Heterologous expression of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes from a novel pyrene-degrading Betaproteobacterium

David R. Singleton[1], Jing Hu[1], Michael D. Aitken[1]

[1]Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599-7431, USA

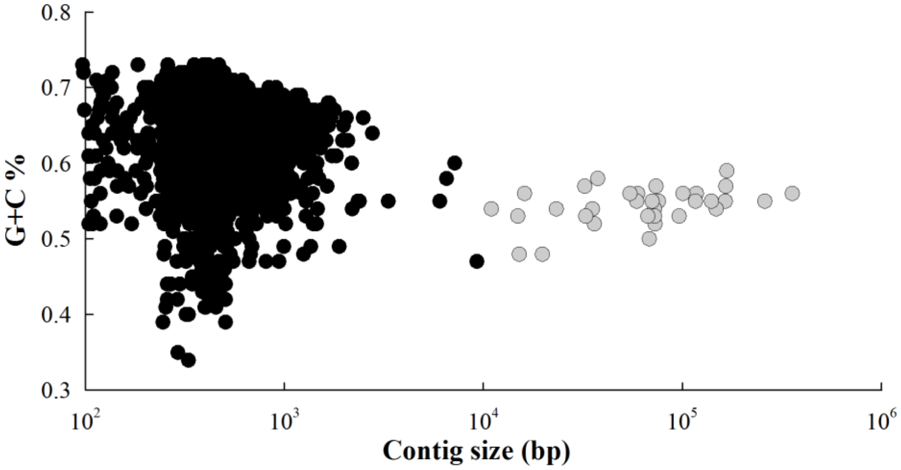

Figure S1. Contig size versus G+C% for the metagenomic library. Contigs putatively assigned to PG1 (>10kbp) are light grey circles. All other contigs are black circles.

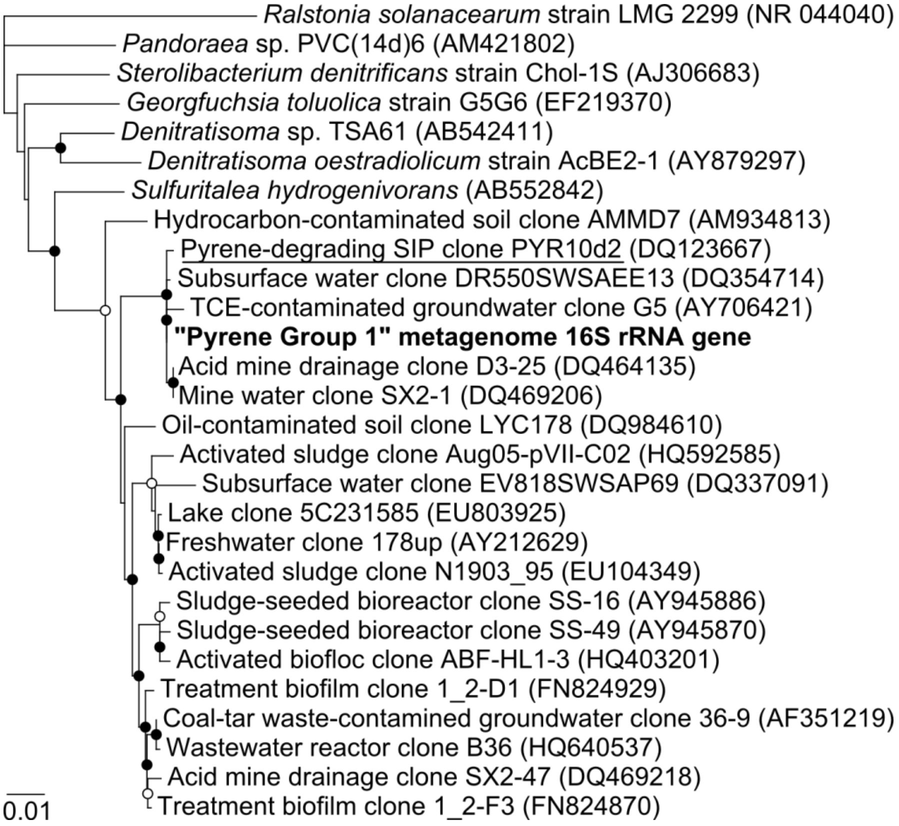

Figure S2. Neighbor-joining 16S rRNA gene tree showing Betaproteobacterial sequences with similarity to the PG1 sequence recovered from the metagenomic library (bold). The PG1 representative sequence from the prior SIP experiment with [U-^{13}C]pyrene is underlined. Open and closed circles at nodes indicate ≥ 95% and ≥ 50% bootstrap support, respectively. GenBank accession numbers are indicated in parentheses. *Ralstonia solanacearum* was used as an outgroup.

