frontiers in Microbiology

ORIGINAL RESEARCH
published: 20 August 2019
doi: 10.3389/fmicb.2019.01862

Catabolism of Alkylphenols in *Rhodococcus via a Meta*-Cleavage Pathway Associated With Genomic Islands

David J. Levy-Booth†, *Morgan M. Fetherolf*†, *Gordon R. Stewart, Jie Liu, Lindsay D. Eltis and William W. Mohn**

Department of Microbiology and Immunology, Life Sciences Institute, The University of British Columbia, Vancouver, BC, Canada

The bacterial catabolism of aromatic compounds has considerable promise to convert lignin depolymerization products to commercial chemicals. Alkylphenols are a key class of depolymerization products whose catabolism is not well-elucidated. We isolated *Rhodococcus rhodochrous* EP4 on 4-ethylphenol and applied genomic and transcriptomic approaches to elucidate alkylphenol catabolism in EP4 and *Rhodococcus jostii* RHA1. RNA-Seq and RT-qPCR revealed a pathway encoded by the *aphABCDEFGHIQRS* genes that degrades 4-ethylphenol *via* the *meta*-cleavage of 4-ethylcatechol. This process was initiated by a two-component alkylphenol hydroxylase, encoded by the *aphAB* genes, which were upregulated ∼3,000-fold. Purified AphAB from EP4 had highest specific activity for 4-ethylphenol and 4-propylphenol (∼2,000 U/mg) but did not detectably transform phenol. Nevertheless, a Δ*aphA* mutant in RHA1 grew on 4-ethylphenol by compensatory upregulation of phenol hydroxylase genes (*pheA1-3*). Deletion of *aphC*, encoding an extradiol dioxygenase, prevented growth on 4-alkylphenols but not phenol. Disruption of *pcaL* in the β-ketoadipate pathway prevented growth on phenol but not 4-alkylphenols. Thus, 4-alkylphenols are catabolized exclusively *via meta*-cleavage in rhodococci while phenol is subject to *ortho*-cleavage. A putative genomic island encoding *aph* genes was identified in EP4 and several other rhodococci. Overall, this study identifies a 4-alkylphenol pathway in rhodococci, demonstrates key enzymes involved, and presents evidence that the pathway is encoded in a genomic island. These advances are of particular importance for wide-ranging industrial applications of rhodococci, including upgrading of lignocellulose biomass.

Keywords: aromatic, catabolism, alkylphenol, *Rhodococcus*, transcriptomics, *meta*-cleavage, genomic island

OPEN ACCESS

Edited by:
Haike Antelmann, Freie Universität Berlin, Germany

Reviewed by:
Keisuke Miyauchi, Tohoku Gakuin University, Japan
Eiji Masai, Nagaoka University of Technology, Japan

***Correspondence:**
William W. Mohn
wmohn@mail.ubc.ca

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *19 June 2019*
Accepted: *29 July 2019*
Published: *20 August 2019*

Citation:
Levy-Booth DJ, Fetherolf MM, Stewart GR, Liu J, Eltis LD and Mohn WW (2019) Catabolism of Alkylphenols in Rhodococcus via a Meta-Cleavage Pathway Associated With Genomic Islands. Front. Microbiol. 10:1862. doi: 10.3389/fmicb.2019.01862

INTRODUCTION

Lignin, a heterogeneous aromatic polymer, accounts for up to 40% dry weight of terrestrial plant biomass (Ragauskas et al., 2014). It is primarily composed of *p*-hydroxyphenyl (H), guaiacyl (G), and sinapyl (S) subunits, polymerized by ether and C–C bonds (Boerjan et al., 2003). Lignin's heterogeneity and recalcitrant bonds create substantial barriers to its efficient microbial and chemical degradation. Industrial lignin depolymerization is gaining traction as a means to produce fuels and chemicals historically derived from petroleum (Ragauskas et al., 2014; Beckham et al., 2016). Yet, heterogeneous depolymerization products can require extensive separation and

purification (Linger et al., 2014; Ragauskas et al., 2014). Bacterial biocatalysts provide a means of transforming mixtures of aromatic compounds to single compounds due to the convergent nature of their catabolic pathways, whereby substrates are transformed to central metabolites *via* shared intermediates, such as catechols (Linger et al., 2014; Eltis and Singh, 2018). Harnessing this biological funneling to refine lignin to high-value chemicals (Linger et al., 2014; Beckham et al., 2016; Eltis and Singh, 2018) is limited in part by a lack of knowledge of the catabolism of lignin-derived monomers.

Alkylphenols are a major class of aromatic compounds generated by a variety of lignin depolymerization technologies. For example, solvolysis of corn lignin produced 24 wt.% alkylated monolignins, 46% of which was 4-ethylphenol derived from H-subunits (Jiang et al., 2014). Alkylphenols were also major pyrolysis products of wheat straw black liquor lignin fractions (Guo et al., 2017). Existing depolymerization strategies can require multiple stages of pre-processing and depolymerization, high heat, or corrosive chemicals, and can produce dozens of alkylphenol and aromatic products (Ye et al., 2012; Kim et al., 2015; Asawaworarit et al., 2019). One promising depolymerization strategy that produces a narrow stream of alkylphenols is reductive catalytic fractionation (RCF) (Pepper and Lee, 1969). 4-Ethylphenol was a major RCF product of corn stover, comprising up to 16.4% of the resulting aromatic monomers (Anderson et al., 2016).

Two bacterial pathways for the aerobic catabolism of 4-ethylphenol have been reported, initially involving either oxidation of the alkyl side chain or hydroxylation of the aromatic ring. In *Pseudomonas putida* JD1, the alkyl side chain is oxidized by 4-ethylphenol methylhydroxylase to eventually yield hydroquinone (Darby et al., 1987; Hopper and Cottrell, 2003). In contrast, *Pseudomonas* sp. KL28 hydroxylates 4-ethylphenol to 4-ethylcatechol (Jeong et al., 2003). In these pathways, the hydroquinone and 4-ethylcatechol undergo *meta*-cleavage (Darby et al., 1987; Jeong et al., 2003). In KL28, the alkylphenol hydroxylase is a six-component enzyme encoded by genes organized in a single co-linear transcriptional unit within a *meta*-cleavage pathway gene cluster (Jeong et al., 2003). A homologous pathway degrades phenol in *Comamonas testosteroni* TA441 (Arai et al., 2000).

Rhodococcus is a genus of mycolic acid-producing Actinobacteria that catabolize a wide variety of aromatic compounds (Yam et al., 2010), including phenols (Kolomytseva et al., 2007; Gröning et al., 2014). These bacteria also have considerable potential as biocatalysts for the industrial production of compounds ranging from nitriles to steroids and high-value lipids (Alvarez et al., 1996; Round et al., 2017; Shields-Menard et al., 2017; Sengupta et al., 2019). In *Rhodococcus*, phenol catabolism is initiated by a two-component flavin-dependent monooxygenase (PheA1A2; Saa et al., 2010) to generate a catechol. PheA1A2 homologs in *Rhodococcus opacus* 1CP can also hydroxylate chlorophenols and 4-methylphenol (Gröning et al., 2014) to produce the corresponding catechols, which undergo *ortho*-cleavage (Maltseva et al., 1994; Kolomytseva et al., 2007) and subsequent transformation to central metabolites *via* the β-ketoadipate pathway. In *rhodococci*, the β-ketoadipate pathway is confluent, with branches responsible for protocatechuate and catechol catabolism converging at PcaL, a β-ketoadipate enol-lactonase (Patrauchan et al., 2005; Yam et al., 2010). However, some *rhodococci* appear to have pathways responsible for the catabolism of alkylated aromatic compounds *via meta*-cleavage (Jang et al., 2005). Elucidating 4-alkylphenol metabolism in *rhodococci* will improve our understanding of Actinobacterial aromatic degradation and support the development of *Rhodococcus* strains as platforms for industrial lignin upgrading.

Genomic islands (GIs) are DNA segments likely to have been acquired by horizontal gene transfer. They are characterized by altered nucleotide characteristics (e.g., GC content), syntenic conservation, and frequent presence of mobility genes [transposases, insertion sequences (IS), and integrases] (Hacker and Kaper, 2000; Juhas et al., 2009). They can be further identified by the absence of genomic regions in closely related strains (Hacker et al., 1990). GIs can confer resistance, virulence, symbiosis, and catabolic pathways (Dobrindt et al., 2004; Juhas et al., 2009). For example, the self-transferable *clc* element enabling 3- and 4-chlorocatechol and 2-aminophenol catabolism was identified as a GI in several *Gamma-* and *Betaproteobacteria* strains (Gaillard et al., 2006). Recent horizontal gene transfer may have played less of a role in shaping the *Rhodococcus jostii* RHA1 genome than in other bacteria such as *Burkholderia xenovorans* LB400, which has a similarly sized genome (McLeod et al., 2006). Further, although RHA1 contains a high number of aromatic pathways, genes encoding these pathways are slightly underrepresented in the identified GIs. GIs can ameliorate in host genomes through nucleotide optimization or loss of mobility elements (Lawrence and Ochman, 1997; Juhas et al., 2009), reducing our effectiveness at predicting ancestral genomic additions. However, examination of GIs in multiple related genomes with an ensemble of predictive software can improve our understanding of the role of GIs in the evolution of bacterial catabolic pathways.

