Environmental Microbiology

ENVIRONMENTAL MICROBIOLOGY

MINI REVIEW OPEN ACCESS

Phylogenetic and Functional Diversity of Soluble Di-Iron Monooxygenases

Sui Nin Nicholas Yang[1] | Michael A. Kertesz[1] | Nicholas V. Coleman[2]

[1]School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia | [2]Australian Genome Foundry and ARC Centre of Excellence in Synthetic Biology, Macquarie University, Sydney, New South Wales, Australia

Correspondence: Nicholas V. Coleman (nicholas.coleman@mq.edu.au)

Received: 12 September 2024 | **Revised:** 23 December 2024 | **Accepted:** 15 January 2025

Funding: This work was supported by University of Sydney.

Keywords: biocatalysis | biodegradation | bioremediation | hydrocarbon | monooxygenase

ABSTRACT

Monooxygenase (MO) enzymes are responsible for the oxidation of hydrocarbons and other compounds in the carbon and nitrogen cycles, are important for the biodegradation of pollutants and can act as biocatalysts for chemical manufacture. The soluble di-iron monooxygenases (SDIMOs) are of interest due to their broad substrate range, high enantioselectivity and ability to oxidise inert substrates such as methane. Here, we re-examine the phylogeny and functions of these enzymes, using recent advances in the field and expansions in sequence diversity in databases to highlight relationships between SDIMOs and revisit their classification. We discuss the impact of horizontal gene transfer on SDIMO phylogeny, the potential of SDIMOs for the biodegradation of pollutants and the importance of heterologous expression as a tool for understanding SDIMO functions and enabling their use as biocatalysts. Our analysis highlights current knowledge gaps, most notably, the unknown substrate ranges and physiological roles of enzymes that have so far only been detected via genome or metagenome sequencing. Enhanced understanding of the diversity and functions of the SDIMO enzymes will enable better prediction and management of biogeochemical processes and also enable new applications of these enzymes for biocatalysis and bioremediation.

1 | Introduction

Hydrocarbons are released into the environment by both natural and industrial processes and are of interest as intermediates in global carbon cycles (Koo and Rosenzweig 2021; Tucci and Rosenzweig 2024), as persistent pollutants (Shennan 2005) and as useful substrates for biocatalysis (Donadio et al. 2015). Bacteria can use most hydrocarbons as carbon and energy sources, provided that environmental conditions are appropriate, but they require specialised enzymes for their metabolism. In the case of aromatic hydrocarbons, dioxygenase enzymes are the typical initial catalysts under aerobic conditions (Gibson and Parales 2000), while for aliphatic hydrocarbons (alkanes and alkenes), monooxygenase (MO) enzymes play this role (Shennan 2005; Guo et al. 2023). This review will focus on the MO enzymes, which have significant roles in biogeochemistry (Greening and Grinter 2022) and many interesting applications in biotechnology (Canada et al. 2002; Donadio et al. 2015).

Bacteria expressing MO enzymes are very important for the bioremediation of pollutants. A recent, well-studied example is the breakdown of 1,4-dioxane, a common groundwater pollutant, which occurs via both co-metabolic (Wang et al. 2021b) and growth-linked (Mahendra and Alvarez-Cohen 2005; He et al. 2017) processes in MO-containing bacteria. There is also very extensive research on the use of MO enzymes for the biodegradation and bioremediation of chloroethenes such as trichloroethene (TCE), which are also common groundwater pollutants (Shennan 2005; Mattes, Alexander, and Coleman 2010). As with 1,4-dioxane, both co-metabolic (Ojo et al. 2023) and

growth-linked (Jin and Mattes 2008) physiologies are associated with the degradation of chloroethenes, depending on the substrate and the bacterial host.

MOs are prized for their applications in biocatalysis, due to their ability to add oxygen to substrates with high chemo-, regio- and stereo-specificity under mild reaction conditions (Constable et al. 2007; Que Jr. and Tolman 2008). As a result, there is great interest in discovering and creating novel MOs that can catalyse these reactions for industrial and medicinal chemistry purposes (Constable et al. 2007; Leak et al. 2009; Bryan et al. 2018; Petkevičius et al. 2019). As will become increasingly clear throughout this review, a key theme for MO enzymes is 'diversity'. This applies to both the types of MOs that are useful for biocatalysis (these include iron, copper and flavin-containing enzymes) (Torres Pazmiño et al. 2010) and also the types of reactions that are catalysed (hydroxylation, epoxidation and desaturation) (Keener and Arp 1994).

Several different classes of MO can be found in Bacteria and Archaea, and a huge diversity of sequence types exists within each of these classes. Bacterial MOs have broad substrate ranges and provide diverse selective benefits. They can therefore be found in physiologically and taxonomically diverse bacteria (Osborne and Haritos 2019) and are often subject to horizontal gene transfer (HGT) via plasmids (Zou et al. 2021). The major exceptions to this rule are the particulate and soluble methane MOs (pMMOs, sMMOs), which are specialised for methane oxidation. These are more restricted in their taxonomic distribution and physiological roles (Kalyuzhnaya, Gomez, and Murrell 2019; Khider, Brautaset, and Irla 2021). For many years, it was believed that pMMO was the only example of a copper-containing membrane-located MO (CuMMO), but it is now clear that both CuMMOs and the related ammonia MOs (AMOs) also exist in non-methanotrophic lineages of Bacteria and Archaea (Coleman et al. 2012; Diamond et al. 2022).

The MOs are typically categorised based on a combination of their cellular location and the cofactors they require (Torres Pazmiño et al. 2010; Coleman et al. 2011), leading to the following classes: haem-containing MOs (cytochrome p450) (Bernhardt 2006), flavin-dependent MOs (van Berkel, Kamerbeek, and Fraaije 2006), CuMMOs (Koo et al. 2022; Tucci et al. 2023), pterin-dependent MOs (Zhao et al. 1994), di-iron membrane-located MOs (AlkB) (Ji et al. 2013; Guo et al. 2023), cofactor-independent MOs (Fetzner 2002) and soluble di-iron monooxygenases (SDIMOs) (Leahy, Batchelor, and Morcomb 2003). Alternatively, the MOs can be categorised based on their applications (biogeochemistry and/or bioremediation and/or biocatalysis, as described above) or their substrate range, but these classifications are less useful, since MOs typically have significant roles across multiple application areas and tend to have broad substrate ranges.

The relationship of MOs to the physiology of the host cell is complex. Some MO reactions are linked to productive metabolism (growth-linked), while others are incidental and do not yield carbon or energy for the cell (cometabolism) (Horvath 1972; Semprini 1997). These distinctions are significant for how MO-containing bacteria are deployed for different applications. The MOs covered in this review almost always act as the first step in a metabolic pathway, and so we have defined their 'primary substrate' (where known) as the compound used for the enrichment and isolation of the host bacterium. The primary substrate (as defined here) will therefore always be growth-linked, but this does not necessarily imply that this is the 'best' substrate in terms of enzyme affinity or turnover rate (e.g., sMMO oxidises ethylene more rapidly than methane (Colby, Stirling, and Dalton 1977), likely due to the chemical lability of ethylene vs. methane). Further, complexity is introduced when we consider a broad variety of compounds that can act as inducers of MO enzymes (i.e., inducing expression of the corresponding genes). These are often growth substrates for the respective bacteria, but they can also be downstream metabolites, co-metabolic substrates or unrelated compounds. These conceptual relationships are summarised in Figure 1.

The SDIMOs are multicomponent enzymes, requiring at least three components: a hydroxylase, a reductase and a coupling protein (Leahy, Batchelor, and Morcomb 2003; Osborne and Haritos 2019). The hydroxylase component is itself complex, containing two or three subunits, arranged as $\alpha_2\beta_2$ or $\alpha_2\beta_2\gamma_2$. Some SDIMOs also contain a ferredoxin or other accessory components (Zhou et al. 1998). The reductase extracts electrons from NADH and transfers them to the di-iron core in the hydroxylase α subunit, with assistance from the coupling protein (and ferredoxin, if present) (Wang et al. 2014). The hydroxylase then activates one oxygen atom in O_2 to a high-energy state, and this then attacks the substrate, while the other oxygen atom is reduced to H_2O (Stainthorpe et al. 1990; Cardy et al. 1991; Sullivan, Dickinson, and Chase 1998; Banerjee, Jones, and Lipscomb 2019). The product(s) generated are dependent on the substrate and the type of SDIMO.

The SDIMOs have been reviewed previously (Leahy, Batchelor, and Morcomb 2003; Notomista et al. 2003; Coleman, Bui, and Holmes 2006; Osborne and Haritos 2019) and classified into groups according to their protein sequences, the number of subunits, the operon arrangement and their substrate specificity. Recent massive expansions of sequence databases warrant a re-investigation of these enzymes. An up-to-date overview of SDIMO sequence diversity, groups and operon structures is given in Figure 2. A better understanding of the sequence diversity and evolutionary relationships of SDIMOs is needed to enable better predictions of their substrates, physiological roles and ecological significance. This review aims to update our definitions of SDIMO groups, highlight recent discoveries and experimental advances and pinpoint knowledge gaps for future work. We will first focus on phylogenetic analysis, with sections organised based on the previously defined groups 1 to 7 (Leahy, Batchelor, and Morcomb 2003; Notomista et al. 2003; Yang et al. 2024). After exploring the phylogeny of the SDIMOs, we will summarise progress on heterologous expression (this is an important tool for the development of biocatalysts). We will examine evidence for HGT of SDIMO genes, since this plays a large role in the evolution of hydrocarbon-degrading bacteria (Osborne and Haritos 2018; Zou et al. 2021). Finally, we will highlight recent examples of the use of SDIMOs for bioremediation of pollutants.

Some comments on nomenclature are important before beginning the review. First, there is a discrepancy in the

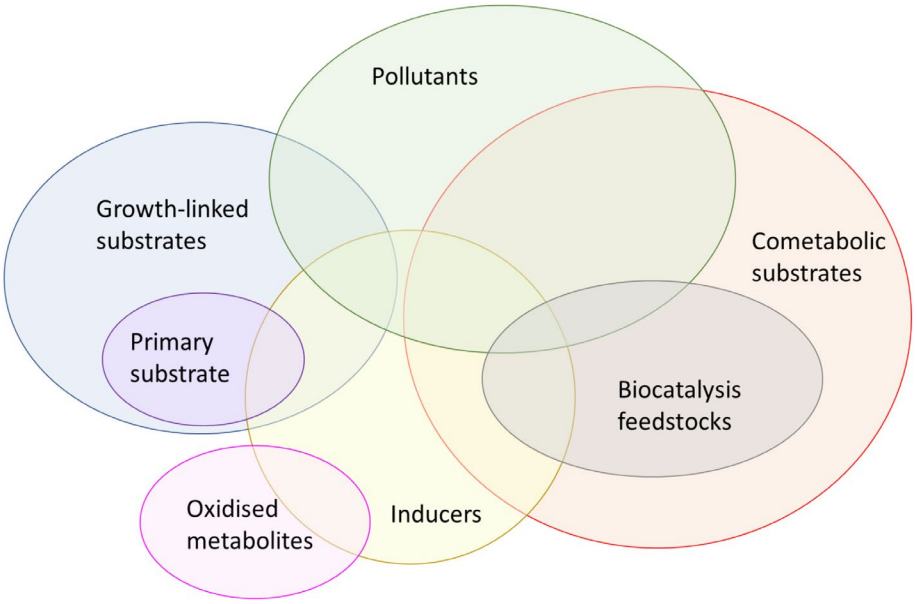

FIGURE 1 | Relationships between compounds that are substrates for monooxygenase enzymes, with respect to host cell physiology. These include substrates that support the growth of the host (including the 'primary substrate' that was used for initial isolation of the host bacterium) and those that are transformed co-metabolically. Environmental pollutants that are MO substrates may fall into either of these groups. Although many biocatalysis feedstocks are growth substrates, in a biomanufacturing situation, the MO reaction product is intentionally diverted from further metabolism, and so these are shown as not growth-linked. Inducers of MO gene expression include compounds from both groups of substrates but also metabolites and other compounds.

literature between SDIMO groups 1 and 2. These group numbers (Notomista et al. 2003) were switched in a later paper (Holmes and Coleman 2008), and this change was propagated in the literature (Ren et al. 2022; Yang et al. 2024). We will use the more recent naming scheme, where group 1 are the 'toluene MOs' (TMOs) and group 2 are the 'phenol MOs' (PhMOs). Second, to simplify language, we will assume that protein sequences encoded in genomes and metagenomes are expressed and functional, i.e., we will refer to them as 'enzymes'. Thirdly, we will use the phylogeny of the alpha subunit as a proxy for the holoenzymes, although typically this only holds true for the two parts of the hydroxylase component (aβ) and less so for the reductase (Leahy, Batchelor, and Morcomb 2003). Finally, we have used the most recent and correct terms for different phyla, so 'Proteobacteria' are now 'Pseudomonadota', 'Firmicutes' are 'Bacillota' and 'Actinobacteria' are 'Actinomycetota' (Oren and Garrity 2021). We have used the older genus name '*Mycobacterium*' here (Meehan et al. 2021) rather than splitting this genus into five new genera (Gupta, Lo, and Son 2018).

2 | Methods

A reference sequence from each SDIMO group (Table S1) was used for protein similarity searches (BLASTp) against the NCBI Protein database (Johnson et al. 2008), retaining all sequence results with over 80% coverage to the query sequence. The results were filtered further by establishing a percent amino acid (aa) identity value cutoff for each group. These cutoffs were set manually after inspection of the BLASTp results and depended on the level of homology within each SDIMO group (Table S1). In cases where >30 BLAST matches were retrieved, these sequences were clustered using CD-HIT (Fu et al. 2012), and one representative sequence of each cluster was retained. In cases where CD-HIT analysis yielded an uncharacterised SDIMO as a reference sequence, this was replaced with the sequence of a characterised enzyme with sequence as close as possible to the one suggested by CD-HIT. The sequences were assembled in FASTA format, aligned using ClustalX (Larkin et al. 2007) and then exported to GeneDoc (Nicholas and Nicholas 1997) for manual trimming to give uniform lengths. Internal gaps were retained. PhyML 3.0 online execution (Guindon et al. 2010) was used to generate the trees using the Smart Model Selection option (Lefort, Longueville, and Gascuel 2017) and the aLRT SH-like fast-likelihood-based method for branch supports. MEGA11 (Tamura, Stecher, and Kumar 2021) was then used to visualise and annotate the trees. See Table S1 for further methods details. The accession numbers for all represented alpha hydroxylase sequences in each phylogenetic tree can be found in Table S3.

