Biodegradation (2012) 23:585–596
DOI 10.1007/s10532-012-9535-z

ORIGINAL PAPER

Multiplicity of genes for aromatic ring-hydroxylating dioxygenases in *Mycobacterium* isolate KMS and their regulation

Chun Zhang · Anne J. Anderson

Received: 9 May 2011 / Accepted: 9 January 2012 / Published online: 4 February 2012

Abstract *Mycobacterium sp.* strain KMS has bioremediation potential for polycyclic aromatic hydrocarbons (PAHs), such as pyrene, and smaller ring aromatics, such as benzoate. Degradation of these aromatics involves oxidation catalyzed by aromatic ring-hydroxylating dioxygenases. Multiple genes encoding dioxygenases exist in KMS: ten genes encode large-subunits with homology to phenylpropionate dioxygenase genes, sixteen pairs of adjacent genes encode alpha- and beta-subunits of dioxygenase and two genes encode beta-subunits. These genes include orthologs of *nid* genes essential for degradation of multi-ring PAHs in *M. vanbaalenii* isolate PYR-1. The multiplicity of genes in part is explained by block duplication that results in two or three copies of certain genes on the chromosome, a linear plasmid, and a circular plasmid within the KMS genome. Quantitative real-time PCR showed that four dioxygenase beta-subunit *nid* genes from operons with almost identical promoter sequences otherwise unique in the genome were induced by pyrene to similar extents. No induction occurred with benzoate. Unlike isolate PYR-1, isolate KMS has an operon specifying benzoate catabolism and the expression of the alpha-subunit dioxygenase gene was activated by benzoate but not pyrene. These studies showed that isolate KMS had a genome well adapted to utilization of different aromatic compounds.

Keywords Aromatic ring-hydroxylating dioxygenase · Benzoate-degrading gene cluster · Phylogenetics

Electronic supplementary material The online version of this article (doi:10.1007/s10532-012-9535-z) contains supplementary material, which is available to authorized users.

C. Zhang · A. J. Anderson (✉)
Biology Department, Utah State University, Logan, UT 84322-5305, USA
e-mail: anne.anderson@usu.edu

Introduction

Aromatics range from single ring, such as benzoate, to multiple fused-ring polycyclic aromatic hydrocarbons (PAHs). Benzoate naturally occurs in plants and is utilized as an antimicrobial agent in the food industry (Orlova et al. 2006; Zengin et al. 2010). PAHs are generated from incomplete combustion of organics, and are notable for their toxic, mutagenic, and carcinogenic properties (U.S. Department of Health and Human Services 1995; Tannheimer et al. 1997). Aerobic degradation of aromatic hydrocarbons frequently is initiated by oxidation catalyzed by ring-hydroxylating dioxygenases (Butler and Mason 1997; Peng et al. 2010). The dioxygenases work in conjunction with small electron transport chains so that a multi-protein structure is involved in the oxygenation process (Butler and Mason 1997; Correll et al. 1993; Jiang et al. 1999; Kweon et al. 2008; Peng et al. 2010). Such dioxygenases generally are complexes of a large alpha-

subunit and a small beta-subunit (Butler and Mason 1997; Parales and Resnick 2006), although a subset only has one alpha-subunit (Parales and Resnick 2006). Kweon et al. (2008) derived a new classification scheme based on homology of the oxygenases and the composition of the electron transport chains. The benzoate dioxygenase is found in Type III which has a simpler electron transport chain than the Type V dioxygenases where the pyrene-degrading dioxygenases are found (Kweon et al. 2008).

Working with *Mycobacteriunm vanbaalenii* PYR-1, Cerniglia's group demonstrates that the initial ring oxidation of high molecular weight PAHs, such as pyrene and fluoranthene, involves the dioxygenases encoded by chromosomal genes, *nidB/nidA*, *nidB2* and *nidA3/nidB3* (Khan et al. 2001; Kim et al. 2006; Kim et al. 2007; Kweon et al. 2010). Specificity for the ring structures that are oxidized is conditioned by the alpha-subunits more than the beta-subunits (Parales and Resnick 2006), although contributions from the beta-subunits are postulated (Hurtubise et al. 1998; Jiang et al. 1999). Kweon et al. (2010) elegantly demonstrate that the preferred activity for a substrate correlates with the size of the substrate pocket in the alpha subunit. However, they also comment that the substrate plays a role in regulating the production of the enzyme. Proteomic analysis reveals that the enzyme that uses pyrene as the preferred substrate NidA/B is induced in pyrene but not fluoranthene cultures and that NidA3/B3, which better degrades fluoranthene than pyrene, is produced to a greater extent when fluoranthene is the carbon source (Kweon et al. 2010).

Our group found that the environmental *Mycobacterium sp.* strain KMS from PAH-contaminated soils in Montana (Miller et al. 2004) has genes encoding orthologs of *nidB*, *nidB2* and *nidA* (Zhang and Anderson 2012). These genes are embedded in segments of the chromosome that have synteny with other PAH-degrading mycobacterium isolates (Zhang and Anderson 2012). In this paper we determined the full complement of dioxygenase genes present in the KMS genome. Previous studies with isolate PYR-1 focus only on the Nid proteins (Kweon et al. 2008, 2010). We provide evidence that block duplication events have caused certain of these genes to be present as two and three copies with locations on the chromosome as well as on circular and linear plasmids. Bioinformatics predicts that one set of paired alpha and beta dioxygenase genes could be involved in benzoate degradation in isolate KMS; isolate PYR-1 lacks such genes and does not grow on benzoate (Kim et al. 2008). We verified growth of isolate KMS on benzoate and examined the influence of the growth substrate, benzoate versus pyrene, on dioxygenase gene expression. We compared growth substrate regulation of the *benA* gene, encoding the benzoate dioxygenase alpha subunit, with the expression of the chromosomal copy of the *nidA* gene. We also examined expression from the homologous *nidB* and *nidB2* genes located on the chromosome and the circular plasmid in isolate KMS. These findings were correlated with identity in the sequences of the predicted promoters for each of the dioxygenase genes. Comparison of the benzoate-degrading gene clusters of isolate KMS as well as those in four other mycobacterium isolates to those in other benzoate-degrading microbes, the Gram-positive *Rhodococcus* (Kitagawa et al. 2001) and Gram-negative *Acinetobacter* isolates (Collier et al. 1997) was performed. The mycobacterium isolates were from the same USA site as isolate KMS, isolates JLS and MCS, and from two other locations in the USA, isolates PYR-1 and PYR-GCK, as well as an isolate, Spyr1, from Greece (Karabika et al. 2009). These studies showed that the arrangement of benzoate catabolism genes in the mycobacterium isolates differs from those of the other microbes but is conserved among these isolates from different geographic areas.