This study sought to identify catabolic pathways required for 4-alkylphenol catabolism. We report on the genomic, transcriptomic, and enzymatic characterization of 4-alkylphenol catabolism in a newly isolated 4-ethylphenol-degrading bacterium, *Rhodococcus rhodochrous* EP4 (**Figure 1A**), as well as RHA1. The activity of a novel two-component alkylphenol monooxygenase (AphAB) was characterized. Gene deletion analysis was employed to identify the subsequent route of catechol catabolism. Genomic analysis identified a putative *aph* GI, providing new evolutionary insight into the *aph meta*-cleavage pathway in *Rhodococcus.* Knowledge gained by this study will facilitate the efficient valorization of lignin following its depolymerization.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Liquid enrichment cultures were inoculated with either ∼4-month-aged agricultural compost (∼25 cm depth) from The University of British Columbia farm (49° 14' 57.8904" N, 123° 14' 0.0492" W) or forest soil from Pacific Spirit Park in Vancouver,

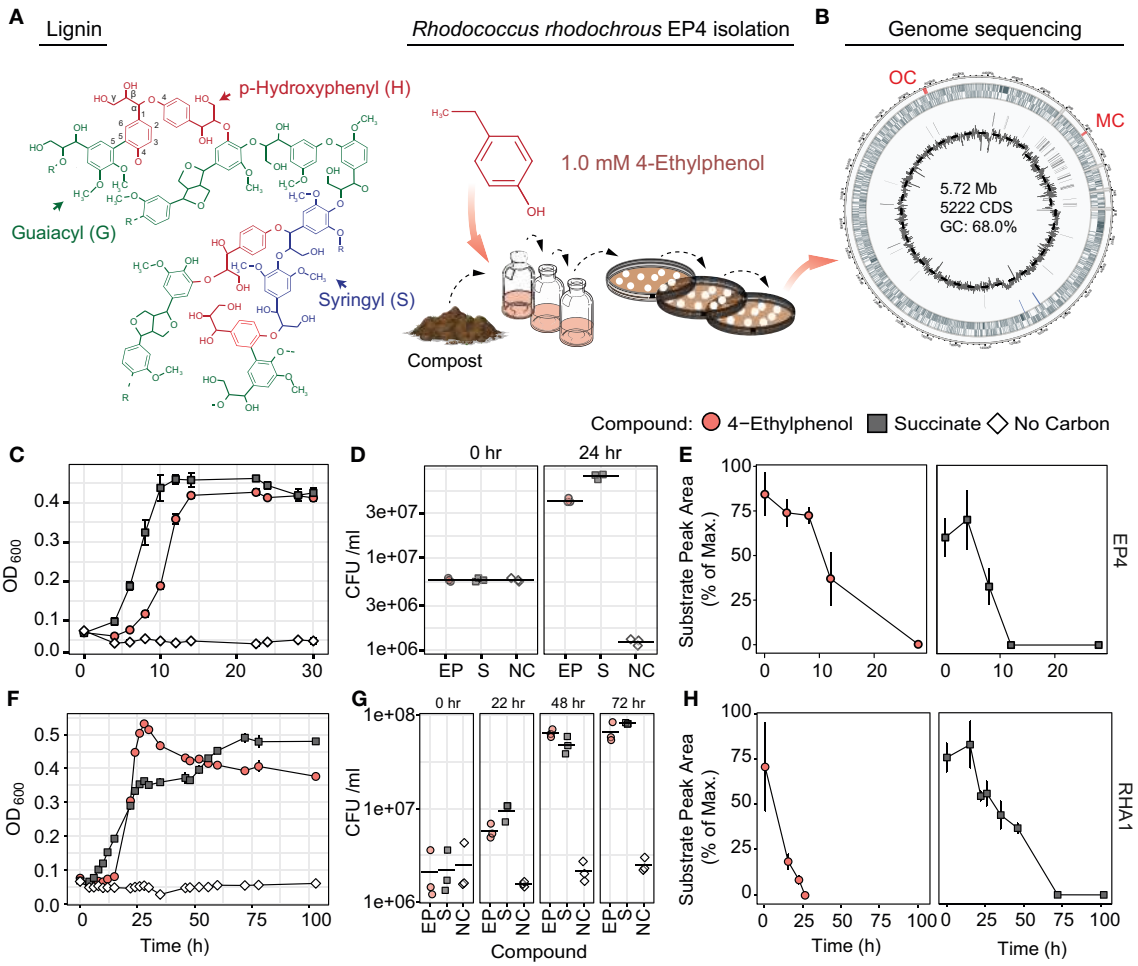

FIGURE 1 | Isolation and growth of 4-ethylphenol-catabolizing strain *R. rhodochrous* EP4. **(A)** Schematic of EP4 isolation from compost on 4-ethylphenol, which is produced from reductive lignin depolymerization. **(B)** Circular genome map for EP4. Outer genomic track: coding sequences by strand; inner track: insertion sequences; lines: GC content (%). MC, *meta*-cleavage pathway gene cluster; OC, *ortho*-cleavage pathway gene cluster. **(C)** Growth of EP4 on 1 mM 4-ethylphenol or 2 mM succinate controls as optical density at 600 nm (OD_{600}). Points and error bars reflect mean and standard error ($n = 3$). **(D)** Colony-forming units (CFU) during growth in **(B)**. Points and horizontal bar indicate individual measurements and mean. **(E)** Removal of 4-ethylphenol and succinate in EP4 cultures. **(F)** Growth of RHA1 on 1 mM 4-ethylphenol or 2 mM succinate. **(G)** CFU during growth in **(F)**. **(H)** Removal of 4-ethylphenol and succinate in RHA1 cultures.

Canada. The cultures contained 1.0 mM 4-ethylphenol (≥97.0% Sigma-Aldrich, St. Louis, U.S.A.) as sole organic substrate in M9-Goodies (Bauchop and Elsden, 1960; Elder, 1983). The cultures were incubated at 30°C with shaking at 200 rpm for 2 weeks. Removal of 4-ethylphenol was monitored by GC-MS, after which cultures were transferred to fresh medium, 0.5% inocula. After three serial transfers, isolates were obtained by plating on homologous medium solidified with 1.5% purified agar. Colonies appeared in 10 days. Individual colonies were transferred to liquid media and replated on solid media for colony isolation. *R. rhodochrous* strain DSM43241 was purchased from DSMZ (Braunschweig, Germany).

EP4 and RHA1 cultures for RNA extraction were grown overnight at 30°C on LB broth (200 rpm), diluted to 0.05 OD_{600}, and washed thrice in M9 with no supplement, and grown to mid-log phase on 50 ml M9-Goodies with 1.0 mM 4-ethylphenol or succinate. EP4 was additionally grown on 1.0 mM benzoate (99%, Sigma-Aldrich). Additional screening of EP4, *R. rhodochrous* DSM43241, RHA1, and mutant strains used 5 ml of M9-Goodies with 1.0 mM phenol (≥99%, VWR International, Ltd., Mississauga, Canada), 3-methylphenol (*m*-cresol; 99% Sigma-Aldrich), 4-methylphenol (99% Sigma-Aldrich), 4-propylphenol (> 99%, TCI), 3,4-dimethylphenol (DMP) (99% Sigma-Aldrich), 2,4-DMP (98% Sigma-Aldrich), 4-hydroxyphenylacetate (HPA) (98% Sigma-Aldrich), 4-hydroxybenzioic acid (HBA) (99%, Sigma-Aldrich), or 0.5 mM 4-nitrophenol (NP) (≥99%, Sigma-Aldrich), incubated for 24 h. Cells were lysed by boiling in 1 M NaOH and protein quantified using the Micro BCA™ Protein Assay (Thermo Fisher Scientific Inc., Waltham, U.S.A.) and a VersaMax™ microplate reader (Molecular Devices LLC, San Jose, U.S.A.).

DNA Manipulation, Plasmid Construction, and Gene Deletion

DNA was isolated, manipulated, and analyzed using standard protocols (Sambrock and Russel, 2001). *E. coli* and RHA1 were transformed with DNA by electroporation using a MicroPulser with GenePulser cuvettes (Bio-Rad). To produce N-terminal His_6-tagged $AphA_{EP4}$, $AphB_{EP4}$, and $AphC_{RHA1}$ (see **Table 1**), C6369_RS01585, C6369_RS01550, and RHA1_RS18750 were amplified from genomic DNA using Phusion Polymerase™ with the oligonucleotides listed in **Supplementary Table 1**. The nucleotide sequence of the cloned genes was verified. The Δ*aphA* and Δ*aphC* mutants were constructed using a *sacB* counter selection system (van der Geize et al., 2007). Five hundred-base pair flanking regions of RHA1_RS18785 and RHA1_RS18750 were amplified from RHA1 genomic DNA using the primers listed in **Supplementary Table 1**. The resulting amplicons were inserted into pK18mobsacB linearized with EcoRI using Gibson Assembly. The nucleotide sequences of the resulting constructs were verified. Kanamycin-sensitive/sucrose-resistant colonies were screened using PCR, and the gene deletion was confirmed by sequencing.

Enzyme Production and Purification

The production and purification of $AphA_{EP4}$, $AphB_{EP4}$, and $AphC_{RHA1}$ are described in **Supplementary Methods**.

GC/MS Analysis

Growth substrate depletion was analyzed in culture supernatants using an Agilent Technologies (Santa Clara, U.S.A.) 6890N gas chromatograph equipped with a 30-m Agilent 190915-433 capillary column and a 5973 mass-selective detector (GC/MS). Briefly, 400-μl samples of culture supernatant were amended with 3-chlorobenzoic acid (as internal standard), extracted with ethyl acetate (1:1 v/v), and dried under a nitrogen stream. The extract was suspended in pyridine and derivatized with trimethylsilyl for 1 h at 60°C. One-microliter injections were analyzed using the following parameters: transfer line temperature of 325°C, run temperature of 90°C for 3 min, and then ramped to 290°C at 12°C min^{-1} with a 10-min final hold. Peaks from raw trace files were aligned and integrated using *xcms* in R 3.4.4 (R Core Team, 2016) against 4-ethylphenol and succinate standards. Values were normalized to the area of the internal standard and expressed as a percent of maximum peak area.

Nucleic Acid Extraction and Sequencing

RNA was extracted from cellular pellets from 15 ml of EP4 and RHA1 cultures using TRIzol™ (Thermo-Fisher) and Turbo™ DNase (Thermo-Fisher). Quality and quantity of nucleic acids were assessed using 1% [w/v] agarose gel electrophoresis and Qubit fluorometric quantitation (Thermo-Fisher), prior to storage at −80°C. Approximately 1 μg of RNA underwent Ribo-Zero rRNA removal (Illumina, San Diego, U.S.A.), TruSeq LT (Illumina) library preparation, and sequencing using HiSeq4000 2 × 100 bp. Genomic DNA was extracted using CTAB. Fifteen micrograms was pulse-field electrophoresis size-selected and sequenced with one Pacific Biosciences (PacBio) RS II SMRT cell.

Bioinformatics

A *de novo* draft genome was assembled with HGAP in the SMRT Analysis v2.3 pipeline (Chin et al., 2013) and MeDuSa 1.6 scaffolding (Bosi et al., 2015). Circularizing the genome sequence was attempted using Circlator 1.5.5 (Hunt et al., 2015), and plasmid detection was attempted using PlasmidFinder 2.0.1 (Carattoli et al., 2014). Annotation used the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 4.4 and BLASTp against the Protein Data Bank (e-value 10^{-3}). Quality control, filtering, and trimming of RNA reads used Trimmomatic 0.3.6 defaults (Bolger et al., 2014). Assembly used Trinity 2.4.0 (Grabherr et al., 2011). Transcript quantification used HTSeq 0.9.1 (Anders et al., 2015), FeatureCounts 1.5.0-p3 (Liao et al., 2014), and Salmon 0.8.1 (Patro et al., 2017). Differential expression was analyzed using *DeSeq2* 1.18.1 (Love et al., 2014) with false discovery rate (*fdr*) correction. RT-qPCR conditions are in **Supplementary Methods**. All data visualization used *ggplot2* 3.1.0 unless otherwise noted.

Phylogenomic Characterization

Protein sequences were structurally aligned with T-Coffee 11.00 Expresso (Armougom et al., 2006); maximum-likelihood trees were generated using the best of 100 RAxML 8.0.0 iterations using the PROTGAMMALG model (Stamatakis, 2014) and visualized with iTOL v4 (Letunic and Bork, 2016). GI regions were predicted with IslandViewer 4 (Hsiao et al., 2003; Bertelli et al., 2017). All available *Rhodococcus* genomes ($n = 325$) were downloaded from NCBI RefSeq. Genomic alignment with a subset of genomes (**Supplementary Table 2**) used *nucmer* 3.1 (Marçais et al., 2018) and visualized with *circlize* 0.4.5 (Gu et al., 2014). Ribosomal protein trees were constructed as in Hug et al. (2016).

Enzyme Assays

$AphAB_{EP4}$ activity was measured spectrophotometrically by following the *meta*-cleavage of the produced 4-ethylcatechol in a coupled assay with $AphC_{RHA1}$ at 25 ± 0.5°C. The standard assay was performed in 200 μl of air-saturated 20 mM MOPS and 90 mM NaCl ($I = 0.1$ mM, pH 7.2) containing 0.5 mM 4-ethylphenol, 5 μM $AphA_{EP4}$, 1 μM $AphB_{EP4}$, 0.2 μM$AphC_{RHA1}$, 1,000 U ml^{-1} of catalase, 1 mM NADH, and 2.5 μM FAD. Components were incubated for 30 s, and then the reaction was initiated by adding NADH and was monitored at 400 nm. Absorbance was monitored using a Varian Cary 5000 spectrophotometer controlled by WinUV software. One unit of activity, *U*, was defined as the amount of enzyme required to hydroxylate of 1 nmol substrate per minute. Extinction coefficients for methyl-, ethyl-, and propylcatechol cleavage at 400 nm were 18,600, 15,100, and 19,400 M^{-1} cm^{-1}, respectively, calculated by differences in liberation of O_2 from alkylcatechol cleavage by 0.2 nmol $AphC_{RHA1}$ monitored using a Clark-type polarographic O_2 electrode OXYG1 (Hansatech, Pentney, UK) connected to a circulating water bath. Details of additional enzyme end point assays are provided in the **Supplementary Methods**.

TABLE 1 | Genes in the alkylphenol *meta*-cleavage pathway.

Gene	EP4[a]	RHA1[b]	Product[c]	Best Hit[d]	% ID[e]	References
aphA	RS01585	RS18785 (89%)	Alkylphenol hydroxylase, oxygenase	4-Nitrophenol 2-monooxygenase, oxygenase (NphA1) Q8RQQ0	87	Takeo et al., 2003
aphB	RS01580	RS18780 (85%)	Alkylphenol hydroxylase, reductase	NADH-dependent flavin reductase (NphB1) Q8RQP9	81	Takeo et al., 2003
aphC	RS01550	RS18750 (87%)	Alkylcatechol 2,3-dioxygenase	Biphenyl-2,3-diol 1,2-dioxygenase (BphC)[f] Q0S9X1	87	PDB entry, unpublished
aphD	RS01565	RS18765 (89%)	5-Alkyl-2-hydroxy-muconate-6-semialdehyde dehydrogenase	4-Hydroxymuconic-semialdehyde dehydrogenase (DmpC) P19059	45	Nordlund and Shingler, 1990
aphE	RS01575, RS01600	RS18775, RS18820 (69%)	5-Alkyl-2-hydroxymuconate tautomerase	2-Hydroxymuconate tautomerase (DmpI) P49172	38	Shingler et al., 1992
aphF	RS01605	RS18825 (85%)	Enol 5-alkyl-2-oxalocrotonate decarboxylase	4-Oxalocrotonate decarboxylase (NahK, DmpH) Q1XGK3	85	Tsuda and Iino, 1990
aphG	RS01610	RS18830 (85%)	2-Keto-4-alkylpentenoate hydratase	2-Keto-4-pentenoate hydratase (MhpD) P77608	42	Pollard and Bugg, 1998
aphH	RS01560	RS18760 (93%)	4-Hydroxy-2-alkylketopentenoate aldolase	4-Hydroxy-2-oxovalerate aldolase (DmpG, MhpE) P51016	48	Shingler et al., 1992
aphI	RS01555	RS18755 (91%)	Alkylacetaldehyde dehydrogenase	Acetaldehyde dehydrogenase (HsaG, MphF) P9WQH3	57	Carere et al., 2013
aphR	RS01590	RS18790 (55%)	Aph transcriptional regulator	AraC family transcriptional regulator Q88H39	39	Nelson et al., 2002
aphQ	RS01595	RS18810 (64%)	Aph transcriptional regulator	AraC family transcriptional regulator Q88H39	36	Nelson et al., 2002
aphS	RS01615	RS18835 (70%)	Aph transcriptional regulator	IclR family transcriptional regulator Q0SH23	31	Pouyssegur and Stoeber, 1974

[a] *Locus in EP4 (C6369_RSXXXXX).*
[b] *Locus in RHA1 (RHA1_RSXXXXX). Percent identity with EP4 homolog in parentheses.*
[c] *Where "alkyl" represents the variable-length 4-alkyl side chain.*
[d] *Gene name and Uniprot identifier of closest characterized homolog.*
[e] *Percent identity of best hit and EP4 homolog determined by Clustal Omega alignment.*
[f] *Also annotated as catechol 2,3-dioxygenase (DmpB).*

RESULTS

Isolation and Genomic Characterization of a 4-Ethylphenol-Degrading *Rhodococcus* Strain

In order to isolate environmental strains capable of degrading alkylphenols, enrichment cultures with 4-ethylphenol as a sole organic growth substrate were inoculated with either forest soil or compost and incubated at 30°C. Those cultures inoculated with compost demonstrated superior potential for 4-ethylphenol degradation and were used for subsequent isolation of strain EP4. The 16S rRNA gene (27F-1492R; Lane, 1991) of EP4 shared 100% sequence identity with that of *R. rhodochrous* NBRC 16069. *De novo* assembly produced a 5.72-Mb, high-quality, single-scaffold EP4 genome sequence (**Figure 1B**) containing 5,198 predicted genes: 4,942 protein coding sequences, 12 rRNAs, 54 tRNAs, three other RNAs, and 187 pseudogenes. Only a single origin of replication (*oriC*) was found, at 1,863,144 bp, and no plasmids were detected (**Supplementary Table 2**). The lack of PacBio long reads overlapping the 5′ and 3′ genome regions indicated that the EP4 genome is linear.

EP4 grew on 1.0 mM 4-ethylphenol to stationary phase within 14 h in shake flasks (**Figure 1C**). Growth on 4-ethylphenol was verified by plating CFUs (**Figure 1D**). GC-MS analysis indicated that 4-ethylphenol was completely removed from the medium during growth (**Figure 1E**), with no metabolites detected.

Quasi-Mapping-Based Quantification of Prokaryotic Gene Expression

We used transcriptomics to identify 4-ethylphenol catabolic genes in EP4 without *a priori* bias. Transcriptome reads were aligned strand-wise to predict transcriptional start sites in the EP4 genome (**Figure 2A**). Read alignment is

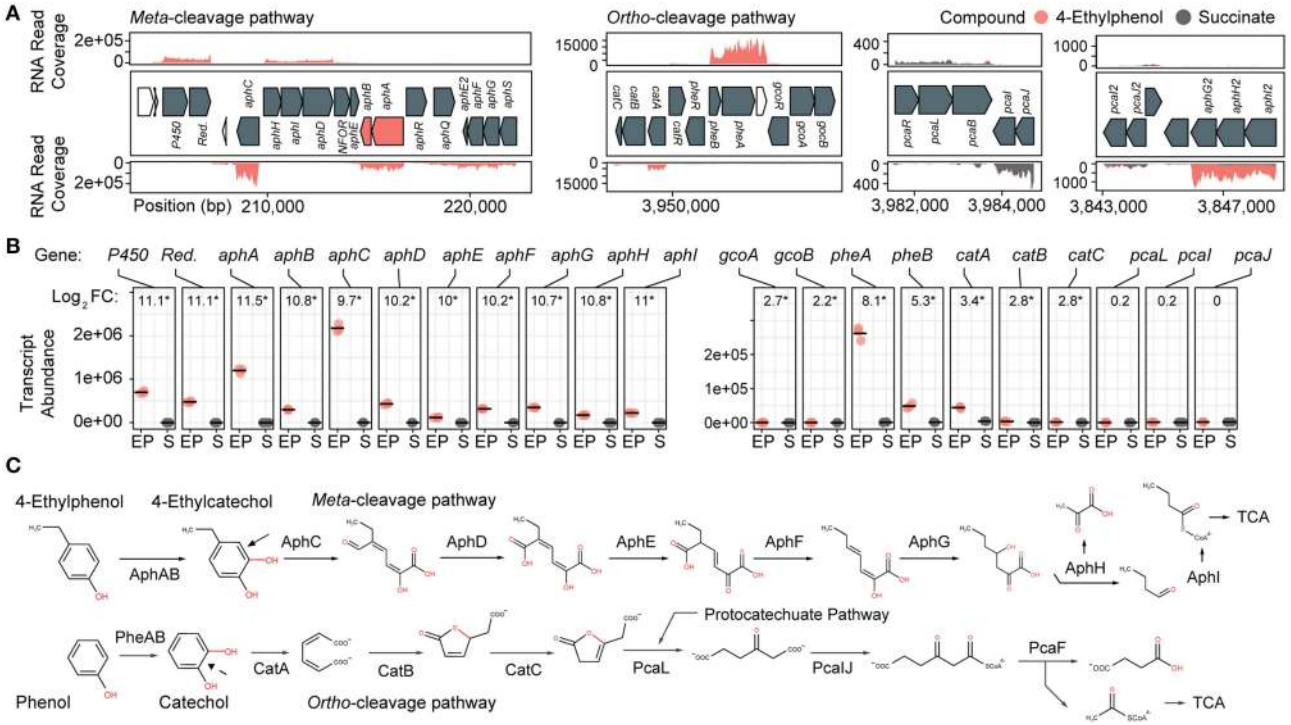

FIGURE 2 | Transcriptomic and genomic identification of the 4-ethylphenol catabolic pathway genes in *R. rhodochrous* EP4. **(A)** RNA reads from cells grown on 4-ethylphenol or succinate mapped to the EP4 gene clusters encoding catechol *meta*-cleavage and *ortho*-cleavage. **(B)** *Deseq2* differential-expression analysis showing log2 fold-change (FC) on 4-ethylphenol vs. succinate (FDR-corrected *p* values: $^*p_{fdr} < 0.001$). Points show values for $n = 3$ replicates; horizontal bar indicates mean. P450, cytochrome P450 gene; Red., P450 reductase. **(C)** Proposed funneling of 4-ethylphenol into the alkylcatechol *meta*-cleavage pathway (upper) and not the phenol *ortho*-cleavage pathway (lower). TCA, tricarboxylic acid cycle.

a common but time-consuming step in prokaryotic RNA-Seq pipelines (**Supplementary Figure 1A**). We therefore compared quasi-mapping to genomic coding regions using Salmon (Patro et al., 2017) with alignment-based read counting software. Salmon results were numerically (**Supplementary Figure 1B**) and statistically equivalent ($p_{adj} = 0.21$) (**Supplementary Figure 1C**) to FeatureCounts, with strong correlation to RT-qPCR expression ($R^2_{adj} = 0.91$, $p < 0.001$) (**Supplementary Figure 1D**). Salmon was about eight times faster than FeatureCounts and superior to htseq in terms of total counts and accuracy and was therefore used for gene quantification prior to differential expression analysis using *DESeq2*.

Transcriptomic Analysis of 4-Ethylphenol Metabolism *via meta*-Cleavage

Growth on 4-ethylphenol vs. succinate significantly modulated expression of 559 genes with $p_{fdr} < 0.001$. Nine of the 16 most upregulated genes occurred in a cluster encoding a proposed alkylphenol catabolic pathway, *aphABCDEFGHIQRS* (**Table 1**). This cluster includes *aphAB*, encoding a two-component alkylphenol hydroxylase discussed below, and *aphC*, encoding an extradiol dioxygenase that we subsequently identified as alkylcatechol 2,3-dioxygenase. The *aphA* gene displayed the highest upregulation on 4-ethylphenol. The gene cluster is organized as four putative operons based on transcriptomic data and operon prediction with BPROM: *aphAB*, *aphHIDE*, *aphE2FGS*, and *aphG2H2I2* (**Figure 2A**). The deduced Aph pathway catabolizes 4-ethylphenol to pyruvate and butyryl-CoA (**Figure 2C**), similar to the Dmp pathway of *Pseudomonas* sp. strain CF600 that catabolizes dimethylphenols (Shingler et al., 1992) and phenol (Powlowski and Shingler, 1994). Based on the transcriptomic data, the resultant butyryl-CoA is degraded to central metabolites by an aerobic fatty acid degradation pathway (Jimenez-Diaz et al., 2017) encoded by butyryl-CoA dehydrogenase genes (locus tags: C6369_RS06395, C6369_RS20140, C6369_RS07820, C6369_RS05465), enoylCoA hydratase (C6369 _RS19405, C6369_RS19860), 3-hydroxybutyryl-CoA dehydrogenase (C6369_RS03325, C6369_RS06400), and acetyl-CoA acyltransferase (C6369_RS17095, C6369_RS15900, C6369_RS19850) (**Supplementary Figure 2**).

The *catABC* cluster encoding catechol 1,2-dioxygenase and other enzymes feeding into the β-ketoadipate pathway was also significantly upregulated during growth on 4-ethylphenol, although much less highly than the *aph* genes. No *ortho*-cleavage metabolites were detected in the culture supernatants (**Figure 1D**), and the genes encoding the downstream β-ketoadipate pathway, *pcaBLIJ*, were not upregulated (**Figure 2B**). Overall, the data suggest that 4-ethyphenol is catabolized *via meta*-cleavage.

Characterization of a Two-Component Akylphenol Hydroxylase, AphAB

We hypothesized that the highly upregulated *aphA* gene (L_2FC = 11.5) encodes the oxygenase component of a novel alkylphenol monooxygenase, based on its location within the *aph* cluster as well as the phylogenetic and functional data presented below. The *aphB* gene, encoding a flavin reductase was co-transcribed with *aphA* (**Figures 2A,B**). The upregulation of *aphA* and *aphC* genes in EP4 during growth on 4-ethylphenol was confirmed using RT-qPCR (**Supplementary Figure 3**).

To establish the physiological role of AphAB from EP4, the oxygenase, and reductase components were each overproduced in *E. coli* and purified to apparent homogeneity. The reconstituted $AphAB_{EP4}$ hydroxylated 4-ethylphenol to 4-ethylcatechol (**Figure 3A**). The enzyme also catalyzed the hydroxylation of 4-methylphenol, 4-propylphenol (**Figure 3B**), and, to a much lesser extent, 4-NP (**Figure 3C**). However, $AphAB_{EP4}$ did not detectably transform phenol (**Figure 3B**) or 4-HPA (**Figure 3C**). In an assay measuring cytochrome *c* reduction, $AphB_{EP4}$ preferentially utilized NADH and flavin adenine dinucleotide (FAD) (**Figure 3D**), as reported for PheB homologs (Straube, 1987; Saa et al., 2010; Gröning et al., 2014).

Annotation of Additional Genes

During growth of EP4 on 4-ethylphenol, genes encoding phenol hydroxylase oxygenase and reductase, *pheA* and *pheB*, respectively, adjacent to the *cat* gene cluster, were additionally upregulated in a single putative regulon (**Figure 2B** and **Supplementary Figure 3**). In structure-based alignments, $PheA_{EP4}$ and $AphA_{EP4}$ clustered with separate 4-NP hydroxylases, rather than the clade of characterized phenol hydroxylases (**Figure 3E**). More specifically, $PheA_{EP4}$ and $AphA_{EP4}$ clustered most closely, respectively, with NphA1 from *Rhodococcus* sp. BUBNP1 (WP_059382681.1; Sengupta et al., 2019) and NphA1 from *Rhodococcus* sp. PN1 (Q8RQQ0; Takeo et al., 2003). Despite 100% sequence identity with $NphA1_{BUBNP1}$, $PheA_{EP4}$ was annotated based on sequence similarity to known phenol hydroxylases (**Figure 3E**). In support of this annotation, EP4 lacks a 4-NP catabolism gene cluster and was unable to grow on 4-NP, while it did grow on phenol (discussed below). $PheA_{EP4}$ shares 82% identity with PheA1(1) (ABS30825.1) in *Rhodococcus erythropolis* UPV-1 (Saa et al., 2010) and 65% identity with a chlorophenol 4-monooxygenase (Q8GMG6) from *Streptomyces globisporus* (Liu et al., 2002; **Supplementary Table 3**; **Supplementary Figure 4**). These similarities suggest that $PheA_{EP4}$ may have broad substrate specificity.

In EP4, genes encoding AraC-family transcriptional regulators (TRs) were found directly adjacent to and in the opposite orientation as *aphAB* and *pheAB* (**Figure 2A** and **Supplementary Figure 5**). These AraC-family TRs were annotated as AphR and PheR, respectively. Another AraC-family TR is encoded by a gene immediately downstream of *aphR*, which has a distinct phylogeny from AphR (**Supplementary Figure 5**) and was annotated as AphQ. Finally, an IclR-family TR is encoded by the last gene of the *aphE2FGS* operon.

Syntenic Conservation of EP4 *aph* Gene Cluster in Rhodococci

The above comparative analyses of hydroxylase proteins revealed homologs of the EP4 *aphA* gene in several other rhodococci. In RHA1, a putative *aphA* gene (**Table 1**) was previously annotated as an aromatic ring hydroxylase possibly involved in steroid degradation (McLeod et al., 2006) and there were three previously identified *pheA* homologs (**Supplementary Table 3**; Gröning et al., 2014). Local alignment of the 13 Aph proteins against proteins predicted from all 325 *Rhodococcus* genomes identified 75 strains with full or partial (≥7 genes) putative Aph pathways, including RHA1 (**Supplementary Figure 6A**). Related pathways were also found in other Actinobacteria, but this study focused on the rhodococcal pathway. RHA1 *aph* genes displayed syntenic conservation with the EP4 *aph* cluster (**Figure 4**), except for an additional butyryl-CoA dehydrogenase gene (RHA1_RS18815). The *aphCHIDE* region was conserved in all *aph*-containing genomes based on *nucmer* alignment (**Figure 4** and **Supplementary Figure 6B**). The EP4 and RHA1 Aph proteins shared 55% to 93% identity (median = 86%) (**Table 1**).

