Leucoagaricus gongylophorus Produces Diverse Enzymes for the Degradation of Recalcitrant Plant Polymers in Leaf-Cutter Ant Fungus Gardens

Frank O. Aylward, University of Wisconsin
Kristin E. Burnum-Johnson, Pacific Northwest National Laboratory
Susannah G. Tringe, Department of Energy Joint Genome Institute
Clotilde Teiling, Roche Diagnostics
Daniel M. Tremmel, University of Wisconsin
Joseph A. Moeller, University of Wisconsin
Jarrod J. Scott, University of Wisconsin
Kerrie W. Barry, Department of Energy Joint Genome Institute
Paul D. Piehowski, Pacific Northwest National Laboratory
Carrie D. Nicor, University of Wisconsin

Only first 10 authors above; see publication for full author list.

Journal Title: Applied and Environmental Microbiology
Volume: Volume 79, Number 12
Publisher: American Society for Microbiology | 2013-06-01, Pages 3770-3778
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1128/AEM.03833-12
Permanent URL: https://pid.emory.edu/ark:/25593/s78zz

Final published version: http://dx.doi.org/10.1128/AEM.03833-12

Copyright information:
© 2013, American Society for Microbiology.

Accessed December 12, 2018 6:22 PM EST
**Leucoagaricus gongylophorus** Produces Diverse Enzymes for the Degradation of Recalcitrant Plant Polymers in Leaf-Cutter Ant Fungus Gardens


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 *Atta cephalotes* 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 1X phosphate-buffered saline (PBS) and shaken for 2 min in a bead beater. Lyed 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 PSI-BLAST (RPS-BLAST) (30) (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 WoLPSort (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 S1 in the supplemental material. A list of the most abundant *L. gongylophorus* lignocellulases identified in the metaproteomic data can be found in Table 2.

### 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 1X phosphate-buffered saline (PBS) and shaken for 2 min in a bead beater. Lyed 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 PSI-BLAST (RPS-BLAST) (30) (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 WoLPSort (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 S1 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 metaproteomic 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

| Table 1 L. gongylophorus draft genome sequencing and annotation statistics |
|-----------------------------|----------------|
| **Statistic**               | **Value**     |
| No. of contigs in the L. gongylophorus assembly | 92,785 |
| Total size (bp) in assembly | 101,584,475 |
| N50 contig size (bp)       | 1,793 |
| No. of predicted proteins  | 12,132 |
| No. of proteins verified by spectral mapping | 4,567 |
| % with KOG annotations     | 71.1 |
| % with Pfam annotations    | 71.4 |
| Total no. of mass spectra mapped onto proteins | 484,059 |
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 to both *Ac. echinatior* and *At. cephalotes* samples.

