UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCIÊNCIAS

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
(ÁREA: MICROBIOLOGIA APLICADA)

MARIANA DE OLIVEIRA BARCOTO

FUNGUS-GROWING INSECTS HOST A CONVERGENT MICROBIOME WITH FUNCTIONAL SIMILARITIES TO OTHER LIGNOCELLULOSE-FEEDING INSECTS

Rio Claro, SP
2017

MARIANA DE OLIVEIRA BARCOTO

FUNGUS-GROWING INSECTS HOST A CONVERGENT MICROBIOME WITH FUNCTIONAL SIMILARITIES TO OTHER LIGNOCELLULOSE-FEEDING INSECTS

Dissertação apresentada ao Instituto de Biociências, Câmpus de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Mestre em Ciências Biológicas (Área: Microbiologia Aplicada).

Orientador: Prof. Dr. André Rodrigues

Co-orientador: Prof. Dr. Maurício Bacci Jr.

Rio Claro, SP

2017

589.2 Barcoto, Mariana de Oliveira
B244f Fungus-growing insects host a convergent microbiome with functional similarities to other lignocellulose-feeding insects / Mariana de Oliveira Barcoto. - Rio Claro, 2017
49 f. : il., figs., gráfs., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências de Rio Claro
Orientador: André Rodrigues
Coorientador: Maurício Bacci Jr.

1. Fungos. 2. Host-associated microbiota. 3. Functional convergence. 4. Host-microbiome association. 5. Fungiculture. 6. Metagenome. 7. CAZy. I. Título.

Ficha Catalográfica elaborada pela STATI - Biblioteca da UNESP
Campus de Rio Claro/SP

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Rio Claro

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: FUNGUS-GROWING INSECTS HOST A CONVERGENT MICROBIOME WITH FUNCTIONAL SIMILARITIES TO OTHER LIGNOCELLULOSE-FEEDING INSECTS

AUTORA: MARIANA DE OLIVEIRA BARCOTO
ORIENTADOR: ANDRE RODRIGUES

Aprovada como parte das exigências para obtenção do Título de Mestra em CIÊNCIAS BIOLÓGICAS (MICROBIOLOGIA APLICADA), pela Comissão Examinadora:

Prof. Dr. ANDRE RODRIGUES
Departamento de Bioquímica e Microbiologia / IB-Rio Claro

Profa. Dra. MICHELE DE CÁSSIA PEREIRA E SILVA
Departamento de Solos e Nutrição de Plantas / ESALQ / USP - Piracicaba,SP

Profa. Dra. MARINA CORREA CORTES
Departamento de Ecologia / Instituto de Biociências de Rio Claro - SP

Rio Claro, 05 de maio de 2017

Aos meus pais, pelo apoio ao estabelecer objetivos, traçar caminhos (mesmo que isso envolva me lançar ao desconhecido) e lidar com a saudade, dedico.

AGRADECIMENTOS

"Do not think that time simply flies away. Do not understand 'flying' as the only function of time. If time simply flew away, a separation would exist between you and time. So, if you understand time as only passing, then you do not understand the time being. To grasp this truly, every being that exist in the entire world is linked together as moments in time, and at the same time they exist as individual moments of time." (Dōgen Zenji, Uji)

As vezes penso que não vi o caminho que percorri, agora que estou chegando ao final. Posso não ter lembranças de muitos detalhes, mas isso não significa que não observei esse trajeto; mais importante que observar é entender as mudanças. Eu mudei a forma pela qual entendo a vida, e isso só é possível devido às experiências associadas a esse percurso. Se me perguntassem pelo resultado final, diria que aprendi muito. Aprendi com pessoas, ouvindo e explicando; aprendi com (muitos) livros e artigos; aprendi com experiências, expectativas e medos. O aprendizado é uma consequência do meio no qual estamos imersos, e a todos os que participaram desse processo, eu gostaria de agradecer:

Aos meus pais, por ensinarem a importância da dedicação e do trabalho árduo como única maneira de atingir mudanças; a vida é pesada e não vai sair do lugar se não houver determinação em empurrá-la. Por ensinarem a lidar com as decisões com responsabilidade, o valor do caráter e por estarem sempre ao meu lado, incentivando, amparando, lidando com dificuldades, saudades e vitórias. E especialmente por terem acolhido meu "white orc". Em particular: mãe, obrigada por ligar todos os dias; isso me faz sentir perto de casa, independentemente de onde eu estiver. Obrigada pelo exemplo profissional, por ouvir e tentar entender um mundo que nem sempre faz parte da sua realidade, pela amizade, pelo apoio nas horas mais difíceis; você faz tudo ficar mais leve. Ao meu irmão, por mostrar que o meu ponto de vista é só uma forma de enxergar a pluralidade do mundo; sua jovialidade e leveza são inspiradoras.

