WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/131954

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

warwick.ac.uk/lib-publications

Functional genomic analysis of bacterial lignin degraders: diversity in mechanisms of lignin oxidation and metabolism

Rommel Santiago Granja-Travez[1#], Gabriela Felix Persinoti[2], Fabio M. Squina[3], and Timothy D.H. Bugg[1*]

1. Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK
2. Laboratório Nacional de Biorrenováveis (LNBR/CNPEM), Campinas, Brazil
3. Programa de Processos Tecnológicos e Ambientais, Universidade de Sorocaba, Sorocaba, Brazil

#Current affiliation: Facultad de Ciencias de la Ingeniería e Industrias, Universidad UTE, Quito, Ecuador

Corresponding author: Prof. Timothy D.H. Bugg, Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom. Email T.D.Bugg@warwick.ac.uk, tel 44-2476-573018.

Abstract:

Although several bacterial lignin-oxidising enzymes have been discovered in recent years, it is not yet clear whether different lignin-degrading bacteria use similar mechanisms for lignin oxidation and degradation of lignin fragments. Genome sequences of thirteen bacterial lignin-oxidising bacteria, including new genome sequences for *Microbacterium phyllosphaerae* and *Agrobacterium sp*., were analysed for the presence of lignin-oxidising enzymes and aromatic degradation gene clusters that could be used to metabolise the products of lignin degradation. Ten bacterial genomes contain DyP-type peroxidases, and ten bacterial strains contain putative multi-copper oxidases (MCOs), both known to have activity for lignin oxidation. Only one strain lacks both MCOs and DyP-type peroxidase genes. Eleven bacterial genomes contain aromatic degradation gene clusters, of which ten contain the central β-ketoadipate pathway, with variable numbers and types of degradation clusters for other aromatic substrates. Hence there appear to be diverse metabolic strategies used for lignin oxidation in bacteria, while the β-ketoadipate pathway appears to be the most common route for aromatic metabolism in lignin-degrading bacteria.

Keywords

Bacterial lignin degradation; genome sequences; aromatic degradation pathways; DyP-type peroxidase; multi-copper oxidase.

Introduction

The aromatic heteropolymer lignin is a major component of the lignocellulose cell wall in plants, comprising 15-40% in dry weight of lignocellulose (Ragauskas et al, 2014). Although lignin is refractory to degradation, due to the presence of ether carbon-oxygen bond linkages, and carbon-carbon bond linkages, some micro-organisms have the ability to break down lignin. White-rot fungi are able to break down lignin via the action of oxidative secreted enzymes, mainly peroxidases and laccases (Camarero et al, 2014; Cragg et al, 2015). Brown-rot fungi can efficiently degrade lignocellulose, but they are able only to modify lignin (Cragg et al, 2015). Several soil bacteria are also reported to degrade lignin, especially members of the phyla Actinobacteria, α-Proteobacteria and γ-Proteobacteria, some of which have also been reported in termite guts and wood-boring insects (Bugg et al, 2011b; Brown & Chang, 2014; Tian et al, 2014; de Gonzalo et al, 2016).

There is considerable current interest in lignin degradation for biotechnology, using either chemocatalysis or biocatalysis (Ragauskas et al, 2014). Metabolic engineering studies have been published using engineered strains of *Rhodococcus jostii* RHA1 and *Pseudomonas putida* KT2440 to produce bioproducts from lignin breakdown, including vanillin (Sainsbury et al, 2013), pyridine-dicarboxylic acids (Mycroft et al, 2015), polyhydroxyalkanoic acids (Linger et al, 2014) and *cis,cis*-muconic acid (Vardon et al, 2016). In order to achieve high product yields via metabolic engineering, a better understanding is needed of the biochemical and genetic basis for lignin oxidation, uptake of lignin fragments, and metabolism of aromatic fragments vis aromatic degradation pathways. Since a number of genomes of bacterial lignin degraders have now been published, in this article we have surveyed the genomes of twelve bacteria reported to break down lignin: Actinobacteria *Rhodococcus jostii* RHA1 (Sainsbury et al, 2013, McLeod et al, 2006), *Microbacterium phyllosphaerae* (Taylor et al, 2012), *Arthrobacter* sp. (Moraes et al, 2018); Firmicutes *Paenibacillus* sp. (Rashid et al, 2017; Granja-Travez et al, 2018) and *Lysinibacillus sphaericus* (Rashid et al, 2017; Persinoti et al, 2018); γ–Proteobacterium *Pseudomonas putida* KT2440 (Linger et al, 2014); β–Proteobacteria *Cupriavidus basilensis* B-8 (Shi et al, 2013), *Comamonas serinivorans* (Zhang et al, 2017) and *Comamonas testosteroni* (Rashid et al, 2017); α–Proteobacteria *Ochrobactrum* sp. (Taylor et al, 2012; Granja-Travez et al, 2018) and *Agrobacterium* sp. (Rashid et al, 2017), and Bacteroides *Sphingobacterium* sp. T2 (Taylor et al, 2012; Rashid et al, 2013). We have also

included the genome of a thirteenth bacterial strain *Burkholderia multivorans* LB400, a well-studied polychlorinated biphenyl degrader (Denef et al, 2004), since there are reports of lignin-degrading *Burkholderia* strains (Akita et al, 2016). The genome sequences for *Microbacterium phyllosphaerae* (accession RAIJ00000000) and *Agrobacterium* sp. (accession VTPB00000000) have been deposited by the authors at DDBJ/ENA/Genbank.

We have analysed for bacterial enzymes reported to oxidise lignin, firstly multi-copper oxidases, such as laccases from *Streptomyces coelicolor* (Majumdar et al, 2014), CueO from *Ochrobactrum* sp (Granja-Travez et al, 2018) and CopA from *Pseudomonas putida* (Granja-Travez & Bugg, 2018), which have been reported to show activity for lignin oxidation. These sequences have been compared with laccases from non-lignin-degrading bacterial strains. Secondly, dye-decolorizing peroxidases, of B-class (Ahmad et al, 2011) and C-class (Brown et al, 2012), since these classes of DyPs have been reported to show activity for lignin oxidation. DyP-type peroxidases have been classified into classes DyPA-D (Ogola et al, 2009), but bacterial MCOs have not been classified, so we have carried out a more detailed bioinformatics analysis for bacterial MCOs. Thirdly, other enzymes reported to show activity towards lignin, such as manganese superoxide dismutase from *Sphingobacterium* sp. T2 (Rashid et al, 2015), and β-etherase enzymes LigDEF found in *Sphingobium* SYK-6 and *Novosphingobium* (Masai et al, 2003; Gall et al, 2014; Ohta et al, 2015). Bioinformatic analysis was performed using CLC workbench software (https://www.qiagenbioinformatics.com/).

The aromatic component of lignin is degraded via aromatic *ortho*- and *meta*-cleavage pathways, which have been reviewed (Bugg et al, 2011a), but it is not clear which of these pathways are used for lignin degradation. From metabolic engineering studies, there appears to be a central role for the vanillate degradation pathway & beta-ketoadipate pathway (Sainsbury et al, 2013; Linger et al, 2014; Mycroft et al, 2015; Vardon et al, 2016). Since some lignin-degrading bacteria contain up to 15 aromatic degradation gene clusters, it is not clear whether other pathways are also used, hence a survey of existing aromatic degradation gene clusters was also carried out.

Multicopper oxidases (MCOs)

Multicopper oxidases (MCOs) and their role in lignin degradation

MCOs are a superfamily of proteins comprising laccases (EC 1.10.3.2), ascorbate oxidase (EC 1.10.3.3), ferroxidase (EC 1.16.3.1), nitrite reductase (EC 1.7.2.1), and ceruloplasmin (EC 1.16.3.1). The biological function of laccases varies: in fungi these enzymes have been reported in almost all wood degraders, where they have been related with

morphogenesis, pathogen-host interaction, stress defence, and lignin degradation; while in plants, laccases seem to be involved in lignin biosynthesis (Bao et al, 1993); in bacteria, laccases appear to be involved in morphogenesis, pigmentation, oxidative protection and copper homeostasis (Sharma et al, 2007), while in insects, they play a role in sclerotization of the cuticle (Giardina et al, 2010). Laccases contain four atoms of copper, ligated to well conserved residues. These residues are arranged in four highly conserved motifs, HXHG, HXH, HXXHXH and HCHXXXHXXXXM/L/F, common in all MCOs (Reiss et al, 2013). However, there is no distinctive sequence belonging solely to laccases that can differentiate them from another subgroup of MCOs, thus the only way to define unambiguously a MCOs as a laccase is by experimental approaches, mostly by confirming its activity towards siringaldezine (SGZ) (Reiss et al, 2013). The ability of laccases to attack, degrade and modify lignin has been well documented in white-rot fungi. Most of the white-rot fungi secrete laccases when grown on lignin, while brown-rot fungi, which are unable to grow on lignin, do not secrete laccases extracellularly (Youn et al, 2001). Laccases are involved in both lignin polymerization and depolymerization, besides other chemical modifications, so the precise role of laccase enzymes in lignin breakdown is yet to be fully understood. (Munk et al, 2015).

Multicopper oxidases (MCOs) found in non-lignin-degrading strains, that have shown laccase activity

Laccase-like genes have been identified in the genome of a large variety of bacterial species, sourced from diverse ecological environments. A search of 2200 complete and draft bacteria genomes and four metagenomic datasets, found more than 1200 putative genes for laccase-like multicopper oxidases, among which 76% contained a predicted signal peptide sequence, suggesting that most of these proteins are exported from the cytoplasm (Ausec et al, 2011). Despite the potential enclosed in these interesting proteins, only few bacterial laccases have been characterized, and most of them have been obtained from bacteria non-related with lignin oxidation activity. Hence, a comparison of the biochemical and sequence properties between laccases from bacteria non-related and related with lignin degradation, could shed some light on how these proteins are similar and whether there is any functional difference among them. With this aim, this section presents a brief description of some multicopper oxidases (MCOs) obtained from bacteria nonrelated with lignin oxidation, that experimentally have shown laccase activity.

