Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 322 (2004) 133–146
BBRC
www.elsevier.com/locate/ybbrc

Molecular characterization of a phenanthrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1

Robin L. Stingley, Ashraf A. Khan*, Carl E. Cerniglia

National Center for Toxicological Research, US FDA, 3900 NCTR Road, Jefferson, AR 72079, USA

Received 1 July 2004
Available online 7 August 2004

Abstract

Mycobacterium vanbaalenii PYR-1 is capable of degrading a number of polycyclic aromatic hydrocarbons (PAHs) to ring cleavage metabolites via multiple pathways. Genes for the large and small subunits of a pyrene dioxygenase, *nidA* and *nidB*, respectively, were previously identified in *M. vanbaalenii* PYR-1 [Appl. Environ. Microbiol. 67 (2001) 3577]. A library of the *M. vanbaalenii* PYR-1 genome was constructed in a fosmid vector to identify additional genes involved in PAH degradation. Twelve fosmid clones containing *nidA* were identified by Southern hybridization. Sequence analysis of one *nidA*-positive clone, pFOS608, revealed a number of additional genes involved in PAH degradation. At this locus, one putative operon contained genes involved in phthalate degradation, and another contained genes encoding a putative ABC transporter(s). A number of the genes found in this region are homologous to those involved in phenanthrene degradation via the phthalic acid pathway. The majority of phenanthrene degradation genes were located between putative transposase genes. In *Escherichia coli*, pFOS608 converted phenanthrene into phenanthrene *cis*-3,4-dihydrodiol, and converted 1-hydroxy-2-naphthoic acid into 2′-carboxybenzalpyruvate, 2-carboxybenzaldehyde, and phthalic acid. A subclone containing *nidA* and *nidB* converted phenanthrene into phenanthrene *cis*-3,4-dihydrodiol, suggesting that the NidAB dioxygenase is responsible for an initial attack on phenanthrene. This study is the first to identify genes responsible for the degradation of phenanthrene via the phthalic acid pathway in *Mycobacterium* species.
Published by Elsevier Inc.

Keywords: *Mycobacterium vanbaalenii*; PAH degradation; Phenanthrene degradation

Mycobacterium vanbaalenii PYR-1 was isolated from oil-contaminated sediment based upon its ability to mineralize pyrene [8]. This organism is capable of degrading several other aromatic hydrocarbons, including high molecular weight polycyclic aromatic hydrocarbons (PAHs). *M. vanbaalenii* PYR-1 biodegradation pathways for biphenyl, naphthalene, anthracene, phenanthrene, fluoranthene, 1-nitropyrene, benzo[*a*]pyrene, and 7,12-dimethylbenz[*a*]anthracene have been proposed based on the metabolites observed from each [9,13,18,20–23,26].

Mycobacterium vanbaalenii PYR-1 utilizes multiple pathways to degrade a number of PAHs, which suggests that the organism uses multiple dioxygenases and monooxygenases [13,21,23]. Our laboratory previously identified and characterized two *M. vanbaalenii* PYR-1 genes that encode the large and small subunits of a dioxygenase, *nidA* and *nidB*, respectively [14]. The NidAB dioxygenase is responsible for initiation of the pyrene degradation pathway that proceeds through pyrene-4,5-dihydrodiol [14]. This pathway may join the phenanthrene degradation pathway represented in Fig. 1 at 3,4-dihydroxyphenanthrene [16].

Two additional PAH-degradation genes were identified within the 5.3 kb fragment containing *nidBA*, an aldehyde dehydrogenase, *nidD*, and the 5′ portion of another putative dehydrogenase [14]. A number of genes encoding enzymes involved in PAH degradation are grouped together in a similar manner in other

* Corresponding author. Fax: +1 870 543 7307.
E-mail address: ashraf@nctr.fda.gov (A.A. Khan).

0006-291X/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.bbrc.2004.07.089

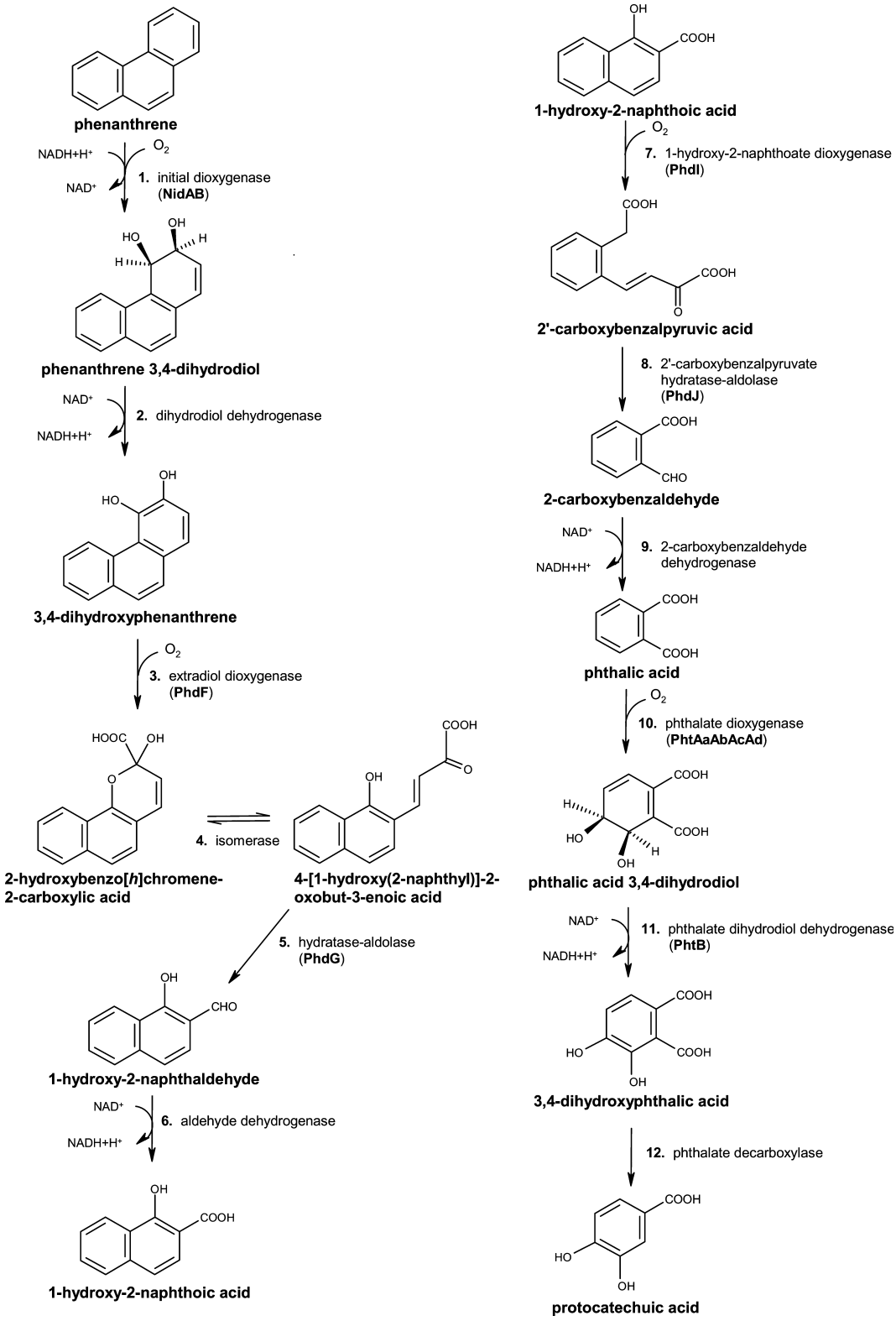

Fig. 1. Proposed phenanthrene degradation pathway in *M. vanbaalenii* PYR-1. Each enzymatic step is numbered and each enzyme is named. Proposed *M. vanbaalenii* PYR-1 enzymes are indicated in parentheses.

bacterial species, including *Nocardioides* sp. KP7, *Burkholderia* sp. strain RP007, and *Sphingomonas* spp. [2,11,15,17,27–29].

Mycobacterium vanbaalenii PYR-1 utilizes three degradation routes to metabolize phenanthrene [21]. Initial attack in the K region, presumably by a dioxygenase

and a monooxygenase, leads to the formation of *cis*- and *trans*-9,10-dihydrodiol phenanthrene, respectively [21]. Another dioxygenase attack at the C-3 and C-4 positions forms *cis*-3,4-dihydrodiol phenanthrene. The presence of the metabolites 1-hydroxy-2-naphthoic acid and phthalic acid indicates that *M. vanbaalenii* PYR-1 likely metabolizes phenanthrene through the phthalate degradation pathway, which has been reported in other bacterial species [7,12,29] (Fig. 1).

The present study focuses on molecular cloning, sequencing, and functional characterization of additional genes in the *nidDBA* region of the *M. vanbaalenii* PYR-1 genome. A number of genes encoding putative enzymes involved in the degradation of phenanthrene were identified within a 37 kb region of the *M. vanbaalenii* PYR-1 genome.

Materials and methods

Genomic DNA isolation. *Mycobacterium vanbaalenii* PYR-1 was grown on Middlebrook 7H10 agar at 25 °C for 4 days. Genomic DNA was isolated according to the protocol for gram-positive bacteria with the Qiagen DNeasy Kit (Qiagen, Valencia, CA).