Table S1. PCR primers used during this study

Name	Sequence (5' to 3')[a]	Source	R.E.[b]	Target
341FGC	[GC clamp] CCT ACG GGA GGC AGC AG	Muyzer et al., 1993[c]	-	Bacterial 16S rRNA gene for DGGE
517R	ATT ACC GCG GCT GCT GG		-	
Ferre_F	GAG AGG ATC CTA GAG AAG GAA ATT GAT GTA TGA ATA CTC GTG TAA AAG	This study	BamHI	ORF_129290
Ferre_R	TCT CAG ATC TTT ACG CTT CCA CTT CCG CCT C		BglIII	Ferredoxin
Reduc_F	GAG AAG ATC TAG AAG GAT CCT GAT TTA TGG AAA CTG GTT TTT CGA TC	This study	BglIII	ORF_129230
Reduc_R	TCT CCT GCA GTT AAT AAA AAC GGT CGA AAT GC		PstI	Reductase
RHD-1_F	GAG ACT GCA GAG AAG GAA GCC CTA TCA TGT TAA ATA TTG AAA GTC TG	This study	PstI	ORF_129020, ORF_128980
RHD-1_R	CTC TCC ATG GTT ACA ACA TGA AGT TGA GAT TTT TAC C		NcoI	RHD α/β subunits
RHD-2_F	GAG ACT GCA GAG AAG AAG GAG TTG ATA TGA CAA ACG TGA GTT CGT TA	This study	PstI	ORF_11620, ORF_11630
RHD-2_R	CTC TCC ATG GTT AAA ATA AAG TGT TCA TGT TGC TAT C		NcoI	RHD α/β subunits
RHD-3_F	GAG ACT GCA GAG AAG AAG AGA GTG AAA TGG TCG ATG TAA ATA GTC TG	This study	PstI	ORF_11730, ORF_11740
RHD-3_R	CTC TCC ATG GCT ACA TCG GGA AGT TAT GGT TCT TGC C		NcoI	RHD α/β subunits
RHD-4_F	GAG ACT GCA GAG AAG AAG GAG TTG GCA TGA ATA AGC TTG AAG GCT TG	This study	PstI	ORF_92190, ORF_92180
RHD-4_R	CTC TCC ATG GTT AAA AAA ATG TGT TCA TGT TGC TGT C		NcoI	RHD α/β subunits
RHD-5_F	GAG ACT GCA GAG AAG AAA GGT GAA TAA TGA CAA GCT TTA ACT ATC AG	This study	PstI	ORF_117570, ORF_117560
RHD-5_R	CTC TCC ATG GTC AGA AAA ACA TGT TCA GAT TTT TAT C		NcoI	RHD α/β subunits
RHD-6_F	GAG ACT GCA GAG AAG AAG ATA CTG CAA TGA AAA TTG AAA AAT TCA AG	This study	PstI	ORF_124010, ORF_124000
RHD-6_R	CTC TCC ATG GCT ACA AAT ATA TGT TGA GGT TTT TGC T		NcoI	RHD α/β subunits
RHD-7_F	GAG ACT GCA GAG AAG AAA TTA CAC ATA TGA CTT CGC TGA AGC CAA CC	This study	PstI	ORF_129790, ORF_129800
RHD-7_R	CTC TCC ATG GTT AGA GAA TGC TGG CCA TAT TGG GCA T		NcoI	RHD α/β subunits
RHD-8_F	GAG ACT GCA GAG AAG AAA GGC GGT GCA TGG GTG ATT TCG ACG CGT TA	This study	PstI	ORF_129740, ORF_129750
RHD-8_R	CTC TCC ATG GTT ACA GGA TAA ACA ACA AAT TTT TTC C		NcoI	RHD α/β subunits

[a] Underlined regions indicate restriction sites

[b] Restriction enzyme targeting the indicated site

[c] Muyzer, G., E. C. de Waal, and A. G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59:695-700.