Escherichia coli contains two multicopper oxidases that have oxidase activity, named CueO and PcoA, and both are related to copper resistance. CueO is a "blue protein" and shows

phenol oxidase activity, enhanced in presence of additional copper (Roberts et al, 2003). It can oxidize catechols, siderophores, and Fe(II) (Kim et al, 2001). For its part, PcoA also shows oxidase activity and, despite having the conserved motifs associated with MCOs, its spectral properties are not available in literature, thus PcoA is not classified as a "blue protein" in this work. PcoA has been mostly characterized in the context of copper resistance, in association with a smaller PcoC protein which is thought to act in cooperation with PcoA by transferring copper ions to it (Huffman et al, 2002). PcoA resembles the MCO CopA, which also is thought to act in cooperation with a smaller CopC protein. Both, CopA and CopC are forming part of a copper resistance operon found in some bacteria, e.g. *Pseudomonas syringae* (Cha et al, 1993). Additionally, CopA has been identified as a ligninolytic enzyme by a biosensor developed to find novel genes involved in lignin degradation, in metagenomics libraries (Strachan et al, 2013). Interestingly, in bacterial lignin-degraders genomes, when a *copA* putative gene is found, the genome in question often lack the rest of genes associated with the copper resistance operon, including *copC* genes. Conversely, if a putative *copC* genes is found, putative *copA* genes are absent. Another multicopper oxidases that has been experimentally characterized is CotA, a protein associated with endospore pigment formation. CotA from *Bacillus subtilis* (Martins et al, 2002) and CotA from *Bacillus licheniformis* (Koschorreck et al, 2008) are both "blue proteins". These proteins accept a similar range of substrates, including ABTS and SGZ. They have shown also activity towards some phenolic substrates such as sinapic acid, caffeic acid and ferulic acid, yielding dimeric products.

Another strain of *Bacillus* genus, *Bacillus halodurans*, a strain adapted to alkaline conditions containing unique genes when compared with *Bacillus subtilis* (Takami et al, 1999), contains a laccase-like gene that has been cloned and expressed in *E. coli*. The resulting protein showed laccase activity towards ABTS and SGZ, moreover, its activity was enhanced in presence of chloride, distinguishing it from conventional laccases (Ruijssenaars et al, 2004). Other bacterial laccase enzymes with unusual properties that have been characterized include a MCO found in *Thermus thermophilus* (Miyazaki, 2005), a "blue protein" that was reported as the most thermophilic laccase. A polyphenol oxidase (PPO) from *Streptomyces lavendulae* REN-7 was also reported as a laccase with high thermal stability (Suzuki et al, 2003); and an MCO found in *Marinomonas mediterranea* showed both laccase activity and a capacity to catalyse tyrosine hydroxylation (Sanchez-Amat et al, 2001). These protein sequences were used to perform a multiple sequence alignment with sequences of laccase-like multicopper oxidases found in lignin-degrading bacterial genomes (see Figure 1).

Multi-copper oxidases (MCOs) found in bacteria related with lignin degradation

As there is no distinctive sequence motif able to distinguish laccases with lignin oxidation activity from the wider group of MCOs, some sequences with known or suspected lignin oxidase activity were selected to perform a sequence alignment, along with sequences of laccases contained in non-lignin-degrading bacteria. Firstly, the sequences of Oc-CueO from *Ochrobactrum sp*. (Granja-Travez et al, 2018), CopA-I from *P. putida* (Granja-Travez & Bugg, 2018), and the *R. jostii* RHA1 McoA ro02377 were used to perform an initial alignment, whose consensus sequence was used to perform a BLAST search against the genomes of the thirteen bacterial strains listed above, in order to identify MCOs with potential activity towards lignin. The sequences identified are shown in Table 1A. The retrieved sequences were used to perform a bigger alignment, including the MCO sequences obtained from bacteria not related with lignin oxidation (Table 1B). The resulting sequence alignment is shown in Figure 1, while its corresponding phylogenetic tree is shown in Figure 2.

The phylogenetic tree results suggest the presence of four different groups among these bacterial strains. The first group (Group A) are related to CueO and CotA type proteins, and are therefore likely to be blue laccase enzymes (Granja-Travez et al, 2018). Recombinant Oc-CueO from *Ochrobactrum* sp. has been shown to be active for oxidation of lignin model compounds and lignosulfonate (Granja-Travez et al, 2018; Martins et al, 2002; Koschorreck et al, 2008). Other sequences from bacterial lignin degraders include *Lysinibacillus* Mco1343, *R. jostii* McoC, and *P. xenovorans* Mco49782249. Another possible group of "blue laccases" has been named as A.1 group, and it contains the thermophilic laccase from *Thermus thermophilus,* and the multipotent laccase from *Marinomonas mediterranea*.

A second group, Group B, contains four proteins from bacterial lignin degraders, and one from bacteria not related with lignin degradation (*Bacillus halodurans* laccase), the latter having been characterized experimentally, showing alkaline tolerance and enhancement of activity in presence of chloride (Ruijssenaars et al, 2004). Group C include proteins related with copper homeostasis, such as PcoA and CopA proteins, which are likely to be colourless pseudo-laccases (Granja-Travez & Bugg, 2018). Three proteins from bacterial lignin degraders, and one (*E. coli* PcoA) from bacteria not related with lignin degradation are listed in this group. The fourth and last group (Group D) are a distinct group of six uncharacterised MCOs found in α- and β-Proteobacteria, and we note that there are no sequences of laccases from bacteria not related with lignin degradation listed in this group, although these proteins lack the presence of conserved residues found in fungal and bacterial laccases, that may be important for laccase activity (Granja-Travez et al, 2018).

Table 1, Figure 1, Figure 2 here.

DyP-type peroxidases

Dye-decolorizing peroxidases (DyP) and their role in lignin degradation

DyP peroxidases are heme-containing enzymes, that were first identified in fungi as a dye degrading enzyme (Kim at al., 1999). All DyP-type peroxidases contain a well-conserved motif GXXDG in their primary sequences (Sugano et al, 2007). Dyp-type peroxidases are further classified into A, B C and D subgroups based on their phylogenetic differences (Ogola et al, 2009). Several DyP-type peroxidases have been reported to oxidise lignin model compounds or polymeric lignin: DypB from *Rhodococcus jostii* RHA1 (Ahmad et al, 2011), Dyp1B from *Pseudomonas fluorescens* Pf-5 (Rahmanpour and Bugg, 2015), and DyP from *Irpex lacteus* showed lignolytic activity (Salvachúa et al, 2013). C-type DyP peroxidases are reported to show high Mn^{2+} oxidation activity, comparable with fungal MnP enzymes (Brown et al, 2012). DyPA enzymes contain a Tat signal sequence for protein export, consequently are likely to be extracellular proteins. DyPB and DypC enzymes do not contain a Tat signal sequence, hence appear to be intracellular enzymes, although DypB from *Rhodococcus jostii* RHA1 and some other DypB enzymes are targeted towards an encapsulin protein nanocompartment (Rahmanpour and Bugg, 2013). The subgroup DyPD include mostly fungal variants of Dyp (Colpa et al., 2014).

DyP-type peroxidases found in non-lignin degrading strains

According to PeroxiBase database (http://peroxibase.toulouse.inra.fr/), thus far only 237 sequences are forming part of the DyP-type peroxidase superfamily group, comprising 26 DyPA proteins, 49 DyPB proteins, 24 DyPC proteins, and 138 fungal DyPD proteins. In this work we have focused our analysis only on DyP-type proteins of A, B and C subgroups. The distribution of each subgroup in the different class of organism is shown in Figure 3. Despite the relative low number of sequences available, some interesting tendencies can be inferred from this data. For instance, all DyPA peroxidases belong to bacterial strains; whereas DyPB enzymes seems to be equally distributed between the bacterial and eukaryote domains. Most peroxidases from the DyPC subgroup are found in bacterial strains, though some proteins have been found in the fungal basidiomycota phylum. Moreover, it is clear that DyP-type peroxidases are mostly found in Actinobacteria and γ-Proteobacteria, which include bacteria that have been related with lignin degradation. A list of bacterial DyP-type peroxidases from

non-lignin degrading strains is shown in Table 2.B. Bacterial DyPA proteins have been characterised from *Escherichia coli* (Sturm et al., 2006), and *Bacillus subtilis* (Jongbloed et al., 2004). A comparison between DyPA from *B. subtilis* and DyPB from *P. putida* revealed that though DyPA is more thermostable, DyPB was more active and accepted a broader range of substrates (Santos et al., 2014). DyPA proteins contain a Tat-signal sequence for protein exportation which is not present in neither DyPB nor DyPC proteins. However, as noted above, a tight association has been reported between encapsulating and DyPB genes (Dailey et al., 2011; Rahmanpour & Bugg, 2013), which indicates a possible extracellular function for this subgroup of DyP peroxidases. The list of bacterial species containing DyPB and DyPC reported by PeroxiBase database include several species related with lignin degradation (Supporting information Tables S1 and S2).

Figure 3 here

Bacterial DyP-type peroxidases found in lignin degrading strains

A search for DyP-type peroxidases in the genome of lignin degrading strains was performed. Firstly, a total of 10 sequences for each subgroup of DyP enzymes (A, B and C) were selected from the PeroxiBase database. Bacterial strains and Uniprot accession codes used are shown in Supporting information Table S3. Then, the resulting consensus sequences of each subgroup DyPA, DyPB and DyPC, were used to carry out a BLAST (Basic Local Alignment Search Tool) against the thirteen bacterial lignin degrading genomes.

Annotated Dyp-type peroxidase proteins were identified in the genomes of ten out of the thirteen lignin-oxidising bacterial strains, as shown in Table 2. DyPB peroxidases, reported to show activity for lignin oxidation (Ahmad et al, 2011) are found in *R. jostii*, *P. putida*, *C. testosteroni*, *C. serinivorans*, and *P. xenovorans*. C-type DyP peroxidases, reported to show high Mn^{2+} oxidation activity (Brown et al, 2012), are found in *Ochrobactrum* sp., and *Agrobacterium* sp. DyPA peroxidases are found in *Ochrobactrum* sp., *R. jostii*, *Lysinibacillus sp.*, *P. xenovorans*, *M. phyllosphaerae* and *Paenibacillus sp*.

In order to compare DyP-type enzymes from bacterial lignin degrading and non-lignin degrading strains, a multiple sequence alignment of DyP-type peroxidases was done using sequences from both type of strains (Supporting Information Figure S1). A phylogenetic tree was generated from this alignment (Figure 4). As expected, three main subgroups are observed corresponding with the already described subgroups for DyP-type peroxidases: A, B and C. Sequences from each subgroup are found in genomes of bacterial lignin degraders, with

DypA most common (7 examples), followed by DypB (5 examples), then DypC (2 examples), but there are also sequences from non-lignin-degrading bacteria in each subgroup, suggesting that these enzymes have multiple functions in different bacteria.

