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

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

regulate photosynthetic capacity in green fruit through their regulation of chlorophyll accumulation and chloroplast development and ultimately contribute to sugars that accumulate in ripe fruit.

As in many other plants, two *GLK* genes are present and expressed in tomato, but in fruit, *SlGLK2* mRNA predominates and accumulates in a spatial pattern consistent with chlorophyll biosynthesis and chloroplast development. All *u*/*u* cultivars examined contain a *Slglk2* allele encoding a truncated loss-of-function GLK protein. Our results suggest that breeding selections for the *u* fruit trait that is helpful for harvesting methods may have had an unintended negative impact on fruit quality because suboptimal chloroplasts develop, and consequently, ripe fruit sugar and lycopene levels decrease. Manipulation of *GLK* levels or spatial expression patterns represents an opportunity to recover and enhance production and quality traits in tomato and other crop species.

References and Notes

1. L. Butler, *J. Hered.* **43**, 25 (1952).
2. A. F. Yeager, *Proc. Am. Soc. Hort. Sci.* **33**, 512 (1935).
3. S. M. Kinzer, S. J. Schwager, M. A. Mutschler, *Theor. Appl. Genet.* **79**, 489 (1990).
4. C. M. Rick, *Hilgardia* **42**, 493 (1974).
5. S. D. Tanksley, J. Hewitt, *Theor. Appl. Genet.* **75**, 811 (1988).
6. G. A. Kemp, I. L. Nonnecke, *Can. J. Plant Sci.* **40**, 306 (1960).
7. C. M. Rick, L. Butler, *Adv. Genet. Incorp. Mol. Gen. Med.* **8**, 267 (1956).
8. M. T. Waters, J. A. Langdale, *EMBO J.* **28**, 2861 (2009).
9. S. Hetherington, R. Smillie, W. Davies, *J. Exp. Bot.* **49**, 1173 (1998).
10. M. M. Blanke, F. Lenz, *Plant Cell Environ.* **12**, 31 (1989).
11. S. Carrara, A. Pardossi, G. F. Soldatini, F. Tognoni, L. Guidi, *Photosynthetica* **39**, 75 (2001).
12. A. J. Matas *et al.*, *Plant Cell* **23**, 3893 (2011).
13. T. Manzara, P. Carrasco, W. Gruissem, *Plant Mol. Biol.* **21**, 69 (1993).
14. M. Sugita, W. Gruissem, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7104 (1987).
15. L. A. Wanner, W. Gruissem, *Plant Cell* **3**, 1289 (1991).
16. B. Piechulla, W. Gruissem, *EMBO J.* **6**, 3593 (1987).
17. B. Piechulla, R. E. Glick, H. Bahl, A. Melis, W. Gruissem, *Plant Physiol.* **84**, 911 (1987).
18. D. W. Fitter, D. J. Martin, M. J. Copley, R. W. Scotland, J. A. Langdale, *Plant J.* **31**, 713 (2002).
19. J. A. Langdale, *Plant Cell* **23**, 3879 (2011).
20. M. T. Waters *et al.*, *Plant Cell* **21**, 1109 (2009).
21. M. T. Waters, E. C. Moylan, J. A. Langdale, *Plant J.* **56**, 432 (2008).
22. H. Nakamura *et al.*, *Plant Cell Physiol.* **50**, 1933 (2009).
23. A. Bravo-Garcia, Y. Yasumura, J. A. Langdale, *New Phytol.* **183**, 133 (2009).
24. Y. Yasumura, E. C. Moylan, J. A. Langdale, *Plant Cell* **17**, 1894 (2005).
25. A. H. Paterson *et al.*, *Nature* **335**, 721 (1988).
26. S. D. Tanksley, M. A. Mutschler, C. M. Rick, in *Genetic Maps*, S. J. O'Brien, Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1987), pp. 655–669.
27. Y. Eshed, D. Zamir, *Genetics* **141**, 1147 (1995).
28. Materials and methods are available as supplementary materials on *Science* Online.
29. J. Rohrmann *et al.*, *Plant J.* **68**, 999 (2011).
30. E. M. A. Enfissi *et al.*, *Plant Cell* **22**, 1190 (2010).
31. Y. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9897 (2004).

