University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

US Department of Energy Publications | US Department of Energy

3-1-2008

The Genome of *Laccaria bicolor* Provides Insights into Mycorrhizal Symbiosis

F. Martin
UMR 1136, INRA-Nancy Universite´, Interactions Arbres/Microorganismes, INRA-Nancy, 54280 Champenoux, France

A. Aerts
US DOE Joint Genome Institute, Walnut Creek, California 94598, USA

D. Ahren
Microbial Ecology, Lund University, SE-223 62 Lund, Sweden

A. Brun
UMR 1136, INRA-Nancy Universite´, Interactions Arbres/Microorganismes, INRA-Nancy, 54280 Champenoux, France

E. G. J. Danchin
Architecture et Fonction des Macromole´cules Biologiques, UMR 6098 CNRS-Universite´s Aix-Marseille I & II, 13288 Marseille Cedex 9, France

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/usdoepub

Part of the Bioresource and Agricultural Engineering Commons

Martin, F.; Aerts, A.; Ahren, D.; Brun, A.; Danchin, E. G. J.; Duchaussoy, F.; Gibon, J.; Kohler, A.; Lindquist, E.; Pereda, V.; Salamov, A.; Shapiro, H. J.; Wuyts, J.; Blaudez, D.; Buee, M.; Brokstein, P.; Canback, B.; Cohen, D.; Courty, P. E.; Coutinho, P. M.; Delaruelle, C.; Detter, J. C.; Deveau, A.; DiFazio, S.; Duplessis, S.; Fraissinet-Tachet, L.; Lucic, E.; Frey-Klett, P.; Fourrey, C.; Feussner, I.; Gay, G.; Grimwood, J.; Hoegger, P. J.; Jain, P.; Kilaru, S.; Labbe, J.; Lin, Y. C.; Legue, V.; Le Tacon, F.; Marmeisse, R.; Melayah, D.; Montanini, B.; Muratet, M.; Nehls, U.; Niculita-Hirzel, H.; Oudot-Le Secq, M. P.; Peter, M.; Quesneville, H.; Rajashekar, B.; Reich, M.; Rouhier, N.; Schmutz, J.; Yin, T.; Chalot, M.; Henrissat, B.; Kues, U.; Lucas, S.; Van de Peer, Y.; Podila, G. K.; Polle, A.; Pukkila, P. J.; Richardson, P. M.; Rouze, P.; Sanders, I. R.; Stajich, J. E.; Tunlid, A.; Tuskan, G.; and Grigoriev, I. V., "The Genome of *Laccaria bicolor* Provides Insights into Mycorrhizal Symbiosis" (2008). *US Department of Energy Publications*. Paper 26.
http://digitalcommons.unl.edu/usdoepub/26

This Article is brought to you for free and open access by the US Department of Energy at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in US Department of Energy Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
F. Martin, A. Aerts, D. Ahren, A. Brun, E. G. J. Danchin, F. Duchaussoy, J. Gibon, A. Kohler, E. Lindquist, V. Pereda, A. Salamov, H. J. Shapiro, J. Wuyts, D. Blaudez, M. Buee, P. Brokstein, B. Canback, D. Cohen, P. E. Courty, P. M. Coutinho, C. Delaruelle, J. C. Detter, A. Deveau, S. DiFazio, S. Duplessis, L. Fraissinet-Tachet, E. Lucic, P. Frey-Klett, C. Fourrey, I. Feussner, G. Gay, J. Grimwood, P. J. Hoegger, P. Jain, S. Kilaru, J. Labbe, Y. C. Lin, V. Legue, F. Le Tacon, R. Marmeisse, D. Melayah, B. Montanini, M. Muratet, U. Nehls, H. Niculita-Hirzel, M. P. Oudot-Le Secq, M. Peter, H. Quesneville, B. Rajashekar, M. Reich, N. Rouhier, J. Schmutz, T. Yin, M. Chalot, B. Henrissat, U. Kues, S. Lucas, Y. Van de Peer, G. K. Podila, A. Polle, P. J. Pukkila, P. M. Richardson, P. Rouze, I. R. Sanders, J. E. Stajich, A. Tunlid, G. Tuskan, and I. V. Grigoriev

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/usdoepub/26

LETTERS

The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis

F. Martin[1], A. Aerts[2], D. Ahrén[3], A. Brun[1], E. G. J. Danchin[4], F. Duchaussoy[1], J. Gibon[1], A. Kohler[1], E. Lindquist[2], V. Pereda[1], A. Salamov[2], H. J. Shapiro[2], J. Wuyts[1,5], D. Blaudez[1], M. Buée[1], P. Brokstein[2], B. Canbäck[3], D. Cohen[1], P. E. Courty[1], P. M. Coutinho[4], C. Delaruelle[1], J. C. Detter[2], A. Deveau[1], S. DiFazio[6], S. Duplessis[1], L. Fraissinet-Tachet[8], E. Lucic[1], P. Frey-Klett[1], C. Fourrey[1], I. Feussner[7], G. Gay[8], J. Grimwood[9], P. J. Hoegger[10], P. Jain[11], S. Kilaru[10], J. Labbé[1], Y. C. Lin[5], V. Legué[1], F. Le Tacon[1], R. Marmeisse[8], D. Melayah[8], B. Montanini[1], M. Muratet[11], U. Nehls[12], H. Niculita-Hirzel[13], M. P. Oudot-Le Secq[1], M. Peter[1,14], H. Quesneville[15], B. Rajashekar[3], M. Reich[1,10], N. Rouhier[1], J. Schmutz[9], T. Yin[16], M. Chalot[1], B. Henrissat[4], U. Kües[10], S. Lucas[2], Y. Van de Peer[5], G. K. Podila[11], A. Polle[10], P. J. Pukkila[17], P. M. Richardson[2], P. Rouzé[5,18], I. R. Sanders[13], J. E. Stajich[19], A. Tunlid[3], G. Tuskan[16] & I. V. Grigoriev[2]