À toda minha família, por mostrar a importância de raízes firmes e manter os elos unidos mesmo com a distância. Por todo incentivo, preocupação, e por tornarem a volta ao lar sempre mais aconchegante e divertida. Vocês trazem significado à jornada e fazem o caminho menos árduo.

Matheus, agradeço por todos os anos de parceria, por crescer comigo e acima de tudo, por se adaptar ao meu universo. Eu sempre soube que é um universo inóspito, e mesmo assim, você permanece e aprende a lidar (essa plasticidade é criticamente importante). Pelo apoio, incentivo, paciência e amizade. Sua companhia pelos caminhos por onde passamos juntos tornou essa jornada muito mais interessante.

Ao Prof. Dr. André Rodrigues, pela orientação ao longo dos anos em que tenho trabalhado no Laboratório de Ecologia e Sistemática de Fungos (LESF), pela dedicação ao projeto, por compartilhar e discutir ideias, pelas oportunidades, por impulsionar meu desenvolvimento acadêmico e pensamento crítico, e por ensinar o respeito ao pensamento científico. Muito do meu desenvolvimento científico se deve ao trabalho no LESF, da bancada à discussão de ideias, aprendendo a entender resultados e a ouvir sugestões.

Ao Prof. Dr. Maurício Bacci Jr, pela co-orientação, pelas diretrizes e sugestões, e também por disponibilizar o laboratório para o desenvolvimento do projeto.

Ao Prof. Cameron Currie, pela atenção e interesse no projeto durante os meses em que trabalhei na UW-Madison; pelas sugestões, trocas de ideias, suporte, pelo ambiente cientificamente empolgante e pela oportunidade de desenvolver minhas pesquisas no Currie lab.

Aos demais professores da Pós-Graduação em Microbiologia Aplicada, com os quais pude aprender muitos outros assuntos e aplicações dentro da Microbiologia, muito além das "ômicas" pelas quais optei por focar minha carreira acadêmica.

Aos integrantes do LESF (atuais e de gerações passadas), Ana Maria, Danilo, Lucas, Tássio, Júlio, Karina, Lorena, Bruna, Quimi e todos os outros colegas com os quais trabalhei no LESF. Em particular, à Irina, pela interação simbiótica "aprendizado-ensinamento", na qual é sempre incerta a função de cada componente. Aprendi muito com você, profissional e pessoalmente, principalmente a criar um ambiente de livre expressão de ideias, planos e divagações.

Ao Prof. Dr. Luiz Carlos Forti e Nilson S. Nagamoto pela ajuda com as coletas.

Aos integrantes do Laboratório de Evolução Molecular, Sérgio Kakazu (Miagui) pela ajuda com a purificação do DNA, e Dra. Milene Ferro pela ajuda com os procedimentos iniciais de controle de qualidade das sequências e pelas sugestões para o projeto.

Aos integrantes do Currie lab, Lily, Heidi, Caitlin, Hongjie, Rob, Marc, Jeniffer, Brandon, pela troca de ideias, sugestões, discussões científicas e hospitalidade. Agradeço à

Huan, pela ajuda com as reconstruções filogenéticas dos metagenomas, e à Camila, pelo apoio em todo o período em que trabalhei no Currie lab (desde ajudar a encontrar um lugar para ficar às inúmeras sugestões no projeto e discussão de ideias).

Agradeço também ao Departamento de Bioquímica e Microbiologia, e ao Centro de Estudos de Insetos Sociais (CEIS), pelo suporte técnico concedido para o desenvolvimento desse trabalho.

Ao CNPq, pela bolsa concedida nos primeiros meses de mestrado. À FAPESP, por conceder a Bolsa de Mestrado Regular (Processo n. 2015/16415-0) e Bolsa de Estágio e Pesquisa no Exterior (Processo n. 2016/02767-5).

"Look again at that dot. That's here. That's home. That's us."