Acknowledgments: Minimum Information About a Microarray Experiment (MIAME)–compliant microarray data are available at http://ted.bti.cornell.edu. and at http://www.ebi.ac.uk/arrayexpress/(accession E-MEXP-3652). F. Carrari and A. Fernie provided *S. pennellii SlGLK2*, and J. Maloof provided *S. habrochaites SlGLK2* sequences. The U.S. Department of Agriculture (USDA)/National Institute of Food and Agriculture Solanaceae Coordinated Agricultural Project provided potato data. We are grateful to the Tomato Genome Consortium and the SOL Genomics Network for prepublication access to the tomato genome sequence. The *S. pennellii* introgression lines were provided by the C. M. Rick Tomato Genetics Resource Center; the *S. pimpinellifolium* populations were provided by the Instituto de Hortofruticultura Subtropical y Mediterranea "La Mayora," Consejo Superior de Investigaciones Cientificas; and both populations are available by request from the sources. The *AtGLK*-expressing lines were provided by Mendel Biotechnology and Seminis/Monsanto Vegetable Seeds. *SlGLK2*, the corresponding lines, and the F2 10-1 IL x M82 population lines and seeds are available from J.J.G. without restriction. Seminis/Monsanto will make available, upon request, and under a material transfer agreement indicating they are to be used for noncommercial purposes, the following lines: LexA:AtGLK1:p35S:LexA-Gal4; LexA:AtGLK1:pLTP:LexA-Gal4; LexA:AtGLK1:pRbcS:LexA-Gal4; LexA:AtGLK1:pPDS:LexA-Gal4; LexA:AtGLK2:p35S:LexA-Gal4; LexA:AtGLK2:pLTP:LexA-Gal4; LexA:AtGLK2:pRbcS:LexA-Gal4; LexA:AtGLK2:pPDS:LexA-Gal4; plus the T63 control line. Other biological materials are available by request from A.L.T.P. or J.J.G. A.L.T.P., T.H., K.L.-C., R.F.-B., and A.B.B. have filed a provisional U.S. patent application UC #2011-841, "Introduction of wild species *GLK* genes for improved ripe tomato fruit quality," through the University of California. A.L.T.P. and A.B.B. have filed the U.S. patent application #2010/0154078, "Transcription factors that enhance traits in plant organs," through Mendel Biotechnology. Assistance from B. Blanco-Ulate, S. Phothiset, S. Reyes, A. Abraham, L. Gilani, and G. Arellano is gratefully acknowledged. J. Langdale provided helpful advice regarding *GLK* phylogeny and nomenclature. G. Adamson and P. Kysar, Electron Microscopy (EM) Laboratory, University of California Davis Medical Center did the EM work. University of California Discovery and partners funded the pepper analysis and the initial investigations of the *Arabidopsis GLK*s. The Vietnam Education Foundation supported C.N. Fundación Genoma España ESPSOL Project provided partial funding to A.G. USDA–Agricultural Research Service, USDA–National Research Initiative (2007-02773), and NSF (Plant Genome Program IOS-0923312) provided support to J.J.G.

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6089/1711/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S8
References (*32*–*58*)

21 March 2012; accepted 5 June 2012
10.1126/science.1222218

The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes

Dimitrios Floudas,[1] Manfred Binder,[1] Robert Riley,[2] Kerrie Barry,[2] Robert A. Blanchette,[3] Bernard Henrissat,[4] Angel T. Martínez,[5] Robert Otillar,[2] Joseph W. Spatafora,[6] Jagjit S. Yadav,[7] Andrea Aerts,[2] Isabelle Benoit,[8,9] Alex Boyd,[6] Alexis Carlson,[1] Alex Copeland,[2] Pedro M. Coutinho,[4] Ronald P. de Vries,[8,9] Patricia Ferreira,[10] Keisha Findley,[11] Brian Foster,[2] Jill Gaskell,[12] Dylan Glotzer,[1] Paweł Górecki,[13] Joseph Heitman,[11] Cedar Hesse,[6] Chiaki Hori,[14] Kiyohiko Igarashi,[14] Joel A. Jurgens,[3] Nathan Kallen,[1] Phil Kersten,[12] Annegret Kohler,[15] 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 Patyshakuliyeva,[9] Antonis Rokas,[21] Francisco J. Ruiz-Dueñas,[5] Grzegorz Sabat,[22] Asaf Salamov,[2] Masahiro Samejima,[14] 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,[9] Darcy Young,[1] Antonio Pisabarro,[25] Daniel C. Eastwood,[26] Francis Martin,[15] Dan Cullen,[12] Igor V. Grigoriev,[2]* David S. Hibbett[1]*

