

ORIGINAL ARTICLE

Open Access



Thermophilic bacteria are potential sources of novel Rieske non-heme iron oxygenases

Joydeep Chakraborty, Chiho Suzuki-Minakuchi, Kazunori Okada and Hideaki Nojiri* 

Abstract

Rieske non-heme iron oxygenases, which have a Rieske-type [2Fe–2S] cluster and a non-heme catalytic iron center, are an important family of oxidoreductases involved mainly in regio- and stereoselective transformation of a wide array of aromatic hydrocarbons. Though present in all domains of life, the most widely studied Rieske non-heme iron oxygenases are found in mesophilic bacteria. The present study explores the potential for isolating novel Rieske non-heme iron oxygenases from thermophilic sources. Browsing the entire bacterial genome database led to the identification of 45 homologs from thermophilic bacteria distributed mainly among Chloroflexi, Deinococcus–Thermus and Firmicutes. Thermostability, measured according to the aliphatic index, showed higher values for certain homologs compared with their mesophilic relatives. Prediction of substrate preferences indicated that a wide array of aromatic hydrocarbons could be transformed by most of the identified oxygenase homologs. Further identification of putative genes encoding components of a functional oxygenase system opens up the possibility of reconstituting functional thermophilic Rieske non-heme iron oxygenase systems with novel properties.

Keywords: Rieske non-heme iron oxygenase, Oxidoreductase, Thermophiles, Aromatic hydrocarbons, Biotransformation

Introduction

Rieske non-heme iron oxygenases (ROs) constitute a large family of oxidoreductase enzymes involved primarily in the oxygenation of various aromatic compounds. Although Gibson et al. (1968) first detected the involvement of such an enzyme system in an alkylbenzene-degrading *Pseudomonas* sp., the family has since garnered a great deal of attention for two major reasons. First, ROs are key enzymes responsible for the initial attack on otherwise inert aromatic nuclei, thereby making them targets of a cascade of downstream enzymes, leading to their complete mineralization (Gibson and Subramanian 1984; Allen et al. 1995; Gibson and Parales 2000; Mallick et al. 2011). Secondly, regio- and stereoselective *cis*-dihydroxylation of aromatic compounds, catalyzed by ROs, generate impressive chiral intermediates in the synthesis

of a wide array of agrochemically and pharmaceutically important compounds (Ensley et al. 1983; Wackett et al. 1988; Hudlicky et al. 1999; Bui et al. 2002; Newman et al. 2004; Boyd et al. 2005; Zezula and Hudlicky 2005).

Members of the RO family are usually either two- or three-component systems in which one or two soluble electron transport (ET) proteins (such as ferredoxin and reductase) transfer electrons from reduced nucleotides, such as NAD(P)H, to the terminal oxygenase component (a large α -subunit, often accompanied by a small β -subunit), which in turn catalyzes the di- or mono-oxygenation of the aromatic nucleus of the substrate (Mason and Cammack 1992; Ferraro et al. 2005). Numerous ROs have been identified and characterized from bacteria, thereby enriching the available information on their diversity in terms of both sequence and function (Habe and Omori 2003; Iwai et al. 2010, 2011; Chakraborty et al. 2012). Although found in all three domains of life, studies have shown that ROs occur more commonly in bacteria compared with archaea and eukaryotes (Chakraborty

*Correspondence: anojiri@mail.ecc.u-tokyo.ac.jp
Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi,
Bunkyo-ku, Tokyo, Japan

et al. 2012). Homologs of the large (α) subunit of RO terminal oxygenase (RO_{ox}) have also been investigated in certain plant species, such as *Arabidopsis thaliana*, *Zea mays*, *Pisum sativum*, *Oryza sativa*, *Physcomitrella patens*, *Amaranthus tricolor*, *Ocimum basilicum* and *Spinacia oleracea* (Caliebe et al. 1997; Meng et al. 2001; Reinbothe et al. 2004; Berim et al. 2014), as well as in insects, nematodes and vertebrates (Rottiers et al. 2006; Yoshiyama et al. 2006; Yoshiyama-Yanagawa et al. 2011). ROs from these taxa, however, have entirely different functions from those of bacterial aromatic ring-hydroxylating ROs. They either act as proteintranslocins, facilitating transport across the chloroplastic envelope membranes during chlorophyll biosynthesis, or are involved in flavone and hormone metabolism in plants. They have also been suggested to be involved in regulation of cholesterol metabolism or trafficking during steroid synthesis in insects (Caliebe et al. 1997; Meng et al. 2001; Reinbothe et al. 2004; Rottiers et al. 2006; Yoshiyama et al. 2006; Yoshiyama-Yanagawa et al. 2011; Berim et al. 2014).

Interestingly, few bacterial RO homologs with novel functions, such as oxidative cyclization during biosynthesis of certain antibiotics, hydroxylation and desaturation of short-chain tertiary alcohols and alkane monooxygenation, have been reported in recent years (Sydor et al. 2011; Schäfer et al. 2012; Li et al. 2013). This suggests that ROs bear much more catalytic potential than previously realized. Almost all bacterial ROs characterized biochemically to date have been isolated from mesophilic bacteria, with the sole exception of polychlorinated biphenyl degrading ring-hydroxylating dioxygenase from *Geobacillus* sp. JF8 (Mukerjee-Dhar et al. 2005; Shintani et al. 2014). As such, very little is known about RO homologs present in bacteria that live in extreme environments. Extremophiles, and in particular their enzymes, have proved to be a potentially valuable resource in the development of novel biotechnological processes. The most well-studied extremophiles include thermophiles and hyperthermophiles, and enzymes isolated from such microorganisms are often extremely thermostable and resistant to proteolysis, chemical denaturants, detergents, and organic solvents (Vieille and Zeikus 2001). Apart from enzymatic stability at high temperatures, which is often desired in industrial processes, there are several advantages of thermophilic systems in bioremediation studies. Owing to the poor aqueous solubility of aromatic hydrocarbons, biodegradation studies often encounter problems related to bioavailability. These issues can be overcome at elevated temperatures, since bioavailability tends to increase with temperature owing to increases in solubility (Margesin and Schinner 2001; Feitkenhauer and Märk 2003; Perfumo et al. 2007). Thermophilic microorganisms may thus be attractive

candidates for sources of novel thermostable ROs with potential utility in industrial biosynthesis and bioremediation at elevated temperatures.

In recent years, microbial genome sequencing projects have generated an enormous quantity of data for public databases. Since publication of the genome sequence of the first extremophile in 1996 (Bult et al. 1996), there has been a substantial increase in the number of extremophilic genome sequences. Metagenomics and single-cell genomics further add to this repertoire (Hedlund et al. 2014). The present study explored all available genome sequences of thermophilic bacteria for the presence of RO homologs and predicted their suitability as novel RO candidates for biotechnological applications.

Materials and methods

Screening thermophilic bacterial genomes for the presence of RO_{ox} α -subunit homologs

Functionally characterized ROs have been categorized into five different similarity classes (A, B, C, D and D*) based on their phylogenetic distribution, substrate preferences and mode of attack on aromatic nuclei (Chakraborty et al. 2012). The National Center for Biotechnology Information (NCBI) 'genome' and 'taxonomy' databases were searched to characterize the distribution of thermophilic bacteria among different bacterial lineages and the availability of their genome sequences. Representative RO_{ox} α -subunit sequences from each class were used as query probes (Table 1) to perform blastp (Altschul et al. 1990) searches against the translated set of genome sequences. Blast searches were also performed using each thermophilic RO as a query against the NCBI non-redundant database to characterize their distribution among thermophiles (and/or other extremophiles) and mesophiles.

Phylogenetic clustering and prediction of substrate preferences

The RHObase server (Chakraborty et al. 2014) was used to categorize each candidate thermophilic RO_{ox} α -subunit into a similarity class and to obtain the closest biochemically characterized homologs. The substrate prediction module of RHObase was further used to predict the substrate preference of the thermophilic homologs and the possible sites of oxygenation. ClustalX v1.81 (Thompson et al. 1997) was used to obtain multiple sequence alignments and to eliminate redundancy among sequences. The default settings were retained for all parameters, with the exception of the matrix (BLOSUM series) used for both pairwise and multiple alignments. Phylogenetic trees were constructed based on distance data using the neighbor-joining method (Saitou and Nei 1987) implemented in ClustalX. The trees were visualized and manipulated using the program TreeExplorer v2.12 (Tamura et al. 2007).

