The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes
Dimitrios Floudas et al.
Science 336, 1715 (2012);
DOI: 10.1126/science.1221748

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of June 29, 2012):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/336/6089/1715.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2012/06/28/336.6089.1715.DC1.html

A list of selected additional articles on the Science Web sites related to this article can be found at:
http://www.sciencemag.org/content/336/6089/1715.full.html#related

This article cites 24 articles, 11 of which can be accessed free:
http://www.sciencemag.org/content/336/6089/1715.full.html#ref-list-1

This article has been cited by 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/336/6089/1715.full.html#related-urls

This article appears in the following subject collections:
Evolution
http://www.sciencemag.org/cgi/collection/evolution
The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes

Dimitrios Floudas,1 Manfred Binder,2 Robert Riley,2 Kerrie Barry,2 Robert A. Blanchette,3 Bernard Aerts,4 Angel T. Martinez,2 Robert Otiliar,2 Joseph W. Spatafora,6 Jagjit S. Yadav,7 Andrea Aerts,7 Isabelle Benoit,8,9 Alex Boyd,2 Alexis Carlson,1 Alex Copeland,2 Pedro M. Coutinho,4 Ronald P. de Vries,6,9 Patricia Ferreira,10 Keisha Findley,11 Brian Foster,2 Jill Gaskell,12 Dylan Glotzer,1 Pawel Górecki,13 Joseph Heitman,11 Cedar Hesse,6 Chiaki Hon,14 Kiyohiko Igarashi,14 Joel A. Jurgens,3 Nathan Kallen,1 Phil Kersten,12 Annegret Kohler,5 Ursula Kües,16 T. K. Arun Kumar,17 Alan Kuo,2 Kurt LaButti,2 Luis F. Larrondo,18 Erika Lindquist,2 Albee Ling,1 Vincent Lombard,4 Susan Lucas,2 Taina Lundell,19 Rachael Martin,1 David J. McLaughlin,17 Ingo Morgenstern,20 Emanuelle Morin,15 Claude Murat,15 Laszlo G. Nagy,1 Matt Nolan,2 Robin A. Ohm,2 Aleksandrina Patyshakulyeva,2 Antonis Rokas,21 Francisco J. Ruiz-Dueñas,5 Grzegorz Sabat,22 Asaf Salamov,2 Masahiro Samejima,4 Jeremy Schmutz,23 Jason C. Slot,21 Franz St. John,12 Jan Stenlid,24 Hui Sun,2 Sheng Sun,11 Khajamohiddin Syed,7 Adrian Tsang,20 Ad Wiebenga,6 Darcy Young,1 Antonio Pisabarro,25 Daniel C. Eastwood,26 Francis Martin,15 Dan Cullen,26 Igor V. Grigoriev,2* David S. Hibbett1*

Wood is a major pool of organic carbon that is highly resistant to decay, owing largely to the presence of lignin. The only organisms capable of substantial lignin decay are white rot fungi in the Agaricomycetes, which also contains non-lignin-degrading brown rot and ectomycorrhizal species. Comparative analyses of 31 fungal genomes (12 generated for this study) suggest that lignin-degrading peroxidases expanded in the lineage leading to the ancestor of the Agaricomycetes, which is reconstructed as a white rot species, and then contracted in parallel lineages leading to brown rot and mycorrhizal species. Molecular clock analyses suggest that the origin of lignin degradation might have coincided with the sharp decrease in the rate of organic carbon burial around the end of the Carboniferous period.

Lignin is a heterogeneous polymer that provides strength and rigidity to wood, protects cellulose and hemicellulose from microbial attack, and is the major precursor of coal (1). Genomic studies of wood decay organisms have focused on model fungal systems for white rot (in which all plant cell wall components are degraded), such as Phanerochaete scaringae.
chrysosporium (2), and brown rot (in which lignin is modified but not appreciably degraded), such as Postia placenta (3) and Serpula lacrymans (4). However, these species represent just two of the 18 recognized orders of Agaricomycetes, of which five contain brown rot taxa. To reconstruct the evolution of lignin decay mechanisms, we analyzed 31 diverse fungal genomes, including 12 newly sequenced species of Agaricomycota (Table 1). The new genomes comprise six white rot species, five brown rot species, and one mycoparasite, representing nine orders (Fig. 1 and figs. S1 to S5) (5).

To estimate phylogenetic relationships among these taxa, we constructed data sets using 71 or 26 single-copy genes, with varying alignment criteria and treatments for fast-evolving sites, yielding matrices of 10,002 to 34,257 amino acids, which we analyzed with maximum likelihood (ML) and Bayesian methods (5–7). All but six nodes receive maximal support values in all analyses, and the rest are strongly supported (bootstrap >99%).