#### 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>No. of spectra mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Ac. echinatior</em></td>
</tr>
<tr>
<td>CAZymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG_992</td>
<td>GH15</td>
<td>Glycoamylase, glycodextranase</td>
<td>3,812</td>
</tr>
<tr>
<td>LAG_4755</td>
<td>PL1</td>
<td>Pectin/pectate lyase</td>
<td>3,425</td>
</tr>
<tr>
<td>LAG_3581</td>
<td>CE5</td>
<td>Acetyl-xylan esterase, cutinase</td>
<td>305</td>
</tr>
<tr>
<td>LAG_1450</td>
<td>CE8</td>
<td>Pectin methylsterase</td>
<td>998</td>
</tr>
<tr>
<td>LAG_543</td>
<td>GH28</td>
<td>Polygalacturonase</td>
<td>1,050</td>
</tr>
<tr>
<td>LAG_3369</td>
<td>PL4</td>
<td>Rhamnogalacturonan lyase</td>
<td>889</td>
</tr>
<tr>
<td>LAG_3001</td>
<td>GH35</td>
<td>$\beta$-Galactosidase</td>
<td>495</td>
</tr>
<tr>
<td>LAG_81</td>
<td>PL4</td>
<td>Rhamnogalacturonan lyase</td>
<td>116</td>
</tr>
<tr>
<td>LAG_420</td>
<td>GH18</td>
<td>Chitinase, acetylglucosaminidase</td>
<td>186</td>
</tr>
<tr>
<td>LAG_2062</td>
<td>GH3</td>
<td>Glucosidase, xylodase</td>
<td>325</td>
</tr>
<tr>
<td>LAG_1651</td>
<td>GH78</td>
<td>Rhamnosidase</td>
<td>163</td>
</tr>
<tr>
<td>LAG_2564</td>
<td>GH3</td>
<td>Glucosidase, xylodase</td>
<td>116</td>
</tr>
<tr>
<td>LAG_2638</td>
<td>GH13, CBM20</td>
<td>Amylase, pullulanase, glucosidase</td>
<td>449</td>
</tr>
<tr>
<td>LAG_5098</td>
<td>GH3</td>
<td>Glucosidase, xylodase</td>
<td>335</td>
</tr>
<tr>
<td>LAG_4224</td>
<td>GH10</td>
<td>Xylanase</td>
<td>249</td>
</tr>
<tr>
<td>FOLymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG_2404</td>
<td>LO1</td>
<td>Laccase</td>
<td>3,483</td>
</tr>
<tr>
<td>LAG_2639</td>
<td>LO1</td>
<td>Laccase</td>
<td>2,562</td>
</tr>
<tr>
<td>LAG_2522</td>
<td>LDA3</td>
<td>Glyoxal oxidase</td>
<td>1,356</td>
</tr>
<tr>
<td>LAG_5549</td>
<td>LO1</td>
<td>Laccase</td>
<td>264</td>
</tr>
<tr>
<td>Proteases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG_2622</td>
<td>A01A</td>
<td>Aspartyl protease</td>
<td>7,717</td>
</tr>
<tr>
<td>LAG_3716</td>
<td>M36</td>
<td>Metalloprotease</td>
<td>2,861</td>
</tr>
<tr>
<td>LAG_2011</td>
<td>S09X</td>
<td>Serine protease</td>
<td>1,998</td>
</tr>
<tr>
<td>LAG_2465</td>
<td>S53</td>
<td>Serine protease</td>
<td>1,578</td>
</tr>
<tr>
<td>LAG_3735</td>
<td>S08A</td>
<td>Serine protease</td>
<td>1,599</td>
</tr>
<tr>
<td>LAG_981</td>
<td>S10</td>
<td>Serine protease</td>
<td>1,038</td>
</tr>
<tr>
<td>LAG_5096</td>
<td>S08A</td>
<td>Serine protease</td>
<td>471</td>
</tr>
<tr>
<td>LAG_439</td>
<td>A01A</td>
<td>Aspartyl protease</td>
<td>433</td>
</tr>
<tr>
<td>LAG_7402</td>
<td>A01A</td>
<td>Aspartyl protease</td>
<td>379</td>
</tr>
<tr>
<td>LAG_3725</td>
<td>M28E</td>
<td>Metalloprotease</td>
<td>388</td>
</tr>
<tr>
<td>LAG_2527</td>
<td>S53</td>
<td>Serine protease</td>
<td>307</td>
</tr>
<tr>
<td>LAG_3512</td>
<td>S08A</td>
<td>Serine protease</td>
<td>145</td>
</tr>
<tr>
<td>LAG_1757</td>
<td>S10</td>
<td>Serine protease</td>
<td>243</td>
</tr>
</tbody>
</table>

*a*The spectral abundances shown represent the sum recovered from the top, middle, and bottom strata for both the *Ac. echinatior* and *At. cephalotes* samples.

---

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 PRD-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 success&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG_Peru1</td>
<td>Peru</td>
<td><em>Acromyrmex sp.</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR1</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_FN1</td>
<td>Panama</td>
<td><em>Acromyrmex octospinosus</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR2</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR3</td>
<td>Costa Rica</td>
<td><em>Acromyrmex echinatior</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR4</td>
<td>Costa Rica</td>
<td><em>Acromyrmex echinatior</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR5</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_Pe2</td>
<td>Peru</td>
<td><em>Acromyrmex sp.</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_CR6</td>
<td>Costa Rica</td>
<td><em>Atta cephalotes</em></td>
<td>X X X X</td>
</tr>
<tr>
<td>LG_ARG1</td>
<td>Argentina</td>
<td><em>Acromyrmex laticeps</em></td>
<td>X X X</td>
</tr>
<tr>
<td>LG_ARG2</td>
<td>Argentina</td>
<td><em>Acromyrmex niger</em></td>
<td>X X X</td>
</tr>
</tbody>
</table>

<sup>a</sup> ID, identifying designation.
<sup>b</sup> All genes could be amplified, but only those genes marked with an “X” were successfully sequenced.
TABLE 4 Nucleotide and amino acid similarities of four CAZymes sequenced from 11 L. gongylophorus isolates