Mycorrhizal symbioses—the union of roots and soil fungi—are universal in terrestrial ecosystems and may have been fundamental to land colonization by plants[1,2]. Boreal, temperate and montane forests all depend on ectomycorrhizae[1]. Identification of the primary factors that regulate symbiotic development and metabolic activity will therefore open the door to understanding the role of ectomycorrhizae in plant development and physiology, allowing the full ecological significance of this symbiosis to be explored. Here we report the genome sequence of the ectomycorrhizal basidiomycete *Laccaria bicolor* (Fig. 1) and highlight gene sets involved in rhizosphere colonization and symbiosis. This 65-megabase genome assembly contains ~20,000 predicted protein-encoding genes and a very large number of transposons and repeated sequences. We detected unexpected genomic features, most notably a battery of effector-type small secreted proteins (SSPs) with unknown function, several of which are only expressed in symbiotic tissues. The most highly expressed SSP accumulates in the proliferating hyphae colonizing the host root. The ectomycorrhizae-specific SSPs probably have a decisive role in the establishment of the symbiosis. The unexpected observation that the genome of *L. bicolor* lacks carbohydrate-active enzymes involved in degradation of plant cell walls, but maintains the ability to degrade non-plant cell wall polysaccharides, reveals the dual saprotrophic and biotrophic lifestyle of the mycorrhizal fungus that enables it to grow within both soil and living plant roots. The predicted gene inventory of the *L. bicolor* genome, therefore, points to previously unknown mechanisms of symbiosis operating in biotrophic mycorrhizal fungi. The availability of this genome provides an unparalleled opportunity to develop a deeper understanding of the processes by which symbionts interact with plants within their ecosystem to perform vital functions in the carbon and nitrogen cycles that are fundamental to sustainable plant productivity.

The 65-megabase genome of *Laccaria bicolor* (Maire) P. D. Orton is the largest sequenced fungal genome published so far[3–7] (Table 1). Although no evidence for large-scale duplications was observed within the *L. bicolor* genome, tandem duplication occurred within multigene families (Supplementary Fig. 4). Transposable elements comprised a higher proportion (21%) than that identified in the other sequenced fungal genomes and may therefore account for the relatively large genome of *L. bicolor* (Supplementary Table 3). Approximately 20,000 protein-coding genes were identified by combined gene predictions (Supplementary Information Section 2). Expression of nearly 80% (~16,000) of the predicted genes was detected in free-living mycelium, ectomycorrhizal root tips or fruiting bodies (Supplementary Table 4) using NimbleGen custom-oligoarrays (Supplementary Information Section 9). Most genes are activated in almost all tissues, whereas other more specialized genes are only activated in some specific developmental stages, such as the free-living mycelium, ectomycorrhizae or the fruiting body (Supplementary Table 5).

Only 14,464 *L. bicolor* proteins (70%) showed sequence similarity (BLASTX, cut-off *e*-value >0.001) to documented proteins. Most homologues were found in the sequenced basidiomycetes *Phanerochaete chrysosporium*[4], *Cryptococcus neoformans*[5], *Ustilago maydis*[6] and *Coprinopsis cinerea*[7] (Supplementary Table 6). The percentage of proteins found in multigene families was related to genome size and was the largest in *L. bicolor* (Fig. 2). This was mainly owing to the expansion of protein family size, but was also because of the larger number of protein families in *L. bicolor* compared with the other basidiomycetes (Supplementary Table 7). Expansion of protein family sizes in *L. bicolor* was prominent in the lineage-specific multigene families. Marked gene family expansions occurred in those genes predicted to have roles in protein–protein interactions (for example, WD40-domain-containing proteins) and in signal transduction