We searched all 31 genomes for 27 gene families encoding oxidoreductases and carbohydrate-active enzymes (CAZymes) that have been implicated in wood decay (Table 1). CAZymes, particularly those acting on crystalline cellulose, are abundant in white rot genomes, whereas brown rot genomes have 32 to 68 copies (average 46) from 9 to 12 families. The ectomycorrhizal (ECM) Laccaria bicolor resembles brown rot species in this regard, possessing 28 CAZyme genes in eight families (Table 1). Notably, glycoside hydrolase (GH) families GH6 and GH7, which include cellobiohydrolases that are involved in the attack of crystalline cellulose (8), are present in all white rot lineages, but they are absent in brown rot lineages (except Boletales) and L. bicolor. Similar patterns of enrichment in white rot genomes are shown by genes encoding GH61 enzymes, which have a copper-dependent oxidative mechanism for disrupting crystalline cellulose (9), and cellulose binding modules (CBM1), which effectively increase the concentration of the enzymes on the surface of crystalline cellulose (10) (Table 1).

To gain access to cellulose, wood-decaying fungi must overcame or circumvent lignin; thus, we focused on fungal class II peroxidases (PODs), which degrade lignin in P. chrysosporium and other species (11) (figs. S7 to S19). We classified PODs into four major groups, including three ligninolytic forms—lignin peroxidase (LiP), manganese peroxidase (MnP), and versatile peroxidase (VP)—and a fourth POD type, defined here as “generic peroxidase” (GP), which is expected to include nonligninolytic low-redox potential peroxidases with catalytic properties similar to those of the peroxidase of Coprinopsis cinerea or the product of the nopA gene in P. chrysosporium (5, 12). LiPs possess a tryptophan residue on the surface of the enzyme corresponding to Trp171 in P. chrysosporium LiP-H8 that enables direct oxidation of lignin compounds via long-range electron transfer; MnPs possess two or three residues corresponding to Glu53, Glu56, and Asp175 of P. chrysosporium MnPI that function in binding Mn (13). VPs possess both the Trp171 homolog and Mn-binding residues, whereas all are lacking in GPs.

Consistent with a central role for PODs in lignin degradation, white rot species have 5 to 26 copies (average 14) of genes encoding ligninolytic PODs, but all brown rot species lack these enzymes, as do the ECM L. bicolor, the soil saprotroph C. cinerea, and Schizopyllum commune, which has been regarded as a white rot fungus but has a limited capacity to degrade lignin (14). Moreover, analyses of gene diversification with binary state speciation analysis (15) confirmed that the rate of duplication of POD genes is elevated in white rot lineages versus non-white rot lineages (3).

To reconstruct the functional evolution of PODs, we performed Bayesian and ML analyses using the GPs of Ascomycota as outgroups, and we estimated the ancestral states of the key residues of ligninolytic PODs using BayesTraits (17). Our results indicate that the ancestor of all PODs likely lacked the Mn-binding and Trp171 residues of MnP, LiP, and VP, suggesting that it was nonligninolytic (Fig. 1B). The most recent ancestor of all ligninolytic Agaricomycete PODs is reconstructed as an MnP, suggesting that there was a single origin of LiP (gain of Trp171 and loss of Mn-binding residues), with parallel expansions in the P. chrysosporium and Trametes versicolor (Polyporales, each with 10 LiP copies; Fig. 1B and figs. S7 and S17). We also identified two origins of VP in the Polyporales, where T. versicolor and Dichomitus squalens each have three VP copies (Fig. 1B and fig. S7). VPs are also produced in the “oyster mushroom” Pleurotus ostreatus (Agaricales) (18), indicating further convergent evolution of this class of enzymes.

To localize the diversification of PODs in the organismal phylogeny, we performed gene tree/species tree reconciliation analyses using CAFE (19), Notung (20), and DrML (21). All methods suggest that a single POD gene copy was present in the common ancestor of Basidiomycota, with parallel losses in lineages leading to the Pucciniomycota, Ustilaginomycota, Tremellomycetes, and Dacryopinax sp. (Fig. 1A). Diversification of PODs occurred in the lineage leading to the most recent common ancestor of the Agaricomycetes (node “A” in Fig. 1A), which is reconstructed as having to two seven POD gene copies in the various analyses. In addition, reconciliation analyses suggest that the ancestor of the Agaricomycetes possessed one or two genes encoding dye-decolorizing peroxidases (DyPs), which are heme peroxidases that have been shown to degrade lignin model compounds (22), as well five to eight genes encoding oxidases (including glyoxal oxidase) involved in peroxide generation (5, 23). Collectively, these results suggest that the ancestor of Agaricomycetes was a white rot species that possessed a ligninolytic system with PODs, DyPs, and multiple pathways for H2O2 production.