3 | Overview of Phylogeny and Functions of Different SDIMO Groups

3.1 | Group 1 SDIMOs—TMOs

The group 1 SDIMOs (Figures 2 and 3) are commonly referred to as TMOs. We have kept this abbreviation for convenience, but it is important to note that their substrate range is very broad and also includes propene, isoprene, isobutene (2-methyl-propene), phenol, *o*-xylene and 1,4-dioxane as physiological (i.e., growth-supporting) substrates (Table S2). The TMO operons are composed of six genes, arranged as follows: α-hydroxylase, γ-hydroxylase, ferredoxin, coupling protein, β-hydroxylase and reductase. The TMOs are abundantly represented in sequence databases (>600 sequences, Table S1) and are well-characterised

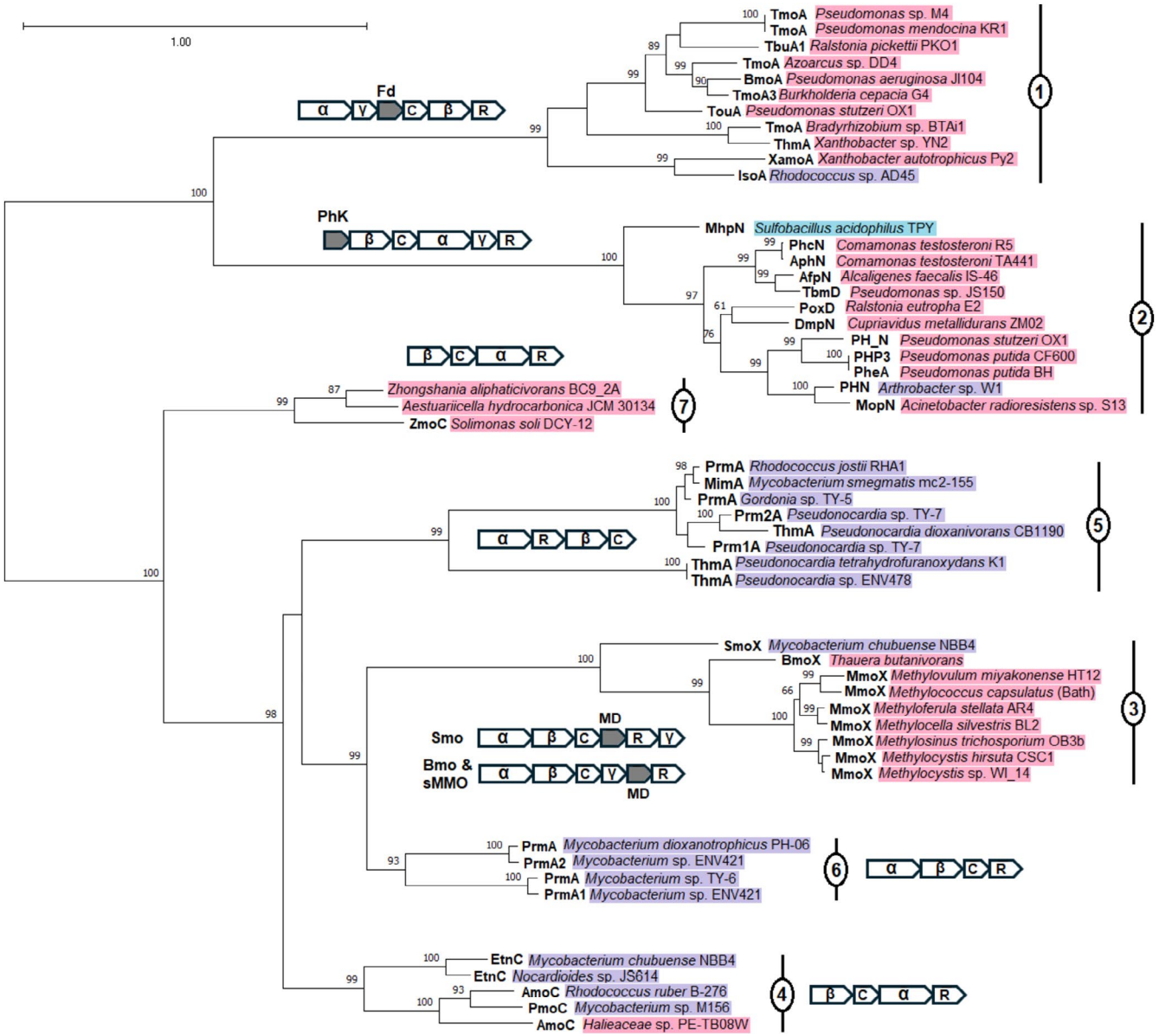

FIGURE 2 | Evolutionary relationships of SDIMO enzymes. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (494 amino acid length), with numbers at nodes indicating percentage bootstrap values. Group numbers are shown in ovals at the right. Only experimentally characterised SDIMOs are shown in the tree, with the exceptions of the predicted enzymes from *Zhongshania aliphativorans* and *Aestuaricella hydrocarbonica* in group 7. The operon structures are shown as block arrow diagrams, with subunits as follows: α = alpha hydroxylase, β = beta hydroxylase, γ = gamma hydroxylase, C = coupling protein, Fd = ferredoxin, MD = MMOD or MMOD-like, PhK = PhK-like, R = reductase. Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota, light blue = Bacillota. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site.

in the scientific literature. High sequence diversity exists in the TMO group, as indicated by the low % aa identity values used in our analyses here as cutoffs for group membership (35%) and CD-HIT clustering (65%) (Table S1). Further details of TMOs that have been experimentally characterised can be found in Table S2.

We have defined five subgroups within the TMOs (A–E) in Figure 3 in order to facilitate discussion regarding patterns of distribution and function. These subgroups are all supported by high bootstrap values (> 90%). Subgroups A and B contain uncharacterised enzymes encoded in genomes and metagenome-assembled genomes (MAGs) from Pseudomonadota, Acidobacteria, Actinomycetota, Planctomycetota and, most intriguingly, the candidate phylum Binatota, which contains many putative hydrocarbon oxidizers (Chuvochina et al. 2019; Murphy et al. 2021). There is no experimental evidence yet for the substrates of the enzymes in subgroups A and B, but given their placement in the group 1 SDIMOs, it is plausible that their substrates are aromatic hydrocarbons. Subgroup C includes most of the well-studied TMOs, which have toluene as the primary substrate and are found typically in the Pseudomonadota. Subgroup D contains dioxane- and toluene-oxidising enzymes, which are found in diverse phyla, i.e., Bacillota, Actinomycetota, Thermomicrobiota and Cyanobacteriota. Subgroup E contains isoprene MOs, in addition to uncharacterised enzymes

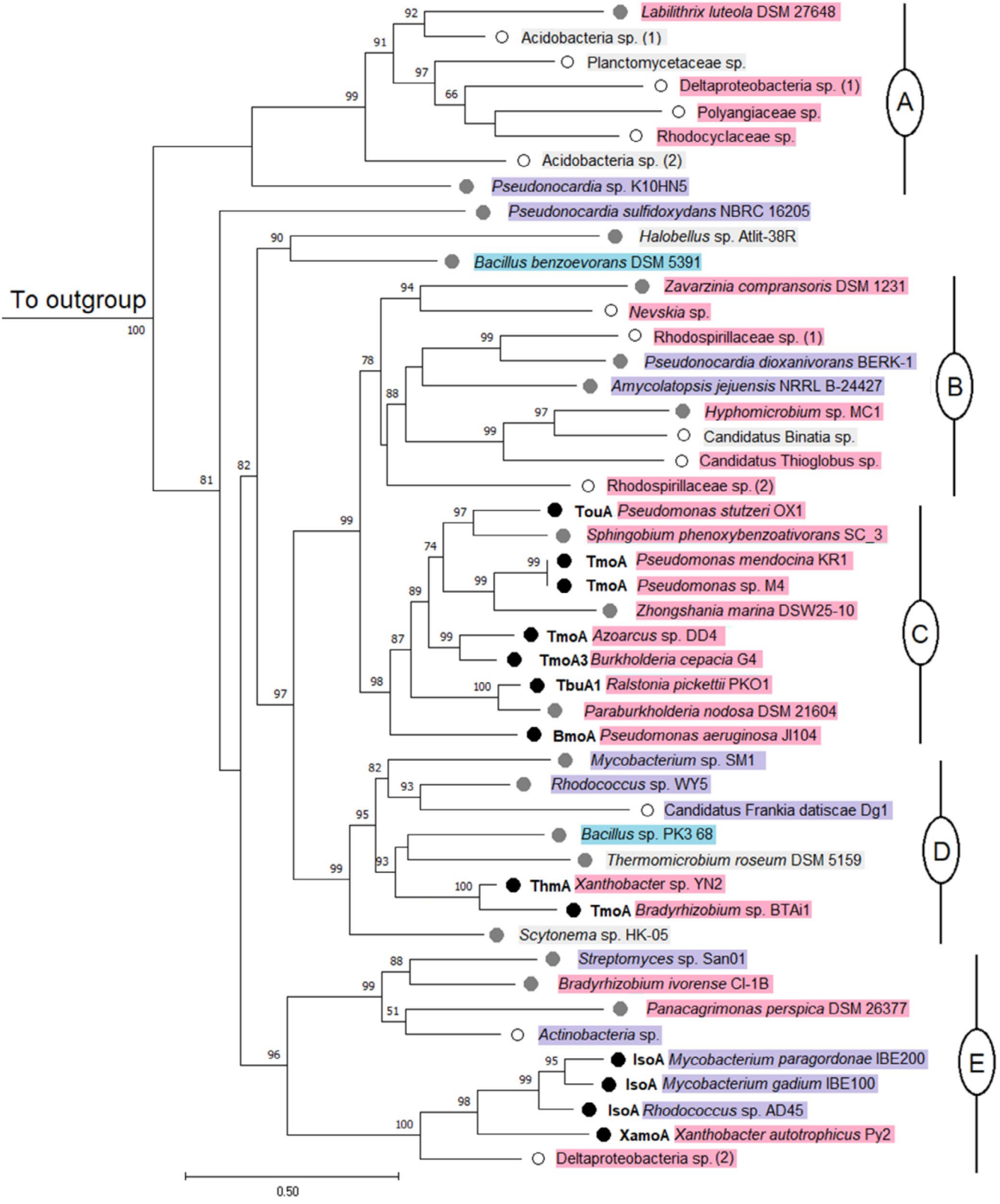

FIGURE 3 | Evolutionary relationships of the group 1 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (533 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from metagenome-assembled genome (MAG). Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota, light blue = Bacillota, grey = other.

from Pseudomonadota and Actinomycetota. From Figure 3, it is clear that the majority of TMO sequences in genomes and metagenomes have unknown functions and significance. This is testament to the high sequence diversity of this group and emphasises the wide knowledge gap between sequences and functions within the SDIMOs more generally.

3.2 | Group 2 SDIMOs—PhMOs

The group 2 SDIMOs (Figures 2 and 4) are commonly referred to as phenol hydroxylases (PHs) or the phenol MOs (PhMOs). The physiological substrates of the PhMOs are diverse, including not only phenol but also cresols, halogenated aromatic compounds, toluene, tetrahydrofuran (THF), benzene and xylene (Table S2). The PhMO operons are composed of six genes, arranged as follows: accessory protein, β-hydroxylase, coupling protein, α-hydroxylase, γ-hydroxylase and reductase. The PhMOs are abundant in sequence databases (>1700 sequences, Table S1) and are very diverse, as indicated by our analyses, which suggested 35% aa identity for group membership and 70% aa identity as a CD-HIT cutoff for clustering analyses (Table S1). Details of experimentally characterised PhMOs can be found in Table S2.

Due to the high diversity of PhMOs and the more gradual transitions seen between the sequences relative to other groups (e.g., compare to the group 3 SDIMO tree, Figure 5), meaningful subgroups within the PhMO phylogenetic tree were more difficult to establish. The experimentally characterised PhMOs are all represented within subgroups A and B (these subgroups are supported by bootstrap values >90%) and are harboured by Pseudomonadota, except for MhpLMNOQBP from *S. acidophilus* TPY (Bacillota) and PhmKLMNOP from *Arthrobacter* sp. W1 (Actinomycetota). The PhMO in *S. acidophilus* TPY is unique, since it is the only characterised PhMO outside of subgroups A and B and the only example from a thermoacidophile. Its operon structure is notably different to the other PhMOs (Zhou et al. 2016). Most of the PhMOs have a narrow substrate range compared to their TMO relatives (Table S2), although some members do perform cometabolic oxidations of various pollutants (see section below and, e.g., *Pseudomonas* CF600 (Shingler et al. 1989, Norlund, Powlowski, and Shingler 1990)). It is intriguing that the deeper branches of the group 2 SDIMO tree are in the Bacillota and Actinomycetota, while the shallower branches are in the Pseudomonadota, suggesting an origin in a Gram-positive lineage and then later transfer into Gram-negatives (see also section below on HGT).

3.3 | Group 3 SDIMOs—sMMOs and Relatives

For many years, the group 3 SDIMOs (Figures 2 and 5) were represented solely by the sMMO enzymes in methanotrophic bacteria, and this is still the most common group 3 SDIMO found in genomes and metagenomes (two-thirds of the sequences in the NCBI database using our search method). The butane MO (BMO) of *Thauera butanivorans* (BmoXYBZDC) provided the first example of a different group 3 SDIMO, in terms of substrate range and host type (Sluis, Sayavedra-Soto, and Arp 2002), later followed by even more divergent examples in the gaseous alkane MOs (Smo) of *Mycobacterium* strain NBB4 (Martin, Ozsvar, and Coleman 2014) and *Rhodococcus* strain ZPP (Zou et al. 2021). The sMMOs are arguably the best-characterised SDIMOs and have well-defined biochemistry, physiology and structures (Sakai, Yurimoto, and Shima 2023). However, a major barrier that has hindered more detailed molecular studies of the sMMOs has been the lack of a good heterologous expression system (see section below). This problem also limits their applications in biocatalysis. The physiological substrate range of the sMMOs is limited (methane and short-chain alkanes) but their cometabolic oxidation range is very wide (Table S2).

The sMMO and BMO operons are composed of six genes in this order: α-hydroxylase, β-hydroxylase, coupling protein, γ-hydroxylase, accessory protein and reductase. The Smo operons have five genes in this order: α-hydroxylase, β-hydroxylase, coupling protein, γ-hydroxylase and reductase. The significance of the different operon structures between the different subgroups is not yet known, but it is important to note that the five-component Smo enzyme of *Mycobacterium* NBB4 was shown to be functional in a heterologous host, despite lacking a homologue of the accessory protein subunit seen in sMMO and BMO (Martin, Ozsvar, and Coleman 2014). The sequence diversity of the group 3 SDIMOs is low compared to groups 1 and 2, as demonstrated by the relatively high cutoff for group membership (50% aa identity, Table S1) and the high clustering value used for CD-HIT analysis (95% aa identity, Table S1). The details of experimentally characterised group 3 SDIMOs are found in Table S2.

The group 3 SDIMOs have been divided into three subgroups (A–C) here to facilitate discussion (Figure 5). All of these subgroups are supported by high bootstrap values (>98%). Subgroup A contains the Smo enzymes and is dominated by Actinomycetota (primarily *Mycobacterium* spp), with one notable exception of a predicted enzyme encoded in the genome of an *Oleomonas* strain (Alphaproteobacteria). Subgroup B enzymes are found in Pseudomonadota and include BMO, which is thus far the only characterised member of this subgroup. Subgroup C enzymes are also found only in the Pseudomonadota. Most strains harbouring subgroup C enzymes are canonical methanotrophs that harbour sMMOs and utilise methane as their physiological substrate. The availability of many new sequences in subgroups A and B adds depth to the group 3 SDIMO tree and allows us to make stronger predictions about the origins of methanotrophy. Specifically, this tree strongly supports the previous proposal that the sMMO in methane-oxidising specialists arose from an earlier C_2–C_4 alkane-oxidising Smo-like enzyme in a generalist host (Osborne and Haritos 2019). The group 3 SDIMO tree shows some similar large-scale patterns to the group 1 SDIMO tree, in that the deeper branches are in the Actinomycetota (with one exception) and the shallower branches are in the Pseudomonadota.