Table 1 List of queries used for blast analysis against genomes of thermophilic bacteria

Target gene	Query	Type	Organism	GenBank accession no.
RO α -subunit	Naphthalene dioxygenase (NahAc)	A-IIIa β	<i>Pseudomonas putida</i> NCIB 9816-4	AAO64274
	Benzoate dioxygenase (BenA)	B-IIa β	<i>Rhodococcus jostii</i> RHA1	BAB70698
	Salicylate 5-hydroxylase (NagG)	C-IIIa β	<i>Ralstonia</i> sp. U2	AAD12607
	Carbazole dioxygenase (CarAall)	D-VIIa	<i>Sphingomonas</i> sp. KA1	YP_717942
	3-Ketosteroid 9 α -hydroxylase (KshA)	D-Ia	<i>Mycobacterium tuberculosis</i> H37rv	NP_218043
Ferredoxin	AntAb associated with anthranilate dioxygenase	Rieske type [2Fe–2S]	<i>Sphingomonas</i> sp. KA1	YP_717959
	CarAcl associated with carbazole dioxygenase	Plant type [2Fe–2S]	<i>Sphingomonas</i> sp. KA1	YP_717977
	PhtA3 associated with phthalate dioxygenase	[3Fe–4S] type	<i>Terrabacter</i> sp. DBF63	BAC54160
	Rub1 associated with naphthalene 1,2-dioxygenase	Rubredoxin	<i>Rhodococcus</i> sp. P200	AAR05110
Reductase	KshB associated with 3-Ketosteroid 9 α -hydroxylase	FNRc type	<i>Rhodococcus erythropolis</i> PR4	BAH32483
	CarAd associated with carbazole dioxygenase	FNRn type	<i>Pseudomonas</i> sp. XLDN4-9	AA56344
	Fdrl associated with carbazole dioxygenase	GR type	<i>Sphingomonas</i> sp. KA1	YP_718026
	HpaC associated with 4-hydroxyphenylacetate 3-monooxygenase	Flavin reductase	<i>Thermus thermophilus</i> HB8	2ECU_A
	Rubredoxin-NAD(+) reductase	Rubredoxin reductase	<i>Pseudomonas aeruginosa</i> PAO1	Q9HTK9

FNRc ferredoxin-NAD reductase fused with a plant-type [2Fe–2S] domain at the C-terminus, *FNRn* ferredoxin-NAD reductase fused with a plant-type [2Fe–2S] domain at the N-terminus, *GR* glutathione reductase

Verification of the integrity of conserved motifs and domain architecture

The RO_{ox} α -subunit homologs obtained from the genomes of thermophiles were subjected to ScanProsite (De Castro et al. 2006) and NCBI conserved domain database searches (Marchler-Bauer et al. 2002) to verify the presence of conserved sequence motifs. The relevant motifs were C-X-H-X_n-C-X₂-H, corresponding to the N-terminal Rieske [2Fe–2S] center, and D-X₂-H-X_{3,4}-H-X_n-D, corresponding to the C-terminal conserved 2-His-1-carboxylate motif preceded by a conserved aspartate (involved in electron transport), as these are the functional prerequisites of ROs (Jiang et al. 1996; Parales 2003). The motifs were compared with those of phylogenetically close mesophilic ROs. For each protein, the aliphatic index (relative volume occupied by aliphatic side chains) (Ikai 1980) was calculated using ProtParam (Gasteiger et al. 2005).

Identification of putative ET components

Genomes exhibiting the presence of RO_{ox} α -subunits were searched (using blastp) for genes putatively encoding ET components (both ferredoxin and reductase) using a set of queries (Table 1), followed by manual inspection of each genomic loci when necessary. The queries included the oxidoreductase sequences (e.g., ferredoxin-NAD reductases and glutathione reductase-type reductases and ferredoxins) commonly associated with ROs, as well as other possible oxidoreductases (e.g., flavin reductase and rubredoxin reductase).

Results

Distribution of RO homologs among thermophiles

Browsing the bacterial taxonomy database revealed the existence of several thermophilic genera belonging to different classes/orders. These taxa were concentrated mainly among the phyla Thermotogae, Deinococcus–Thermus, Chloroflexi, Aquificae, Firmicutes, and to some extent, Bacteroidetes/Chlorobi, Actinobacteria and Proteobacteria. Blast searches against all thermophile genomes initially led to the identification of 95 putative RO_{ox} α -subunit homologs distributed among 20 different genera (data not shown). Among 45 non-redundant sequences (Table 2), the one obtained from *Alicyclobacillus acidoterrestris* ATCC 49025 (GenBank: EPZ42375) was found to be truncated at the N-terminal end and was therefore excluded from further analysis. Analysis of the distribution of the remaining candidate ROs among both thermophiles and mesophiles revealed that they were present mainly among thermophilic strains belonging to the phyla Chloroflexi, Deinococcus–Thermus, Firmicutes and Thermotogae (Fig. 1). However, distant homologs were abundant among mesophilic strains belonging to the phyla Actinobacteria, Firmicutes and Proteobacteria.

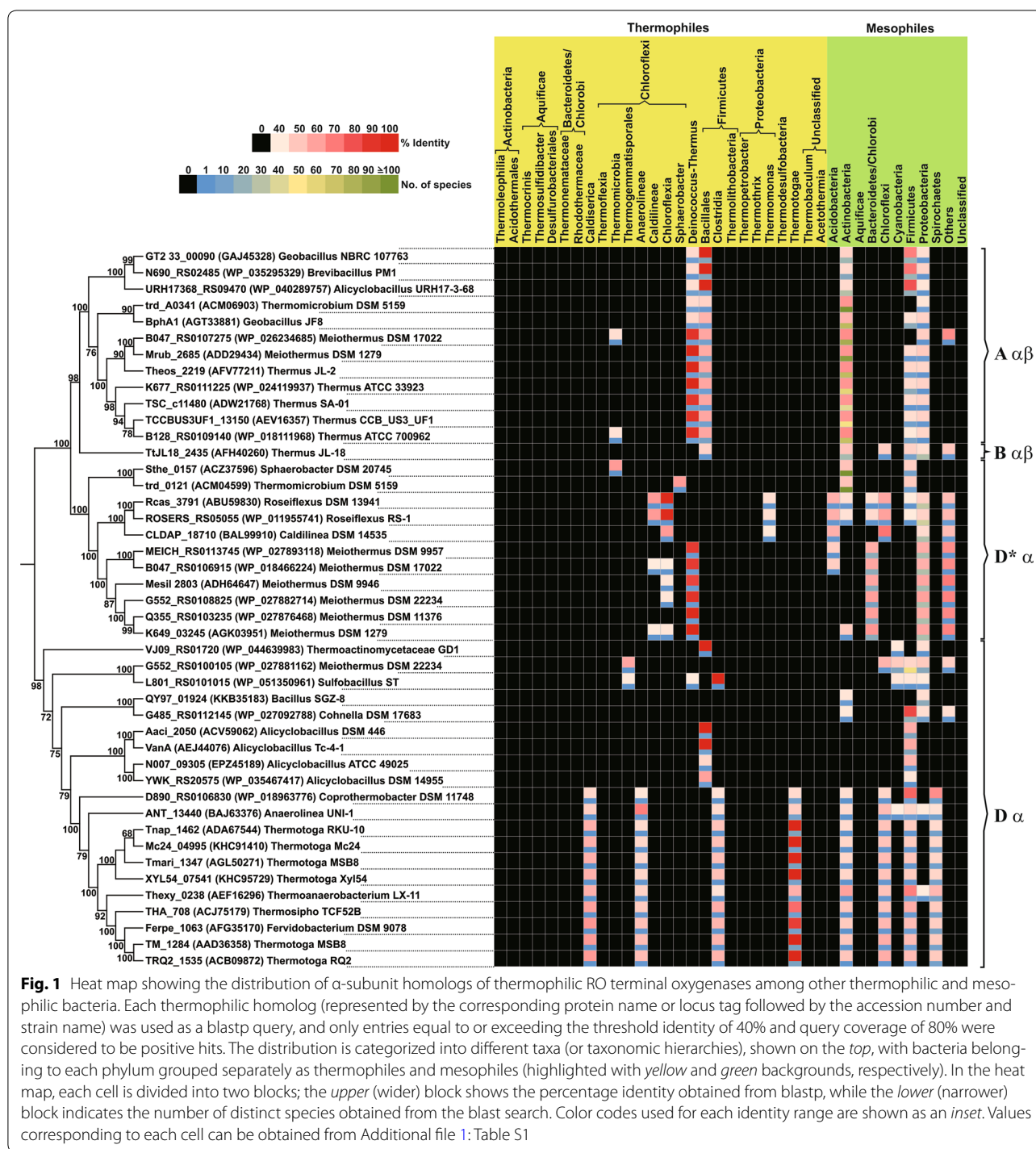
Functional clustering of thermophilic RO_{ox} α -subunit homologs

The candidate α -subunit sequences from thermophiles were subjected to phylogenetic studies to assess their relatedness with functionally characterized ROs from other bacteria, as well as from eukaryotes. The

Table 2 Putative RO terminal oxygenase α -subunit homologues obtained from blastp search against thermophilic genomes