Table S2. Mismatches of PG1 RHD α-subunit nucleotide sequences to recommended PCR primer pairs

For comprehensive targeting of dioxygenase clades (Iwai et al., 2011. Appl. Environ. Microbiol. 77:3551-7)

DP1, Rieske_f

	T G Y M G N C A Y M G N G G	*mismatches*
PG1_pahAc1	– – – G C – – – – – – – – –	2
PG1_pahAc2	– – – – C – – – – – – – – –	1
PG1_pahAc3	– – – – C – – – – – – – – –	1
PG1_pahAc4	– – – – C – – – – – – – – –	1
PG1_pahAc5	– – – – C – – – – – – – – –	1
PG1_pahAc6	– – – – C – – – – – – – – –	1
PG1_pahAc7	– – – – C – – – – – – – – –	1
PG1_pahAc8	– – – T C – – – – – – – – –	2
PpG7_nahAc	– – – – – – – – – – – – – –	0
Mv_nidA	– – – – – – – – – – – – – –	0

ARHD2R

	A A N T K Y T C N G C N G S N R M Y T T C C A	*mismatches*
PG1_pahAc1	– – – – – – – – – – – – – – – – T – – – – – –	1
PG1_pahAc2	– C – – – – – – – A – – – A – – – – – – – – –	3
PG1_pahAc3	– – – – – – – – – A – – – – – – T – – – – – –	2
PG1_pahAc4	– C – – – – – – – A – – – – – – G – – – – – –	3
PG1_pahAc5	G C – – – – – – – A – – – – – – – – – – – – –	3
PG1_pahAc6	– – – – – – – – – A – – C – – – – – – – – – –	2
PG1_pahAc7	– T – – – – – – – – – – – – – – G – – – – – –	2
PG1_pahAc8	– – – – – – – – – C – – – – – – T – – – – – –	2
PpG7_nahAc	– –	0
Mv_nidA	– – – – – G – – – – – – – – – – T – – – – – –	2

For determination of a Rieske Center (Iwai et al., 2011. Appl. Environ. Microbiol. 77:3551-7)

DP1, Rieske_f

	T G Y M G N C A Y M G N G G	*mismatches*
PG1_pahAc1	– – – G C – – – – – – – – –	2
PG1_pahAc2	– – – – C – – – – – – – – –	1
PG1_pahAc3	– – – – C – – – – – – – – –	1
PG1_pahAc4	– – – – C – – – – – – – – –	1
PG1_pahAc5	– – – – C – – – – – – – – –	1
PG1_pahAc6	– – – – C – – – – – – – – –	1
PG1_pahAc7	– – – – C – – – – – – – – –	1
PG1_pahAc8	– – – T C – – – – – – – – –	2
PpG7_nahAc	– – – – – – – – – – – – – –	0
Mv_nidA	– – – – – – – – – – – – – –	0

Nah-for

	T G C M V N T A Y C A Y G G Y T G G	*mismatches*
PG1_pahAc1	– – T – – – – – – – – – – – – – – –	1
PG1_pahAc2	– – – – – – – – – – – – – – – – – –	0
PG1_pahAc3	– – T – – – – – – – – – – – – – – –	1
PG1_pahAc4	– – T – – – – – – – – – – – – – – –	1
PG1_pahAc5	– – – T – – – – – – – – – – – – – –	1
PG1_pahAc6	– – – – – – – – – – – – – – – – – –	0
PG1_pahAc7	– – – – – – – – – – – – – – – – – –	0
PG1_pahAc8	– – – – – – – – – – – – – – – – – –	0
PpG7_nahAc	– – – – – – – – – – – – – – – – – –	0
Mv_nidA	– – – – – – – – – – – – – – – – – –	0

For PAH-Gram Negative-type dioxygenases (Iwai et al., 2011. Appl. Environ. Microbiol. 77:3551-7)

Ac596R

	C R G G T G Y C T T C C A G T T G	*mismatches*
PG1_pahAc1	– C – – – – – T – – – – – – – – –	2
PG1_pahAc2	– – – A C A A – – – – – – – – – –	4
PG1_pahAc3	– C – – G – – T – – – – – – – – –	3
PG1_pahAc4	– C – – C A G T – – – – – – – – –	5
PG1_pahAc5	– – – – C A A T – – – – – – – – –	4
PG1_pahAc6	– – C C G A A – – – – – – – – – A	6
PG1_pahAc7	– C – C A A G – – – – – – – – – –	5
PG1_pahAc8	– C – – C – – T – – – – – – – – –	3
PpG7_nahAc	– – – – – – – – – – – – – – – – –	0
Mv_nidA	– – – C – – – – – – – – – – – – C	2