3.4 | Group 4 SDIMOs—Alkene MOs

The group 4 SDIMOs (Figures 2 and 6) are commonly referred to as alkene MOs. Unlike many other SDIMO groups, their physiological substrate range is narrow and focused on the utilisation of short-chain alkenes, i.e., ethene (Coleman et al. 2011), propene (Saeki and Furuhashi 1994; Woodland, Matthews, and Leak 1995; Saeki

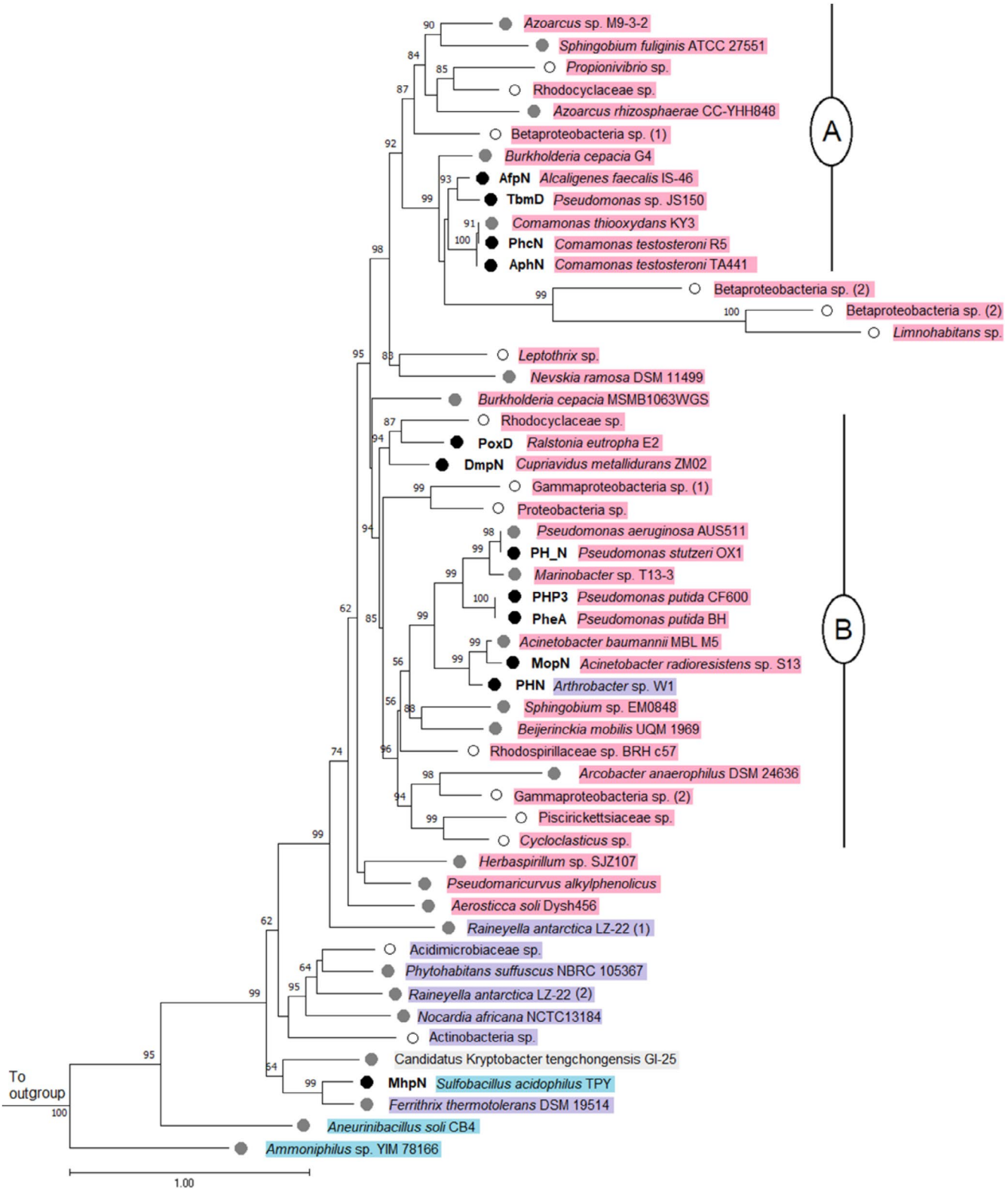

FIGURE 4 | Evolutionary relationships of the group 2 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (470 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota, light blue = Bacillota, grey = other.

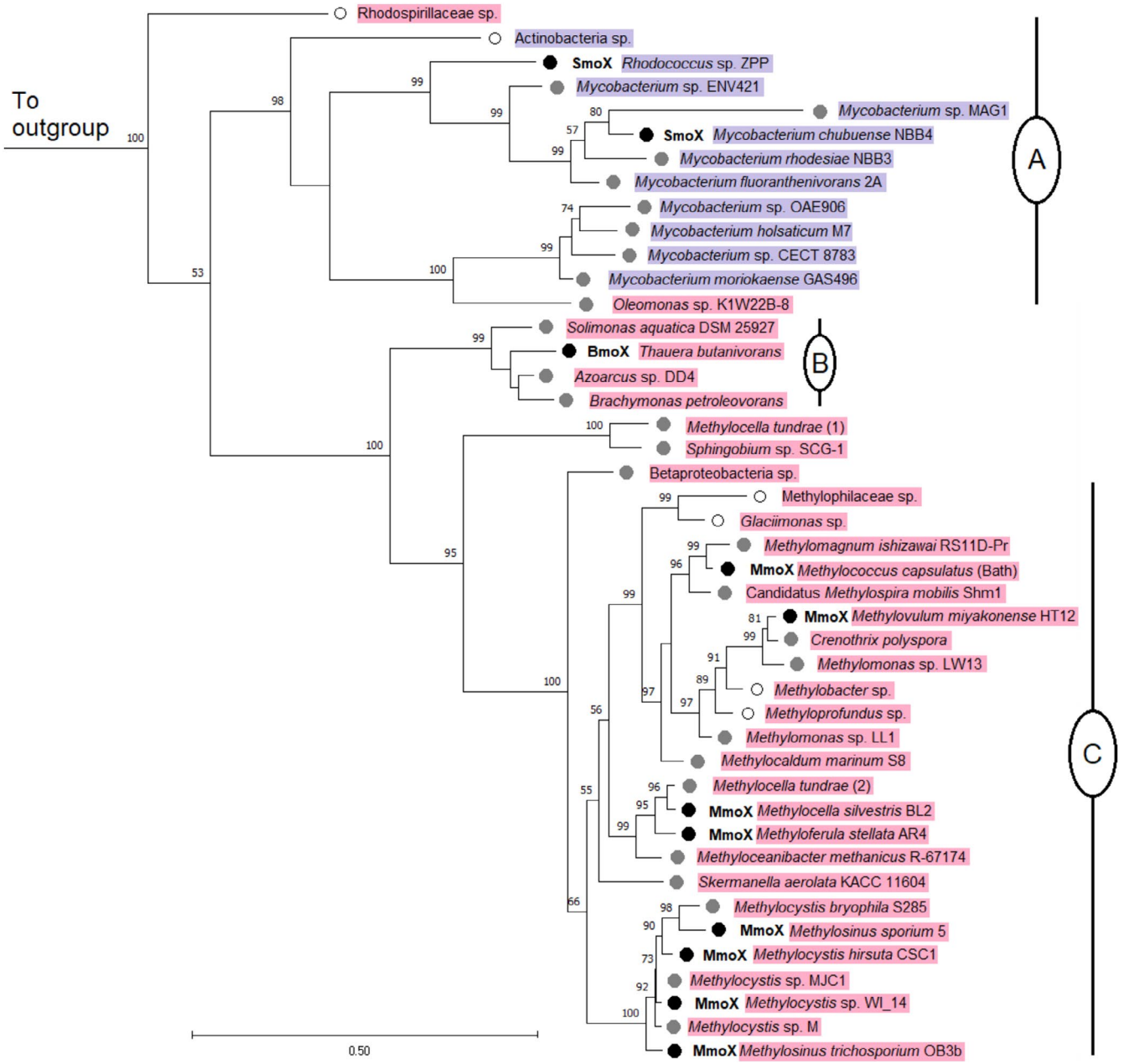

FIGURE 5 | Evolutionary relationships of the group 3 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (495 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota.

et al. 1999), 1-butene (Suzuki et al. 2019) and vinyl chloride (VC) (Mattes et al. 2005) (Table S2). The alkene MO operons are composed of four genes, which are arranged as follows: β-hydroxylase, coupling protein, α-hydroxylase and reductase. Compared to other SDIMO groups, alkene MOs are not well-represented in sequence databases (only 22 total sequences, Table S1) and are not as well-characterised in the scientific literature. The group 4 SDIMOs are exclusively found in the Actinomycetota, with one exception to date (*Halieaceae* strain PE-TB08W, a propene-oxidising member of the Gammaproteobacteria). Alkene MOs have moderate sequence diversity, as shown by the 50% aa identity value used in our analysis here as the cutoff for group membership. Due to the small number of enzymes in group 4, clustering via CD-HIT analysis was not done, and thus, Figure 6 represents all the predicted alkene MOs in sequence databases at the time of writing. The details of the alkene MOs that have been experimentally characterised are found in Table S2.

We have defined here three distinct subgroups (A–C) within alkene MOs, which are supported by bootstrap values > 97% (Figure 6). These correspond quite well to the physiological substrates of the host cells, where this is known. Subgroup A contains

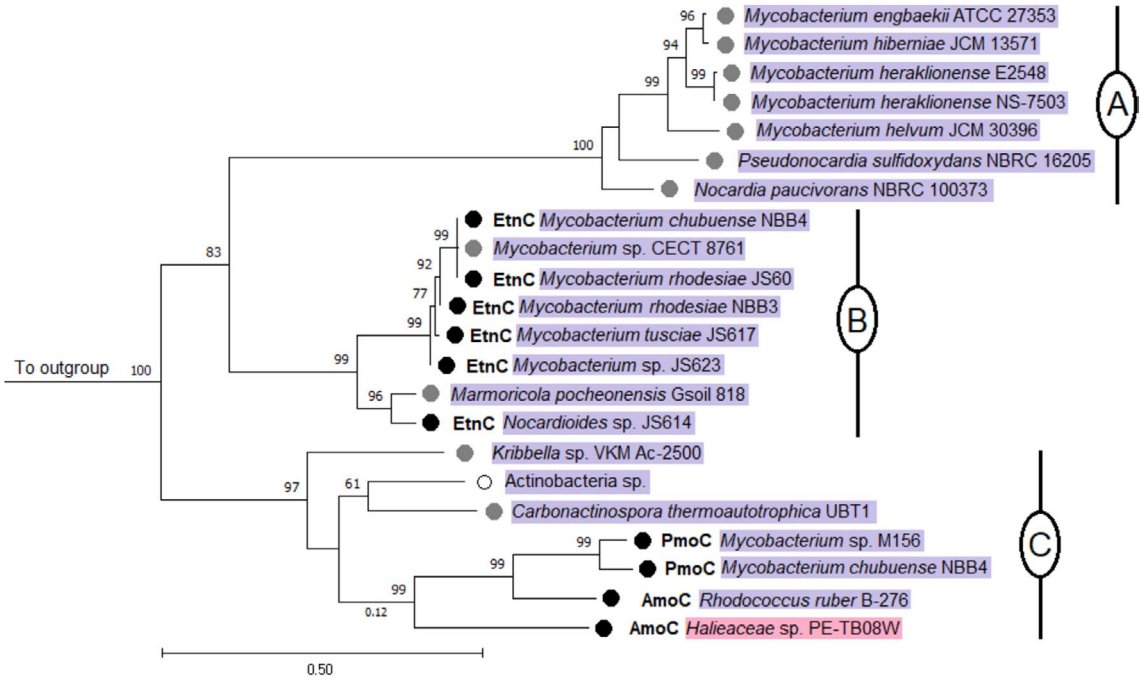

FIGURE 6 | Evolutionary relationships of the group 4 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (497 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: Purple = Actinomycetota, pink = Pseudomonadota.

no experimentally characterised MOs and is very divergent from subgroups B and C, suggesting that these enzymes may have a different substrate range and physiological role to the known alkene MOs. Subgroup B contains ethene MOs (EtnABCD). These have several characterised members in the genera *Mycobacterium* and *Nocardioides* (Coleman and Spain 2003b; Mattes et al. 2005; Chuang and Mattes 2007; McCarl et al. 2018). Subgroup C contains propene MOs (AmoABCD, PmoABCD), with characterised members in *Mycobacterium*, *Rhodococcus* and *Halieaceae*. Alkene MOs have been characterised primarily by heterologous expression and reverse-transcription PCR, but these enzymes are less well-studied than the SDIMOs in groups 1, 2 and 3. For example, no knockout mutants have been made, only one of the proteins has been purified (AmoABCD from *Rhodococcus ruber* B-276 (Miuran and Dalton 1995, Gallagher, Cammack, and Dalton 1997)), and no crystal structures are available. More work is needed on both the partly characterised group 4 SDIMOs and the unstudied members of this family in order to better understand their ecological significance and potential applications. In the latter case, this is important due to the uniquely high stereoselectivity of oxidations catalysed by alkene MOs (Cheung et al. 2013).

3.5 | Group 5 SDIMOs: Propane-2-MOs and Relatives

The group 5 SDIMOs (Figures 2 and 7) are commonly referred to as propane MOs, or more correctly, propane-2-MOs (Pr2MOs), but their physiological substrate range covers diverse chemical families including THF, 1,4-dioxane, acetone, methylethylketone and *N*-nitrosodimethylamine (Table S2). The group 5 SDIMO operons are composed of 4 genes arranged as follows: α-hydroxylase, reductase, β-hydroxylase and coupling protein. These are represented by a very large number of sequences found in databases (> 1000, Table S1) and some members are moderately well-characterised. The group 5 SDIMOs have high intra-group sequence diversity, as shown by the 40% aa identity value used as a cutoff for group membership and the relatively low clustering cutoff (80%) for CD-HIT analysis (Table S1). The details of experimentally characterised group 5 SDIMOs can be found in Table S2.