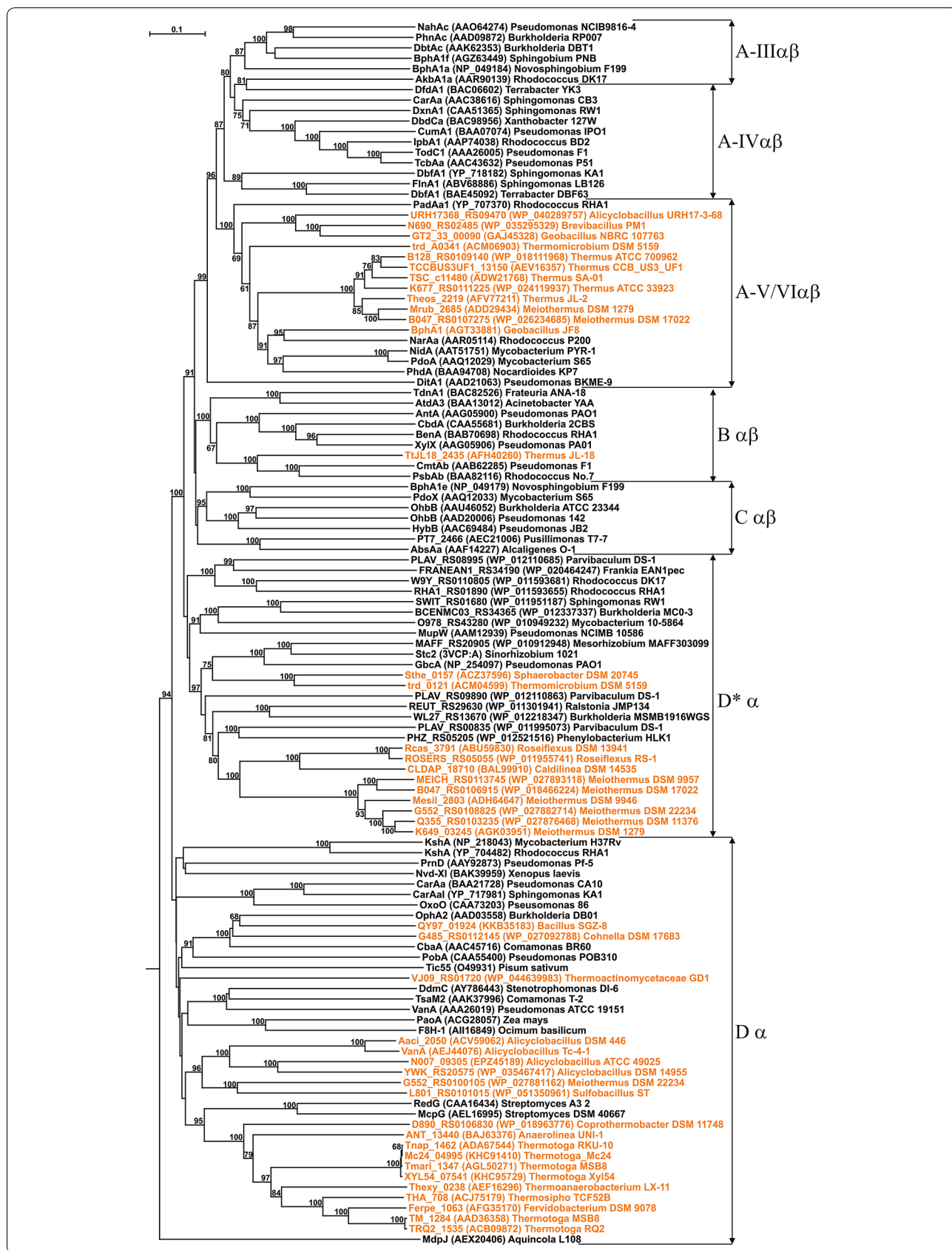
Organism	Annotation	NCBI accn. no.
Phylum: Chloroflexi		
<i>Anaerolinea thermophila</i> UNI-1	Putative oxidoreductase	BAJ63376
<i>Caldilinea aerophila</i> DSM 14535	Putative oxidoreductase	BAL99910
<i>Roseiflexus castenholzii</i> DSM 13941	Rieske (2Fe–2S) domain protein	ABU59830
<i>Roseiflexus</i> sp. RS-1	(2Fe–2S)-binding protein	WP_011955741
<i>Sphaerobacter thermophilus</i> DSM 20745	Rieske (2Fe–2S) domain protein	ACZ37596
<i>Thermomicrobium roseum</i> DSM 5159	Rieske 2Fe–2S domain protein	ACM04599
<i>Thermomicrobium roseum</i> DSM 5159	3-Phenylpropionate (digoxigenin) dioxygenase	ACM06903
Phylum: Deinococcus–Thermus		
<i>Meiothermus cereberus</i> DSM 11376	Choline monooxygenase	WP_027876468
<i>Meiothermus chliarophilus</i> DSM 9957	Choline monooxygenase	WP_027893118
<i>Meiothermus ruber</i> DSM 1279	Aromatic-ring-hydroxylating dioxygenase, α -subunit	ADD29434
<i>Meiothermus ruber</i> DSM 1279	Choline monooxygenase	AGK03951
<i>Meiothermus rufus</i> DSM 22234	Hypothetical protein	WP_027881162
<i>Meiothermus rufus</i> DSM 22234	Choline monooxygenase	WP_027882714
<i>Meiothermus silvanus</i> DSM 9946	Rieske (2Fe–2S) iron–sulfur domain protein	ADH64647
<i>Meiothermus timidus</i> DSM 17022	Choline monooxygenase	WP_018466224
<i>Meiothermus timidus</i> DSM 17022	Ribosomal subunit interface protein	WP_026234685
<i>Thermus igniterrae</i> ATCC 700962	Ribosomal subunit interface protein	WP_018111968
<i>Thermus oshimai</i> JL-2	Ring-hydroxylating dioxygenase, large terminal subunit	AFV77211
<i>Thermus scotoductus</i> SA-01	Biphenyl dioxygenase, subunit alpha	ADW21768
<i>Thermus</i> sp. CCB_US3_UF1	Aromatic-ring-hydroxylating dioxygenase, α -subunit	AEV16357
<i>Thermus thermophilus</i> ATCC 33923	Ribosomal subunit interface protein	WP_024119937
<i>Thermus thermophilus</i> JL-18	Rieske (2Fe–2S) domain-containing protein	AFH40260
Phylum: Firmicutes		
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> DSM 446	Rieske (2Fe–2S) iron–sulfur domain protein	ACV59062
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> Tc-4-1	2Fe–2S ferredoxin	AEJ44076
<i>Alicyclobacillus acidoterrestris</i> ATCC 49025	Hypothetical protein	EPZ45189
<i>Alicyclobacillus acidoterrestris</i> ATCC 49025	Hypothetical protein ^a	EPZ42375
<i>Alicyclobacillus hesperidum</i> URH17-3-68	3-Phenylpropionate dioxygenase	WP_040289757
<i>Alicyclobacillus pomorum</i> DSM 14955	Hypothetical protein	WP_035467417
<i>Bacillus thermotolerans</i> SGZ-8	Phthalate 4,5-dioxygenase oxygenase subunit	KKB35183
<i>Brevibacillus thermoruber</i> PM1	3-Phenylpropionate dioxygenase	WP_035295329
<i>Cohnella thermotolerans</i> DSM 17683	Rieske (2Fe–2S) protein	WP_027092788
<i>Coprothermobacter platensis</i> DSM 11748	Hypothetical protein	WP_018963776
<i>Geobacillus</i> sp. JF8	Large subunit of biphenyl dioxygenase	AGT33881
<i>Geobacillus thermoglucosidasius</i> NBRC 107763	Putative naphthalene 1,2-dioxygenase large subunit	GAJ45328
<i>Sulfobacillus thermosulfidoxidans</i> ST	Hypothetical protein	WP_051350961
Thermoactinomycetaceae bacterium GD1	Rieske (2Fe–2S) protein	WP_044639983
<i>Thermoanaerobacterium xylanolyticum</i> LX-11	Rieske (2Fe–2S) iron–sulfur domain protein	AEF16296
Phylum: Thermotogae		
<i>Fervidobacterium pennivorans</i> DSM 9078	Ring-hydroxylating dioxygenase, large terminal subunit	AFG35170
<i>Thermosipho africanus</i> TCF52B	Oxidase-related protein	ACJ75179
<i>Thermotoga maritima</i> MSB8	Oxidase-related protein	AAD36358
<i>Thermotoga maritima</i> MSB8	Rieske (2Fe–2S) domain protein	AGL50271
<i>Thermotoga naphthophila</i> RKU-10	Rieske (2Fe–2S) iron–sulfur domain protein	ADA67544
<i>Thermotoga</i> sp. Mc24	Rieske (2Fe–2S) iron–sulfur domain-containing protein	KHC91410
<i>Thermotoga</i> sp. RQ2	Rieske (2Fe–2S) domain protein	ACB09872
<i>Thermotoga</i> sp. Xyl54	Rieske (2Fe–2S) iron–sulfur domain-containing protein	KHC95729

^a Protein truncated at the N-terminal end



phylogenetic tree (Fig. 2) showed that thermophilic homologs were unevenly distributed among all classes of ROs, being clustered in a few specific regions of the tree, again suggesting a radical diversification followed by independent evolution of these genes in thermophiles. As can be seen in Fig. 2, 12 of the sequences clustered

with Class A (especially Type A-V α β and A-VI α β) ROs, 1 with Class B ROs, 20 with Class D ROs, and the remaining 11 with Class D* ROs. In all sequences, the conserved N-terminal Rieske [2Fe-2S] motif and C-terminal 2-His-1-carboxylate motif (preceded by a conserved aspartate involved in electron transport), necessary for proper



(See figure on previous page.)

Fig. 2 Neighbor-joining tree representing the phylogenetic relationship of putative α -subunits of oxygenase components of ROs obtained from thermophilic bacteria (*orange font*) with homologous sequences from other bacteria and eukaryotes. Each entry is represented by the corresponding protein name or locus tag, followed by the accession number (within *parentheses*) and the strain name. Values at each node indicate the level of bootstrap support based on 100 resampled datasets, while bootstrap values below 60% are not shown. The *bar* represents 0.1 substitutions per amino acid. The sequences have been clustered according to similarity class as defined in Chakraborty et al. (2014)

functioning of an RO (Jiang et al. 1996; Parales 2003), is consistent with phylogenetically related, previously characterized ROs (Fig. 3).

The aliphatic index is regarded as a positive factor for the increased thermostability of globular proteins (Ikai 1980). Therefore, we calculated this index for each protein as a measure of thermostability. Not all proteins showed significantly higher values compared with their mesophilic homologs (Fig. 3). However, the average value was found to be higher (80.88) in the thermophilic clusters as compared to those of the mesophiles (75.25), with some proteins (e.g., ROs obtained from Thermoactinomycetaceae bacterium GD1 [GenBank ID WP_044639983], *Meiothermus rufus* DSM 22234 [WP_027881162], *Sphaerobacter thermophilus* DSM 20745 [ACZ37596] and *Thermomicrobium roseum* DSM 5159 [ACM04599]) showing values as high as 90 (Fig. 3).

