Lignin-degrading peroxidases in Polyporales: an evolutionary survey based on 10 sequenced genomes

Francisco J. Ruiz-Dueñas
Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040-Madrid, Spain

Taina Lundell
Department of Food and Environmental Sciences, Viikki Biocenter, P.O. Box 56, University of Helsinki, FIN-00014 Helsinki, Finland

Dimitrios Floudas
Laszlo G. Nagy
Biology Department, Clark University, Worcester, Massachusetts 01610

Jose M. Barassa
Departamento de Ciencias de la Vida, Facultad de Biología, Ciencias Ambientales y Química, Universidad de Alcalá, Carretera de Barcelona Km 33.6, E-28871 Alcalá de Henares, Madrid, Spain

David S. Hibbett
Biology Department, Clark University, Worcester, Massachusetts 01610, USA

Angel T. Martínez1
Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040-Madrid, Spain

Abstract: The genomes of three representative Polyporales (Bjerkandera adusta, Phlebia brevispora and a member of the Ganoderma lucidum complex) were sequenced to expand our knowledge on the diversity of ligninolytic and related peroxidase genes in this Basidiomycota order that includes most wood-rotting fungi. The survey was completed by analyzing the heme-peroxidase genes in the already available genomes of seven more Polyporales species representing the antrodia, gelatoporia, core polyporoid and phlebioid clades. The study confirms the absence of ligninolytic peroxidase genes from the manganese peroxidase (MnP), lignin peroxidase (LiP) and versatile peroxidase (VP) families, in the brown-rot fungal genomes (all of them from the antrodia clade), which include only a limited number of predicted low redox-potential generic peroxidase (GP) genes. When members of the heme-thiolate peroxidase (HTP) and dye-decolorizing peroxidase (DyP) superfamilies (up to a total of 64 genes) also were considered, the newly sequenced B. adusta appears as the Polyporales species with the highest number of peroxidase genes due to the high expansion of both the ligninolytic peroxidase and DyP (super)families. The evolutionary relationships of the 111 genes for class-II peroxidases (from the GP, MnP, VP, LiP families) in the 10 Polyporales genomes is discussed including the existence of different MnP subfamilies and of a large and homogeneous LiP cluster, while different VPs mainly cluster with short MnPs. Finally, ancestral state reconstructions showed that a putative MnP gene, derived from a primitive GP that incorporated the Mn(II)-oxidation site, is the precursor of all the class-II ligninolytic peroxidases. Incorporation of an exposed tryptophan residue involved in oxidative degradation of lignin in a short MnP apparently resulted in evolution of the first VP. One of these ancient VPs might have lost the Mn(II)-oxidation site being at the origin of all the LiP enzymes, which are found only in species of the order Polyporales.

Key words: ancestral state reconstruction, brown-rot fungi, electron transfer, genome sequencing, Mn(II)-oxidation site, white-rot fungi

INTRODUCTION

Although many Basidiomycota species in the orders Agaricales, Boletales and Russulales, among others, and some Ascomycota species contribute to lignocellulose decay, the ability to degrade wood is a typical feature of the members of the order Polyporales (class Agaricomycetes). This capability was an essential evolutionary trait for recycling the large carbon amounts fixed by land plant photosynthesis, which was estimated to be acquired by ancestral basidiomycetes in the later Carboniferous period (Floudas et al. 2012). Nowadays, fungal decay of wood represents a natural model for the sustainable use of plant resources in the lignocellulose biorefineries (Ragauskas et al. 2006, Martínez et al. 2009).

The first basidiomycete genome to be sequenced in 2004 was that of a Polyporales species, Phanerochaete chrysosporium (Martinez et al. 2004), due to the interest on this white-rot fungus as a model lignin-degrading organism (Kersten and Cullen 2007). After the discovery of the first “ligninase” (lignin peroxidase, LiP) in P. chrysosporium (Tien and Kirk 1983), enzymatic degradation of lignin has been thoroughly investigated due to the biotechnological interest to gain access to lignocellulose
carbohydrates in industrial applications (Martínez et al. 2009). Other enzymes involved are manganese peroxidase (MnP), first discovered also in P. chrysosporium, and versatile peroxidase (VP), discovered in the Agaricales species Pleurotus eryngii, together with various oxidases providing \( \text{H}_2\text{O}_2 \) to peroxidases. Ecological impact of white-rot fungal oxidative decay of lignin and wood is immense in the forest ecosystems (Lundell et al. 2010).

Brown-rot fungi have developed an alternative strategy, based on Fenton chemistry, to overcome the lignin barrier in wood decay (Balldrian 2006). The Fenton reaction results in formation of hydroxyl free radical being able to access, oxidize and depolymerize cellulose with a limited modification of lignin (Martínez et al. 2011). In 2009 the Rhodonia placenta (synonym: Postia placenta) genome was sequenced as the model brown-rot fungus to increase our understanding of this wood-decay pattern (Martínez et al. 2009).

Wood decay by P. chrysosporium is characterized by the simultaneous degradation of both lignin and carbohydrates, in contrast with a few white-rot Polyporales species that are able to remove lignin with a minimal depletion of cellulose (Otjen and Blanchette 1986). This degradation pattern has high potential for delignification of wood and non-wood plant materials in biopulping and bioethanol production applications (Akhtar et al. 1997, Salvachúa et al. 2011). As a representative species for the selective wood delignification pattern, the genome of a third Polyporales species, Gelatoporia subvermispora (synonym: Ceriporiopsis subvermispora), has been sequenced (Fernández-Fueyo et al. 2012a).

In the present paper, three representative wood-rotting Polyporales, Bjerkaandra adusta, Ganoderma sp. (from the Ganoderma lucidum complex) and Phlebia brevispora, were selected for sequencing to extend the genomic information on the distinct taxonomic clades of Polyporales. Among them, B. adusta is a strong lignin-degrading species colonizing a variety of angiosperms and also some coniferous trees (Ryvarden and Gilbertson 1994). Together with LiP and MnP (Heinfling et al. 1998) the lignin-attacking system of B. adusta includes VP, which was described in this polypore (Mester and Field 1998) and in the agaric P. eryngii (Martínez et al. 1996) as a MnP-LiP hybrid enzyme. This third ligninolytic peroxidase family has not been isolated yet from many lignin-degrading Polyporales species, such as P. chrysosporium and G. subvermispora. Ganoderma species include selective white-rot fungi being able to cause extensive wood delignification, which has been considered a model for biotechnological applications (Agosín et al. 1990, Martínez et al. 1995). Finally, the white-rot species P. brevispora, together with G. subvermispora, was investigated for biomechanical pulping resulting in significant energy savings due to selective lignin removal (Akhtar et al. 1993). Species of the genus Phlebia are efficient producers of LiP and MnP enzymes (Niku-Paavola et al. 1990, Lundell et al. 2010, Mäkelä et al. 2013).

In addition to the three newly sequenced genomes, other recently sequenced Polyporales genomes, up to a total of 10, were subjected to peroxidase gene annotation and comparative sequence analyses. The study is part of a wider project, initiated by a phylogenetic and phylogenomic overview of the order Polyporales (Binder et al. 2013) and will continue with other parallel studies based on the same 10 Polyporales genomes. More than 100 ligninolytic peroxidase-encoding genes are compared to establish their distribution and evolutionary history in the Polyporales clades, in connection with white-rot and brown-rot wood-decay patterns. The study focuses on fungal ligninolytic peroxidases in class II (hereinafter POD) of the superfamily of plant-fungal-bacterial (non-animal) heme-peroxidases (Welinder 1992). In addition to PODs, the distribution and evolutionary history of two recently discovered heme-peroxidase superfamilies, the so-called heme-thiolate peroxidases (HTP) and dye-decolorizing peroxidases (DyP) (Sugano 2009, Hofrichter et al. 2010, Ruiz-Dueñas and Martínez 2010), are investigated for the first time in the order Polyporales.

**Materials and methods**

*Genome sequencing.—* The genomic DNA sequences of B. adusta (HHB-12826-SP), P. brevispora (HHB-7050-SS6) and Ganoderma sp. (10597-SS1, from the G. lucidum complex) were obtained from the Joint Genome Institute (JGI) of the US Department of Energy (DOE) in Walnut Creek, California, as part of the Saprotrrophic Agaricomycotina Project of the JGI Fungal Genome Program. The genomes were produced as described in the accompanying article of Binder et al. (2013). The gene prediction and annotation results for the three genomes are available in Mycocosm (Grigoriev et al. 2012) at the JGI portal (http://jgi.doe.gov/fungi) or directly through http://genome.jgi-psf.org/Bjea1L1, http://genome.jgi-psf.org/Gansp1 and http://genome.jgi-psf.org/Phlb1L respectively.

*Peroxidase gene and protein model analyses.—* The complete inventory of the heme-peroxidase genes present in the three new genomes was obtained in a process including: (i) screening the automatically annotated gene models at the JGI portal; (ii) revising and manually curating, when necessary, the positions of introns and the predicted N and C termini of the corresponding protein models; (iii) comparing the amino-acid sequence identities with related protein sequences; and (iv) confirming the presence of
characteristic residues at the heme pocket and substrate oxidation sites, after protein structural homology modeling using peroxidase crystal structures available at the RCSB Protein Data Bank (www.pdb.org) as templates.

The additional bioinformatic methods used in the analyses were: BLAST (basic local alignment tool) (Altschul et al. 1990) programs at the National Center for Biotechnology Information (NCBI, USA) for searching the NCBI and UniProt nucleotide and protein databases; MEGA 5 molecular evolutionary package (Tamura et al. 2011) and Geneious 5.5.5 (from Biomatters, available from http://www.geneious.com/) software platform for nucleotide and protein sequence multiple alignments during gene model searches and manual curation; and SignalP 3.0 software (Bendtsen et al. 2004) for prediction of signal peptides, their cleavage signal peptide sites and protein targeting. N-terminal secretion signal peptide predictions were validated by multiple alignment (with MEGA 5, Clustal W or Muscle algorithms) of the 487 basidiomycete heme-peroxidase sequences available, which include sequences whose mature protein N-termini were sequenced.

Theoretical molecular 3D structural models of 64 peroxidases identified in the three new genomes were generated at the Swiss Institute of Bioinformatics SWISS-MODEL automated protein homology-modeling server (Bordoli et al. 2009) using the following crystal structures as templates: (i) *P. chrysosporium* LiP (PDB 1L7L) and MnP (PDB 1LYD), *P. eryngii* VP (PDB 3FJW) and *Coprinopsis cinerea* (synonym: *Coprinus cinereus*) peroxidase, also described as *Arthromyces ramosus* “nomen nudum” peroxidase (PDB 1ARX), as POD templates; (ii) *Leptoxypinus fumago* (conidial state: *Calderariomyces fumago*) chloroperoxidase (PDB 1CPO) and unpublished *Agrocybe aegeria* peroxygenase (from K. Piontek) as HTP templates; and (iii) *B. adusta* DyP, formerly described as *Thanatophilus cucumeris* DyP (PDB 2D3Q) and recently published (Strittmatter et al. 2012) *Auricularia auricula-judae* DyP (PDB 4AU9) as DyP templates.

Phylogenetic analyses.—The evolutionary history of the 179 peroxidase sequences identified in 10 Polyporales genomes, including the new *B. adusta*, *Ganoderma* sp. and *P. brevispora* genomes plus the genomes of *P. chrysosporium*, *Dickomitus squaleus*, *Trametes versicolor*, *G. subvermispora*, *Fomitopsis pinicola*, *R. placenta* and *Wolfiporia cocos* already available at JGI, was established. Amino-acid sequences of the three peroxidase gene superfamilies (corresponding to POD, HTP and DyP enzymes) were aligned separately with PRANK with default settings (Loytynoja and Goldman 2008), followed by minor manual adjustments. Highly ambiguous regions of the alignments were excluded from the analyses and ultrametric trees were inferred with Bayesian MCMC in BEAST (Drummond and Rambaut 2007) with 10 000 000 generations in two replicates, default priors, the wag model of protein evolution and a gamma distribution to account for rate heterogeneity. The *Auricularia delicata* sequence No. 101509, annotated as putative class-I heme peroxidase in the genome (http://genome.jgi.doe.gov/Aurde1/Aurde1), was included for rooting purposes. Burn-in was determined on the basis of likelihood convergence, using the Tracer 1.5 program of the BEAST 1.7.22 package.