Lignin oxidation activity has been reported for B- and C-type dye-decolorizing peroxidase enzymes in bacteria (Ahmad et al, 2011; Brown et al, 2012), furthermore, up-regulation of *dyp* genes in response to lignin has been reported in *Pseudomonas putida* A514 (Lin et al, 2016; Lin et al, 2019). In our set of 13 bacteria, ten contained DyP-type peroxidase genes, out of which 5 organisms contained B-type DyPs, and 3 contained C-type DyPs. Roles for bacterial multi-copper oxidases in lignin oxidation have also been reported (Majumdar et al, 2014; Granja-Travez et al, 2018; Granja-Travez & Bugg, 2018), and multi-copper oxidase genes are found in ten of the bacterial genomes. There appear to be four groups of bacterial multi-copper oxidase sequences, of which groups A and C contain enzymes shown to oxidise lignin (see Figure 2), but groups A-C also contain MCOs from bacteria that do not degrade lignin, suggesting that MCOs perform multiple roles in different organisms. Eight bacteria contain both DyPs and MCO genes, with the remaining bacteria containing either only DyP genes (*M. phyllosphaerae, Paenibacillus* sp,), only MCO genes (*Arthrobacter* sp., *C. basilensis*), or neither (*Sphingobacterium* sp.). Hence an important conclusion is that *there is diversity in the lignin oxidation mechanisms used in different lignin-degrading bacteria*, there is not one class of enzyme found in all organisms. Diversity in lignin oxidation gene is also observed in wood-degrading Basidiomycete fungi, for example, *Phanerochaete chrysosporium* contains 15 lignin peroxidase genes but no laccase genes, whereas *Stereum hirsutum* contains 15 laccase genes but only 5 lignin peroxidase genes, and other white-rot fungi contain variable numbers of each class (Riley et al, 2014).

Table 2, Figure 4 here.

Other lignin degradation genes

Sphingobacterium sp. T2 is unusual in utilising two specialised extracellular manganese superoxide dismutase enzymes for lignin oxidation (Rashid et al, 2015). This strain does not contain any putative MCO or DyP peroxidase genes. *Agrobacterium* sp. contains one LigE β-etherase gene, related to β-etherase LigE from *Sphingobium* SYK-6 (Masai et al, 2003), although no other annotated β-etherase genes were found in this genome. The β-etherase family of enzymes are a specialised type of glutathione *S*-transferase, hence we note that the bacterial genomes surveyed contained variable numbers of glutathione *S*-transferase (GST)

genes. While most genomes contained 1-2 GST genes, *P. putida* KT2440 contains 14 GST genes, *Ochrobactrum* sp. contains 9, and *Agrobacterium* sp. 8, so there is a possibility that some of these additional GST genes might encode additional β-etherase enzyme activities.

Since the genome of *Sphingobacterium* sp. T2 does not contain any aromatic degradation gene clusters, but does contain many genes for hemi-cellulose breakdown, it seems possible that this bacterium is able to remove lignin in order to attack the polysaccharide component of lignocellulose, in a similar fashion to brown-rot fungi (Kerem et al, 1999). *Agrobacterium* sp. also contains one *ligE* β-etherase gene, but does not appear to contain other *lig* genes reported in *Sphingobium* SYK-6 (Masai et al., 2003).

Aromatic degradation pathways

We have surveyed the twelve bacterial genomes for the presence of aromatic degradation gene clusters, which are listed in Table 3. Nine of the genomes contain the β-ketoadipate pathway (*pca* genes, Harwood & Parales, 1996) for conversion of protocatechuic acid to β-ketoadipate, which is thought to be a central pathway for degradation of lignin fragments (Sainsbury et al, 2013, Linger et al, 2014), though in *Cupriavidus basilensis* B-8 the genes for this pathway are not clustered (Shi et al, 2013). Five genomes (*R. jostii, P. putida, Cupriavidus basilensis, Comamonas testosteroni, Paraburkholderia xenovorans*) also contain the catechol *ortho*-cleavage pathway, which also leads to β-ketoadipate. Five bacteria contain the 4-hydroxyphenylacetate (or homoprotocatechuate) pathway (*hpc* genes, Jenkins & Cooper, 1988), and *P. putida* KT2440 also contains the *pha* gene cluster for phenylacetic acid breakdown. Since 4-hydroxyphenyl units are present in grass lignins, it is possible that the *hpc* pathway could be used for degradation of lignin fragments, but there is currently no experimental evidence for that hypothesis.

Biphenyl degradation clusters (*bph* genes, Ohtsubo et al, 2000) are present in *R. jostii* RHA1 and *P. multivorans,* both PCB degraders, and also in *Arthrobacter* sp. Since biphenyl units are found in lignin, the *bph* pathway could potentially be used to degrade biphenyl units found in lignin fragments, but only 3 out of 12 organisms contain this pathway. Five bacteria (*R. jostii, Arthrobacter sp., C. basilensis, C. testosteroni, and P. xenovorans*) contain the gentisate degradation pathway (Zhou et al, 2001), while *R. jostii* RHA1 and *P. putida* KT2440 also contain the homogentisate degradation pathway (Arias-Barrau et al, 2004). Three bacteria (*R. jostii, Arthrobacter sp., P. multivorans*) contain the phenylpropionic acid degradation pathway (Burlingame & Chapman, 1983); although phenylpropionic acid contains the aryl C_3 skeleton found in lignin fragments, the catecholic intermediate on the pathway is 2,3-

dihydroxyphenylpropionic acid, whose substitution pattern does not match that found in lignin.

Other pathways include the phthalic acid degradation pathways (3 bacteria, Chang & Zylstra, 1998), L-tryptophan degradation (3 bacteria, Colabroy & Begley, 2005), hydroxyquinol degradation pathway (2 bacteria, Armengaud et al, 1999), benzoic acid degradation pathway (4 bacteria, Harayama et al, 1987), phenol oxidation (5 bacteria, Nordlund et al, 1993), salicylate degradation (1 bacterium, Dennis & Zylstra, 2004), p-cumate degradation (1 bacterium, Eaton, 1996). Three bacteria contain *mhq* genes, which are believed to be involved in hydroquinone degradation (Tago et al, 2005).

Table 3 here

Metabolism of the low molecular weight aromatic products of lignin breakdown occurs via aromatic degradation pathways that are well-studied in soil bacteria (Bugg et al, 2011a). Several pieces of evidence implicate the β-ketoadipate pathway from protocatechuic acid to β-ketoadipate, preceded by the conversion of vanillin and vanillic acid to protocatechuic acid, as important pathways for lignin fragment degradation: 1) lignin-degrading bacteria often generate vanillic acid as a metabolite, and can usually grow on vanillic acid as a sole carbon source (Taylor et al, 2012); 2) gene deletion of the *vdh* gene encoding vanillate dehydrogenase in *R. jostii* RHA1 gives a mutant strain that accumulates vanillin when grown on minimal media containing lignin or lignocellulose (Sainsbury et al, 2013); 3) metabolic engineering studies in *P. putida* KT2440 have generated other bioproducts from engineering of the β-ketoadipate pathway (Linger et al, 2014; Vardon et al, 2016); 4) the *pca* genes in *P. putida* A514 are upregulated in response to lignin (Lin et al, 2016). Of the 12 bacteria examined, 9 contain the *pca* gene cluster encoding the β-ketoadipate pathway, which supports a central role for the β-ketoadipate pathway in degradation of lignin fragments. Bacteria lacking the the β-ketoadipate pathway are *Sphingobacterium* sp. T2 discussed above, and the two Firmicutes *Paenibacillus* sp. and *Lysinibacillus* sp., implying that the latter bacteria either do not fully metabolise the aromatic lignin fragments, or use other pathways to do so. Furthermore, 6 bacteria contain genes for vanillin oxidation (*vdh* gene) or vanillic acid demethylation (*vanAB*), which have been characterised in *R. jostii* RHA1 (Chen et al, 2012), and which generate protocatechuic acid, at the start of the β-ketoadipate pathway.

Ferulic acid and *p*-coumaric acid are also found in grass lignin structures, and are generated as breakdown products, and are converted in some bacteria such as *Pseudomonas putida* to

vanillin via the *fcs* and *ech* genes, or in other bacteria such as *Rhodococcus jostii* RHA1 to vanillic acid via the *couLMN* genes (Otani et al, 2014). Upregulation of the ferulate degradation genes in response to lignin has been observed in *P. putida* A514 (Lin et al, 2016). In the 12 genomes examined, 4 contained genes for ferulic acid degradation: *P. putida* and *C. basilensis* contained genes for conversion to vanillin, while *R. jostii* RHA1 and *Ochrobactrum* sp. contained genes for conversion to vanillic acid.

Gene clusters encoding several other aromatic degradation pathways were observed, but with variable frequency, the most common being the *hpc* cluster for breakdown of 4-hydroxyphenylacetic acid, whose structure could potentially be derived from H units in lignin, and was observed in 6 bacteria. The *hmg* cluster encoding the homogentisate pathway in *P. putida* A514 was reported to be upregulated in response to lignin (Lin et al, 2016), but this cluster is found in only 2 of the bacteria surveyed. Potentially some of these pathways could be linked to the degradation of lignin fragments, as shown in Figure 5, but there is currently little experimental evidence linking these pathways to lignin degradation.

Figure 5 here

In conclusion, analysis of genome sequences for bacterial lignin degraders, whose enzymology has emerged relatively recently, reveals diversity in mechanisms of lignin oxidation, as found in lignin-degrading fungi. Metabolic routes for aromatic metabolism of lignin fragments indicates that the β-ketoadipate pathway is a common pathway for aromatic metabolic, but that other aromatic degradation pathways may potentially contribute.

Acknowledgements

This work was supported by research grants from BBSRC (research grant BB/P01738X/1) and FAPESP (research grant 2015/50590-4), and a PhD studentship (to RSGT) funded by Secretaria de Educacion Superior, Ciencia, Tecnologia e Innovacion (SENESCYT), Ecuador. We thank Brazilian Biorenewable National Laboratory LNBR NGS Sequencing Facility for generating the DNA sequencing of *Agrobacterium* sp. strain.

Compliance with Ethical Standards

Funding: this study was funded by research grants from BBSRC (research grant BB/P01738X/1) and FAPESP (research grant 2015/50590-4), and a PhD studentship (to

RSGT) funded by Secretaria de Educacion Superior, Ciencia, Tecnologia e Innovacion (SENESCYT), Ecuador.

Conflict of interest: the authors declare that they have no conflict of interest for this publication.

Ethical approval: This article does not contain any studies with human participants oir animals performed by any of the authors.

References

Ahmad M, Roberts JN, Hardiman EM, Singh R, Eltis LD, Bugg TDH (2011) Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase. Biochemistry 50:5096-5107.

Akita H, Kimura ZI, Mohd Yusoff MZ, Nakashima N, Hoshino T (2016) Isolation and characterisation of *Burkholderia* sp. strain CCA53 exhibiting lignolytic potential. SpringerPlus 5:article 596.

Arias-Barrau E, Olivera ER, Luengo JM, Fernández C, Galán B, Garcia JL, Diaz E, Miñambres B (2004). The homogentisate pathway: a central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphewnylacetate in *Pseudomonas putida*. 186:5062-5077.

Armengaud J, Timmis KN, Wittich RM (1999). A functional 4-hydroxysalicylate/ hydroxyquinol degradative pathway gene cluster is linked to the initial dibenzo-*p*-dioxin pathway genes in *Sphingomonas* sp. strain RW1. J. Bacteriol. 181:3452-3461.

Ausec L, Zakrzewski M, Goesmann A, Schlüter A, Mandic-Mulec, I (2011). Bioinformatic analysis reveals high diversity of bacterial genes for laccase-like enzymes. PLoS One, 6:e25724.

Bao W, O'Malley DM, Whetten R, Sederoff RR (1993). A laccase associated with lignification in loblolly pine xylem. Science, 260:672-674.

Brown ME, Barros T, Chang MCY (2012) Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium. ACS Chem. Biol., 7:2074-2081.

Brown ME, Chang MC (2014). Exploring bacterial lignin degradation. Curr. Opin. Chem. Biol*.* 19:1-7.

Bugg TDH, Ahmad M, Hardiman EM, Rahmanpour R (2011). Pathways for degradation of lignin in bacteria and fungi. Nat. Prod. Rep*.* 28:1883-1896.

Bugg TDH, Ahmad M, Hardiman EM, Singh R (2011). The emerging role for bacteria in lignin degradation and bio-product formation. Curr. Opin. Biotech. 22:394-400.

Burlingame R, Chapman PJ (1983). Catabolism of phenylpropionic acid and its 3-hydroxy derivative by *Escherichia coli*. J. Bacteriol. 155:113-121.

Camarero S, Martínez MJ, Martínez AT (2014). Understanding lignin biodegradation for the improved utilization of plant biomass in modern biorefineries. Biofuels, Bioproducts and Biorefining*.* 8:615-625.

Chang HK, Zylstra GJ (1998). Novel organization of the genes for phthalate degradation from *Burkholderia cepacia* DB01. J. Bacteriol. 180:6529-6537.

Chen HP, Chow M, Liu CC, Lau A, Liu J, Eltis LD (2012). Vanillin catabolism in *Rhodococcus jostii* RHA1. Appl. Environ. Microbiol. 78:586-588.

Cha JS, Cooksey DA (1993). Copper hypersensitivity and uptake in *Pseudomonas syringae* containing cloned components of the copper resistance operon. Appl. Environ. Microbiol., 59:1671-1674.

Colabroy KL, Begley TP (2005). Tryptophan catabolism: identification and characterization of a new degradative pathway. J. Bacteriol. 187:7866-7869.

Cragg SM, Beckham GT, Bruce NC, Bugg TDH, Distel DL, Dupree P, Etxabe AG, Goodell BS, Jellison J, McGeehan JE (2015). Lignocellulose degradation mechanisms across the Tree of Life. Curr. Opin. Chem. Biol. 29:108-119.

Denef VJ, Park J, Tsoi TV, Rouillard JM, Zhang H, Wibbenmeyer JA, Verstraete W, Gulari E, Hashsham SA, Tiedje JM (2004). Biphenyl and benzoate metabolism in a genomic context: outlining genome-wide metabolic networks in *Burkholderia xenovorans* LB400. Appl. Environ. Microbiol. 70:4961-4970.

Dailey HA, Septer AN, Daugherty L, Thames D, Gerdes S, Stabb EV, Dunn AK, Dailey TA, Phillips JD (2011). The *Escherichia coli* protein YfeX functions as a porphyrinogen oxidase, not a heme dechelatase. MBio, 2:e00248-11.

Dennis JJ, Zylstra GJ (2004). Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NCIB 9816-4. J. Mol. Biol. 341:753-768.

Eaton RW (1996). p-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the *cmt* operon. J. Bacteriol. 178:1351-1362.

Gall DL, Kim H, Lu F, Donohue TJ, Noguera DR, Ralph J (2014). Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-aryl etherase pathway. J. Biol. Chem. 289:8656-8667.

Giardina P, Faraco V, Pezzella C, Piscitelli A, Vanhulle S, Sannia G (2010). Laccases: a never-ending story. Cell. Mol. Life Sciences, 67:369-385.

de Gonzalo G, Colpa DI, Habib MH, Fraaije MW (2016). Bacterial enzymes involved in lignin degradation. J. Biotechnol. 236:110-119.

Granja-Travez RS, Bugg TDH (2018) Characterisation of multi-copper oxidase CopA from *Pseudomonas putida* KT2440 and *Pseudomonas fluorescens* Pf-5: involvement in bacterial lignin oxidation. Arch. Biochem. Biophys. 660:97-107.

Granja-Travez RS, Wilkinson RC, Persinoti GF, Squina FM, Fülöp V, Bugg TDH (2018). Structural and functional characterisation of multi-copper oxidase CueO from lignin-

degrading bacterium *Ochrobactrum* sp. reveal its activity towards lignin model compounds and lignosulfonate. FEBS J. 285:1684-1700.

Harayama S, Mermod N, Rekik M, Lehrbach PR, Timmis KN (1987). Roles of the divergent branches of the meta-cleavage pathway in the degradation of benzoate and substituted benzoates. J. Bacteriol. 169:558-564.

Harwood CS, Parales RE (1996). The beta-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. 50:553-590.

Huffman D L, Huyett J, Outten FW, Doan PE, Finney LA, Hoffman BM, O'Halloran TV (2002). Spectroscopy of Cu (II)-PcoC and the multicopper oxidase function of PcoA, two essential components of *Escherichia coli pco* copper resistance operon. Biochemistry, 41:10046-10055.

Jenkins JR, Cooper RA (1988). Molecular cloning, expression and analysis of the genes of the hoimoprotocatechuate catabolic pathway of *Escherichia coli* C. J. Bacteriol. 170:5317-5324.

Kerem Z, Jensen KA, Hammel KE (1999). Biodegradative mechanism of the brown rot basidiomycete *Gloeophyllum trabeum*: evidence for an extracellular hydroquinone-driven Fenton reaction. FEBS Lett. 446:49-54.

Kim C, Lorenz WW, Hoopes JT, Dean JF (2001). Oxidation of phenolate siderophores by the multicopper oxidase encoded by the *Escherichia coli yacK* gene. J. Bacteriol., 183:4866-4875.

Koschorreck K, Richter SM, Ene AB, Roduner E, Schmid RD, Urlacher VB (2008). Cloning and characterization of a new laccase from *Bacillus licheniformis* catalyzing dimerization of phenolic acids. Appl. Microbiol. Biotechnol. 79:217-224.

Lin L, Cheng Y, Pu Y, Sun S, Li X, Jin M, Pierson EA, Gross DC, Dale BE, Dai SY, Ragauskas AJ, Yuan JS (2016). Systems biology-guided biodesign of consolidated lignin conversion. Green Chemistry 18:5536-5547.

Lin L, Wang X, Cao L, Xu M (2019). Lignin catabolic pathways reveal unique characteristics of dye-decolorizing peroxidases in *Pseudomonas putida*. Environ. Microbiol. 21:1847-1863.

Linger JG, Vardon DR, Guarneri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT, Beckham GT (2014). Lignin valorization through integrated biological funnelling and chemical catalysis. Proc. Natl. Acad. Sci. USA, 111:12013-12018.

Majumdar S, Lukk T, Solbiati JO, Bauer S, Nair SK, Cronan JE, Gerlt JA (2014). Roles of small laccases from *Streptomyces* in lignin degradation. Biochemistry 53:4047-4058.

Masai E, Ichimura A, Sato Y, Miyauchi K, Katayama Y, Fukuda M (2003). Roles of the enantioselective glutathione *S*-transferases in cleavage of β-aryl ether. J. Bacteriol. 185:1768 – 1775

Martins LO, Soares CM, Pereira MM, Teixeira M, Costa T, Jones G, Henriques AO (2002). Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the *Bacillus subtilis* endospore coat. J. Biol. Chem. 277:18849-18859.

McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006). The complete genome sequence of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc. Natl. Acad. Sci. USA 103:15582-15587.

Miyazaki K (2005). A hyperthermophilic laccase from *Thermus thermophilus* HB27. Extremophiles 9:415-425.

Moraes EC, Alvarez TM, Persinoti GF, Tomazetto G, Brenelli LB, Paixão DAA, Ematsu GC, Aricetti JA, Caldana C, Dixon N, Bugg TDH, Squina FM (2018). Lignolytic-consortium omics analyses reveal novel genomes and pathways involved in lignin modification and valorization. Biotechnol. Biofuels 11: article 75.

Munk L, Sitarz AK, Kalyani DC, Mikkelsen JD, Meyer AS (2015). Can laccases catalyze bond cleavage in lignin? Biotechnol. Adv. 33:13-24.

Mycroft Z, Gomis M, Mines P, Law P, Bugg TDH (2015). Biocatalytic conversion of lignin to aromatic dicarboxylic acids in *Rhodococcus jostii* RHA1 by re-routing aromatic degradation pathways. Green Chem. 17:4974-4979.