Prediction of substrate preferences

Table 3 lists the closest biochemically characterized homolog of each candidate RO_{ox} α -subunit, as obtained from RHObase (Chakraborty et al. 2014). Preferable substrate(s) for most candidate ROs could be predicted using the RHObase substrate prediction module (Fig. 4). RO_{ox} α -subunit homologs belonging to Class A showed a preference for polycyclic aromatic hydrocarbons and heterocyclic polyaromatic hydrocarbons. For one of the Class A ROs, obtained from *Alicyclobacillus acidocaldarius* subsp. *acidocaldarius* Tc-4-1, ketosteroid was found to be its putative substrate. The predicted substrates for the Class B RO_{ox} α -subunit from *Thermus thermophilus* JL-18 were carboxylated aromatics, such as *p*-cumate, while members of Class D showed a preference for carboxylated aromatics such as phthalate, chlorobenzoate, vanillate and phenoxybenzoates, as well as for toluene-4-sulfonate. However, owing to the lack of information regarding the function of Class D* ROs, the substrate

preference of these ROs could not be predicted. Apart from MupW and GbcA, involved in the mupirocin (El-Sayed et al. 2003) and glycine betaine (Wargo et al. 2008) biosynthetic pathways, respectively, all other sequences belonging to this class have been derived from whole genome annotations and lack complete information regarding their biochemical function. This makes Class D* the 'dark matter' of Rieske oxygenases.

Reconstitution of functional RO systems

As discussed earlier, the oxygenase α -subunit is often accompanied by a small β -subunit, and these subunits function in combination with one or two ET component(s). All observations discussed thus far concern the α -subunits of RO_{ox}. However, to reconstitute a functional RO system, all of the above components must work together in a coordinated manner. Whenever present, the genes encoding both α - and β -subunits are usually co-localized. Thus, the genome of each thermophile bearing the candidate RO_{ox} α -subunit homologs was screened for the presence of genes putatively encoding the ET component(s). Several putative genes (listed in Table 4) were identified and, in most cases, were located at a distance from the terminal oxygenase genes. As expected, putative ferredoxin and reductase components, along with the β -subunit of RO_{ox}, could be identified in most organisms bearing Class A and B ROs. The genome of *Thermus thermophilus* ATCC 33923 had no adjacent β -subunit and did not yield any ferredoxin hits with the queries used. *Alicyclobacillus hesperidum* URH17-3-68 and *Brevibacillus thermoruber* PM1 were also found to lack ferredoxin. Among Class D ROs, D-IV α and D-VII α type ROs form three-component systems containing both ferredoxin and reductase (Chakraborty et al. 2012). Similarly, putative ferredoxin and reductase components could be identified in several thermophiles bearing Class D ROs. The only exceptions were *Bacillus thermotolerans* SGZ-8 and Thermoactinomycetaceae bacterium GD1,

(See figure on next page.)

Fig. 3 Comparison of conserved N-terminal Rieske [2Fe–2S] and C-terminal 2-His-1-carboxylate motifs among α -subunits of oxygenase components of putative thermophilic ROs obtained from thermophilic bacteria (*orange font*) and those obtained from mesophilic bacteria and eukaryotes. The *horizontal bars* represent the aliphatic index of each sequence. *Blue and orange vertical dotted lines* indicate the average aliphatic indices obtained for mesophilic (75.25) and thermophilic (80.88) RO homologs, respectively. All those homologs which showed an aliphatic index ≥ 80.88 are indicated by an *arrow*, while clades representing only the thermophilic homologs are denoted by *asterisks*

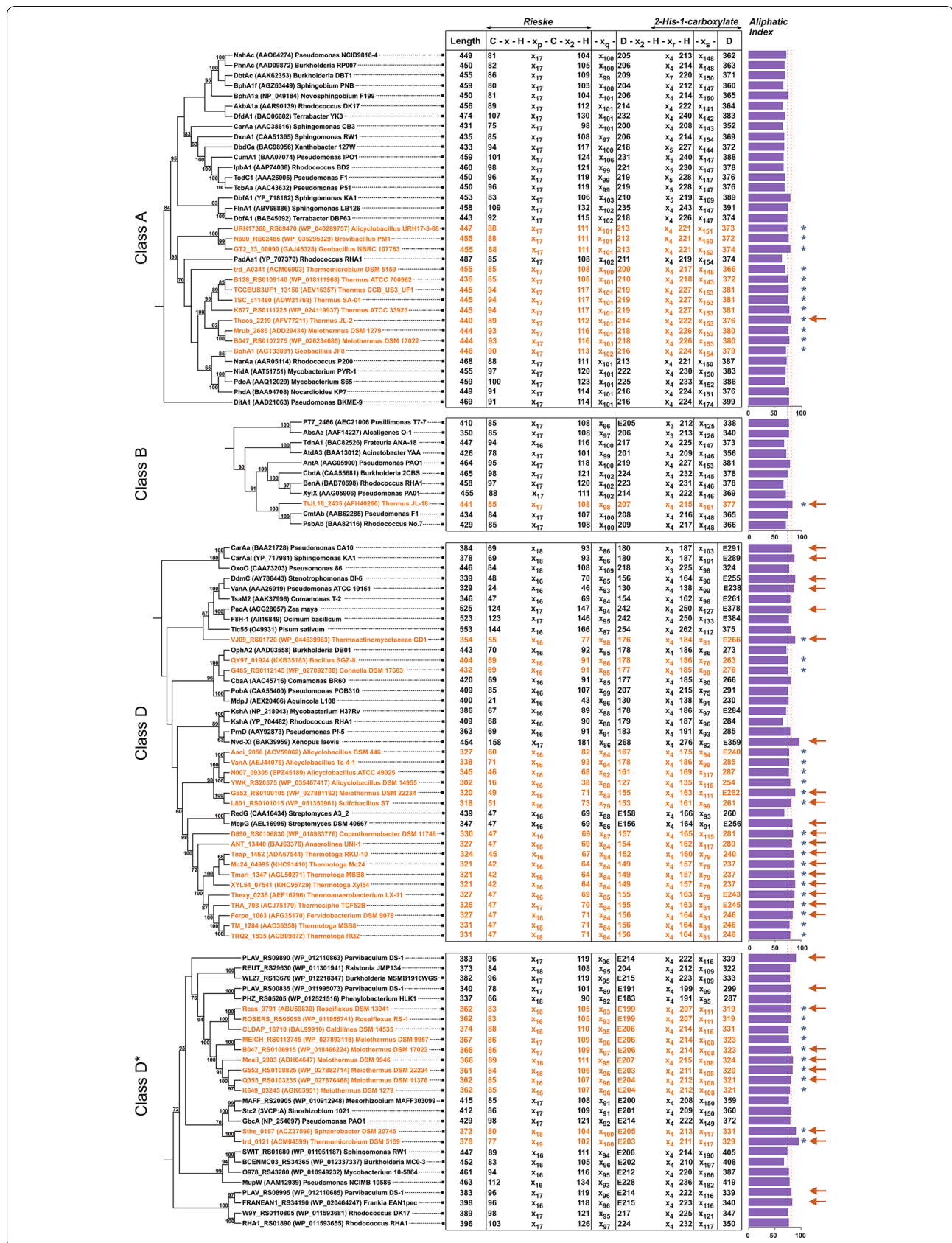


Table 3 The closest characterized homologue of each thermophilic RO terminal oxygenase α -subunit sequence obtained from RHObase