NAPH-2R

	A T S T C T T T T T C B A C	*mismatches*
PG1_pahAc1	– – – – T C – – – – – A – –	3
PG1_pahAc2	– – – G – C – – – – – A G –	4
PG1_pahAc3	– – – – T C – – – – – A – –	3
PG1_pahAc4	– – – G – – – – – – – – – –	1
PG1_pahAc5	– – A – – – C G C – – A C A	7
PG1_pahAc6	– – A – T C – – G – T – – –	5
PG1_pahAc7	– A T – – – – – G – – – – –	3
PG1_pahAc8	– – A – – – – – – – – – – –	1
PpG7_nahAc	– – – – – – – – – – – – – –	0
Mv_nidA	G C A – – G C G – – – – – –	6

For PAH-Gram Positive-type dioxygenases (Iwai et al., 2011. Appl. Environ. Microbiol. 77:3551-7)

nidA-forward

	ATGACCACCGAAACAACCGGAACAGC
PG1_pahAc1	
PG1_pahAc2	
PG1_pahAc3	
PG1_pahAc4	
PG1_pahAc5	
PG1_pahAc6	–TGAA––TT–A
PG1_pahAc7	A
PG1_pahAc8	
PpG7_nahAc	
Mv_nidA	––––––––––––––––––––––––––

Blanks indicate no equivalent nucleotide position in those genes

pdo1-r

	CTGACCCATGTATTCCAGCC
PG1_pahAc1	TCTG ––––––GC–AT
PG1_pahAc2	GCAT ––GC––TT–––
PG1_pahAc3	TCAG AA–C––GC–AT
PG1_pahAc4	GCAT ––G–––TT–––
PG1_pahAc5	–CAT CGCA–Γ(C–––
PG1_pahAc6	TCCG– A–TCAT CAAG
PG1_pahAc7	–CTG– A–G–AT TCGG
PG1_pahAc8	–CAG– A–T–C– T–GT
PpG7_nahAc	––––– ––G–G––T–GT
Mv_nidA	––––––––––––––––––––

Blanks indicate no equivalent nucleotide position in those genes

For PAH-Gram negative-type dioxygenases (Cébron et al., 2008. J. Microbiol. Methods 73:148-59)

PAH-RHDα GN F

	GAGATGCATACCACGTKGGTTGGA	*mism.*
PG1_pahAc1	–C–––––C––––––––C––C––––	4
PG1_pahAc2	–C––C–GC––––––C––––––––T	6
PG1_pahAc3	–T––C––C–––––T–––––C––––	5
PG1_pahAc4	–C––––GC–––––TT–A––C–––T	8
PG1_pahAc5	–G––C–GC––––––––C–CC–––G	8
PG1_pahAc6	–T–––––C––T––––CC–––––––	5
PG1_pahAc7	CT–––A–T––T––T–––TT–GCTG	12
PG1_pahAc8	–C–––ATG––T––––––––C––––	6
PpG7_nahAc	–G––––––––––––––––––––––	1
Mv_nidA	–C––CT–C––––––ACCCTC–TTG	13

PAH-RHDα GN R

	AGCTGTTGTTCGGGAAGAYWGTGCMGTT	*mism.*
PG1_pahAc1	––GA–––A––A–––––T––G––CG–A––	9
PG1_pahAc2	CTGCA A–C–G–––––A––C––CA––––	12
PG1_pahAc3	––GA–––A–––––A––A––C–ACG–A––	10
PG1_pahAc4	CTGCA G–C–G–––––C––––––A––––	10
PG1_pahAc5	TTGCA G–––G–––––C––GCCAATA–G	16
PG1_pahAc6	–CGA–ACA––––––––A––C–ACG––GA	13
PG1_pahAc7	–AGAA AA––––––––C––CCC–––––C	11
PG1_pahAc8	–TGAA–––––G–––––G––C–––––A–C	9
PpG7_nahAc	T–––––––––––––––A––G––––––––	3
Mv_nidA	T–TG––CC–G––C––G–G–CAC–TT–A–	15

Blanks indicate no equivalent nucleotide position in those genes

PpG7_nahAc	:	*Pseudomonas putida* G7, nahAc gene
Mv_nidA	:	*Mycobacterium vanbaalenii* PYR-1, nidA gene

Table S3. Proteins from GenBank with highest similarity to predicted RHD protein sequences and other ORFs

Contig	ORF	Figure Label[a]	Strongest BLASTP Match	GenBank Accession	% a.a. Identity	% a.a. Similarity	Gene[b]
00018	11620	A	*Novosphingobium* sp. PP1Y; RHD α-subunit	CCA92463	72	84	*pahAc2*
	11630	B	*Sphingobium yanoikuyae* ; RHD β-subunit	ABM79808	56	71	*pahAd2*
	11640	C	*Burkholderia* sp. Ch1-1; Rieske (2Fe-2S) iron-sulfur domain protein	EFG65990	67	81	
	11650	D	*Pseudoxanthomonas spadix* BD-a59; small subunit aromatic oxygenase	EHB93918	54	73	
	11660	E	*Pseudoxanthomonas spadix* BD-a59; small subunit aromatic oxygenase	EHB93991	53	72	
	11670	F	*Sphingomonas* sp. CHY-1; salicylate 1-hydroxylase α-subunit	CAG17582	63	77	
	11680	G	Rhizobiales bacterium HF4000_48A13; phenylpropionate dioxygenase α-subunit	ADI19681	55	71	
	11700	H	*Novosphingobium* sp. PP1Y; aromatic-ring-hydroxylating dioxygenase β-subunit	CCA92441	57	73	
	11710	I	*Alteromonas* sp. SN2; hypothetical protein ambt_17905	AEF05080	37	58	
	11720	J	*Alteromonas* sp. SN2; protocatechuate 4,5-dioxygenase β-subunit	AEF05079	52	73	
	11730	K	*Burkholderia* sp. DBT1; DbtAc	AAK62353	65	82	*pahAc3*
	11740	L	*Burkholderia* sp. DBT1; DbtAd	AAK62354	50	68	*pahAd3*
05319	92200	A	*Thauera* sp. MZ1T; short-chain dehydrogenase/reductase SDR	ACR01697	49	61	
	92190	B	*Novosphingobium* sp. PP1Y; RHD α-subunit	CCA92463	74	85	*pahAc4*
	92180	C	*Sphingobium yanoikuyae* ; RHD β-subunit	ABM79808	58	73	*pahAd4*
	92170	D	*Burkholderia* sp. Ch1-1; Rieske (2Fe-2S) iron-sulfur domain protein	EFG65990	70	82	
	92160	E	*Pseudoxanthomonas spadix* BD-a59; small subunit aromatic oxygenase	EHB93918	52	71	
	92150	F	*Pseudoxanthomonas spadix* BD-a59; small subunit aromatic oxygenase	EHB93991	55	75	
	92140	G	*Sphingomonas* sp. CHY-1; salicylate 1-hydroxylase α-subunit	CAG17582	64	80	
	92130	H	*Acinetobacter baumannii* AB0057; RHD Rieske	ACJ41568	41	57	
	92120	I	*Sphingomonas wittichii* RW1; aromatic RHD, β-subunit	ABQ68235	39	53	
	92110	J	*Lutiella nitroferrum* 2002; carboxymuconolactone decarboxylase	EEG08881	36	57	
	92100	K	*Halanaerobium hydrogeniformans* ; hypothetical protein Halsa_0924	ADQ14370	27	43	
05407	118500	A	*Ralstonia pickettii* 12D; sigma54 specific transcriptional regulator Fis family	ACS64822	48	65	
	117610	B	*Burkholderia* sp. Ch1-1; Rieske (2Fe-2S) iron-sulfur domain protein	EFG68719	32	48	
	117600	C	*Frankia* sp. EuI1c; hypothetical protein FraEuI1c_5528	ADP83514	38	63	
	117590	D	*Caulobacter* sp. K31; short-chain dehydrogenase/reductase SDR	ABZ71049	52	66	
	117580	E	*Methylobacterium nodulans* ORS 2060; glutathione S-transferase domain protein	ACL56876	40	56	
	117570	F	*Novosphingobium* sp. PP1Y; RHD α-subunit	CCA92463	52	69	*pahAc5*
	117560	G	*Sphingopyxis macrogoltabida* ; RHD β-subunit	AAN26444	37	59	*pahAd5*
	117550	H	*Bordetella petrii* ; similar to glucose 1-dehydrogenase	CAP41850	50	66	
	117540	I	*Rhodococcus jostii* RHA1; flavin-binding monooxygenase	ABG95240	66	79	
05425	124030	A	*Methylovorus* sp. MP688; DNA helicase/exodeoxyribonuclease V, subunit B	ADQ84475	38	54	
	124020	B	*Dechloromonas aromatica* RCB; UvrD/REP helicase	AAZ45225	48	61	
	124010	C	*Sphingobium* sp. HV3; large subunit aromatic oxygenase	CAM32403	79	86	*pahAc6*
	124000	D	*Novosphingobium* sp. PP1Y; 3-phenylpropionate dioxygenase β-subunit	CCA93888	71	84	*pahAd6*
	123990	E	*Novosphingobium* sp. PP1Y; fatty acid desaturase	CCA93889	46	61	

Contig	ORF	Figure Label[a]	Strongest BLASTP Match	GenBank Accession	% a.a. Identity	% a.a. Similarity	Gene[b]
05431	129310	A	*Pigmentiphaga* sp. NDS2; putative LysR-type transcriptional regulator	BAC53588	50	71	
	129290	B	*Burkholderia* sp. DBT1; initial dioxygenase ferrodoxin DbtAb	AAK96190	61	80	*pahAb*
	129280	C	*Burkholderia* sp. DBT1; putative monoxygenase α-subunit	AAK96191	85	93	
	129270	D	*Burkholderia* sp. DBT1; putative hydrolase	AAK96192	70	82	
	129230	E	*Cycloclasticus* sp. P1; PAH dioxygenase component ferredoxin reductase	ABF56512	52	69	*pahAa*
	129130	F	*Burkholderia* sp. Ch1-1; dihydrodipicolinate synthetase	EFG65992	80	91	
	129090	G	*Acidovorax* sp. NA3; putative 2-hydroxychromene-2-carboxylate isomerase	ACG70975	52	68	
	129020	H	*Burkholderia* sp. DBT1; DbtAc	AAK62353	67	82	*pahAc1*
	128980	I	*Burkholderia* sp. DBT1; DbtAd	AAK62354	50	70	*pahAd1*
	129200	J	*Methyloversatilis universalis* FAM5; putative cox2 cytochrome oxidase subunit 2	EGK69739	49	62	
	129300	K	*Methyloversatilis universalis* FAM5; hypothetical protein METUNv1_03703	EGK69740	50	65	
05432	129870	A	*Methylibium petroleiphilum* PM1; conserved hypothetical protein	ABM95224	44	64	
	129840	B	*Sphingopyxis macrogoltabida*; cis-dihydrodiol dehydrogenase	AAN26445	66	80	
	129770	C	*Azoarcus* sp. BH72; conserved hypothetical flavin reductase	CAL94579	59	74	
	129780	D	*Burkholderia sartisoli*; catechol 2,3-dioxygenase	AAF02426	75	85	
	129790	E	*Sphingomonas* sp. P2; RHD α-subunit	BAC65448	42	57	*pahAc7*
	129800	F	*Azotobacter vinelandii* DJ; Aromatic-RHD, β-subunit	ACO80345	35	51	*pahAd7*
	129810	G	*Mycobacterium vanbaalenii* PYR-1; succinate dehydrogenase/fumarate reductase	AAT51748	67	81	
	129820	H	*Mycobacterium* sp. KMS; 4Fe-4S ferredoxin, iron-sulfur binding domain protein	ABL90868	63	72	
	129830	I	*Mycobacterium rhodesiae* NBB3; hypothetical protein MycrhNDRAFT_2578	EHB48945	51	68	
	129850	J	*Mycobacterium rhodesiae* NBB3; Rieske (2Fe-2S) iron-sulfur domain protein	EHB48946	66	80	
	129860	K	*Sphingomonas wittichii* RW1; aromatic RHD, β-subunit	ABQ67262	47	62	
	129720	L	*Mycobacterium rhodesiae* NBB3; L-threonine 3-dehydrogenase	EHB48944	54	68	
	129730	M	*Burkholderia sartisoli*; unknown	AAF02431	64	74	
	129740	N	*Delftia* sp. Cs1-4; naphthalene 1,2-dioxygenase	AEF88772	52	71	*pahAc8*
	129750	O	*Burkholderia sartisoli*; dioxygenase β-subunit PhnAd	AAD09873	44	62	*pahAd8*
	129760	P	*Pseudoxanthomonas spadix* BD-a59; xylene monooxygenase	EHB93984	36	52	

[a] As designated in Figure 3.

[b] Name of gene from this study.