Gene tree/species tree reconciliation.—The ML tree from each analysis with bootstrap values was rooted and reconciled using Notung 2.6 (Durand et al. 2006, Vernot et al. 2008). The settings for the edge weight threshold were 90 and 75 during performance of rearrangements, while the default cost of duplications and losses was used.

Ancestral state reconstruction (ASR).—One hundred trees were subsampled randomly from the Bayesian tree sample (excluding the burn-in) with Mesquite (Maddison and Maddison 2011) and were used in the ASR analysis. The existence of a Mn(II)-oxidation site (three conserved acidic amino-acid residues near the internal propionate of heme) and the presence or absence of the exposed tryptophan residue related to lignin degradation as the starting point for long range electron transfer (LRRET), homologous to *P. chrysosporium* LiP Trp171 (Blödsg et al. 2001) and *P. eryngii* VP Trp164 (Pérez-Boada et al. 2005), were coded as discrete traits. Mn(II)-oxidation sites were coded as typical (consisting of three amino acids), atypical (two amino acids) or absent (non-functional, 0–1 amino acid/s). Ancestral states were estimated at 16 nodes of the Class-II peroxidase phylogeny, representing the backbone of the tree and important nodes with respect to changes in the catalytic tryptophan and Mn(II)-oxidation sites. The BayesTraits package (Pagel et al. 2004) was used for ASR, under the ML criterion, with an asymmetric model of trait evolution and 10 ML attempts per tree. To estimate which state is the most plausible at the designated nodes, each node was fixed in one state at a time, and the likelihood values were recorded for each of the 100 trees. Next, the fit of alternative states was compared with 2 log likelihood units as a measure of significance (Pagel et al. 2004). This metric approximates the significance threshold of likelihood ratio tests, which cannot be applied in these cases due to the non-nested nature of the models being compared.

RESULTS

Inventory of heme-peroxidase genes in three new genomes of Polyporales.—A total of 64 heme-peroxidase encoding genes, including 39 POD, 9 HTP and 16 DyP models, were identified after manual annotation and analysis in the genomes of *B. adusta*, *Ganoderma* sp. (*G. lucidum* complex) and *P. brevispora* (Table I; references in Supplementary Table I). The gene
model prediction was based on multiple alignments of the deduced open-reading frame-translated protein sequences and inspection of the enzyme structural molecular models described below for characteristic heme environments and catalytic residues, as described in Ruiz-Duenas and Martinez (2010) and Fernandez-Fueyo et al. (2012c). Together with the above PODs (class II fungal peroxidases), four class I peroxidase genes, also belonging to the superfamily of plant-fungal-bacterial peroxidases, were identified representing one putative cytochrome c peroxidase per genome and one hybrid ascorbate-cytochrome c peroxidase gene in the *Ganoderma* sp. genome.

However, the class I peroxidases are not included in the study due to their prokaryotic origin and enzymatic functions for purposes other than oxidation and degradation of lignin.

Of the three new basidiomycete genomes sequenced, that of *B. adusta* has the highest number (a total of 34) and diversity of peroxidase genes (Table I) with representatives for the four POD families namely six MnPs, 12 LiPs, one VP and one non-ligninolytic generic peroxidase (GP), as well as for the HTP (four genes) and DyP (10 genes) superfamilies. The total peroxidase gene numbers in the two other genomes are significantly lower (16 and 14 genes in *P. brevispora* and *Ganoderma* sp. respectively), both genomes lacking GP genes. Notably, when five LiP genes are predicted in the *P. brevispora* genome no VP genes are found whereas in the *Ganoderma* sp. genome two VP but no LiP genes are annotated. Concerning the MnP enzymes, currently three subfamilies of short, long and extra-long MnPs have been proposed (Floudas et al. 2012). In the new fungal genomes, protein models of the short MnP subfamily were annotated while long MnP models were found only in the *P. brevispora* genome. No extra-long MnP models were predicted in any of the three genomes. On the other hand, one of the *Ganoderma* sp. models (No. 116446) corresponds to an atypical VP with the Mn(II)-oxidation site formed by only two acidic residues, while the third residue (homologous to Asp183 in FIG. 1C) is a histidine.

Among the 64 peroxidase genes identified in the new genomes (see SUPPLEMENTARY TABLE I) only a few in *B. adusta* (Nos. 159724 LiP, 43095 VP, 72253 DyP) correspond to previously cloned genes/cDNAs (with NCBI GenBank protein sequence references 1906181A, AAO47909, BAA77283 respectively).

### Table I. Inventory of genes (175) of peroxidase types (from the LiP, VP, GP families; the short, long and extra-long MnP subfamilies; and the HTP and DyP superfamilies) in the genomes of 10 Polyporales (*BJEAD, B. adusta*; *PHLBR, P. brevispora*; *PHACH, P. chrysosporium*; *DICSQ, D. squalens*; *GANSP, Ganoderma sp.* [ *G. lucidum complex*]; *TRAVE, T. versicolor*; *GELSU, G. subvermispora*; *FOMPI, F. pinicola*; *RHOPLE, R. placenta*; and *WOLCO, W. cocos*) producing white-rot and brown-rot wood decay

<table>
<thead>
<tr>
<th>Taxonomy* →</th>
<th>- - - Phlebioid - - -</th>
<th>- - - Core polyporoid -</th>
<th>Gelatopo-</th>
<th>- - - Antrodia - - -</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJE AD</td>
<td><strong>6</strong></td>
<td><strong>3</strong></td>
<td><strong>0</strong></td>
<td><strong>5</strong></td>
</tr>
<tr>
<td>PHL BR</td>
<td><strong>0</strong></td>
<td><strong>3</strong></td>
<td><strong>5</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>PHA CH</td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>4</strong></td>
</tr>
<tr>
<td>DIC SO</td>
<td><strong>12</strong></td>
<td><strong>5</strong></td>
<td><strong>10</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>GAN SP</td>
<td><strong>1</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>3</strong></td>
</tr>
<tr>
<td>TRA VE</td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>GEL SU</td>
<td><strong>20</strong></td>
<td><strong>11</strong></td>
<td><strong>16</strong></td>
<td><strong>12</strong></td>
</tr>
<tr>
<td>FOM PI</td>
<td><strong>HTP</strong></td>
<td><strong>4</strong></td>
<td><strong>2</strong></td>
<td><strong>4</strong></td>
</tr>
<tr>
<td>RHO PL</td>
<td><strong>DyP</strong></td>
<td><strong>10</strong></td>
<td><strong>3</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>WOL CO</td>
<td><strong>All peroxidases</strong></td>
<td><strong>34</strong></td>
<td><strong>16</strong></td>
<td><strong>18</strong></td>
</tr>
</tbody>
</table>