Organism (NCBI accession no. of the putative RO terminal oxygenase α -subunit)	Closest match obtained from RHObase	Accession no.	Identity (%)	Query coverage (%)
Class A				
<i>Thermomicrobium roseum</i> DSM 5159 (ACM06903)	Biphenyl 2,3-dioxygenase (BphA1) from <i>Geobacillus</i> sp. JF8	BAC79226	46.10	91
<i>Meiothermus ruber</i> DSM 1279 (ADD29434)	Dibenzofuran dioxygenase (NarAa) from <i>Rhodococcus opacus</i> SAO101	BAD02377	50.11	92
<i>Meiothermus timidus</i> DSM 17022 (WP_026234685)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	49.89	98
<i>Thermus igniterrae</i> ATCC 700962 (WP_018111968)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	51.48	99
<i>Thermus oshimai</i> JL-2 (AFV77211)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	50.34	98
<i>Thermus scotoductus</i> SA-01 (ADW21768)	Dibenzofuran dioxygenase (NarAa) from <i>Rhodococcus opacus</i> SAO101	BAD02377	51.14	99
<i>Thermus</i> sp. CCB_US3_UF1 (AEV16357)	Dibenzofuran dioxygenase (NarAa) from <i>Rhodococcus opacus</i> SAO101	BAD02377	51.26	97
<i>Thermus thermophilus</i> ATCC 33923 (WP_024119937)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	51.03	97
<i>Alicyclobacillus hesperidum</i> URH17-3-68 (WP_040289757)	Dibenzofuran dioxygenase (NarAa) from <i>Rhodococcus opacus</i> SAO101	BAD02377	43.02	93
<i>Brevibacillus thermoruber</i> PM1 (WP_035295329)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	44.59	92
<i>Geobacillus</i> sp. JF8 (AGT33881)	Biphenyl 2,3-dioxygenase (BphA1) from <i>Geobacillus</i> sp. JF8	BAC79226	100.00	100
<i>Geobacillus thermoglucosidasius</i> NBRC 107763 (GAJ45328)	Indene dioxygenase (NidA) from <i>Rhodococcus</i> sp. I24	AAD25395	45.81	93
Class B				
<i>Thermus thermophilus</i> JL-18 (AFH40260)	<i>p</i> -Cumate dioxygenase (CmtAb) from <i>Pseudomonas putida</i> F1	AAB62285	45.37	94
Class D				
<i>Anaerolinea thermophila</i> UNI-1 (BAJ63376)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.04	31
<i>Meiothermus rufus</i> DSM 22234 (WP_027881162)	Vanillate <i>O</i> -demethylase (VanA) from <i>Pseudomonas</i> sp. ATCC19151	AAA26019	32.10	47
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> DSM 446 (ACV59062)	Toluene-4-sulfonate methyl monooxygenase (TsaM2) from <i>Comamonas testosteroni</i> T-2	AAK37996	32.45	32
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> Tc-4-1 (AEJ44076)	3-Ketosteroid 9 α -hydroxylase (KshA) from <i>Rhodococcus opacus</i> B-4	BAH52700	30.54	54
<i>Alicyclobacillus acidoterrestris</i> ATCC 49025 (EPZ45189)	Toluene-4-sulfonate methyl monooxygenase (TsaM2) from <i>Comamonas testosteroni</i> T-2	AAK37996	30.22	48
<i>Alicyclobacillus pomorum</i> DSM 14955 (WP_035467417)	3-Chlorobenzoate-3,4/4,5-dioxygenase (CbaA) from <i>Comamonas testosteroni</i> BR60	AAC45716	27.56	50
<i>Bacillus thermotolerans</i> SGZ-8 (KKB35183)	Phthalate 4,5-dioxygenase (Pht3) from <i>Pseudomonas putida</i>	BAA02511	37.86	94
<i>Cohnella thermotolerans</i> DSM 17683 (WP_027092788)	Phthalate 4,5-dioxygenase (OphA2) from <i>Burkholderia cepacia</i> DBO1	AAD03558	36.24	97
<i>Coprothermobacter platensis</i> DSM 11748 (WP_018963776)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.93	51
<i>Sulfobacillus thermosulfidooxidans</i> ST (WP_051350961)	Toluene-4-sulfonate methyl monooxygenase (TsaM2) from <i>Comamonas testosteroni</i> T-2	AAK37996	30.12	46
Thermoactinomycetaceae bacterium GD1 (WP_044639983)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	34.78	54
<i>Thermoanaerobacterium xylanolyticum</i> LX-11 (AEF16296)	Vanillate <i>O</i> -demethylase (VanA) from <i>Pseudomonas</i> sp. HR199	CAA72287	31.84	48
<i>Fervidobacterium pennivorans</i> DSM 9078 (AFG35170)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	34.13	47
<i>Thermosipho africanus</i> TCF52B (ACJ75179)	Vanillate <i>O</i> -demethylase (VanA) from <i>Pseudomonas</i> sp. HR199	CAA72287	33.92	46
<i>Thermotoga maritima</i> MSB8 (AAD36358)	3-Chlorobenzoate-3,4/4,5-dioxygenase (CbaA) from <i>Comamonas testosteroni</i> BR60	AAC45716	25.35	65
<i>Thermotoga maritima</i> MSB8 (AGL50271)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.72	64
<i>Thermotoga naphthophila</i> RKU-10 (ADA67544)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.32	73
<i>Thermotoga</i> sp. Mc24 (KHC91410)	Dicamba <i>O</i> -demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.72	73

Table 3 continued

Organism (NCBI accession no. of the putative RO terminal oxygenase α -subunit)	Closest match obtained from RHObase	Accession no.	Identity (%)	Query coverage (%)
<i>Thermotoga</i> sp. RQ2 (ACB09872)	3-Chlorobenzoate-3,4/4,5-dioxygenase (CbaA) from <i>Comamonas testosteroni</i> BR60	AAC45716	25.35	65
<i>Thermotoga</i> sp. Xyl154 (KHC95729)	Dicamba O-demethylase (DdmC) from <i>Stenotrophomonas maltophilia</i> DI-6	AAV53699	32.72	73
Class D*				
<i>Caldilinea aerophila</i> DSM 14535 (BAL99910)	Phenylpropionate dioxygenase from <i>Phenylbacterium zucineum</i> HLK1 ^a	WP_012521516	36.50	71
<i>Roseiflexus castenholzii</i> DSM 13941 (ABU59830)	(2Fe-2S)-binding protein from <i>Parvibaculum lavamentivorans</i> DS-1T ^a	WP_011995073	33.92	75
<i>Roseiflexus</i> sp. RS-1 (WP_011955741)	(2Fe-2S)-binding protein from <i>Parvibaculum lavamentivorans</i> DS-1T ^a	WP_011995074	33.63	81
<i>Sphaerobacter thermophilus</i> DSM 20745 (ACZ37596)	Aromatic oxygenase (GbcA) from <i>Pseudomonas aeruginosa</i> PAO1	NP_254097	31.36	88
<i>Thermomicrobium roseum</i> DSM 5159 (ACM04599)	Aromatic oxygenase (GbcA) from <i>Pseudomonas aeruginosa</i> PAO1	NP_254097	29.47	91
<i>Meiothermus cerbereus</i> DSM 11376 (WP_027876468)	(2Fe-2S)-binding protein from <i>Parvibaculum lavamentivorans</i> DS-1T ^a	WP_011995073	26.91	80
<i>Meiothermus chliarophilus</i> DSM 9957 (WP_027893118)	(2Fe-2S)-binding protein from <i>Parvibaculum lavamentivorans</i> DS-1T ^a	WP_011995073	28.25	77
<i>Meiothermus ruber</i> DSM 1279 (AGK03951)	Rieske (2Fe-2S) domain-containing protein from <i>Parvibaculum lavamentivorans</i> DS-1 ^a	WP_011995073	40.68	76
<i>Meiothermus rufus</i> DSM 22234 (WP_027882714)	Rieske (2Fe-2S) domain-containing protein from <i>Parvibaculum lavamentivorans</i> DS-1 ^a	WP_011995073	26.84	84
<i>Meiothermus silvanus</i> DSM 9946 (ADH64647)	Rieske (2Fe-2S) domain-containing protein from <i>Parvibaculum lavamentivorans</i> DS-1 ^a	WP_011995073	29.83	80
<i>Meiothermus timidus</i> DSM 17022 (WP_018466224)	Rieske (2Fe-2S) domain-containing protein from <i>Parvibaculum lavamentivorans</i> DS-1 ^a	WP_011995073	27.81	78

^a Biochemically uncharacterized protein

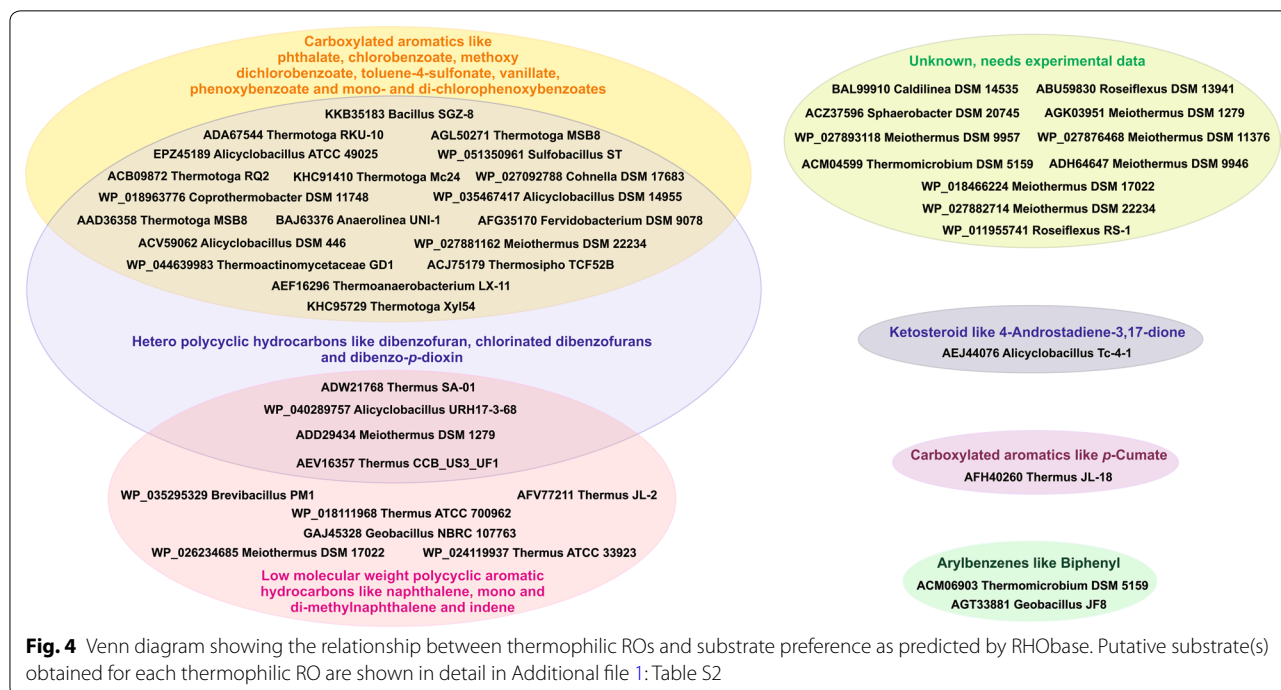


Table 4 Constituent components, obtained from each thermophile genome, that can be coupled with α -subunit of the terminal oxygenase component in order to reconstitute functional RO system