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>CAZy family</th>
<th>Length of nucleotide alignment</th>
<th>Length of amino acid alignment</th>
<th>Avg nucleotide identity (%)</th>
<th>Avg no. of gaps in nucleotide alignment</th>
<th>Avg amino acid identity (%)</th>
<th>Avg no. of gaps in amino acid alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>XynI</td>
<td>GH10</td>
<td>647</td>
<td>94</td>
<td>99.1</td>
<td>0.92</td>
<td>99.17</td>
<td>0</td>
</tr>
<tr>
<td>XynII</td>
<td>GH10</td>
<td>429</td>
<td>125</td>
<td>98.9</td>
<td>1.04</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>CBHI</td>
<td>GH7</td>
<td>567</td>
<td>125</td>
<td>99.8</td>
<td>1.02</td>
<td>98.94</td>
<td>0</td>
</tr>
<tr>
<td>CBHIII</td>
<td>GH6</td>
<td>838</td>
<td>237</td>
<td>99.5</td>
<td>0.91</td>
<td>99.9</td>
<td>0</td>
</tr>
</tbody>
</table>

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 spectrum 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 lignocellulase 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 strata 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 fungus gardens.
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).
and Table 2; see also Fig. S1 and S2 in the supplemental material). Future transcriptomic analyses and additional genome sequencing will likely identify more enzymes used by this fungus for biomass degradation in leaf-cutter ant fungus gardens.

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 sera 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* proteomes 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 coinhabiting 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 ant, 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. gongyl-

ACKNOWLEDGMENTS

We thank the staff at Roche 454, the DOE Joint Genome Institute, Pacific Northwest National Laboratories, 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. Clemons, 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 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-AC02-05W001264. Metaproteomics 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

hints from external sources. BMC Bioinformatics 7:62. doi:10.1186/1471-
2105-7-62.
28. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin
EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolayska AY, Rao BS,
Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA.
2003. The COG database: an updated version includes eukaryotes. BMC
29. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G,
/nar/gkm960.
31. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Hen-
ribsat B. 2009. The Carbohydrate-Active EnZymes database (CAZY): an
32. Aylward FO, Tremmel DM, Starrett GJ, Bruce DC, Chain P, Chen A,
Davenport K, Detter C, Han CS, Han J, Huntemann M, Ivanova NN,
Kyrpides NC, Markowitz V, Mavromatis K, Nolan M, Pagani I, Pati A,
Pitluck S, Teshima H, Deshpande S, Goodwin L, Woyke T, Currie CR.
2013. Complete genome of Serratia sp. strain FGI 94, a strain associated
33. Aylward FO, Tremmel DM, Bruce DC, Chain P, Chen A, Walston
Davenport K, Detter C, Han CS, Han J, Huntemann M, Ivanova NN,
Kyrpides NC, Markowitz V, Mavromatis K, Nolan M, Pagani I, Pati A,
Complete genome of Enterobacteriaceae bacterium strain FGI 57, a strain
associated with leaf-cutter ant fungus gardens. Genome Announc. 1(1):
e0023812. doi:10.1128/genomeA.00238-12.
34. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpi-
des NC. 2009. IMG ER: a system for microbial genome annotation expert
35. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ,
119.
36. Edwards IP, Upchurch RA, Zak DR. 2008. Isolation of fungal cellobio-
hydrolase I genes from sporocarps and forest soils by PCR. Appl. Environ.
Microbiol. 74:3481–3489.
37. Pruitt KD, Tatusova T, Maglott DR. 2007. NCBI reference sequences
(RefSeq): a curated non-redundant sequence database of genomes, tran-
38. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accu-
29. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G,
/nar/gkm960.
31. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Hen-
ribsat B. 2009. The Carbohydrate-Active EnZymes database (CAZY): an
32. Aylward FO, Tremmel DM, Starrett GJ, Bruce DC, Chain P, Chen A,
Davenport K, Detter C, Han CS, Han J, Huntemann M, Ivanova NN,
Kyrpides NC, Markowitz V, Mavromatis K, Nolan M, Pagani I, Pati A,
Pitluck S, Teshima H, Deshpande S, Goodwin L, Woyke T, Currie CR.
2013. Complete genome of Serratia sp. strain FGI 94, a strain associated
33. Aylward FO, Tremmel DM, Bruce DC, Chain P, Chen A, Walston
Davenport K, Detter C, Han CS, Han J, Huntemann M, Ivanova NN,
Kyrpides NC, Markowitz V, Mavromatis K, Nolan M, Pagani I, Pati A,
Complete genome of Enterobacteriaceae bacterium strain FGI 57, a strain
associated with leaf-cutter ant fungus gardens. Genome Announc. 1(1):
e0023812. doi:10.1128/genomeA.00238-12.
34. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpi-
des NC. 2009. IMG ER: a system for microbial genome annotation expert
35. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ,