Leucoagaricus gongylophorus produces diverse enzymes for the degradation of recalcitrant plant polymers in leaf-cutter ant fungus gardens

George M. Weinstock
Washington University School of Medicine in St. Louis
et al

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
http://digitalcommons.wustl.edu/open_access_pubs/2003

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Leucoagaricus gongylophorus Produces Diverse Enzymes for the Degradation of Recalcitrant Plant Polymers in Leaf-Cutter Ant Fungus Gardens


Updated information and services can be found at: [http://aem.asm.org/content/79/12/3770](http://aem.asm.org/content/79/12/3770)

**SUPPLEMENTAL MATERIAL**

Supplemental material

**REFERENCES**

This article cites 56 articles, 19 of which can be accessed free at: [http://aem.asm.org/content/79/12/3770#ref-list-1](http://aem.asm.org/content/79/12/3770#ref-list-1)

**CONTENT ALERTS**

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](http://aem.asm.org/content/79/12/3770#ref-list-1)

Information about commercial reprint orders: [http://journals.asm.org/site/misc/reprints.xhtml](http://journals.asm.org/site/misc/reprints.xhtml)

To subscribe to another ASM Journal go to: [http://journals.asm.org/site/subscriptions/](http://journals.asm.org/site/subscriptions/)
**Leucoagaricus gongylophorus** Produces Diverse Enzymes for the Degradation of Recalcitrant Plant Polymers in Leaf-Cutter Ant Fungus Gardens


Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin, USA; Department of Energy Great Lakes Bioenergy Research Center, University of Wisconsin—Madison, Madison, Wisconsin, USA; Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA; Smithsonian Tropical Research Institute, Balboa, Ancon, Panama; Department of Energy Joint Genome Institute, Walnut Creek, California, USA; Department of Biology, Emory University, Atlanta, Georgia, USA; Roche Diagnostics, Indianapolis, Indiana, USA; The Genome Center, Washington University School of Medicine, St. Louis, Missouri, USA

**Abstract**

Plants represent a large reservoir of organic carbon comprised primarily of recalcitrant polymers that most metazoans are unable to deconstruct. Many herbivores gain access to nutrients in this material indirectly by associating with microbial symbionts, and leaf-cutter ants are a paradigmatic example. These ants use fresh foliar biomass as manure to cultivate gardens composed primarily of **Leucoagaricus gongylophorus**, a basidiomycetous fungus that produces specialized hyphal swellings that serve as a food source for the host ant colony. Although leaf-cutter ants are conspicuous herbivores that contribute substantially to carbon turnover in Neotropical ecosystems, the process through which plant biomass is degraded in their fungus gardens is not well understood. Here we present the first draft genome of **L. gongylophorus**, and, using genomic and metaproteomic tools, we investigate its role in lignocellulose degradation in the gardens of both *Atta cephalotes* and *Acromyrmex echinatior* leaf-cutter ants. We show that *L. gongylophorus* produces a diversity of lignocellulases in ant gardens and is likely the primary driver of plant biomass degradation in these ecosystems. We also show that this fungus produces distinct sets of lignocellulases throughout the different stages of biomass degradation, including numerous cellulases and laccases that likely play an important role in lignocellulose degradation. Our study provides a detailed analysis of plant biomass degradation in leaf-cutter ant fungus gardens and insight into the enzymes underlying the symbiosis between these dominant herbivores and their obligate fungal cultivar.

**The ecology and evolution of metazoans are shaped, at least in part, by microbial symbionts (1, 2).** The formation of symbiotic associations with beneficial microbes that confer novel physiological capacities on their hosts has been described as a form of evolutionary innovation and even allows some animals to occupy ecological niches that would otherwise be unavailable (1–4). The ability of animals to gain access to nutrients in plant biomass, an abundant energy source in terrestrial ecosystems, is facilitated by the formation of symbioses with microbes, since plant cell walls are largely composed of recalcitrant polymers that most animals are unable to deconstruct (5, 6). For herbivorous mammals, especially ruminants, the role of microbes in mediating plant biomass degradation has been described (6, 7). In contrast, detailed studies of lignocellulolytic microbes associated with insect herbivores, the most species-diverse and dominant plant-feeding animals in most ecosystems, are limited.

Leaf-cutter ants are hallmark examples of insect herbivores that gain access to nutrients in plant material through symbioses with microbes (Fig. 1A and B). Through the cultivation of fungus-bacterium “gardens” on fresh foliar material, these ants are able to access nutrients in plant biomass that would otherwise be unavailable (8). Despite the central importance of plant biomass degradation to the ecology and evolution of leaf-cutter ants, this process is poorly understood. It has long been assumed that the dominant basidiomycetous cultivar in these gardens, **Leucoagaricus gongylophorus**, can readily degrade plant biomass (9). However, reports that *L. gongylophorus* can grow on cellulose in pure culture have varied (10, 11), and the full extent of this organism’s lignocellulolytic capabilities has not been determined. Culture-based studies have shown that a diversity of bacteria, yeasts, and microfungi can also be found in ant gardens (12–14), and recent culture-independent investigations have revealed that a distinct community of bacteria resides in the fungus gardens of leaf-cutter ants (15, 16). Although the genetic potential of bacteria in *Atta* gardens has been shown to be consistent with the degradation of less-recalcitrant plant material (17), the relative contribution of these microbes to biomass degradation remains unknown.

The symbiotic gardens of leaf-cutter ants are meticulously tended by their hosts to maintain a stable microbial assemblage (18, 19). Fresh biomass is integrated into the top strata of fungus gardens and progressively degraded as it moves into lower strata. After 4 to 6 weeks, the ants remove spent fungus garden material from the bottom strata of gardens. Thus, gardens represent a gradient of biomass degradation whereby fresh biomass, composed...
of diverse polymers, plant toxins, and other compounds, is degraded in top strata while partially degraded plant biomass, from which many usable nutrients have been exhausted, comprises the bottom strata (Fig. 1C) (20). These gardens are continuously tended by the ants to facilitate biomass degradation, the removal of pests, and the prevention of infection from fungal pathogens (19). Although the two genera of leaf-cutter ants, *Atta* and *Acromyrmex*, cultivate the same phylogenetic group of *L. gongylophorus* as the primary cultivar in their gardens (21), these ant genera have distinct life histories. While species of *Acromyrmex* maintain colonies containing ~100,000 ants, *Atta* species can achieve colonies with millions of workers and hundreds of fungus garden chambers (8). Both *Atta* and *Acromyrmex* are broadly distributed throughout Neotropical ecosystems, with species of *Atta* often dominating herbivorous niches (22).

In this study, we used genomic and metaproteomic tools to investigate plant biomass degradation in the fungus gardens of both *Atta cephalotes* and *Acromyrmex echinatior* leaf-cutter ants. We sequenced the first draft genome of an *L. gongylophorus* Ac12 strain used for genome sequencing was isolated from a nest of *At. cephalotes* in Gamboa, Panama, in July 2010. Cultures of *L. gongylophorus* were maintained in the dark at 30°C on plates of potato dextrose agar (BD Difco, Franklin Lakes, NJ). DNA was purified using a modified cetyltrimethylammonium bromide (CTAB) extraction. Briefly, mycelium was scraped directly from plates into 2-ml tubes containing 0.25 mg of silica beads and 1× phosphate-buffered saline (PBS) and shaken for 2 min in a bead beater. Lysed fungal cells were then combined with CTAB buffer, and DNA was extracted as previously described (23). DNA was sequenced using one plate each of shotgun and 8-kb paired-end Roche 454 Titanium pyrosequencing (24) and subsequently assembled using the program Newbler v. 2.1 (details are in Table 1).

Proteins were predicted from the assembly using the programs GeneMark-ES (25) and Augustus (26) (for details, see the supplemental material). Each protein prediction data set was annotated independently, and proteins predicted by both tools were identified using BLASTP (27). Only proteins identified in our metaproteomic analyses (at least one peptide matching with <10-ppm error; see below for details) were included in the final protein prediction data set. Eukaryotic Orthologous Group (KOG) (28) and Pfam (29) annotations were constructed using reverse database (downloaded 1 November 2012) using BLASTP (E value, 1e−5), carbohydrate-active enzymes (CAZymes) (31) were identified using methods previously described (16), fungal oxidative lignin-degrading enzymes (FOLymes) (32) were identified by comparing all predictions to proteins on the FOLy database (downloaded 1 November 2012) using BLASTP (E value, 1e−10), and proteases were predicted using the MEROPS database (33). The annotations for all predicted CAZymes, FOLymes, and proteases were inspected manually. Signal peptide and protein localization predictions were generated for all fungal proteins using the programs WoLPsPort (34) and SignalP (35) (both the hidden Markov model and neural network implementation). All FOLymes and proteases not predicted to be secreted were excluded from subsequent analyses. Annotations for all proteins confirmed through spectral mapping can be found in Data Set SI in the supplemental material. A list of the most abundant *L. gongylophorus* lignocellulases identified in the metaproteomic data can be found in Table 2.

**Metaproteomic sample preparation, processing, and analysis.** Metaproteomic analysis was conducted with garden material collected from one *At. cephalotes* and one *Ac. echinatior* leaf-cutter ant colony reared in the laboratory on diets of oak (*Quercus*) and maple (*Acer*). Three samples were collected across the vertical gradient of biomass degradation in each fungus garden (top, middle, and bottom strata). Details for proteomic sample preparation can be found in the supplemental material. Briefly, 150 mg of each sample was added to an ice-cold mortar with liquid nitrogen and ground for ~2 min with a pestle. Then, 2 ml of water was added to the sample with continuous grinding until the mixture was

coside hydrolases sequenced from isolates associated with different ant hosts and geographic areas.

**MATERIALS AND METHODS**

**Fungal genome sequencing and annotation.** The *Leucoagaricus gongylophorus* Ac12 strain used for genome sequencing was isolated from a nest of *At. cephalotes* in Gamboa, Panama, in July 2010. Cultures of *L. gongylophorus* were maintained in the dark at 30°C on plates of potato dextrose agar (BD Difco, Franklin Lakes, NJ). DNA was purified using a modified cetyltrimethylammonium bromide (CTAB) extraction. Briefly, mycelium was scraped directly from plates into 2-ml tubes containing 0.25 mg of silica beads and 1× phosphate-buffered saline (PBS) and shaken for 2 min in a bead beater. Lysed fungal cells were then combined with CTAB buffer, and DNA was extracted as previously described (23). DNA was sequenced using one plate each of shotgun and 8-kb paired-end Roche 454 Titanium pyrosequencing (24) and subsequently assembled using the program Newbler v. 2.1 (details are in Table 1).

Proteins were predicted from the assembly using the programs GeneMark-ES (25) and Augustus (26) (for details, see the supplemental material). Each protein prediction data set was annotated independently, and proteins predicted by both tools were identified using BLASTP (27). Only proteins identified in our metaproteomic analyses (at least one peptide matching with <10-ppm error; see below for details) were included in the final protein prediction data set. Eukaryotic Orthologous Group (KOG) (28) and Pfam (29) annotations were constructed using reverse database (downloaded 1 November 2012) using BLASTP (E value, 1e−5), carbohydrate-active enzymes (CAZymes) (31) were identified using methods previously described (16), fungal oxidative lignin-degrading enzymes (FOLymes) (32) were identified by comparing all predictions to proteins on the FOLy database (downloaded 1 November 2012) using BLASTP (E value, 1e−10), and proteases were predicted using the MEROPS database (33). The annotations for all predicted CAZymes, FOLymes, and proteases were inspected manually. Signal peptide and protein localization predictions were generated for all fungal proteins using the programs WoLPsPort (34) and SignalP (35) (both the hidden Markov model and neural network implementation). All FOLymes and proteases not predicted to be secreted were excluded from subsequent analyses. Annotations for all proteins confirmed through spectral mapping can be found in Data Set SI in the supplemental material. A list of the most abundant *L. gongylophorus* lignocellulases identified in the metaproteomic data can be found in Table 2.
thawed. The samples were pipetted into 0.6-ml centrifuge tubes with 0.1-mm zirconia/silica beads and bead beaten for 3 min. The tubes were centrifuged at 4°C at 10,000×g for 5 min, and the supernatant was removed. The collected supernatant was then centrifuged at 4000×g for 5 min to remove debris. Two proteomic samples were prepared from each biological sample, one using a Rapigest SF surfactant (Waters, Milford, MA) protocol and the other using a filter-aided sample preparation (FASP) (36) protocol, and resulting peptides were analyzed by liquid chromatography–high-mass-accuracy tandem mass spectrometry (LC-MS/MS) (LTQ Orbitrap Velos and LTQ Orbitrap; Thermo Fisher Corporation, San Jose, CA). The peptide tandem mass spectra resulting from these procedures were then mapped onto appropriate protein prediction data sets. An outline of the metaproteomic workflow is provided in the supplemental material (Fig. S4). Peptide matches were filtered using Sequest scores (37), MS-GF software spectral probabilities (38), and false discovery rates, as previously described (17). Moreover, all peptides with a 10-ppm mass error were discarded. Proteins with only a single unique peptide matching were retained, but those mapping to lignocellulases were inspected manually. These peptides are listed in Data Set S4.

Protein prediction databases used for spectral mapping. Both the Augustus and GeneMark-ES protein predictions of the draft L. gongylophorus genome were used independently for spectral mapping. To compare fungal proteins to those of bacteria in fungus gardens, 10 databases corresponding to bacterial proteins were also used. Four of these databases correspond to bacterial metagenomes constructed from Atta colombica, At. cephalotes, and Ac. echinatior fungus gardens, while 6 correspond to the genomes of bacteria isolated from fungus gardens (Enterobacter strain FGI 35, Pseudomonas strain FGI 182, Serratia strain FGI 94 [39], Enterobacteriaceae strain FGI 157 [40], Klebsiella variicola AT-22, and Pantoea strain at-9b). Proteins were predicted using the software programs IMG-ER (41) and Prodigal (42) for the metagenomic and genomic data sets, respectively, and CAZymes were predicted from these data sets in a manner identical to that of the L. gongylophorus protein predictions.

Statistical analyses of spectral profiles for L. gongylophorus lignocellulases. The numbers of mass spectra mapped to L. gongylophorus lignocellulases were compared between samples to identify enzymes enriched in a given stratum. The numbers of spectra mapping to both GeneMark-ES and Augustus proteins were combined, if predicted by both programs, and final spectral profiles were clustered using Pearson’s r values calculated in the R software environment (http://www.R-project.org). In the case of individual lignocellulases (see Fig. S1 and S2 in the supplemental material), only those proteins having at least 10 spectra mapping were included in the analysis.

