types of soil and petroleum hydrocarbon contamination. Functional genes were enumerated by the replicate limited dilution (RLD) polymerase chain reaction (PCR) technique. RLD-PCR was further compared to real-time PCR measurements for *nahAc* and *xylE* for some samples. At a heating oil-contaminated site, octacosane mineralization rates were higher (on average 0.0015 day^{-1}) when compared to aerobic naphthalene and toluene mineralization (on average 0.00003 and 0.0007 day^{-1}). The corresponding gene abundances measured by RLD-PCR were on average 0.95, 0.3, and 0.13×10^3 gene copies g^{-1} soil for *alkB*, *nahAc*, and *xylE*, respectively. At a site contaminated with gasoline, the situation was the opposite: Toluene mineralization was the highest (on average 0.0031 day^{-1}), and only *xylE* genes could be detected (on average 0.13×10^3 gene copies g^{-1} soil by RLD-PCR). *XylE* and *nahAc* gene abundances were correlated with the ^{14}C-toluene and naphthalene mineralization activities, respectively, in samples from aerobic layers. *AlkB* gene abundances were not correlated with the octacosane mineralization. Real-time PCR was a more sensitive method than RLD-PCR by a factor of 1,200 for *nahAc* and 300 for *xylE*. In conclusion, functional gene abundances seemed to reflect the type of the contamination. With optimized assays, the gene abundances can be used to assess bioremediation efficacy.

J. M. Salminen · P. M. Tuomi · K. S. Jørgensen (✉)
Finnish Environment Institute, P.O. Box 140, 00251 Helsinki, Finland
e-mail: kirsten.jorgensen@ymparisto.fi

Present address:
J. M. Salminen
Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 Amsterdam, The Netherlands

Present address:
P. M. Tuomi
Golder Associates Oy, Ruosilankuja 3 E, 00390 Helsinki, Finland

Keywords Functional gene · *alkB* · *nahAc* · *xylE* · Oil-contaminated sites · Bioremediation

Introduction

Intrinsic soil microorganisms possess a great potential to degrade various anthropogenic organic compounds. Today, this potential is widely used in passive and active in situ bioremediation of petroleum hydrocarbon-contaminated soils and groundwater. One of the greatest remaining challenges in the use of in situ bioremediation is how to provide evidence on the progress of the contaminant biodegradation in the subsurface [1]. Ideally, decreasing trends of the contaminants could be shown by soil and groundwater sampling and consecutive chemical analyses. However, the collection of this information requires monitoring over several years. With that said, there is the problem that the heterogeneity of the soil matrix may lead to difficulties in obtaining comparable samples for the creation of a time series of the contaminant concentration. To overcome these problems, evidence for the contaminant-destructive processes, i.e., biodegradation, is often collected, and microcosm studies have traditionally been used for this purpose [2]. Biodegradation rates obtained from microcosm studies may not, however, estimate in situ biodegradation correctly due to sampling bias [2]. As a result, novel methods such as isotope fractionation studies and molecular biology methods, which do not require laboratory incubations, are being developed as monitoring technologies [3, 4]. Though, the applicability and feasibility of such methods remain to be evaluated. Furthermore, as microorganisms tend to attach to surfaces, samples containing both solid and aqueous phases of the aquifer material ought to be used when studying the biodegradation potential in contaminated aquifers [5]. However, the application of molecular biology methods to soil samples is often more challenging than it is to aqueous samples [6].

Deoxyribonucleic acid (DNA) extraction from soil is currently easy to perform, and the resulting DNA extract reflects the intrinsic microbial communities because no laboratory cultivation is needed. By the quantification of functional genes, the presence of the biodegradation potential for petroleum hydrocarbons has been demonstrated [7–9]. By contrast, the comparison of gene numbers and actual degradation rates in soil samples has been less seldom demonstrated: *NahAc* abundance and *nahAc* messenger ribonucleic acid transcript levels were shown to correlate with naphthalene mineralization activity in polycyclic aromatic hydrocarbon-contaminated soils [10–12], and bioaugmented soil samples showed concurrent naphthalene mineralization and *nahAc* gene abundance [13]. Information on the interrelation of degradation gene abundances and microbial activity measured by traditional methods such as contaminant disappearance in microcosms is also needed [12, 14].

Crude oil and refined oil products are complex mixtures of aliphatic (e.g., *n*-alkanes, branched alkanes, and cycloalkanes), aromatic, and polyaromatic compounds. Gasoline typically contains hydrocarbons ranging in size from 4 to 12 carbon atoms. It is also rich in monoaromatic BTEX (that is, benzene, toluene, ethylbenzene, and xylenes) compounds [15]. Lightweight fuels are classified as middle distillates and include hydrocarbons with approximately 9–20 carbon atoms. Lubricants are heavy fuel oils and consist of compounds with between 14 and 30 carbon atoms [15]. In middle distillates and heavy fuel oils, BTEX compounds are found only in trace concentrations. Due to the varying chemical composition of oil products in general, a number of functional genes is involved in the biodegradation of the various oil components. Thus, an array of genes is needed when using this methodology for the estimation of the biodegradation potential. For example, the genes *xylA*, *xylE*, and *tomA* have been used to study the degradation potential for monoaromatic,

nahH, *nahA*, and *phnAc* for polyaromatic and *alkB* for aliphatic petroleum hydrocarbons [7, 9–11, 16–18].

In this study, we aimed at establishing whether the determination of known functional genes from cultured microbes namely, *nahAc*, *alkB*, and *xylE*, which are involved in the biodegradation of naphthalenes, *n*-alkanes, and toluene, respectively, can be used as a reliable tool for assessing the oil biodegradation activity at different sites undergoing natural attenuation. We evaluated how well the abundances of these genes reflected the oil biodegradation potential that we determined using mineralization assays. The microbial communities present at these sites most like also include a large number of hitherto uncultured microbes. We also studied the interrelation between the abundances of the genes, the type of the contamination, and the characteristics of the studied environments (e.g., oxygen conditions). For this purpose, five petroleum hydrocarbon-contaminated sites with different characteristics (extent of microbial activity, contaminant composition, soil type, organic matter content, oxygen conditions, etc.), were chosen.

Materials and Methods

Study Sites

We studied five different petroleum hydrocarbon contaminated sites in southern Finland. The quality of the contaminating petroleum hydrocarbons and the soil type at each site are listed in Table 1. The distribution of the contamination at each site is shown in Fig. 1.

Prior to our investigation, no active remediation had been carried out at sites 1 (Trollberget), 3 (Vuosaari), and 4 (Lempäälä). Site 1 was a closed dump site, where oily wastes had been deposited. The site is described in detail in [19]. Site 3 was located by a closed wastewater treatment plant. At this site, oily wastewater sludge had been deposited into the ground. Site 4 was a former gasoline station, where leakage from underground storage tanks had caused the contamination. At site 2 (Etna), heating oil had leaked under a factory hall. The site had been treated with aeration and nutrient amendments during June 1999–September 2000. Since then, only soil vacuum extraction had been performed, with the residual contamination at the site being left to attenuate naturally. Soil samples from site 2 were taken in November 2000, 2 months after the termination of aeration and nutrient amendment. At that time, oil biodegradation at the site had declined from the onset of the remedial treatments in June 1999. From site 5 (Sköldvik), we obtained highly contaminated surface soil samples. At this site, the deposition of waste from an oil refinery onto an abandoned agricultural field has been ongoing for decades. In addition, the surface soil is fertilized to enhance microbial activity in the surface soil of this site by land farming [12].

Soil Sampling

Table 1 shows how the soil samples were taken from each site. All soil samples except the samples from site 5 were taken by an engine-driven auger at 0.5–1-m-depth intervals. Immediately after withdrawing the samples from the ground, the soil was homogenized within each depth interval and sieved through an 8-mm sieve. At sites 1, 3, and 4, both unsaturated and saturated zones were sampled. At site 2, samples were obtained only from the unsaturated zone. From site 5, surface soil samples were taken with a shovel. All together, 48 soil samples were obtained. All the soil samples, including the microcosms, were transported to the laboratory at +4 to +10 °C.

Table 1 Sampling information, contaminants, and the characteristics of the study sites 1–5.

Site	Name	Contamination	Oxic layer (m)	Soil type	Sampling information			
					Time	Point	Profile depth (m)	Number of samples in profile
1	Trollberget	Diesel fuel, lubrication oil	0–1	Sand, filling	July 2000	G1	0–3	5
			0–1.5	Sand, filling	July 2000	G15	0.5–3.5	6
			0–1.2	Sand, silt	May 2001	G17	0.5–3.5	6
			0–1.5	Sand, filling	May 2001	G18	0.5–5.5	10
2	Etna	Heating oil	0–11	Sand, till, silt	Nov 2000	P51	2–11	7
3	Vuosaari	C10-C30 PHs, wastewater sludge	0–1	Deposited waste water sludge	Mar 2002	2	1.5–2.5	2
			0–1			3	1.5–2.5	2
			0–1			4	1.5–2.0	1
4	Lempäälä	Gasoline	0–1.5	Silt, clay	May 2002	7	2.5–3.5	2
			0–1.5			9	2.0–4.5	3
5	Sköldvik	Refinery waste	0–0.3	Agricultural soil + waste	Aug 2002		0–0.3	3 (no profile)
Clean soil	Trollberget	None	0–4	Sand	May 2001	G4	0.5–1.5	2

Physical and Chemical Analyses

Soil gas composition (CO_2, O_2, CH_4) at the study sites was measured from gas-monitoring wells at sites 1–4 using a portable Dräger Multiwarn II to estimate the in situ microbial activities and oxygen conditions.

Dry weight (DW) of the soil was determined after drying the soil sample for 15–20 h at 105 °C, and the organic matter content (% of DW) was determined thereafter by burning the sample at 550 °C for 2 h.

For the mineral oil analysis, samples were distributed into tightly closed glass bottles and then frozen at −18 °C until the analysis. Total petroleum hydrocarbons (PHC) in the range C_{10}–C_{40} were extracted using acetone–heptane, and the extract was analyzed using gas chromatography–flame ionization detection according to standard procedures [20].

The BTEX, methyl-*tert*-buthyl ether, *tert*-amylether, and naphthalene analyses from the soil samples were performed as previously described [19]. Approximately 10 g of soil was distributed into a 70-ml glass bottle into which 20 ml of methanol was added for preservation. The soil sample for these analyses was taken from the auger before the sieving to avoid volatilization of the compounds to be analyzed.

Soil Respiration

Aerobic respiration experiments were performed in microcosms as previously described [19]. Briefly, soil subsamples were incubated in serum bottles at 8 °C under aerobic conditions.

Mineralization of ^{14}C-Naphthalene, ^{14}C-Octacosane, and Toluene

Mineralization of ^{14}C-naphthalene, ^{14}C-octacosane, and toluene was measured using the method described in [12]. Briefly, 10-g subsamples of soil were put into air-tight 110-ml bottles. One killed control and three replicates were prepared for each soil sample. To every sample, approximately 100,000 DPM of *UL*-^{14}C-naphthalene (31 mCi/mmol, Sigma), 14,15-^{14}C-octacosane (20.5 mCi/mmol), or *UL*-^{14}C-toluene (2.8 mCi/mmol) was added.

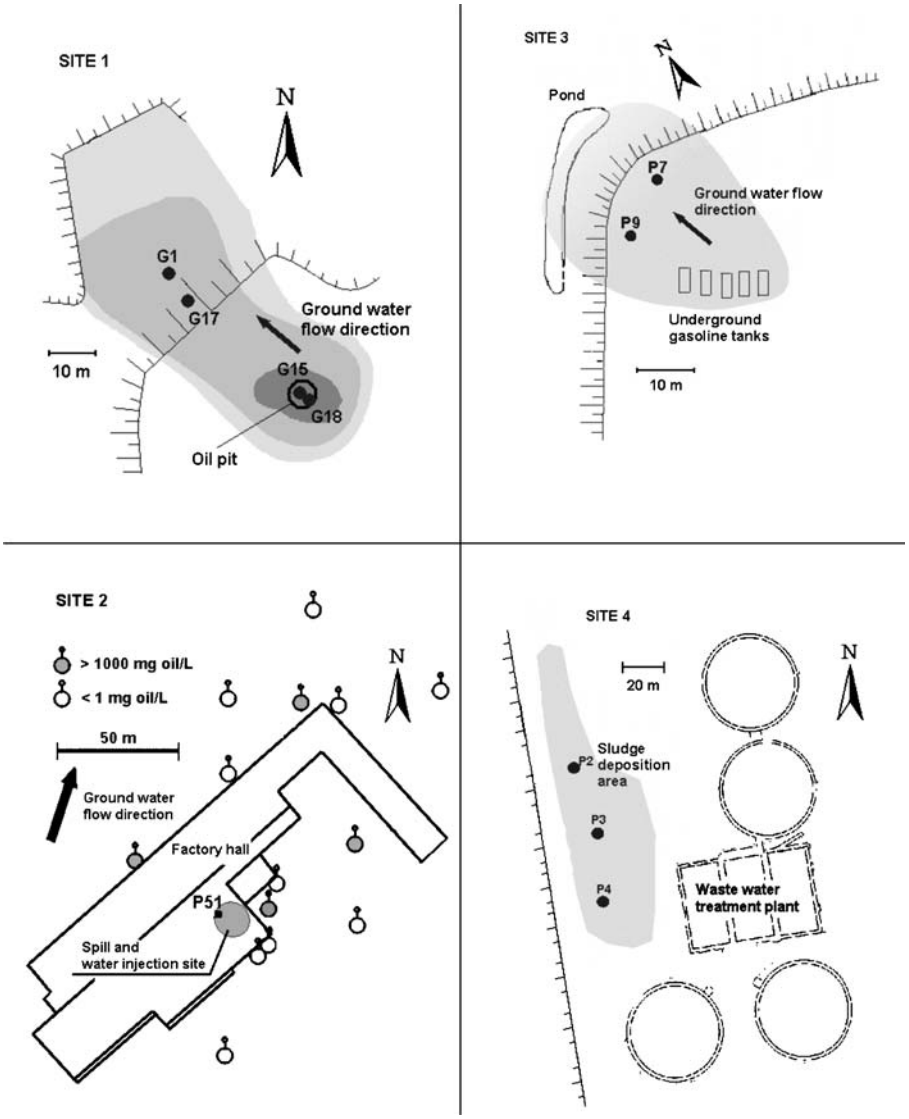

Fig. 1 Spatial distribution of the petroleum hydrocarbon contamination at site 1 (Trollberget), site 2 (ETNA), site 3 (Vuosaari), and site 4 (Lempäälä)

The samples were incubated at +8±1 °C, except for the samples from site 5, which were incubated at 18 °C. The radioactive $^{14}CO_2$ produced in the samples was determined using a liquid scintillation counter.

DNA Extraction from Soil and Control Bacterial Strains

For DNA extraction, 2×0.7-g soil samples were stored at −70 °C to await further preparation. From the soil samples, DNA was extracted directly using FastDNA® SPIN Kit for Soil (BIO 101), which purifies genomic DNA from soil using the FastPrep Instrument

(Savant Instruments). Positive control strains *Pseudomonas putida* G7 (Pp G7, German Resource Centre for Biological Material [DSMZ] no. 4476), *P. putida* mt-2 (Pp mt-2, DSMZ no. 3931, strain from Martin Romantschuk), and *Pseudomonas oleovorans* (Po, ATCC no. 29347) were grown according to DSMZ and American Type Culture Collection instructions, respectively. From pure culture cells, DNA was extracted using MasterPure™ Complete DNA Purification Kit (Epicentre Technologies). This DNA was further purified with Wizard® DNA Clean-Up System (Promega). The DNA yield was quantified using MassRuler DNA Ladder, High Range (MBI Fermentas), or EZ Load Precision Molecular Mass Ruler (BioRad). Three and 5 μl of the mass ruler and DNA sample were run in a 1.5% (wt/vol) agarose gel (0.5× Tris–borate–ethylenediamine tetraacetic acid [TBE]) for 30 min at 150 V. The amount of DNA in samples was quantified using a Gel Doc 2000 image analysis system and a Quantity One program (BioRad).

Replicate Limited Dilution PCR

In this work, the functional genes that were studied were *nahAc* for naphthalene dioxygenase [21], *alkB* for alkane hydroxylase [22], and *xylE* for catechol-2,3-dioxygenase [23]. *NahAc* gene was amplified using a degenerate primer pair Ac114F and Ac956R (product of the size 482 bp), which amplifies nah-type genes (*nah*, *ndo*, *pah*, *dox*) [24]. *AlkB* was amplified using a primer pair alkB703for and alkB1572rev (product size 869 bp) designed for the OCT plasmid in *P. oleovorans* ATCC 29347 [8]. For the amplification of *xylE*, a primer pair xylEbf and xylEbr was used [25]. The sequences of the primers are given in Table 2. The primers are targeting genes known from cultured organisms, and they will not reveal unknown or more distantly related genes with the same function.

The abundance of genes was determined using a replicate limited dilution (RLD) polymerase chain reaction (PCR) technique [26]. From each soil sample, two replicate DNA samples were extracted, and from both of these, replicate 3× dilution series were prepared. PCR was run on all four dilution series according to [7]. PCR amplification was carried out in 0.2-ml thin-walled PCR tubes. The final volume of 50 ml contained 1 mM of each primer, 200 mM of each deoxyribonucleoside triphosphate, 0.2 U of *Taq* DNA polymerase (MBI Fermentas, 1 U/ml), 75 mM Tris–HCl, 20 mM $(NH_4)_2SO_4$, 0.01% Tween 20, 1.5 mM $MgCl_2$, and 1% nuclease-free bovine serum albumin (BioLabs). The PCR was

Table 2 Oligonucleotide sequences of primers used for the enumeration of petroleum hydrocarbon degradation genes in soil samples by RLD-PCR.

Target gene	Target compound degradation	Primer name	Sequence	Product size (bp)	Reference
alkB	Alkane degradation	alkB703for	5′-TGGCCGGCTACTCCGATGATC GGAATCTGG-3′	869	[8]
		alkB1572rev	5′-CGCGTGGTGATCCGAGTGCCG CTGAAGGTG-3′		
nahAc	Naphthalene degradation	Ac114F	5′-CTGGC(T/A)(T/A)TT(T/C)CTCAC(T/C) CAT-3′	482	[24]
		Ac596R	5′-C(G/A)GGTG(C/T)CTTCCAGTTG-3′		
xylE	Toluene degradation	xylEbf (482)	5′-AGGTATGGCGGCTGTGCGTTTC-3′	469	[25]
		xylEbr (950)	5′-TTCGTTGAGAATGCGGTCGTGG-3′		

run in a thermal cycler (MJ Reserach PTC-100 or PTC-200). With primer pair Ac114F and Ac596R, a touchdown PCR program was run according to [24]. With primer pair alkB703for and alkB1572rev, PCR was run with 45 cycles of 1 min at 95 °C, 1 min at 60 °C, 1 min at 72 °C, and a final extension of 10 min at 72 °C. With primer pair xylEbf and xylEbr, PCR was run with 35 cycles of 1 min at 95 °C, 1 min at 68 °C, 1 min at 72 °C, and a final extension of 5 min at 72 °C. The PCR products were verified in a 1.5% (wt/vol) agarose gel in 0.5% TBE by running samples for 30 min at 150 V. The number of gene copies (per milliliter of DNA extract or per gram of soil) was calculated from the highest dilution in which the PCR product was observed. The final gene copy abundances were an average of the four dilution series prepared from each soil sample.

The detection limit of the PCR was 0.2–21 pg DNA or 100–4,700 cell equivalents assuming a bacterial DNA content of 5 fg per cell. The PCR inhibition in nondiluted soil DNA was checked by spiking the PCR samples with 1 ml of pure culture DNA (20–60 pg/μl).

Real-time PCR

To compare the RLD-PCR method with real-time PCR for gene abundance quantification, we selected 11 soil samples from sites 1 to 5. These samples were analyzed for the abundance of *nahAc* and *xylE* using the same primers as for RLD-PCR. The PCR product for *alkB* using the RLD-PCR primers was too long to be assayed by real-time PCR. From the DNA extract of pure culture cells of the two control strains *P. putida* G7 and *P. putida* mt-2, the *nahAc* and *xylE* gene fragments were amplified by PCR as described above, and the concentration of the DNA fragment was determined by a spectrophotometer (U-2000, Hitachi, Tokyo, Japan) at 260 nm. The number of the target copies was calculated based on the concentration and molecular weight of the DNA fragment. From the purified PCR product, separate dilution series were prepared, which contained 400 to 4×10^8 *nahAc* target copies per microliter or 30 to 3×10^8 *xylE* target copies per microliter. These dilution series were used to make standard curves for the quantification of the *nahAc* and *xylE* genes by the real-time quantitative PCR. The standards together with 3-, 9-, and 27-fold dilutions of the DNA extracts were run in triplicates on a thermal cycler (Chromo4 Four-color Real-time PCR System, MJ Research) to quantify the number of the *nahAc* and *xylE* gene copies in the 11 selected soil samples. The total volume of the real-time PCR reaction was 20 μl and contained 10 μl of the 2× Quantitect SYBR Green mastermix (Qiagen GmbH, Hilden, Germany), 2.5 μM of each of the respective forward and reverse primers, and 1 μl of the template DNA. The temperature program of the real-time PCR run for *nahAc* was 95 °C, 15 min (first cycle only), 94 °C, 15 s, 52 °C, 30 s, and 72 °C, 30 s, altogether 40 cycles with a final extension at 72 °C for 6 min. The temperature program of the real-time PCR run for *xylE* was 95 °C, 15 min (first cycle only), 94 °C, 15 s, 68 °C, 30 s, and 72 °C, 30 s, altogether 40 cycles with a final extension at 72 °C for 5 min.

Enumeration of Bacteria

Microorganisms were extracted from 10 g of soil with a solution containing 90 ml of 0.9% NaCl, 1 ml of 10.4% $Na_5P_3O_{10}$, and 0.1 ml of 2% Tween 80. The soil slurry was homogenized for 2 min in a homogenizer. A 10-ml subsample was taken from the slurry and fixed with 0.2 ml of 37% formaldehyde. For microscopy, first, 0.5–1.5 ml of the fixed sample was filtered through a 5-μm pore size filter, which was then rinsed with an equal volume of 1 M KCl. Bacteria from the filtered sample were harvested onto black

nucleopore membranes (pore size 0.2 μm), which were then stained with 4′,6-diamidino-2-phenylindole and prepared for epifluorescence microscopy as described in [27].

Statistical Analysis

The data were analyzed by correlation analysis using SPSS.

Results

Petroleum Hydrocarbon Concentrations, Organic Matter Content, and Oxygen Conditions at the Sites

The study sites differed with respect to the contaminating petroleum products, soil type, organic matter content, the contribution of the petroleum hydrocarbons to the total organic matter, and soil gas composition (Table 3). At sites 1, 3, and 4, anoxic conditions prevailed below a depth of 1.0–1.5 m depth in the most contaminated areas. By contrast, in the subsurface of site 2, oxygen was found both in shallow and deeper layers (Table 1). Organic matter content was highest at site 3 (on average 48% of DW), and the contribution of PHC to that content was minor (on average 1%). At site 5, organic matter content was on average 19% with oil contributing on average 16%. At the other sites, organic matter contents were considerably lower (average=2.7%, range=0.4–9.0% of DW) and the proportion of PHC ranged from 0.2% to 100% (average=22%) at these sites (Table 3).

Respiration and Mineralization

The rates of the aerobic respiration and mineralization are summarized in Table 3. The highest respiration and ^{14}C-petroleum hydrocarbon mineralization rates were found in the surface soil of site 5. Compared with the other sites, mineral oil concentration and in situ temperature (ambient temperature +20±5 °C) were also the highest at this site. By contrast, at sites 1–4, where the contamination was located deeper in the subsurface and where in situ temperatures were lower (+8 °C), the highest aerobic respiration activity was found in the samples from site 3. These results reflect the fact that at site 3, the organic matter content is higher than at sites 1, 2, and 4 (Table 3).

The ^{14}C-petroleum hydrocarbon mineralization rates varied among the sites and between the three compounds we studied: The highest naphthalene mineralization activity was found at site 1, the highest toluene mineralization rate at BTEX-polluted site 4, and the highest octacosane mineralization rate at the heating oil-polluted site 2 (Table 3). The correlation between the activity parameters, total petroleum hydrocarbon concentration, organic matter content, and bacterial cell numbers is shown in Table 4.

Bacterial Cell and Gene Abundances

The highest average bacterial cell numbers were recorded at site 3. The number of *nahAc*, *xylE*, and *alkB* genes detected was 3×10^1–9×10^4, 2×10^1–2×10^6, and 5×10^1–3×10^4 copies per gram of soil (DW), respectively. The highest gene abundances were found in the samples from site 5, where also mineral oil concentration and aerobic respiration were the highest (Table 3). Of the subsurface samples from the other sites, the highest abundances

Table 3 Average (range in parenthesis) of the measured parameters in soil samples at the study sites 1–5.

Site	PHC C_{10}–C_{40}, average (range) (mg kg^{-1} DW)	BTEX or volatile pHs (mg kg^{-1} DW)	Organic matter, average (range) (% of DW)	TPH of organic matter, average (range) (%) (*w*/*w*)	Respiration rate, average (range) (μg C g^{-1} DW h^{-1})	Bacteria average (range) (10^6 cells g^{-1} DW)	Gene abundances by RLD-PCR			^{14}C-mineralisation rate		
							NahAc average (range) (10^3 gene copies g^{-1} DW)	*XylE* average (range) (10^3 gene copies g^{-1} DW)	*AlkB* average (range) (10^3 gene copies g^{-1} DW)	Naphthalene average (day^{-1})	Toluene average (day^{-1})	Octacosane average (day^{-1})
1, *n*=27	4,100 (<50–17,000)	BTEX<6.2 per each compound	2.9 (0.4–9.0)	14 (1 – 68)	0.27 (0.02–0.81)	78 (20–210)	6.8 (0.07–87)	0.08 (0.02–0.14)	0.01 (<0.05–0.14)	0.0030	0.0005	0.0012
2, *n*=7	1,300 (50–5,400)	not determined	0.6 (0.4–0.9)	45 (1 – 100)	0.02 (<0.01–0.04)	35 (18–55)	0.30 (0.03–0.80)	0.13 (<0.11–0.13)	0.94 (0.05–2.7)	0.00003	0.0007	0.0015
3, *n*=5	3,000 (1,700–5,200)	C_6–C_{10} PHs 100–1,000	48 (39–58)	1 (0.8 - 1.2)	1.4 (0.81–1.7)	670 (380–1300)	<0.20	0.20 (<0.15–0.20)	0.31 (<0.15–0.31)	0.0023	0.0021	0.0003
4, *n*=5	3,300 (50–12,000)	BTEX 0.6–3,000	5.1 (3.5–7.3)	9 (1 – 27)	0.74 (0.07–2.23)	410 (180–810)	<0.15	0.13 (0.08–0.16)	<0.15	0.0020	0.0031	0.0008
5, *n*=3	21,000 (10,000–42,000)	Not determined	19 (16–21)	12 (5 – 27)	5.9 (3.1–8.7)	280 (260–300)	14 (1–40)	624 (10–1,800)	34 (<0.20–34)	0.0776	0.009	0.048
Clean, *n*=2	<50	Not determined	0.5	0	<0.05	13	<0.12	<0.12	<0.05	0.0003	0.00003	0.0001

Table 4 Correlation between the measured parameters.

	Naphthalene mineralization rate	Toluene mineralization rate	Octacosane mineralization rate	Mineral oil concentration	Organic matter content	Respiration rate	Bacterial cell number
NahAc gene abundance RLD-PCR	+	+		+			+
XylE gene abundance RLD-PCR		+	+		+		
AlkB gene abundance RLD-PCR							
Naphthalene mineralization rate		+		+	+	+	+
Toluene mineralization rate				+	+		+
Octacosane mineralization rate					+		
Mineral oil concentration					+		
Organic matter content						+	+
Respiration rate							
Bacterial cell number							

Only those samples from sites 1–5 that were judged aerobic based on soil gas measurements are included. Significant correlation is indicated with +.

were found at site 1 where the *nahAc* gene was the most abundant one. In the clean soil (from site 1), none of the functional genes was detected, and the mineralization rate of the ^{14}C-labeled compounds was at least one order of magnitude slower compared to the contaminated sites.

When statistically analyzing the data from all the sites, the *xylE* gene abundance correlated with the ^{14}C-toluene mineralization rates, but *nahAc* and *alkB* gene abundances did not correlate with the ^{14}C-napthalene and ^{14}C-octacosane mineralization rates, respectively. When analyzing only those samples that were assumed to be aerobic in situ, *xylE* and *nahAc* gene abundances correlated with the ^{14}C-toluene and naphthalene mineralization activity, respectively, as well as with some other parameters measured (Table 4). *AlkB* gene abundance did not correlate with any of the parameters measured.

When statistically analyzing all gene numbers and all mineralization rate constants analyzed from the samples, a significant correlation (r=0.488, p=0.01) was found (Fig. 2). When the mineralization potential was high, the gene numbers were high as well. It is

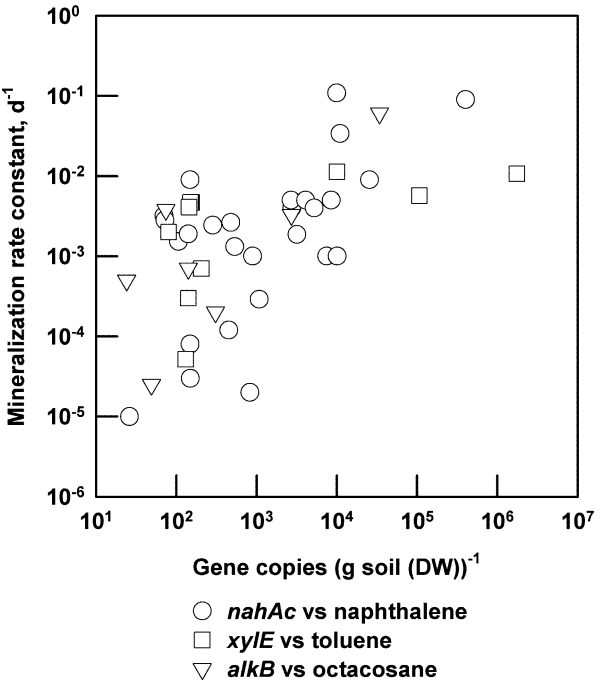

Fig. 2 Relation between gene abundances measured by RLD-PCR and mineralization rates in the samples from sites 1–5

interesting to note that high numbers of genes were found even when the mineralization potential was low.

Comparison of RLD-PCR and Real-time PCR

For the selected 12 samples, the gene abundances of the *nahAc* gene were on average 1,200 times higher, and the *xylE* gene abundances were 300 times higher than the RLD-PCR result (Table 5). Furthermore, genes that could not be detected with the RLD-PCR could be observed using real-time PCR. At the gasoline-contaminated site 4, the RLD-PCR method revealed only the *xylE* gene, but with real-time PCR, also the *nahAc* could be detected and quantified. When comparing the gene abundances of the two methods in a statistical test (Person correlation), it was found that the abundances were correlated. For the *nahAc* gene, the significance level was 0.05, and for *xylE*, it was 0.01.

Discussion

We proved effective mineralization of the three tested compounds naphthalene, octacosane and xylene at all five sites. This demonstrates that the degradation potential for different types of petroleum hydrocarbons is widely distributed at northern sites undergoing natural attenuation. Organic matter content and the contribution of PHC to that content varied within and among the sites. As expected, the highest rates of respiration and mineralization of ^{14}C-substrates were measured from the sites with the highest organic matter contents (sites 3 and 5; Table 3). At site 5, PHC concentrations were high and made a significant part of the total organic matter. Subsequently, both the gene abundances and the mineralization

Table 5 Comparison of RLD-PCR and real-time PCR for *nahAc* and *xylE* quantification for 12 selected soil samples.

Site	Sample	*nahAc*			*xylE*		
		qPCR	RLD-PCR	qPCR/RLD-PCR	qPCR	RLD-PCR	qPCR/RLD-PCR
		Gene copies g^{-1} DW			Gene copies g^{-1} DW		
1 (Trollberget)	G18, 0.8–1.3 m	60,900	100	575	400	<110	
	G18, 1.3–1.8 m	287,500	1,100	268	100	<110	
	G18, 1.8–2.3 m	885,100	3,200	278	30	<140	
2 (Etna)	P51, 3.0–3.5 m	473,000	800	573	14,100	<150	
	P51, 4.5–5.0 m	120,200	30	4623	151,600	130	1,157
3 (Vuosaari)	P2, 1.3–1.8 m	184,800	<150		5,400	<150	
	P3, 1.3–1.8 m	32,100	<210		2,700	210	13
4 (Lempäälä)	P7, 3.0–3.5 m	25,400	<160		12,400	160	80
	P9, 1.8–2.3 m	4,300	<150		34,500	150	230
5 (Sköldvik)	0.0–0.3 m	1,385,000	1,000	1394	Not determined	1,755,000	
	0.0–0.3 m	561,700	1,100	512	5,677,300	106,600	53

rates were high at this site. By contrast, at site 3, PHC contributed only by 1% to the organic matter content, although mineral oil concentrations up to 5,200 mg kg^{-1} DW were measured. At this site, the presence of organic matter other than PHC was reflected in the abundances of the three genes studied: They were detectable in low numbers and only in two samples out of five. By contrast, the average rates of aerobic naphthalene and toluene mineralization were relatively high at site 3 in comparison with the other sites. The organic matter content at site 2 was only 0.6%, and the contribution of oil to that content was relatively high (on average 36%). Moreover, the aerobic respiration rate at this site was 70 times slower than at site 3. Nonetheless, the heating oil contamination seemed to support *nahAc*-, *xylE*-, and *alkB*-carrying populations.

We found a correlation (r=0.488, p<0.05) between gene abundance and the aerobic mineralization potential when the whole data set was analyzed (Fig. 2). This suggests that functional gene abundances do reflect the degradation potential in soils, which is encouraging for the use of gene abundances in the monitoring of bioremediation. Within the individual sites or gene abundances, the observed trends were less obvious. For instance, *alkB* abundance did not correlate with the octacosane mineralization rate. However, at the heating oil-contaminated site 2, aliphatic hydrocarbons from C_{10} to C_{24} were major contaminants, and *alkB* was the most abundant gene at this site as determined by RLD-PCR. Similarly, of the three genes studied, only the *xylE* gene was found by RLD-PCR at the BTEX-contaminated site 4, although the gene abundances recorded at this site did not reflect the BTEX concentrations measured from the samples. Possible explanations to these observations can be suggested. For instance, the *alkB* gene is known to be involved in the degradation of medium-chain-length alkanes (C_6–C_{12}) [28], while octacosane has 28 carbon atoms, and the ^{14}C-label was located at the 14,15-position in the substrate we used. Consequently, both parameters may reveal the degradation potential of aliphatic petroleum hydrocarbons, although they do not target exactly the same compounds. At the BTEX-contaminated site, high contaminant concentrations were found. It has been observed in another study [16] that the abundance of the *xylE* gene was a sensitive indicator of contamination only at low levels of gasoline contamination. Similar indications were found in our work: The *xylE* copy numbers were in the same order of magnitude in highly and

slightly contaminated samples at site 4 (Table 5). The lack of correlation may also be explained by the possibility that other unknown degradation genes, e.g., from not-yet cultured microorganisms, were present in the samples.

In general, *nahAc* was the most abundant of the three genes that we studied in our samples. At sites 3 and 4, however, *nahAc* genes were undetectable by RLD-PCR but detectable in low amounts by real-time PCR. Clear differences in the mineralization and gene abundance patterns were also found between the clean and the contaminated samples. In the contaminated samples, low mineralization activities were occasionally measured for an individual compound (Table 3). In the clean sample, all mineralization rates were low (Table 3). The same applies to the gene abundances, which all were below the limit of detection in the clean soil as determined by RLD-PCR. By contrast, in the contaminated soils, even low mineralization potential (10^{-5}–10^{-4} day^{-1}) was accompanied by degradation gene abundances of 10–1,000 copies g^{-1} soil as determined by RLD-PCR. This indicates that even low concentrations of contaminants may favor or enrich microbial populations that can divert energy from their degradation. The gene abundances and mineralization rates were not in any systematical way correlated with respiration rates and bacterial cell numbers, which are both measures of general microbial activity. This indicates that mineralization rates and gene abundances are more specific tools to assess the biodegradation of specific compounds than general microbial activity measurements, which are often used to assess bioremediation efficacy.

The comparison of the RLD-PCR and real-time PCR for 12 (out of 48 samples) selected samples for the genes *nahAc* and *xylE* showed that the real-time PCR method was far more sensitive. This means that the values reported here for gene abundances by RLD-PCR are underestimated. The fact that the gene abundances measured by the two methods correlated with each other indicates that the conclusions drawn on the correlation between the other parameters measured in this study are valid. For future studies, real-time PCR is recommended. Real-time PCR, however, requires a specific piece of equipment with the continuous detection of PCR products in contrast to RDL-PCR, which can be performed with a basic PCR thermal cycler. Furthermore, there are differences in the principles for the primer design. Real-time PCR using SYBR cannot detect longer PCR products. The primer design and the validation of the targeted functional genes from different species have to be optimized in each application. In another study [29], real-time PCR was compared to hybridization with gene probes in a microarray, and they also found that real-time PCR was more sensitive and that the gene abundances obtained by those two methods were correlated.

Anoxic conditions often prevail at oil-contaminated sites undergoing natural attenuation [5, 30]. This was, for instance, the case at sites 1, 3, and 4 in this study. Moreover, anaerobic mineral oil degradation has been found in the samples from sites 1 [19] and 3 (unpublished data). However, the primers that we used here assessed only genes that are part of the aerobic degradation pathways thus excluding the genetic potential for the anaerobic degradation potential. Nonetheless, aerobic degradation genes were present not only in the oxic but also the anoxic zones of the studied sites. Thus, the potential for aerobic degradation seems to persist in contaminated areas after the onset of anaerobic conditions and supports the degradation potential under changing environmental conditions. The fact that the soils studied in this work were mostly taken from the unsaturated zone has relevance for the observations made. The unsaturated horizons are much more susceptible to variations in oxygen conditions compared to water-saturated soils [1, 31]. Consequently, aerobic microorganisms may survive for long periods and revive their activity when suitable conditions again are prevailing. One future challenge is to apply real-time PCR

methods to detect and quantify the genes that are involved in anaerobic petroleum hydrocarbon degradation pathways in anoxic zones of contamination. This is because several degradation pathways and physiological types of microorganisms are involved in the anaerobic degradation of petroleum hydrocarbons [32].

In this study, we successfully assessed the natural aerobic petroleum hydrocarbon biodegradation potential at five sites with different types of oil contamination. The mineralization tests using three model compounds representing different types of oil contamination proved the actual complete mineralization of these compounds in the soil samples. We found that the abundances of genes encoding for the corresponding compound degradation enzymes were correlated with the mineralization rates in aerobic samples for two of these compounds. The existence of other genes than *alkB* that encode for alkane degradation may explain this, and different genes might need to be assessed for alkane degradation. In conclusion, functional gene abundances seemed to reflect the type of the contamination. With optimized assays, the gene abundances can be used to assess bioremediation efficacy. Since the DNA-based PCR methods are faster than long laboratory incubation for mineralization assays, they provide an additional efficient tool in the monitoring of bioremediation efficiency.

Acknowledgments We are grateful to Marina Heino of the City of Hanko, Katarina Leminen of the City of Helsinki, Henrik Westerholm of Neste Oil, Kati Vaajasaari of the Pirkanmaa Regional Environmental Center, and Jaakko Heinolainen of Uusimaa Regional Environmental Center for the cooperation with the sites. We also thank Anna-Mari Suortti for her input with the petroleum hydrocarbon analyses. The technical personnel at SYKE laboratory is acknowledged for excellent technical assistance in the field and in the laboratory. This work was supported by grants from the Neste Oy Foundation and from the Academy of Finland (63527 and 73513).

References

1. Jørgensen, K. S. (2007). *Advances in Applied Microbiology*, *61*, 285–305.
2. Madsen, E. L. (1991). *Environmental Science & Technology*, *25*, 1663–1672.
3. Morasch, B., Richnow, H. H., Schink, B., Vieth, A., & Meckenstock, R. U. (2002). *Applied and Environmental Microbiology*, *68*, 5191–5194.
4. Stapleton, R. D., Ripp, S., Jimenez, L., Cheol-Koh, S., Fleming, J. T., Gregory, I. R., et al. (1998). *Journal of Microbiological Methods*, *32*, 165–178.
5. Bekins, B. A., Godsy, E. M., & Warren, E. (1999). *Microbial Ecology*, *37*, 263–275.
6. Ogram, A. (2000). *Soil Biology & Biochemistry*, *32*, 1499–1504.
7. Chandler, D. P., & Brockman, F. J. (1996). *Applied Biochemistry and Biotechnology*, *57/58*, 971–982.
8. Whyte, L. G., Greer, C. W., & Inniss, W. E. (1996). *Canadian Journal of Microbiology*, *42*, 99–106.
9. Siciliano, S. D., Germida, J. J., Banks, K., & Greer, C. W. (2003). *Applied and Environmental Microbiology*, *69*, 483–489.
10. Fleming, J. T., Sanseverino, J., & Sayler, G. S. (1993). *Environmental Science & Technology*, *34*, 1068–1074.
11. Sanseverino, J., Applegate, B. M., King, J. M. H., & Sayler, G. S. (1993). *Applied and environmental microbiology*, *59*, 1931–1937.
12. Tuomi, P. M., Salminen, J. M., & Jørgensen, K. S. (2004). *FEMS Microbiology, Ecology*, *51*, 99–107.
13. Piskonen, R., Nyyssönen, M., Rajamäki, T., & Itävaara, M. (2005). *Biodegradation*, *16*, 127–134.
14. Park, J. W., & Crowley, D. E. (2006). *Applied Microbiology and Biotechnology*, *72*, 1322–1329.
15. USEPA (1999). http://www.epa.gov/oilspill/oiltypes.htm
16. Guo, C., Sun, W., Harsh, J. B., & Ogram, A. (1997). *Microbial Ecology*, *34*, 178–187.
17. Stapleton, R. D., & Sayler, G. S. (1998). *Microbial Ecology*, *36*, 349–361.
18. Laurie, A. D., & Lloyd-Jones, G. (2000). *Applied and Environmental Microbiology*, *66*, 1814–1817.
19. Salminen, J. M., Tuomi, P. M., Suortti, A. -M., & Jørgensen, K. S. (2004). *Biodegradation*, *15*, 29–39.
20. ISO (2004). ISO 16703:2004.
21. Simon, M. J., Osslund, T. D., Saunders, R., Ensley, B. D., Suggs, S., Hartcourt, A., et al. (1993). *Gene*, *127*, 31–37.

22. Kok, M., Oldenhuis, R., van der Linden, M. P. G., Raatjes, P., Kingma, J., van Lelyveld, P. H., et al. (1989). *Journal of Biological Chemistry*, *264*, 5435–5441.
23. Suzuki, M., Hayakawa, T., Shaw, J. P., Rekik, M., & Harayama, S. (1991). *Journal of Bacteriology*, *173*, 1690–1695.
24. Wilson, M. S., Bakermans, C., & Madsen, E. L. (1999). *Applied and Environmental Microbiology*, *65*, 80–87.
25. Jussila, M. M., Zhao, J., Suominen, L., & Lindström, K. (2007). *Environmental Pollution*, *146*, 510–524.
26. Chandler, D. P. (1998). *Journal of Industrial Microbiology & Biotechnology*, *21*, 128–140.
27. Porter, K. G., & Feig, Y. S. (1980). *Limnology and Oceanography*, *25*, 943–948.
28. van Beilen, J. B., Wubbolts, M. G., & Witholt, B. (1994). *Biodegradation*, *10*, 161–174.
29. Rhee, S. -K., Liu, X., Wu, L., Chong, S. C., Wan, X., & Zhou, J. (2004). *Applied and Environmental Microbiology*, *70*, 4303–4317.
30. Bolliger, C., Schoenholzer, F., Schroth, M. H., Hahn, D., Bernasconi, S. M., & Zeyer, J. (2000). *Bioremediation*, *4*, 359–371.
31. Huesemann, M. H., & Truex, M. J. (1996). *Journal of Hazardous Materials*, *51*, 93–113.
32. Heider, J., Spormann, A. M., Beller, H. R., & Widdel, F. (1999). *FEMS Microbiology Reviews*, *22*, 459–473.