Materials and methods

Growth of isolate KMS on pyrene or sodium benzoate as the sole carbon source

Cells of *Mycobacterium sp.* strain KMS were stored at −80°C in 15% glycerol. The cells were cultured on modified minimal medium (Kim et al. 2003) plates amended with 2.5 mM pyrene or 5 mM sodium benzoate (Sigma Aldrich, St. Louis, MO, USA). Cells from the plates were transferred to liquid medium with defined concentrations of pyrene or benzoate to determine growth rate. The pyrene was dissolved in methanol at 5 mg ml^{-1} and 5 ml were transferred to each flask. The methanol was evaporated before adding 50 ml medium. The sodium benzoate was

sterilized by filtration through a 0.2 μm pore size Whatman filter (Whatman Inc., NJ, USA) and added separately to the minimal medium. The cells in liquid medium were shaken at 220 rpm at 25°C. Cell growth was monitored at OD 600 nm, and colony forming units per ml (CFU ml^{-1}) were assessed by dilution plating on Luria–Bertani (LB) agar medium; colonies were counted after 7–10 days growth at 25°C.

Bioinformatic analysis of genes encoding ring-hydroxylating dioxygenases in isolate KMS

MicrobesOnline (http://www.microbesonline.org) was used to identify genes encoding the alpha- and beta-subunits of aromatic ring dioxygenases in the genome of *Mycobacterium* isolate KMS. These loci were confirmed using the JGI-IMG website (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). The predicted amino acid sequences were used in BLASTp to determine matches to proteins in other microbes. The phylogenetic relationships at the amino acid level were examined using the software program MEGA4 (http://www.megasoftware.net/) using neighbor joining analysis (Saitou and Nei 1987; Tamura et al. 2007). This methodology has been used in other studies of bacterial gene duplication (Gu et al. 2009). The analysis was performed with 1,000 replications.

The potential promoter regions for the forty-four genes encoding aromatic ring-hydroxylating dioxygenase alpha- and beta-subunits were predicted by the online software Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) (Reese 2001). The minimum promoter score was set as 0.4, from the range of 0 to 1, where 1 represents the highest possibility of being a promoter region. The position of the promoters also was based on the predicted operons in MicrobesOnline and by searching through Artimis (Rutherford et al. 2000). The sequences totaling 60 bp from 50 bp upstream from the predicted transcriptional start sites for all the ring-hydroxylating dioxygenase alpha- and beta-subunits genes were aligned by ClustalX2 acquired from http://www.clustal.org/.

Induction of different dioxygenase genes by pyrene or benzoate

Total RNA was isolated by bead beating using TRI-reagent from middle-log phase cells (OD 600 nm at 0.15–0.20 for pyrene-induced cultures grown on 2.5 mM pyrene and OD 600 nm at 0.25–0.30 for cultures grown with 5 mM sodium benzoate (Molecular Research Center, Inc. OH, USA). The RNA was treated by RNase-free DNase (Promega Corporation, Madison, WI, USA) and purified by extraction with phenol:chloroform:isoamyl alcohol (25:24:1) to remove the DNase. The concentration of the purified RNA was determined by using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE, USA). PCR was performed with the DNase-treated RNA as the template to ensure the complete digestion of DNA. First-strand cDNA was synthesized by using specific reverse primers for the *benA* gene Mkms_1391, the chromosomal *nidA* gene (Mkms_1668), or the 16S rRNA genes with 1 μg purified total RNA. The primer sequences are listed in Supplemental Table S1. The positions of the specific primers for two *nidB* and two *nidB2* genes are shown in Supplemental Fig. S1: these primers were constructed around the regions of difference in their nucleotide sequences and consequently yield products of different sizes. The specificity of the primers was verified using genomic DNA and sequencing of the product prior to transcript analysis. The protocol for M-MuLV reverse transcriptase (Promega Corporation, Madison WI, USA) was followed. The synthesized first strand cDNA was used in a 20 μl volume end-point PCR reaction containing: 11.3 μl H_2O, 2 μl $MgCl_2$ (25 mM), 2 μl 10× Taq DNA polymerase buffer, 1 μl dNTP mix (10 mM each), 1 μl forward primer (final concentration 1 μM), 1 μl reverse primer (final concentration 1 μM), 1 μl cDNA, 0.3 μl Taq DNA polymerase (1.5 unit) (Fermentas Inc., MD, USA). The PCR was performed in an Eppendorf Mastercycler gradient PCR machine (Certified GeneTool, Inc., CA, USA). The reactions were heated at 94°C for 2 min followed by 35 cycles of 94°C 30 s, 58°C 40 s, and 72°C 40 s, and an additional 72°C 10 min was performed before the end of the PCR. The PCR products were checked by running a 0.8% agarose gel under 80 V in 1× TAE buffer. The gel was examined by a UV transilluminator (UVP Inc., CA, USA) and images were recorded photographically (AlphaEase™ version 3.25, Alpha Innotech Corporation, CA, USA).

Quantification of change of transcriptional accumulation of four dioxygenase beta-subunit genes with homologous promoters

Mycobacterium sp. strain KMS was cultured in a modified basal salt medium (BSM^+) as described by Liang et al. (2006) with or without pyrene at a final concentration 100 μM. Methods described above were used to generate first strand cDNA from isolated total RNA.

DNA products amplified from KMS genomic DNA using primers for two *nidB* genes (Mkms_1667 and Mkms_5625), two *nidB2* genes (Mkms_1660 and Mkms_5621), and 16S rRNA genes (Supplemental Table S1) were used as controls and to standardize the quantitative real time PCR (Q-PCR). The primers were designed to be specific for each of the genes through using the differences in base pair sequences that exist between the homologs as described in Zhang and Anderson (2012). As shown in Supplemental Table S1 the products generated were of different sizes for each of these genes. The specificity was confirmed by sequencing the DNA products derived from PCR with genomic DNA. Six 10-fold dilutions were made for each of the DNA products to make a concentration gradient used in the Q-PCR to create standards.

Each Q-RT-PCR reaction contained 1 μl DNA, 0.4 μl of primers (final concentration 1 μM) for the relevant gene, 10 μl 2 × master mix from DyNAmo™ HS SYBR® Green Q-PCR Kit (Invitrogen, Carlsbad, CA, USA), and 8.6 μl H_2O. The PCR was performed in a DNA Engine Opticon 2 with continuous fluorescence detection system (MJ Research, Incorporated, MA, USA).The samples were heated at 95°C for 15 min before cycling for 40 cycles of 94°C 30 s, 58°C 40 s, 72°C 1 min, 73°C 1 s. For each cDNA sample, triplicate reactions were performed. The average Ct value was plotted against the $\log_2$DNA mass and regression analysis was performed. Changes in transcript accumulation of the two *nidB* genes and two *nidB2* genes were calculated based on the method mentioned in Stratagene PCR manual (2007) by using the transcriptional accumulation of 16S rRNA genes as control.

Results

Isolate KMS utilizes both pyrene and benzoate as the sole carbon source

Isolate KMS grew on minimal medium agar plates with either 2.5 mM pyrene or 5.0 mM sodium benzoate as the sole carbon sources. Growth in liquid medium from plate-grown inocula (Fig. 1) showed the lag phase was more extended as the concentration of benzoate increased from 5 to 25 mM benzoate and no growth was observed at 50 mM. The final culture density increased as the substrate concentration was raised from 5 to 25 mM benzoate. Growth on pyrene as the sole carbon source was achieved after a longer lag phase (160 h) and then the growth rate was slower

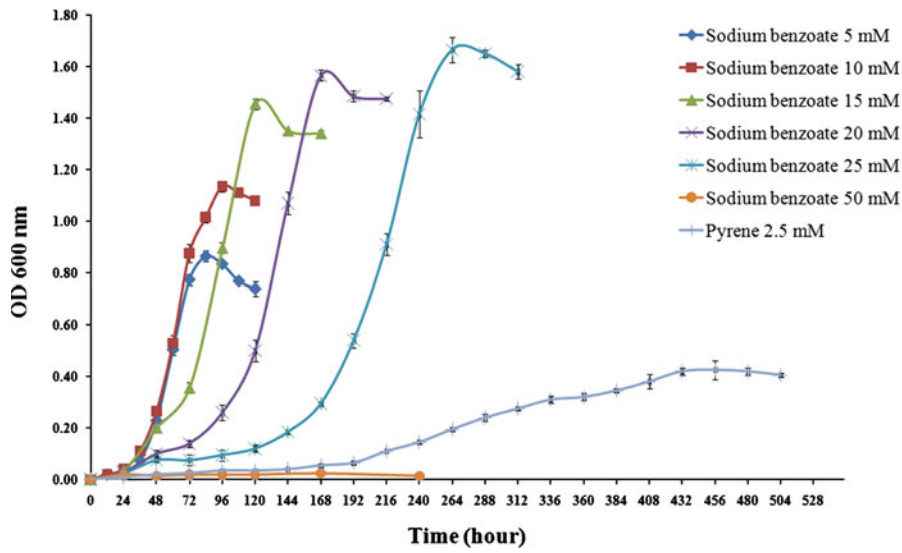

Fig. 1 Growth curves of isolate KMS grown on 5, 10, 15, 20, 25, or 50 mM sodium benzoate or 2.5 mM pyrene as sole carbon sources in liquid minimal medium (Kim et al. 2003) with shaking at 220 rpm at 25°C. Data and the standard error are the means from two replicates

than on benzoate (Fig. 1). From CFU ml^{-1} data, the doubling time of KMS grown with 5–25 mM sodium benzoate was 5.8 ± 0.6 h compared with the doubling time on 2.5 mM pyrene of 57.3 ± 11.5 h.

Distribution of genes encoding alpha- and beta-subunits of dioxygenases within the genome of *Mycobacterium* isolate KMS

Forty four genes encoding alpha- and beta-subunits of dioxygenase were detected on the chromosome, circular plasmid, and linear plasmid of isolate KMS (Fig. 2). We have used the designations *adi* and *bdi* followed by a distinguishing letter to specify the genes encoding dioxygenase alpha- and beta-subunits respectively. We examined the annotations from MicrobesOnline where domains are predicted based on the COG and IPR characterization (Tatusov et al. 2000; Mulder et al. 2005). The alpha-subunits were recognized by COG4638 (large subunits of phenylpropionate dioxygenase and ring-hydroxylating dioxygenases), or IPR001663 and the beta-subunits by COG5517 (small subunits of ring hydroxylating dioxygenases). Genes encoding alpha-subunits outnumbered the genes encoding the beta-subunits; ten large subunit genes were lone alpha-subunit genes without an adjacent beta-subunit gene. Domain recognition by MicrobesOnline revealed all of the lone alpha-subunit genes as being associated