### TABLE 2 Fungal lignocellulases with high spectral abundance in the metaproteomic data sets

<table>
<thead>
<tr>
<th>LAG ID</th>
<th>Protein family</th>
<th>Annotation</th>
<th>Ac. echinatior</th>
<th>At. cephalotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAG_992</td>
<td>CAZymes</td>
<td>GH15 Glycoamylase, glycodeltranase</td>
<td>3,812</td>
<td>1,428</td>
</tr>
<tr>
<td>LAG_4755</td>
<td></td>
<td>PL1 Pectin/pectate lyase</td>
<td>3,425</td>
<td>1,121</td>
</tr>
<tr>
<td>LAG_3581</td>
<td></td>
<td>CE5 Acetyl-xylan esterase, cutinase</td>
<td>305</td>
<td>1,440</td>
</tr>
<tr>
<td>LAG_1450</td>
<td></td>
<td>CE8 Pectin methylsterase</td>
<td>998</td>
<td>453</td>
</tr>
<tr>
<td>LAG_543</td>
<td></td>
<td>GH28 Polygalacturonase</td>
<td>1,050</td>
<td>545</td>
</tr>
<tr>
<td>LAG_3349</td>
<td></td>
<td>PL4 Rhamnogalacturonan lyase</td>
<td>889</td>
<td>462</td>
</tr>
<tr>
<td>LAG_3001</td>
<td></td>
<td>GH35 β-Galactosidase</td>
<td>495</td>
<td>557</td>
</tr>
<tr>
<td>LAG_81</td>
<td></td>
<td>PL4 Rhamnogalacturonan lyase</td>
<td>116</td>
<td>791</td>
</tr>
<tr>
<td>LAG_420</td>
<td></td>
<td>GH18 Chitinase, acetylglucosaminidase</td>
<td>186</td>
<td>688</td>
</tr>
<tr>
<td>LAG_2062</td>
<td></td>
<td>GH3 Glucosidase, xylidosidase</td>
<td>325</td>
<td>252</td>
</tr>
<tr>
<td>LAG_1651</td>
<td></td>
<td>GH78 Rhamnosidase</td>
<td>163</td>
<td>403</td>
</tr>
<tr>
<td>LAG_2564</td>
<td></td>
<td>GH3 Glucosidase, xylidosidase</td>
<td>116</td>
<td>445</td>
</tr>
<tr>
<td>LAG_2638</td>
<td></td>
<td>GH13, CBM20 Amylase, pullulanase, glucosidase</td>
<td>449</td>
<td>58</td>
</tr>
<tr>
<td>LAG_5098</td>
<td></td>
<td>GH3 Glucosidase, xylidosidase</td>
<td>335</td>
<td>131</td>
</tr>
<tr>
<td>LAG_4224</td>
<td></td>
<td>GH10 Xylanase</td>
<td>249</td>
<td>146</td>
</tr>
<tr>
<td>LAG_2404</td>
<td>FOlymes</td>
<td>LO1 Laccase</td>
<td>3,483</td>
<td>1,947</td>
</tr>
<tr>
<td>LAG_2639</td>
<td></td>
<td>LO1 Laccase</td>
<td>2,562</td>
<td>2,013</td>
</tr>
<tr>
<td>LAG_2522</td>
<td></td>
<td>LDA3 Glyoxal oxidase</td>
<td>1,356</td>
<td>641</td>
</tr>
<tr>
<td>LAG_5549</td>
<td></td>
<td>LO1 Laccase</td>
<td>264</td>
<td>373</td>
</tr>
<tr>
<td>LAG_2622</td>
<td>Proteases</td>
<td>A01A Aspartyl protease</td>
<td>7,717</td>
<td>3,371</td>
</tr>
<tr>
<td>LAG_3716</td>
<td></td>
<td>M36 Metalloprotease</td>
<td>2,861</td>
<td>1,564</td>
</tr>
<tr>
<td>LAG_2011</td>
<td></td>
<td>S09X Serine protease</td>
<td>1,998</td>
<td>1,262</td>
</tr>
<tr>
<td>LAG_2465</td>
<td></td>
<td>S53 Serine protease</td>
<td>1,578</td>
<td>712</td>
</tr>
<tr>
<td>LAG_3735</td>
<td></td>
<td>S08A Serine protease</td>
<td>1,599</td>
<td>337</td>
</tr>
<tr>
<td>LAG_981</td>
<td></td>
<td>S10 Serine protease</td>
<td>1,038</td>
<td>310</td>
</tr>
<tr>
<td>LAG_5096</td>
<td></td>
<td>S08A Serine protease</td>
<td>471</td>
<td>783</td>
</tr>
<tr>
<td>LAG_439</td>
<td></td>
<td>A01A Aspartyl protease</td>
<td>433</td>
<td>693</td>
</tr>
<tr>
<td>LAG_7402</td>
<td></td>
<td>A01A Aspartyl protease</td>
<td>379</td>
<td>220</td>
</tr>
<tr>
<td>LAG_3725</td>
<td></td>
<td>M28E Metalloprotease</td>
<td>398</td>
<td>120</td>
</tr>
<tr>
<td>LAG_2527</td>
<td></td>
<td>S53 Serine protease</td>
<td>307</td>
<td>122</td>
</tr>
<tr>
<td>LAG_3512</td>
<td></td>
<td>S08A Serine protease</td>
<td>145</td>
<td>247</td>
</tr>
<tr>
<td>LAG_1757</td>
<td></td>
<td>S10 Serine protease</td>
<td>243</td>
<td>91</td>
</tr>
</tbody>
</table>

The spectral abundances shown represent the sum recovered from the top, middle, and bottom strata for both the Ac. echinatior and At. cephalotes samples.
were considered in order to reduce the chances of spurious clustering due to low spectral coverage. In the case of lignocellulase families, all spectra mapping to proteins of a given family were combined, and all CAZy, FOLy, and MEROPS protein families are shown (Fig. 2). Fisher’s exact test was performed on the raw number of spectra mapping to enzymes in different samples to identify lignocellulases enriched in one sample compared to another. Comparisons were performed only between samples corresponding to the same strata in *Ac. echinatior* and *At. cephalotes* gardens.