Organism (NCBI accession no. of the putative RO terminal oxygenase α -subunit)	Protein annotation [accession no.]	Putative function
Class A		
<i>Thermomicrobium roseum</i> DSM 5159 (ACM06903)	3-Phenylpropionate dioxygenase subunit beta [ACM06517]	Beta subunit
	Rieske protein [ACM07109]	Ferredoxin
	4Fe–4S ferredoxin, iron–sulfur binding domain protein [ACM06911]	
	Ferredoxin–NADP reductase; thioredoxin reductase [ACM06977]	Reductase
	Putidaredoxin reductase [ACM07176] Thioredoxin reductase [ACM06488]	
<i>Meiothermus ruber</i> DSM 1279 (ADD29434)	Aromatic-ring-hydroxylating dioxygenase beta subunit [ADD29433]	Beta subunit
	Rieske (2Fe–2S) domain protein [ADD27546]	Ferredoxin
	Dihydroliipoamide dehydrogenase [ADD29074]	Reductase
<i>Meiothermus timidus</i> DSM 17022 (WP_026234685)	Aromatic-ring-hydroxylating dioxygenase subunit beta [WP_018466297]	Beta subunit
	Diguanylate cyclase [WP_026234838]	Ferredoxin
	Dihydroliipoamide dehydrogenase [WP_026234742]	Reductase
	Dihydroliipoamide dehydrogenase [WP_026234716]	
	hypothetical protein [WP_018465799]	
<i>Thermus igniterrae</i> ATCC 700962 (WP_018111968)	Aromatic-ring-hydroxylating dioxygenase subunit beta [WP_018111967]	Beta subunit
	<i>Thermus oshimai</i> JL-2 (AFV77211)	Small subunit of phenylpropionate dioxygenase [AFV77212]
Ferredoxin subunit of nitrite reductase and ring-hydroxylating dioxygenase [AFV75193]		Ferredoxin
NAD(FAD)-dependent dehydrogenase [AFV75199]		Reductase
Dihydroliipoamide dehydrogenase [AFV77091]		
Flavoprotein oxygenase [WP_016329217]		
<i>Thermus scotoductus</i> SA-01 (ADW21768)	3-Phenylpropionate dioxygenase, subunit beta [ADW21767]	Beta subunit
	Ferredoxin subunit of phenylpropionate dioxygenase [ADW21026]	Ferredoxin
	4Fe–4S Ferredoxin/formate dehydrogenase, nitrate-inducible, iron–sulfur subunit [ADW21629]	
	NADH oxidase/coenzyme A disulfide reductase [ADW21032]	Reductase
<i>Thermus</i> sp. CCB_US3_UF1 (AEV16357)	Phenol hydroxylase component B [ADW21668]	
	<i>Thermus</i> sp. CCB_US3_UF1 (AEV16357)	Naphthalene dioxygenase small subunit [AEV16358]
Initial dioxygenase ferredoxin subunit [AEV16981]		Ferredoxin
4-Hydroxybenzoate 3-monooxygenase [AEV16360]		Reductase
FAD-dependent pyridine nucleotide-disulfide oxidoreductase [AEV16975]		
4-Hydroxyphenylacetate 3-monooxygenase reductase (Flav_red) [AEV16285]		
<i>Thermus thermophilus</i> ATCC 33923 (WP_024119937)	NADH oxidase [WP_024119527]	Reductase
	<i>Alicyclobacillus hesperidum</i> URH17-3-68 (WP_040289757)	Thioesterase superfamily protein [EJY55215]
Dihydropteridine reductase [WP_006448064]		Reductase
Dihydroliipoamide dehydrogenase [WP_006446492]		
NADH dehydrogenase [WP_006448257] Thioredoxin reductase [WP_006447474];		
<i>Brevibacillus thermoruber</i> PM1 (WP_035295329)	Aromatic-ring-hydroxylating dioxygenase [WP_035295713]	Beta subunit
	Dihydroliipoamide dehydrogenase [WP_035297436]	Reductase
<i>Geobacillus</i> sp. JF8 (AGT33881)	Small subunit of biphenyl dioxygenase [AGT33882]	Beta subunit
	Nitrite reductase NAD(P)H small subunit [AGT32189]	Ferredoxin
	Thioredoxin reductase [AGT33412]	Reductase
<i>Geobacillus thermoglucosidasius</i> NBRC 107763 (GAJ45328)	Putative naphthalene 1,2-dioxygenase small subunit [GAJ45329]	Beta subunit
	Hypothetical protein GT2_34_00240 [GAJ45383]	Ferredoxin
	Thioredoxin reductase [GAJ43870]	Reductase
	Coenzyme A disulfide reductase [GAJ42754]	
	Putative oxidoreductase [GAJ45381]	

Table 4 continued

Organism (NCBI accession no. of the putative RO terminal oxygenase α -subunit)	Protein annotation [accession no.]	Putative function
Class B		
<i>Thermus thermophilus</i> JL-18 (AFH40260)	Small subunit of phenylpropionate dioxygenase [AFH40261]	Beta subunit
	Ferredoxin subunit of nitrite reductase and ring-hydroxylating dioxygenase [AFH39800]	Ferredoxin
	NAD(FAD)-dependent dehydrogenase [AFH39794] NAD(FAD)-dependent dehydrogenase [AFH40137]	Reductase
Class D		
<i>Anaerolinea thermophila</i> UNI-1 (BAJ63376)	Putative ferredoxin [BAJ63818]	Ferredoxin
	Phytoene dehydrogenase [BAJ63379]	Reductase
<i>Meiothermus rufus</i> DSM 22234 (WP_027881162)	Dehydrogenase [WP_027881168]	Reductase
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> DSM 446 (ACV59062)	Rieske (2Fe–2S) iron–sulphur domain protein [ACV58029]	Ferredoxin
	Arsenate reductase (thioredoxin) [ACV59045]	Reductase
	Flavin reductase domain protein FMN-binding [ACV59470] Flavin reductase domain protein FMN-binding [ACV58921]	
<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> Tc-4-1 (AEJ44076)	Rieske (2Fe–2S) domain protein [AEJ42957]	Ferredoxin
	Dihydroliipoamide dehydrogenase [AEJ42545]	Reductase
	Flavin reductase domain protein FMN-binding protein [AEJ43919]	
<i>Alicyclobacillus acidoterrestris</i> ATCC 49025 (EPZ45189)	Hypothetical protein N007_08015 [EPZ45711]	Reductase
	Hypothetical protein [EPZ45711]	
<i>Alicyclobacillus pomorum</i> DSM 14955 (WP_035467417)	Dihydropteridine reductase [WP_035467306]	Reductase
<i>Bacillus thermotolerans</i> SGZ-8 (KKB35183)	Absent	
<i>Cohnella thermotolerans</i> DSM 17683 (WP_027092788)	Ferredoxin [WP_027093665]	Ferredoxin
<i>Coprothermobacter platensis</i> DSM 11748 (WP_018963776)	2-Enoate reductase [WP_018963737]	Reductase
	Hypothetical protein [WP_018963777]	
<i>Sulfobacillus thermosulfidooxidans</i> ST (WP_051350961)	(2Fe–2S)-binding protein [WP_037913512]	Ferredoxin
	Hypothetical protein, partial [WP_040767264]	Reductase
Thermoactinomycetaceae bacterium GD1 (WP_044639983)	Absent	
<i>Thermoanaerobacterium xylanolyticum</i> LX-11 (AEF16296)	Rubredoxin-type Fe(Cys) ₄ protein [AEF16485]	Ferredoxin
	Pyruvate ferredoxin/flavodoxin oxidoreductase [AEF16297]	Reductase
	CoA-disulfide reductase [AEF16270] Thioredoxin reductase [AEF16306]	
<i>Fervidobacterium pennivorans</i> DSM 9078 (AFG35170)	Rubredoxin [AFG34604]	Ferredoxin
	2-Polyprenylphenol hydroxylase-like oxidoreductase [AFG34955]	Reductase
	Phytoene dehydrogenase-like oxidoreductase [AFG35173]	
	Thioredoxin-disulfide reductase [AFG35278]	
	FAD binding protein [AFG35124] Thioredoxin reductase [AFG34984]	
	Dihydroliipoamide dehydrogenase [AFG34941] Thioredoxin reductase [AFG35081]	
<i>Thermosipho africanus</i> TCF52B (ACJ75179)	Rubredoxin [ACJ76024]	Ferredoxin
	Oxidoreductase/NADH: ubiquinone oxidoreductase, na translocating, f-subunit [ACJ75165]	Reductase
	Oxidoreductase/tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA [ACJ75137]	
	Ferredoxin/anaerobic glycerol 3-phosphate dehydrogenase [ACJ75136]	
<i>Thermotoga maritima</i> MSB8 (AAD36358)	Rubredoxin [AGL49584]	Ferredoxin
	Oxidoreductase [AAD35836]	Reductase
<i>Thermotoga maritima</i> MSB8 (AGL50271)	Rubredoxin [AGL49584]	Ferredoxin
	Oxidoreductase [AAD35836]	Reductase
<i>Thermotoga naphthophila</i> RKU-10 (ADA67544)	Rubredoxin-type Fe(Cys) ₄ protein [ADA66551]	Ferredoxin
	FAD-dependent pyridine nucleotide-disulphide oxidoreductase [ADA67462]	Reductase

Table 4 continued

Organism (NCBI accession no. of the putative RO terminal oxygenase α -subunit)	Protein annotation [accession no.]	Putative function
<i>Thermotoga</i> sp. Mc24 (KHC91410)	Rubredoxin-type Fe(Cys) ₄ protein [KHC90328]	Ferredoxin
	Oxidoreductase [KHC90231]	Reductase
<i>Thermotoga</i> sp. RQ2 (ACB09872)	Rubredoxin-type Fe(Cys) ₄ protein [ACB08629]	Ferredoxin
	FAD-dependent pyridine nucleotide-disulphide oxidoreductase [ACB08534]	Reductase
<i>Thermotoga</i> sp. Xyl54 (KHC95729)	Rubredoxin-type Fe(Cys) ₄ protein [KHC96420]	Ferredoxin
Class D*		
<i>Caldilinea aerophila</i> DSM 14535 (BAL99910)	Hypothetical protein [BAL99023]	Ferredoxin
	Putative mercuric reductase [BAM01871]	Reductase
	Putative flavin reductase [BAL99068]	
<i>Roseiflexus castenholzii</i> DSM 13941 (ABU59830)	Rieske (2Fe–2S) domain protein [ABU56444]	Ferredoxin
	Dihydroliipoamide dehydrogenase [ABU58189]	Reductase
	Flavin reductase domain protein FMN-binding [ABU60217]	
	Flavin reductase domain protein FMN-binding [ABU58489]	
<i>Roseiflexus</i> sp. RS-1 (WP_011955741)	Rieske (2Fe–2S) domain protein [ABQ89132]	Ferredoxin
	FAD-dependent pyridine nucleotide-disulphide oxidoreductase [ABQ91503]	Reductase
	Dihydroliipoamide dehydrogenase [ABQ91367]	
	Flavin reductase domain protein, FMN-binding [ABQ92660]	
	Flavin reductase domain protein, FMN-binding [ABQ90950]	
<i>Sphaerobacter thermophilus</i> DSM 20745 (ACZ37596)	Rieske (2Fe–2S) iron–sulphur domain protein [ACZ40633]	Ferredoxin
	FAD-dependent pyridine nucleotide-disulphide oxidoreductase [ACZ40632]	Reductase
	Flavin reductase domain protein FMN-binding protein [ACZ39977]	
<i>Thermomicrobium roseum</i> DSM 5159 (ACM04599)	Rieske protein [ACM07109]	Ferredoxin
	Xylene monooxygenase electron transfer subunit [ACM06651]	Reductase
	Ferredoxin–NADP reductase; thioredoxin reductase [ACM06977]	
	Putidaredoxin reductase [ACM07176]	
	Thioredoxin reductase [ACM06488]	
<i>Meiothermus cerbereus</i> DSM 11376 (WP_027876468)	Diguanylatecyclase [WP_027876040]	Ferredoxin
	Dihydroliipoamide dehydrogenase [WP_027876580]	Reductase
	MFS transporter [WP_027877386]	
<i>Meiothermus chliarophilus</i> DSM 9957 (WP_027893118)	Dihydroliipoamide dehydrogenase [WP_027893501]	Reductase
	MFS transporter [WP_036218580]	
<i>Meiothermus ruber</i> DSM 1279 (AGK03951)	Dihydroliipoamide dehydrogenase [ADD29074]	Reductase
	Flavin reductase domain-containing FMN-binding protein [AGK04562]	
<i>Meiothermus rufus</i> DSM 22234 (WP_027882714)	Dehydrogenase [WP_027881168]	Reductase
<i>Meiothermus silvanus</i> DSM 9946 (ADH64647)	Rieske (2Fe–2S) iron–sulfur domain protein [ADH62911]	Ferredoxin
	FAD-dependent pyridine nucleotide-disulfide oxidoreductase [ADH64702]	Reductase
	Flavin reductase domain protein FMN-binding protein [ADH64172]	
<i>Meiothermus timidus</i> DSM 17022 (WP_018466224)	Dihydroliipoamide dehydrogenase [WP_026234742]	Reductase
	Dihydroliipoamide dehydrogenase [WP_026234716]	
	Hypothetical protein [WP_018465799]	

neither of which contained any known component apart from the α -subunit.

Discussion

Owing to their extensive presence among Proteobacteria and Actinobacteria, all extant ROs are postulated to have originated and evolved within these lineages (Chakraborty et al. 2014). Though the thermophilic RO homologs identified in this study were present among taxonomically close organisms, their distributions were very specific for certain phyla, often with very low

abundance (Fig. 1). Thus, it is quite likely that the thermophiles also acquired RO_{ox} genes from Proteobacteria or Actinobacteria and further evolved separately. The role of Firmicutes in the evolution of ROs in thermophiles cannot be ruled out, as several RO homologs were found among mesophilic strains of Firmicutes (Fig. 1), especially within the order Bacillales.

Similarly, it would be difficult to claim that only the distantly located ET components identified in this study complement the putative α - and β -subunit genes. It is highly likely that unknown gene(s) vicinal to those

encoding the terminal oxygenase(s) are responsible for the proper functioning of the ROs. Similar observations have previously been made in *Rhodococcus opacus* TKN14, in which rubredoxin and another hypothetical protein were found to be crucial for the oxidation of *o*-xylene (Maruyama et al. 2005). In another study, the purified large subunit of a novel alkane monooxygenase (belonging to Class B ROs), identified from a cold-tolerant *Pusillimonas* sp. T7-7, showed NADH-dependent alkane monooxygenase activity (Li et al. 2013). Transformation of aromatic hydrocarbons has also been attained by heterologous expression of the terminal oxygenase components using non-specific ET proteins complemented by the host strain (Mukerjee-Dhar et al. 2005). Although the present study indicates that the RO homologs present in these organisms are either cryptic in nature or are involved in some other physiological function, we cannot rule out the possibility of reconstituting a thermostable functional RO system with novel properties (in terms of substrate preference or mode of catalysis) by combining the terminal oxygenase genes along with all possible combinations of ET components. Integrity of the motif signatures and predicted enhanced thermostability (Fig. 3) further strengthens this hypothesis.

The existence of unexplored microbial diversity, together with the availability of whole genomes, represents a large pool for future industrial catalysts. Thermostable ROs may be attractive candidates for carrying out efficient biotransformation at elevated temperatures. Apart from enhancing our understanding of the distribution of ROs in nature, the present study may aid in designing new bioremediation strategies or industrial biosynthetic processes. Based on the information provided here, functional RO systems can be reconstituted from each organism by cloning both terminal oxygenase and ET genes into suitable vectors and performing biotransformation assays using the predicted substrates (Fig. 4). On the other hand, gene knockout studies can be performed (provided that appropriate genetic tools are available) to help elucidate the physiological role of RO homologs with unknown functions in thermophilic bacteria.

Additional file

Additional file 1. Additional tables.

Abbreviations

RO: Rieske non-heme iron oxygenase; RO_{ox}: large (α) subunit of RO terminal oxygenase; ET: electron transport; NCBI: National Center for Biotechnology Information; NAD(P)H: reduced form of nicotinamide adenine dinucleotide phosphate.

Authors' contributions

All the authors conceived the research work. JC conducted the analyses. JC, CSM, KO and HN analyzed the data. JC drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

JC extends his thanks to the Biotechnology Research Center, The University of Tokyo, for providing research facilities.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional file.

Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Funding

This study was funded by the Japan Society for the Promotion of Science (JSPS). JC was supported by postdoctoral fellowship for foreign researchers from the JSPS.

Received: 8 November 2016 Accepted: 23 December 2016

Published online: 04 January 2017

References

- Allen CR, Boyd DR, Dalton H, Sharma ND, Brannigan I, Kerley NA, Shelldrake NG, Taylor SC (1995) Enantioselective bacterial biotransformation routes to *cis*-diol metabolites of monosubstituted benzenes, naphthalene and benzocycloalkenes of either absolute configuration. *Chem Commun* 2:117–118. doi:10.1039/C3995000117
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. doi:10.1016/S0022-2836(05)80360-2
- Berim A, Park JJ, Gang DR (2014) Unexpected roles for ancient proteins: flavone 8-hydroxylase in sweet basil trichomes is a Rieske-type, PAO-family oxygenase. *Plant J* 80:385–395. doi:10.1111/tpj.12642
- Boyd DR, Sharma ND, O'Dowd CR, Carroll JG, Loke PL, Allen CCR (2005) *cis*-Dihydrodiol, arene oxide and phenol metabolites of dictamnine: key intermediates in the biodegradation and biosynthesis of furoquinoline alkaloids. *Chem Commun* 31:3989–3991. doi:10.1039/b506944k
- Bui VP, Hudlicky T, Hansen TV, Stenstrom Y (2002) Direct biooxidation of arenes to corresponding catechols with *E. coli* JM109(pDTG602). Application to synthesis of combretastatins A-1 and B-1. *Tetrahedron Lett* 43:2839–2841. doi:10.1016/S0040-4039(02)00389-1
- Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghegan NS, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273:1058–1073. doi:10.1126/science.273.5278.1058
- Caliebe A, Grimm R, Kaiser G, Lübeck J, Soll J, Heins L (1997) The chloroplastic protein import machinery contains a Rieske-type iron–sulfur cluster and a mononuclear iron-binding protein. *EMBO J* 16:7342–7350. doi:10.1093/emboj/16.24.7342
- Chakraborty J, Ghosal D, Dutta A, Dutta TK (2012) An insight into the origin and functional evolution of bacterial aromatic ring-hydroxylating oxygenases. *J Biomol Struct Dyn* 30:419–436. doi:10.1080/07391102.2012.682208
- Chakraborty J, Jana T, Saha S, Dutta TK (2014) Ring-hydroxylating oxygenase database: a database of bacterial aromatic ring-hydroxylating oxygenases in the management of bioremediation and biocatalysis of aromatic compounds. *Environ Microbiol Rep* 6:519–523. doi:10.1111/1758-2229.12182
- de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34:W362–W365. doi:10.1093/nar/gkl124

- El-Sayed AK, Hothersall J, Cooper SM, Stephens E, Simpson TJ, Thomas CM (2003) Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. *Chem Biol* 10:419–430. doi:10.1016/S1074-5521(03)00091-7
- Ensley BD, Ratzkin BJ, Osslund TD, Simon MJ, Wackett LP, Gibson DT (1983) Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* 222:167–169. doi:10.1126/science.6353574
- Feitkenhauer H, Märk H (2003) Biodegradation of aliphatic and aromatic hydrocarbons at high temperatures. *Water Sci Technol* 47:123–130
- Ferraro DJ, Gakhar L, Ramaswamy S (2005) Rieske business: structure-function of Rieske non-heme oxygenases. *Biochem Biophys Res Commun* 338:175–190. doi:10.1016/j.bbrc.2005.08.222
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana Press, New York City, pp 571–607. doi:10.1385/1-59259-890-0:571
- Gibson DT, Parales RE (2000) Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr Opin Biotechnol* 11:236–243. doi:10.1016/S0958-1669(00)0090-2
- Gibson DT, Subramanian V (1984). Microbial degradation of aromatic hydrocarbons. In: Gibson DT (ed) *Microbial degradation of organic compounds*. Dekker, Inc., New York, pp 181–252. doi:10.1002/actp.1986.010371127
- Gibson DT, Koch JR, Kallio RE (1968) Oxidative degradation of aromatic hydrocarbons by microorganisms. I. Enzymatic formation of catechol from benzene. *Biochemistry* 7:2653–2662. doi:10.1021/bi00847a031
- Habe H, Omori T (2003) Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. *Biosci Biotechnol Biochem* 67:225–243. doi:10.1271/bbb.67.225
- Hedlund BP, Dodsworth JA, Murugapiran SK, Rinke C, Woyke T (2014) Impact of single-cell genomics and metagenomics on the emerging view of extremophile “microbial dark matter”. *Extremophiles* 18:865–875. doi:10.1007/s00792-014-0664-7
- Hudlicky T, Gonzalez D, Gibson DT (1999) Enzymatic dihydroxylation of aromatics in enantioselective synthesis: expanding asymmetric methodology. *Aldrichimica Acta* 32:35–62
- Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88:1895–1898
- Iwai S, Chai B, Sul WJ, Cole JR, Hashsham SA, Tiedje JM (2010) Gene-targeted metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. *ISME J* 4:279–285. doi:10.1038/ismej.2009.104
- Iwai S, Johnson TA, Chai B, Hashsham SA, Tiedje JM (2011) Comparison of the specificities and efficacies of primers for aromatic dioxygenase gene analysis of environmental samples. *Appl Environ Microbiol* 77:3551–3557. doi:10.1128/AEM.00331-11
- Jiang H, Parales RE, Lynch NA, Gibson DT (1996) Site-directed mutagenesis of conserved amino acids in the alpha subunit of toluene dioxygenase: potential mononuclear non-heme iron coordination sites. *J Bacteriol* 178:3133–3139
- Li P, Wang L, Feng L (2013) Characterization of a novel Rieske-type alkane monooxygenase system in *Pusillimonas* sp. strain T7-7. *J Bacteriol* 195:892–1901. doi:10.1128/JB.02107-12
- Mallick S, Chakraborty J, Dutta TK (2011) Role of oxygenases in guiding diverse metabolic pathways in the bacterial degradation of low-molecular weight polycyclic aromatic hydrocarbons: a review. *Crit Rev Microbiol* 37:64–90. doi:10.3109/1040841X.2010.512268
- Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res* 30:281–283. doi:10.1093/nar/gkg087
- Margesin R, Schinner F (2001) Biodegradation and bioremediation of hydrocarbons in extreme environments. *Appl Microbiol Biotechnol* 56:650–663. doi:10.1007/s002530100701
- Maruyama T, Ishikura M, Taki H, Shindo K, Kasai H, Haga M, Inomata Y, Misawa N (2005) Isolation and characterization of o-xylene oxygenase genes from *Rhodococcus opacus* TKN14. *Appl Environ Microbiol* 71:7705–7715. doi:10.1128/AEM.71.12.7705-7715.2005
- Mason JR, Cammack R (1992) The electron-transport proteins of hydroxylating bacterial dioxygenases. *Annu Rev Microbiol* 46:277–305. doi:10.1146/annurev.mi.46.100192.001425
- Meng YL, Wang YM, Zhang B, Nii N (2001) Isolation of a choline monooxygenase cDNA clone from *Amaranthus tricolor* and its expressions under stress conditions. *Cell Res* 11:187–193. doi:10.1038/sj.cr.7290085
- Mukerjee-Dhar G, Shimura M, Miyazawa D, Kimbara K, Hata T (2005) *bph* genes of the thermophilic PCB degrader, *Bacillus* sp. JF8: characterization of the divergent ring-hydroxylating dioxygenase and hydrolase genes upstream of the Mn-dependent BphC. *Microbiology* 151:4139–4151. doi:10.1099/mic.0.28437-0
- Newman LM, Garcia H, Hudlicky T, Selifonov SA (2004) Directed evolution of the dioxygenase complex for the synthesis of furanone flavor compounds. *Tetrahedron* 60:729–734. doi:10.1016/j.tet.2003.10.105
- Parales RE (2003) The role of active-site residues in naphthalene dioxygenase. *J Ind Microbiol Biotechnol* 30:271–278. doi:10.1007/s10295-003-0043-3
- Perfumo A, Banat IM, Marchant R, Vezzulli L (2007) Thermally enhanced approaches for bioremediation of hydrocarbon-contaminated soils. *Chemosphere* 66:179–184. doi:10.1016/j.chemosphere.2006.05.006
- Reinbothe S, Quigley F, Gray J, Schemenewitz A, Reinbothe C (2004) Identification of plastid envelope proteins required for import of protochlorophyllide oxidoreductase A into the chloroplast of barley. *Proc Natl Acad Sci USA* 101:2197–2202. doi:10.1073/pnas.0307284101
- Rottiers V, Motola DL, Gerisch B, Cummins CL, Nishiwaki K, Mangelsdorf DJ, Antebi A (2006) Hormonal control of *C. elegans* dauer formation and life span by a Rieske-like oxygenase. *Dev Cell* 10:473–482. doi:10.1016/j.devcel.2006.02.008
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Schäfer F, Schuster J, Würz B, Härtig C, Harms H, Müller RH, Rohwerder T (2012) Synthesis of short-chain diols and unsaturated alcohols from secondary alcohol substrates by the Rieske nonheme mononuclear iron oxygenase MdpJ. *Appl Environ Microbiol* 78:6280–6284. doi:10.1128/AEM.01434-12
- Shintani M, Ohtsubo Y, Fukuda K, Hosoyama A, Ohji S, Yamazoe A, Fujita N, Nagata Y, Tsuda M, Hata T, Kimbara K (2014) Complete genome sequence of the thermophilic polychlorinated biphenyl degrader *Geobacillus* sp. strain JF8 (NBRC 109937). *Genome Announc*. doi:10.1128/genome.A.01213-13
- Sydor PK, Barry SM, Odulate OM, Barona-Gomez F, Haynes SW, Corre C, Song L, Challis GL (2011) Regio- and stereodivergent antibiotic oxidative carbocyclizations catalysed by Rieske oxygenase-like enzymes. *Nat Chem* 3:388–392. doi:10.1038/nchem.1024
- 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. doi:10.1038/nchem.1024
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882. doi:10.1093/nar/25.24.4876
- Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* 65:1–43
- Wackett LP, Kwart LD, Gibson DT (1988) Benzylic monooxygenation catalyzed by toluene dioxygenase from *Pseudomonas putida*. *Biochemistry* 27:1360–1367. doi:10.1021/bi00404a041
- Wargo MJ, Szwergold BS, Hogan DA (2008) Identification of two gene clusters and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism. *J Bacteriol* 190:2690–2699. doi:10.1128/JB.01393-07
- Yoshiyama T, Namiki T, Mita K, Kataoka H, Niwa R (2006) Neverland is an evolutionarily conserved Rieske-domain protein that is essential for ecdysone synthesis and insect growth. *Development* 133:2565–2574. doi:10.1242/dev.02428
- Yoshiyama-Yanagawa T, Enya S, Shimada-Niwa Y, Yaguchi S, Haramoto Y, Matsuya T, Shiomi K, Sasakura Y, Takahashi S, Asashima M, Kataoka H, Niwa R (2011) The conserved Rieske oxygenase DAF-36/Neverland is a novel cholesterol-metabolizing enzyme. *J Biol Chem* 286:25756–25762. doi:10.1074/jbc.M111.244384
- Zeveloff J, Hudlicky T (2005) Recent progress in the synthesis of morphine alkaloids. *Synlett* 3:388–405. doi:10.1055/s-2005-862383