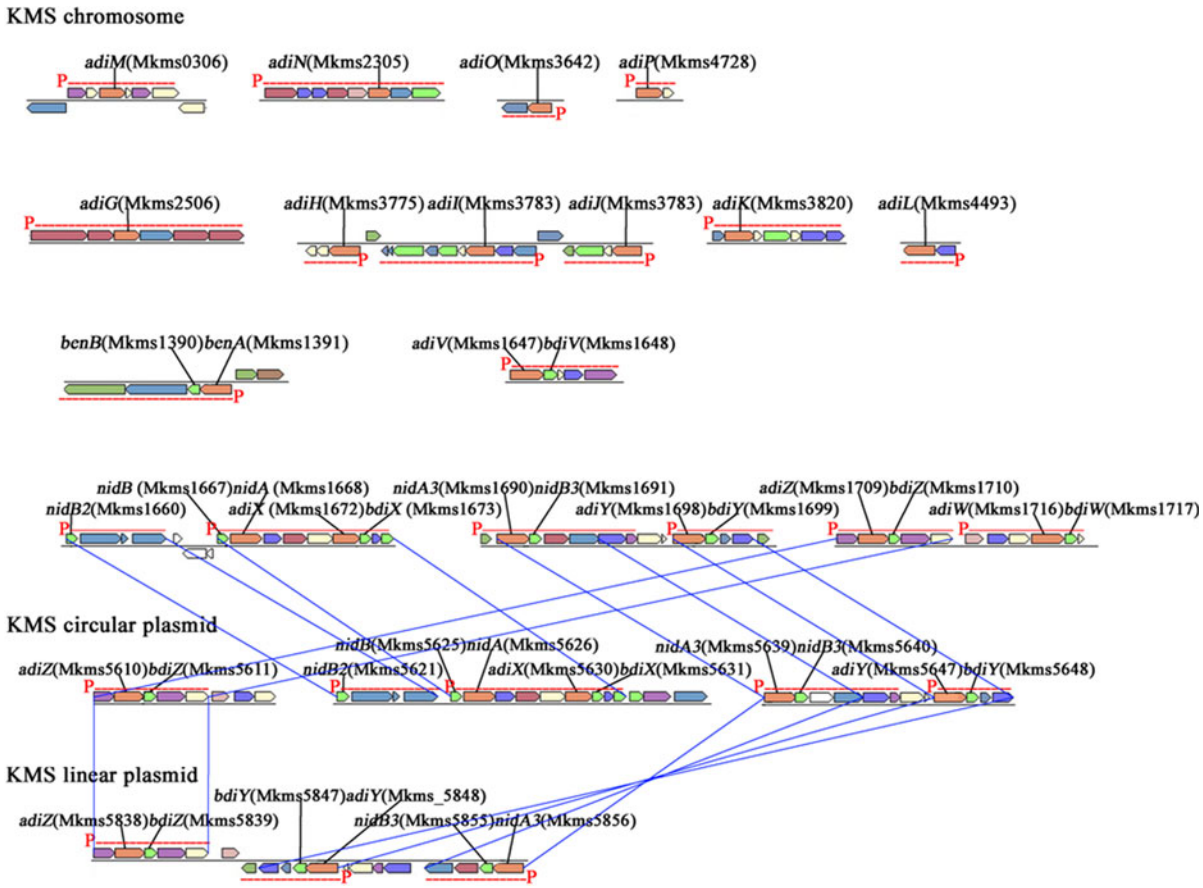

Fig. 2 The locations of the genes encoding lone ring-hydroxylating dioxygenase alpha- and beta-subunit genes and the paired ring-hydroxylating genes in the genome of *Mycobacterium* sp. isolate KMS. Designation of genes *nidA*, *nidB*, *nidB2*, *nidA3*, and *nidB3* followed those used for *M. vanbaalenii* PYR-1 (Khan et al. 2001; Kim et al. 2006, 2007) and designation of genes *benA* and *benB* followed those identified in *Rhodococcus jostii* RHA1 (Kitagawa et al. 2001). Other genes are designated as *adi* (alpha-subunits of dioxygenases) or *bdi* (beta-subunits of dioxygenases) followed by a specifying letter. The locus numbers of the gene is shown. For the on line version, the *red capitalized P* shows the position of the predicted promoter and the *red line* predicts the operon under control of that promoter. Block duplication between the chromosome and plasmids is shown by the *blue lines*. (Color figure online)

with the Rieske [2Fe-2S] iron-sulphur domain (IPR017941). Five of the genes, *adiH*, *adiJ*, *adiJ*, *adiK* and *adiL* had COG4638 as the other major conserved domain characteristic of phenylpropionate dioxygenases; a second group, *adiM*, *adiN*, *adiO* and *adoP* had the IPR001663 domain characteristic of an alpha-subunit dioxygenase. The gene *adiG* contained the domains COG4638, IPR001663, and IPR015879, characteristic of the C terminus of an aromatic ring dioxygenase alpha-subunit.

Orthologs to the *nid* genes (*nidA*, *nidB*, *nidA3*, and *nidB3*) already characterized in isolate PYR-1 by Cerniglia's group for their proven functionality in pyrene degradation (Khan et al. 2001; Kim et al. 2006) were detected in the genome of KMS. These genes were among the eight pairs of dioxygenase genes clustered on the chromosome, from Mkms_1390/1391 to Mkms_1717/1718 (Fig. 2). Only two genes at loci Mkms_1660 and 5621 encoding the beta-subunit, NidB2, were unpaired with an alpha-subunit gene. Five sets of chromosomal genes were duplicated on the linear and/or circular plasmid (Fig. 2); this group included orthologs *nidA/B*, *nidB2* and *nidA3/B3* (Table 1).

Analysis using ALIGN of the nucleotide sequences revealed the *nidB* and *nidB2* were homologs (96.9%) possibly only recently duplicated according to criteria discussed by Gevers et al. (2004). All of the other beta subunit genes have homology with *nidB* greater than 30% for over 150 amino acids characterizing them as homologs according to Gevers et al. (2004): identity of *nidB* to *bdiW* was 57.9%, and *nidB3* 55.8%; other beta-subunit genes had between 48.6 and 52.1% homology to *nidB* (Supplemental Table S2). Similarly all alpha-subunit genes were homologs: compared with the *nidA* sequence all genes had between 58% for *nidA3* and *adiW* to 47.1% homology (Supplemental Table S3). Homology among the lone alpha-subunit genes was between 47.0% (*adiH* and *adiO*) and 70.5% (*adiM* and *adiP*) (Supplemental Table S4).

BLASTp analysis predicts that other dioxygenase genes encode proteins related to utilization of an array of substrates. Genes at loci Mkms_1391 and Mkms_1390 encoded alpha- and beta-subunits of a benzoate dioxygenase respectively and, consequently, we use the designation of *benA* and *benB*. The BenA protein had 71% identity with the alpha-subunit of benzoate dioxygenase (YP_702347.1) and BenB 74%

Table 1 Distribution and location of genes encoding dioxygenase alpha- and beta-subunits in the genome of KMS

	Single copy	Double copy	Triple copy
Lone genes	Mkms_2506 (*adiG*), ch;	Mkms_1660 (*nidB2*), ch;	
	Mkms_3775 (*adiH*), ch;	Mkms_5621(*nidB2*), cp.	
	Mkms_3783 (*adiI*), ch;		
	Mkms_3790 (*adiJ*), ch;		
	Mkms_3820 (*adiK*), ch;		
	Mkms_4493 (*adiL*), ch;		
	Mkms_0306 (*adiM*), ch;		
	Mkms_2305 (*adiN*), ch;		
	Mkms_3642 (*adiO*), ch;		
	Mkms_4728 (*adiP*), ch.		
Paired genes	Mkms_1390/1391 (*benB/benA*), ch;	Mkms_1667/1668 (*nidB/nidA*), ch;	Mkms_1690/1691 (*nidA3/nidB3*), ch;
	Mkms_1647/1648 (*adiV/bdiV*), ch;	Mkms_5625/5626 (*nidB/nidA*), cp;	Mkms 5639/5640 (*nidA3/nidB3*), cp;
	Mkms_1716/1717 (*bdiW/adiW*), ch.	Mkms 1672/1673 (*adiX/bdiX*), ch;	Mkms_5855/5854 (*nidA3/nidB3*), lp;
		Mkms_5630/5631 (*adiX/bdiX*), cp.	Mkms_1698/1699 (*adiY/bdiY*), ch;
			Mkms_5647/5648 (*adiY/bdiY*), cp;
			Mkms_5847/5846 (*adiY/bdiY*), lp;
			Mkms_1709/1710 (*adiZ/bdiZ*), ch;
			Mkms_5610/5611 (*adiZ/bdiZ*), cp;
			Mkms_5838/5839 (*adiZ/bdiZ*), lp.

ch means chromosome, *cp* circular plasmid, *lp* linear plasmid

with the beta-subunit of benzoate dioxygenase of *Rhodococcus jostii* RHA1 (YP_702348.1). The protein AdiV had 95% identity to phthalate 3,4-dioxygenase alpha-subunit (YP_707370.1) from *Rhodococcus jostii* RHA1. AdiW had 90% identity with NidA or NidA3 (ADH94655.1 and ADH94645.1) encoded by *Mycobacterium sp.* strains and other naphthalene-degrading ring-hydroxylating dioxygenases in other bacteria. AdiX had 97% identity to a Rieske (2Fe-2S) domain protein (YP_001261401.1) in *Sphingomonas* and 99% identity to a Rieske (2Fe-2S) domain protein (ZP_06846381.1) in *Burkholderia* strains. AdiY had 96% identity to the alpha-subunit of dibenzofuran dioxygenase (ZP_06829928.1) in a *Terrabacter* strain and 93% identity to a type IV biphenyl 2,3-dioxygenase (YP_002777102.1) in a *Rhodococcus* strain. There was 99% identity between AdiZ and the alpha-subunits of a biphenyl 2,3-dioxygenase alpha-subunit (YP_707265.1) and a benzene dioxygenase (YP_002777102.1) in *Rhodococcus* strains. Thus, this characterization indicates that the putative dioxygenases encoded by the genes could metabolize an array of aromatic ring structures.

Block duplication

Block duplication of segments containing dioxygenase genes in isolate KMS is apparent (Fig. 2). The genes pairs designated as *a/bdiZ* and *a/bdiY* and *nidA3B3* were present in three copies, located on the chromosome, circular plasmid and linear plasmid. The two lone *nidB2* genes (one chromosomal and one plasmid) were grouped in potential operons with the *nidB/nidA* genes showing block duplication on the chromosome and the circular plasmid (Fig. 2). However, rearrangement of the block duplications are noted between the chromosome and plasmids as denoted by using blue lines in Fig. 2. Block duplication was supported by BLAST analysis showing that chromosomal and plasmid duplicated genes had 96–100% homology, with the exception of lower (90%) homology between the *bdiY* genes at loci Mkms_1699 (chromosome) and 5486 (linear plasmid) (Supplemental Table S5).

Phylogenetic analysis

The phylogenetic relationship between alpha- and beta-subunit protein sequences was determined to better understand the divergence of these genes within the single microbe (Fig. 3). The tree for the alpha-subunits (Fig. 3a, b) showed the clustering of sequences from unpaired genes encoding alpha-subunits (AdiG-P) away from the alpha-subunits that function with beta subunits. Further analysis showed that the alpha subunits in isolate KMS could be classified into the different types based on the system published by Kweon et al. (2008). This finding indicates that the proteins are associated with different electron transport systems and have distinct substrate specificities. Supplemental Fig. S2 shows the tree obtained for the KMS alpha subunit proteins using the standard proteins employed in the Type designation. There are no proteins that correspond to Type III$\alpha\beta$ or IV designations within the KMS genome. Also a subset of the proteins, the clustered lone alpha subunits AdiG/H/I/J/K/L and the clustered AdiZ and AdiY proteins have no homology to the standard proteins used for the type classification by Kweon et al. 2010. The phylogenetic relationships for the beta-subunits is simpler and supports that NidB2 and NidB subunits likely arose from duplication of a single gene.

Alignment of the predicted promoters

Sequences totaling 60 bp from 50 bp upstream from the predicted transcriptional start sites for all dioxygenase genes were aligned by ClustalX2 (Fig. 4). The genes *nidA*, *nidB*, *adiX*, *and bdiX* were predicted to be in the same operon and, thus, have the same promoter. The alignment showed that certain promoters had high identity supporting duplication of blocks of genes. There was 75–99% identity for predicted promoters of four operons containing *nidB/A* and *nidB2* genes, 92.5–97.5% identity for the triple copies of *nidA3/B3* genes, 97.5–37.5% for the copies of the *a/bdiZ* genes, and 100% homology for the promoters of the triple copies of *a/bdiY* genes. The promoters of other paired genes *benA/B*, *a/bdiV*, and *a/bdiW* or the lone alpha-subunit genes *adiG/H/I/J/K/L/M/N/O/P* lacked alignments.

Benzoate and pyrene differentially induced ring-hydroxylating dioxygenase genes

The differences between the promoter sequences suggested that they could dictate different regulatory conditions. Consequently, we examined accumulation

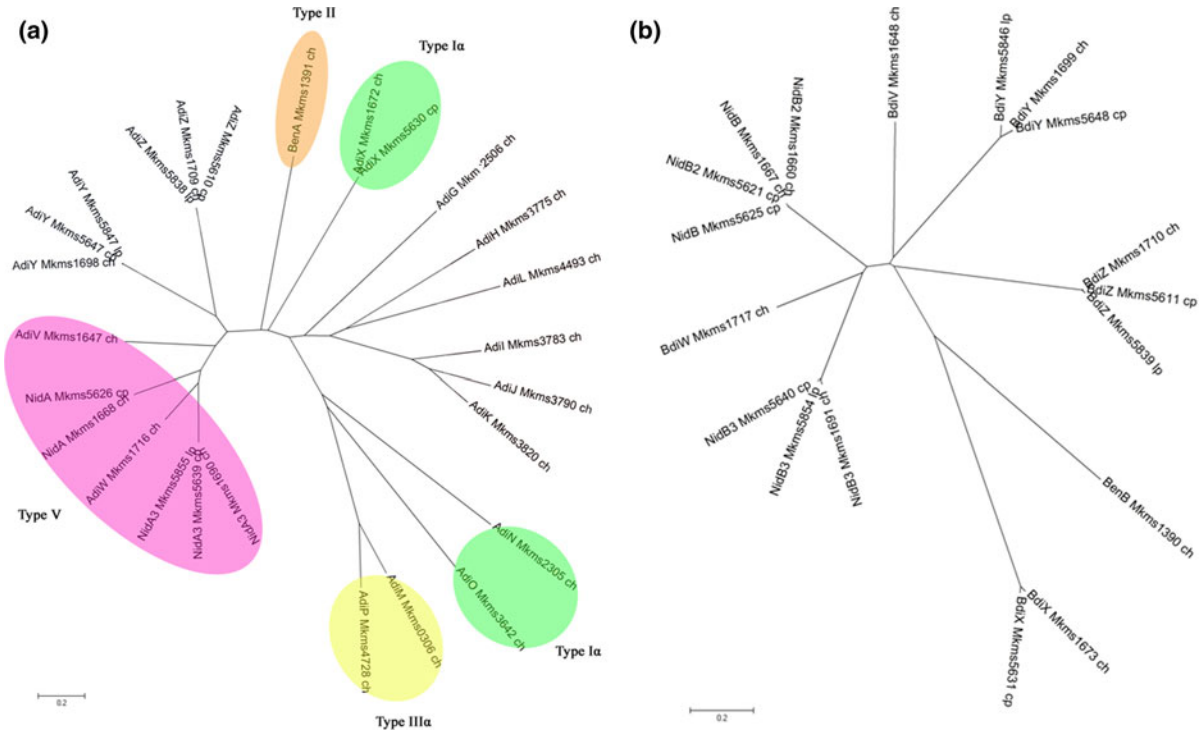

Fig. 3 Phylogenetic trees generated by MEGA 4 for the predicted protein sequences for the alpha-subunits of dioxygenases (**a**) and the beta-subunits of dioxygenase (**b**) for *Mycobacterium* isolate KMS. The tree was obtained by neighbor-joining and represents 1,000 replications. Overlayered on this tree are balloons indicating the classification of the dioxygenases into the group (Type I–V) based on the work of Kweon et al. (2008) as determined by analysis shown in Supplemental Fig S2. Designation of proteins NidA, NidB, NidB2, NidA3, and NidB3 followed those identified in *M. vanbaalenii* PYR-1 (Khan et al. 2001; Kim et al. 2006, 2007). Designation of proteins BenA and BenB was because of homology to proteins in *Rhodococcus jostii* RHA1 (Kitagawa et al. 2001). Other as yet uncharacterized aromatic ring-hydroxylating dioxygenase alpha-subunits are designated as *Adi*, and the beta-subunits as *Bdi*. The term *ch* indicates a chromosomal gene, *cp* a circular plasmid gene, and *lp* a linear plasmid genes. The locus numbers of the genes encoding the proteins are shown. The scale bar means 0.2 changes per amino acid position