Ogola HJO, Kamiike T, Hashimoto N, Ashida H, Ishikawa T, Shibata H, Sawa Y (2009). Molecular characterization of a novel peroxidase from the cyanobacterium Anabaena sp. strain PCC7120. Appl. Environ. Microbiol. 75:7509-7518.

Ohta Y, Nishi S, Hasegawa R, Hatada Y (2015). Combination of six enzymes of a marine *Novosphingobium* converts the stereoisomers of β-*O*-4 lignin model dimers into the respective monomers. Sci Reports 5:article 15105.

Ohtsubo Y, Nagata Y, Kimbara K, Takagi M, Ohta A (2000), Expression of the *bph* genes involved in biphenyl/PCB degradation in *Pseudomonas* sp. KKS102 induced by the biphenyl degradation intermediate, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid. Gene 256:223-228.

Otani H, Lee YE, Casabon I, Eltis LD (2014). Characterization of *p*-hydroxycinnamate catabolism in a soil actinobacterium. J. Bacteriol. 196:4293-4303.

Nordlund I, Powlowski J, Hagström A, Shingler V (1993). Conbservation of structural and regulatory genes for a multi-component phenol hydroxylase within phenol-catabolizing bacteria that utilize a meta-cleavage pathway. Microbiology 139:2695-2703.

Persinoti GF, Paixão DAA, Bugg TDH, Squina FM (2018). Genome sequence of Lysinibacillus sphaericus, a lignin-degrading bacterium isolated from municipal solid waste soil. Genome Announcements 6:e00353-18.

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, Wyman CE (2014). Lignin valorization: improving lignin processing in the biorefinery. Science 344: 1246843.

Rahmanpour R, Bugg TDH (2013). Assembly in vitro of *Rhodococcus jostii* RHA 1 encapsulin and peroxidase DypB to form a nanocompartment. FEBS Journal 280:2097-2104.

Rahmanpour R, Bugg TDH (2015). Characterisation of Dyp-type peroxidases from *Pseudomonas fluorescens* Pf-5: oxidation of Mn(II) and polymeric lignin by Dyp1B. Arch. Biochem. Biophys. 574:93-98.

Rashid GMM, Taylor CR, Liu Y, Zhang X, Rea D, Fülöp V, Bugg TDH (2015). Identification of manganese superoxide dismutase from *Sphingobacterium* sp. T2 as a novel bacterial enzyme for lignin oxidation. ACS Chem. Biol. 10:2286-2294.

Rashid GMM, Durán-Peña MJ, Rahmanpour R, Sapsford D, Bugg TDH (2017). Delignification and enhanced gas release from soil containing lignocellulose by treatment with bacterial lignin degraders. J. Appl. Microbiol. 123:159-171.

Reiss R, Ihssen J, Richter M, Eichhorn E, Schilling B, Thöny-Meyer L (2013). Laccase versus laccase-like multi-copper oxidase: a comparative study of similar enzymes with diverse substrate spectra. PloS one, 8:e65633.

Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otillar R, Lindquist EA, Sun H, LaButti KM, Scmutz J, Jabbour D, Luo H, Baker SE, Pisabarro AG, Walton JD, Blanchette RA, Henrissat B, Martin F, Cullen D, Hibbett DS, Grigoriev IV (2014). Extensive sampling of basidiomycete genomes demonstrated inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. USA 111:9923-9928.

Roberts SA, Wildner GF, Grass G, Weichsel A, Ambrus A, Rensing C, Montfort WR (2003). A labile regulatory copper ion lies near the T1 copper site in the multicopper oxidase CueO. J. Biol. Chem. 278:31958-31963.

Ruijssenaars HJ, Hartmans S (2004). A cloned *Bacillus halodurans* multicopper oxidase exhibiting alkaline laccase activity. Appl. Microbiol. Biotechnol. 65:177-182.

Sainsbury PD, Hardiman EM, Ahmad M, Otani, H, Seghezzi N, Eltis LD, Bugg TDH (2013). Breaking down lignin to high-value chemicals: the conversion of lignocellulose to vanillin in a gene deletion mutant of *Rhodococcus jostii* RHA1. ACS Chem. Biol. 8:2151-2156.

Sanchez-Amat A, Lucas-Elío P, Fernández E, García-Borrón JC, Solano F (2001). Molecular cloning and functional characterization of a unique multipotent polyphenol oxidase from *Marinomonas mediterranea*. Biochim. Biophys. Acta 1547:104-116.

Santos A, Mendes S, Brissos V, Martins LO (2014). New dye-decolorizing peroxidases from *Bacillus subtilis* and *Pseudomonas putida* MET94: towards biotechnological applications. Appl. Microbiol. Biotechnol. 98:2053-2065.

Salvachúa D, Prieto A, Martínez ÁT, Martínez MJ (2013). Characterization of a novel dye-decolorizing peroxidase (DyP)-type enzyme from *Irpex lacteus* and its application in enzymatic hydrolysis of wheat straw. Appl. Environ. Microbiol. 79:4316-4324.

Sharma P, Goel R, Capalash N. (2007). Bacterial laccases. World J. Microbiol. Biotechnol. 23:823-832.

Shi Y, Chai L, Tang C, Yang Z, Zhang H, Chen R, Chen Y, Zheng Y (2013) Characterization and genomic analysis of kraft lignin biodegradation by the beta-proteobacterium *Cupriavidus basilensis* B-8. Biotechnol Biofuels 6:article 1.

Strachan CR, Singh R, Van Insberghe D, Ievdokymenko K, Budwill K, Mohn WW, Eltis LD, Hallam SJ (2014). Metagenomic scaffolds enable combinatorial lignin transformation. Proc. Natl. Acad. Sci. USA 111:10143-10148.

Suzuki T, Endo K, Ito M, Tsujibo H, Miyamoto K, Inamori Y (2003). A thermostable laccase from *Streptomyces lavendulae* REN-7: purification, characterization, nucleotide sequence, and expression. Biosci. Biotech. Biochem. 67:2167-2175.

Tago K, Sato J, Takesa H, Kawagishi H, Hayatsu M (2005). Characterization of methylhydroquinone-metabolizing oxygenase genes encoded on plasmid in *Burkholderia* sp. NF100. J. Biosci. Bioeng. 100:517-523.

Takami H, Horikoshi K (1999). Re-identification of facultatively alkaliphilic *Bacillus* sp. C-125 to *Bacillus halodurans*. Biosci. Biotech. Biochem, 63:943-945.

Taylor CR, Hardiman EM, Ahmad M, Sainsbury PD, Norris PR, Bugg TDH (2012). Isolation of bacterial strains able to metabolize lignin from screening of environmental samples. J. Appl. Microbiol. 113:521-530.

Tian JH, Pourcher AM, Bouchez T, Gelhaye E, Peu P (2014). Occurrence of lignin degradation genotypes and phenotypes among prokaryotes. Appl. Microbiol. Biotechnol. 98:9527-9544.

Vardon DR, Franden MA, Johnson CW, Karp EM, Guarneri MT, Linger JG, Salm MJ, Strathman TJ, Beckham GT (2015). Adipic acid production from lignin. Energy Environ. Sci. 8:617-628.

Yoshida T, Sugano Y (2015). A structural and functional perspective of DyP-type peroxidase family. Arch. Biochem. Biophys. 574:49-55.

Zhang PP, Sun JZ, Xie CX, Zhu DC (2017). Lignin degradation by *Comamonas serinivorans* C35. Microbiol China, 44:1131-1137.

Youn HD, Hah YC, Kang SO (1995). Role of laccase in lignin degradation by white-rot fungi. FEMS Microbiol. Lett. 132:183-188.

Zhou NY, Fuenmayor SL, Williams PA (2001). *Nag* genes of *Ralstonia* (formerly *Pseudomonas*) sp. strain U2 encoding enzymes for gentisate catabolism. J. Bacteriol. 183:700-708.

Table 1. Bacterial multi-copper oxidase enzymes found in genomes of (A) lignin degrading strains and (B) non-lignin degrading strains. Key: # Enzymes reported to have lignin oxidation activity; * Annotated as "outer membrane-bounded periplasmic space" by Uniprot database; ** Annotated as "Capsid protein, imported" by Uniprot database. NA, not available.

Organism genome	**Name**	**UniProt code**	**Signal peptide? SignalP-5.0 server**	**Sequence Group**
A. Lignin-degraders				
Ochrobactrum sp. (MYb71)				
	CueO#	A0A2S8TY82	Yes (Tat/SPI)	A
Pseudomonas putida KT2440	CopA-I	Q88KT4	Yes (Tat/SPI)	C
	CopA-II#	Q88C03	Yes (Tat/SPI)	C
Rhodococcus jostii RHA1	McoA	Q0SE54	Yes (Tat/SPI)	B
	McoB	Q0RV38	Yes (Tat/SPI)	B
	McoC 50787466	Q0SGE3 (partial)	Yes (Tat/SPI)	A
Lysinibacillus sp.	Mco1343	A0A2S5D0F4	Yes (Sec/SPII)	A
Arthrobacter HW13 (*Paenarthrobacter*)	Mco 0018	NA	Yes (Other)	B
Comamonas testosteroni TK102	Mco 43372992	A0A076PMU7	Yes (Tat/SPI)	C
	Mco 43371778	A0A076PMU5	Yes (Tat/SPI)	D
Comamonas serinivorans C35	Mco 87276968	A0A1Y0EJD8	Yes (Tat/SPI)	D
Agrobacterium sp.	Mco 5546	A0A0Q8G0G0	Yes (Tat/SPI)	B
	Mco 5559	A0A0Q8FL38	Yes (Tat/SPI)	D
	Mco 3876	A0A1S7UA41	Yes (Tat/SPI)	D
Paraburkholderia xenovorans LB400	Mco 11491053	Q13QE6	Yes (Tat/SPI)	D
	Mco 49782249 (Bilirubin oxidase)	Q13RH6	Yes (Other)	A
Cupriavidus basilensis B-8	Mco 43356901	A0A0C4YQC6	Yes (Tat/SPI)	D
B. Non-lignin degraders				
Escherichia coli K12	CueO	P36649	Yes (Tat/SPI)	A
Escherichia coli	PcoA	Q47452	Yes (Tat/SPI)	C
Bacillus subtilis 168	CotA	P07788	Yes (Other)*	A
Bacillus licheniformis	CotA	Q65MU7	Yes (Other)**	A
Bacillus halodurans	McO	Q9KB49	Yes (Sec/SPII)	B
Thermus thermophilus HB27	McO	YP_005339 (GB)	Yes (Tat/SPI)	A.1
Streptomyces lavendulae	PPO	AB092576 (GB)	Yes (Other)	A
Marinomonas mediterranea	MCO	F2JUM9	Yes (Sec/SPII)	A.1

Table 2. DyP-type peroxidases identified in bacterial genomes of lignin degrading strains. Sequence groups are illustrated in Figure 4. Key: * Targeted to encapsulin nanocompartment. NA: not available.