*Class I peroxidases of prokaryotic origin and two GP and three HTP alleles from *R. placenta* (SUPPLEMENTARY TABLE I; Figs. 2, 3) are not included here.*

Molecular structural modeling of the POD, HTP and DyP enzymes.—Six representative molecular models for the different peroxidase families/superfamilies found in the three new genomes, namely LiP, MnP, VP and GP families (collectively referred as PODs), and HTP and DyP superfamilies, are illustrated (Fig. 1).

All PODs share a common structure and folding. This includes two structural Ca^{2+} ions and generally four disulfide bridges (except for long/extra-long MnPs that contain a fifth, C-terminal disulphide bridge) resulting in tight packing of the proteins. A
FIG. 1. Molecular models for representative members of the different peroxidase families and superfamilies found in the *B. adusta, Ganoderma* sp. (*G. lucidum* complex) and *P. brevispora* genomes. A–D. Models for the MnP (*Ganoderma* sp. No. 142454), LiP (*P. brevispora* No. 150901), VP (*B. adusta* No. 43095) and GP (*B. adusta* No. 173495) POD families respectively. E. Model for the HTP superfamily (*Ganoderma* sp. No. 64073). F. Model for the DyP superfamily (*P. brevispora* No. 75878). The molecular models for the peroxidases identified in the Polyporales genomes were built from their predicted amino-acid sequences using crystal structures of related proteins as templates.
conserved heme environment, with a histidine residue acting as the fifth iron ligand, also was found. The MnP and the VP structural models (Fig. 1A, C) have a characteristic Mn(II)-binding and oxidation site, as formed by the carboxylates of one aspartate and two glutamate side chains and the internal propionate of heme (i.e. the propionate occupying the most internal position with respect to the main heme access channel). In addition, the VP models share with the LiP models the presence of an exposed tryptophan residue (Fig. 1B, C) responsible for oxidation of lignin and other high redox-potential aromatic substrates by LRET to heme. Finally, GP models lack all these specific amino-acid residues (Fig. 1D) and their only catalytic site is apparently the main heme access channel, where low redox-potential substrates would be oxidized in direct contact with the heme edge. Among the 17 LiP genes identified in the three new genomes (see Supplementary Table I), three *B. adusta* genome models (Nos. 172246, 184697, 499082) indicate variations in the heme pocket topology and another three (Nos. 41982, 121664, 269524) also would present an atypical catalytic tryptophan orientation. Therefore, no functional molecular models could be predicted for these LiPs with the crystal templates available.

The protein structural models for HTPs are different from the POD models not only in their general structure but also in the presence of a cysteine instead of a histidine as the heme iron proximal ligand, together with other differences in the heme environment (Fig. 1E). Among the nine predicted HTP models (see Supplementary Table I), one (*B. adusta* No. 81882) represents an unusually short (only 169 amino acids) chloroperoxidase-like protein. Finally, the DyP structural models exhibit a histidine heme ligand (Fig. 1F) likewise in PODs but their general molecular architectures and heme environments are very different. Among the 16 DyP models predicted (see Supplementary Table I) two of *B. adusta* (Nos. 37610, 371576) contain unusually long N-terminal tail, which was not removable by manual curation of the corresponding gene models.

**Peroxidase distribution in 10 Polyporales genomes.**—In addition to the three new genomes, seven more species of Polyporales with genomes available at the JGI fungal genomics program site were analyzed. These species include representative of the antrodia (*F. pinicola*, *R. placenta*, *W. cocos*), gelatoporia (*G. subvermispora*), core polyporoid (*D. squalens*, *T. versicolor*) and phlebioid (*P. chrysosporium*) clades (Binder et al. 2005).

The number of genes of the peroxidase families and superfamilies in each of the 10 genomes is illustrated (Table I), up to a total of 175 genes (whose references are provided in Supplementary Table I). Among the Polyporales genomes, the three of the antrodia clade correspond to brown-rot species, while the other seven species cause white-rot decay of wood. Of note, the MnP, LiP and VP families were absent from the brown-rot genomes but present in all the white-rot species. Namely, all the white-rot genomes possess genes from one or several MnP subfamilies (up to 13 copies in *G. subvermispora*) and either LiP (in *P. brevispora* and *P. chrysosporium*) or VP (in *D. squalens* and *Ganoderma* sp.) or both LiP and VP genes (in the other genomes). In contrast, GPs are present in the three brown-rot fungal genomes but only in some of the white-rot genomes analyzed. The opposite was observed for DyPs, which are more widespread among the white-rot genomes with up to 10 genes annotated in *B. adusta*. Finally, the HTP genes are the only peroxidase genes present in all the genomes analyzed, irrespective of the taxonomic clade or wood degradation pattern leading to either white-rot and brown-rot decay.

**Evolutionary history of heme-peroxidases in the Polyporales genomes.**—The history of the 179 peroxidases identified in the 10 Polyporales genomes was established by comparing their predicted amino-acid sequences with ML methods. Because of the large phylogenetic distance between POD, HTP and DyP proteins, the three superfamilies were analyzed separately. Concerning HTPs, the 42 sequences found were grouped into the four clusters illustrated in the phylogram (Fig. 2A). One of them, cluster A formed by only three protein sequences from *G. subvermispora*, positioned far from the other clusters. The other 42 sequences form clusters B, C and D, the two latter appearing as the most related. HTP sequences from the 10 analyzed genomes are included in clusters B and D, whereas protein representatives of the core polyporoid clade are absent from cluster C. Concerning the DyP superfamily, the evolutionary history of the 21 sequences present in six of the 10 genomes analyzed shows four main clusters (Fig. 2C). Their composition reveals that DyPs from species of the same clade are more closely related to each other than to sequences from species of other clades. In this way clusters C and D are formed only by putative DyP sequences from *P. brevispora* and *B. adusta* respectively.