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*

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 Agaricomycotina (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% or posterior probability ≥0.99) in at least three analyses. The tree topology is consistent with prior analyses and resolves four independent brown rot lineages (Fig. 1A and fig. S6).

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, which have 61 to 148 (average 87) copies of genes encoding CAZymes, representing 14 to 17 gene families, 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 overcome or circumvent lignin; thus, we focus 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 *nop*A gene in *P. chrysosporium* (*5*, *12*). LiPs possess a tryptophan residue on the surface of the enzyme corresponding to Trp^{171} 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 Glu^{35}, Glu^{39}, and Asp^{175} of *P. chrysosporium* MnP1 that function in binding Mn (*13*). VPs possess both the Trp^{171} 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 *Schizophyllum 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 (*5*).

To reconstruct the functional evolution of PODs, we performed Bayesian and ML analyses (*6*, *16*) 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 Trp^{171} 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 Trp^{171} 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 Pucciniomycotina, Ustilaginomycotina, 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 two to 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 (DyP), 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 H_2O_2 production.

The "backbone" nodes in the Agaricomycete phylogeny (labeled "B" in Fig. 1A) are reconstructed as having 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,

[1]Biology Department, Clark University, Worcester, MA 01610, USA. [2]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA. [3]Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108, USA. [4]Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 6098, 13288 Marseille Cedex 9, France. [5]Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040 Madrid, Spain. [6]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA. [7]Environmental Genetics and Molecular Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA. [8]Microbiology and Kluyver Centre for Genomics of Industrial Fermentation, Utrecht University, Padualaan 8, 3584 CH Utrecht, Netherlands. [9]CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, Netherlands. [10]Department of Biochemistry and Molecular and Cellular Biology and Institute of Biocomputation and Physics of Complex Systems, Zaragoza University, 50009, Zaragoza, Spain. [11]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA. [12]USDA Forest Products Laboratory, Madison, WI 53726, USA. [13]Institute of Informatics, Warsaw University, Warsaw, 02-097, Poland. [14]Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. [15]INRA, UMR 1136, INRA-Nancy Université, Interactions Arbres/Microorganismes, 54280 Champenoux, France. [16]Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, Büsgenweg 2, D-37077 Göttingen, Germany. [17]Department of Plant Biology, University of Minnesota, St. Paul, MN 55108, USA. [18]Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontifica Universidad Católica de Chile, Casilla 114-D, Santiago 833-1010, Chile. [19]Department of Applied Chemistry and Microbiology, Viikki Biocenter, P.O. Box 56, Biocenter 1, University of Helsinki, FIN-00014 Helsinki, Finland. [20]Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke Street West, Montreal, Quebec H4B 1R6, Canada. [21]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA. [22]University of Wisconsin Biotechnology Center, Madison, WI 53726, USA. [23]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA. [24]Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Box 7026, Ulls v 26A, 750 07 Uppsala, Sweden. [25]Genetics and Microbiology Research Group, Public University of Navarre, 31006 Pamplona, Spain. [26]College of Science, University of Swansea, Singleton Park, Swansea SA2 8PP, UK.

*To whom correspondence should be addressed. E-mail: ivgrigoriev@lbl.gov (I.V.G.); dhibbett@clarku.edu (D.S.H.)