Organism genome	Name	UniProt code	Signal peptide? SignalP-5.0 server	DyP-type subgroup
A. Lignin degraders				
Ochrobactrum sp. (MYb71)	DyP 1901	NA	Yes (Tat/SPI)	A
	DyP 564	NA	Yes (other)	C
Pseudomonas putida KT2440	DyP 3248	Q88HV5	Yes (other)*	B
Rhodococcus jostii RHA1	DyP ro05773	Q0S4I5	Yes (Tat/SPI)	A
	DyP ro02407	Q0SE24	Yes (other)*	B
Lysinibacillus sp.	DyP 1547	NA	Yes (Tat/SPI)	A
Comamonas testosteroni TK102	DyP 43373522	NA	Yes (other)*	B
Comamonas serinivorans C35	DyP 087281313	A0A1Y0ENS5	Yes (Other)*	B
Agrobacterium sp.	DyP 062	NA	Yes (other)	C
Paraburkholderia xenovorans LB400	DyP 11491480	Q13P55	Yes (Tat/SPI)	A
	DyP 7179030	Q144S4	No	B
Microbacterium phyllosphaerae	DyP 353	NA	Yes (Tat/SPI)	A
	DyP 1011	NA	Yes (Tat/SPI)	A
Paenibacillus sp.	DyP 01700	NA	Yes (Tat/SPI)	A
B. Non-lignin degraders				
Bacillus subtilis 168	DyP-A	P39597	Yes (Tat/SPI)	A
Escherichia coli K12	DyP-A	P31545	Yes (Tat/SPI)	A
Escherichia coli UTI89	DyP-B	Q1R8U0	Yes (Other)*	B
Acaryochloris marina MBIC 11017	DyP-C	A8ZLS7	Yes (Other)	C

Table 3. Genes for lignin degradation and aromatic degradation present in bacterial genomes.

Strain	Phylum	Activity	DyP-type peroxidases	Multi-copper oxidases	Other lignin degrading genes	Aromatic degradation gene clusters		
						β-keto-adipate	vanillate, ferulate	Other gene clusters
Rhodococcus jostii RHA1	Actinobacteria	Aromatic & lignin degrader	*dypB ro02407* *dypA ro05773*	*mcoA ro02377* *mcoB ro11201* *mcoC ro01580*	-	*pca ro01333* *catechol ro02371*	*vdh ro02986* *vanAB ro04163* *couLMN ro05122*	*hpc ro02849* *mhp ro00515* *bph ro08044, ro10121* *gent ro01863* *hmg ro02308* *phth ro08161, ro08175* *Trp ro01800* *HQ ro01857, ro11309* *ben ro02383* *phenol ro02513, ro08076*
Microbacterium phyllosphaerae		Lignin degrader from soil	*dypA 353* *dypA 1011*	-	-	*pca 1399*	-	*hpc 3755* *phenol 1371*
Arthrobacter sp.		Lignin degrader from soil	-	*mco0018*	-	*pca 1472*	*vanB 941*	*hpc 2196* *mhp 1610* *bph 1591* *gent 740* *phth 939* *Trp 2238* *hpc/mhp 730* *phenol 927*
Paenibacillus sp.	Firmicutes	Kraft lignin degrader	*dypA 01700*	-	-	-	-	*mhq 2216, 3760, 4617*
Lysinibacillus sp.		Lignin degrader from waste	*dypA 1547*	*mco1343*	-	-	-	*phenol 3252*
Pseudomonas putida KT2440	Gamma-proteobacteria	Aromatic & lignin degrader	*dypB PP3248*	*copA PP2205* *copAII PP5380*	-	*pca PP1566, PP4457* *catechol PP4236*	*vdh PP3357* *vanAB PP3736* *fcs/ech PP3356*	*pha PP3328* *hga PP5241* *ben PP3580*
Cupriavidus basilensis B-8	Beta-proteobacteria	Kraft lignin degrader	-	*mco 43356901*	-	*pca catABCD* *catechol meta-*	*hcaABC* *vanAB*	*mhb cluster* *salicylate*
Comamonas testosteroni TK102		Aromatic & lignin degrader	*dypB 43373522*	*mco 43372992 43371778*	-	*pca 2740* *catechol 3144*	-	*gent 2777* *ben 65* *phenol 3158*
Comamonas serinivorans C35		Lignin degrader	*dypB 087281313*	*mco 87276968*	-	*catechol 07050*	-	*benz 07055*
Paraburkholderia xenovorans LB400		Aromatic degrader	*dypA 11491480* *dypB 7179030*	*mco 11491053 49782249*	-	*pca 572/642/ 2775* *catechol 2107*	-	*hpc 2027* *mhp 2323* *bph 1186* *gent 2625* *hmg 2723* *phth 2747* *Trp 733, 1107* *ben 912* *phenol 1127*
Ochrobactrum sp	Alpha-proteobacteria	Lignin degrader from soil	*dypC 564* *dypA 1901*	*cueO 4352*	-	*pca 4452*	*vdh 814* *vanB 1060* *couL 262*	*hpc 96/1962* *mhq 2570*
Agrobacterium sp		Lignin degrader from waste	*dypC 62*	*mco5546* *mco5559* *mco3876*	*ligE 1577*	*pca 4311*	*vdh 04175*	*HQ 120* *mhq 196, 661*
Sphingobacterium sp. T2	Bacteroides	Lignin degrader from soil	-	-	*sod1 87* *sod2 3423*	-	-	-

Figure Legends

Figure 1. Partial amino acid sequence alignment showing bacterial multicopper oxidases found in lignin degrading strains. Numbering is based on *Ochrobactrum sp.* CueO sequence as reference. Well-conserved residues found in all MCOs are highlighted in blue, whereas new conserved residues found in fungal and bacterial laccases (Granja-Travez et al, 2018) are highlighted in red. Uniprot accession codes are listed in Table 1.1 and Table 1.2. Key: # enzymes reported to have lignin oxidation activity; a, MCOs found in non-lignin degrading strains for which recombinant enzyme has been characterized.

Figure 2. Phylogenetic tree of bacterial multicopper oxidases (MCO) sequences found in the genome of lignin degrading strains. Bootstrap analysis is shown. Key: # enzymes reported to have lignin oxidation activity; a, MCOs found in non-lignin degrading strains for which recombinant enzyme has been characterized.

Figure 3. Distribution of DyP-type peroxidases from A, B and C subgroups, by bacterial class. Key: * Archea; ** Eukaryota.

Figure 4. Phylogenetic tree of DyP-type peroxidases sequences found in the bacterial lignin degrading and non-lignin-degrading strains. Bootstrap analysis is shown. Key: a, Bacterial DyP-Type peroxidases found in non-lignin-degrading strains.

Figure 5. Aromatic degradation pathways found in bacterial lignin degraders. A. Pathways for which experimental evidence links to lignin degradation, as described in text. B. Other aromatic degradation pathways that could potentially be linked to lignin degradation.

	126	166	212	256
CueO (*Ochrobactrum sp.*)[#]	TLHWHGLFVPSI--LDGG	PWFHPHLHGNTARQAHL--GIA	LQDRR	IVRLRIL
CopA-I (*P. putida*)	SIHWHGIILPAN--MDGV	PWYHSHSGLQEQ------AGVY	FSDWS	TVRLRLI
CopA-II (*P. putida*)[#]	SIHWHGIILPAN--MDGV	PWYHSH----SGFQEQV--GVY	LSDWT	KIRLRFI
McoA ro02377 (*R. jostii*)	TVHWHGLAIR--NDMDGV	PWYHTHGDLQRGR------GLY	LSDWL	RARIRLI
McoB (*R. jostii*)	SVHWHGIALR--NDMDGA	SWAHPHVGLDTDY------GLY	LDDWT	RVRIRII
McoC 50787466 (*R. jostii*)	ATHLHGGVNPP--QFDGH	PWYHDHAMGVT--RLNIYAGLV	LQEKM	LYRMRLV
Mco 1343 (*Lysinibacillus*)	TFHWHGLEIPGNG--DGG	PWFHPHLEGATAEQ--VYKGLA	FQDRQ	KVRLRLL
Mco 0018 (*Arthrobacter*)	TLHWHGYDVP--NAMDGV	AWYHTHQDSAEG----VRKGLY	GHD--	LVRLRLI
Mco 43372992 (*C. testosteroni*)	SIHWHGIILPY--QMDGV	PWYHSHSGMQE------LTGMY	FSDWT	RVRLRMV
Mco 43371778 (*C. testosteroni*)	TIHWHGQRLP--NGMDGV	GMYHPH----SDEMVQMAMGMM	MTD--	KVRIRVG
Mco 87276968 (*C. serinivorans*)	TVHWHGQRLP--NGMDGV	GMYHPH----ADEMVQMAMGMM	MTD--	RVRIRVG
Mco 5546 (*Agrobacterium sp.*)	LIHWHGLTPPW--ELDGV	PWMHAH---TLQEQSLLAAPL-	LHDFS	RVRLRII
Mco 5559 (*Agrobacterium sp.*)	TIHWHGMILP--SGMDGV	GMYHPH----SDEMVQMAMGMM	MTD--	KVRVRVG
Mco 3876 (*Agrobacterium sp.*)	TIHWHGMILP--SGMDGV	GMYHPH----SDEMVQMAMGMM	MTD--	RVRVRVG
Mco 11491053 (*P. xenovorans*)	TVHWHGMLLP--CGMDGV	GMYHPH----ADEMVQMAMGMM	MTD--	RVRIRVG
Mco 49782249 (*P. xenovorans*)	TVHLHGAELPA--AYDGD	PWFHDHALGVT--RLNVYAGLA	VQDRM	-RRYRLH
Mco (*C. basilensis*)	TVHWHGVILP--SGMDGV	GMYHPH----SDEMVQMAMGMM	MTE--	RVRIRIG
PcoA (*E. coli*)[a]	SIHWHGIILPAN--MDGV	PWYHSHSGLQEQE------GVY	LSDWT	KIRLRFI
CueO (*E. coli*)[a]	TLHWHGLEVP--GEVDGG	PWFHPHQHGKTGR--QVAMGLA	VQDKK	WLRLRLL
McO (*T. thermophilus*)[a]	NLHWHG--LPISPKVDD-	-WYHPHLHGRVA--PQLFAGLL	LKD--	TLRLRLL
Lacc (*B. halodurans*)[a]	ALHLHGFPVP--NEMDGV	PWYHSHQDGAT----QVDRGLY	VID--	RVKLRFV
CotA (*B. licheniformis*)[a]	VVHLHGGCTPADS--DGY	PWYHDHAMAIT--RLNVYAGLV	IQDKS	KYRFRIL
CotA (*B. subtilis*)[a]	VVHLHGGVTPDDS--DGY	PWYHDHAMALT--RLNVYAGLV	ITDRT	KYRFRVI
PPO (*S. lavendulae*)[a]	VTHLHGAQTGGGN--DGW	AWYHDHAMNVT--RWNVHTGLY	IADRN	WHRFRLV
McO (*M. mediterranea*)[a]	NFHTHGLWVSPSKNSDNV	-WYHAHVHGSTA--LQVSSGMA	AYDCN	IQRWRFI
Consensus	TXHWHGLILP----MDGV	PWYHPH--G-TD---QVAXGLY	LTD--	RVRLRLI