We have divided the group 5 SDIMOs into three subgroups (A–C) for the purposes of discussion. Subgroup A contains predominantly sequences from Actinomycetota MAGs, along with three characterised THF-oxidising enzymes (ThmADBC) from *Pseudonocardia* and *Arthrobacter*. Subgroup B is also comprised exclusively of enzymes from Actinomycetota and includes all the characterised Pr2MOs (typically annotated PrmABCD) and one THF-oxidising enzyme (ThmADBC). Subgroup C is found primarily in Pseudomonadota and also includes one Bacillota host and one Verrucomicrobia host. This subgroup does not contain any experimentally characterised enzymes. Like all the previously discussed SDIMO groups, there is a large knowledge gap between sequences and functions within the group 5 SDIMOs. In the case of the subgroup B enzymes, this is especially intriguing, since they are found in the genomes of many Actinomycetota, regardless of their isolation substrate. We hypothesise that acetone or related ketones could be the usual

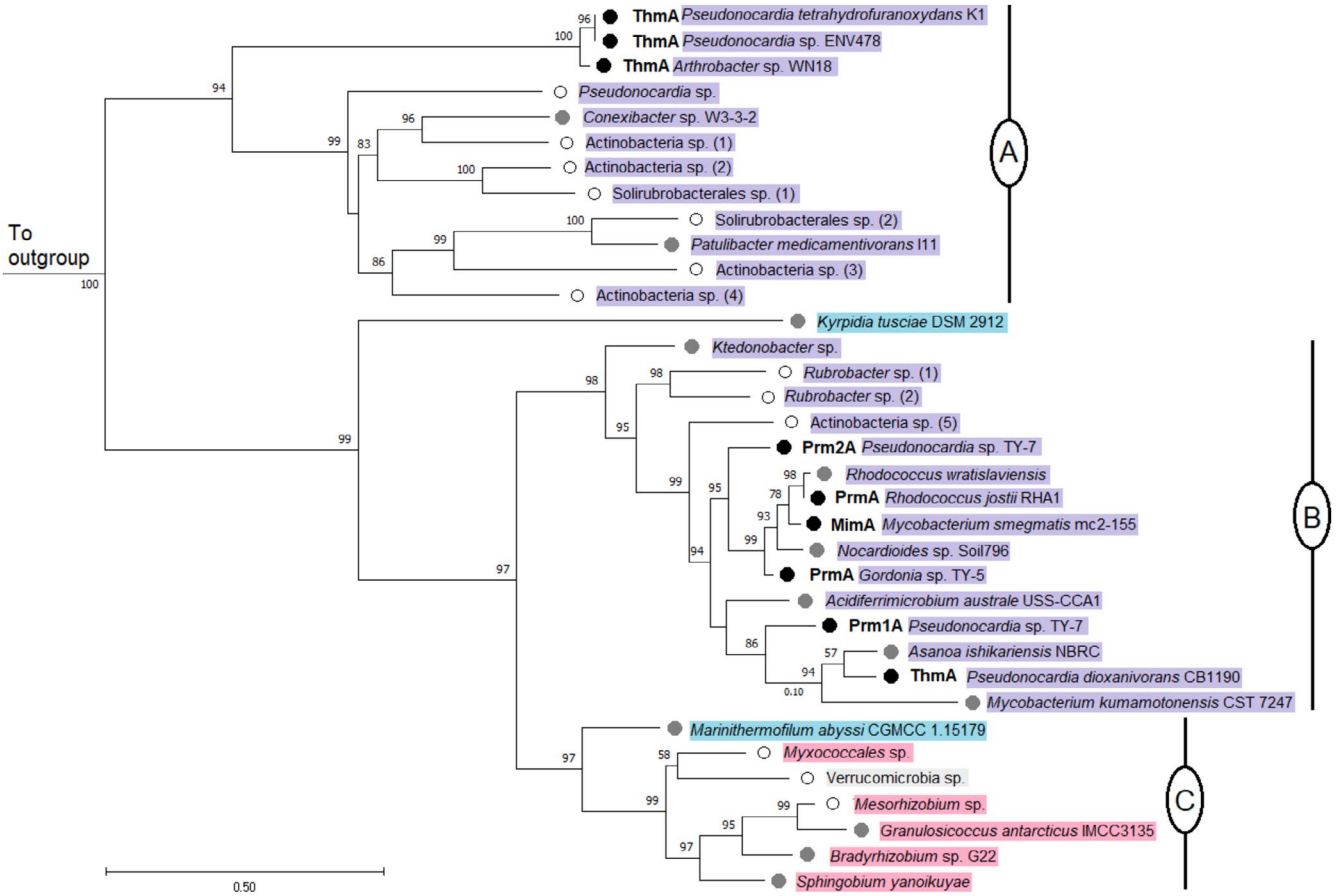

FIGURE 7 | Evolutionary relationships of the group 5 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (439 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota, light blue = Bacillota, grey = other.

physiological substrates of these enzymes, since ketones are common metabolites in many bacteria, unlike hydrocarbons.

3.6 | Group 6 SDIMOs: Propane-1-MOs and Relatives

The group 6 SDIMOs (Figures 2 and 8) are commonly referred to as propane MOs, or more correctly, propane-1-MOs (Pr1MOs). While their primary substrate is propane (Kotani et al. 2006; Masuda et al. 2012), the physiological substrates of the group 6 SDIMOs also include *n*-butane (Kotani et al. 2006), 1,4-dioxane and THF (He et al. 2017) (Table S2). A key functional difference between the group 5 and group 6 SDIMOs is that the former oxidise the subterminal carbon of alkanes, whereas the latter oxidise the terminal carbon (Kotani et al. 2006). The group 6 SDIMO operons are composed of four genes arranged as follows: α-hydroxylase, β-hydroxylase, coupling protein and reductase. These genes are not common in sequence databases (44 total sequences, Table S1) and they have low sequence diversity as suggested by the high CD-HIT clustering cutoff (90%) used here for tree construction (Table S1).

The properties and significance of the group 6 SDIMOs are not well-understood, since only a small number have been experimentally characterised, and even in those cases, only limited evidence for their biochemistry and physiology has been obtained. The details of the Pr1MOs that have been characterised can be found in Table S2. Unlike the group 5 SDIMOs, which tend to be chromosomal, implying a place in the core metabolism of the hosts, the group 6 enzymes in *Mycobacterium* sp. NBB4 (Coleman et al. 2011), *Mycobacterium* sp. ELW1 (Kottegoda, Waligora, and Hyman 2015), *Mycobacterium gadium* IBE100 (Helbich et al. 2023) and *Mycobacterium paragordonae* IBE200 (Helbich et al. 2023) are harboured on large plasmids. This is consistent with a role for the group 6 SDIMOs in oxidising less-frequently encountered substrates such as gaseous alkanes.

We have defined three subgroups (A–C) of group 6 SDIMOs (Figure 8) to facilitate discussion. Subgroup A contains predominantly sequences from Pseudomonadota MAGs and does not contain any characterised enzymes. Subgroup B contains a mixture of Actinomycetota (mostly Mycobacteria), an uncharacterised Deltaproteobacteria sp. and, most intriguingly, sequences from Candidatus Binataceae and Chloroflexi. Both

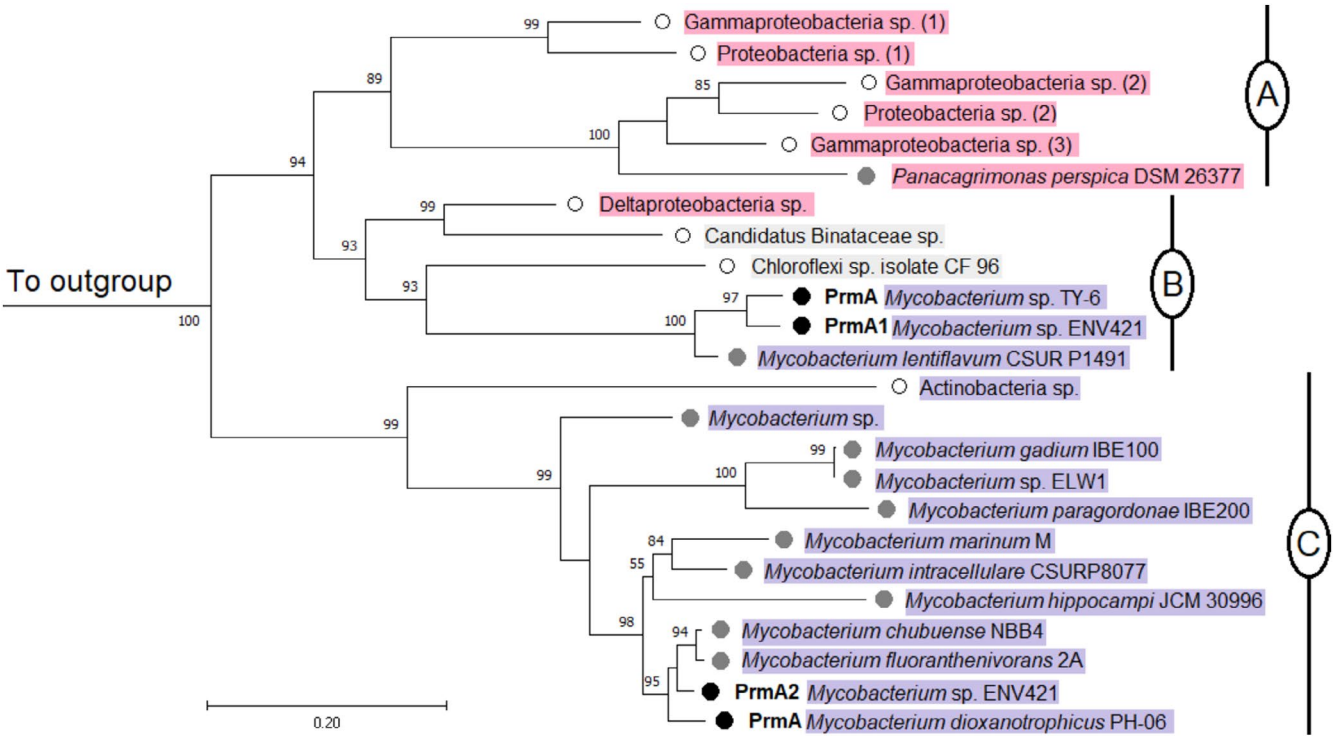

FIGURE 8 | Evolutionary relationships of the group 6 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (459 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. ZmoC of *Solimonas soli* was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: purple = Actinomycetota, pink = Pseudomonadota, grey = other.

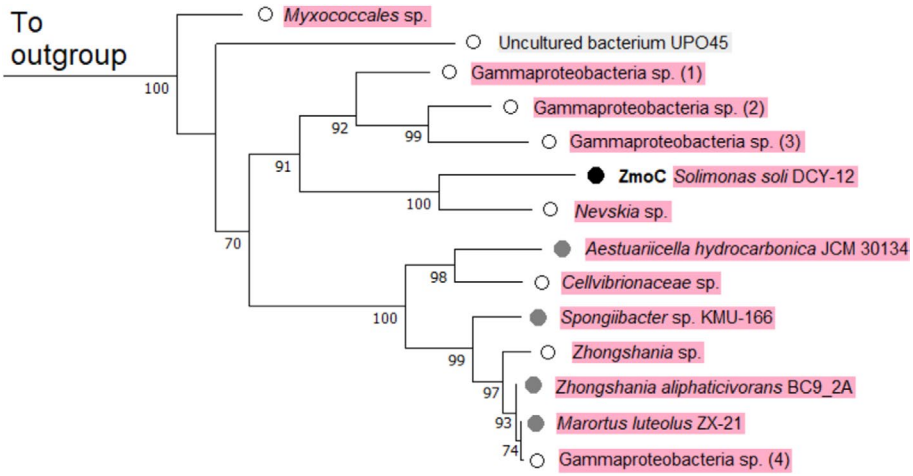

FIGURE 9 | Evolutionary relationships of the group 7 SDIMOs. The maximum likelihood tree was made from a trimmed alignment of alpha-subunit sequences (455 aa length), with numbers at nodes indicating percentage bootstrap values. Bootstrap values less than 50% are not shown. Scale bar units are amino acid substitutions per site. PoxD of *Ralstonia* sp. E2 was used as the outgroup to root the tree. Dots preceding organism names indicate the following: black circle = characterised SDIMO, grey circle = uncharacterised SDIMO from pure culture genome, white circle = uncharacterised SDIMO from MAG. Coloured shading behind organism names indicates their phylum as follows: pink = Pseudomonadota, grey = other.

these phyla are known to harbour putative hydrocarbon degraders (Chuvochina et al. 2019; Dong et al. 2020). Subgroup C is mostly made up of *Mycobacterium* sequences, except for one sequence from an Actinomycetota MAG. Again, it is clear that most of the sequence diversity of the group 6 SDIMOs represents unexplored functions, especially in subgroup A.

3.7 | The Group 7 SDIMOs: A New Subgroup of Unclear Function and Significance

The group 7 SDIMOs (Figures 2 and 9) are the most recently identified subgroup (Yang et al. 2024), and we know very little about their biochemistry, physiology and ecological significance.

The first member of this subgroup was identified in 2019 in the genome of *Solimonas soli* (Osborne and Haritos 2019). The *S. soli* SDIMO (ZmoABCD) can be expressed in *Pseudomonas putida* KT2440 and was able to oxidise C_2–C_4 alkenes, 1-hexene, 1-octene, isoprene, acrylic acid, VC and *cis*-1,2-DCE (Yang et al. 2024) (Table S2). However, that study could not identify the physiological substrate of the enzyme. At the time of writing, only 14 group 7 SDIMO sequences can be identified in sequence databases. All of these have the same operon structure as the group 4 SDIMOs (β-hydroxylase, coupling protein, α-hydroxylase and reductase). The group 7 SDIMOs form a coherent clade, which is distant from the existing SDIMO groups (Figure 2), justifying the proposal for a new SDIMO group.

The group 7 SDIMOs are found almost exclusively in Gammaproteobacteria (Figure 9), with one exception in the Myxococcota, and one in a taxonomically indeterminate MAG. As mentioned above, the primary substrate of this SDIMO group remains unknown (Table S2), as do most of the relevant properties of the enzymes, such as the nature of the inducer(s), their protein structures and their associated metabolic pathways. One interesting trend in this subgroup is that the majority of the bacterial hosts are marine bacteria, which suggests that the physiological substrate(s) of this SDIMO type may be specific to this habitat.

4 | SDIMOS in Biodegradation & Bioremediation

Biodegradation is the process by which organic compounds are broken down by microbes, while bioremediation refers to the technology of harnessing these biodegradation processes to effectively cleanup organic pollutants found in contaminated sites (Wackett and Hershberger 2001). SDIMO-containing bacteria are popular candidates for pollutant cleanup due to their wide substrate range. The use of SDIMOs for bioremediation has previously been reviewed, especially with respect to halogenated alkanes and alkenes (Shennan 2005), halogenated aromatic hydrocarbons (Pimviriyakul et al. 2020) and cyclic ethers (Skinner et al. 2024). We will therefore provide only a brief overview in this section, focussing on some recently discovered SDIMOs that have not been reviewed elsewhere. The reader is referred to Table S2 for specific details of pollutants metabolised by different SDIMOs.

4.1 | Impacts of Cometabolism and Physiology on Bioremediation

Prior to implementation of an SDIMO for bioremediation, it is helpful to first understand the place of the enzyme in its host's physiology, since a purely biochemical understanding of the enzyme's activity may be insufficient for successful deployment of the bacteria for bioremediation (Skinner et al. 2024). It is especially important to understand the process limitations imposed by cometabolism (metabolism of a compound in a manner that provides no growth benefit (Boethling and Alexander 1979, Schmidt, Simkins, and Alexander 1985)). In cometabolic bioremediation processes, the growth substrate competes with the pollutant for the enzyme active site, and this impacts the enzyme kinetics (Alvarez-Cohen and Speitel Jr. 2001). While high rates of turnover are possible using cometabolic reactions, the lack of carbon and energy yield from the pollutant and the need for an auxiliary substrate causes numerous problems for process implementation and sustainability. The growth substrate may need sequential (rather than continuous) addition (Frascari et al. 2012), a different inducer may also need to be added (Suttinun, Muller, and Luepromchai 2009) and the cometabolic substrate may itself be toxic or may give rise to toxic metabolites (van Hylckama Vlieg and Janssen 2001; Halsey et al. 2005). For more detail, the reader is referred to an excellent recent review (Skinner et al. 2024), which covers issues of cometabolism and SDIMOs in depth.