The “backbone” nodes in the Agaricomycete phylogeny (labeled “B” in Fig. 1A) are reconstructed as having to 3 to 16 POD gene copies, which suggests that the white rot mechanism was retained throughout the early evolution of Agaricomycetes. Subsequently, all reconciliation analyses suggest that there were parallel expansions of POD genes in terminal lineages, leading to white rot species in five orders (Auriculariales, Hymenochaetales, Corticiales, Russulales, and Polyporales). In contrast, parallel contractions of PODs are resolved within lineages leading to the brown rot Dacryopinax sp., Gloeophyllum trabeum, the Boletales, and the brown rot Polyporales.
suggesting that these lineages lost PODs as they shifted to a non ligninolytic mode of wood decay (Fig. 1A).

To place the origin of lignin degradation in the context of geologic time, we performed Bayesian relaxed molecular clock analyses using BEAST (16) and PhyloBayes (7), with fossil-based calibrations at three nodes, including the ancestors of the Boletales, Agaricales, and Ascomycota (5). The mean age of the Agaricomycetes is ~290 Ma (95% highest posterior density interval (hpd) = 222 to 372 Ma), consistent with basidiomycete fossils from the Permian (~260 Ma) and the oldest definitive white rot fossils from the Permian (~260 Ma) and Triassic (~230 Ma) (25).

Organic carbon accumulated at an exceptionally high rate during the Carboniferous and Permian, resulting in the formation of vast coal deposits, derived primarily from lignin (26). A frequently cited explanation for this phenomenon is that decay was inhibited in the anoxic sediments of widespread coastal swamp forests. Our results are consistent with a complementary hypothesis (1), which posits that the sharp decline in the rate of organic carbon burial at the

Table 1. Gene contents in 11 oxidoreductase and 17 CAZyme families in the genomes of 20 Agaricomycotina and 11 other fungi. Species: New genomes: Ad, Auricularia delicata; Cp, Coniochaeta puteana; Da, Dacryopinax sp., Ds, Dichomitus squalens; Fm, Fomitopsis mediterranea; Fp, Fomitopsis pinicola; Gt, Gloeophyllum trabeum; Pu, Punctularia strigosozonata; Sh, Stereum hirsutum; Tm, Tremella mesenterica, Tv, Trametes versicolor; Wc, Wolfiporia cocos. Others: An, Aspergillus niger; Bd, Batrachochytrium dendrobatidis; Cc, Coprinopsis cinerea; Cc, Cryptococcus neoformans; Gc, Cryptococcus paratetrasis; Ha, Helaterobasidion annulosum (has been reclassified as H. irregularis); Lb, Laccaria bicolor; Mg, Malassezia globosa; Mw, Melampsora larici-populina; Pb, Phanerochaete chrysosporium; Pp, Postia placenta; Ps, Picha stiptis; Sc, Schizophyllum commune; Sl, Serpula lacrymans; Sn, Stagonospora nodorum; Sr, Sparobolomyces roseus; Tr, Trichoderma reesei; Um, Ustilago maydis. Ecologies: WR, white rot; BR, brown rot; ECM, mycorrhiza; S, non-wood decay saprotroph; PP, plant pathogen; Y, yeast. Genes: GH, glycoside hydrolases; CAZ, carbohydrate esterases; POD, class II peroxidases; MCO, multicopper oxidases; CRO, copper-containing oxidases; P450, cytochromes P450.