	284	304	469	498	524
CueO (*Ochrobactrum sp.*)[#]	ASDGG	GER YE	MP-HPFHI HGA	WKDTA	MF HCHVLEHEDV GM--
CopA-I (*P. putida*)	AADGL	-LR IA	MT-HPIHL HGM	TVD--	AY HCHLLYHMET GM--
CopA-II (*P. putida*)[#]	AADGQ	-FR IA	MT-HPIHL HGM	TID--	AY HCHLLFHMEM GM--
McoA ro02377 (*R. jostii*)	DTDGF	GER YD	MF-HPMHL HGH	RKDTV	IT HCHNDYHLAA GM--
McoB (*R. jostii*)	HTDGF	GER YD	MW-HPMHL HGH	RKDTV	ML HCHNAYHQES GM--
McoC 50787466 (*R. jostii*)	GGDNG	AER VD	-P-HPVHL HLA	FKDTV	VW HCHILDHEDH EM--
Mco 1343 (*Lysinibacillus*)	ATDGG	GER AE	MT-HPFHI HGT	WKDTV	MF HCHVLEHEDN GM--
Mco 0018 (*Arthrobacter*)	AVDGT	GGR YD	EP-HPMHP HGH	VLDTV	MD HCHNLDHAAQ GM--
Mco 43372992 (*C. testosteroni*)	HVDGV	-FR FG	MT-HPMHL HGM	RRHTL	AW HCHLMFHMDA GM--
Mco 43371778 (*C. testosteroni*)	-----	--- --	---HPIHI HGH	TTD-V	AF HCHKSHHTMG AMGH
Mco 87276968 (*C. serinivorans*)	-----	--- --	---HPIHL HGH	TTD-V	AF HCHKSHHTMG AMGH
Mco 5546 (*Agrobacterium sp.*)	AVDGM	GQR LD	MT-HPMHL HGH	VRDTV	PF HCHHLYHMAA GM--
Mco 5559 (*Agrobacterium sp.*)	-----	--- --	---HPIHM HGY	SID-I	AI HCHKSHHTMN AMGH
Mco 3876 (*Agrobacterium sp.*)	-----	--- --	---HPIHM HGY	SID-I	AI HCHKSHHTMN AMGH
Mco 11491053 (*P. xenovorans*)	-----	--- --	---HPIHL HGY	TAD-I	AF HCHKSHHTMN AMGH
Mco 49782249 (*P. xenovorans*)	GADDN	GER AD	AS-HPVHI HLS	WKDTA	VW HCHVITHEDN EM--
Mco (*C. basilensis*)	-----	--- --	---HPIHL HGV	TAD-V	AF HCHKSHHTMN AMGH
PcoA (*E. coli*)[a]	AADGQ	--R IA	MT-HPIHL HGM	TIDVP	AY HCHLLYHMEM GM--
CueO (*E. coli*)[a]	ASDGG	GER FE	ML-HPFHI HGT	WKDTV	MA HCHLLEHEDT GM--
McO (*T. thermophilus*)[a]	AADGG	GER AE	MD-HPFHL HVH	WKDVV	VF HCHIVEHEDR GM--
Lacc (*B. halodurans*)[a]	HYDGQ	AER YD	FD-HPMHL HGD	LKDTL	MF HCHEFHHASG GMVA
CotA (*B. licheniformis*)[a]	GTDGG	AER CD	---HPIHL HLV	LKDTV	VW HCHILEHEDY DMMR
CotA (*B. subtilis*)[a]	GSDGG	AER YD	---HPIHL HLV	WKDTI	VW HCHILEHEDY DMMR
PPO (*S. lavendulae*)[a]	GSDGG	AER MD	---HPMHI HLA	YKDVF	MY HCHLLEHEDM GMMR
McO (*M. mediterranea*)[a]	ESNAI	QLN CD	APMHPYHI HVN	WRDTL	VL HCHILDHEDQ GMMQ
Consensus	A-DGG	-ER YD	M--HPIHL HGH	XKDTV	AF HCHXLXHEDN GM--

Figure 1.

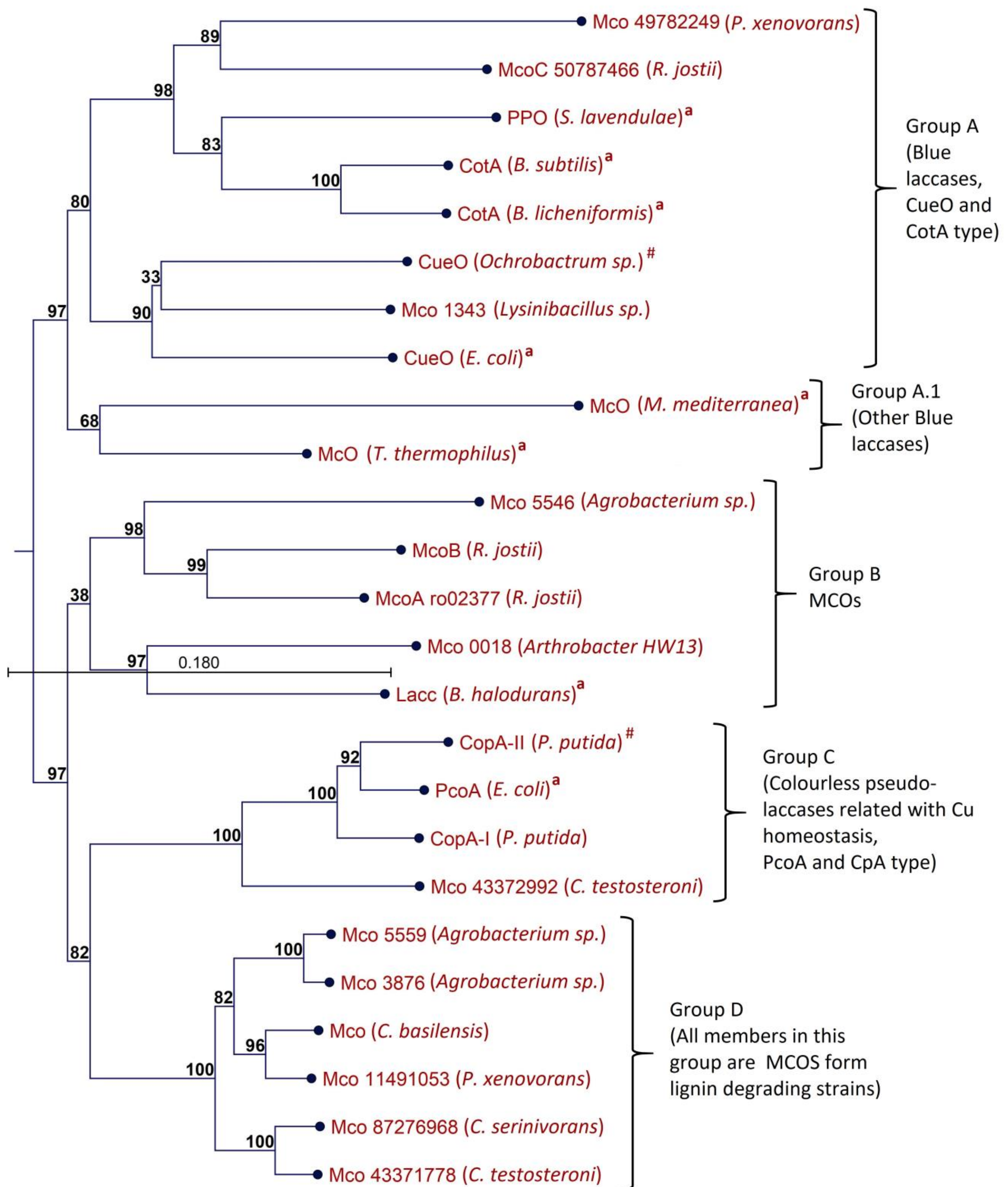

Figure 2.

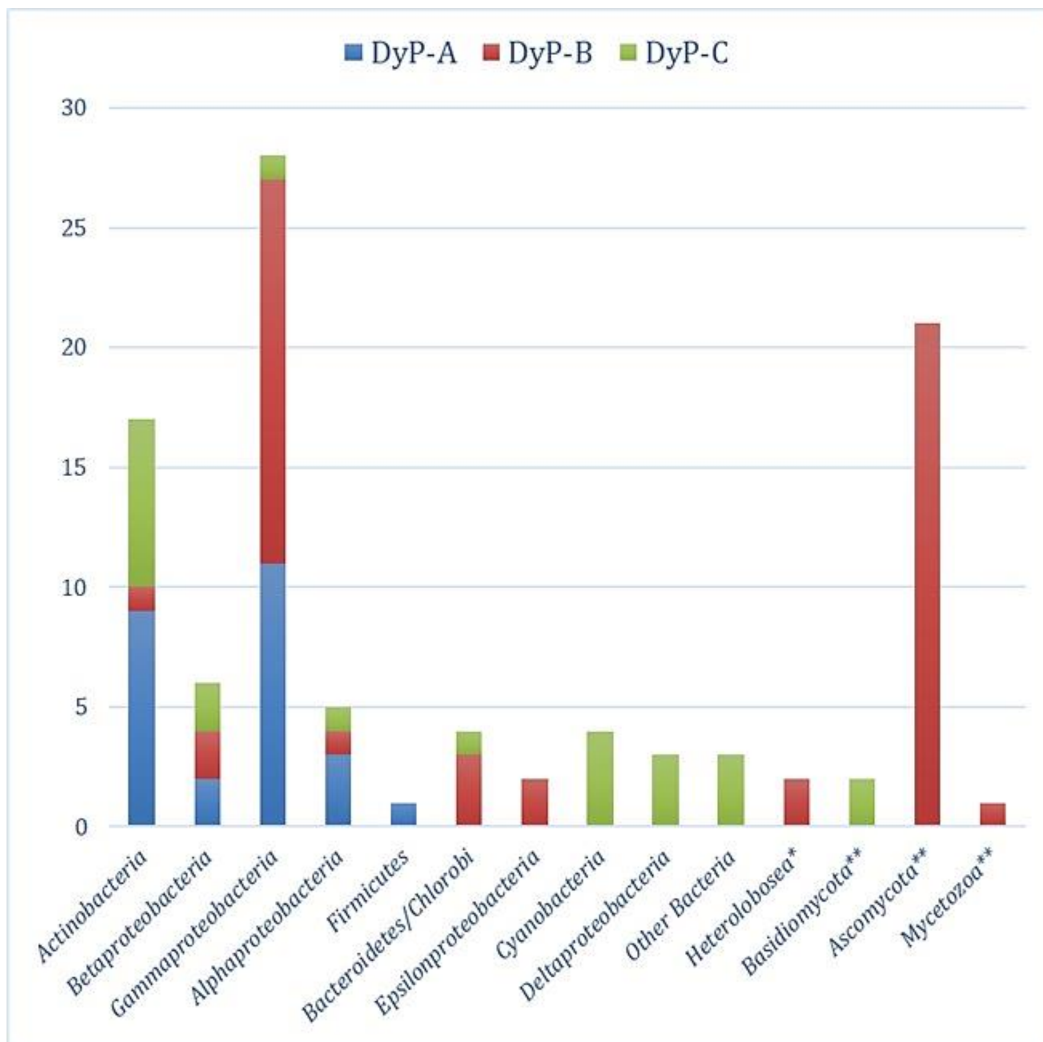

Figure 3.

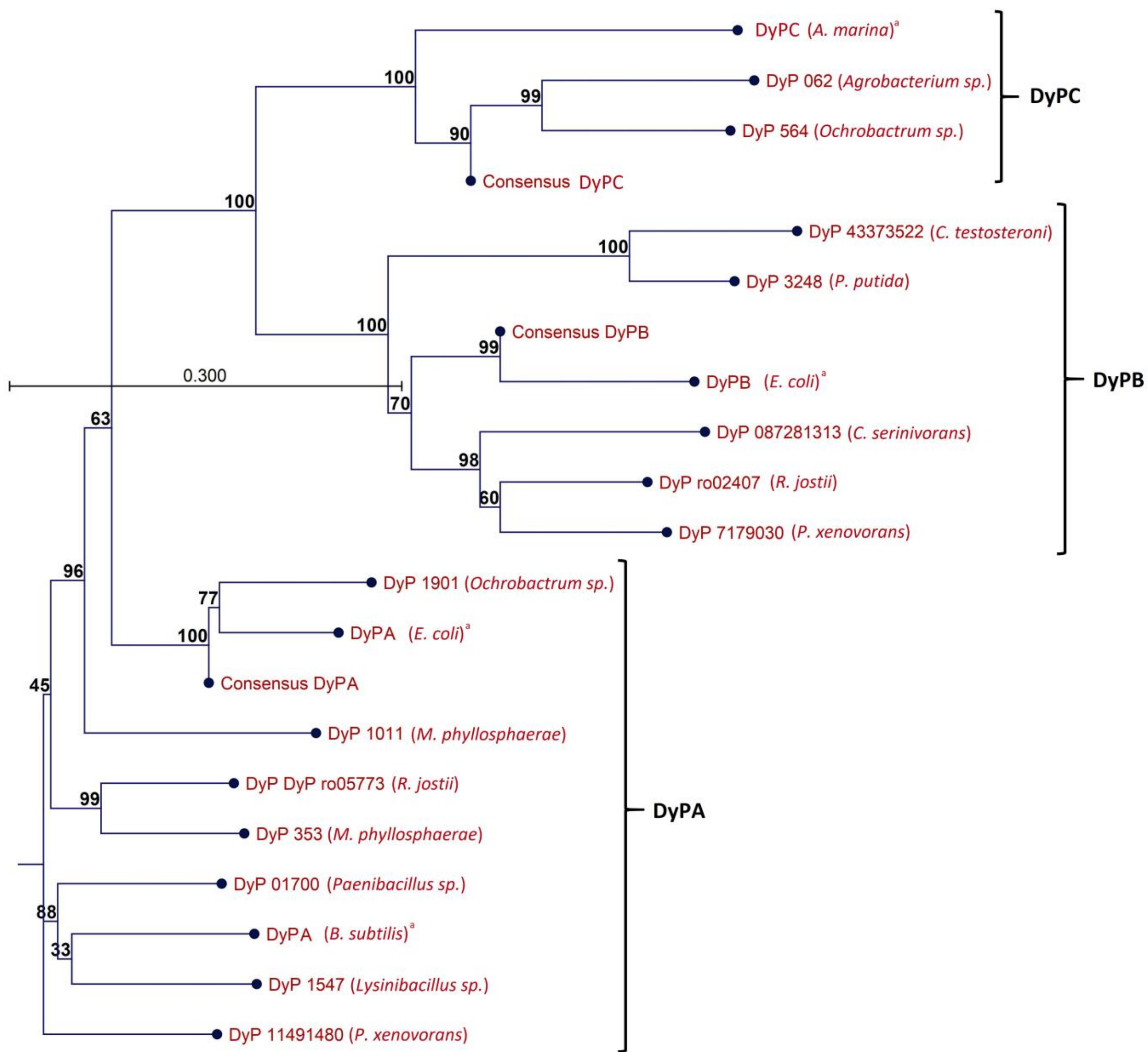

Figure 4.

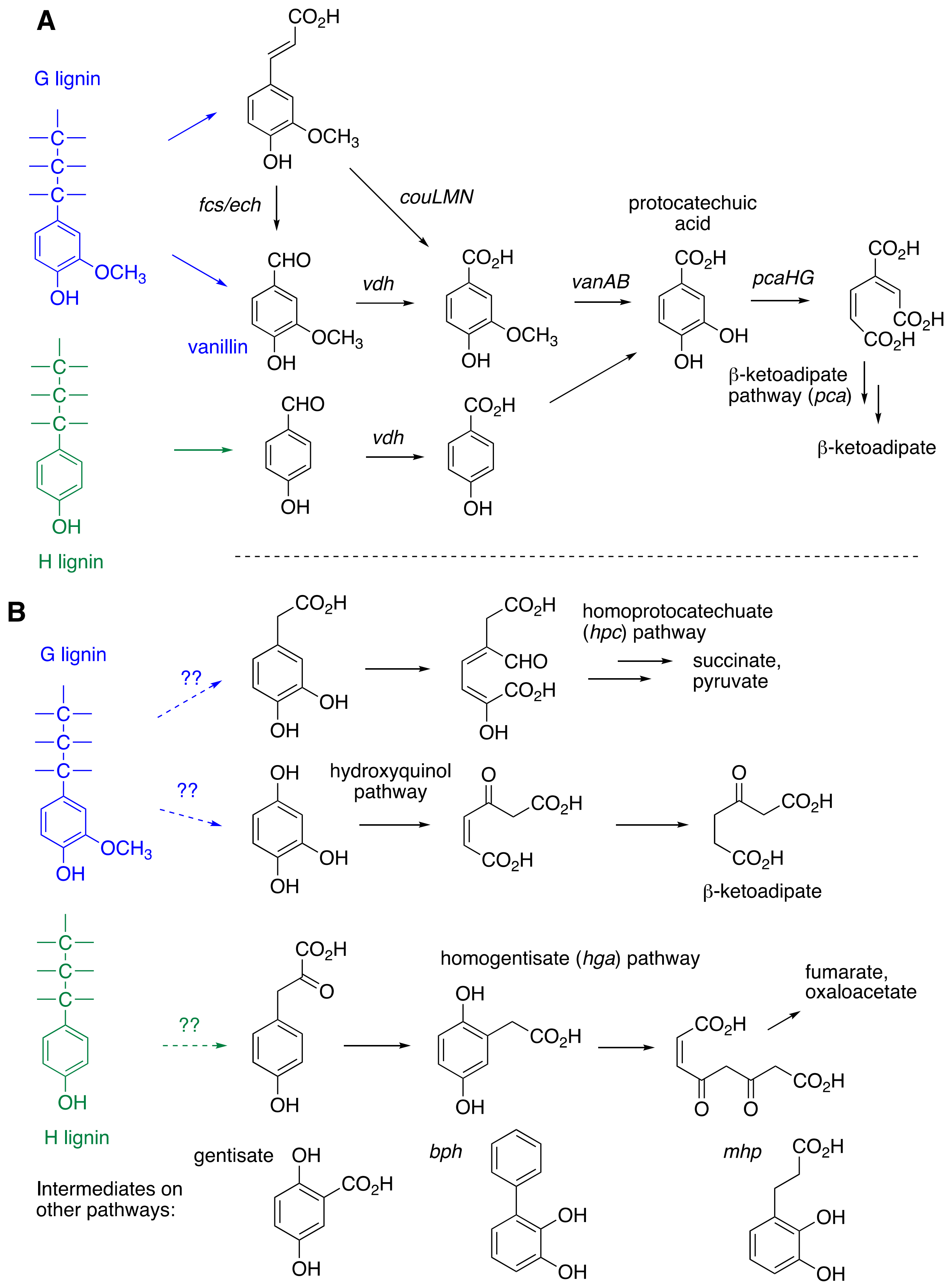

Figure 5.