4.2 | SDIMO-Mediated Biodegradation of Halogenated Hydrocarbons

Halogenated alkanes, alkenes and aromatics are common groundwater pollutants, and thus, there has been much interest in finding and characterising bacteria that can biodegrade these chemicals (Shennan 2005; Pimviriyakul et al. 2020). Nearly all of the SDIMO groups contain members that are capable of cometabolic oxidation of haloalkenes (Table S2), with some well-studied examples including the PhMO of *Burkholderia cepacia* G4 (Shim and Wood 2000) and the TMO of *Pseudomonas stutzeri* OX1 (Ryoo et al. 2001) (Table S2). The alkene MOs (EtnABCD, group 4) found in mycobacteria are noteworthy, since these enable growth on a chlorinated ethene (VC), not just cometabolism. One such strain, *Mycobacterium* JS623, provided the first evidence for the specific mutations associated with adaptation to growth on chlorinated hydrocarbons (VC) from a cometabolic ancestor (Jin et al. 2010). The mycobacteria have proved a rich source of unusual SDIMOs. Strain NBB4, for example, has 4 SDIMOs (Coleman et al. 2011), including SmoXYC1B1Z, an unusual group 3 SDIMO, which can attack VC and 1,2-dichloroethane (Martin, Ozsvar, and Coleman 2014). This enzyme is notable for being the only group 3 SDIMO to date that is amenable to heterologous expression in a generalist bacterial host (see section below).

Two interesting recently discovered SDIMOs capable of biodegradation of chlorinated ethenes are the TmoABCDEF (group 1) and PrmABCD (group 5) in *Azoarcus* sp. DD4 (Deng et al. 2020; Li et al. 2021). These SDIMOs are unusual because they are found in a single bacterium and both have activity on propane. This is unique for a group 1 SDIMO. Strain DD4 has been used for bioremediation of trichloroethylene (TCE) in groundwater, by employing an anaerobic consortium to convert TCE to *cis*-dichloroethylene (cDCE) and VC and then using *Azoarcus* to cometabolise cDCE and VC during growth on propane (Li et al. 2021). Another recently discovered SDIMO with potential applications for bioremediation is ZmoABCD from *Solimonas soli* DCY12. This SDIMO can attack VC and cDCE (Table S2) and is interesting as the first example of an SDIMO with a useful application but unknown primary substrate. The only SDIMO that is known to be involved in growth on a halogenated aromatic is Tbm (group 2 SDIMO) from *Pseudomonas* sp. JS150, which utilises *p*-chlorobenzene (Spain and Nishino 1987). However, other group 2 SDIMOs can readily oxidise halogenated phenols

by cometabolism (e.g., phenol-grown cells of *Pseudomonas putida* CF600, which will attack 4-chlorophenol (Nowak and Mrozik 2016)).

4.3 | SDIMO-Mediated Biodegradation of Cyclic Ethers

Cyclic ethers are another major category of groundwater pollutants (Miao et al. 2023). Two of the most well-studied are THF and 1,4-dioxane, which have been used as isolation substrates and/or cometabolic substrates for many bacteria that contain group 5 SDIMOs (Table S2). A recent interesting ether-degrading bacterium is *Xanthobacter* sp. YN2, which in contrast contains a group 1 SDIMO (ThmABCDEF) and can utilise 1,4-dioxane, 1,3-dioxane, 1,4-dioxene and THF as growth substrates (Ma et al. 2021) (Table S2). This SDIMO is remarkable, since it is the first example of an enzyme from group 1 linked to ether metabolism, and also because it enables growth on a broad range of cyclic ethers. The nomenclature of this enzyme is a difficult issue. The sequence of the *Xanthobacter* sp. YN2 enzyme shows it to be more closely related to Tmo enzymes than to other Thm enzymes, but its substrate range looks more like the latter. This problem occurs in numerous places in the SDIMO family tree (Figure 2). The reader is advised to be wary of assumptions about genetic relatedness or enzymatic functions based on SDIMO names and to critically assess the prior literature when naming new SDIMOs. We recommend that new enzymes should be named based on sequence identity, not substrate range, since the latter is extremely fluid with SDIMOs.

Another recently reported cyclic ether degrader is *Cupriavidus metallidurans* ZM02 (Ren et al. 2022). This isolate is unusual, since it oxidises THF using a group 2 SDIMO, DmpKLMNOP (Table S2). In this case, the enzyme has been (we believe) correctly named based on its sequence identity to dimethylphenol (Dmp) MOs, and it is notable that strain ZM02 does also grow on phenol. The activity of DmpKLMNOP towards THF was confirmed by knockout experiments (these knockout strains lost the ability to grow on both THF and phenol) and also by heterologous expression in *Cupriavidus* strain JMP134. It was remarkable that the recombinants in the latter experiments could not only degrade THF but could also grow on this compound. This is a powerful reminder that acquisition of SDIMO genes alone can sometimes modify the growth substrate range of the host and impact its fitness in different environments—these considerations are relevant for understanding the selective forces impacting HGT of SDIMO genes (below).

Arthrobacter strain WN18 uses a group 5 SDIMO (ThmADBC) to grow on THF and co-metabolically degrade 1,4-dioxane (Wang et al. 2021a; Wang et al. 2021b) (Table S2). Strain WN18 is typical of most previously described dioxane degraders, which are Actinomycetota-containing group 5 SDIMOs that use THF or propane as primary substrates. It is worth noting that to date, relatively few organisms have been shown to grow on dioxane (Inoue et al. 2016; He et al. 2017). THF-utilising bacteria provide a good example of the challenges associated with cometabolic biodegradation processes, since the issues relate not only to the supply of carbon and energy but also to the need for THF as an inducer of the genes in the wild-type bacteria. It is notable that the environmental tolerances of the WN18 isolate (pH, temperature, salinity and resistance to other pollutants) were characterised (Wang et al. 2021b). This was vital, as these properties of a potential inoculant are important in determining the success of bioremediation.

Mycobacterium strain PH-06 is another interesting recent isolate (He et al. 2017; Deng, Li, and Li 2018). This is distinguished by its possession of a group 6 SDIMO PrmABCD, which attacks both THF and 1,4-dioxane and enables growth on both ethers as carbon and energy sources (Table S2). While Prm is upregulated by both THF and 1,4-dioxane, strain PH-06 preferentially metabolises THF. This is in accordance with the preference for THF vs. dioxane seen in other ether degraders (Table S2) and is consistent with the fact that THF is a naturally occurring molecule (Parod 2014), while 1,4-dioxane is xenobiotic. Rigorous evidence for the involvement of the Prm enzyme in ether metabolism was obtained through heterologous expression of the SDIMO in *Mycobacterium smegmatis* mc^2-155. Such recombinants may be useful for bioremediation, since they allow uncoupling of the pollutant-degrading enzymes from the metabolism of the host organism and enable the use of inducers that do not compete with pollutants for the enzyme active site. These approaches however do bring their own technical, legal and ethical issues (de Lorenzo 2009; Kelle 2013; Dvořák et al. 2017).

5 | Heterologous Expression: A Key Tool for Understanding and Applying SDIMO Activities

The expression of SDIMO genes in an alternative host bacterium (heterologous expression) is a very useful method for testing the function of the encoded enzymes and holds promise for applying them in bioremediation and biocatalysis. Ideally, fast-growing and easy-to-manipulate hosts like *Escherichia coli* are used for heterologous expression experiments (Rosano and Ceccarelli 2014), but in the case of SDIMOs, this is often not possible, for reasons that are not clear. There is evidence that the phylogenetic distance between the native and heterologous hosts is one predictor of success (a more distant relationship typically causes problems), and it also seems that SDIMO expression causes physiological stresses to the heterologous hosts, either due to the proteins themselves misfolding or due to reactive oxygen species generated in the active site. Below, we highlight some relevant strategies that have enabled successful heterologous expression of SDIMOs.

The first factor that affects heterologous expression of SDIMOs is the phylogenetic relationship between the native and heterologous hosts, with greater success observed with closely related hosts. There are many reports of successful expression of group 1 and 2 SDIMOs using standard cloning vectors in *E. coli* (Shields et al. 1995; Bertoni et al. 1996; Horinouchi et al. 1999; Tao et al. 2004). These PHs and TMOs are harboured by hosts such as *Pseudomonas*, *Burkholderia* and *Ralstonia*, which are Gammaproteobacteria like *E. coli*. These group 1 and 2 SDIMOs can also be successfully expressed in *Pseudomonas* hosts (Kukor and Olsen 1990; Norlund, Powlowski, and Shingler 1990; Bertoni et al. 1996) (see Table S2 for more details). Similar patterns are seen in other SDIMO groups, for example, the group 5 SDIMOs from *Pseudonocardia tetrahydrofuranoxydans* K1

and *Pseudonocardia dioxanivorans* CB1190 can be expressed in *Rhodococcus jostii* RHA1 (Sales et al. 2013). In these cases, both native and expression hosts are Actinomycetota. The host phylogeny effects are at least partly due to codon usage differences (Gustafsson, Govindarajan, and Minshull 2004), and codon optimisation of the target genes can therefore help to improve expression. This was employed to good effect with the group 5 SDIMO MimABCD of *Mycobacterium smegmatis*, enabling expression in *E. coli* (Furuya, Hayashi, and Kino 2013) (Table S2), although co-expression of chaperone proteins (see below) was also needed. One recent advance that has yet to be applied to the case of SDIMOs is the use of codon harmonisation (Angov et al. 2008) rather than optimisation. In this approach, codons with similar rarity in both native and expression hosts are used rather than just the most rapidly translated codons. It will be interesting to see if this approach helps SDIMO expression.

Another factor that may impede the success of SDIMO expression is the physiological stress that oxygenase enzymes place on the host cell. This reflects at least two distinct underlying problems; first, the challenge of correctly assembling a large multi-subunit enzyme, which requires numerous cofactors (Izzo et al. 2011; McCarl et al. 2018), and second, the generation of damaging reactive oxygen species in the active sites of oxygenases (Lee 1999; Sazykin et al. 2019; Bopp et al. 2022). One approach to combat these problems is to express the SDIMO in a host that already has their own native MO, e.g., sMMO of *Methylosinus* was successfully expressed in a pMMO-expressing host (Lloyd et al. 1999), and the SMO from *Mycobacterium chubuense* NBB4 was able to be expressed in *M. smegmatis* mc^2-155 (Martin, Ozsvar, and Coleman 2014) (Table S2). Using a MO-containing host is not ideal, however, due to potential difficulties in assigning activities to the relevant MO. In addition, the heterologous expression of MOs in MO-containing hosts while impressive may not be viable as an avenue for deployment of these enzymes, as the hosts are likely less amenable to in situ deployment or industrial processes. Another approach is to co-express chaperone proteins that aid protein folding. This approach is informed by the presence of chaperone genes next to many MO gene clusters, e.g., *mimG* near *mimABCD* in *M. smegmatis* mc^2-155 (Furuya et al. 2012) and *bmoG* near *bmoXYBZDC* in *Thauera butanivorans* (Kurth et al. 2008). Co-expression of chaperones has enabled heterologous expression of Actinomycetota SDIMOs in Pseudomonadota hosts (e.g., EtnABCD from *M. chubuense* NBB4 in *P. putida* KT2440 (McCarl et al. 2018), MimABCD of *M. smegmatis* mc^2-155 in *E. coli* (Furuya, Hayashi, and Kino 2013)) (Table S2) and the notoriously difficult to heterologously express sMMO in *E. coli* (sMMO from *Methylomonas methanica* MC09) (Zill et al. 2022).

The successes described above, and other examples found in Table S2, have enabled important advances in our understanding of SDIMO function and laid the foundations for more detailed molecular work. For example, the ease of expression of the PHs and TMOs in *E. coli* has enabled site-directed mutagenesis and directed evolution studies to create enzymes with altered kinetics and substrate ranges (Canada et al. 2002; Vardar and Wood 2004; Vardar and Wood 2005). The co-expression of chaperone proteins and the use of codon optimisation strategies have enabled the SDIMOs from Actinomycetota to be expressed in more industrially useful and genetically tractable hosts like *E. coli* and expanded the potential applications of these enzymes. Other general methods to reduce the stress of MO expression on heterologous hosts include the use low copy number vectors (Yen et al. 1991) and inducible promoters (Yang et al. 2024). Despite all these advances, there still appear to be unknown factors that are preventing heterologous expression of some SDIMOs (especially sMMO) in 'workhorse' bacteria, and even SDIMOs that are amenable to heterologous expression in *E. coli* give recombinants with much lower activities than the original host cells (Martin, Ozsvar, and Coleman 2014; McCarl et al. 2018). Strategies such as transposon mutagenesis or TraDIS (Cain et al. 2020) could be used to pinpoint these unknown factors in the native hosts, as a prequel to their co-expression alongside the SDIMOs in heterologous hosts.

6 | HGT

Many studies have documented the importance of mobile genetic elements (MGEs) such as plasmids and transposons in the evolution of pollutant-degrading bacteria (Tan 1999; Top, Springael, and Boon 2002; Top and Springael 2003), but less is known about the specific HGT events that have involved SDIMO genes. Understanding these events may help us piece together the missing links in the story of how SDIMO genes have moved between different taxa and evolved different functions suited to different environmental niches. An early seminal review of SDIMO evolution (Leahy, Batchelor, and Morcomb 2003) drew several key conclusions. First, the α and β subunits are paralogs (i.e., duplicated from a common ancestor), second, the reductase subunits share ancestry with oxidoreductases from diverse oxygenase families, third, the common ancestor of the SDIMOs most likely had a very broad substrate range (both alkenes and aromatics) and finally, HGT played an important role in SDIMO evolution, both within gene clusters (acquisition and shuffling of subunits) and between taxa (movement of whole operons between diverse hosts).

We believe that the conclusions from the older literature regarding HGT of SDIMOs are still valid, and it is clear from the examination of the figures in this review that the SDIMO genes have moved over large taxonomic distances (i.e., Pseudomonadota to Actinomycetota) numerous times. This poses questions about the mechanisms of HGT, since the plasmid types that carry SDIMOs are traditionally considered to be specific to either Gram-negative bacteria (e.g., IncP-2 for PhMOs (Bartilson, Nordlund, and Shingler 1990)) or Gram-positive bacteria (e.g., large linear plasmids for alkene MOs (Saeki et al. 1999)). Some resolutions to this conundrum may be found in recent studies, which suggest that, first, novel MGEs with very broad host ranges exist in yet-to-be-cultured environmental bacteria (Smalla, Jechalke, and Top 2015) and, second, well-studied plasmids may have broader host ranges than previously believed, when these are tested using more innovative and/or systematic approaches (Klümper et al. 2015; Pesce et al. 2019).

The group 1 SDIMOs in particular appear to have been subject to extensive HGT, since the phylogeny of their α subunit sequences is incongruent with the phylogeny of their bacterial hosts (Figure 3). A good example of this can be seen in the isoprene-oxidising enzymes from *Xanthobacter* Py2 (XamoABCDEF)

and *Rhodococcus* AD45 (IsoABCDEF), which are closely related (70% aa identity in α subunit) despite the hosts belonging to different phyla (Pseudomonadota vs. Actinomycetota). In both these cases, the SDIMOs are plasmid-encoded (Zhou, Chan Kwo Chion, and Leak 1996; Crombie et al. 2015), but little is known of the biology of these plasmids. There is indirect evidence that large linear catabolic plasmids may be able to move between different phyla (Fetzner, Kolkenbrock, and Parschat 2007), but specific experiments to test this are lacking. The reasons for the apparent higher mobility of the group 1 SDIMOs relative to the other groups are not known, although it could be speculated that this is related to their relative ease of heterologous expression. There is also some evidence for the inter-phylum transfer of group 2 SDIMOs in our analyses (Figure 4), e.g., the PhMOs from *Acinetobacter* strain S13 and *Arthrobacter* strain W1 are close neighbours (85% aa identity), but most HGT with the group 2 SDIMOs seems to be occurring over smaller taxonomic distances (e.g., between different families and classes within the Pseudomonadota).

While the sMMO genes (group 3 SDIMOs) are considered to be stable chromosomal features of methanotrophs, these enzymes are likely to have had a complex evolutionary history involving numerous HGT events (Osborne and Haritos 2018). It is especially notable that the more 'divergent' group 3 SDIMOs (i.e., non-sMMO-like) discovered in recent years tend to be encoded on plasmids (Martin, Ozsvar, and Coleman 2014; Zou et al. 2021). There is evidence for some major taxonomic divides within the deeper branches of group 3 SDIMO tree (Figure 4), where it appears that these genes have moved between the Actinomycetota and Pseudomonadata phyla on more than one occasion. It seems very likely from inspection of this tree that the sMMOs evolved from Smo-like ancestors that oxidised larger substrates. Notably, this still allows for the hypothesis that the methane-oxidising enzymes were moved via HGT into methanol-utilising hosts to give rise to the first aerobic methanotrophs (Kang, Dunfield, and Semrau 2019).

A case for HGT in the group 4 SDIMOs can also be made, but in a different way to the groups discussed above. All but one of the known group 4 SDIMOs are found in Actinomycetota (Figure 6), with only one case of apparent long-range HGT, i.e., AmoABCD in *Halieaceae* sp. PE-TB08W (Gammaproteobacteria), which has 68% aa identity in the alpha subunit to the AmoC from *Rhodococcus* B-276. However, three of the four experimentally characterised alkene MOs are known to be plasmid encoded (Saeki et al. 1999; Chan Kwo Chion, Askew, and Leak 2005; Chuang and Mattes 2007; Coleman et al. 2011), and there is indirect evidence based on pulsed-field gel electrophoresis experiments that many of the others are too (Coleman and Spain 2003a). The genome sequences of a set of ethene- and VC-oxidising mycobacteria further indicate that the alkene assimilation genes are found on large plasmids or plasmid-like contigs in nearly all cases (strains *Nocardioides* sp. JS614 (NCBI Reference Sequence: NC_008697.1), *Mycobacterium rhodesiae* JS60 (NCBI Accession Number: AY243034), *Mycobacterium rhodesiae* NBB3 (Reference Sequence: NC_016604.1) and *Mycobacterium chubuense* NBB4 (NCBI Reference Sequence: NC_018022.1)), although knockout or curing experiments are needed to confirm this link. The picture that emerges of the group 4 SDIMOs is that they are highly mobile within the Actinomycetota thanks to the activities of large catabolic plasmids and are also more rarely involved in HGT events over further taxonomic distances. Much more work is needed to understand the biology of the large catabolic plasmids found in the Actinomycetota, which are poorly studied compared to corresponding catabolic plasmids in Gram-negative bacteria.

As mentioned above, group 5 SDIMOs are common components of the chromosomes of many Nocardioform species, and the corresponding SDIMO tree is dominated by the Actinomycetota (Figure 7). At first glance therefore, these are not strong candidates for HGT. However, looking at the deeper branches in this tree reveals some taxonomic anomalies suggestive of HGT, including the presence of two Bacillota spp. (*Kyrpidia tusciae* and *Marinithermofilum abyssi*), alongside several Pseudomonadota species in subgroup C. It would be especially interesting to further probe the differences between the well-studied subgroup B enzymes from Actinomycetota and the related group C enzymes from Pseudomonadota. None of the latter have been experimentally characterised yet, but their phylogeny suggests they may have a different substrate and/or ecological niche relative to the better-studied Prm and Thm enzymes. At least some of the group 5 SDIMOs are plasmid-borne (Thiemer, Andreesen, and Schrader 2003; Sales et al. 2013), providing a mechanism for HGT in at least a few cases.

There is only very limited literature available on the group 6 SDIMOs. However, it is likely that at least some HGT has occurred in this lineage between closely related genera. For example, the uncharacterized group 6 SDIMOs in *Mycobacterium gadium* IBE100, *Mycobacterium paragordonae* IBE200 and *Mycobacterium* sp. ELW1 appear to have been recently shared among these species by mobile elements such as plasmid pELW1-1 (Helbich et al. 2023). The group 6 SDIMO of *Mycobacterium chubuense* NBB4 is plasmid-encoded (NCBI Reference Sequence: NC_018022.1), as is the group 6 SDIMO from *Mycobacterium dioxanotrophicus* PH-06 (Deng, Li, and Li 2018). The latter is known to be located between two insertion sequences, which is a significant finding, since it provides a mechanism for the SDIMO genes to move between different replicons.

7 | Conclusions

This review presents a comprehensive and updated phylogenetic analysis of the SDIMOs and supports their division into seven groups based on the sequence identity of the catalytic alpha subunit. We have examined in detail the relationship between enzymes, bacterial hosts and substrates and attempted to find robust links between these where possible. It is also important to acknowledge that these relationships are in a state of constant flux due to forces such as HGT. We have given an overview of the issue of HGT as it applies to SDIMOs and shown via phylogenetic analyses that HGT is common for SDIMO genes, although the patterns of HGT differ among different SDIMO groups. We have discussed heterologous expression as a key tool for understanding the functions and significance of SDIMOs. We believe that this is still the best approach for investigating the substrate range of these enzymes, since it separates them from the genome of the original host (which may contain multiple

MOs) and allows access to enzymes from genomes and metagenomes in the absence of information about growth substrates or inducers. However, we must also acknowledge the importance of native expression systems and their role in our understanding of the basic assembly of SDIMO subunits and cofactors that are required by the SDIMO in the context that is physiologically relevant. Heterologous expression and native expression studies should therefore be utilised in tandem to maximise our understanding of SDIMOs where possible. We have highlighted some recent examples of SDIMOs used for pollutant biodegradation and provided a comprehensive list of these biodegradative activities in the Table S2. In the cases of both chlorinated ethenes and cyclic ethers, SDIMO-containing bacteria are still the lead candidates for bioremediation efforts under aerobic conditions.

Several large knowledge gaps in the SDIMO literature have become apparent as a result of compiling this review, leading us to propose several priority research areas. Perhaps most urgent is the need to experimentally characterise representatives of the deep-branching and thus-far-uncharacterised SDIMO lineages found in genomes and metagenomes (e.g., subgroups 1A, 4A, 5C, 6A). These enzymes may have significant environmental impacts and/or useful biotechnology applications. Advances in high-throughput synthetic biology methods may facilitate this work (Casas, Bultelle, and Kitney 2024), by allowing the rapid construction and screening of large numbers of recombinant clones. These possibilities depend to some extent on closing the second major knowledge gap, which is to understand why heterologous expression of some SDIMOs is so difficult, and to pinpoint the factors in the wild-type hosts that enable high-level expression of functional SDIMO enzymes. The last knowledge gap worth mentioning is the ongoing lack of good genetic tools and methods for making knockouts and other modifications in environmental bacteria. We are hopeful that recent advances in CRISPR (Burbano et al. 2024) and similar technologies combined with the increased availability and affordability of DNA synthesis will yield breakthroughs on this front.

Author Contributions

Sui Nin Nicholas Yang: investigation, writing – original draft, methodology, formal analysis, data curation. **Michael A. Kertesz:** writing – review and editing, supervision, project administration, formal analysis. **Nicholas V. Coleman:** writing – review and editing, conceptualization, supervision, formal analysis, project administration.

Acknowledgements

This research was supported by internal funding from the University of Sydney. Open access publishing facilitated by Macquarie University, as part of the Wiley - Macquarie University agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in GenBank at https://www.ncbi.nlm.nih.gov/genbank/. Nucleotide accession numbers have been provided for all sequences and are listed in the Tables S1–S3.

References

Alvarez-Cohen, L., and G. E. Speitel Jr. 2001. "Kinetics of Aerobic Cometabolism of Chlorinated Solvents." *Biodegradation* 12: 105–126.

Angov, E., C. J. Hillier, R. L. Kincaid, and J. A. Lyon. 2008. "Heterologous Protein Expression Is Enhanced by Harmonizing the Codon Usage Frequencies of the Target Gene With Those of the Expression Host." *PLoS One* 3: e2189.

Banerjee, R., J. C. Jones, and J. D. Lipscomb. 2019. "Soluble Methane Monooxygenase." *Annual Review of Biochemistry* 88: 409–431.

Bartilson, M., I. Nordlund, and V. Shingler. 1990. "Location and Organization of the Dimethylphenol Catabolic Genes of Pseudomonas CF600." *Molecular and General Genetics* 220: 294–300.

Bernhardt, R. 2006. "Cytochromes P450 as Versatile Biocatalysts." *Journal of Biotechnology* 124: 128–145.

Bertoni, G., F. Bolognese, E. Galli, and P. Barbieri. 1996. "Cloning of the Genes for and Characterization of the Early Stages of Toluene and o-Xylene Catabolism in *Pseudomonas stutzeri* OX1." *Applied and Environmental Microbiology* 62: 3704–3711.

Boethling, R. S., and M. Alexander. 1979. "Microbial Degradation of Organic Compounds at Trace Levels." *Environmental Science & Technology* 13: 989–991.

Bopp, C. E., N. M. Bernet, H. E. Kohler, and T. B. Hofstetter. 2022. "Elucidating the Role of O(2) Uncoupling in the Oxidative Biodegradation of Organic Contaminants by Rieske Non-heme iron Dioxygenases." *ACS Sustainable Chemistry & Engineering* 2: 428–440.

Bryan, M. C., P. J. Dunn, D. Entwistle, et al. 2018. "Key Green Chemistry Research Areas From a Pharmaceutical manufacturers' Perspective Revisited." *Green Chemistry* 20: 5082–5103.

Burbano, D. A., C. Kiattisewee, A. V. Karanjia, et al. 2024. "CRISPR Tools for Engineering Prokaryotic Systems: Recent Advances and New Applications." *Annual Review of Chemical and Biomolecular Engineering* 15: 389–430.

Cain, A. K., L. Barquist, A. L. Goodman, I. T. Paulsen, J. Parkhill, and T. van Opijnen. 2020. "A Decade of Advances in Transposon-Insertion Sequencing." *Nature Reviews Genetics* 21: 526–540.

Canada, K. A., S. Iwashita, H. Shim, and T. K. Wood. 2002. "Directed Evolution of Toluene Ortho-Monooxygenase for Enhanced 1-Naphthol Synthesis and Chlorinated Ethene Degradation." *Journal of Bacteriology* 184: 344–349.

Cardy, D. L. N., V. Laidler, G. P. C. Salmond, and J. C. Murrell. 1991. "Molecular Analysis of the Methane Monooxygenase (MMO) Gene Cluster of Methyiosinus Trichosporium OB3b." *Molecular Biotechnology* 5: 335–342.

Casas, A., M. Bultelle, and R. Kitney. 2024. "An Engineering Biology Approach to Automated Workflow and Biodesign." *Synthetic Biology* 9: ysae009.

Chan Kwo Chion, C. K., S. E. Askew, and D. J. Leak. 2005. "Cloning, Expression, and Site-Directed Mutagenesis of the Propene Monooxygenase Genes From Mycobacterium Sp. Strain M156." *Applied and Environmental Microbiology* 71: 1909–1914.

Cheung, S., V. McCarl, A. J. Holmes, N. V. Coleman, and P. J. Rutledge. 2013. "Substrate Range and Enantioselectivity of Epoxidation Reactions Mediated by the Ethene-Oxidising Mycobacterium Strain NBB4." *Applied and Environmental Microbiology* 97: 1131–1140.

Chuang, A. S., and T. E. Mattes. 2007. "Identification of Polypeptides Expressed in Response to Vinyl Chloride, Ethene, and Epoxyethane in Nocardioides Sp. Strain JS614 by Using Peptide Mass Fingerprinting." *Applied and Environmental Microbiology* 73: 4368–4372.

Chuvochina, M., C. Rinke, D. H. Parks, et al. 2019. "The Importance of Designating Type Material for Uncultured Taxa." *Systematic and Applied Microbiology* 42: 15–21.

Colby, J., D. I. Stirling, and H. Dalton. 1977. "The Soluble Methane Mono-Oxygenase of *Methylococcus capsulatus* (Bath). Its Ability to Oxygenate n-Alkanes, n-Alkenes, Ethers, and Alicyclic, Aromatic and Heterocyclic Compounds." *Biochemical Journal* 165: 395–402.

Coleman, N. V., N. B. Bui, and A. J. Holmes. 2006. "Soluble di-iron Monooxygenase Gene Diversity in Soils, Sediments and Ethene Enrichments." *Environmental Microbiology* 8: 1228–1239.

Coleman, N. V., N. B. Le, M. A. Ly, et al. 2012. "Hydrocarbon Monooxygenase in Mycobacterium: Recombinant Expression of a Member of the ammonia Monooxygenase Superfamily." *ISME Journal* 6: 171–182.

Coleman, N. V., and J. C. Spain. 2003a. "Distribution of the Coenzyme M Pathway of Epoxide Metabolism Among Ethene- and Vinyl Chloride-Degrading Mycobacterium Strains." *Applied and Environmental Microbiology* 69: 6041–6046.

Coleman, N. V., and J. C. Spain. 2003b. "Epoxyalkane: Coenzyme M Transferase in the Ethene and Vinyl Chloride Biodegradation Pathways of Mycobacterium Strain JS60." *Journal of Bacteriology* 185: 5536–5545.

Coleman, N. V., S. Yau, N. L. Wilson, et al. 2011. "Untangling the Multiple Monooxygenases of *Mycobacterium chubuense* Strain NBB4, a Versatile Hydrocarbon Degrader." *Environmental Microbiology Reports* 3: 297–307.

Constable, D. J. C., P. J. Dunn, J. D. Hayler, et al. 2007. "Key Green Chemistry Research Areas—A Perspective From Pharmaceutical Manufacturers." *Green Chemistry* 9: 411–420.

Crombie, A. T., M. E. Khawand, V. A. Rhodius, et al. 2015. "Regulation of Plasmid-Encoded Isoprene Metabolism in Rhodococcus, a Representative of an Important Link in the Global Isoprene Cycle." *Environmental Microbiology* 17: 3314–3329.

de Lorenzo, V. 2009. "Recombinant bacteria for Environmental Release: What Went Wrong and What We Have Learnt From It." *Clinical Microbiology and Infection* 15: 63–65.

Deng, D., F. Li, and M. Li. 2018. "A Novel Propane Monooxygenase Initiating Degradation of 1,4-Dioxane by Mycobacterium Dioxanotrophicus PH-06." *Environmental Science & Technology Letters* 5: 86–91.

Deng, D., D. N. Pham, F. Li, and M. Li. 2020. "Discovery of an Inducible Toluene Monooxygenase That Cooxidizes 1,4-Dioxane and 1,1-Dichloroethylene in Propanotrophic Azoarcus Sp. Strain DD4." *Applied and Environmental Microbiology* 86: e01163-20.

Diamond, S., A. Lavy, A. Crits-Christoph, et al. 2022. "Soils and Sediments Host Thermoplasmata archaea Encoding Novel Copper Membrane Monooxygenases (CuMMOs)." *ISME Journal* 16: 1348–1362.

Donadio, G., C. Sarcinelli, E. Pizzo, et al. 2015. "The Toluene o-Xylene Monooxygenase Enzymatic Activity for the Biosynthesis of Aromatic Antioxidants." *PLoS One* 10: e0124427.

Dong, X., J. E. Rattray, D. C. Campbell, et al. 2020. "Thermogenic Hydrocarbon Biodegradation by Diverse Depth-Stratified Microbial Populations at a Scotian Basin Cold Seep." *Nature Communications* 11: 5825.

Dvořák, P., P. I. Nikel, J. Damborsky, and V. de Lorenzo. 2017. "Bioremediation 3.0: Engineering Pollutant-Removing bacteria in the Times of Systemic Biology." *Biotechnology Advances* 35: 845–866.

Fetzner, S. 2002. "Oxygenases Without Requirement for Cofactors or Metal Ions." *Applied Microbiology and Biotechnology* 60: 243–257.

Fetzner, S., S. Kolkenbrock, and K. Parschat. 2007. "Catabolic Linear Plasmids." In *Microbial Linear Plasmids*, vol. 7, 63–98. Berlin, Heidelberg: Springer.

Frascari, D., M. Cappelletti, S. Fedi, et al. 2012. "Application of the Growth Substrate Pulsed Feeding Technique to a Process of Chloroform Aerobic Cometabolism in a Continuous-Flow Sand-Filled Reactor." *Process Biochemistry* 47: 1656–1664.

Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li. 2012. "CD-HIT: Accelerated for Clustering the Next-Generation Sequencing Data." *Bioinformatics* 28: 3150–3152.

Furuya, T., M. Hayashi, and K. Kino. 2013. "Reconstitution of Active Mycobacterial Binuclear iron Monooxygenase Complex in *Escherichia coli*." *Applied and Environmental Microbiology* 79: 6033–6039.

Furuya, T., M. Hayashi, H. Semba, and K. Kino. 2012. "The Mycobacterial Binuclear iron Monooxygenases Require a Specific Chaperonin-Like Protein for Functional Expression in a Heterologous Host." *FEBS Journal* 280: 817–826.

Gallagher, S. C., R. Cammack, and H. Dalton. 1997. "Alkene Monooxygenase From Nocardia corallina B-276 Is a Member of the Class of Dinuclear iron Proteins Capable of Stereospecific Epoxygenation Reactions." *European Journal of Biochemistry* 247: 635–641.

Gibson, D. T., and R. E. Parales. 2000. "Aromatic Hydrocarbon Dioxygenases in Environmental Biotechnology." *Current Opinion in Biotechnology* 11: 236–243.

Greening, C., and R. Grinter. 2022. "Microbial Oxidation of Atmospheric Trace Gases." *Nature Reviews Microbiology* 20: 513–528.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. "New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0." *Systematic Biology* 59: 307–321.

Guo, X., J. Zhang, L. Han, et al. 2023. "Structure and Mechanism of the Alkane-Oxidizing Enzyme AlkB." *Nature Communications* 14: 2180.

Gupta, R. S., B. Lo, and J. Son. 2018. "Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium Into an Emended Genus Mycobacterium and Four Novel Genera." *Frontiers in Microbiology* 9: 67.

Gustafsson, C., S. Govindarajan, and J. Minshull. 2004. "Codon bias and Heterologous Protein Expression." *Trends in Biotechnology* 22: 346–353.

Halsey, K. H., L. A. Sayavedra-Soto, P. J. Bottomley, and D. J. Arp. 2005. "Trichloroethylene Degradation by Butane-Oxidizing bacteria Causes a Spectrum of Toxic Effects." *Applied Microbiology and Biotechnology* 68: 794–801.

He, Y., J. Mathieu, Y. Yang, P. Yu, M. L. B. da Silva, and P. J. J. Alvarez. 2017. "1,4-Dioxane Biodegradation by Mycobacterium Dioxanotrophicus pH-06 Is Associated With a Group-6 Soluble di-iron Monooxygenase." *Environmental Science & Technology Letters* 4: 494–499.

Helbich, S., I. Barrantes, L. G. Dos Anjos Borges, et al. 2023. "The 2-Methylpropene Degradation Pathway in Mycobacteriaceae Family Strains." *Environmental Microbiology* 25: 2163–2181.

Holmes, A. J., and N. V. Coleman. 2008. "Evolutionary Ecology and Multidisciplinary Approaches to Prospecting for Monooxygenases as Biocatalysts." *Antonie Van Leeuwenhoek* 94: 75–84.

Horinouchi, M., T. Yoshida, H. Nojiri, H. Yamane, and T. Omori. 1999. "Oxidation of Dimethyl Sulfide by Various Aromatic Compound Oxygenases From bacteria." *Biotechnology Letters* 21: 929–933.

Horvath, R. S. 1972. "Microbial Co-Metabolism and the Degradation of Organic Compounds in Nature." *Bacteriological Reviews* 36: 146–155.

Inoue, D., T. Tsunoda, K. Sawada, et al. 2016. "1,4-Dioxane Degradation Potential of Members of the Genera Pseudonocardia and Rhodococcus." *Biodegradation* 27: 277–286.

Izzo, V., G. Leo, R. Scognamiglio, L. Troncone, L. Birolo, and A. Di Donato. 2011. "PHK From Phenol Hydroxylase of Pseudomonas Sp. OX1. Insight Into the Role of an Accessory Protein in Bacterial Multicomponent Monooxygenases." *Archives of Biochemistry and Biophysics* 505: 48–59.

Ji, Y., G. Mao, Y. Wang, and M. Bartlam. 2013. "Structural Insights Into Diversity and n-Alkane Biodegradation Mechanisms of Alkane Hydroxylases." *Frontiers in Microbiology* 4: 58.

Jin, Y. O., S. Cheung, N. V. Coleman, and T. E. Mattes. 2010. "Association of Missense Mutations in Epoxyalkane Coenzyme M Transferase With Adaptation of Mycobacterium Sp. Strain JS623 to Growth on Vinyl Chloride." *Applied and Environmental Microbiology* 76: 3413–3419.

Jin, Y. O., and T. E. Mattes. 2008. "Adaptation of Aerobic, Ethene-Assimilating Mycobacterium Strains to Vinyl Chloride as a Growth Substrate." *Environmental Science & Technology* 42: 4784–4789.

Johnson, M., I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, and T. L. Madden. 2008. "NCBI BLAST: A Better Web Interface." *Nucleic Acids Research* 36: W5–W9.

Kalyuzhnaya, M. G., O. A. Gomez, and J. C. Murrell. 2019. "The Methane-Oxidizing bacteria (Methanotrophs)." In *Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes*, edited by T. J. McGenity, 1–34. Cham: Springer International Publishing.

Kang, C. S., P. F. Dunfield, and J. D. Semrau. 2019. "The Origin of Aerobic Methanotrophy Within the Proteobacteria." *FEMS Microbiology Letters* 366: fnz096.

Keener, W. K., and D. J. Arp. 1994. "Transformations of Aromatic Compounds by *Nitrosomonas europaea*." *Applied and Environmental Microbiology* 60: 1914–1920.

Kelle, A. 2013. "Beyond Patchwork Precaution in the Dual-Use Governance of Synthetic Biology." *Science and Engineering Ethics* 19: 1121–1139.

Khider, M. L. K., T. Brautaset, and M. Irla. 2021. "Methane Monooxygenases: Central Enzymes in Methanotrophy With Promising Biotechnological Applications." *World Journal of Microbiology and Biotechnology* 37: 72.

Klümper, U., L. Riber, A. Dechesne, et al. 2015. "Broad Host Range Plasmids Can Invade an Unexpectedly Diverse Fraction of a Soil Bacterial Community." *ISME Journal* 9: 934–945.

Koo, C. W., and A. C. Rosenzweig. 2021. "Biochemistry of Aerobic Biological Methane Oxidation." *Chemical Society Reviews* 50: 3424–3436.

Koo, C. W., F. J. Tucci, Y. He, and A. C. Rosenzweig. 2022. "Recovery of Particulate Methane Monooxygenase Structure and Activity in a Lipid Bilayer." *Science* 375: 1287–1291.

Kotani, T., Y. Kawashima, H. Yurimoto, N. Kato, and Y. Sakai. 2006. "Gene Structure and Regulation of Alkane Monooxygenases in Propane-Utilizing Mycobacterium Sp. TY-6 and Pseudonocardia Sp. TY-7." *Journal of Bioscience and Bioengineering* 102: 184–192.

Kottegoda, S., E. Waligora, and M. Hyman. 2015. "Metabolism of 2-Methylpropene (Isobutylene) by the Aerobic Bacterium Mycobacterium Sp. Strain ELW1." *Applied and Environmental Microbiology* 81: 1966–1976.

Kukor, J. J., and R. H. Olsen. 1990. "Molecular Cloning, Characterization, and Regulation of a *Pseudomonas pickettii* PKO1 Gene Encoding Phenol Hydroxylase and Expression of the Gene in *Pseudomonas aeruginosa* PAO1c." *Journal of Bacteriology* 172: 4624–4630.

Kurth, E. G., D. M. Doughty, P. J. Bottomley, D. J. Arp, and L. A. Sayavedra-Soto. 2008. "Involvement of BmoR and BmoG in n-Alkane Metabolism in Pseudomonas butanovora." *Microbiology* 154: 139–147.

Larkin, M. A., G. Blackshields, N. P. Brown, et al. 2007. "Clustal W and Clustal X Version 2.0." *Bioinformatics* 23: 2947–2948.

Leahy, J. G., P. J. Batchelor, and S. M. Morcomb. 2003. "Evolution of the Soluble Diiron Monooxygenases." *FEMS Microbiology Reviews* 27: 449–479.

Leak, D. J., R. A. Sheldon, J. M. Woodley, and P. Adlercreutz. 2009. "Biocatalysts for Selective Introduction of Oxygen." *Biocatalysis and Biotransformation* 27: 1–26.

Lee, K. 1999. "Benzene-Induced Uncoupling of Naphthalene Dioxygenase Activity and Enzyme Inactivation by Production of Hydrogen Peroxide." *Journal of Bacteriology* 181: 2719–2725.

Lefort, V., J. E. Longueville, and O. Gascuel. 2017. "SMS: Smart Model Selection in PhyML." *Molecular Biology and Evolution* 34: 2422–2424.

Li, F., D. Deng, L. Zeng, S. Abrams, and M. Li. 2021. "Sequential Anaerobic and Aerobic Bioaugmentation for Commingled Groundwater Contamination of Trichloroethene and 1,4-Dioxane." *Science of the Total Environment* 774: 145118.

Lloyd, J. S., P. De Marco, H. Dalton, and J. C. Murrell. 1999. "Heterologous Expression of Soluble Methane Monooxygenase Genes in Methanotrophs Containing Only Particulate Methane Monooxygenase." *Archives of Microbiology* 171: 364–370.

Ma, F., Y. Wang, J. Yang, H. Guo, D. Su, and L. Yu. 2021. "Degradation of 1,4-Dioxane by Xanthobacter Sp. YN2." *Current Microbiology* 78: 992–1005.

Mahendra, S., and L. Alvarez-Cohen. 2005. "*Pseudonocardia dioxanivorans* Sp. Nov., a Novel Actinomycete That Grows on 1,4-Dioxane." *International Journal of Systematic and Evolutionary Microbiology* 55: 593–598.

Martin, K. E., J. Ozsvar, and N. V. Coleman. 2014. "SmoXYB1C1Z of Mycobacterium Sp. Strain NBB4: A Soluble Methane Monooxygenase (sMMO)-like Enzyme, Active on C2 to C4 Alkanes and Alkenes." *Applied and Environmental Microbiology* 80: 5801–5806.

Masuda, H., K. McClay, R. J. Steffan, and G. J. Zylstra. 2012. "Characterization of Three Propane-Inducible Oxygenases in Mycobacterium Sp. Strain ENV421." *Letters in Applied Microbiology* 55: 175–181.

Mattes, T. E., A. K. Alexander, and N. V. Coleman. 2010. "Aerobic Biodegradation of the Chloroethenes: Pathways, Enzymes, Ecology, and Evolution." *FEMS Microbiology Reviews* 34: 445–475.

Mattes, T. E., N. V. Coleman, J. C. Spain, and J. M. Gossett. 2005. "Physiological and Molecular Genetic Analyses of Vinyl Chloride and Ethene Biodegradation in Nocardioides Sp. Strain JS614." *Archives of Microbiology* 183: 95–106.

McCarl, V., M. V. Somerville, M. A. Ly, et al. 2018. "Heterologous Expression of Mycobacterium Alkene Monooxygenases in Gram-Positive and Gram-Negative Bacterial Hosts." *Applied and Environmental Microbiology* 84: e00397-18.

Meehan, C. J., R. A. Barco, Y. E. Loh, S. Cogneau, and L. Rigouts. 2021. "Reconstituting the Genus Mycobacterium." *International Journal of Systematic and Evolutionary Microbiology* 71: 004922.

Miao, Y., T. Zhou, X. Zheng, and S. Mahendra. 2023. "Investigating Biodegradation of 1,4-Dioxane by Groundwater and Soil Microbiomes: Insights Into Microbial Ecology and Process Prediction." *ACS ES&T Water* 4: 1046–1060.

Miuran, A., and H. Dalton. 1995. "Purification and Characterization of the Alkene Monooxygenase From Nocardia corallina B-276." *Bioscience, Biotechnology, and Biochemistry* 59: 853–859.

Murphy, C. L., A. Sheremet, P. F. Dunfield, et al. 2021. "Genomic Analysis of the Yet-Uncultured Binatota Reveals Broad Methylotrophic, Alkane-Degradation, and Pigment Production Capacities." *MBio* 12: e00985-21.

Nicholas, K. B., and H. B. J. Nicholas. 1997. "GeneDoc: A Tool for Editing and Annotating Multiple Sequence Alignments." Distributed by the authors.

Norlund, I., J. Powlowski, and V. Shingler. 1990. "Complete Nucleotide Sequence and Polypeptide Analysis of Multicomponent Phenol Hydroxylase From Pseudomonas Sp. Strain CF600." *Journal of Bacteriology* 172: 6823–6833.

Notomista, E., A. Lahm, A. Di Donato, and A. Tramontano. 2003. "Evolution of Bacterial and Archaeal Multicomponent Monooxygenases." *Journal of Molecular Evolution* 56: 435–445.

Nowak, A., and A. Mrozik. 2016. "Facilitation of Co-Metabolic Transformation and Degradation of Monochlorophenols by

Pseudomonas Sp. CF600 and Changes in Its Fatty Acid Composition." *Water, Air, & Soil Pollution* 227: 83.

Ojo, A. O., J. Castillo, E. D. Cason, and A. Valverde. 2023. "Biodegradation of Chloroethene Compounds Under Microoxic Conditions." *Biotechnology and Bioengineering* 121: 1036–1049.

Oren, A., and G. M. Garrity. 2021. "Valid Publication of the Names of Forty-Two phyla of Prokaryotes." *International Journal of Systematic and Evolutionary Microbiology* 71: 005056.

Osborne, C. D., and V. S. Haritos. 2018. "Horizontal Gene Transfer of Three Co-Inherited Methane Monooxygenase Systems Gave Rise to Methanotrophy in the Proteobacteria." *Molecular Phylogenetics and Evolution* 129: 171–181.

Osborne, C. D., and V. S. Haritos. 2019. "Beneath the Surface: Evolution of Methane Activity in the Bacterial Multicomponent Monooxygenases." *Molecular Phylogenetics and Evolution* 139: 106527.

Parod, R. J. 2014. "Tetrahydrofuran." In *Encyclopedia of Toxicology*, edited by P. Wexler, vol. 4, 505–508. Cambridge, MA, USA: Academic Press.

Pesce, C., R. Oshone, S. G. Hurst IV, V. A. Kleiner, and L. S. Tisaa. 2019. "Stable Transformation of the Actinobacteria Frankia Spp." *Applied and Environmental Microbiology* 85: e00957-19.

Petkevičius, V., J. Vaitekūnas, D. Vaitkus, N. Čėnas, and R. Meškys. 2019. "Tailoring a Soluble Diiron Monooxygenase for Synthesis of Aromatic N-Oxides." *Catalysts* 9: 356.

Pimviriyakul, P., T. Wongnate, R. Tinikul, and P. Chaiyen. 2020. "Microbial Degradation of Halogenated Aromatics: Molecular Mechanisms and Enzymatic Reactions." *Microbial Biotechnology* 13: 67–86.

Que, L., Jr., and W. B. Tolman. 2008. "Biologically Inspired Oxidation Catalysis." *Nature* 455: 333–340.

Ren, H., H. Wang, Y. Wang, Y. Chen, and Z. Lu. 2022. "A Multicomponent THF Hydroxylase Initiates Tetrahydrofuran Degradation in *Cupriavidus metallidurans* ZM02." *Applied and Environmental Microbiology* 88: e01880-21.

Rosano, G. L., and E. A. Ceccarelli. 2014. "Recombinant Protein Expression in *Escherichia coli*: Advances and Challenges." *Frontiers in Microbiology* 5: 172.

Ryoo, D., H. Shim, F. L. Arenghi, P. Barbieri, and T. K. Wood. 2001. "Tetrachloroethylene, Trichloroethylene, and Chlorinated Phenols Induce Toluene-o-Xylene Monooxygenase Activity in *Pseudomonas stutzeri* OX1." *Applied and Environmental Microbiology* 56: 545–549.

Saeki, H., M. Akira, K. Furuhashi, B. Averhoff, and G. Gottschalk. 1999. "Degradation of Trichloroethene by a Linear-Plasmid-Encoded Alkene Monooxygenase in *Rhodococcus corallinus* (Nocardia corallina) B-276." *Microbiology* 145: 1721–1730.

Saeki, H., and K. Furuhashi. 1994. "Cloning and Characterization of a Nocardia corallina B-276 Gene Cluster Encoding Alkene Monooxygenase." *Journal of Fermentation and Bioengineering* 78: 399–406.

Sakai, Y., H. Yurimoto, and S. Shima. 2023. "Methane Monooxygenases; Physiology, Biochemistry and Structure." *Catalysis Science & Technology* 13: 6342–6354.

Sales, C. M., A. Grostern, J. V. Parales, R. E. Parales, and L. Alvarez-Cohen. 2013. "Oxidation of the Cyclic Ethers 1,4-Dioxane and Tetrahydrofuran by a Monooxygenase in Two Pseudonocardia Species." *Applied and Environmental Microbiology* 79: 7702–7708.

Sazykin, I., M. Makarenko, L. Khmelevtsova, E. Seliverstova, A. Rakin, and M. Sazykina. 2019. "Cyclohexane, Naphthalene, and Diesel Fuel Increase Oxidative Stress, CYP153, sodA, and recA Gene Expression in *Rhodococcus erythropolis*." *Microbiology* 8: e00855.

Schmidt, S. K., S. Simkins, and M. Alexander. 1985. "Models for the Kinetics of Biodegradation of Organic Compounds Not Supporting Growth." *Applied and Environmental Microbiology* 50: 323–331.

Semprini, L. 1997. "Strategies for the Aerobic Co-Metabolism of Chlorinated Solvents." *Current Opinion in Biotechnology* 8: 296–308.

Shennan, J. L. 2005. "Utilisation of C2–C4 Gaseous Hydrocarbons and Isoprene by Microorganisms." *Journal of Chemical Technology & Biotechnology* 81: 237–256.

Shields, M. S., M. J. Reagin, R. R. Gerger, R. Campbell, and C. Somerville. 1995. "TOM, a New Aromatic Degradative Plasmid From Burkholderia (Pseudomonas) Cepacia G4." *Applied and Environmental Microbiology* 61: 1352–1356.

Shim, H., and T. K. Wood. 2000. "Aerobic Degradation of Mixtures of Chlorinated Aliphatics by Cloned Toluene-o-Xylene Monooxygenase and Toluene o-Monooxygenase in Resting Cells." *Biotechnology and Bioengineering* 70: 693–698.

Shingler, V., F. C. Franklin, M. Tsuda, D. Holroyd, and M. Bagdasarian. 1989. "Molecular Analysis of a Plasmid-Encoded Phenol Hydroxylase From Pseudomonas CF600." *Journal of General Microbiology* 135: 1083–1092.

Skinner, J., A. G. Delgado, M. Hyman, and M. J. Chu. 2024. "Implementation of In Situ Aerobic Cometabolism for Groundwater Treatment: State of the Knowledge and Important Factors for Field Operation." *Science of the Total Environment* 925: 171667.

Sluis, M. K., L. A. Sayavedra-Soto, and D. J. Arp. 2002. "Molecular Analysis of the Soluble Butane Monooxygenase From 'Pseudomonas Butanovora'." *Microbiology* 148: 3617–3629.

Smalla, K., S. Jechalke, and E. M. Top. 2015. "Plasmid Detection, Characterization, and Ecology." *Microbiology Spectrum* 3: PLAS-0038-2014.

Spain, J. C., and S. F. Nishino. 1987. "Degradation of 1,4-Dichlorobenzene by a Pseudomonas Sp." *Applied and Environmental Microbiology* 53: 1010–1019.

Stainthorpe, A. C., V. Lees, G. P. C. Salmond, H. Dalton, and J. C. Murrell. 1990. "The Methane Monooxygenase Gene Cluster of *Methylococcus capsulatus* (Bath)." *Gene* 91: 27–34.

Sullivan, J. P., D. Dickinson, and H. A. Chase. 1998. "Methanotrophs, *Methylosinus trichosporium* OB3b, sMMO, and Their Application to Bioremediation." *Critical Reviews in Microbiology* 24: 335–373.

Suttinun, O., R. Muller, and E. Luepromchai. 2009. "Trichloroethylene Cometabolic Degradation by Rhodococcus Sp. L4 Induced With Plant Essential Oils." *Biodegradation* 20: 281–291.

Suzuki, T., T. Yazawa, N. Morishita, A. Maruyama, and H. Fuse. 2019. "Genetic and Physiological Characteristics of a Novel Marine Propylene-Assimilating Halieaceae Bacterium Isolated From Seawater and the Diversity of Its Alkene and Epoxide Metabolism Genes." *Microbes and Environments* 34: 33–42.

Tamura, K., G. Stecher, and S. Kumar. 2021. "MEGA11: Molecular Evolutionary Genetics Analysis Version 11." *Molecular Biology and Evolution* 38: 3022–3027.

Tan, H.-M. 1999. "Bacterial Catabolic Transposons." *Applied Microbiology and Biotechnology* 51: 1–12.

Tao, Y., A. Fishman, W. E. Bentley, and T. K. Wood. 2004. "Oxidation of Benzene to Phenol, Catechol, and 1,2,3-Trihydroxybenzene by Toluene 4-Monooxygenase of *Pseudomonas mendocina* KR1 and Toluene 3-Monooxygenase of *Ralstonia pickettii* PKO1." *Applied and Environmental Microbiology* 70: 3814–3820.

Thiemer, B., J. R. Andreesen, and T. Schrader. 2003. "Cloning and Characterization of a Gene Cluster Involved in Tetrahydrofuran Degradation in Pseudonocardia Sp. Strain K1." *Archives of Microbiology* 179: 266–277.

Top, E. M., and D. Springael. 2003. "The Role of Mobile Genetic Elements in Bacterial Adaptation to Xenobiotic Organic Compounds." *Current Opinion in Biotechnology* 14: 262–269.

Top, E. M., D. Springael, and N. Boon. 2002. "Catabolic Mobile Genetic Elements and Their Potential Use in Bioaugmentation of Polluted Soils and Waters." *FEMS Microbiology Ecology* 42: 199–208.

Torres Pazmiño, D. E., M. Winkler, A. Glieder, and M. W. Fraaije. 2010. "Monooxygenases as Biocatalysts: Classification, Mechanistic Aspects and Biotechnological Applications." *Journal of Biotechnology* 146: 9–24.

Tucci, F. J., R. J. Jodts, B. M. Hoffman, and A. C. Rosenzweig. 2023. "Product Analogue Binding Identifies the Copper Active Site of Particulate Methane Monooxygenase." *Nature Catalysis* 6: 1194–1204.

Tucci, F. J., and A. C. Rosenzweig. 2024. "Direct Methane Oxidation by Copper- and iron-Dependent Methane Monooxygenases." *Chemical Reviews* 124: 1288–1320.

van Berkel, W. J., N. M. Kamerbeek, and M. W. Fraaije. 2006. "Flavoprotein Monooxygenases, a Diverse Class of Oxidative Biocatalysts." *Journal of Biotechnology* 124: 670–689.

van Hylckama Vlieg, J. E. T., and D. B. Janssen. 2001. "Formation and Detoxification of Reactive Intermediates in the Metabolism of Chlorinated Ethenes." *Journal of Biotechnology* 85: 81–102.

Vardar, G., and T. K. Wood. 2004. "Protein Engineering of Toluene-o-Xylene Monooxygenase From *Pseudomonas stutzeri* OX1 for Synthesizing 4-Methylresorcinol, Methylhydroquinone, and Pyrogallol." *Applied and Environmental Microbiology* 70: 3253–3262.

Vardar, G., and T. K. Wood. 2005. "Protein Engineering of Toluene-o-Xylene Monooxygenase From *Pseudomonas stutzeri* OX1 for Enhanced Chlorinated Ethene Degradation and o-Xylene Oxidation." *Applied Genetics and Molecular Biotechnology* 68: 510–517.

Wackett, L. P., and C. D. Hershberger. 2001. "General Concepts in Biodegradation and Biocatalysis." In *Biocatalysis and Biodegradation: Microbial Transformation of Organic Compounds*, edited by L. P. Wackett and C. D. Hershberger, vol. 1, 1–5. Washington, DC, USA: ASM Press.

Wang, P., T. Cui, J. Li, N. Liu, and M. Hong. 2021a. "1,4-Dioxane Degradation Performance of Tetrahydrofuran-Grown Arthrobacter Sp. WN18." *Water, Air, & Soil Pollution* 232: 473.

Wang, P., F. Li, W. Wang, et al. 2021b. "Cometabolic Degradation of 1,4-Dioxane by a Tetrahydrofuran-Growing Arthrobacter Sp. WN18." *Ecotoxicology and Environmental Safety* 217: 112206.

Wang, W., R. E. Iacob, R. P. Luoh, J. R. Engen, and S. J. Lippard. 2014. "Electron Transfer Control in Soluble Methane Monooxygenase." *Journal of the American Chemical Society* 136: 9754–9762.

Woodland, M. P., C. S. Matthews, and D. J. Leak. 1995. "Properties of a Soluble Propene Monooxygenase From Mycobacterium Sp. (Strain M156)." *Archives of Microbiology* 163: 231–234.

Yang, S. N. N., V. Haritos, M. A. Kertesz, and N. V. Coleman. 2024. "A Novel Soluble di-iron Monooxygenase From the Soil Bacterium *Solimonas soli*." *Environmental Microbiology* 26: e16567.

Yen, K. M., M. R. Karl, L. M. Blatt, et al. 1991. "Cloning and Characterization of a *Pseudomonas mendocina* KR1 Gene Cluster Encoding Toluene-4-Monooxygenase." *Journal of Bacteriology* 173: 5315–5327.

Zhao, G., T. Xia, J. Song, and R. A. Jensen. 1994. "*Pseudomonas aeruginosa* Possesses Homologues of Mammalian Phenylalanine Hydroxylase and 4 Alpha-Carbinolamine Dehydratase/DCoH as Part of a Three-Component Gene Cluster." *Proceedings of the National Academy of Sciences of the United States of America* 91: 1366–1370.

Zhou, N.-Y., C. K. Chan Kwo Chion, and D. J. Leak. 1996. "Cloning and Expression of the Genes Encoding the Propene Monooxygenase From Xanthobacter Py2." *Applied and Environmental Microbiology* 44: 582–588.

Zhou, N.-Y., A. Jenkins, C. K. Chan Kwo Chion, and D. J. Leak. 1998. "The Alkene Monooxygenase From Xanthobacter Py2 Is a Binuclear Non-haem iron Protein Closely Related to Toluene 4-Monooxygenase." *FEMS Letters* 430: 181–185.

Zhou, W., W. Guo, H. Zhou, and X. Chen. 2016. "Phenol Degradation by *Sulfobacillus acidophilus* TPY via the meta-Pathway." *Microbiological Research* 190: 37–45.

Zill, D., E. Lettau, C. Lorent, F. Seifert, P. K. Singh, and L. Lauterbach. 2022. "Crucial Role of the Chaperonin GroES/EL for Heterologous Production of the Soluble Methane Monooxygenase From *Methylomonas methanica* MC09." *Chembiochem* 23: e202200195.

Zou, B., Y. Huang, P. P. Zhang, X. M. Ding, H. J. M. Op den Camp, and Z. X. Quan. 2021. "Horizontal Gene Transfer of Genes Encoding Copper-Containing Membrane-Bound Monooxygenase (CuMMO) and Soluble di-iron Monooxygenase (SDIMO) in Ethane- and Propane-Oxidizing Rhodococcus bacteria." *Applied and Environmental Microbiology* 87: e0022721.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.