The evolutionary analysis of the 113 PODs found in the 10 genomes (Fig. 3A) identified four main clusters in the POD phylogram, some of them divided into several subclusters. This is the case of the large cluster D, which includes a LiP subcluster where the *B. adusta* (12 sequences), *P. chrysosporium* (10
FIG. 2. ML phylogenetic analyses of 45 HTPs (A) and 21 DyPs (C) from 10 Polyporales genomes with bootstrap values of ≥75, and estimated range of gene copies for HTP (B) and DyP (D) at the ancestral nodes of the represented Polyporales phylogeny based on Notung reconciliation analyses for edge weight threshold 90 and 75 (see Table I for species abbreviations). Presence of a single number at a node indicates that both settings estimated the same number of gene copies for that node. Three HTP alleles in *R. placenta* are indicated with asterisks, and they have not been considered as duplication events of the lineage leading to *R. placenta*. The phylogenetic trees included in B and D were taken from Binder et al. (2013).
Fig. 3. ML phylogenetic analysis of PODs (113) from 10 Polyporales genomes (A) with bootstrap values of $\geq 75$, and estimated range of gene copies (B) at the ancestral nodes of the represented Polyporales phylogeny based on Notung reconciliation analyses for edge weight threshold 90 and 75 (see Table I for species abbreviations). Presence of a single number at a node indicates that both settings estimated the same number of gene copies for that node. Two GP alleles in *R. placenta* are indicated with asterisks, and they have not been considered as duplication events of the lineage leading to *R. placenta*. The phylogenetic tree included in B was taken from Binder et al. (2013).
sequences), *P. brevispora* (five sequences) and *T. versicolor* (10 sequences) LiP sequences form four homogeneous groups (although the latter also includes *T. versicolor* VP No. 43289 and *Ganoderma* sp. VP No. 116056). This fact reveals the high evolutionary relationships among the LiP isoenzymes in each of the Polyporales analyzed (5–12 per genome) providing evolutionary support for gene duplications within each species.

Concerning MnPs, the evolutionary analysis shows that all the so-called long (15 sequences involved) and extra-long (nine sequences) MnPs, found in four species (TABLE I), are closely related, forming a well defined cluster C (FIG. 3A). The so-called short MnPs appear more heterogeneous from an evolutionary point of view. Sixteen of the short MnPs (from *D. squalens*, *Ganoderma* sp., *T. versicolor*, *B. adusta*, *P. brevispora*) form the second subcluster of cluster D that also includes one *B. adusta* VP (No. 43095). The rest of the short MnPs are included in cluster B (11 sequences) or unclustered in cluster D (five sequences) together with one GP (B. adusta No. 173495) and two putative VP (No. 99382) and LiP (No. 118677) models from *G. subvermispora*. Cluster B is heterogeneous because, together with the 11 short MnPs (from five species), it also comprises six VP sequences (from *D. squalens*, *T. versicolor*, *Ganoderma* sp.) including two atypical VPs lacking one of the Mn(II)-oxidation residues. Contrary to the LiPs, which form a homogeneous subcluster with four species-specific groups, the high heterogeneity of the short MnPs spreads the isoenzymes to separate enzyme clusters. *D. squalens*, *Ganoderma* sp., *T. versicolor* and *B. adusta* have short MnPs in both clusters B and D. The situation is similar for VPs that also are included in both cluster B (six sequences) and D (four sequences), suggesting strong evolutionary relationships between VP and short MnP enzymes.

Finally, the small cluster A is homogeneous, being formed by seven GP sequences from the three species of the antrodia clade and *G. subvermispora*. Two additional GPs are included in cluster B (*P. chrysosporium* No. 6250 corresponding to the GP-type product of the *nop* gene) and the large cluster D (see above).

**Peroxidase gene duplication and loss during diversification of Polyporales.**—The reconciliation results (FIGS. 2, 3) reveal different patterns of gene duplications and losses between the three peroxidase superfamilies and corresponding gene types. The Notung results suggest few duplication and loss events for the HTP genes along the core polyoporoid and phlebioid clades (FIG. 2B), but expansion of the HTP gene number are suggested for the gelatoporia and antrodia clades. In the DyP superfamily, only one or two genes were maintained throughout the diversification of Polyporales (FIG. 2D) while some white-rot (*G. subvermispora*, *P. chrysosporium*) and brown-rot species (*W. cocos*, *F. pinicola*) have lost all the DyP genes. Exception to this is the lineage leading to *B. adusta*, with eight estimated gene duplications.

In contrast to the HTP and DyP superfamilies, the numbers of POD gene copies in the genomes sampled deviate significantly from the numbers estimated for the last common ancestor of Polyporales (FIG. 3B). Thus, this common ancestor genome potentially possessed 5–6 POD genes, but the species included in this study encode 1–25 POD genes, indicating significant expansions and contractions in the gene numbers. Notung results suggest that losses of all the ligninolytic POD genes occurred in the antrodia clade, associated with a transition from a white-rot to a brown-rot wood-decay strategy. These losses would take place after the separation of the gelatoporia and antrodia clades but most likely before the diversification of the latter clade. On the other hand, both the phlebioid and core polyoporoid clades show multiple duplications of POD genes. Some of these duplications appear to have taken place before the origin of the core polyposes, but most of the duplication events map on the terminal branches.

**ASR of POD encoding genes.**—The study of ancestral POD genes suggests that the first ligninolytic POD was a MnP (node 12 in FIG. 4 and SUPPLEMENTARY TABLE II) while other ligninolytic peroxidases have arisen multiple times from the MnP ancestor genes. VP-type peroxidases seem to have evolved twice from the MnP ancestors by acquiring the codon for the exposed catalytic tryptophan involved in LRET from the protein surface from nodes 14 to 16 and nodes 6 to 4 (FIG. 4, TABLE II). Two additional transitions have occurred in *G. subvermispora* and *B. adusta* because node 8 also is reconstructed as a MnP.

Concerning LiP genes, the ASR results suggest that their emergence also has happened at least two times. The first case is the appearance of a single LiP gene in the genome of *G. subvermispora* after node 9, which is reconstructed as a MnP (SUPPLEMENTARY TABLE II). The second case includes nodes 3 and 5, which appear to be transitional. The ASR analysis cannot resolve the type of genes present at these intermediate nodes, and thereby both the presence and absence of Mn(II)-oxidation site are equally supported. This means that, depending on the presence or absence of a Mn(II)-oxidation site at these nodes, the emergence of the rest of LiPs (up to a total of 38 genes in the 10 genomes analyzed) either could have happened only once before the diversification of the
core polyporoid and phlebioid clades (node 3) or twice, with VP genes independently giving rise to LiP genes in both clades.

**DISCUSSION**

**PODs in Polyporales.**—Ligninolytic peroxidases from the LiP and MnP families were isolated and characterized for the first time near 30 y ago from cultures of *P. chrysosporium* (Glenn et al. 1983, Tien and Kirk 1983, Glenn and Gold 1985). Then more enzymes of these two families were described from cultures of *T. versicolor* (Dodson et al. 1987, Johansson and Nyman 1987), *Phlebia radiata* (Niku-Paavola et al. 1988, Karhunen et al. 1990) and *B. adusta* (Heinfling et al. 1998a, ten Have et al. 1998), among other species of Polyporales (Lundell et al. 2010). Moreover MnPs but no LiP were described in *G. subvermispora* (Rüttimann et al. 1992, Lobos et al. 1994) and other Polyporales degrading lignin selectively (Maeda et al. 2001, Hakala et al. 2005, Elissetche et al. 2007). More recently, peroxidases of the VP family, which have been thoroughly characterized in *P. eryngii* from the order Agaricales (Ruiz-Duénas et al. 2009), also were isolated from the polypore *B. adusta* (Heinfling et al. 1998b, Mester and Field 1998). Finally, no GP type enzymes, similar to the well known GP from *C. cinerea* (order Agaricales) (Morita et al. 1988), have been isolated from species of the order Polyporales. In summary, the total number of POD genes cloned to date from species of Polyporales is fewer than 30 (Lundell et al. 2010, Ruiz-Duénas and Martínez 2010). However, the current bioinformatic search for ligninolytic peroxidases reveals the existence of around 175 genes in the 10 Polyporales genomes, representing the four POD families.

**HTPs in Polyporales.**—The chloroperoxidase from the ascomycete *L. fumago* for years was the only known fungal HTP enzyme (Thomas et al. 1970). The discovery of a related enzyme in the wood- and litter-colonizing Agaricales species *A. aegerita*, with low haloperoxidase and predominant peroxygenase activity on aromatic (Ullrich et al. 2004, Ullrich and Hofrichter 2005) and aliphatic compounds (Gutiérrez et al. 2011, Peter et al. 2011), constituted the first report of this type of enzymes in Basidiomycota. HTP enzymes share the heme-pocket architecture (with a proximal cysteine as the fifth ligand to the heme iron)

---

![Maximum clade credibility chronogram of PODs (113) from the genomes of 10 Polyporales species, with inclusion of the *A. delicata* (AURDE) genome class-I peroxidase sequence No. 101509 for rooting purposes](image-url)

For each node we reconstructed the presence or absence of LRET tryptophan as well as the ancestral type of Mn(II)-oxidation site. For each node the first colored circle indicates the ancestral Mn(II)-oxidation site, while the second indicates whether LRET tryptophan was inferred as present or absent (see inset legend). Ambiguities are displayed by colored circles divided in two semicircles corresponding to both plausible states (see SUPPLEMENTARY TABLE II for the likelihood values at the different nodes).
and substrate oxygenation mechanism with the cytochrome P450 monooxygenases. Therefore, they constitute a distinct enzyme superfamily (Hofrichter et al. 2010, Ruiz-Dueñas and Martínez 2010). The P450 enzymes require an auxiliary flavooxidase or flavin domain mediating their activation by molecular oxygen and a source of reducing power (NAD[P]H). However, the A. aegerita HTP peroxygenase presents the biotechnological advantage of acting as a self-sufficient monooxygenase directly activated by hydrogen peroxide (Hofrichter and Ullrich 2006).

HTPs were isolated from the litter-degrading agaric Marasmius rotula (Gröbe et al. 2011) and more than 20 HTP genes were identified in the genome/transcriptome of Agaricus bisporus (Morin et al. 2012), but HTPs have not yet been isolated from Polyporales species. However, the genomic screening indicates that HTPs are widespread in Polyporales, with 2–9 copies per genome and forming several well defined sequence-based clusters. This fact, together with variations in the modeled Polyporales HTP active site and heme-access channel architectures, strongly suggest distinct catalytic functions and substrate specificities, which remain to be investigated. Oxidation of the LiP/VP substrate veratryl alcohol (Ullrich et al. 2004) and cleavage of lignin model dimers (Kinne et al. 2011) were reported for the A. aegerita peroxygenase (Kinne et al. 2011). However, the ether-bond cleavage activity of this HTP is more adapted to oxygenation of smaller compounds than to a direct action on polymeric lignin (Hofrichter and Ullrich 2006).

This agrees with the similar average number of HTP gene numbers found in the non-ligninolytic brown-rot (6.0) and ligninolytic white-rot (6.4) Polyporales genomes.

DyPs in Polyporales.—The DyP-type peroxidases represent a divergent protein superfamily (Zubieta et al. 2007, Sugano 2009, Hofrichter et al. 2010), biochemically characterized by its dye-decolorizing ability. DyP first was described from a conidial fungus identified as Geotrichum sp. due to whitish colonies and arthroconidial propagation (Kim and Shoda 1999). Ribosomal DNA sequencing reveals that the fungus is a conidial state of B. adusta (Ruiz-Dueñas et al. 2011, Yoshida et al. 2011). This strain initially was misidentified as T. cucumeris (the Rhizoctonia solani sexual state) (Sato et al. 2004) due to several errors in the ITS sequence origin (fungal taxon) information at the NCBI database, as discussed by Ruiz-Dueñas et al. (2011).

DyP-type enzymes are reported for a variety of fungal and bacterial species, including cyanobacteria (Hofrichter et al. 2010). Among fungi, DyP enzymes were isolated from cultures of wood-decay Agarico-mycetes species, such as A. auricula-judae of the order Auriculariales (Liers et al. 2010) and from the Agaricales Pleurotus ostreatus (Shin et al. 1997, Faraco et al. 2007) and Mycetinis scorodonius (synonym: Marasmius scorodonius) (Scheibner et al. 2008). With the only exception of the B. adusta DyP cited above, no other reports for DyP active enzymes or DNA sequences were provided until now from species of the order Polyporales. However, our bioinformatic analyses retrieved 21 DyP genes from genomes of B. adusta, P. brevispora, D. squalens, Ganoderma sp., T. versicolor and R. placenta (with P. chrysosporium and G. subvermispora as the only white-rot polypos without a single DyP gene).

As in the case of HTPs, the evolutionary changes observed in the number of DyP genes are small, with gene loss in four genomes. Noticeably and exceptionally recent gene duplications leading to expansion of DyPs have occurred in the genome of B. adusta (up to 10 genes). In accordance to HTPs, DyP protein sequences form several clusters with different evolutionary relationships and probably deviating functions that require further investigations. Of note, the A. auricula-judae DyP demonstrates lignin-degradation ability in oxidation of non-phenolic lignin model dimers (Liers et al. 2010). This fact agrees with the higher average number of DyP genes in the ligninolytic white-rot species (2.6) than is calculated for the antrodia clade brown-rot Polyporales (0.7).

Evolutionary relationships of ligninolytic and other PODs.—A more in depth evolutionary analysis of the POD families and subfamilies could be performed (Fig. 3). This was due to the large number of MnP, LiP and VP genes available, which expanded during the evolution of white-rot Polyporales up to 19–25 genes in B. adusta and T. versicolor, and the great amount of biochemical information available. In contrast, most members of the DyP and HTP superfamilies are known only from genomes and thereby lack catalytic studies.

All GPs, except two sequences that clustered with LiP and VP/MnP-short sequences, appear evolutionarily apart from the rest of PODs (forming cluster A). This agrees with their primitive nature presented by ASR analyses (Fig. 4) presenting them as the precursors for ligninolytic peroxidases. The low redox-potential of GPs prevents their oxidation of lignin (Fernández-Fueyo et al. 2012c). However, they may contribute to iron reduction by phenol redox-cycling (Gómez-Toribio et al. 2009) for generation of hydroxyl radical involved in cellulose depolymerization. This hypothesis is in agreement with the conservation of the corresponding genes in the three brown-rot polypos genomes. The ASR analysis
demonstrated that the first ligninolytic peroxidase was a MnP (node 12 in Fig. 4) apparently derived from an ancestral GP, a fact that agrees with the presence of MnP-type peroxidases in all of the three other main POD groups (clusters B, C, D).

Concerning the MnP subfamilies, cluster C formed by the so-called long and extra-long MnPs (Floudas et al. 2012), is the most homogeneous and distant from the rest of the POD sequences (Fig. 3). These two subfamilies are closely related (intermixed in the tree) and have significantly diverged from the short MnPs, and other ligninolytic peroxidases (clusters B, D). Long MnPs were the first MnP enzymes described from cultures of P. chrysosporium (Glenn and Gold 1985). The so-called extra-long MnPs correspond to enzymes presenting a longer C-terminal tail, first isolated from the polypore D. squalens and reported as significantly more thermostable (Li et al. 2001). The third MnP subfamily, called short MnPs, was described more recently, but it includes the larger number (up to 33 gene models) of the MnPs identified from the 10 Polyporales genomes. These MnPs are characterized by a shorter C-terminal tail than in the typical P. chrysosporium MnPs and slightly different catalytic properties. For example, short MnPs are able to oxidize some low redox-potential compounds in the absence of Mn(II), as described for P. radiata MnP3 (Hildén et al. 2005, Hofrichter et al. 2010) and also for short MnPs from the genome of P. ostreatus (Fernández-Fueyo et al. 2012b). Two additional issues concerning short MnPs are their evolutionary divergence in Polyporales, being present in both clusters B and D (Fig. 3), and the fact that in both cases they appear intermixed (cluster D) or closely related (cluster B) with VP sequences.

VPs, which are enzymatically and structurally characterized by the presence of an exposed tryptophan involved in oxidation of lignin similar to that found in LiPs, and a Mn(II)-oxidation site similar to that of MnPs, have been extensively characterized in P. eryngii (Ruiz-Dueñas et al. 2009). Their presence in the order Polyporales, including phlebioid (B. adusta), core polyporoid (T. versicolor, Ganoderma sp., D. squalens) and gelatoporia (G. subvermispora) clade genomes, indicates two possible separate origin for VP genes in the Agaricomycetes. The VP origin also could be multiple among Polyporales being produced (Fig. 3 clusters B, comprising VPs of the antrodia clade) typically grow on coniferous wood, in agreement with the general tendency observed for the brown-rot Agaricomycetes (Hibbett and Donoghue 2001). The opposite tendency was observed for the white-rot Polyporales in the

Of note, all the Polyporales LiPs (37 models in the 10 genomes), with only one exception, form a homogeneous subcluster (cluster D, Fig. 3) with four groups corresponding to four typical LiP-producing Polyporales (B. adusta, P. chrysosporium, P. brevispora, T. versicolor). The subcluster also includes two VPs (from T. versicolor and Ganoderma sp.) that retained the potentiality of the VP ancestor of the LiP subcluster. One last LiP from G. subvermispora (No. 118677) is included in this cluster, but the sequence appears unrelated to the other LiP sequences. This unusual LiP, together with a related hypothetical VP from the same fungus (No. 99382), were described as VP-LiP transitional enzymes after their heterologous expression and biochemical characterization (Fernández-Fueyo et al. 2012c). It is important to mention that, among the different PODs discussed above, LiPs are the only enzymes being restricted to the order Polyporales, taking into account isolation, protein characterization and gene cloning (Ruiz-Dueñas et al. 2009), together with wider genomic analysis (Floudas et al. 2012). This suggests a central role of LiPs in the white-rot wood-decay pattern by species of the order Polyporales.

The phylogenetic placement of some VP sequences together with LiPs and the corresponding ASR analysis suggest that evolution of a LiP gene from an ancestral short MnP gene is a process that may have included two steps: (i) a transition of a MnP to a VP due to gain of a catalytic tryptophan and (ii) the subsequent loss of the Mn(II)-oxidation site to give rise to a LiP gene. The presence of atypical VPs (with only two, instead of three, acidic residues at the above site) in some of the Polyporales genomes gives support to this two-step evolution, thus indicating that for some VP genes there is a gradual codon substitution leading to deterioration of the Mn(II)-oxidation site.

**Wood decay by Polyporales.**—Concerning the ecology and feeding strategies of Polyporales, most of the species in the order, including all 10 species analyzed in the present study, are saprotrophic organisms that cause either brown-rot or white-rot decay of wood. The three brown-rot species (F. pinicola, R. placenta, W. cocos of the antrodia clade) typically grow on coniferous wood, in agreement with the general tendency observed for the brown-rot Agaricomycetes (Hibbett and Donoghue 2001). The opposite tendency was observed for the white-rot Polyporales in the
gelatoporia, core polyporoid and phlebioid clades, where five species (G. subvermispora, T. versicolor, P. brevispora, P. chrysosporium, B. adusta) colonize both conifer and angiosperm wood and only one (D. squalens) causes selective white-rot decay in wood of conifers (Hibbett and Donoghue 2001).

The present POD encoding gene survey in the 10 Polyporales genomes confirms the preliminary conclusions resulting from the comparison of the first sequenced white-rot and brown-rot fungal genomes of P. chrysosporium (Martinez et al. 2004) and R. placenta (Martinez et al. 2009) respectively. According to our phylogenomic and evolutionary analyses, brown-rot decay in the antrodia clade of the Polyporales has resulted from the absence of POD genes while the three other Polyporales clades apparently have maintained their white-rot feeding strategy as an inherited and preserved trait from the common ancestor of the class Agaricomycetes. In a similar way brown-rot decay patterns would have independently appeared in four other classes of Agaricomycetes from the white-rot initial ancestor (Floudas et al. 2012). The evolutionary tendency to the loss of POD genes in the brown-rot species often is associated with growth on coniferous wood. In these species a decay strategy for direct chemical attack to cellulose, based on hydrogen peroxide-producing oxidases and iron-reducing enzymes resulting in hydroxyl radical generation, substitutes the (peroxidase-oxidase) enzymatic ligninolytic strategy of the white-rot fungi (Martinez et al. 2009).

The latter strategy would predominate in the fungal decay of angiosperm wood and, as a consequence, the number of POD genes (of the LiP, VP, MnP families) expanded in the white-rot species (even up to 19 and 25 genes in the B. adusta and T. versicolor genomes respectively). High redox-potential POD genes also are present in the genomes of wood and forest litter white-rot Agaricales, such as Pleurotus ostreatus (with up to a total of nine short MnPs and VPs) (Ruiz-Dueñas et al. 2011) and A. bisporus (with only two short MnPs) (Morin et al. 2012) respectively. Further genomic studies in this order most likely will confirm the presence of POD genes in other saprotrophic Agaricales species causing white-rot decay of wood and forest litter.

The highest POD expansions among the genomes analyzed, with the presence of the different ligninolytic peroxidase families, were found in the three more generalistic white-rot species (G. subvermispora, T. versicolor, B. adusta) in agreement with their high lignin-degrading ability and adaptation to a wide range of both conifer and angiosperm wood (and lignin) types. Moreover, it seems that species growing only on conifers, such as D. squalens, produce VP (and MnP) and have no LiP while the opposite (production of LiP in the absence of VP) is observed for the species mainly growing on woody angiosperms, such as P. chrysosporium and P. brevispora. In this way, wood-decay evolution within the Polyporales might have involved different strategies: (i) an ancestral white-rot strategy, at first on angiosperm wood, caused by the phlebioid clade species; (ii) a second, derived white-rot strategy on gymnosperm wood, caused by the core polyporoid clade species; and (iii) a third, more derived brown-rot strategy on gymnosperm wood, caused by the species of the antrodia clade. Of course species of the phlebioid, antrodia and core polyporoid clades can colonize and live on wood of both gymnosperms and angiosperms. Therefore the generality of this inference needs to be tested with new sequenced genomes across the Polyporales, representing different feeding strategies and growth-substrate specificities.

Concerning the ability of a few white-rot Polyporales to degrade lignin selectively (Otjen and Blanchette 1986), some interesting clues can be obtained from the analysis of the genome of G. subvermispora, a species studied for wood delignification in paper pulp manufacture (Akhtar et al. 1993). First, the existence of LiP/VP-type peroxidases in this fungus, which had been a matter of debate due to unsuccessful enzyme activity detection (Srebotnik et al. 1994), is revealed by the genomic analysis. The lignin-degrading ability of these enzymes is confirmed by studies using synthetic lignin and lignin model dimers after their heterologous expression (Fernández-Fueyo et al. 2012c). Second, the noteworthy expansion of the long and extra-long MnP genes (12 copies) in the G. subvermispora genome suggests that these enzymes also may play a key role for efficient degradation of wood lignin, in agreement with biochemical results showing the secretion of MnP enzymes during wood delignification by other white-rot Polyporales (Hakala et al. 2005, Mäkelä 2009).

The present genomic survey demonstrates important findings of the evolutionary history of the fungal ligninolytic class-II peroxidase gene families and simultaneously provides strong support for their involvement in lignin degradation resulting in white-rot decay, the most characteristic feeding strategy of the wood-rotting Basidiomycota species of the class Agaricomycetes, order Polyporales.

ACKNOWLEDGMENTS

This work was supported by the PEROXICATS (www.peroxicats.org) project KBBE-2010-4-265397 of the European Union (to ATM), by the HIPOP project BIO2011-26694 (to FJR-D) and project CGL-2009-07316 (to JMB) of the Spanish
Ministry of Economy and Competitiveness (MINECO), by the OX-RED project (AP-138331) of the Academy of Finland (to TL), and by the PolyPEET project (DEB-0933081) of the US National Science Foundation (to DSH). The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231. FJR-D thanks a Ramon y Cajal contract of the Spanish MINECO. Klaus Piontek (Freiburg University, Germany) is acknowledged for providing access to the unpublished crystal structure of A. aegerita HTP.

LITERATURE CITED


RUIZ-DUEÑAS ET AL.: PEROXIDASE GENES IN POLYPORALES GENOMES


———, Martínez AT. 2010. Structural and functional features of peroxidases with a potential as industrial