[1]UMR 1136, INRA-Nancy Université, Interactions Arbres/Microorganismes, INRA-Nancy, 54280 Champenoux, France. [2]US DOE Joint Genome Institute, Walnut Creek, California 94598, USA. [3]Microbial Ecology, Lund University, SE-223 62 Lund, Sweden. [4]Architecture et Fonction des Macromolécules Biologiques, UMR 6098 CNRS-Universités Aix-Marseille I & II, 13288 Marseille Cedex 9, France. [5]Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology (VIB), Ghent University, B-9052 Ghent, Belgium. [6]Department of Biology, West Virginia University, Morgantown, West Virginia 26506, USA. [7]Department for Plant Biochemistry, Georg-August-Universität Göttingen, 37077 Göttingen, Germany. [8]Université Lyon 1, UMR CNRS - USC INRA d'Ecologie Microbienne, 69622 Villeurbanne, France. [9]Stanford Human Genome Center, Department of Genetics, Stanford University School of Medicine, 975 California Avenue, Palo Alto, California 94304, USA. [10]Institute of Forest Botany, Georg-August-Universität, 37077 Göttingen, Germany. [11]Department of Biological Sciences, University of Alabama, Huntsville, Alabama 35899, USA. [12]Eberhard-Karls-Universität, Physiologische Oekologie der Pflanzen, 72076 Tübingen, Germany. [13]Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland. [14]Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland. [15]Unité de Recherches en Génomique-Info, INRA-Evry, 91034 Évry Cedex, France. [16]Environmental Science Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA. [17]Department of Biology, The University of North Carolina, Chapel Hill, North Carolina 27599-3280, USA. [18]Laboratoire Associé de l'INRA, Ghent University, B-9052 Gent, Belgium. [19]Department of Plant and Microbial Biology, University of California, Berkeley, California 94720-3102, USA.

This article is a U.S. government work, and is not subject to copyright in the United States.
©2008 Nature Publishing Group

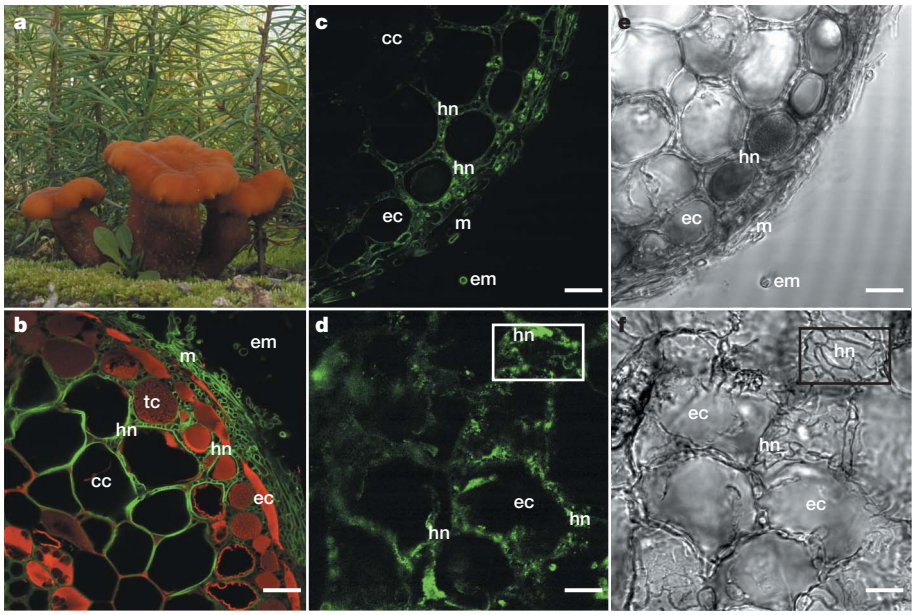

Figure 1 | Ectomycorrhizal symbiosis and the localization of the SSP MISSP7. a, Fruiting bodies of *L. bicolor* colonizing seedlings of Douglas fir (photograph courtesy of D. Vairelles). **b**, Laser-scanning confocal microscopy image of a transverse section of *Pseudotsuga menziesii–L. bicolor* ectomycorrhizae showing extramatrical mycelium (em), the mantle (m) and Hartig net hyphae (hn) between epidermal (ec), tannin (tc) and cortical (cc) root cells. Scale, 10 μm. **c–f**, Immunofluorescent localization of MISSP7. Transverse (**c**, **e**) and longitudinal (**d**, **f**) sections of *P. trichocarpa–L. bicolor* ectomycorrhizae. MISSP7 was detected in the hyphae of the mantle (m) and the Hartig net (hn) ensheathing epidermal cells (ec). Rectangles in **d** and **f** show the finger-like, labyrinthine hyphal system accumulating a large amount of MISSP7. **e**, **f**, Phase contrast images. Scale, 10 μm.

mechanisms (Supplementary Table 7). Two new classes of GTPase α genes were found and may be candidates for the complex communication that must occur between the mycobiont and its host plant during mycorrhizae establishment (Supplementary Table 8). Several transcripts coding for expanded and lineage-specific gene families were upregulated in symbiotic and fruiting body tissues, suggesting a role in tissue differentiation (Supplementary Tables 5 and 9).

In our analysis of annotated genes, in particular that of paralogous gene families, we highlighted processes that may be related to the biotrophic and saprotrophic lifestyles of *L. bicolor*. Twelve predicted proteins showed a similarity to known haustoria-expressed secreted proteins of the basidiomycetous rusts *Uromyces fabae*[8] and *Melampsora lini*[9], which are involved in pathogenesis (Supplementary Table 10). Out of the 2,931 proteins predicted to be secreted by *L. bicolor*, most (67%) cannot be ascribed a function, and 82% of these predicted proteins are specific to *L. bicolor*. Within this set, we found a large number of genes that encode cysteine-rich products that have a predicted size of <300 amino acids. Of these 278 SSPs, 69% belong to multigene families, but only nine groups comprising a total of 33 SSPs co-localized in the genome (Supplementary Fig. 5). The structure of two of these clusters is shown in Supplementary Fig. 6. Other SSPs are scattered all over the genome, and we found no correlation between SSP and transposable element genome localization (Supplementary Fig. 5). Transcript profiling revealed that the expression of several SSP genes is specifically induced in the symbiotic interaction (Table 2 and Supplementary Fig. 10). Five of the 20 most highly upregulated fungal transcripts in ectomycorrhizal root tips code for SSPs (Supplementary Table 5). These mycorrhiza-induced cysteine-rich SSPs (MISSPs) belong to *L. bicolor*-specific orphan gene families. Within the MISSPs, we found a family of secreted proteins with a CFEM domain (INTERPRO IPR014005) (Supplementary Figs 7 and 8), as previously identified in the plant pathogenic fungi *M. lini*[9] and *Magnaporthea grisea*[10] (Supplementary Table 10), and proteins with a gonadotropin (IPR0001545) or snake-toxin-like (SSF57302) domains related to the cysteine-knot domain. Expression of several SSPs was downregulated in ectomycorrhizal root tips (cluster E in Supplementary Fig. 10), suggesting a complex interplay between these secreted proteins in the symbiosis interaction.

The rich assortment of MISSPs may therefore act as effector proteins to manipulate host cell signalling or to suppress defence pathways during infection, as suggested for pathogenic rusts[8,9], smuts[6] (*U. maydis*) and *Phytophthora*[11] species. To have a role in symbiosis development, MISSPs should be expressed in *L. bicolor* hyphae colonizing the root tips. To test this assertion, we determined the tissue distribution of the mycorrhiza-induced cysteine-rich SSP of 7 kDa (MISSP7) (JGI identification number 298595) showing the highest induction in ectomycorrhizal tips (Table 2 and Supplementary Table 5). Two peptides, one of which is located in the amino-terminal and the other in the carboxy-terminal part of the mature protein, were selected as antigens for the production of anti-MISSP7 antibodies. The selected peptides were not found in the deduced protein sequences of other *L. bicolor* gene models, nor in the *Populus trichocarpa* genome[12]. MISSP7

Table 1 | Genome characteristics of *L. bicolor* and other basidiomycetes

Genome characteristics	*L. bicolor*	*C. cinerea*[7]	*P. chrysosporium*[4]	*C. neoformans*[5]	*U. maydis*[6]
Strain	S238N-H82	Okayama7#130	RP78	H99	521
Sequencing institution	JGI	Broad	JGI	Broad	Broad
Genome assembly (Mb)	64.9	37.5	35.1	19.5	19.7
GC content (%)	46.6	51.6	53.2	48.2	54
Number of protein-coding genes	20,614	13,544	10,048	7,302	6,522
Coding sequence <300 bp	2,191	838	163	313	58
Average gene length (bp)	1,533	1,679	1,667	1,828	1,935
Average coding sequence length (bp)	1,134	1,352	1,366	1,502	1,840
Average exon length (nt)	210.1	251	232	253	1,051
Average intron length (nt)	92.7	75	117	66	127

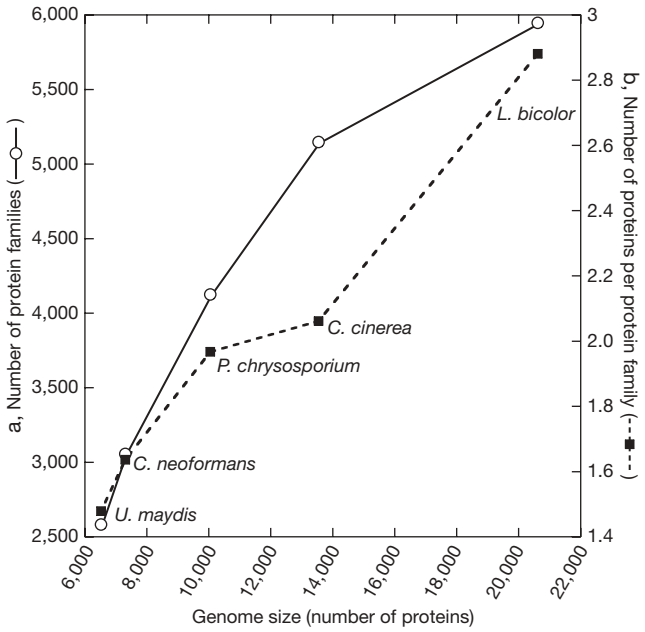

Figure 2 | Expansion of protein families in *L. bicolor*. **a**, Relationship between genome size and the number of protein families. **b**, Relationship between genome size and protein family sizes in five sequenced basidiomycetes. Protein sequences predicted from the genome sequences of *L. bicolor*, *C. cinerea*, *P. chrysosporium*, *C. neoformans* and *U. maydis* were clustered into families using the TRIBE-MCL algorithm (see Supplementary Information Section 5 for details).

localization in *L. bicolor–P. trichocarpa* ectomycorrhizal root tips by indirect immunofluorescence is illustrated in Fig. 1 and Supplementary Fig. 11. Control images in which the ectomycorrhizae sections were obtained by replacing primary anti-MISSP7 antibodies with pre-immune immunoglobulin (Ig)G are shown in Supplementary Fig. 12. For cases in which ectomycorrhizae were treated with anti-MISSP7 antibody followed by fluorescent-labelled secondary antibody, fluorescence was localized in the hyphae colonizing short roots (Fig. 1 and Supplementary Fig. 11) and was not detected in the free-living mycelium (Supplementary Fig. 12). Although MISSP7 was detected in the hyphal mantle layers ensheating the root tips, the protein mainly accumulated in the finger-like, labyrinthine branch hyphal system (Hartig net), which provides a very large area of contact between cells of the two symbionts. It accumulated in the cytosol and cell wall of the fungal cells. The MISSP7 protein could therefore interact with the plant components after secretion. MISSP7 shares no sequence similarity or protein motif with other SSPs.