Cloning, sequencing, and sequence analyses. A genomic library was prepared in pCC1FOS according to the manufacturer's protocol (Epicentre, Madison, WI). Briefly, *M. vanbaalenii* PYR-1 genomic DNA was randomly sheared to approximately 40 kb fragments and the ends were blunted and ligated into pCC1FOS. The ligated DNA was packaged with MaxPlax Lambda Packaging Extracts to form the fosmid library. The library was transduced into EPI300 *E. coli* and spread on Luria–Bertani (LB) agar plates containing 12.5 μg/ml chloramphenicol (USB, Cleveland, OH).

The clones were transferred to nylon membranes (GE Osmonics Labstore, Minnetonka, MN) and probed for *nidA* as described in the hybridization section. Fosmid DNA was isolated using the Qiagen Plasmid Miniprep Kit (Qiagen). *Bam*HI restriction fragments of fosmid clone # 608 (pFOS608) were subcloned into pGEM-11z f(+) (Promega, Madison, WI) and introduced into DH5α *E. coli*. For sequencing, subclone DNA was isolated with the Qiagen Miniprep Spin Kit (Qiagen). Subclones pRS9, pRS12, pRS14, pRS19, pRS22, pRS35, pRS42, pRS50, and pRS59 were sequenced with walking primers. The region containing subclone pRS68 was further divided into pRS68-A, pRS68-B, and pRS68-C to facilitate sequencing (Fig. 3). Subclone pRS68-A was generated by *Sac*I digest of pRS68 and subsequent religation with T4 DNA ligase (Promega). Subclone pRS68-C was generated by cloning a 2.5 kb *Eco*RI fragment from pFOS608 into pGEM-11z. To generate pRS68-B, primers PRS68-R10 and *Eco*RI-F-R1 (Table 1) were used in high fidelity PCR to amplify the region between pRS68-A and pRS68-C. Approximately 600 bp at either end of subclone pRS53, a previously sequenced 5.3 kb *Bam*HI fragment containing the *nidBA* region [14], were sequenced to confirm that the region was contained in pFOS608. All junctions between subclones were amplified by PCR using clone-specific primers and PCR SuperMix High Fidelity (Invitrogen), and subsequently sequenced. Each PCR consisted of initial denaturation at 94 °C for 2 min followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, and a final extension step at 72 °C for 7 min. The pRS68-B product was generated using AccuPrime *Pfx* DNA Polymerase (Invitrogen) under the same conditions, except that cycle extensions were at 68 °C for 2 min.

DNA sequencing was performed on an Applied Biosystems Model 377 DNA sequencer at the University of Arkansas for Medical Sciences. Sequences were compiled, translated, and analyzed using Lasergene software (DNASTAR, Madison, WI) and compared to similar genes and proteins using blastn, blastx, blastp, and Conserved Domain Database searches [19].

Hybridizations. For Southern hybridizations, *M. vanbaalenii* PYR-1 fosmid DNA *Bam*HI restriction digests were separated on agarose gels stained with ethidium bromide for visualization, then transferred and crosslinked to positively charged nylon membranes (Roche, Indianapolis, IN). The resulting blots were incubated at 65 °C for at least 30 min in DIG Easy Hyb (Roche) prior to addition of digoxigenin (DIG)-labeled DNA probes (denatured at 95 °C for 10 min) specific for *nidA*. Hybridization was carried out at 65 °C for 16–18 h. DIG-labeled probe was detected with alkaline phosphate-conjugated anti-DIG antibody (Roche) and the chemiluminescent substrate disodium 3-(4-methoxyspiro (1,2-dioxetane-3,2′-(5′-chloro)tricyclo[3.3.1]decan)-4-yl)phenylphosphate (CSPD, Roche).

Enzyme activity assays. Overnight cultures of pFOS608 and pFOS570 in EPI300 *E. coli* were used to inoculate 50 ml LB broth supplemented with 12.5 μg/ml chloramphenicol. The fosmids were induced to high copy number with 50 μl induction solution (Epicentre). EPI300 *E. coli* were grown under the same conditions, omitting the chloramphenicol. Phenanthrene was added to induce cultures at a final concentration of 6 μg/ml. After 24 or 7 h of growth at 37 °C, phenanthrene or 1-hydroxy-2-naphthoate, respectively, was added to the cultures at a final concentration of 0.15 mM. Neutral and acidic (pH 2–2.5) extractions were performed twice with an equal volume of ethyl acetate 24 h after addition of substrate. A 50 ml overnight culture of pRS53 in LB that had been induced with phenanthrene was centrifuged to pellet the cells which were then resuspended in MBS broth. Phenanthrene was added to a final concentration of 0.15 mM, and neutral and acidic extractions were performed after approximately 24 h at 37 °C.

Analytical chemistry procedures. Phenanthrene metabolites were separated by HPLC using a Hewlett–Packard model 1050 pump system (Hewlett–Packard, Palo Alto, CA) with a Hewlett–Packard diode array model 1040A detector at 254 nm and a 4.6- by 250-mm 5-μm C18 Inertsil ODS-3 column (MetaChem Technologies, Torrance, CA) at a flow rate of 1 ml/min. UV absorbance spectra were obtained online. The compounds were eluted using a linear gradient of 40–95% methanol/water over 40 min.

Mass spectrometry was performed by the direct injection of extracted samples dissolved in acetonitrile before or after the derivatization using a Regis reagent including Regisil + 1% TMSC [bis(trimethylsilyl)trifluoroacetamide + 1% trimethylchlorosilane] (Regis Technologies, Morton Grove, IL). For the TMS derivatization, 100 μl of 0.2 μm membrane-filtrated samples was incubated at 60 °C for 1 h with the same volume of the Regis reagent, and then maintained at room temperature overnight in the dark.

Direct exposure probe/electrospray mass spectrometry (DEP/EI-MS) was performed on a Finnigan TSQ 700 triple quadrupole mass spectrometer (ThermoFinnigan, San Jose, CA) operated in a single quadrupole analyzer mode. The ion source was maintained at 150 °C and the electron energy was 70 V. The DEP current was increased to 800 mA in a linear mode at 5 mA/s.

Gas chromatography electron ionization mass spectrometry (GC/EI-MS) was performed using a Varian 3400 Gas Chromatograph equipped with septum-equipped, temperature programmable injector (SPI), Finnigan temperature programmable transfer line, J&W DB-5ms capillary GC column (30 m × 0.25 mm i.d. × 0.25 μm film thickness), and a Finnigan TSQ 700 triple quadrupole mass spectrometer. Helium carrier gas flew at 15 psig. The injector was initially held at 70 °C for 0.1 min, and the temperature was increased up to 280 °C at 200 °C min^{-1} and held at 280 °C for 26 min. The column was held at 95 °C for initial 1 min, and the temperature was increased up to 280 °C at 10 °C min^{-1} and held at 280 °C for 10 min. The transfer line was kept at 280 °C for initial 15 min, and the temperature was increased up to 320 °C at 10 °C min^{-1} and held at 320 °C for 10 min. The mass spectrometer conditions were: electron ionization at 70 eV electron energy,

Table 1
Primers for PCR

Primer	Sequence (5′–3′)	Relative position[a]	Junction
pFOS-RP	CTCGTATGTTGTGTGGAATTGTGAGC	(pCC1FOS-reverse)	Insert site
PRS59-F1	AAGTCGGGCGGGTGTAGTT	215–233	
PRS59-R3	TCGTTCATCATCGCTCGTG	5351–5369	pRS59-14
PRS14-R1	CAGTTCGCTTTTACCCGCCAAC	5733–5754	
PRS14-F2	GCTCACAGAGGACATTCAC	6804–6822	pRS14-X-9-50
PRS50-R1	CTGTGTTCGTCGGGTTCAATTCC	8307–8329	
PRS9-R1	GTCATCGTTCTATCCAGGAC	7945–7964	pRS9-50
PRS50-R1	CTGTGTTCGTCGGGTTCAATTCC	8307–8329	
PRS50-F2	CATCACGGAGAGCGCAAC	12616–12633	pRS50-68
PRS68-R2	ATGACGCAACCCCGAAGGATC	12781–12801	
PRS68-F3	GAACCAGGCACAGGAGAGCTT	19299–19319	pRS68-53
PRS53-R2	ACTCATCGTTGTCCGCGTAG	19626–19645	
PRS53-F1	GTTTGCCTCAGCGATGTTG	24504–24522	pRS53-19
PRS19-R1	TCTGGTCCTGCGAAACGTTGCATC	24950–24955	
PRS19-F2	GAACAGAACCAGCGCACTTG	26451–26470	pRS19-12
PRS12-R1	TCGGTGCTGCGGTAAATGC	26784–26802	
PRS12-F1	ACTTTCGCGAGGAGGACTAC	27686–27705	pRS12-35
PRS35-R1	CAGTTGAGCAGTTTGAACGC	27934–27953	
PRS35-F1	CTACGGTCAATCGATCAGCTC	30956–30976	pRS35-42
PRS42-F1	GACTTGCGACATCTGGTTG	31133–31151	
PRS42-R2	CCAAGAATCCATCAAGGGTCTGG	34986–35008	pRS42-22
PRS22-R3	CGAACCAGAAGGTGATTCTG	35690–35709	
PRS22-F4	GATGTCGAAACCCTTCAG	36977–36991	Insert site
pFOS-FP	GGATGTGCTGCAAGGCGATTAAGTTGG	(pCC1FOS-forward)	
PRS68-R10	CGAATGTTGGACACGAAGTG	17684–17703	pRS68-B
*Eco*R1-F-R1	CATGCGTTCCGCATCACATTG	14416–14436	

[a] Relative position on the pFOS608 sequence (GenBank Accession No. AY365117).

150 °C ion source temperature, and Quadrupole Q1 scanned from *m/z* 50 to 450 in 0.6 s.

The liquid chromatography mass spectrometry (LC/MS) analyses were performed on a Finnigan TSQ 7000 triple quadrupole mass spectrometer with an instrument operated in a single quadrupole analyzer mode. Electrospray ionization was utilized in both positive ion and negative ion modes. Specific conditions included 275 °C capillary temperature, 70 psi sheath gas pressure, 7 auxiliary gas flow, and 4 kV spray voltage. An in-source energy of 15 eV was applied to reduce solvent-adduct formation. Chromatographic system consisted of a model 1100 degasser, quaternary pump, autosampler, and diode array detector (Agilent Technologies, Palo Alto, CA). The column was a Polaris C18-A5μ, 2 × 250 mm (Ansys-Metachem Technologies, Lake Forest, CA). The injection volume was 10 μl, and the mobile phase was delivered at 0.2 ml/min. Solvent A (acetonitrile/water/formic acid = 5:95:0.1, v/v/v) was held at 100% for 1 min before solvent B (acetonitrile/water/formic acid = 95:5:0.1,v/v/v) was linearly increased from 0% to 100% in 45 min and held at 100% for 4 min. The diode array detector was scanned from 190 to 400 nm in 2 nm steps with a 1 nm slit width.

GenBank accession number. M. vanbaalenii PYR-1 phenanthrene degradation pathway (pFOS608 sequence) AY365117.

Results

Mycobacterium vanbaalenii PYR-1 genomic library construction and screening

A genomic library of *M. vanbaalenii* PYR-1 was constructed in a fosmid vector to search for additional genes near *nidA* and *nidB* that may be involved in PAH degradation. From 1120 clones, 12 were positive for *nidA* in colony hybridization. Fosmid DNA was isolated from these clones and *Bam*HI digests of the DNA were analyzed by gel electrophoresis followed by hybridization with *nidA*. *Bam*HI sites within pCC1FOS are near the

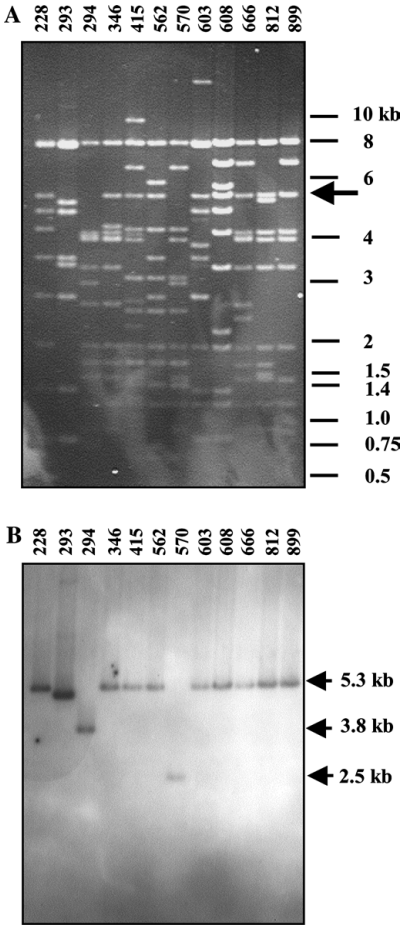

Fig. 2. *Bam*HI digested *nidA*-positive fosmids. (A) Agarose gel. Molecular weights are indicated to the right. The arrow indicates the previously characterized 5.3 kb *M. vanbaalenii* PYR-1 *Bam*HI fragment containing *nidA* [14]. (B) Southern hybridization with *nidA*-specific probe. The approximate molecular weights of *nidA*-hybridizing bands are indicated.

insert sites and produce the linear fosmid band at 8.1 kb for all fosmid clones (Fig. 2A). In addition to the fosmid band all clones, except pFOS293, −294, and −570, contain the previously characterized 5.3 kb *Bam*HI fragment (Fig. 2A). Most of the *nidA*-positive fosmids also contain a number of common bands, including those at approximately 4.2, 4.0, 3.2, 1.9, 1.6, 1.45, and 1.2 kb, suggesting that they overlap in the same region of the genome (Fig. 2A).

Southern hybridizations confirmed that the clones contained *nidA* (Fig. 2B). Eight clones contained the previously characterized 5.3 kb *Bam*HI fragment [14]. Clones pFOS293, pFOS294, and pFOS570 each had a single band that hybridized with *nidA* at approximately 5.2, 3.8, and 2.5 kb, respectively (Fig. 2B). These data indicate that inserts for these fosmid clones were likely sheared within the original 5.3 kb *Bam*HI fragment.

A single fosmid clone, pFOS608, was chosen from among the *nidA*-positive clones for sequencing. Eleven subclones (pRS9, −12, −14, −19, −22, −35, −42, −50, −53, −59, and −68) were generated using *Bam*HI fragments from pFOS608 cloned into pGEM-11z (Fig. 3). Subclones were initially sequenced using T7 and Sp6 primers from the vector, followed by walking primers in both directions from the insert sequences. Subclone junctions were identified using insert-specific primers in high-fidelity PCR, with subsequent sequencing. An uncloned *Bam*HI fragment of 267 bp, identified as "X," was discovered between subclones pRS14 and pRS9 (Fig. 3). Due to sequencing difficulties, pRS68 was further divided into pRS68−A, −B, and −C, as described in the Materials and methods section (Fig. 3).

Putative open reading frames in pFOS608

The sequence obtained from pFOS608 subclones was analyzed by blastn and blastx searches and projected protein sequences were analyzed by comparison to the Conserved Domain Database [19]. Putative open reading frames (orfs) were identified using the DNASTAR software package. Blast data were used to identify the putative genes and proteins listed in Table 2 and shown in Fig. 4. The pFOS608 insert was 37.237 kbp in length and contained 32 complete and three partial putative orfs. Among the orfs putatively involved in PAH-degradation were ten oxygenase genes, five dehydrogenase genes, two hydratase–aldolase genes, a ferredoxin, and a ferredoxin reductase. Additional genes putatively en-

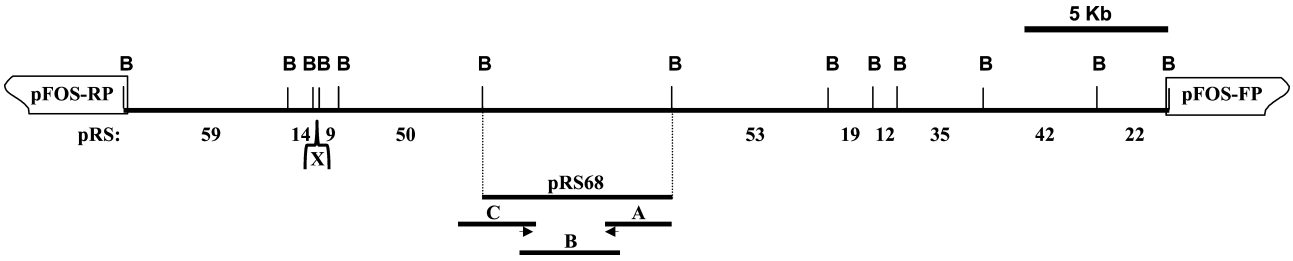

Fig. 3. Schematic representation of pFOS608. Fosmid sequence is indicated on either end of the insert by pFOS-RP (reverse primer) and pFOS-FP (forward primer). *Bam*HI sites are represented above the insert by "B." Numbers below the insert refer to pRS subclones. "X" represents a 267 bp *Bam*HI fragment that was not subcloned, but was discovered between pRS14 and pRS9 using high-fidelity PCR and subsequent sequencing to determine subclone junctions. Subclone pRS68, detailed below the insert, was further divided into pRS68-A, -B, and -C to facilitate sequencing.

Table 2
Putative *M. vanbaalenii* PYR-1 gene products from pFOS608

Putativegene	Relative position[a]	Deduced MW (a.a. residues)	Deduced *p*I	Gene product	Related proteins, GenBank Accession No. % identity–% positive[b]
tnpA	1–273	(partial, 91 residues)		Transposase	Putative ISmav2 family transposase from *Streptomyces avermitilis* MA-4680, NP_824918, 62–73% over 67 residues; putative transposase from *Mycobacterium avium* subsp. *paratuberculosis*, AAG44885, 56–71% over 64 residues
tnpB	333–1967	59,406 (544)	8.0	Transposase	Putative transposase from *M. smegmatis*, AAC38198, 39–56% over 541 residues; transposase from *Thermoanaerobacter tengcongensis*, NP_623662, 25–45% over 380 residues
tnpC	2048–2953	32,474 (301)	11.1	Transposase	Hypothetical protein of *Streptomyces coelicolor* A3,NP_624440, 37–54% over 237 residues; putative Ismav2 family transposase from *S. avermitilis* MA-4680, NP_824918, 34–49% over 320 residues
phtR	3884–4690 (complement)	28,971 (268)	8.0	*pht* operon regulator	Regulatory protein, from *Terrabacter* sp. DBF63, AB55883, 55–69% over 245 residues; putative pht operon regulator, *Arthrobacter keyseri* 12B, AF331043_19, 53–71% over 225 residues
phtAa	5153–6616	54,484 (487)	5.4	Phthalate dioxygenase large subunit	Aromatic ring hydroxylation dioxygenase E from *Rhodococcus* sp. RHA1, BAB62289, 78–88% over 356 residues; oxygenase large subunit of phthalate dioxygenase from *Terrabacter* sp. DBF63, BAC54156, 71–82% over 477 residues
phtAb	6613–7212	22,563 (199)	6.0	Phthalate dioxygenase small subunit	Phthalate dioxygenase small subunit from *A. keyseri* 12B, AF331043_15, 68–80% over 199 residues; oxygenase small subunit of phthalate dioxygenase from *Terrabacter* sp. DBF63, BAC54157, 67–79% over 191 residues
phtU	7209–7514	10,414 (101)	4.2	Unknown	Unknown protein of *Terrabacter* sp. DBF63, BAC54158, 57–71% over 91 residues; hypothetical protein from *Mycobacterium* sp. 6PY1, CAD38645, 44–63% over 61 residues
phtB	7511–8314	28,296 (267)	5.6	Dihydrodiol dehydrogenase	3,4-Dihydroxy-3,4-dihydrophthalate dehydrogenase from *Terrabacter* sp. DBF63, BAC54159, 65–76% over 263 residues; *cis*-naphthalene dihydrodiol dehydrogenase from *Rhodococcus* sp. NCIMB12038, AAD30203, 49–66% over 258 residues
phtAc	8342–8542	7200 (66)	4.4	Phthalate dioxygenase ferredoxin subunit	Phthalate dioxygenase ferredoxin from *A. keyseri* 12B, AF331043_16, 69–77% over 61 residues; ferredoxin of phthalate dioxygenase from *Terrabacter* sp. DBF63, BAC54160, 55–66% over 55 residues
phtAd	8539–9780	44,562 (413)	4.9	Phthalate dioxygenase ferredoxin reductase	Phthalate dioxygenase reductase subunit from *A. keyseri* 12B, AF331043_17, 58–73% over 403 residues; ferredoxin reductase of phthalate dioxygenase from *Terrabacter* sp. DBF63, BAC54161, 57–71% over 398 residues
phdI	10333–11418	40638 (361)	5.1	1,2-Dioxygenase	1-Hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7, BAA31235, 46–62% over 360 residues; gentisate 1,2-dioxygenase from *Bacillus halodurans*, NP_242868, 36–55% over 348 residues
phdJ	11437–12441	31,079 (286)	4.8	Hydratase–aldolase	Trans-2′-carboxybenzalpyruvate hydratase–aldolase from *Nocardioides* sp. KP7, BAA24966, 55–71% over 322 residues; Trans-*o*-hydroxybenzylidenepyruvate hydratase–aldolase from *Pseudomonas putida*, AAO64280, 45–66% over 330 residues

phdF	12460–13350	32,527 (296)	5.1	Extradiol dioxygenase	Extradiol dioxygenase from *Nocardioides* sp. KP7, BAA94706, 83–91% over 296 residues; 2,3-dihydroxybiphenyl 1,2-dioxygenase from *R. erythopolis* TA421, BAA25614, 42–60% over 296 residues
Orf14	13396–13782	13,863 (128)	4.4	Unknown (sterol carrier)	Orf131 from *Nocardioides* sp. KP7, BAA94707, 75–87% over 128 residues; Fox2 protein from *Glomus mosseae*, CAB55552, 39–54% over 64 residues
phdG	13951–14976	36,789 (341)	5.1	Hydratase–aldolase	Hydratase–aldolase from *Nocardioides* sp. KP7, BAA94711, 87–94% over 339 residues; hydratase–aldolase from *P. putida*, BAA20397, 35–54% over 329 residues
tnpD	15776–17105 (complement)			Truncated transposase (stop mutations and frameshift)	Probable transposase from *M. ulcerans*, AAC72813, 87–93% over 444 residues; hypothetical protein from *M. tuberculosis* H37Rv, NP_215435, 69–81% over 437 residues. (Values disregard mutations and frameshift)
nidB2	17405–17914	19,388 (169)	5.0	Dioxygenase small subunit	Dioxygenase small β subunit from *M. gilvum* BB1, AAN78317, 100–100% over 169 residues; naphthalene-inducible dioxygenase small subunit from *M. vanbaalenii* PYR-1, AF249302, 99–99% over 169 residues
Orf18	18055–19812	65,261 (595)	5.4	Dehydrogenase/reductase	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit from *Methanopyrus kandleri* AV19, NP_614111, 30–43% over 565 residues; putative succinate dehydrogenase flavoprotein subunit from *S. avermitilis* MA-4680, NP_824574, 27–43% over 578 residues
nidD	20295–21755	51,886 (486)	4.9	Aldehyde dehydrogenase	Aldehyde dehydrogenase from *M. vanbaalenii* PYR-1, AF249300
nidB	22313–22822	19,446 (169)	4.9	Dioxygenase small subunit	Naphthalene-inducible dioxygenase small subunit from *M. vanbaalenii* PYR-1, AF249302
nidA	22883–24250	50,141 (455)	5.3	Dioxygenase large subunit	Naphthalene-inducible dioxygenase large subunit from *M. vanbaalenii* PYR-1, AF249301
Orf22	24306–25097	27,324 (264)	4.9	Dehydrogenase/reductase	Probable oxidoreductase, short chain dehydrogenase/reductase family from *Rhizobium etli*, NP_659924, 39–56% over 250 residues; 3-oxoacyl-(acyl carrier protein) reductase from *Thermotoga maritima*, NP_229523, 38–56% over 242 residues
Orf23	25155–26180	35,889 (341)	5.2	Zinc-dependent dehydrogenase	2,3-Butanediol dehydrogenase from *P. putida*, AAB58982, 38–55% over 353 residues; hypothetical protein from *Azotobacter vinelandii*, ZP_00088750, 38–54% over 353 residues
Orf24	26195–27280	39,134 (361)	6.1	Membrane protein	VPS10 domain from *Bacillus anthracis* A2012, NP_655627, 29–50% over 226 residues; BNR repeat domain protein from *B. anthracis* strain Ames, NP_844185, 29–50% over 226 residues
Orf25	27295–28458	44,377 (387)	5.8	Aromatic oxygenase large subunit	Hypothetical protein from *Sphingomonas* sp., CAA11187, 59–73% over 385 residues; large subunit aromatic oxygenase from *Novosphingobium aromaticans*, NP_049179, 43–60% over 375 residues
Orf26	28500–28949	17,035 (149)	4.2	Aromatic oxygenase small subunit	Small subunit aromatic oxygenase from *N. aromaticans*, NP_049178, 40–53% over 153 residues; small subunit of oxygenase from *Sphingomonas* sp. P2, BAC65454, 38–53% over 153 residues
Orf27	28978–29376	14,336 (132)	5.0	Unknown (sterol carrier)	Putative protein from *Aquifex aeolicus*, NP_213060, 29–52% over 123 residues; putative protein from *A. aeolicus*, NP_213005, 28–46% over 133 residues
Orf28	29379–29897	18,912 (172)	8.3	Dioxygenase	Protocatechuate 3,4-dioxygenase β subunit from *Corynebacterium glutamicum* ATCC 13032, NP_601598, 34–50% over 144 residues; protocatechuate 3,4-dioxygenase alpha subunit from *S. coelicolor* A3 (2), NP_630772, 39–52% over 156 residues

(*continued on next page*)

Table 2 (*continued*)

Putativegene	Relative position[a]	Deduced MW (a.a. residues)	Deduced *p*I	Gene product	Related proteins, GenBank Accession No. % identity–% positive[b]
trnB1	30049–30858	28,223 (269)	8.5	Membrane protein (ABCT permease 1)	Conserved hypothetical protein from *M. tuberculosis* H37Rv, NP_336476, 61–79% over 243 residues, conserved hypothetical integral membrane protein YRBE4A from *M. bovis* subsp. *bovis* AF2122/97, NP_857170, 58–77% over 245 residues
trnB2	30855–31712	29,747 (285)	9.6	Membrane protein (ABCT permease 2)	Conserved hypothetical protein from *M. tuberculosis* H37Rv, NP_336477, 57–72% over 262 residues; hypothetical protein from *S. avermitilis* MA-4680, NP_827077, 46–66% over 259 residues
trnC1	31709–33064	47,089 (448)	7.3	Membrane protein (ABCT periplasmic 1)	mce3 from *M. tuberculosis* H37Rv, NP_216482, 57–72% over 425 residues; mce4 from *M. tuberculosis* H37Rv, NP_218016, 35–56% over 351 residues
trnC2	33122–34150	36,705 (342)	5.4	Membrane protein (ABCT periplasmic 2)	Virulence factor mce family protein from *M. tuberculosis* CDC1551, NP_336479, 65–77% over 338 residues; mce-family protein mce2B from *M. bovis* subsp. *bovis* AF2122/97, NP_854265, 46–62% over 343 residues
trnC3	34135–35475	46,416 (446)	4.8	Membrane protein (ABCT periplasmic 3)	Virulence factor mce family protein from *M. tuberculosis* CDC1551, NP_336480, 61–77% over 384 residues; hypothetical protein from *M. tuberculosis* H37Rv, NP_216484, 62–77% over 380 residues
trnC4	35472–36818	47,395 (448)	4.9	Membrane protein (ABCT periplasmic 4)	Virulence factor mce family protein from *M. tuberculosis* CDC1551, NP_336481, 59–71% over 449 residues; mce-family protein from *M. bovis* subsp. *bovis* AF2122/97, NP_854267, 40–56% over 393 residues
trnC5	36815–37237	(partial, 130 residues)		Membrane protein (ABCT periplasmic 5)	LprM from *M. tuberculosis* H37Rv, NP_216486, 64–77% over 127 residues; virulence factor mce family protein from *M. tuberculosis* CDC1551, NP_334588, 46–62% over 130 residues

[a] Relative position in bases on pFOS608 sequence, Accession No. AY365117.
[b] Percentages based on Blastx alignments.

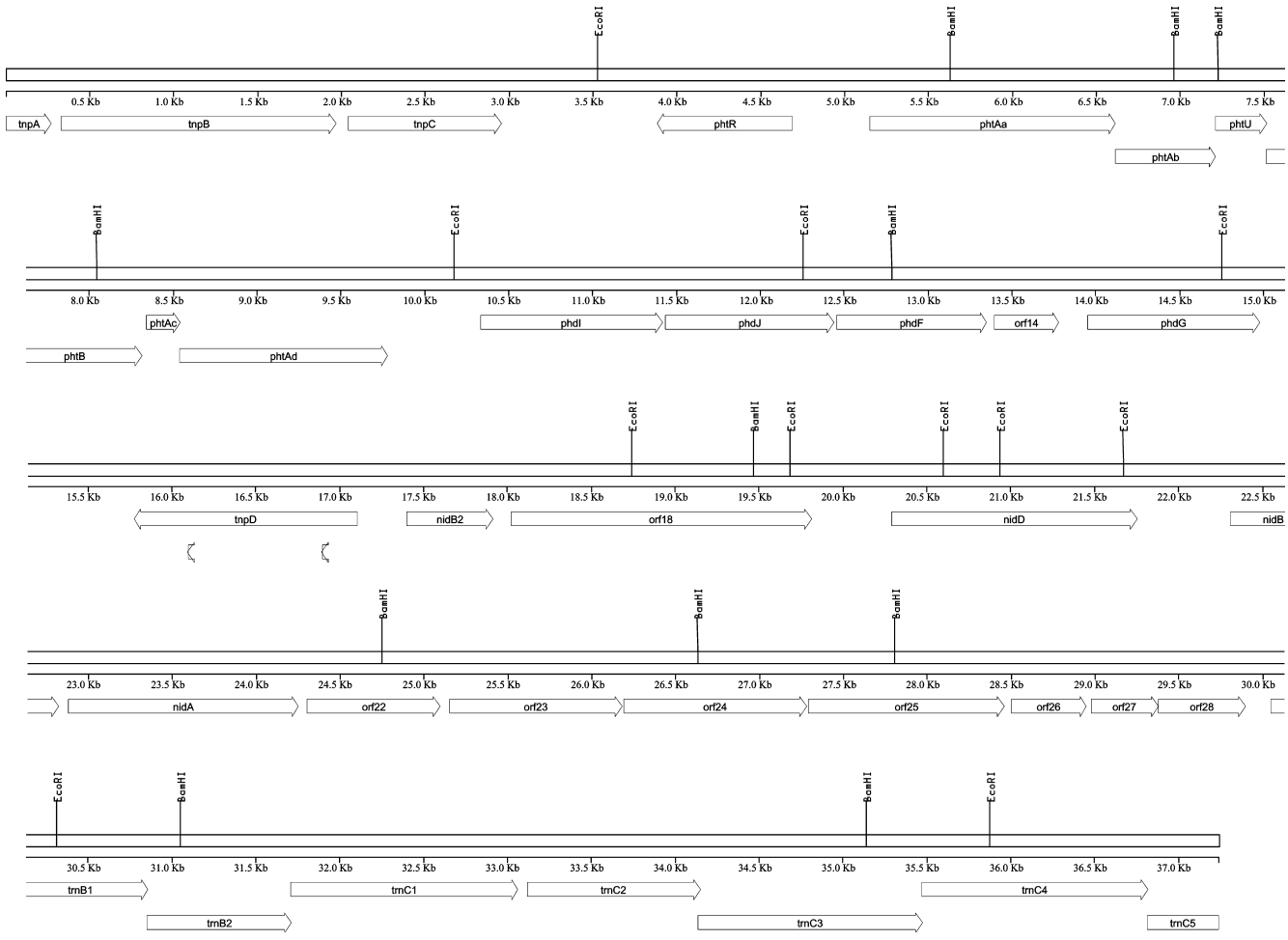

Fig. 4. Scaled representation of putative open reading frames (orfs) in pFOS608. Open arrows represent putative orfs and indicate their orientation. Symbols beneath the *tnpD* orf indicate attenuation by an amber stop (nt 160976–8) and a +1 frameshift (within nt 16097–16104). Two putative operons, the *pht* operon and the *trn* operon, contain some overlapping orfs, which are shown as staggered arrows. The representation was created using the MapDraw program from Lasergene software.

coded one partial, one attenuated, and two complete transposases, one partial and six complete membrane proteins, a transcriptional regulator, and four unknown proteins. Two putative operons were discovered in the region. One operon, which spanned approximately nucleotides (nt) 3886–9780, encoded enzymes involved in phthalate degradation (Fig. 4, Table 2) [30]. The other operon encoded genes for a putative ATP-binding cassette (ABC) transporter(s), and appeared to extend from approximately nt 30049 to beyond the pFOS608 insert region.

Ring-hydroxylating dioxygenases

Among the putative oxygenase genes were three sets of large and small subunit ring-hydroxylating dioxygenase genes. The large subunits contained conserved Rieske and aromatic ring-hydroxylating α-dioxygenase domains. A small subunit gene with an aromatic ring-hydroxylating β-dioxygenase domain was adjacent to each large subunit gene within the pFOS608 region. Genes *phtAa* and *phtAb* encoded putative phthalate dioxygenase large and small subunits, respectively [30]; *nidB* and *nidA* encoded putative naphthalene-inducible small and large subunits, respectively [14]; and orfs 25 and 26 encoded putative aromatic oxygenase large and small subunits, respectively (Table 2). The nucleotide sequence of *nidB2* was 97% identical to that of *nidB*. The putative proteins encoded by the two genes differ in a single amino acid residue (*nidB2* Ala-16 and *nidB* Glu-16).

While the large subunit ring-hydroxylating dioxygenases shared conserved domain regions, their nucleotide sequences shared no significant similarity and their deduced amino acid sequences shared limited homology. Amino acid sequences for the product of orf25 shared 28% identity and 40% positive similarity with NidA and 32% identity and 43% positive similarity with PhtAa; and NidA and PhtAa shared 41% identity and

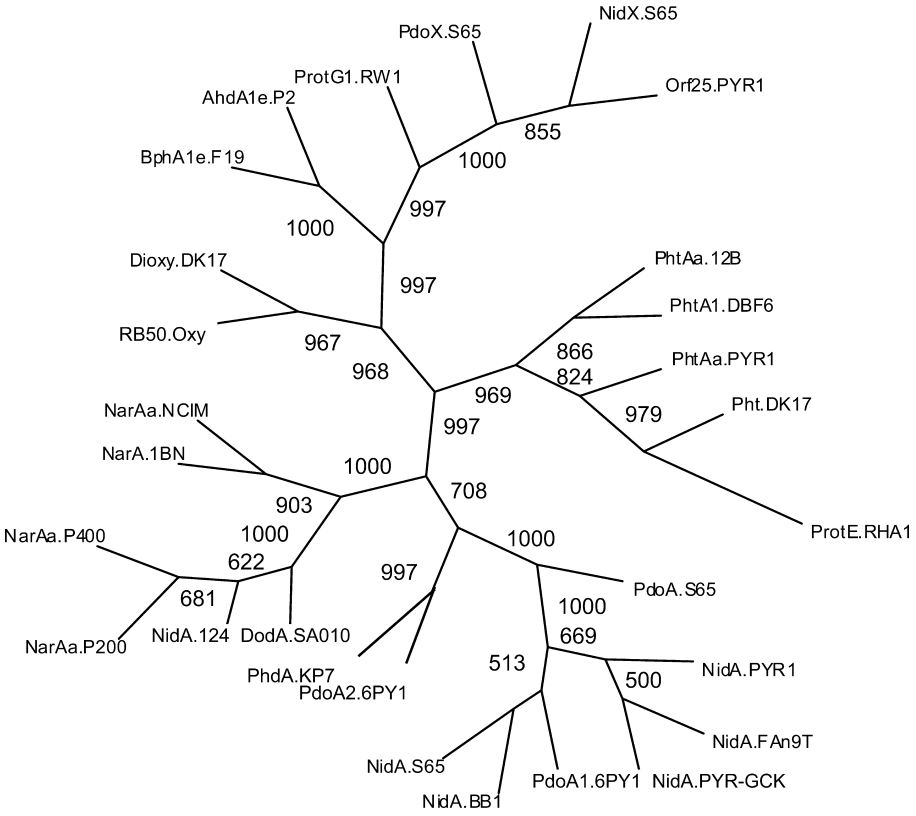

Fig. 5. Phylogenetic analysis of ring-hydroxylating large subunit dioxygenases. Protein sequences with the greatest homology to *M. vanbaalenii* PYR-1 large subunit dioxygenase proteins, based on Blastp results, were used in the analysis. The multiple-alignment analysis was performed using the PHYLIP software package, and the phylogenetic unrooted tree was drawn using TreeView. The numbers on some branches refer to the percentage confidence, estimated by a bootstrap analysis with 1000 replications. The scale bar indicates percentage divergence. GenBank accession numbers are as follows: Orf25 of *M. vanbaalenii* PYR-1, AY365117; PhtAa of *M. vanbaalenii* PYR-1, AY365117; NidA of *M. vanbaalenii* PYR-1, AAF75991; Protein E of *Rhodococcus* sp. RHA1, BAB62289; PhtA1 of *Terrabacter* sp. DBF63, BAC54156; PhtAa of *Arthrobacter keyseri* 12B, AAK16534; NidA of *Rhodococcus* sp. I24, AAD25395; NarAa of *Rhodococcus* sp. NCIMB12038, AAD28100; NarA of *Rhodococcus* sp. 1BN, CAC14063; PhdA of *Nocardioides* sp. KP7, BAA84712; NidA of *M. gilvum* BB1, AAN78316; NidA of *M. flavescens* PYR_GCK,AAN78312; PdoA1 of *Mycobacterium* sp. 6PY1, CAD38647; NidA of *M. frederiksbergense* FAn9T, AAN78314; PdoA2 of *Mycobacterium* sp. 6PY1, CAD38643; NidA of *Mycobacterium* sp. S65, AAQ12023; PdoA of *Mycobacterium* sp. S65, AAQ12029; DodA of *R. opacus* SA0101, BAD02377; NarAa of *Rhodococcus* sp. P400, AAR05106; NarAa of *Rhodococcus* sp. P200, AAR05114; NidX of *Mycobacterium* sp. S65, AAQ12027; PdoX of *Mycobacterium* sp. S65, AAQ12033; ProtG1 of *Sphingomonas* sp. RW1, CAA11187; BphA1e of *Novosphingobium aromaticivorans* F199, NP_049179; AhdA1e of *Sphingomonas* sp. P2, BAC65453; Oxygenase of *Bordetella bronchiseptica* RB50, NP_888136; dioxy of *Rhodococcus* sp. DK17, AAR90187; and Pht of *Rhodococcus* sp. DK17, AAR90178.

56% positive similarity. Phylogenetic analysis of the large subunit ring-hydroxylating dioxygenases and the top proteins in blast searches with each indicated that these three dioxygenases are distantly related. Each of the *M. vanbaalenii* PYR-1 large subunit dioxygenases was part of a different clade in an unrooted phylogenetic tree (Fig. 5).

Transposases

Based on blastx data, putative transposases were encoded by *tnpA* (partial), *tnpB*, *tnpC*, and *tnpD* (attenuated) (Table 2). The portion of *tnpA* that is contained in pFOS608 potentially encoded a partial protein that aligned with the 3′ portion of an integrase domain between residues 2 and 65. The protein encoded by *tnpB* aligned with the 3′ portion of a transposase domain between residues 195 and 484, and with the 3′ portion of a transposase DDE domain between residues 314 and 486. The transposase DDE domain contains three carboxylate residues, which are likely involved in coordinating metal ions for catalysis [19]. Two long direct repeats of 141 nt each flank *tnpB*, overlapping the 3′ end of *tnpA* and the 5′ end of *tnpC*. The *tnpC* protein contained no known conserved domains, but shared limited homology with known transposases (Table 2). Finally, *tnpD* appeared to have been attenuated by mutations causing both premature stops and a frameshift. A blastx search with *tnpD* (nt 15,776–16,774) indicated that protein sequence from this region shared more than 85% identity and 91% positive similarity with a probable transposase from *Mycobacterium ulcerans*. However, the region con-

tained an amber stop site at nt 16896–8 and a +1 frameshift within nt 16,097–16,104. Compensation for the frameshift exposed an ochre stop site at nt 16095–7.

Putative phenanthrene lower-pathway enzymes

The region between *tnpC* and *tnpD* contained a number of orfs encoding putative proteins with homology to those that are involved in the degradation pathway from 3,4-dihydroxyphenanthrene to 3,4-dihydroxyphthalic acid (Fig. 1). Only genes encoding isomerase (step 4), dehydrogenase (step 6), 2-carboxybenzaldehyde dehydrogenase (step 9), and phthalate decarboxylase (step 12) were missing from this region.

Orfs *phdF* and *phdG* encoded an extradiol dioxygenase and a hydratase–aldolase, respectively. *M. vanbaalenii* PYR-1 PhdF and PhdG shared 83% and 87% identity with their counterparts in *Nocardioides* sp. KP7, respectively. Orfs *phdI* and *phdJ* encoded 1-hydroxy-2-naphthoate dioxygenase and another hydratase–aldolase, respectively. *M. vanbaalenii* PYR-1 PhdI and PhdJ shared 46% and 55% identity with their counterparts in *Nocardioides* sp. KP7, respectively.

The putative *pht* operon contained genes with significant homology to phthalate dioxygenase large (*phtAa*) and small (*phtAb*) subunits, ferredoxin (*phtAc*), ferredoxin reductase (*phtAd*), and phthalate dihydrodiol dehydrogenase (*phtB*) [30] (Table 1). The operon also contained a putative gene encoding an unknown protein (*phtU*). The putative transcriptional regulator (*phtR*) for the operon was divergently encoded upstream of the operon (Fig. 4).

Dehydrogenases

Orf18 encoded a putative protein with a succinate dehydrogenase/fumarate reductase domain between residues 36 and 558. The orf18 protein also contained a fumarate reductase/succinate dehydrogenase flavoprotein C-terminal domain between residues 463 and 577. The *M. vanbaalenii* PYR-1 aldehyde dehydrogenase is encoded by *nidD* and has been described [14]. Orf22 encoded a putative protein with a short chain dehydrogenase domain between residues 27 and 251, and shared limited homology with oxidoreductases and short chain dehydrogenases (Table 2). Orf23 encoded a putative protein with limited homology to butanediol dehydrogenases (Table 2) and a zinc-dependent dehydrogenase domain between residues 1 and 319.

Putative ABC-transporter operon

A putative operon for an ABC-transport system contained *trnB1*, *B2*, *C1*, *C2*, *C3*, *C4*, and *C5* (partial). The stop sites of *trnB1*, *trnC3*, and *trnC4* overlap with the start sites of *trnB2*, *trnC4*, and *trnC5*, respectively, with each overlap consisting of the four bases: ATGA. In addition, *trnC2* and *trnC3* overlap by 16 bases. Blastx data for *trnC1*, *trnC2*, *trnC3*, and *trnC4* showed sequence homology with mycobacterial cell entry (mce) proteins, which compose a secretion system that is associated with virulence in *M. tuberculosis* [3,6]. Conserved Domain data indicated that each of the putative proteins contained a domain related to an ABC-transport system involved in resistance to organic solvents (Table 3).

Table 3
Putative ABC-transport operon conserved domains

Orf	Putative protein size (amino acid residues)	Conserved domains[a]
trnB1	269	Residues 20–268 aligned with COG0767, Ttg2B, ABC-type transport system involved in resistance to organic solvents, permease component (secondary metabolite biosynthesis, transport, and catabolism)
trnB2	285	Residues 79–279 aligned with Ttg2B
trnC1	448	Residues 8–310 aligned with COG1463, Ttg2C, ABC-type transport system involved in resistance to organic solvents, periplasmic component (secondary metabolite biosynthesis, transport, and catabolism)
		Residues 15–155 aligned with pfam02470, mycobacterial cell entry (mce) related protein
trnC2	342	Residues 12–342 aligned with Ttg2C
		Residues 11–152 aligned with mce
trnC3	446	Residues 10–356 aligned with Ttg2C
		Residues 16–152 aligned with mce
trnC4	448	Residues 8–364 aligned with Ttg2C
		Residues 8–146 aligned with mce
		Residues 143–291 aligned with portion of COG0840, Tar, methyl-accepting chemotaxis protein
		Residues 25–219 aligned with portion of COG0643, CheA, chemotaxis protein histidine kinase
		Residues 147–330 aligned with portion of pfam04582, Reovirus sigma C capsid protein
trnC5	130 (partial)	Residues 1–128 aligned with portion of Ttg2C
		Residues 1–127 aligned with mce

[a] Conserved domains based on blastp conserved domain search [19].

Putative unknown proteins

Orf28 encoded a putative protein with a dioxygenase domain between residues 18 and 154. This region of the putative protein also shared limited homology with the 3′ end of a protocatechuate β subunit domain (Table 2). Orfs 14 and 27 encoded putative proteins with homology to proteins of unknown function (Table 2). Both putative proteins, however, contained sterol transfer and sterol carrier domains. The protein encoded by orf14 aligned with the 3′ region of a sterol transfer domain between residues 48 and 117, and with the 3′ region of a sterol carrier domain between residues 45 and 108. The orf27 protein aligned with the sterol carrier domain between residues 7 and 117, and with the sterol transfer domain between residues 30 and 117. The two putative sterol carrier proteins shared no significant sequence homology. Orf24 encoded an unknown protein with limited homology to known proteins (Table 2), and only 11% alignment with the 5′ portion of a vacuolar protein sorting (VPS) domain between residues 137 and 210.

Biotransformation of phenanthrene by the enzymes of pFOS608

To determine whether the putative genes encoded on the pFOS608 insert were capable of phenanthrene degradation, pFOS608 was induced to high copy number. Production of putative PAH-degrading enzymes was simultaneously induced with 6 μg/ml of phenanthrene. After approximately 24 h of growth at 37°C, phenanthrene substrate was added to a final concentration of 26.6 μg/ml. A control culture of EPI300 cells with no fosmid was treated in the same manner. After a 24 h incubation at 37 °C, neutral and acidic ethyl acetate extractions were analyzed by HPLC. Three peaks were present in the pFOS608 neutral extract and not in the EPI300 extract, however, no differences were observed in the acid extracts. A compound eluting at 18.3 min of the pFOS608 neutral extract had a UV absorbance spectrum corresponding to that of phenanthrene *cis*-3,4-dihydrodiol [21]. A similar peak and UV spectrum was obtained with pRS53, which contains *nidDBA*, suggesting that the metabolite was formed by the NidAB dioxygenase. GC/MS analysis of the pFOS608 phenanthrene metabolite supported its identification as phenanthrene *cis*-3,4-dihydrodiol [21]. The major ions were detected at *m/z* 212 (100, M^+), 194 (22, M − 18), 181 (76, M − 31), and 165 (10). The GC mass analysis of the trimethylsilylated product showed one compound at 14.6 min which is consistent with a di-trimethylsilylated ether of phenanthrene *cis*-3,4-dihydrodiol. The major fragment ions were 357 (8, M + 1), 356 (29, M^+), 341 (24, M − 15), 325 (10, M − 31), 309 (5, M − 47), 266 (24, M − 90), 253, 251, 235, 178, 165, 147, and 73. Approximately 98% of the phenanthrene metabolite formed was phenanthrene *cis*-3,4-dihydrodiol with a conversion ratio of 25%. The phenanthrene *cis*-9,10-dihydrodiol was not detected. The other two minor compounds had M^+ at *m/z* 194 and retention times identical to phenanthrols suggesting dehydration of the phenanthrene *cis*-3,4-dihydrodiol.

Biotransformation of 1-hydroxy-2-naphthoic acid by the enzymes of pFOS608

To determine whether the lower pathway of phenanthrene degradation was performed by products from pFOS608, the clone was induced to high copy number and putative PAH-degrading enzyme production was simultaneously induced with 6 μg/ml of phenanthrene. After approximately 7 h of growth, 40 μg/ml of 1-hydroxy-2-naphthoic acid was added and the cells were incubated at 37 °C for approximately 24 h. Preliminary sequencing, *Bam*HI digestion patterns, and Southern hybridization (Fig. 2) suggest that pFOS570 overlaps pFOS608 but does not contain genes for the lower pathway enzymes, so this clone was used as a negative control. Neutral and acidic extractions were analyzed by negative ion HPLC/ESI MS. Neutral samples from the two clones were similar, but acidic samples from pFOS608 contained at least three peaks that were not present in the pFOS570 control.

LC/ESI MS in the negative electrospray mode of the pF0S608-1-hydroxy-2-naphthoic acid metabolites showed one major metabolite at 17.02 min and two minor metabolites at 4.35 and 16.32 min, respectively. The compound at 17.02 min showed ions at *m/z* 219 [M − H^-] and *m/z* 129 and UV spectral properties (274 and 281 nm) which is consistent with 2-carboxybenzalpyruvate [1]. The relative amounts of 2-carboxybenzalpyrate formed were approximately 277% in relation to pFOS570 control. The compounds at 4.35 min, *m/z* 165 [M − H^-], and 16.32 min, *m/z* 149 [M − H^-] were consistent with phthalic acid and carboxybenzaldehyde. GCMS analysis of the trimethylsilylated derivatives of these compounds further confirmed them as metabolites of 1-hydroxy-2-naphthoic acid.

Discussion

Mycobacterium vanbaalenii PYR-1 is capable of degrading a wide range of hydrocarbons including high molecular weight PAHs. The organism utilizes multiple pathways to degrade some PAHs, therefore its genome likely encodes a number of enzymes involved in these pathways. We have identified several potential PAH-degradation pathway genes within a 37 kb region of previously characterized dioxygenase large and small subunit genes from *M. vanbaalenii* PYR-1 [14].

Based on protein sequence homology and conserved domain data, a number of genes contained in pFOS608 appeared to be involved in the phenanthrene degradation pathway initiated by the dioxygenase attack at the C-3 and C-4 positions. These genes encode putative enzymes that are capable of performing the majority of the steps involved in phenanthrene degradation to 3,4-dihydroxyphthalate (Fig. 1).

The initial dioxygenase step in this phenanthrene degradation pathway is likely performed by the NidAB dioxygenase. NidAB attacks pyrene in the K region, producing the 4,5-dihydrodiol [14]. This pyrene degradation pathway likely intersects the phenanthrene pathway described here at the 3,4-dihydroxyphenanthrene step [4] (Fig. 1). Considering the pattern of NidAB attack for pyrene, we had anticipated that this dioxygenase would attack phenanthrene at the C-9 and C-10 positions, initiating a different degradation pathway that has been observed in *M. vanbaalenii* PYR-1 [21]. Based upon the UV and mass spectra of the phenanthrene dihydrodiol produced by both pFOS608 and pRS53 (*nidDBA* only), we believe that NidAB actually attacks phenanthrene at the C-3 and C-4 positions, initiating the pathway shown in Fig. 1. These data suggest that the multiple dioxygenases are utilized by *M. vanbaalenii* PYR-1 for the degradation of PAHs.

Blastp and Conserved Domain Database searches indicated that the proteins encoded by orfs 25 and 26 were large and small subunits, respectively, of an aromatic ring hydroxylating dioxygenase. A number of proteins identified in blastp searches belonged to gram-negative organisms and were responsible for the oxygenation of benzoate compounds including salicylate, anthranilate, and 2-chlorobenzoate [5,10,24,31]. AdhA1e and AdhA2e of *Sphingobium* sp. strain P2 were among the top blastp results for the products of orfs 25 and 26, respectively, and with a ferredoxin and ferredoxin reductase converted 5-methylsalicylate to 4-methylcatechol in *E. coli* [24].

Phenanthrene degradation by pFOS608 produced only one metabolite, phenanthrene *cis*-3,4-dihydrodiol, suggesting that the dehydrogenase required to form 3,4-dihydroxyphenanthrene is not present in this region of the genome (Fig. 1, step 2). Putative proteins encoded by *phdF*, orf14, *phdG*, *phdI*, and *phdJ* share sequence homology to *phdF*, orf131, *phdG*, *phdI*, and *phdJ* of *Nocardioides* sp. KP7, respectively (Table 2), and likely share their respective roles in phenanthrene degradation [11,28]. Aldehyde dehydrogenase steps (Fig. 1, steps 6 and 9) may be performed by the aldehyde dehydrogenase, NidD, or by one of the other dehydrogenases in the region. Since the upper pathway step 2 of phenanthrene degradation (Fig. 1) was not performed by the clone, the precise functions of PhdF and PhdG could not be confirmed. Degradation of 1-hydroxy-2-naphthoate by pFOS608 produced 2-carboxybenzalpyruvic acid, 2-carboxybenzaldehyde, and phthalic acid, suggesting that *M. vanbaalenii* PYR-1 PhdI and PhdJ are indeed responsible for their respective lower pathway degradation steps (Fig. 1, steps 7–8), and that NidD, or one of the other dehydrogenases in the region, is responsible for step 9 in Fig. 1. The putative *pht* operon region is capable of degrading phthalic acid to 3,4-dihydroxyphthalic acid in *E. coli* [30]. No genes that encoded homologs to an isomerase (Fig. 1, step 4) or to a decarboxylase (Fig. 1, step 12) were found in the pFOS608 sequence.

A number of genes in the region appeared to be part of an operon for an ABC-transport system. Genes in the putative operon encoded enzymes with domains similar to those for proteins of an ABC-transporter involved in secondary metabolite biosynthesis, transport, and catabolism (Table 3). These data and the proximity of the operon to a number of PAH-degradation genes suggest that the putative transporter may play a role in the uptake of hydrocarbons or in the efflux of metabolites.

In addition to PAH-degradation enzymes, several transposases were discovered in the sequenced region. Long direct repeats (145 nt) were identified on either side of *tnpB*. Some insertion elements form relatively long variable-length direct repeats of the target DNA during transposition, causing the insertion element to be flanked by the direct repeats [25,32]. The majority of the genes involved in the pathway between 3,4-dihydroxy phenanthrene and 3,4-dihydroxy phthalic acid fall in the region between three tandem transposases (*tnpABC*) and a single attenuated transposase (*tnpD*) (Fig. 4).

This report is the first to identify genes in a *Mycobacterium* species that are involved in the degradation of phenanthrene via the phthalic acid pathway. This pathway was previously observed in *M. vanbaalenii* PYR-1 [21]. The current study describes a number of *M. vanbaalenii* PYR-1 genes that encode enzymes responsible for the pathway. The genes were identified based upon the limited homology of their products to proteins of other organisms. The overall organization of the phenanthrene degradation genes in *M. vanbaalenii* PYR-1 differs significantly from that of similar genes in other bacterial genera. Further investigation will likely discover many additional genes required to perform the numerous PAH degradation pathways utilized by *M. vanbaalenii* PYR-1.

Acknowledgments

This work was supported in part by an appointment to the Postgraduate Research Program at the National Center for Toxicological Research administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the US Food and Drug Administration.

The authors thank Mr. Allen Gies at the University of Arkansas for Medical Sciences for sequencing, Dr. James P. Freeman for GCMS analyses, Ms. Joanna Moody and Ms. Lisa Mullis for technical support, and Dr. Mark Hart and Dr. Christopher Elkins for critical review of the manuscript.

References

[1] K. Adachi, T. Iwabuchi, H. Sano, S. Harayama, Structure of the ring cleavage product of 1-hydroxy-2-naphthoate, an intermediate of the phenanthrene-degradative pathway of *Nocardioides* sp. strain KP7, J. Bacteriol. 181 (1999) 757–763.

[2] J. Armenguad, B. Happe, K.N. Timmis, Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. Strain RW1: catabolic genes dispersed on the genome, J. Bacteriol. 180 (1998) 3954–3966.

[3] S. Arruda, G. Bomfim, R. Knights, T. Huima-Byron, L.W. Riley, Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells, Science 261 (1993) 1454–1457.

[4] C.E. Cerniglia, Biodegradation of polycyclic aromatic hydrocarbons, Biodegradation 3 (1992) 351–368.

[5] H.K. Chang, P. Mohseni, G.J. Zylstra, Characterization and regulation of the genes for a novel anthranilate 1,2-dioxygenase from *Burkholderia cepacia* DBO1, J. Bacteriol. 185 (2003) 5871–5881.

[6] S. Chitale, S. Ehrt, I. Kawamura, T. Fujimura, N. Shimono, N. Anand, S. Lu, L. Cohen-Gould, L.W. Riley, Recombinant *Mycobacterium tuberculosis* protein associated with mammalian cell entry, Cell. Microbiol. 3 (2001) 247–254.

[7] H.P. Doddamani, H.Z. Ninnekar, Biodegradation of phenanthrene by a *Bacillus* species, Curr. Opin. Microbiol. 41 (2000) 11–14.

[8] M.A. Heitkamp, W. Franklin, C.E. Cerniglia, Microbial metabolism of polycyclic aromatic hydrocarbons: isolation and characterization of a pyrene-degrading bacterium, Appl. Environ. Microbiol. 54 (1988) 549–555.

[9] M.A. Heitkamp, J.P. Freeman, D.W. Miller, C.E. Cerniglia, Pyrene degradation by a *Mycobacterium* sp.: identification of ring oxidation and ring fission products, Appl. Environ. Microbiol. 54 (1988) 2556–2565.

[10] W.J. Hickey, G. Sabat, A.S. Yuroff, A.R. Arment, J. Perez-Lesher, Cloning, nucleotide sequencing, and functional analysis of a novel, mobile cluster of biodegradation genes from *Pseudomonas aeruginosa* strain JB2, Appl. Environ. Microbiol. 67 (2001) 4603–4609.

[11] T. Iwabuchi, S. Harayama, Biochemical and genetic characterization of *trans*-2′-carboxybenzalpyruvate hydratase–aldolase from a phenanthrene-degrading *Nocardioides* strain, J. Bacteriol. 180 (1998) 945–949.

[12] H. Kang, S.Y. Hwang, Y.M. Kim, E. Kim, Y.S Kim, S.K. Kim, S.W. Kim, C.E. Cerniglia, K.L. Shuttleworth, G.J. Zylstra, Degradation of phenanthrene and naphthalene by a *Burkholderia* species strain, Can. J. Microbiol. 49 (2003) 139–144.

[13] I. Kelley, J.P. Freeman, F.E. Evans, C.E. Cerniglia, Identification of metabolites from the degradation of fluoranthene by *Mycobacterium* sp. strain PYR-1, Appl. Environ. Microbiol. 59 (1993) 800–806.

[14] A.A. Khan, R.-F. Wang, W.-W. Cao, D.R. Doerge, D. Wennerstrom, C.E. Cerniglia, Molecular cloning, nucleotide sequence, and expression of genes encoding a polycyclic aromatic dioxygenase from *Mycobacterium* sp. strain PYR-1, Appl. Environ. Microbiol. 67 (2001) 3577–3585.

[15] E. Kim, G.J. Zylstra, Functional analysis of genes involved in biphenyl, naphthalene, phenanthrene, and m-xylene degradation by *Sphingomonas yanoikuyae* B1, J. Ind. Microbiol. Biotechnol. 23 (1999) 294–302.

[16] S. Krivobok, S. Kuony, C. Meyer, M. Louwagie, J.C. Willison, Y. Jouanneau, Identification of pyrene-induced proteins in *Mycobacterium* sp. strain 6PY1: evidence for two ring-hydroxylating dioxygenases, J. Bacteriol. 185 (2003) 3828–3841.

[17] A.D. Laurie, G. Lloyd-Jones, The phn genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism, J. Bacteriol. 181 (1999) 531–540.

[18] C.T. MacLeod, A.J. Daugulis, Biodegradation of polycyclic aromatic hydrocarbons in a two-phase partitioning bioreactor in the presence of a bioavailable solvent, Appl. Microbiol. Biotechnol. 62 (2003) 291–296.

[19] A. Marchler-Bauer, J.B. Anderson, C. DeWeese-Scott, N.D. Fedorova, L.Y. Geer, S. He, D.I. Hurwitz, J.D. Jackson, A.R. Jacobs, C.J. Lanczycki, C.A. Liebert, C. Liu, T. Madej, G.H. Marchler, R. Mazumder, A.N. Nikolskaya, A.R. Panchenko, B.S. Rao, B.A. Shoemaker, V. Simonyan, J.S. Song, P.A. Thiessen, S. Vasudevan, Y. Wang, R.A. Yamashita, J.J. Yin, S.H. Bryant, CDD: a curated Entrez database of conserved domain alignments, Nucleic Acids Res. 31 (2003) 383–387.

[20] J.D. Moody, D.R. Doerge, J.P. Freeman, C.E. Cerniglia, Degradation of biphenyl by *Mycobacterium* sp. strain PYR-1, Appl. Microbiol. Biotechnol. 58 (2002) 364–369.

[21] J.D. Moody, J.P. Freeman, D.R. Doerge, C.E. Cerniglia, Degradation of phenanthrene and anthracene by cell suspensions of *Mycobacterium* sp. strain PYR-1, Appl. Environ. Microbiol. 67 (2001) 1476–1483.

[22] J.D. Moody, J.P. Freeman, P.P. Fu, C.E. Cerniglia, Degradation of benzo[*a*]pyrene by *Mycobacterium vanbaalenii* PYR-1, Appl. Environ. Microbiol. 70 (2004) 340–345.

[23] J.D. Moody, P.P. Fun, J.P. Freeman, C.E. Cerniglia, Regio- and stereoselective metabolism of 7,12-dimethylbenz[*a*]anthracene by *Mycobacterium vanbaalenii* PYR-1, Appl. Environ. Microbiol. 69 (2003) 3924–3931.

[24] O. Pinyakong, H. Habe, T. Yoshida, H. Nojiri, T. Omori, Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of *Sphingobium* sp. strain P2, Biochem. Biophys. Res. Commun. 301 (2003) 350–357.

[25] B.B. Plikaytis, J.T Crawford, T.M. Shinnick, IS1549 from *Mycobacterium smegmatis* forms long direct repeats upon insertion, J. Bacteriol. 180 (1998) 1037–1043.

[26] F. Rafii, A.L. Selby, R.K. Newton, C.E. Cerniglia, Reduction and mutagenic activation of nitroaromatic compounds by a *Mycobacterium* sp, Appl. Environ. Microbiol. 60 (1994) 4263–4267.

[27] M.F. Romine, L.C. Stillwell, K.K. Wong, S.J. Thurston, E.C. Sisk, C. Sensen, T. Gaasterland, J.K. Fredrickson, J.D. Saffer, Complete sequence of a 184-kb catabolic plasmid from *Sphingomonas aromaticivorans* F199, J. Bacteriol. 181 (1999) 1585–1602.

[28] A. Saito, T. Iwabuchi, S. Harayama, Characterization of genes for enzymes involved in the phenanthrene degradation in *Nocardioides* sp. KP7, Chemopshere 38 (1999) 1331–1337.

[29] A. Saito, T. Iwabuchi, S. Harayama, A novel phenanthrene dioxygenase from *Nocardioides* sp. Strain KP7: expression in *Escherichia coli*, J. Bacteriol. 182 (2000) 2134–2141.

[30] R.L. Stingley, B. Brezna, A.A. Khan, C.E. Cerniglia, Novel organization of genes in a phthalate degradation operon of *Mycobacterium vanbaalenii* PYR-1, Microbiology (2004), in press.

[31] T.V. Tsoi, E.G Plotnikova, J.R. Cole, W.F. Guerin, M. Bagdasarian, J.M. Tiedje, Cloning, expression, and nucleotide sequence of the *Pseudomonas aeruginosa* 142 ohb genes coding for oxygenolytic ortho dehalogenation of halobenzoates, Appl. Environ. Microbiol. 65 (1999) 2151–2162.

[32] E.M. Vilei, J. Nicolet, J. Frey, IS1634, a novel insertion element creating long, variable-length direct repeats which is specific for *Mycoplasma mycoides* subsp. *Mycoides* small-colony type, J. Bacteriol. 181 (1999) 1319023.