**PCR amplification, sequencing, and phylogenetic analysis of *L. gongylophorus* CAZymes.** All *L. gongylophorus* strains used for phylogenetic analysis were obtained from leaf-cutter ant colonies collected from different locations in South and Central America and maintained in the Currie laboratory (details are in Table 3). Cultures were isolated on plates of potato dextrose agar (BD Difco, Franklin Lakes, NJ) and maintained at 30°C in the dark. DNA was extracted using a modified phenol-chloroform method. Primers used for the amplification of genes encoding one GH6 and two GH10 glycoside hydrolases were designed in this study from contigs present in previously reported community metagenomes of ant gardens (details are in Table S5), which contained a small amount of *L. gongylophorus* sequences in addition to the bacterial component (17). Amplification of partial-length GH7-encoding genes was performed using the primers fungcbhIF and fungcbhIR, designed in a previous study (43). All sequences were amplified using PCR on an MJ Research PDR-200 Peltier thermal cycler with a 2-min 94°C denaturing step followed by 29 cycles of a 1-min 94°C denaturing step, a 2-min annealing step (see individual temperatures used below), and a 2-min 72°C elongation step, and finally a 6-min 72°C elongation step. The annealing temperatures used for the GH7 and GH6 amplicons and two GH10 amplicons (termed CBH1, CBHII, xynI, and xynII, respectively) were 50°C, 44°C, 52°C, and 45°C, respectively. Amplicons were run on 1% agarose gels containing ethidium bromide (EtBr) for visualization and subsequently sequenced using an ABI 3730xl DNA analyzer. The resulting chromatograms were visualized using the program Sequencer v 4.5 (Gene Codes Inc., Ann Arbor, MI). Positions for which the base calls were ambiguous due to conflicting chromatogram signals were designated "N."

Amino acid sequences for the nucleotide sequences of these CAZyme genes were obtained using a combination of 6-frame translation, the software program FGENESH (Softberry), and BLASTX (27) against the NCBI NR database (44). Nucleotide and amino acid sequences were compared among amplicons using standalone BLASTN and BLASTX (results are in Table 4). A selection of top BLASTP hits was compiled for each CAZyme and used for subsequent phylogenetic analyses. The program MUSCLE (45) was used for amino acid alignments, and FastTree software (46) was used to infer maximum-likelihood phylogenetic trees with corresponding

---

**FIG 2** Heat map representing the relative numbers of mass spectra matching to the *L. gongylophorus* CAZyme (blue), FOLyme (magenta), and MEROPS (brown) protein families in the top, middle, and bottom strata of *Ac. echinatior* and *At. cephalotes* gardens. Rows have been normalized to unity. The dendrogram represents clustering based on the Pearson correlation of the spectral profiles for each protein family.

**TABLE 3** *Leucoagaricus gongylophorus* isolates, their countries of origin, host ant species, and genes that were successfully amplified and sequenced from their purified DNA

<table>
<thead>
<tr>
<th>Isolate ID</th>
<th>Country of origin</th>
<th>Host ant species</th>
<th>Sequencing successa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CBHI</td>
</tr>
<tr>
<td>LG_Peru1</td>
<td>Peru</td>
<td><em>Acromyrmex</em> sp.</td>
<td>X</td>
</tr>
<tr>
<td>LG_CR1</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_PN1</td>
<td>Panama</td>
<td><em>Acromyrmex octospinosus</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_CR2</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_CR3</td>
<td>Costa Rica</td>
<td><em>Acromyrmex echinatior</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_CR4</td>
<td>Costa Rica</td>
<td><em>Acromyrmex echinatior</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_CR5</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_Pe2</td>
<td>Peru</td>
<td><em>Acromyrmex sp.</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_CR6</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_ARG1</td>
<td>Argentina</td>
<td><em>Acromyrmex laticeps</em></td>
<td>X</td>
</tr>
<tr>
<td>LG_ARG2</td>
<td>Argentina</td>
<td><em>Acromyrmex nigra</em></td>
<td>X</td>
</tr>
</tbody>
</table>

a ID, identifying designation.

b All genes could be amplified, but only those genes marked with an "X" were successfully sequenced.
local support values calculated using the Shimodaira-Hasegawa (SH) test (47). A small number of sequences for which regions were trimmed due to ambiguous base calls in the chromatograms were excluded in the phylogenetic analyses, although they were used for the overall comparison of nucleotide and amino acid identity. Trees were visualized using the Interactive Tree of Life (ITOL) (48) software tool.

**Accession numbers.** The draft genome sequence of *L. gongylophorus* has been deposited in DDBJ/EMBL/GenBank under the accession no. ANIS00000000. The version described in this article is the first version, ANIS01000000. *L. gongylophorus* sequences used in the phylogenetic analysis have been submitted to GenBank and are available under accession no. KC476397 to KC476437. All protein prediction data sets used in spectral mapping and all raw data concerning the proteomic experiments can be found at http://omics.pnl.gov/view/publication_1059.html.

**RESULTS**

*L. gongylophorus* draft genome assembly and annotation. Assembly of the reads from two full plates of Roche 454 Titanium pyrosequencing of purified *L. gongylophorus* DNA yielded 92,785 contigs comprising 101 Mbp of total sequence (Table 1). The largest contig in the assembly was 100,988 bp in size, and 2,368 contigs were larger than 5 kb. Analysis of the GC content of all contigs revealed a bimodal distribution with peaks at 29% and 47%, with primarily mitochondrial sequences comprising the former and chromosomal sequences the latter (see Fig. S3 in the supplemental material). Gene prediction using a combination of Augustus and GeneMark-ES yielded 12,132 nonredundant predictions, and the majority of these had homology to proteins in other sequence databases (see Table S1 for details). Individual spectral mapping of metaproteomic data to the raw Augustus and GeneMark-ES protein prediction data sets confirmed 4,567 protein predictions. Upon annotation of these confirmed proteins using the CAZY, PDB, and MEROPS databases, we identified 145 predicted biomass-degrading enzymes, including 81 glycoside hydrolases (GHs), 6 polysaccharide lyases (PLs), 9 carbohydrate esterases (CEs), 9 laccases, 5 glyoxal oxidases, 4 aryl-alcohol oxidases, and 26 secreted proteases (see Tables S2, S3, and S4, respectively). Additionally, we amplified and sequenced a single GH7 cellulase from *L. gongylophorus* that was not identified in the draft genome.

**Metaproteomics.** Our metaproteomic analysis of six samples comprising the top, middle, and bottom strata of *Ac. echinatior* and *At. cephalotes* fungus gardens identified 505,566 spectra that could be confidently mapped to amino acid sequences in at least one of our protein prediction databases (Table 1; see also Data Sets S1 and S3 in the supplemental material). Of these, 484,059 could be mapped to proteins belonging to *L. gongylophorus*, and 4,309 could be mapped to proteins in one of our bacterial databases. Of the 145 lignocellulases identified in the *L. gongylophorus* genome, 137 were represented in the metaproteome of at least one *Ac. echinatior* sample and 138 in at least one *At. cephalotes* sample (for details, see Data Set S2). In total, 44,347 spectra were mapped to all *L. gongylophorus* lignocellulases in the *Ac. echinatior* samples, while 27,878 were mapped to these same enzymes in the *At. cephalotes* samples. Inspection of CAZymes present in the bacterial component of the metaproteomic data identified only a single mass spectrometry mapping to a GH8 (a full list of bacterial proteins identified in the metaproteomic data is given in Data Set S3).

Clustering analysis of spectral profiles mapping to the *L. gongylophorus* lignocellulose protein families yielded distinct patterns throughout both the *Ac. echinatior* and *At. cephalotes* samples (Fig. 2). When the spectral profiles of individual enzymes were analyzed, separate sets of lignocellulases were identified as abundant in the top, middle, and bottom strata in both ant species (see Fig. S1 and S2 in the supplemental material). Over 46% of the lignocellulases identified in the metaproteome were overrepresented in at least one *Ac. echinatior* stratum compared to the corresponding stratum in *At. cephalotes*, while the reverse was found for 30% of the enzymes (Fisher’s exact test, P < 0.05; details are available in Data Set S2). Comparison of the spectral profiles recovered for all *L. gongylophorus* proteins in the six samples resulted in separate clustering of the *At. cephalotes* and *Ac. echinatior* samples (Fig. 3).

The most abundant *L. gongylophorus* biomass-degrading enzymes in the six metaproteomic data sets comprised 15 CAZymes, 4 FOLymes, and 13 proteases (Table 2). Specifically, these enzymes comprised 10 GHs, 3 PLs, 2 CEs, and 2 laccases. Of the abundant CAZymes identified, the majority belong to families predicted to hydrolyze hemicelluloses, pectins, and starch. Inspection of the spectral profiles of the most abundant lignocellulases revealed many to be most abundant in a particular stratum of fungus garden (Fig. 4). Of the two most abundant laccases identified, one (LAG_2404) was found to be most abundant in the top stratum of both *Ac. echinatior* and *At. cephalotes* gardens, while a nearly opposite trend was identified for the other (LAG_2639). Of the three cellulases identified, the cellobiohydrolases (GH6 and GH7) were found to be most abundant in the bottom strata of the garden.
both ant gardens, while no trend could be found for the endoglucanase (GH9). Of the most abundant hemicellulases and pectinases, a glucoamylase (GH15) and xylan esterase/cutinase (CE5) were most abundant in the top strata, while a pectin methylesterase (CE8) displayed an opposite trend. No trend consistent between *Ac. echinatior* and *At. cephalotes* samples could be identified for the other abundant lignocellulases.

**Analysis of fungal lignocellulases from different isolates.** To investigate the similarities of *L. gongylophorus* lignocellulases isolated from different ant hosts and geographic regions, we sequenced and compared the partial-length nucleotide sequences of 4 GHs from 11 fungal strains isolated from *Atta* and *Acromyrmex* colonies collected in Panama, Costa Rica, Peru, and Argentina (Table 3). Comparison of the nucleotide sequences revealed that each gene was on average 98.8 to 99.5% identical to corresponding genes in other isolates (Table 4). Moreover, the corresponding amino acid sequences for the genes were on average 98.9 to 100% identical between isolates. Phylogenetic analysis of these amino acid sequences revealed homology to enzymes encoded by numerous other saprotrophic fungi (see Fig. S5 in the supplemental material).

**DISCUSSION**

In Neotropical ecosystems, leaf-cutter ants and their symbiotic fungus play important roles in nutrient cycling and plant biomass turnover (8). Despite their importance to leaf-cutter ants, it has remained largely unknown how the fungus gardens cultivated by these insects convert fresh foliar biomass into nutrients for their hosts. Here, using genomic and metaproteomic tools, we have provided detailed insight into the lignocellulases used by the fungal cultivar *L. gongylophorus* to degrade plant biomass in the symbiotic gardens of leaf-cutter ants.

The combination of our genomic and metaproteomic analyses supports the hypothesis that *L. gongylophorus* is the dominant driver of biomass degradation in leaf-cutter ant gardens. In our analysis of the draft genome of this fungus and our metaproteomic experiments, we identified a total of 145 lignocellulases, including numerous pectinases, xylanases, amylases, and cellulases (Fig. 2).
Previous metagenomic characterizations of bacterial communities in *Atta* and *Acromyrmex* gardens recovered mainly oligosaccharide-degrading enzymes and few CAZy families associated with the degradation of recalcitrant polysaccharides (17), suggesting that bacteria are not playing a prominent role in biomass degradation in fungus gardens. Moreover, our metaproteomic results revealed that although >81,000 spectra could be confidently mapped to diverse *L. gongylophorus* lignocellulases, only a single spectrum could be mapped to a bacterial GH8, a CAZyme family that in bacteria is often associated with cell wall modification rather than plant biomass degradation (49). Previous work identified a bacterial glucosidase produced in fungus gardens (17), however, and it remains a possibility that bacterial enzymes not present in the protein databases used in this study could be playing a role in biomass degradation.

Most of the highly abundant *L. gongylophorus* enzymes identified in the metaproteomic data are predicted to degrade pectin, xylan, starch, and proteins (Table 2 and Fig. 4A). This is consistent with studies showing that both hemicellulases and proteases are produced by this organism (20, 50, 51) and that activity against these substrates can be consistently detected in whole fungus gardens (52). Moreover, culture-based work has shown that *L. gongylophorus* can produce enzymes active against a variety of polysaccharides when grown in pure culture (53, 54). Our finding that enzymes targeting hemicelluloses are among the most highly abundant in our metaproteomic data sets indicates that these polymers are likely the primary polysaccharides degraded by *L. gongylophorus* in fungus gardens. The abundance of proteases throughout all strata of our samples also suggests that plant proteins may be an important nutrient source for *L. gongylophorus* throughout the degradation process.

In addition to numerous hemicellulases and proteases, our genomic and metaproteomic analyses also identified putative cellulases (GH6, GH7, and GH9) produced by *L. gongylophorus*. Interestingly, this fungus has been reported to be incapable of growth on cellulose alone in pure culture (10). Moreover, recent investigations of cellulose degradation in fungus gardens have given conflicting results (16, 55), raising the question of whether *L. gongylophorus* or other microbes in fungus gardens can degrade this polymer. Our genomic and metaproteomic work is the first sequence-based evidence that *L. gongylophorus* both encodes cellulases and produces them in fungus gardens, suggesting that this fungus is contributing to some amount of cellulose degradation in these ecosystems. Results of our metaproteomic data also indicate that the abundance of the GH6 and GH7 cellulases increases in the lower strata of ant gardens (Fig. 4B), suggesting that they may be produced primarily when less-recalcitrant polymers have been exhausted. Interactions with bacteria or compounds produced by ants in gardens may stimulate *L. gongylophorus* to produce these enzymes, potentially explaining why reports of the cellulolytic ability of this fungus in pure culture have varied (10, 11).

Our identification of numerous laccases and accessory oxidases with secretion signals in the genome of *L. gongylophorus* indicates that these enzymes may also be important for lignocellulose degradation in the ant-fungus symbiosis. Two laccases, an aryl-alcohol oxidase, and a glyoxal oxidase were among the most well-represented enzymes in our metaproteomic data (Table 2 and Fig. 4B), suggesting that *L. gongylophorus* produces large quantities of these enzymes in ant gardens. Similar enzymes in other basidiomycetous fungi have been shown to play important roles in the degradation of plant polymers, especially lignin (56). Although recent work has indicated that significant amounts of lignin are not degraded in fungus gardens (16), it is likely that this polymer is physically linked to polysaccharides in plant cell walls, and partial degradation of lignin may therefore make otherwise unavailable polysaccharides more accessible to other lignocellulases. The use of laccases and redox-active enzymes may thus increase the efficiency of biomass degradation even if lignin itself is not used as a carbon source.

Alternatively, laccases may be necessary for the degradation of secondary metabolites in plant tissue that may be toxic to *L. gongylophorus* or the host ants, as indicated by recent work (57). This may explain why some laccases were found to be more abundant in top garden strata (Fig. 4A), where plant toxins would first be encountered by *L. gongylophorus* and lignin degradation would not yet be necessary to access additional polysaccharides. It has previously been suggested that secondary metabolites produced by plants, or even by endophytic fungi living in plant tissue, may be important factors influencing the choice of foliar biomass foraged by the ants (58). The laccases and redox-active enzymes produced by *L. gongylophorus* may thus play an important role in detoxifying these compounds and broadening the range of plants that can be harvested and efficiently degraded in the ant-fungus symbiosis.

The composition of plant biomass in ant gardens is highly variable due to the diversity of plants foraged by the ants and changes in substrate composition throughout the degradation process (59). Therefore, the ability of *L. gongylophorus* to quickly alter the production and secretion of lignocellulases in response to nutrient availabilities is likely critical for efficient biomass processing. Consistent with this, we found distinct stratification in the lignocellulase profiles of the metaproteomes of both *Ac. echinatior* and *At. cephalotes* gardens, indicating that separate enzymatic cocktails are used by *L. gongylophorus* for the degradation of lignocellulose at different stages of biomass breakdown (Fig. 2; see also Fig. S1 and S2 in the supplemental material). A distinct cocktail of abundant CAZymes, laccases, and proteases was identified in the gardens throughout all stages of biomass degradation (see Fig. S1 and S2), suggesting that *L. gongylophorus* uses different sets of these enzymes to acquire nutrients from different plant substrates depending on the stage of biomass degradation.

Despite the similar overall patterns of stratified lignocellulase profiles between ant species, significantly different quantities of mass spectra were recovered from 65% of the lignocellulases when strata were compared directly (Fisher’s exact test, P < 0.05). Both nests had been fed the same mixture of oak and maple leaves prior to the time of sampling, indicating that this difference is not due to diet. Our clustering analysis of overall mass spectra profiles for all *L. gongylophorus* proteins suggests that the physiology of *L. gongylophorus* may be influenced by the host ants (Fig. 3). Although both *Acromyrmex* and *Atta* species culture *L. gongylophorus*, a number of factors could contribute to differences in the physiology of the fungal symbiont between nests. For example, differences in the hygienic practices of the host ants or interactions with bacteria co-inhabiting fungus gardens are possible explanations.

Species of *Atta* and *Acromyrmex* are distributed across the
Americas in numerous distinct ecosystems, and their selection of plants varies depending on the ant species, location, and season (59). Because of this ecological variability, we explored the possibility that lignocellulases varied with the range or species of ant host. Contrary to our expectations, our comparison of four CAZymes reveals that the lignocellulases encoded by this fungus are highly conserved across both host ant species and geographic range (Tables 3 and 4; see also Fig. S5 in the supplemental material). This is also surprising given previous reports identifying genetic diversity in L. gongylophorus cultures obtained from the same geographic area (60). Interestingly, two of the genes analyzed here encode predicted cellulases (families GH6 and GH7), consistent with previous work indicating that the degradation of cellulose is an important process in fungus gardens (16).

In this work, we have demonstrated that the fungal symbiont of the ants, L. gongylophorus, encodes a diversity of plant biomass-degrading enzymes and is likely the primary driver of lignocellulose degradation in fungus gardens. Our metaproteomic analysis provides evidence that cellulases and redox-active enzymes produced by L. gongylophorus may be playing critical roles in this symbiosis by both degrading recalcitrant plant polymers and detoxifying secondary metabolites in plant tissue. Moreover, our phylogenetic analysis of L. gongylophorus CAZymes indicates that different species of leaf-cutter ants, which may inhabit different ecosystems and have colony sizes differing by millions of workers, appear to use the same highly conserved enzymes of L. gongylophorus to degrade their plant forage and convert it into usable nutrients. Our work highlights the importance of microbes to the herbivory of a dominant herbivore, as well as their importance to nutrient cycling and carbon turnover in Neotropical ecosystems.

ACKNOWLEDGMENTS

We thank the staff at Roche 454, the DOE Joint Genome Institute, Pacific Northwest National Laboratory, and the Smithsonian Tropical Research Institute for their expertise and support in the collection and processing of all samples. In particular, we thank L. Seid, Y. Clemmons, R. Urriola, M. Paz, and O. Arosemena for permit and collection support in Panama. We thank all members of the Currie lab for their comments on the manuscript. We thank two anonymous reviewers for their helpful comments.

Fungal genome sequencing was supported by a Roche 454 Life Sciences 10 GB grant. The U.S. Department of Energy Joint Genome Institute effort was supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231. Metaproteomic measurements were supported by the U.S. Department of Energy’s (DOE) Office of Biological and Environmental Research (OBER) Pan-omics program at Pacific Northwest National Laboratory (PNNL) and performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) Office of Biological and Environmental Research national scientific user facility on the PNNL campus. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under contract DE-AC05-76RL01830. This work was also supported by National Science Foundation grants DEB-0747002, MCB-0702025, and MCB-0731822 to C.R.C. and the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FG02-07ER64494).

REFERENCES