Comparison of the MISSP sequences did not reveal a specific conserved motif that could potentially contribute to their function or to targeting to the host cell, such as the RXLR motif[11] of phytopathogenic *Phytophthora* or the malaria parasite. SSPs with upregulated expression in fruiting bodies (Supplementary Table 5 and Supplementary Fig. 10) may have a role in the differentiation of the sexual tissues and/or the aggregation of sporophore tissues. Interestingly, there is a large set of SSP genes showing significant changes in gene expression in both ectomycorrhizal root tips and fruiting bodies (cluster A in Supplementary Fig. 10), suggesting that both developmental processes recruit similar gene networks (for example, those involved in hyphal aggregation).

Host trees are able to harness the formidable web of mycorrhizal hyphae (which permeates the soil and decaying leaf litter) for their nutritional benefit. A process that is pivotal to the success of ectomycorrhizal interactions is therefore the equitable exchange of nutrients between the symbiont and its host plant[1,2,13]. A comparison with other basidiomycetes (Supplementary Table 12) revealed that the total number of predicted transporters is larger in *L. bicolor* compared with *C. cinerea* and *P. chrysosporium*. Interestingly, *L. bicolor* has multiple ammonia transporters, although it encodes a single nitrate permease. Ammonia is arguably the most important inorganic nitrogen source for ectomycorrhizal fungi[14]. One of the ammonia transporters (*AMT2.2*), for instance, is greatly upregulated in ectomycorrhizae (Supplementary Table 5). Therefore, *L. bicolor* shows an increased genetic potential in terms of nitrogen uptake compared with other basidiomycetes. These capabilities are consistent with *L. bicolor* being exposed to a range of nitrogen sources from the decay of organic matter[15].

Although the *L. bicolor* genome contains numerous genes coding for key hydrolytic enzymes, such as proteases and lipases, we observed an extreme reduction in the number of enzymes involved in the degradation of plant cell wall (PCW) oligosaccharides and polysaccharides. Glycoside hydrolases, glycosyltransferases, polysaccharide lyases, carbohydrate esterases and their ancillary carbohydrate-binding modules were identified using the carbohydrate-active enzyme (CAZyme) classification (http://www.cazy.org/). A comparison of the *L. bicolor* candidate CAZymes with fungal phytopathogens

Table 2 | Changes in expression of transcripts coding for MISSPs

Protein identification (JGI *Laccaria* database)	Family size	Length (amino acids)	Transcript concentration (FLM)	*P. menziesii* ECM/FLM ratio (fold)	*P. trichocarpa* ECM/FLM ratio (fold)	Features
298595	sc	68	ND	21,877	12,913	MISSP7
333839	5	129	ND	7,844	1,931	Glycosyl phosphatidylinositol (GPI)-anchored
298667	2	70	ND	1,906	1,407	
332226	8	181	43	847	780	CFEM domain (INTERPRO IPR014005)
311468	2	59	ND	191	ND	
295737	8	288	131	171	252	
334759	sc	101	ND	109	18	
395403	4	121	24	103	93	
333423	9	120	6	102	72	Gonadotropin domain (IPR0001545)
312262	4	106	85	69	53	
295625	4	199	325	66	48	
325402	8	238	310	49	74	Snake toxin-like (SSF57302)
316998	sc	56	137	29	57	
333197	3	148	266	17	8	
327918	2	154	763	13	4	Homologue in *C. cinerea*
307956	sc	74	336	13	90	Whey acidic domain (IPR008197)
327246	sc	194	1,025	10	18	Homologue in *C. cinerea*
303550	5	98	1,365	10	14	
300377	2	291	5,499	10	8	
293250	sc	224	127	9	10	Homologue in *C. cinerea*
298648	sc	64	1,108	8	12	
298646	2	73	1,028	7	14	
293729	3	210	3,000	7	7	

Transcript profiling was performed on free-living mycelium (FLM) and ectomycorrhizal root tips (ECM) of poplar (*P. trichocarpa*) and Douglas fir (*P. menziesii*). See Supplementary Information Section 9 for details. Abbreviations: ND, not detected; sc, single copy.