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, *Coniophora puteana*; Da, *Dacryopinax* sp., Ds, *Dichomitus squalens*; Fm, *Fomitiporia 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*; Cn, *Cryptococcus neoformans*; Cr, *Cryphonectria parasitica*; Ha, *Heterobasidion annosum* (has been reclassified as *H. irregulare*); Lb, *Laccaria bicolor*; Mg, *Malassezia globosa*; Ml, *Melampsora laricis-populina*; Pb, *Phycomyces blakesleeanus*; Pc, *Phanerochaete chrysosporium*; Pp, *Postia placenta*; Ps, *Pichia stipitis*; Sc, *Schizophyllum commune*; Sl, *Serpula lacrymans*; Sn, *Stagonospora nodorum*; Sr, *Sporobolomyces roseus*; Tr, *Trichoderma reesei*; Um, *Ustilago maydis*. **Ecologies:** WR, white rot; BR, brown rot; ECM, mycorrhiza; S, non-wood decay saprotroph; MP, mycoparasite; AP, animal pathogen/parasite; PP, plant pathogen; Y, yeast. **Genes:** GH, glycoside hydrolases; CE, carbohydrate esterases; POD, class II peroxidases; MCO, multicopper oxidases; CRO, copper-radical oxidases; CDH, cellobiose dehydrogenase; Cytb562, cytochrome b562; OXO, oxalate oxidase/decarboxylases; GLP, Fe(III)-reducing glycopeptides; QRD, quinone reductases; DyP, dye-decolorizing peroxidases; HTP, heme-thiolate peroxidases; P450, cytochromes P450. *P* values indicate strength of rejection of model of random diversification in CAFE analyses.

Taxonomy		Basidiomycota																								Ascomycota						
		Agaricomycotina																				Ust		Pucc		Pez				Sc	Ch	Mu
		Agaricomycetes															Dac	Trem														
		Ha	Sh	Pu	Fm	Ad	Tv	Ds	Pc	Pp	Wc	Fp	Gt	Sl	Cp	Lb	Sc	Cc	Da	Tm	Cn	Mg	Um	Ml	Sr	An	Cr	Tr	Sn	Ps	Bd	Pb
Ecology		WR								BR						ECM	WR	S	BR	MP	AP	AP	PP	PP	Y	AP	PP	S	PP	Y	AP	S
		CAZymes																														
Genes	*P*																															
GH3*	0.000	11	15	12	7	12	11	7	9	5	7	12	9	9	12	2	11	7	8	3	7	1	3	3	3	16	15	11	16	7	1	2
GH5†	0.305	7	6	6	6	8	5	5	5	5	4	5	5	8	8	3	3	6	5	0	0	0	1	7	2	4	7	3	5	0	1	1
GH6	0.916	1	1	1	2	2	1	1	1	0	0	0	0	1	2	0	1	5	0	0	0	0	0	0	0	2	2	1	4	0	0	0
GH7	0.000	1	3	5	2	6	4	4	8	0	0	0	0	0	2	0	2	6	0	0	0	0	0	9	0	2	5	2	5	0	0	0
GH10	0.004	2	6	5	4	4	6	5	6	4	4	2	3	1	3	0	5	6	3	0	0	0	2	6	0	2	4	1	7	1	0	0
GH11	0.687	0	1	1	0	3	0	0	1	0	0	0	0	0	0	0	1	6	0	0	0	0	1	0	0	3	4	3	7	0	0	0
GH12	0.483	4	5	2	3	1	5	3	2	2	2	2	2	1	4	3	1	1	1	0	0	0	0	5	0	3	5	2	4	0	0	0
GH28‡	0.000	8	17	13	16	10	11	7	4	7	9	13	10	7	13	7	3	3	6	0	1	0	1	3	0	22	21	4	4	0	0	9
GH61	0.000	10	16	14	13	19	18	15	15	2	2	4	4	5	10	5	22	35	0	0	1	0	0	0	0	7	12	3	29	0	0	0
GH74	0.923	1	2	2	4	1	1	1	4	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	2	1	0	0	0	0
GH43	0.000	4	10	7	6	26	3	7	4	1	1	7	5	1	6	0	12	4	5	0	0	1	2	4	0	7	14	2	10	0	0	0
CE1	0.129	1	1	2	0	3	3	0	4	0	0	0	1	0	0	0	4	3	0	0	0	0	1	0	0	1	2	0	6	0	0	0
CE16	0.000	5	10	8	6	29	7	10	2	5	6	11	6	3	6	3	10	5	4	1	0	0	0	0	2	3	7	2	2	0	0	2
CE5	0.002	0	1	1	0	3	0	0	0	0	0	0	0	0	1	1	2	6	0	0	0	0	4	14	0	5	15	4	11	0	0	0
CE8	0.347	3	4	6	3	3	2	3	2	1	1	2	2	2	2	4	2	0	3	0	0	0	1	5	0	3	4	0	2	0	0	4
CE12	0.796	2	3	0	2	1	0	2	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	2	0	1	0	0	0
CE15	0.488	1	1	2	1	6	2	2	2	1	1	1	1	0	0	0	2	8	1	0	0	0	0	0	0	0	2	1	1	0	0	0
		Oxidoreductases																														
POD	0.000	8	6	11	17	19	26	12	16	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	5	0	0	0
MCO	0.000	17	20	13	11	10	10	13	5	5	5	7	4	6	8	11	6	17	5	4	5	9	6	20	1	14	16	8	9	2	3	5
CRO§	0.000	5	8	9	4	9	9	9	7	3	4	4	2	3	6	11	2	6	3	2	3	1	3	4	3	0	0	1	2	0	3	1
CDH‖	0.575	1	1	1	1	1	1	1	1	0	0	0	1	2	2	0	1	1	0	0	0	0	0	0	0	2	3	0	3	0	0	0
Cytb562	0.177	1	1	0	0	0	1	1	1	0	0	0	0	2	3	0	2	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0
OXO	0.026	3	3	2	3	3	5	5	7	5	4	5	4	3	2	1	5	1	2	0	0	0	0	1	0	2	2	3	3	0	0	0
GLP	0.000	1	11	6	8	1	2	6	3	5	10	10	6	3	11	1	7	0	3	0	0	0	0	2	0	0	0	0	0	0	0	0
QRD	0.496	2	1	3	3	4	1	1	4	1	1	1	3	2	2	2	4	3	1	1	2	0	1	1	1	1	1	1	1	4	1	4
DyP	0.000	1	2	5	3	11	2	1	0	2	0	0	0	0	0	2	0	4	0	1	1	0	0	2	0	0	0	0	0	0	0	0
HTP	0.000	5	10	8	4	16	3	4	3	5	5	4	6	3	2	5	3	8	6	0	0	0	3	17	0	5	2	4	13	0	0	0
P450	0.000	144	215	144	130	249	190	187	149	250	206	190	130	164	238	101	115	139	126	9	13	7	17	28	7	156	125	71	125	10	9	52

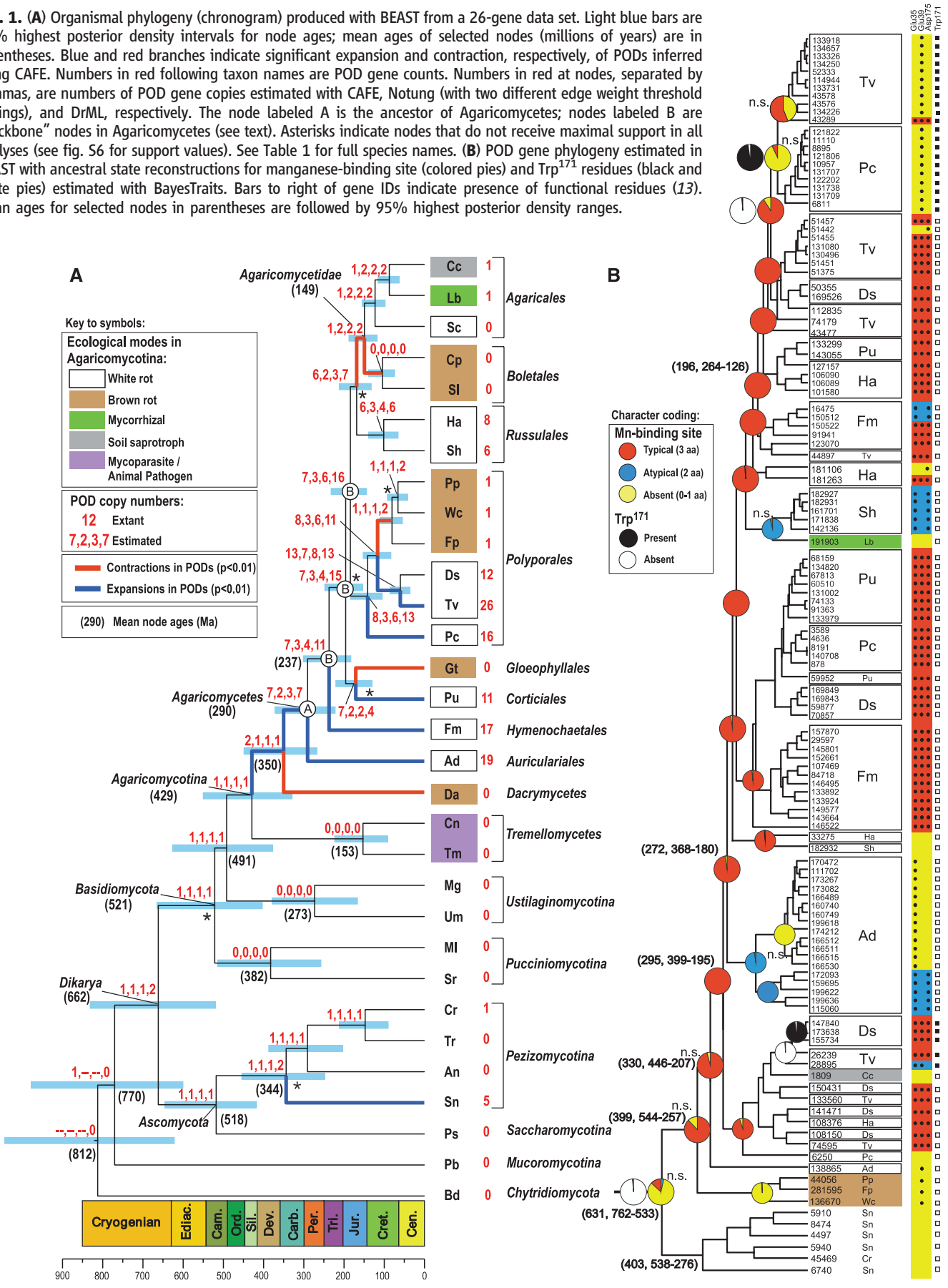
*GH3 does not include β-N-acetylhexosaminidase genes. †GH5 includes only models with similarity to endo-1,4-β-D-glucanases and mannan endo-β-1,4-mannosidases. ‡One model (Fompi1 162677) is a potential pseudogene. §One model in *A. delicata* is a potential pseudogene. ‖One CDH gene in *C. puteana* lacks a cyt domain and may not be functional.

suggesting that these lineages lost PODs as they shifted to a nonligninolytic 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 (millions of years ago) in both BEAST and PhyloBayes analyses [95% highest posterior density interval (hpd) = 222 to 372 Ma], with the mean age of the Agaricomycotina placed at ~430 to 470 Ma (95% hpd = 329 to 557 Ma), consistent with basidiomycete fossils that were not used as calibration points, including hyphae with clamp connections from the Mississippian (*24*) from ~330 Ma. BEAST analyses of the POD genes, calibrated with the split between Ascomycota and Basidiomycota according to the organismal phylogeny, suggest that the first ligninolytic MnP arose at ~295 Ma (95% hpd = 195 to 399 Ma; Fig. 1), which is slightly earlier than (and therefore consistent with) 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

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^{171} 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

1. J. M. Robinson, *Geology* **18**, 607 (1990).
2. D. Martinez *et al.*, *Nat. Biotechnol.* **22**, 695 (2004).
3. D. Martinez *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1954 (2009).
4. D. C. Eastwood *et al.*, *Science* **333**, 762 (2011).
5. See supplementary materials on *Science* Online.
6. A. Stamatakis, *Bioinformatics* **22**, 2688 (2006).
7. N. Lartillot, T. Lepage, S. Blanquart, *Bioinformatics* **25**, 2286 (2009).
8. P. Baldrian, V. Valásková, *FEMS Microbiol. Rev.* **32**, 501 (2008).
9. R. J. Quinlan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15079 (2011).
10. D. Guillén, S. Sánchez, R. Rodríguez-Sanoja, *Appl. Microbiol. Biotechnol.* **85**, 1241 (2010).
11. A. T. Martínez, F. J. Ruiz-Dueñas, M. J. Martínez, J. C. Del Río, A. Gutiérrez, *Curr. Opin. Biotechnol.* **20**, 348 (2009).
12. L. Larrondo, A. González, T. Perez Acle, D. Cullen, R. Vicuña, *Biophys. Chem.* **116**, 167 (2005).
13. F. J. Ruiz-Dueñas *et al.*, *J. Exp. Bot.* **60**, 441 (2009).
14. O. Schmidt, W. Liese, *Holzforschung* **34**, 67 (1980).
15. R. G. FitzJohn, W. P. Maddison, S. P. Otto, *Syst. Biol.* **58**, 595 (2009).
16. A. J. Drummond, A. Rambaut, *BMC Evol. Biol.* **7**, 214 (2007).
17. D. Barker, A. Meade, M. Pagel, *Bioinformatics* **23**, 14 (2007).
18. F. J. Ruiz-Dueñas, E. Fernández, M. J. Martínez, A. T. Martínez, *C. R. Biol.* **334**, 795 (2011).
19. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, *Bioinformatics* **22**, 1269 (2006).
20. D. Durand, B. V. Halldórsson, B. Vernot, *J. Comput. Biol.* **13**, 320 (2006).
21. P. Górecki, G. J. Burleigh, O. Eulenstein, *BMC Bioinformatics* **12** (suppl. 1), S15 (2011).
22. M. Hofrichter, R. Ullrich, M. J. Pecyna, C. Liers, T. Lundell, *Appl. Microbiol. Biotechnol.* **87**, 871 (2010).
23. A. Vanden Wymelenberg *et al.*, *Appl. Environ. Microbiol.* **72**, 4871 (2006).
24. M. Krings, N. Dotzler, G. Galtier, T. N. Taylor, *Mycoscience* **52**, 18 (2011).
25. S. P. Stubblefield, T. N. Taylor, *Bot. Gaz.* **147**, 116 (1986).
26. R. A. Berner, *The Phanerozoic Carbon Cycle: CO_2 and O_2* (Oxford Univ. Press, Oxford, 2004).

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: AFVO00000000, AEIT00000000, AEUS00000000, AEID00000000, AEJJ00000000, AEHC00000000, AFVP00000000, AEGM00000000, AEGX00000000, AEJI00000000, AFVY00000000, and AEHD00000000. 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.5k3t47p0).

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6089/1715/DC1
Materials and Methods
Supplementary Text
Tables S1 to S16
Figs. S1 to S22
References

12 March 2012; accepted 7 May 2012
10.1126/science.1221748

Leucine-tRNA Initiates at CUG Start Codons for Protein Synthesis and Presentation by MHC Class I

Shelley R. Starck,[1] Vivian Jiang,[1] Mariana Pavon-Eternod,[2] Sharanya Prasad,[1] Brian McCarthy,[3] Tao Pan,[2] Nilabh Shastri[1]*

Effective immune surveillance by cytotoxic T cells requires newly synthesized polypeptides for presentation by major histocompatibility complex (MHC) class I molecules. These polypeptides are produced not only from conventional AUG-initiated, but also from cryptic non–AUG-initiated, reading frames by distinct translational mechanisms. Biochemical analysis of ribosomal initiation complexes at CUG versus AUG initiation codons revealed that cells use an elongator leucine-bound transfer RNA (Leu-tRNA) to initiate translation at cryptic CUG start codons. CUG/Leu-tRNA initiation was independent of the canonical initiator tRNA (AUG/Met-tRNA$_i^{Met}$) pathway but required expression of eukaryotic initiation factor 2A. Thus, a tRNA-based translation initiation mechanism allows non–AUG-initiated protein synthesis and supplies peptides for presentation by MHC class I molecules.

In almost all nucleated cells, newly translated polypeptides supply antigenic precursors for loading major histocompatability complex (MHC) class I molecules (*1*). Peptide-loaded MHC class I (pMHC I) molecules reveal the presence of viral or mutated proteins to circulating cytotoxic T cells (CTLs), which bind pMHC I through their T cell receptors to eliminate infected or transformed cells. Antigenic precursors are translated from conventional AUG-initiated open reading frames (ORFs) by the canonical initiator transfer RNA (tRNA), Met-tRNA$_i^{Met}$ (*2*, *3*). Cryptic, non–AUG-initiated ORFs (*4*–*7*) also generate pMHC I during viral infections (*8*–*12*) and oncogenesis (*13*–*15*) by unknown mechanisms. Cryptic CUG start codons can also be decoded with leucine at the initiation stage of translation in mammalian cells (*5*–*7*, *16*). However, this decoding is incompatible with the current model of translation, which indicates that ribosomes are preloaded with initiator Met-tRNA$_i^{Met}$ before recognition of AUG or even non-AUG start codons (*17*).

Our previous study suggested that translation of antigenic precursors from a CUG start codon using leucine represents a distinct initiation pathway (*16*). To determine the molecular mechanism of CUG/leucine initiation, we first screened a series of compounds described as inhibitors of eukaryotic protein synthesis (*18*) using primer extension inhibition analysis (toeprinting) (*19*) of AUG-YL8 and CUG-YL8 mRNA ribosome initiation complexes (fig. S1). We found that NSC119893, which inhibits Met-tRNA$_i^{Met}$ association with eukaryotic initiation factor 2 (eIF2) (*20*), selectively inhibits AUG-YL8 initiation (table S1) in a dose-dependent fashion, whereas CUG-YL8 toeprints were resistant to NSC119893 treatment (Fig. 1A). Structurally unrelated protein synthesis inhibitors, such as suramine and aurin tricarboxylic acid, also inhibited AUG initiation yet enhanced initiation at the CUG start codon (table S1). In contrast, the small molecule acriflavine inhibited CUG initiation more than AUG initiation in a dose-dependent fashion (Fig. 1A). Thus, a structurally diverse set of compounds can distinguish ribosomal recognition of AUG and CUG start codons.

We next assessed the effect of these protein synthesis inhibitors on translation of antigenic precursors in living cells by biochemically analyzing peptides from extracts of cells transfected with the AUG-YL8 or CUG-YL8 plasmids (*5*). As expected, a single peak of antigenic activity—corresponding to the methionine-initiated peptide (MYL8)—was detected from AUG-YL8–transfected cells (Fig. 1B). Yet, CUG-YL8–transfected cells yielded the leucine-initiated peptide (LYL8) as well as the MYL8 peptide, arising from Met-tRNA$_i^{Met}$ "wobble" initiation (Fig. 1C). Although NSC119893 inhibited the expression of MYL8 from AUG and CUG start codons, it did not inhibit decoding of the CUG initiation codon with leucine (LYL8) (Fig. 1, B to D). In contrast, and consistent with the toeprint analysis (Fig. 1A), translation of LYL8 was inhibited by acriflavine, a nucleic acid intercalator, whereas MYL8 initiation from either AUG or CUG codons was unaffected (Fig. 1, B to D). This inhibitor effect was not limited to translation of antigenic precursors, because NSC119893 inhibited AUG-GFP, but not CUG-GFP, expression (fig. S2, A and B). Conversely, CUG-GFP, but not AUG-GFP, expression was inhibited by acriflavine treatment in cultured cells

[1]Division of Immunology and Pathogenesis, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [2]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA. [3]DNA Sequencing Facility, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: nshastri@berkeley.edu