<table>
<thead>
<tr>
<th>Genes</th>
<th>P</th>
<th>H</th>
<th>D</th>
<th>F</th>
<th>M</th>
<th>C</th>
<th>T</th>
<th>Ust</th>
<th>Pucc</th>
<th>Pez</th>
<th>Sc</th>
<th>Ch</th>
<th>Mu</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH3*</td>
<td>0.000</td>
<td>11</td>
<td>15</td>
<td>12</td>
<td>7</td>
<td>12</td>
<td>11</td>
<td>7</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>GH5†</td>
<td>0.307</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>GH6</td>
<td>0.916</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GH7</td>
<td>0.000</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GH10</td>
<td>0.004</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>GH11</td>
<td>0.687</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH12</td>
<td>0.483</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>GH28‡</td>
<td>0.800</td>
<td>8</td>
<td>17</td>
<td>13</td>
<td>19</td>
<td>11</td>
<td>10</td>
<td>11</td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>GH61</td>
<td>0.000</td>
<td>10</td>
<td>16</td>
<td>14</td>
<td>19</td>
<td>13</td>
<td>15</td>
<td>2</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>GH74</td>
<td>0.923</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH43</td>
<td>0.000</td>
<td>4</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>CE1</td>
<td>0.129</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CE16</td>
<td>0.000</td>
<td>5</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>CE5</td>
<td>0.002</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CE8</td>
<td>0.347</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CE12</td>
<td>0.796</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CE15</td>
<td>0.488</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>POD</td>
<td>0.000</td>
<td>8</td>
<td>6</td>
<td>11</td>
<td>17</td>
<td>19</td>
<td>26</td>
<td>12</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MCO</td>
<td>0.000</td>
<td>17</td>
<td>20</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>CRO‡</td>
<td>0.000</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>CDH‡</td>
<td>0.575</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cytb562</td>
<td>0.177</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Oxo</td>
<td>0.026</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>GLP</td>
<td>0.026</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>QRD</td>
<td>0.496</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>DyP</td>
<td>0.000</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HTP</td>
<td>0.000</td>
<td>144</td>
<td>100</td>
<td>200</td>
<td>550</td>
<td>240</td>
<td>430</td>
<td>1435</td>
<td>350</td>
<td>125</td>
<td>75</td>
<td>150</td>
<td>125</td>
</tr>
</tbody>
</table>

*GH3 does not include β-N-acetylhexosaminidase genes. †GH5 includes only models with similarity to endo-1,4-β-D-glucanases and mannan endo-1,4-mannosidases. ‡One model (Fompi1 162677) is a potential pseudogene. ††One model in A. delicata is a potential pseudogene. ‡‡One CDH gene in C. puteana lacks a cyt domain and may not be functional.
Fig. 1. (A) Organismal phylogeny (chronogram) produced with BEAST from a 26-gene data set. Light blue bars are 95% highest posterior density intervals for node ages; mean ages of selected nodes (millions of years) are in parentheses. Blue and red branches indicate significant expansion and contraction, respectively, of PODs inferred using CAFE. Numbers in red following taxon names are POD gene counts. Numbers in red at nodes, separated by commas, are numbers of POD gene copies estimated with CAFE, Notung (with two different edge weight threshold settings), and DrML, respectively. The node labeled A is the ancestor of Agaricomycetes; nodes labeled B are "backbone" nodes in Agaricomycetes (see text). Asterisks indicate nodes that do not receive maximal support in all analyses (see fig. S6 for support values). See Table 1 for full species names. (B) POD gene phylogeny estimated in BEAST with ancestral state reconstructions for manganese-binding site (colored pies) and Trp residues (black and white pies) estimated with BayesTraits. Bars to right of gene IDs indicate presence of functional residues (13). Mean ages for selected nodes in parentheses are followed by 95% highest posterior density ranges.
end of the Permo-Carboniferous was caused, at least in part, by the evolution of lignin decay capabilities in white rot Agaricomycetes.

References and Notes
5. See supplementary materials on Science Online.

Acknowledgments: 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. Also supported by the Assembling the Fungal Tree of Life (AFTOL) project under NSF awards DEB-0732968 (D.S.H.), DEB-0732993 (J.W.S.), and DEB-0732550 (D.J.M.). We thank R. H. Petersen for the strain of A. delicata. The organisms A. delicata and Dacryopinax sp. were obtained in Costa Rica and can only be used for research purposes. Assemblies and annotations of the 12 genomes reported here are available from the JGI fungal portal MycoCosm (http://jgi.doe.gov/fungi) and from DDBJ/EMBL/GenBank under the following accessions: AFV000000000, AEIT00000000, AEUS0000000, AEID00000000, AEJC000000, AEGM00000000, AEGP00000000, AEFD00000000, AEHL00000000, and AEN00000000. Aligned sequence data for organismal and gene family phylogenies and molecular clock analyses, and secretome results are available at DRYAD (http://dx.doi.org/10.5061/dryad.5k347p0).

Supplementary Materials
www.sciencemag.org/cgi/content/full/336/6089/1715/DC1
Materials and Methods
Supplementary Text
Tables S1 to 516
Figs. 51 to 522
References
12 March 2012; accepted 7 May 2012
10.1126/science.1221748