confirms the adaptation of its enzyme repertoire to symbiosis and reveals the strategy used for the interaction with the host (Supplementary Tables 13 and 14). The reduction in PCW CAZymes affects almost all glycoside hydrolase families, culminating in the complete absence of several key families. For instance, there is only one candidate cellulase (glycoside hydrolase 5, GH5) appended to the sole fungal cellulose-binding module (CBM1) found in the genome, and no cellulases from families GH6 and GH7 (Supplementary Table 14). Similar reductions or loss of hemicellulose- and pectin-degrading enzymes were also noted. These observations suggest that the inventory of *L. bicolor* PCW-degrading enzymes underwent massive gene loss as a result of its adaptation to a symbiotic lifestyle, and that this species is now unable to use many PCW polysaccharides as a carbon source, including those found in soil and leaf litter. The remaining small set of secreted CAZymes with potential action on plant polysaccharides (for example, GH28 polygalacturonases) is probably required for cell wall remodelling during fungal tissue differentiation because their expression was upregulated in both fruiting bodies and ectomycorrhizae (Supplementary Table 15 and Supplementary Fig. 13). By contrast, transcripts coding for proteins with an expansin domain were only induced in ectomycorrhizae, suggesting they may be used by *L. bicolor* for penetrating into the root apoplastic space.

To survive before its mycorrhizal association with its host, *L. bicolor* seems to have developed a capacity to degrade non-plant (for example, animal and bacterial) oligosaccharides and polysaccharides; this is suggested by retention of CAZymes from families GH79, polysaccharide lyase 8 (PL8), PL14 and GH88 (Supplementary Table 14). Interestingly, there is no invertase gene in the *L. bicolor* genome, implying that this fungus is unable to use sucrose directly from the plant. This is consistent with earlier observations[16] that *L. bicolor* depends on its host plant to provide glucose in exchange for nitrogen. We also noticed an expansion of CAZymes involved in the fungal cell wall biosynthesis and rearrangement, almost entirely owing to an increased number of putative chitin synthases and enzymes acting on β-glucans (Supplementary Table 14). Several of the corresponding genes are up- or down-regulated in developmental processes requiring cell wall alterations such as formation of fruiting bodies or mycorrhizae (Supplementary Table 15 and Supplementary Fig. 13).

Ectomycorrhizal fungi have an important role in mobilizing nitrogen from well-decomposed organic matter[2,15]. The hyphal network permeating the soil might therefore be expected to express a wide diversity of proteolytic enzymes. The total number of secreted proteases (116 members) identified (Supplementary Fig. 14) is relatively large compared with that in other sequenced saprotrophic basidiomycetes, such as *C. cinerea* and *P. chrysosporium*. Secreted aspartyl-, metallo- and serine-proteases may have a role in degradation of decomposing litter[15], confirming that *L. bicolor* has also the ability to use nitrogen of animal origin, as suggested previously[17]. They may also have a role in developmental processes because the expression of several secreted proteases is up- or down-regulated in fruiting bodies and ectomycorrhizal root tips (Supplementary Table 16). Mycelial mats formed by *L. bicolor* hyphae colonizing organic matter therefore possess the ability to degrade proteins from decomposing leaf litter.

Our analysis of the gene space reveals a multi-faceted mutualistic biotroph equipped to take advantage of transient occurrences of high-nutrient niches (living host roots and decaying soil organic matter) within a heterogeneous, low-nutrient environment. The availability of genomes from mutualistic, saprotrophic[4] and pathogenic[6] fungi, as well as from the mycorrhizal tree *P. trichocarpa*[12], now provides an unparalleled opportunity to develop a deeper understanding of the processes by which fungi colonize wood and soil litter, and also interact with living plants within their ecosystem, to perform vital functions in the carbon and nitrogen cycles[2] that are fundamental to sustainable plant productivity.

METHODS SUMMARY

Genomic sequence. Scaffolds and assemblies for all genomic sequences generated by this project are also available from the Joint Genome Institute (JGI) portal (http://genome.jgi-psf.org/Lacbi1/Lacbi1.download.ftp.html). A genome browser is available from JGI (http://www.jgi.doe.gov/laccaria). BLAST search of the genome is available at JGI (http://www.jgi.doe.gov/laccaria) and INRA LaccariaDB (http://mycor.nancy.inra.fr/IMGC/LaccariaGenome/).

Predicted gene models. Consensus gene predictions, produced by combining several different gene predictors, are available from JGI (http://www.jgi.doe.gov/laccaria) as General Feature Format (GFF) files. These gene models can also be accessed from the Genome Browser in the JGI *L. bicolor* portal (http://www.jgi.doe.gov/laccaria).

Gene annotations. Tables compiling KEGG, PFAM, KOG and best BLAST hits for predicted gene models, transposable element and CAZyme data, as well as Tribe-MCL gene families, are available from INRA LaccariaDB (http://mycor.nancy.inra.fr/IMGC/LaccariaGenome/).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 August; accepted 20 December 2007.

1. Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis* 2nd edn (Academic, London, 1996).
2. Read, D. J. & Perez-Moreno, J. Mycorrhizas and nutrient cycling in ecosystems — a journey towards relevance? *New Phytol.* **157,** 475–492 (2003).
3. Galagan, J. E., Henn, M. R., Ma, L. J., Cuomo, C. A. & Birren, B. Genomics of the fungal kingdom: insights into eukaryotic biology. *Genome Res.* **15,** 1620–1631 (2005).
4. Martinez, D. *et al.* Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnol.* **22,** 695–700 (2004).
5. Loftus, B. J. *et al.* The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans. Science* **307,** 1321–1324 (2005).
6. Kämper, J. *et al.* Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis. Nature* **444,** 97–101 (2006).
7. *Coprinus cinereus* database ⟨http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/Home.html⟩.
8. Wirsel, S. G. R., Voegele, R. T. & Mendgen, K. W. Differential regulation of gene expression in the obligate biotrophic interaction of *Uromyces fabae* with its host *Vicia faba. Mol. Plant Microb. Int.* **14,** 1319–1326 (2001).
9. Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A. & Ellis, J. G. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* **18,** 243–256 (2006).
10. Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends Bioch. Sci.* **28,** 118–121 (2003).
11. Kamoun, S. A. Catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.* **44,** 41–60 (2006).
12. Tuskan, G. A. *et al.* The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313,** 1596–1604 (2006).
13. Martin, F., Kohler, A. & Duplessis, S. Living in harmony in the wood underground: ectomycorrhizal genomics. *Curr. Opin. Plant Biol.* **10,** 204–210 (2007).
14. Chalot, M., Blaudez, D. & Brun, A. Ammonia: a candidate for nitrogen transfer at the mycorrhizal interface. *Trends Plant Sci.* **11,** 263–266 (2006).
15. Lindahl, B. D. *et al.* Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *New Phytol.* **173,** 611–620 (2007).
16. Nehls, U., Grunze, N., Willmann, M., Reich, M. & Küster, H. Sugar for my honey: carbohydrate partitioning in ectomycorrhizal symbiosis. *Phytochemistry* **68,** 82–91 (2007).
17. Klironomos, J. N. & Hart, M. M. Animal nitrogen swap for plant carbon. *Nature* **410,** 651 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The genome sequencing of *L. bicolor* was funded by the US Department of Energy's Office of Science, Biological and Environmental Research Program and the by University of California, the Lawrence Berkeley National Laboratory, the Lawrence Livermore National Laboratory, and the Los Alamos National Laboratory. Annotation and transcriptome analyses were supported by INRA, the US Department of Energy, the US National Science Foundation, the European Commission, the Région Lorraine and the Swedish Research Council. We thank S. Rombauts, L. Sterck, K. Vandepoele, G. Werner and his colleagues, S. Pitluck and K. Zhou, B. Hilselberger and J. Gérard for their assistance. F.M. thanks N. Talbot for critical reading of an early draft of the manuscript.

Author Contributions A.A., D.A., A.B., E.G.J.D., F.D., J.G., A. K., E.L., V.P., A.S., H.J.S. and J.W. contributed equally to this work as second authors. M.C., B.H., U.K., S.L., Y.V.d.P., G.K.P., A.P., P.J.P., P.M.R., P.R., I.R.S., J.E.S., A.T., G.T. and I.V.G. contributed equally to this work as senior authors. Individual contributions were as follows. For the sequencing project, F.M. and G.T. initiated the sequencing project in the wake of the *Populus* genome project. F.M., I.V.G. and P.M.R. developed and coordinated the sequencing, annotation and transcriptome projects. S.L. and sequencing staff at

JGI performed the shotgun sequencing. H.J.S. and his staff at JGI performed the JAZZ assembly of the nuclear and mitochondrial genome. J.G. and J.S. performed the Arachne assembly, closed up repetitive gaps and fixed missasembled regions. A.A., A.S., J.W., M.M., P.R., Y.V.d.P. and I.V.G. did the *ab initio* annotation of protein-coding gene models. A.K., E.L., P.B., C.D., A.D., J.C.D., M.P., G.K.P., A.T. and F.M. provided expressed sequence tag/cDNA information for the *ab initio* and manual annotation. Genome statistics was performed by D.A., F.D., J.W., P.R., I.V.G. and F.M. A.A. and I.V.G., and F.D. and M.P.O.-L.S. were responsible for database design and maintenance at JGI and INRA, respectively. For genome analysis, D.C., M.P. and G.K.P. were responsible for DNA extraction and purification; D.A., F.D., Y.C.L., B.R., Y.V.P., P.R., A.S., J.E.S., A.T., I.V.G. and F.M. for comparative genome analysis; B.C., D.A. and A.T. for genome synteny; J.L., T.Y., G.T., F.M. and F.L.T. for construction of the genetic map; S.D.F. for single nucleotide polymorphisms; A.K., F.D. and F.M. for DNA arrays; A.D., B.C., P.F.K. and F.M. for high-GC sequences; J.W., P.R. and F.M. for tRNA, snRNA and rDNA; E.G.J.D., P.M.C., B.H. for carbohydrate active enzymes; A.D. and P.F.-K. for carbohydrate metabolism; J.G., P.H., U.K. and F.M. for cell wall proteins and secretome; L.F.-T., G.G., D.M. and R.M. for cytoskeleton and motor proteins; M.R., I.F. and A.P. for lipid metabolism; and H.N.-H., U.K. and I.R.S. for mating type genes. F.M. was responsible for genome analysis of the mitochondrion; P.E.C., P.H., M.B., S.K. and U.K. for the multi-copper oxidases; M.B., P.E.C. and F.M. for the proteases; N.R. for the redox genes; S.D., P.J. and G.K.P. for the signal transduction pathway; A.B., D.B., C.F., E.L., B.M., U.N., V.P. and M.C. for transporters; J.L., J.W., P.R., F.M. and H.Q. for transposable elements; V.P., J.G., A.B. and V.L. for immunolocalization of SSP; and J.E.S. and P.J.P. for the *Coprinopsis* genome. F.M.organized co-ordination between different groups. F.M. wrote and edited the paper with input from I.V.G., A.T., D.A., P.R., M.C. and B.H.

Author Information The whole-genome shotgun project has been deposited at GenBank/EMBL/DDBJ under project accession number ABFE00000000. The version described in this paper including assembly and annotation is the first version, ABFE01000000. The complete expression data set is available as a series under accession number GSE9784 at the Gene Expression Omnibus at NCBI (http://www.ncbi.nlm.nih.gov/geo/). Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Correspondence and requests for materials should be addressed to F.M. (fmartin@nancy.inra.fr).

METHODS

Genome sequencing. The haploid genome of the strain S238N-H82 from *L. bicolor* (Maire) P. D. Orton was sequenced with the use of a whole-genome shotgun strategy. All data were generated by paired-end sequencing of cloned inserts using Sanger technology on ABI3730xl sequencers. Supplementary Table 1 gives the number of reads obtained per library.

Genome assembly. The data were assembled using release 1.0.1b of JAZZ, a JGI whole-genome shotgun assembler. On the basis of the number of alignments per read, the main genome scaffolds were at a depth of 9.88. The amount of sequence in the unplaced reads was 6.5 Mb, which is sufficient to cover the main-genome gaps to a mean depth of 9.9. A total of 64.9 Mb are captured in the scaffold assembly (Supplementary Table 2).

Genome annotation. Gene models were predicted using FgenesH[18], homology-based FgenesH$^+$ (ref. 18) and Genewise[19], as well as EuGène[20] and TwinScan[21], and alignments of several complementary DNA resources (Supplementary Information Section 3). The JGI pipeline selected a best representative gene model for each locus on the basis of expressed sequence tag support and similarity to known proteins from other organisms, and predicted 20,614 protein-coding gene models. All predicted genes were annotated using Gene Ontology[22], eukaryotic clusters of orthologous groups[23] and KEGG pathways[24]. Protein domains were predicted using InterProScan[25]. Signal peptides were predicted in 2,931 *L. bicolor* proteins by both the hidden Markov and the neural network algorithms of SignalP[26]. After eliminating predicted transmembrane proteins and removal of transposable element fragments, we selected 278 cysteine-rich secreted proteins with a size of $<$300 amino acids. Gene families were built from proteins in *L. bicolor*, *C. cinerea*, *P. chrysosporium*, *C. neoformans* and *U. maydis* using Tribe-MCL tools[27] with default settings.

Indirect immunofluorescent localization of MISSP7. The peptides LRALGQASQGGDLHR and GPIPNAVFRRVPEPNF located in the N-terminal and C-terminal parts of the MISSP7 sequence (without the signal peptide) were synthesized and used as antigens for the generation of antibodies in rabbits according to the manufacturer's procedures (Eurogentec). The anti-MISSP7IgG fraction was purified using the MAbTrap kit (GE Healthcare) according to the manufacturer's recommendations. Subsequently, the IgG-containing fraction was desalted using a HiTrap desalting column (GE Healthcare). The concentration of purified IgG from pre-immune serum was determined by Bradford assay using a Bio-Rad protein assay. The final concentration of anti-MISSP7 IgG was 0.16 mg ml^{-1}. Immunolocalization was performed essentially as described in refs 28 and 29, with slight modifications (Supplementary Section 10).

Gene expression. Average expression levels of genes in different tissues and conditions were analysed using CyberT statistical framework (http://www.igb.uci.edu/servers/cybert/) and hierarchical clustering with EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/) (Supplementary Information Section 8).

18. Salamov, A. & Solovyev, V. *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res.* **10,** 516–522 (2000).
19. Birney, E., Clamp, M. & Durbin, R. GeneWise and genomewise. *Genome Res.* **14,** 988–995 (2004).
20. Schiex, T., Moisan, A. & Rouzé, P. in: *Computational Biology* vol. 2066 (eds, Gascuel, O. & Sagot, M. F.) Lecture Notes in Computer Science, publisher: Springer, Berlin/Heidleberg, Germany 111–125.
21. Tenney, A. E. *et al.* Gene prediction and verification in a compact genome with numerous small introns. *Genome Res.* **14,** 2330–2335 (2004).
22. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genet.* **25,** 25–29 (2000).
23. Koonin, E. V. *et al.* A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **5,** R7.1–R7.28 (2004).
24. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32,** D277–D280 (2004).
25. Zdobnov, E. M. & Apweiler, R. InterProScan — an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17,** 847–848 (2001).
26. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP, and related tools. *Nature Protocols* **2,** 953–971 (2007).
27. Enright, A. J., van Dongen, S. & Duzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucl. Acids Res.* **30,** 1575–1584 (2002).
28. Blancaflor, E. B., Zhao, L. & Harrison, M. J. Microtubule organization in root cells of *Medicago truncatula* during development of an arbuscular mycorrhizal symbiosis with *Glomus versiforme*. *Protoplasma* **217,** 154–165 (2001).
29. Harrison, M. J., Dewbre, G. R. & Liu, J. Y. A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* **14,** 2413–2429 (2002).