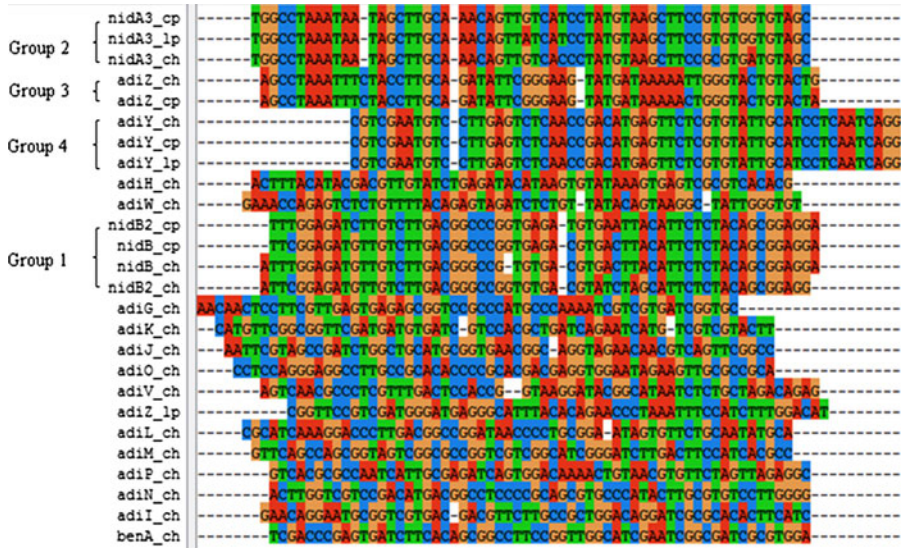

Fig. 4 Alignments of the predicted promoter regions of all the ring-hydroxylating dioxygenase genes for isolate KMS. The alignment by ClustalX2 was of 60 bp total showing 50 bp upstream from the predicted transcriptional start site. The symbol *ch* means chromosomal gene, *cp* circular plasmid gene, and *lp* linear plasmid gene. The promoters with homology are shown in groups

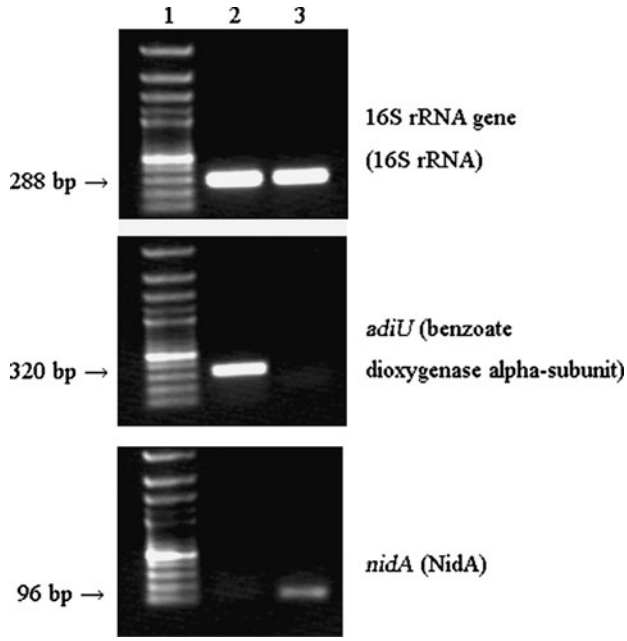

Fig. 5 Transcriptional analysis of genes *adiU* (Mkms_1391) and *nidA* (Mkms_1668) when isolate KMS was cultured on either 5 mM sodium benzoate or 2.5 mM pyrene as the sole carbon source. The expression of 16S rRNA genes was used as control. *Lane 1*, 100 bp gene ladder; *lane 2* product from cDNA of corresponding genes from cells grown on 5 mM sodium benzoate; *lane 3* product from cDNA of corresponding genes from cells grown on 2.5 mM pyrene. Genes used in this study and their products (*in brackets*) are shown. The study is from one of two studies with the same results

of transcripts from the *benA* gene, at locus Mkms_1391, predicted to encode the large subunit for a benzoate dioxygenase and the *nidA* gene, encoding the large subunit of a dioxygenase involved in transformation of pyrene. Reverse transcriptase (RT) PCR analysis revealed accumulation of transcripts from *benA* occurred in cells cultured on benzoate but not pyrene. In contrast the *nidA* gene, from locus Mkms_1668, was expressed during pyrene growth but not during growth on benzoate. The PCR products from these analyses are shown along with PCR products from the 16S rRNA genes used as a control for equal loading of samples in Fig. 5.

Dioxygenase genes with homologous promoters were similarly regulated

The homology between predicted promoters for both the chromosomal and plasmid copies of the *nidB* and *nidB2* genes suggested that these genes would be coordinately expressed. This possibility was tested using Q-RT-PCR to assess the RNA accumulating from these genes during growth on pyrene. The fold increases in expression for each gene, normalized based upon expression of the 16S rRNA genes, had a mean fold increase of 3-fold and, thus, were of similar magnitudes (Table 2).

Table 2 The fold change in expression for genes encoding the beta-subunit of dioxygenase normalized by the expression of 16S rRNA genes with growth of isolate KMS on pyrene

Selected gene and location	Fold change based on the mass of cDNA[a]
Mkms_1660 (*nidB2*) Chromosome	3.3 ± 1.5
Mkms_1667 (*nidB*) Chromosome	3.0 ± 1.0
Mkms_5621 (*nidB2*) Plasmid	2.8 ± 1.1
Mkms_5625 (*nidB*) Plasmid	3.2 ± 0.9

[a] Data were normalized based on the expression of the 16S rRNA genes and standard deviations were calculated based on results from four separate Q-RT-PCR studies

Discussion

The pyrene-degrading *Mycobacterium* isolate KMS had a plethora of genes encoding ring-hydroxylating dioxygenase subunits that were clustered on the chromosome as well as on the circular and linear plasmids. There is evidence of block duplication for the operons containing the genes that based on observations of Cerniglia's group initiated ring

oxidation in the high molecular weight PAHs (Khan et al. 2001; Kim et al. 2006; Kweon et al. 2010). Duplication of genes leading to the existence of families with related function is common in bacteria (Gevers et al. 2004; Serres et al. 2009). These events are postulated to permit adaption to different environments. Andersson and Hughes (2009) posited "Whenever cellular growth is restricted escape from these growth restrictions often occurs by gene duplication-amplification events that resolve the selective problem". They provide examples with *Alcaligenes*, *Pseudomonas* and *Acinetobacter* isolates where gene duplication relates to adapted utilization of contaminants such as alkanes and benzoate. Similarly, Gevers et al. (2004) stated that "gene duplication is an evolutionary response in bacteria exposed to different selection pressures such as starvation". The ability of certain environmental *Mycobacterium* isolates, such as KMS which came from PAH-contaminated soils, to develop the ability to utilize the recalcitrant PAH, pyrene, fits well within this concept of gene duplication to overcome starvation and promote adaption to novel carbon sources. Supporting evidence comes from comparison of the KMS genome with the genomes of pathogenic, *M. tuberculosis.* Dioxygenase genes involved in housekeeping pathways (taurine catabolism, breakdown of nitro-organics) exist in both environmental and pathogenic isolates (Eichhorn et al. 2000; Stiens et al. 2006). However the pathogenic isolates possess limited genes encoding enzymes for modification of complex aromatics (e.g. two genes for large subunit phenylpropionate dioxygenase in *M. tuberculosis* H37Ra among the 25 predicted dioxygenase genes with general metabolic activities). In contrast there is an extensive ring-hydroxylating dioxygenase gene family with specificity for the high molecular PAHs in KMS. This finding is seen also in two other *Mycobacterium* isolates, MCS and JCS, from the same Montana site as KMS, as well as two isolates from other PAH-contaminated sources, *M. vanbaalennii* PYR-1 and M. *gilvum* PYR-GCK (Kim et al. 2008; Zhang and Anderson 2012). Additional genes with the COG4638 signature for phenylpropionate alpha-subunits are found in other *Mycobacterium* isolates, such as the pathogenic *M. marinum*, *M. avium* and environmental *M. smegmatis* according analysis with MicrobesOnline Treebrowser. Thus there is circumstantial evidence that the expanded gene family offers an advantage when the microbes are forced to utilize the multi-aromatic ring PAHs.

In many cases, members of the duplicated genes are modified further (e.g. by mutation or acquisition of novel domains) such that there is clear divergence of function (Iwabe et al. 1989; Notebaart et al. 2005). Indeed extensive superfamilies have been annotated within many bacterial genomes (Serres et al. 2009). For example the genes in the crotonase family in *Escherichia coli* encode for enzymes exhibiting a range of distinct activities (Serres et al. 2009). Clearly this family of dioxygenase genes has not diversified to such a large extent as they retained high homology.

The existence of some of the amplified genes on plasmids in isolate KMS is consistent with similar observations in other microbes. Andersson and Hughes (2009) cited the extensively studied cases of amplification of genes for drug resistance on plasmids. In KMS, the circular and the linear plasmids have few genes with predicted functions and lack genes predicted to be associated with drug resistance. The circular plasmid has genes involved in resistance towards cadmium and copper, suggesting that genes involved in other needed environmental adaptations are present; these metal-resistance genes were not duplicated on the chromosome.

BLASTp analysis of the ring hydroxylating dioxygenase genes revealed homology with genes encoding proteins with assayed functionality. For instance, the additional paired genes we designated as V, W, X, Y, and Z genes in isolate KMS were predicted to encode the alpha and beta subunits of dioxygenase acting on small ring structures (e.g. V, phthalate; W, naphthalene; Y and Z, biphenyl). In this paper we correlated expression from the *benA* gene at locus Mkms_1391 encoding a dioxygenase alpha-subunit with benzoate degradation because transcripts from this locus accumulated when isolate KMS was grown on benzoate. The genes adjacent to this locus are predicted to encode other proteins with functions in benzoate degradation: Mkms_1388 encoded a regulator protein with 97% identity to a LuxP family transcriptional regulator, Mkms_1390 *benB* encoding the beta subunit of the dioxygenase and Mkms_1389 encoded a chimeric protein of a benzoate dioxygenase reductase (BenC) and a diol dehydrogenase (BenD) as shown in Supplemental Fig. S3. This benzoate-degrading gene cluster in KMS differed from two other types of clusters endowing benzoate catabolism

(Supplemental Fig. S3); the operons in strains of a pseudomonad and a *Rhodococcus* species contained five genes (Cowles et al. 2000; Kitagawa et al. 2001) whereas in *Acinetobacter* there were eight genes (Collier et al. 1997). Identical sequences to that in strain KMS were found in four other PAH-degrading mycobacterium strains MCS, JLS, *M. gilvum* PYR-GCK, and Spyr1 (Supplemental Fig. S4) but *M. vanbaalenii* PYR-1 lacked such genes for benzoate catabolism as indicated in Kim et al. (2008). These strains are from diverse geographical sites showing conservation of function.

Our studies illustrated regulated expression from the *benA* gene by benzoate versus pyrene. Similarly the *nidA* gene was not expressed during growth on benzoate but was induced by pyrene. Because the clusters containing *nidA* and *benA* have distinct promoters, these findings are consistent with differential activation of the promoters by aromatic ring compounds. This concept was supported by observing during pyrene growth the very similar levels of activation from the promoters that have high sequence identity of four *nidB* genes, whether these genes were located on the chromosome or plasmid.

Our findings illustrate the adaption of *Mycobacterium sp.* KMS to maximize its effectiveness in using aromatic ring structures. Adaption appears to have evolved gene duplication with divergence of the subunits of the dioxygenases towards different substrate specificity as well as the development of distinct promoter sequences. We demonstrated expression simultaneously from plasmid and chromosomal copies of a duplicated gene and we show functionality of two distinct promoter sequences.

Acknowledgments We acknowledge financial support for this study from the National Science Foundation and from the Utah Agricultural Experiment Station. We thank Dr. Marie Walsh and Dr. Jeff Broadbent for the use of the Q-PCR instrumentation.

References

Andersson DI, Hughes D (2009) Gene amplification and adaptive evolution in bacteria. Annu Rev Genet 43:167–195

Butler CS, Mason JR (1997) Structure–function analysis of the bacterial aromatic ring-hydroxylating dioxygenases. Adv Microb Physiol 38:47–84

Collier LS, Nichols NN, Neidle EL (1997) *benK* encodes a hydrophobic permease-like protein involved in benzoate degradation by *Acinetobacter sp*. strain ADP1. J Bacteriol 179:5943–5946

Correll CC, Ludwig ML, Bruns CM, Karplus PA (1993) Structural prototypes for an extended family of flavoprotein reductases: comparison of phthalate dioxygenase reductase with ferredoxin reductase and ferredoxin. Protein Sci 2:2112–2133

Cowles CE, Nichols NN, Harwood CS (2000) BenR, a XylS homologue, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. J Bacteriol 182:6339–6346

Eichhorn E, van der Ploeg JR, Leisinger T (2000) Deletion analysis of the *Escherichia coli* taurine and alkanesulfonate transport systems. J Bacteriol 182:2687–2795

Gevers D, Vandepoele K, Simillon C, Van de Peer Y (2004) Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Microbiol 12:148–154

Gu J, Neary J, Cai H, Moshfeghian A, Rodriguez SA, Lilburn TG, Wang Y (2009) Genomic and systems evolution in *Vibrionaceae* species. BMC Genomics 10:S11

Hurtubise Y, Barriault D, Sylvestre M (1998) Involvement of the terminal oxygenase beta subunit in the biphenyl dioxygenase reactivity pattern toward chlorobiphenyls. J Bacteriol 180:5828–5835

Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci 86:9355–9359

Jiang H, Parales RE, Gibson DT (1999) The alpha subunit of toluene dioxygenase from *Pseudomonas putida* F1 can accept electrons from reduced ferredoxin$_{TOL}$ but is catalytically inactive in the absence of the beta subunit. Appl Environ Microbiol 65:315–318

Karabika E, Kallimanis A, Dados A, Pilidis G, Drainas C, Koukkou AI (2009) Taxonomic identification and use of free and entrapped cells of a new *Mycobacterium* sp., strain Spyr1 for degradation of polycyclic aromatic hydrocarbons (PAHs). Appl Biochem Biotechnol 159:155–167

Khan AA, Wang RF, Cao WW, Doerge DR, Wennerstrom D, Cerniglia CE (2001) Molecular cloning, nucleotide sequence, and expression of genes encoding a polycyclic aromatic ring dioxygenase from *Mycobacterium* sp. strain PYR-1. Appl Environ Microb 67:3577–3585

Kim YH, Cerniglia CE, Engesser KH (2003) Two polycyclic aromatic hydrocarbon *o*-quinone reductases from a pyrene-degrading *Mycobacterium*. Arch Biochem Biophys 416: 209–217

Kim SJ, Kweon O, Freeman JP, Jones RC, Adjei MD, Jhoo JW, Edmondson RD, Cerniglia CE (2006) Molecular cloning and expression of genes encoding a novel dioxygenase involved in low- and high-molecular-weight polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Appl Environ Microb 72:1045–1054

Kim SJ, Kweon O, Jones RC, Freeman JP, Edmondson RD, Cerniglia CE (2007) Complete and integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on systems biology. J Bacteriol 189:464–472

Kim SJ, Kweon O, Jones RC, Edmondson RD, Cerniglia CE (2008) Genomic analysis of polycyclic aromatic

hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Biodegradation 19:859–881

Kitagawa W, Miyauchi K, Masai E, Fukuda M (2001) Cloning and characterization of benzoate catabolic genes in the gram-positive polychlorinated biphenyl degrader *Rhodococcus sp*. strain RHA1. J Bacteriol 183:6598–6606

Kweon O, Kim SJ, Baek S, Chae JC, Adjei MD, Baek DH, Kim YC, Cerniglia CE (2008) A new classification system for bacterial Rieske non-heme iron aromatic ring-hydroxylating oxygenases. BMC Biochem 9:11

Kweon O, Kim SJ, Jones RC, Freeman JP, Song J, Baek S, Cerniglia CE (2010) Substrate specificity and structural characteristics of the novel rieske nonheme iron aromatic ring-hydroxylating oxygenases NidAB and NidA3B3 from *Mycobacterium vanbaalenii* PYR-1. MBio 15:e 00135-10

Liang Y, Gardner DR, Miller CD, Chen D, Anderson AJ, Weimer BC, Sims RC (2006) Study of biochemical pathways and enzymes involved in pyrene degradation by *Mycobacterium sp.* strain KMS. Appl Environ Microb 72:7821–7828

Miller CD, Hall K, Liang YN, Nieman K, Sorensen D, Issa B, Anderson AJ, Sims RC (2004) Isolation and characterization of polycyclic aromatic hydrocarbon-degrading *Mycobacterium* isolates from soil. Microb Ecol 48:230–238

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bradley P, Bork P, Bucher P, Cerutti L, Copley R, Courcelle E, Das U, Durbin R, Fleischmann W, Gough J, Haft D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lonsdale D, Lopez R, Letunic I, Madera M, Maslen J, McDowall J, Mitchell A, Nikolskaya AN, Orchard S, Pagni M, Ponting CP, Quevillon E, Selengut J, Sigrist CJ, Silventoinen V, Studholme DJ, Vaughan R, Wu CH (2005) InterPro, progress and status in 2005. Nucleic Acids Res 33:D201–D205

Notebaart RA, Huynen MA, Teusink B, Siezen J, Snel B (2005) Correlation between sequence conservation and the genomic context after gene duplication. Nucleic Acids Res 33:6164–6171

Orlova I, Marshall-Colón A, Schnepp J, Wood B, Varbanova M, Fridman E, Blakeslee JJ, Peer WA, Murphy AS, Rhodes D, Pichersky E, Dudareva N (2006) Reduction of benzenoid synthesis in petunia flowers reveals multiple pathways to benzoic acid and enhancement in auxin transport. Plant Cell 18:3458–3475

Parales RE, Resnick SM (2006) Aromatic ring hydroxylating dioxygenases. Pseudomonas 4:287–340

Peng RH, Xiong AS, Xue Y, Fu XY, Gao F, Zhao W, Tian YS, Yao QH (2010) A profile of ring-hydroxylating oxygenases that degrade aromatic pollutants. Rev Environ Contam Toxicol 206:65–94

Reese MG (2001) Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput Chem 26:51–56

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B (2000) Artemis: sequence visualization and annotation. Bioinformatics 16:944–945

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Serres MH, Kerr AR, McCormack TJ, Riley M (2009) Evolution by leaps: gene duplication in bacteria. Boil Direct 4:46

Stiens M, Schneiker S, Keller M, Kuhn S, Pühler A, Schlüter A (2006) Sequence analysis of the 144-Kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. Appl Environ Microbiol 72:3662–3672

Stratagene (2007) Introduction to quantitative PCR Methods and application guide, Stratagene La Jolla, San Digeo

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tannheimer SL, Barton SL, Ethier SP, Burchiel SW (1997) Carcinogenic polycyclic aromatic hydrocarbons increase intracellular Ca^{2+} and cell proliferation in primary human mammary Epithelial cells. Carcinogenesis 18:1177–1182

Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36

U.S. Department Of Health And Human Services (1995) Toxicological profile for polycyclic aromatic hydrocarbons. p 19

Zengin N, Yüzbaşıoğlu D, Unal F, Yılmaz S, Aksoy H (2010) The evaluation of the genotoxicity of two food preservatives: sodium benzoate and potassium benzoate. Food Chem Toxicol. doi:10.1016/j.fct.2010.11.0

Zhang C, Anderson AJ (2012) Polycyclic aromatic hydrocarbon-degrading gene islands in five pyrene-degrading *Mycobacterium* isolates from different geographic locations. Can J Microbiol 58(1):102–111