Identification of *aph* genes in RHA1 motivated us to test its growth on alklyphenols. Consistent with the occurrence of the putative Aph pathway in RHA1, this strain grew on 4-ethylphenol (**Figures 1F– H**, **5A**), with concomitant upregulation of the *aph* genes (**Figure 5B**). Specifically, *aphA* and *aphC* were highly upregulated (L_2FC, 14.3 and 10.0, respectively). In contrast to EP4, when RHA1 grew on 4-ethylphenol, it did not upregulate any of its three *pheA* genes or any of the ring-cleavage dioxygenase genes associated with the *pheA* genes, including *catA* (RHA1_RS11595), *catA2* (RHA1_RS35920), and a plasmid-borne catechol 2,3-dioxygenase gene (RHA1_RS35970; McLeod et al., 2006; **Figure 5B**).

Gene Deletion Analysis of 4-Alkylphenol Ring Cleavage

Efforts to genetically transform EP4 were unsuccessful. However, the above analysis demonstrated the presence of the Aph pathway in RHA1, a strain with existing gene deletion mutants and well-studied systems for genetic manipulation. Because RHA1 is genetically tractable and contains an orthologous Aph pathway, we constructed Δ*aphA* and Δ*aphC* deletion mutants and used these together with an existing Δ*pcaL* mutant to further investigate 4-ethylphenol catabolism. The Δ*aphC* mutant did not grow on either 1.0 mM 4-ethylphenol or 4-propylphenol (**Figures 5A,C**) demonstrating that both compounds are exclusively metabolized by *meta*-cleavage. No metabolites, including 4-ethylcatechol, accumulated in the culture supernatant, possibly due to lack of excretion of any transformed products (data not shown).

However, Δ*aphA* did grow on 4-ethylphenol (**Figure 5A**). It appears that one or more of three PheA homologs from RHA1 may catalyze 4-ethylphenol hydroxylation and compensate for the deletion of *aphA*. While the corresponding *pheA* genes were not upregulated in wild-type RHA1 growing on 4-ethylphenol vs. succinate, they were upregulated 7.6 to 8.8 L_2FC in the Δ*aphA*

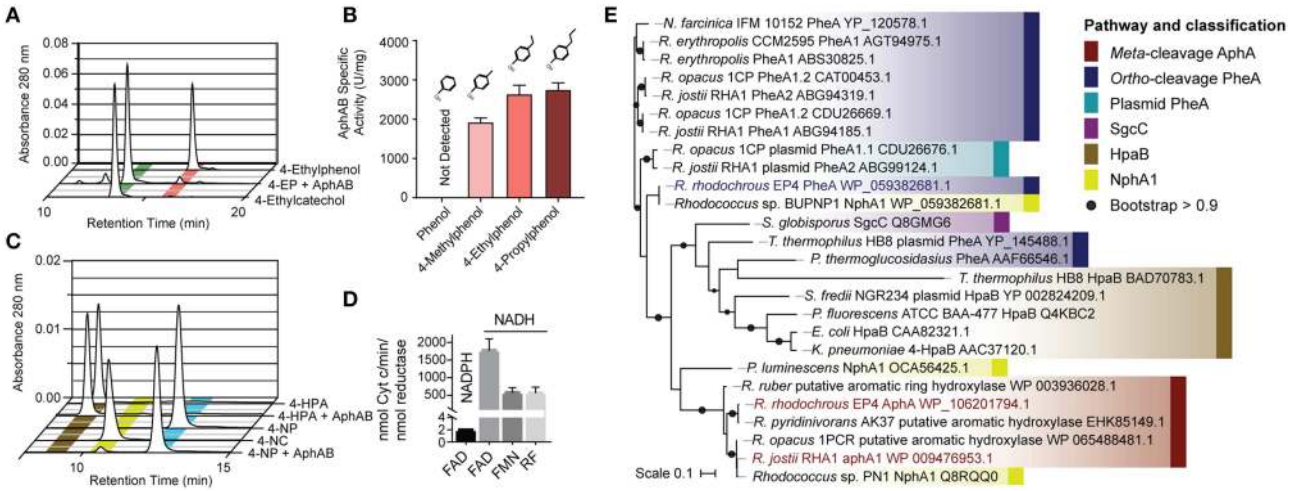

FIGURE 3 | Characterization of $AphAB_{EP4}$. **(A)** Hydroxylation of 4-ethylphenol to 4-ethlycatechol by purified $AphAB_{EP4}$. Reaction mixtures contained 20 μM of each enzyme component and 100 μM substrate, and were incubated overnight. **(B)** Specific activity of $AphAB_{EP4}$ for select phenols. Activity was measured using a coupled, spectrophotometric assay. **(C)** Transformation of 4-HPA and 4-NP by $AphAB_{EP4}$. Conditions as in **(A)**. **(D)** Cofactor and substrate preference of $AphB_{EP4}$. Reductase activity was measured using cytochrome *c*. **(E)** Phylogenetic tree constructed using structural-based alignment and RAxML. HPA, hydroxyphenylacetate; NP, nitrophenol; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; RF, riboflavin.

mutant (**Figure 5B**), while the plasmid-borne *C23D* gene was not upregulated. Finally, the Δ*pcaL* mutant grew on alkylphenols but did not grow on either phenol or 4-HBA, indicating that the latter two substrates are catabolized solely *via ortho*-cleavage pathways (**Figure 5C**).

Identification of a Putative *aph* GI

The *aph* gene cluster (∼17 kb) occurs within 117 and 4.2 kb regions predicted to be two of 61 GIs (or 38 non-overlapping GI regions) identified in EP4 using IslandViewer4 (**Figure 4** and **Supplementary Table 2**). These GI elements do not include *aphAB* or *aphE* in EP4, but do in three other *Rhodococcus* strains. These putative GIs are located near the 3′ end of the EP4 chromosome assembly in a 600-kb region of apparent genomic instability, as it contains high IS density (**Figure 1A**) and 36 predicted GIs (**Figure 4**). The GI-like characteristics of these elements containing the *aph* cluster include a −5.8% deviation from the mean GC content, presence of mobility genes (integrases, transposases, ISs), and absence of the region in closely related genomes following alignment (Langille et al., 2008; **Figure 4** and **Supplementary Figure 6**). They are not located near a tRNA sequence, indicating that it is not likely a mobile integrative and conjugative element (ICE) (Burrus and Waldor, 2004). More generally, 7.4% of the EP4 genome and 9.2% of the RHA1 genome were predicted to occur on GIs (**Supplementary Table 2**).

Analysis of 37 complete, full-length *Rhodococcus* genomes found that 16 carried genes encoding a complete Aph pathway. The *aph* genes are predicted to be fully or partially contained in a GI in six of these strains and to occur immediately downstream of a GI in a seventh, RHA1. This genomic region was conserved in three *Rhodococcus* clades: one containing EP4 and *R. pyridinivorans* strains; one containing *R. jostii*, *R. opacus*, and *Rhodococcus wratislaviensis* strains; and one containing *Rhodococcus* sp. Eu32 (**Supplementary Figure 6**). With the exception of a partial *aph* cluster in *R. rhodochrous* ATCC 21198, the *aph* genes were not found in any of the 13 other *R. rhodochrous* genomes including strain DSM43241. Accordingly, DSM43241 could not grow on 4-ethylphenol and 4-propylphenol, but grew on phenol, 4-HBA, 3-methylphenol, and 4-methylphenol (**Figure 5C**).

DISCUSSION

In this study, we used a variety of approaches to identify an Aph pathway responsible for the catabolism of alkylphenols *via meta*-cleavage in *Rhodococcus*. Catabolism is initiated by AphAB, a two-component hydroxylase that transforms the alkylphenol to the corresponding alkylcatechol (**Figures 2**, **3**). To date, only six-component proteobacterial alkylphenol hydroxylases have been reported (Arai et al., 2000; Jeong et al., 2003). The ensuing Aph pathway generates pyruvate and an acyl-CoA following *meta*-cleavage of 4-alkylcatechol. The length of the acyl-CoA produced depends on the alkyl side chain of the growth substrate. This is in contrast to the Phe and Nph pathways that catabolize phenol and 4-NP, respectively, *via ortho*-cleavage (Takeo et al., 2008; Szoköl et al., 2014; Sengupta et al., 2019).

The activity of $AphAB_{EP4}$ is consistent with its phylogenetic relationship with two-component phenolic hydroxylases. Thus, the clade containing $AphA_{EP4}$ and $AphA_{RHA1}$ includes an NphA1 but no characterized PheA or HpaB (**Figure 3**). In keeping with this classification, $AphAB_{EP4}$ had weak activity with 4-NP (**Figure 3C**) but did not detectably transform phenol or 4-HPA. However, the determinants of substrate specificity of these enzymes are not clear. The catalytic

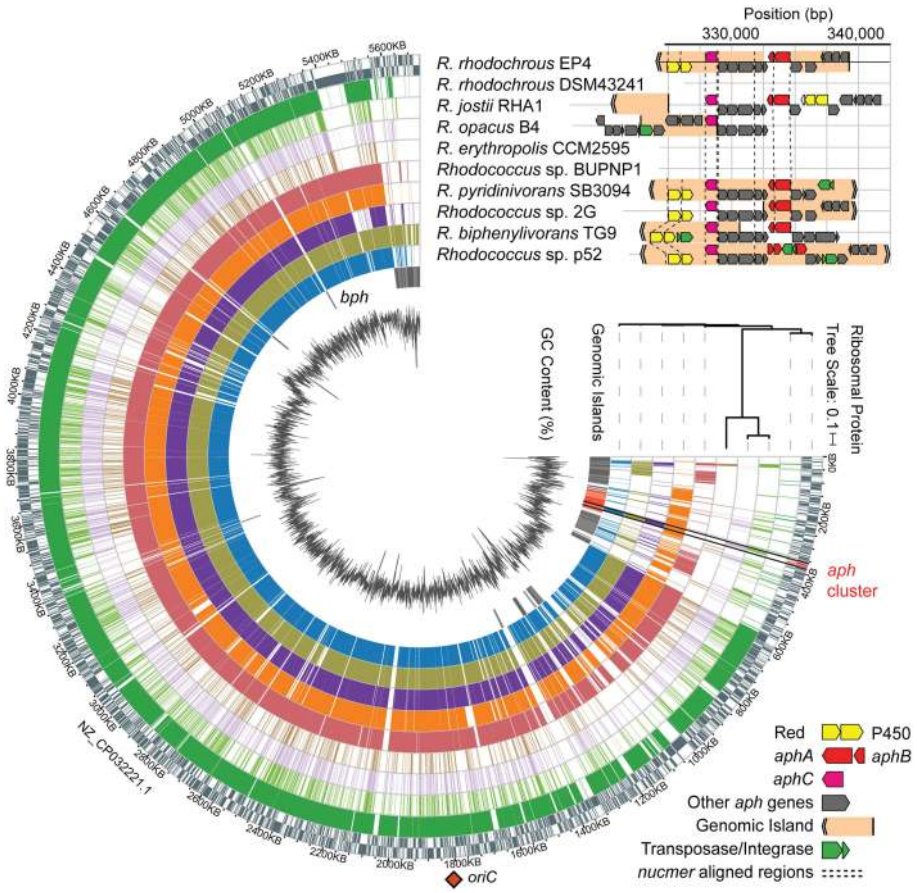

FIGURE 4 | Identification of a putative *aph* genomic island in rhodococci. Alignment of *Rhodococcus* genomes to EP4 reference with *nucmer* ordered by RAxML tree calculated from concatenated alignment of 16 ribosomal protein sequences, predicted RHA1 genomic islands, GC content (%), and syntentic organization of *aph* gene clusters showing genomic islands predicted using IslandViewer4. *Nucmer* alignment regions shown with dashed line. *aphA*, 4-alkylphenol 3-monooxygenase, oxygenase gene; *aphB*, 4-alkylphenol 3-monooxygenase, reductase gene; *aphC*, alkylcatechol 2,3-dioxygenase gene; *bph*, biphenyl catabolism gene cluster; P450, cytochrome P450 gene; Red., P450 reductase. The single origin of replication (*oriC*) shown with orange diamond. Detail of the *aph* region alignments in **Supplementary Figure 6B**.

residues of these hydroxylases (Kim et al., 2007; Chang et al., 2016) are conserved in AphA: Arg119, Tyr123, and His161 ($AphA_{EP4}$ numbering; **Supplementary Figure 4**). In a structurally characterized HpaB:4-HPA binary complex, the substrate's carboxylate is coordinated by Ser197 and Thr198 (Kim et al., 2007). In PheA, NphA, and AphA, these residues are His214 and Tyr215, suggesting that they do not contribute to the enzyme's substrate specificity despite their predicted interaction with the *para*-substituent of the substrate.

The Aph pathway is similar to the Dmp pathway described in *Pseudomonas* sp. strain CF600 (Shingler et al., 1992). However, it is clear that the Aph pathway has a distinct substrate specificity because neither EP4 nor RHA1 grew on 2,4- or 3,4-DMP (**Figure 5C**) and *aph* pathway mutants of RHA1 grew on phenol. We had previously suggested that some of the *aph* genes could be involved in steroid degradation (McLeod et al., 2006) due to their similarity to known steroid catabolic genes (van der Geize et al., 2007). Further, in a recently published structure of $AphC_{RHA1}$, the enzyme was identified as 2,3-dihydroxybiphenyl dioxygenase (**Table 1**). However, AphC is encoded in a gene cluster upregulated on 4-alkylphenols and is essential for growth of RHA1 on those compounds, supporting annotation of this rhodococcal Aph *meta*-cleavage pathway, with AphC as an alkylcatechol 2,3-dioxygenase.

4-Ethylphenol strongly induced *aphAB* expression. This is likely due to positive induction of the AphR TR, just as phenol activates *pheA2A1* expression by PheR in RHA1 (Szoköl et al., 2014), and 4-NP activates *npaA2A1* expression by NphR in *Rhodococcus* sp. PN1 (Takeo et al., 2008). AphR was identified based on conservation or syntentic organization relative to PheRAB, as both TR elements are expressed in opposition to their respective oxygenase components. AphR, PheR, and NphR are all AraC-family TRs. AphR and PheR may play a role in the unexpected ability of the RHA1 Δ*aphA* mutant to grow on 4-ethylphenol. The lack of *pheA1-3* expression in wild-type RHA1 (**Figure 5B**) strikingly contrasts with the upregulation of these genes in the Δ*aphA* mutant (**Figure 5C**). This phenomenon requires further study, as aphR and pheR bear little

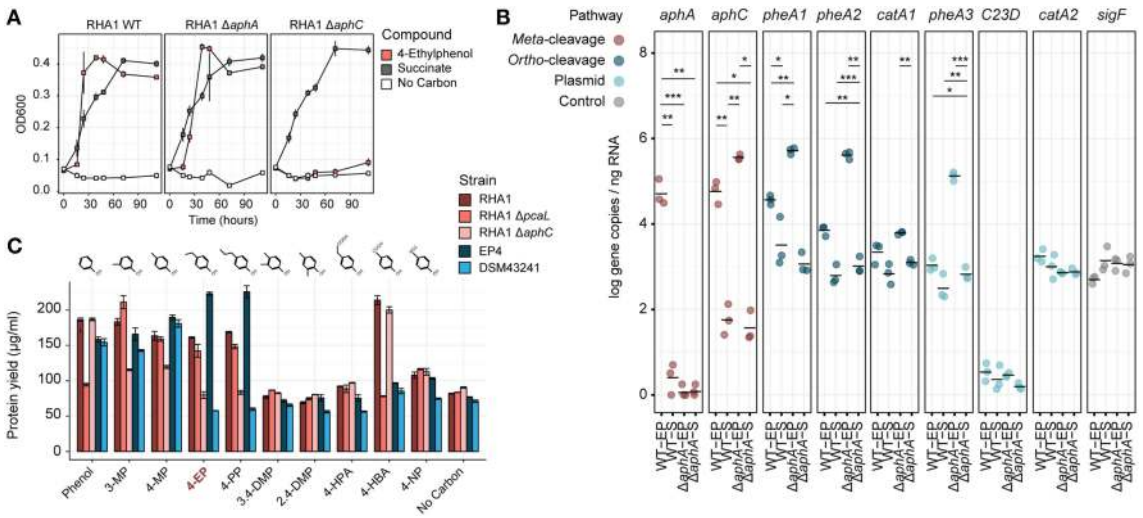

FIGURE 5 | Molecular genetic analysis of 4-ethylphenol catabolism in RHA1. **(A)** Growth of WT, Δ*aphA*, and Δ*aphC1* RHA1 strains on 1 mM 4-ethylphenol or 2 mM succinate. **(B)** Expression of select genes in WT and Δ*aphA* RHA1 strains during growth on 1 mM 4-ethylphenol or 2 mM succinate using RT-qPCR. Colors indicate cleavage pathway. Points and horizontal bars show individual measurements ($n = 3$) and mean. Significance levels following Bonferroni-corrected two-tailed Student's *t*-tests ($^{*}p_{bon} < 0.05$; $^{**}p_{bon} < 0.01$; $^{***}p_{bon} < 0.001$). **(C)** Protein yield of WT, Δ*aphC*, and Δ*pcaL* RHA1 strains as well as EP4 and *R. rhodochrous* DSM43241 on phenolic substrates. Protein measured after 24 h incubation. DMP, dimethylphenol; EP, ethylphenol; HBA, hydroxybenzoic acid; HPA; hydroxyphenylacetate; MP, methylphenol; NP, nitrophenol; PP, propylphenol.

phylogenetic similarity (**Supplementary Figure 5A**). Indeed, the putative *aphA* promotor sequences (−10: CAGGAG; −35: CCGTCT) (**Supplementary Figure 5B**) bear more similarity to the T80 promotor of *Mycobacterium tuberculosis* (Bashyam et al., 1996) than with the rhodococcal *pheA* promotors. Related to this, homologs of PheAB in *R. opacus* 1CP hydroxylated 4-methylphenol with about twice the specific activity as with phenol (Gröning et al., 2014), further suggesting that the RHA1 PheABs may hydroxylate 4-ethylphenol.

In addition to 4-ethylphenol, alkylguaiacols and alkylsyringols commonly occur in lignin depolymerization streams (Ye et al., 2012; Jiang et al., 2014; Kim et al., 2015; Anderson et al., 2016; Guo et al., 2017; Asawaworarit et al., 2019). Interestingly, genes encoding a cytochrome P450 and reductase are linked to the *aph* clusters in some rhodococci (**Figures 2, 4**). Further, these were the second and third most highly upregulated genes in EP4 during growth on 4-ethylphenol versus succinate (both $L_2FC = 11.1$; **Figure 2B**). The P450 shares 65% sequence identity with a guaiacol *O*-demethylase (Mallinson et al., 2018), suggesting that the rhodococcal enzyme has a similar role, and that these strains may also funnel methoxylated compounds into the Aph pathway.

The ability of RHA1 and EP4 to catabolize 4-ethylphenol and other alkylphenols is of potential use in upgrading lignin streams generated by RCF and other depolymerization strategies. The Aph *meta*-cleavage pathway harbored by these strains contrasts with the *ortho*-cleavage pathways targeted to date in the design of biocatalysts for lignin valorization (Abdelaziz et al., 2016; Beckham et al., 2016; Barton et al., 2018). This is largely due to the identified economic potential of some of the *ortho*-cleavage metabolites. For example, *cis*, *cis*-muconate resulting from *ortho*-cleavage can be used to make adipic acid and terephthalic acid (Xie et al., 2014; Beckham et al., 2016; Barton et al., 2018). However, alkylphenols may be funneled through *ortho*-cleavage by oxidizing the *para*-side chain to 4-hydroxyacetophenone or hydroquinone (Darby et al., 1987). Alternatively, oleaginous *Rhodococcus* strains such as RHA1 may be modified to use the Aph pathway to produce lipid-based commodity chemicals (e.g., Round et al., 2017) offering a method for valorization of alkylphenols *via* fatty-acid synthesis in this genus.

We found genes encoding the Aph pathway in several *Rhodococcus* strains, including oleaginous strains such as RHA1 and *R. opacus* B4 (**Figures 3A**, **4**). However, the absence of an *aph* cluster in most *R. rhodochrous* strains (e.g., DSM43241) demonstrates that phylogenetically related strains can have important metabolic differences. Previous studies suggested that recent horizontal gene transfer did not play a large role in generating RHA1's considerable catabolic capabilities (McLeod et al., 2006). In several rhodococci, putative GIs did not contain all of the *aph* genes. This could represent the imprecision of the prediction tools, incomplete amelioration of the element, or reflect that these GIs arose from separate insertion events. Patchwork *aph* GIs are consistent with the theorized modular origins of GIs (Juhas et al., 2009). Fermentation of plant-derived aromatic compounds, including cinnamic acids, by yeasts and lactic acid bacteria can naturally produce alkylphenols (Caboni et al., 2007; Kridelbaugh et al., 2010). The apparent complete loss of the *aph* genes in other *R. rhodochrous* strains may result without selective alkylphenol exposure if it otherwise has a deleterious effect on overall fitness. Testing these regions for their excision capacity was beyond the scope of this work, but remains an intriguing prospect. We posit that the presence of aromatic compounds in compost selected for microorganisms capable of 4-ethylphenol catabolism, including EP4.

In this study, we described a newly isolated, 4-ethylphenol-catabolizing strain, EP4, a novel alkylphenol hydroxylase, AphAB, and its role in funneling alkylphenols into the Aph *meta*-cleavage pathway in some *Rhodococcus* strains. We showed that this pathway is associated with putative GIs, primarily found in strains from contaminated soil environments. Characterizing 4-ethylphenol metabolism in EP4 and RHA1 advances our capacity for bio-refinement of reductively depolymerized lignin subunits from sustainable chemical feedstocks.

DATA AVAILABILITY

The EP4 genome assembly can be downloaded from NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCA_003004765.2). The transcriptome data can be downloaded from SRA at accessions SRR6877528–SRR6877536. All data can be downloaded and the full transcriptomic analysis can be run using a custom pipeline found at: https://github.com/levybooth/Rhodococcus_Transcriptomics.

AUTHOR CONTRIBUTIONS

DL-B performed all genomic and transcriptomic analysis, performed growth experiments, and wrote the manuscript. MF performed growth and enzymological experiments, and edited the manuscript. GS isolated the EP4 strain. JL produced the RHA1 mutants. WM and LE designed the study and edited the manuscript. DL-B and MF contributed equally to this work.

FUNDING

This study was supported by a research contract from Genome BC (SIP004) and a grant from the Natural Sciences and Engineering Research Council of Canada (STPGP 506595-17). LE is the recipient of a Canada Research Chair.

ACKNOWLEDGMENTS

All sequencing was performed at the McGill University and Génome Québec Innovation Center (Montreal, CAN). We thank Andrew Wilson and Alexandra Booth for their assistance with laboratory experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01862/full#supplementary-material

REFERENCES

Abdelaziz, O. Y., Brink, D. P., Prothmann, J., Ravi, K., Sun, M., García-Hidalgo, J., et al. (2016). Biological valorization of low molecular weight lignin. *Biotechnol. Adv.* 34, 1318–1346. doi: 10.1016/j.biotechadv.2016.10.001

Alvarez, H. M., Mayer, F., Fabritius, D., and Steinbüchel, A. (1996). Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. *Arch. Microbiol.* 165, 377–386.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Anderson, E. M., Katahira, R., Reed, M., Resch, M. G., Karp, E. M., Beckham, G. T., et al. (2016). Reductive catalytic fractionation of corn stover lignin. *ACS Sustain. Chem. Eng.* 4, 6940–6950. doi: 10.1021/acssuschemeng.6b01858

Arai, H., Ohishi, T., Chang, M. Y., and Kudo, T. (2000). Arrangement and regulation of the genes for meta-pathway enzymes required for degradation of phenol in *Comamonas testosteroni* TA441. *Microbiology* 146(Pt 7), 1707–1715. doi: 10.1099/00221287-146-7-1707

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., et al. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* 34, W604–W608. doi: 10.1093/nar/gkl092

Asawaworarit, P., Daorattanachai, P., Laosiripojana, W., Sakdaronnarong, C., Shotipruk, A., and Laosiripojana, N. (2019). Catalytic depolymerization of organosolv lignin from bagasse by carbonaceous solid acids derived from hydrothermal of lignocellulosic compounds. *Chem. Eng. J.* 356, 461–471. doi: 10.1016/j.cej.2018.09.048

Barton, N., Horbal, L., Starck, S., Kohlstedt, M., Luzhetskyy, A., and Wittmann, C. (2018). Enabling the valorization of guaiacol-based lignin: Integrated chemical and biochemical production of cis,cis-muconic acid using metabolically engineered *Amycolatopsis* sp ATCC 39116. *Metab. Eng.* 45, 200–210. doi: 10.1016/j.ymben.2017.12.001

Bashyam, M. D., Kaushal, D., Dasgupta, S. K., and Tyagi, A. K. (1996). A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. *J. Bacteriol.* 178, 4847–4853. doi: 10.1128/jb.178.16.4847-4853.1996

Bauchop, T., and Elsden, S. R. (1960). The growth of micro-organisms in relation to their energy supply. *Microbiology* 23, 457–469. doi: 10.1099/00221287-23-3-457

Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D., and Vardon, D. R. (2016). Opportunities and challenges in biological lignin valorization. *Curr. Opinion Biotechnol.* 42, 40–53. doi: 10.1016/j.copbio.2016.02.030

Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoad, G., Winsor, G. L., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 45, W30–W35. doi: 10.1093/nar/gkx343

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M.-F., Lió, P., et al. (2015). MeDuSa: a multi-draft based scaffolder. *Bioinformatics* 31, 2443–2451. doi: 10.1093/bioinformatics/btv171

Burrus, V., and Waldor, M. K. (2004). Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.* 155, 376–386. doi: 10.1016/j.resmic.2004.01.012

Caboni, P., Sarais, G., Cabras, M., and Angioni, A. (2007). Determination of 4-ethylphenol and 4-ethylguaiacol in wines by LC-MS-MS and HPLC-DAD-fluorescence. *J. Agric. Food Chem.* 55, 7288–7293. doi: 10.1021/jf071156m

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14

Carere, J., McKenna, S. E., Kimber, M. S., and Seah, S. Y. K. (2013). Characterization of an aldolase–dehydrogenase complex from the cholesterol degradation pathway of *Mycobacterium tuberculosis*. *Biochemistry* 52, 3502–3511. doi: 10.1021/bi400351h

Chang, C.-Y., Lohman, J. R., Cao, H., Tan, K., Rudolf, J. D., Ma, M., et al. (2016). Crystal structures of SgcE6 and SgcC, the two-component monooxygenase that catalyzes hydroxylation of a carrier protein-tethered substrate during the biosynthesis of the enediyne antitumor antibiotic C-1027 in *Streptomyces globisporus*. *Biochemistry* 55, 5142–5154. doi: 10.1021/acs.biochem.6b00713

Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474

Darby, J. M., Taylor, D. G., and Hopper, D. J. (1987). Hydroquinone as the ring-fission substrate in the catabolism of 4-ethylphenol and 4-hydroxyacetophenone by *Pseudomonas putida* JD1. *Microbiology* 133, 2137–2146. doi: 10.1099/00221287-133-8-2137

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* 2, 414–424. doi: 10.1038/nrmicro884

Elder, R. T. (1983). Cloning techniques. *BioScience* 33, 721–722. doi: 10.2307/1309366

Eltis, L. D. and Singh, R. (2018). "Chapter 11: Biological funneling as a means of transforming lignin-derived aromatic compounds into value-added chemicals," in *Lignin Valorization*, ed G. T. Beckham (London: The Royal Society of Chemistry). doi: 10.1039/9781788010351-00290

Gaillard, M., Vallaeys, T., Vorhölter, F. J., Minoia, M., Werlen, C., Sentchilo, V., et al. (2006). The *clc* element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. *J. Bacteriol.* 188, 1999–2013. doi: 10.1128/JB.188.5.1999-2013.2006

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Gröning, J. A. D., Eulberg, D., Tischler, D., Kaschabek, S. R., and Schlömann, M. (2014). Gene redundancy of two-component (chloro)phenol hydroxylases in *Rhodococcus opacus* 1CP. *FEMS Microbiol. Lett.* 361, 68–75. doi: 10.1111/1574-6968.12616

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812. doi: 10.1093/bioinformatics/btu393

Guo, D., Wu, S., Lyu, G., and Guo, H. (2017). Effect of molecular weight on the pyrolysis characteristics of alkali lignin. *Fuel* 193, 45–53. doi: 10.1016/j.fuel.2016.12.042

Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R., et al. (1990). Deletions of chromosomal regions coding for fimbriae and hemolysins occur *in vitro* and *in vivo* in various extra intestinal *Escherichia coli* isolates. *Microb. Pathog.* 8, 213–225. doi: 10.1016/0882-4010(90)90048-U

Hacker, J., and Kaper, J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54, 641–679. doi: 10.1146/annurev.micro.54.1.641

Hopper, D. J., and Cottrell, L. (2003). Alkylphenol biotransformations catalyzed by 4-ethylphenol methylenehydroxylase. *Appl. Environ. Microbiol.* 69, 3650–3652. doi: 10.1128/AEM.69.6.3650-3652.2003

Hsiao, W., Wan, I., Jones, S. J., and Brinkman, F. S. L. (2003). IslandPath: aiding detection of genomic islands in prokaryotes. *Bioinformatics* 19, 418–420. doi: 10.1093/bioinformatics/btg004

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree of life. *Nat. Microbiol.* 1:16048. doi: 10.1038/nmicrobiol.2016.48

Hunt, M., Silva, N. D., Otto, T. D., Parkhill, J., Keane, J. A., and Harris, S. R. (2015). Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 16:294. doi: 10.1186/s13059-015-0849-0

Jang, J. Y., Kim, D., Bae, H. W., Choi, K. Y., Chae, J.-C., Zylstra, G. J., et al. (2005). Isolation and characterization of a *Rhodococcus* species strain able to grow on *ortho*- and *para*-xylene. *J. Microbiol.* 43, 325–330.

Jeong, J. J., Kim, J. H., Kim, C.-K., Hwang, I., and Lee, K. (2003). 3- and 4-alkylphenol degradation pathway in *Pseudomonas* sp. strain KL28: genetic organization of the *lap* gene cluster and substrate specificities of phenol hydroxylase and catechol 2,3-dioxygenase. *Microbiology* 149, 3265–3277. doi: 10.1099/mic.0.26628-0

Jiang, Z., He, T., Li, J., and Hu, C. (2014). Selective conversion of lignin in corncob residue to monophenols with high yield and selectivity. *Green Chem.* 16, 4257–4265. doi: 10.1039/C4GC00620H

Jimenez-Diaz, L., Caballero, A., and Segura, A. (2017). "Pathways for the degradation of fatty acids in bacteria," in *Aerobic Utilization of Hydrocarbons, Oils and Lipids, Handbook of Hydrocarbon and Lipid Microbiology*, ed. F. Rojo (Cham: Springer International Publishing), 1–23.

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., and Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* 33, 376–393. doi: 10.1111/j.1574-6976.2008.00136.x

Kim, J.-Y., Park, J., Kim, U.-J., and Choi, J. W. (2015). Conversion of lignin to phenol-rich oil fraction under supercritical alcohols in the presence of metal catalysts. *Energy Fuels* 29, 5154–5163. doi: 10.1021/acs.energyfuels.5b01055

Kim, S.-H., Hisano, T., Takeda, K., Iwasaki, W., Ebihara, A., and Miki, K. (2007). Crystal structure of the oxygenase component (HpaB) of the 4-hydroxyphenylacetate 3-monooxygenase from *Thermus thermophilus* HB8. *J. Biol. Chem.* 282, 33107–33117. doi: 10.1074/jbc.M703440200

Kolomytseva, M. P., Baskunov, B. P., and Golovleva, L. A. (2007). Intradiol pathway of *para*-cresol conversion by *Rhodococcus opacus* 1CP. *Biotechnol. J.* 2, 886–893. doi: 10.1002/biot.200700013

Kridelbaugh, D., Hughes, S., Allen, T., and Doerner, K. C. (2010). Production of 4-ethylphenol from 4-hydroxycinnamic acid by *Lactobacillus* sp. isolated from a swine waste lagoon. *J. Appl. Microbiol.* 109, 190–198. doi: 10.1111/j.1365-2672.2009.04642.x

Lane, D. J. (1991). "16S/23S rRNA sequencing," in *Nucleic Acid Techniques in Bacterial Systematic*, eds E. Stackebrandt, and M. Goodfellow (New York, NY: John Wiley and Sons), 115–175.

Langille, M. G., Hsiao, W. W., and Brinkman, F. S. (2008). Evaluation of genomic island predictors using a comparative genomics approach. *BMC Bioinformatics* 9:329. doi: 10.1186/1471-2105-9-329

Lawrence, J. G., and Ochman, H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* 44, 383–397.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Linger, J. G., Vardon, D. R., Guarnieri, M. T., Karp, E. M., Hunsinger, G. B., Franden, M. A., et al. (2014). Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12013–12018. doi: 10.1073/pnas.1410657111

Liu, W., Christenson, S. D., Standage, S., and Shen, B. (2002). Biosynthesis of the enediyne antitumor antibiotic C-1027. *Science* 297, 1170–1173. doi: 10.1126/science.1072110

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Mallinson, S. J. B., Machovina, M. M., Silveira, R. L., Garcia-Borràs, M., Gallup, N., Johnson, C. W., et al. (2018). A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. *Nat. Commun.* 9:2487. doi: 10.1038/s41467-018-04878-2

Maltseva, O. V., Solyanikova, I. P., and Golovleva, L. A. (1994). Chlorocatechol 1,2-dioxygenase from *Rhodococcus erythropolis* 1CP. Kinetic and immunochemical comparison with analogous enzymes from gram-negative strains. *Eur. J. Biochem.* 226, 1053–1061.

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: a fast and versatile genome alignment system. *PLoS Comput. Biol.* 14:e1005944. doi: 10.1371/journal.pcbi.1005944

McLeod, M. P., Warren, R. L., Hsiao, W. W. L., Araki, N., Myhre, M., Fernandes, C., et al. (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15582–15587. doi: 10.1073/pnas.0607048103

Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V., et al. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* 4, 799–808. doi: 10.1046/j.1462-2920.2002.00366.x

Nordlund, I., and Shingler, V. (1990). Nucleotide sequences of the meta-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxymuconic semialdehyde hydrolase from *Pseudomonas* CF600. *Biochim. Biophys. Acta* 1049, 227–230.

Patrauchan, M. A., Florizone, C., Dosanjh, M., Mohn, W. W., Davies, J., and Eltis, L. D. (2005). Catabolism of benzoate and phthalate in *Rhodococcus* sp.

strain RHA1: redundancies and convergence. *J. Bacteriol.* 187, 4050–4063. doi: 10.1128/JB.187.12.4050-4063.2005

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Pepper, J. M., and Lee, Y. W. (1969). Lignin and related compounds. *I.* A comparative study of catalysts for lignin hydrogenolysis. *Can. J. Chem.* 47, 723–727. doi: 10.1139/v69-118

Pollard, J. R., and Bugg, T. D. (1998). Purification, characterisation and reaction mechanism of monofunctional 2-hydroxypentadienoic acid hydratase from *Escherichia coli. Eur. J. Biochem.* 251, 98–106.

Pouyssegur, J., and Stoeber, F. (1974). Genetic control of the 2-keto-3-deoxy-d-gluconate metabolism in *Escherichia coli* K-12: kdg regulon. *J. Bacteriol.* 117, 641–651.

Powlowski, J., and Shingler, V. (1994). Genetics and biochemistry of phenol degradation by *Pseudomonas* sp. CF600. *Biodegradation* 5, 219–236.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., et al. (2014). Lignin valorization: improving lignin processing in the biorefinery. *Science* 344:1246843. doi: 10.1126/science.1246843

Round, J., Roccor, R., Li, S.-N., and Eltis, L. D. (2017). A fatty acyl coenzyme A reductase promotes wax ester accumulation in *Rhodococcus jostii* RHA1. *zppl. Environ. Microbiol.* 83, e00902–e00917. doi: 10.1128/AEM.00902-17

Saa, L., Jaureguibeitia, A., Largo, E., Llama, M. J., and Serra, J. L. (2010). Cloning, purification and characterization of two components of phenol hydroxylase from *Rhodococcus erythropolis* UPV-1. *Appl. Microbiol. Biotechnol.* 86, 201–211. doi: 10.1007/s00253-009-2251-x

Sambrock, J., and Russel, D. W. (2001). *Molecular Cloning: A Laboratory Manual. 3rd Edn.* New York, NY: Cold Spring Harbor Laboratory Press.

Sengupta, K., Swain, M. T., Livingstone, P. G., Whitworth, D. E., and Saha, P. (2019). Genome sequencing and comparative transcriptomics provide a holistic view of 4-nitrophenol degradation and concurrent fatty acid catabolism by *Rhodococcus* sp. strain BUPNP1. *Front. Microbiol.* 9:3209. doi: 10.3389/fmicb.2018.03209

Shields-Menard, S. A., AmirSadeghi, M., Green, M., Womack, E., Sparks, D. L., Blake, J., et al. (2017). The effects of model aromatic lignin compounds on growth and lipid accumulation of *Rhodococcus rhodochrous*. *Int. Biodeterior. Biodegr.* 121, 79–90. doi: 10.1016/j.ibiod.2017.03.023

Shingler, V., Powlowski, J., and Marklund, U. (1992). Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. *J. Bacteriol.* 174, 711–724.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Straube, G. (1987). Phenol hydroxylase from *Rhodococcus* sp. P 1. *J. Basic Microbiol.* 27, 229–232. doi: 10.1002/jobm.3620270415

Szoköl, J., Rucká, L., Šimčíková, M., Halada, P., Nešvera, J., and Pátek, M. (2014). Induction and carbon catabolite repression of phenol degradation genes in *Rhodococcus erythropolis* and *Rhodococcus jostii*. *Appl. Microbiol. Biotechnol.* 98, 8267–8279. doi: 10.1007/s00253-014-5881-6

Takeo, M., Murakami, M., Niihara, S., Yamamoto, K., Nishimura, M., Kato, D., et al. (2008). Mechanism of 4-nitrophenol oxidation in *Rhodococcus* sp. strain PN1: characterization of the two-component 4-nitrophenol hydroxylase and regulation of its expression. *J. Bacteriol.* 190, 7367–7374. doi: 10.1128/JB.00742-08

Takeo, M., Yasukawa, T., Abe, Y., Niihara, S., Maeda, Y., and Negoro, S. (2003). Cloning and characterization of a 4-nitrophenol hydroxylase gene cluster from *Rhodococcus* sp. PN1. *J. Biosci. Bioeng.* 95, 139–145. doi: 10.1016/S1389-1723(03)80119-6

Tsuda, M., and Iino, T. (1990). Naphthalene degrading genes on plasmid NAH7 are on a defective transposon. *Mol. Gen. Genet.* 223, 33–39.

van der Geize, R., Yam, K., Heuser, T., Wilbrink, M. H., Hara, H., Anderton, M. C., et al. (2007). A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *PNAS* 104, 1947–1952. doi: 10.1073/pnas.0605728104

Xie, N.-Z., Liang, H., Huang, R.-B., and Xu, P. (2014). Biotechnological production of muconic acid: current status and future prospects. *Biotechnol. Adv.* 32, 615–622. doi: 10.1016/j.biotechadv.2014.04.001

Yam, K. C., van der Geize, R., and Eltis, L. D. (2010). "Catabolism of aromatic compounds and steroids by *Rhodococcus*," in *Biology of Rhodococcus, Microbiology Monograph,* ed H. M. Alvarez (Heidelberg: Springer Berlin Heidelberg), 133–169. doi: 10.1007/978-3-642-12937-7_6

Ye, Y., Zhang, Y., Fan, J., and Chang, J. (2012). Novel method for production of phenolics by combining lignin extraction with lignin depolymerization in aqueous ethanol. *Ind. Eng. Chem. Res.* 51, 103–110. doi: 10.1021/ie202118d

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.