Carl Sagan

RESUMO

Insetos cultivadores de fungos (formigas, cupins e besouros) evoluíram de forma independente uma associação simbionte com fungos que, ao metabolizar biomassa vegetal recalcitrante, produzem nutrientes disponíveis para seu hospedeiro. Esses sistemas de fungicultura também abrigam microbiomas bacterianos que apresentam importantes impactos fisiológicos na biologia do inseto fungicultor. Nesse trabalho, foram explorados os padrões de convergência funcional da microbiota associada a sistemas de fungicultura. Com o intuito de expandir a distribuição geográfica de microbiomas associados a sistemas de fungicultura, metagenomas de comunidades bacterianas de *Mycocepurus goeldii* (Attini basal) e *Atta sexdens rubropilosa* (Attini derivada, também denominada formiga-cortadeira de folhas) foram sequenciados e anotados. Tais amostras constituem os primeiros microbiomas de formigas Attini da América do Sul. Os gêneros *Pseudomonas, Pantoea, Rhizobium, Enterobacter, Achromobacter*, *Stenotrophomonas* e *Serratia* foram os mais abundantes na comunidade bacteriana associada ao jardim de fungo de *A. sexdens rubropilosa. Pseudomonas* também foi o gênero encontrado em maior abundância na comunidade bacteriana de *M. goeldii*, seguido de *Dysgonomonas, Bacteroides*, *Parabacteroides*, *Prevotella, Comamonas* e *Burkholderia*. A fim de explorar o perfil funcional da comunidade bacteriana, foram realizadas comparações entre microbiomas de Attini basais e derivadas; de insetos fungicultores; de insetos fungicultores e trato intestinal de insetos cuja alimentação é baseada em lignocelulose. As análises comparativas revelaram que os microbiomas associados a insetos fungicultores apresentam uma evidente convergência funcional e taxonômica. É possível verificar a existência de similaridades funcionais entre microbiomas de insetos fungicultores e do trato intestinal de insetos herbívoros, principalmente em relação às rotas envolvidas no metabolismo de carboidratos, aminoácidos, compostos aromáticos, cofatores e vitaminas. No entanto, diferenças taxonômicas notáveis podem ser observadas entre microbiomas de insetos fungicultores e trato intestinal de insetos herbívoros, fornecendo evidências adicionais para a convergência funcional na comunidade bacteriana associada a insetos cultivadores de fungos.

Palavras-chave: Convergência funcional. Associação hospedeiro-microbioma. Fungicultura. Attini. Macrotermitinae. Scolytinae. Metagenoma. CAZy.

ABSTRACT

Fungus-growing insects (ants, termites, and beetles) independently evolved a symbiotic association with fungi that metabolize recalcitrant plant biomass, producing nutrients available to the insect host. These fungicultural systems also harbor bacterial microbiota of important physiological impacts for the host life style. Here, we explore convergence patterns of the microbiota associated with fungiculture systems. For expanding the geographic distribution of microbiomes fungiculture systems available, we sequenced and annotated metagenomes of bacterial communities from *Mycocepurus goeldii* (lower Attini ant) and *Atta sexdens rubropilosa* (higher Attini, a leaf-cutter ant), the first attine ants' microbiomes from South America. *Pseudomonas, Pantoea, Rhizobium, Enterobacter, Achromobacter, Stenotrophomonas* and *Serratia* were the most abundant genera in the bacterial community of *A. sexdens rubropilosa* fungus garden. Similarly, *Pseudomonas* was also the most abundant genus in the bacterial community of *M. goeldii* fungus garden, followed by *Dysgonomonas*, *Bacteroides*, *Parabacteroides*, *Prevotella, Comamonas* and *Burkholderia*. For metabolic profiling, these microbiomes were included in comparisons of several levels: between lower and higher attines, among fungus-growing insects, and between fungus-growing and non-fungus-growing insects. Comparative analysis of fungus-growing insects associated microbiomes support remarkable functional and taxonomic similarities, pointing to convergence in bacterial communities. Metabolic parallels may be found among microbiomes from fungus-growing insects and other lignocellulose-feeding insects, particularly for pathways involved with the metabolism of carbohydrates, amino acids, aromatic compounds, cofactors and vitamins. However, there are substantial taxonomic differences between microbiomes from fungiculture systems and herbivorous insects' gut, giving further evidence for the functional convergence in bacterial microbiota associated with fungus-growing insects.

Keywords: Functional convergence. Host-microbiome association. Fungiculture. Attini. Macrotermitinae. Scolytinae. Metagenome. CAZy

CONTENTS

1 Introduction ……12

2 Methods ……...……15

3 Results ……....18

4 Discussion ……...…22

5 References ……...….26

Figure legends ……34

Table and Figures……...…….36

Supplemental material ……...….43

Fungus-growing insects host a convergent microbiome with functional similarities to other lignocellulose-feeding insects

Mariana O. Barcoto[1], Camila Carlos[2], Huan Fan[2], Mauricio Bacci Jr.[3], Cameron Currie[2], André Rodrigues[1,3]*

[1] *Department of Biochemistry and Microbiology, UNESP – São Paulo State University, Rio Claro, SP, Brazil.*

[2]*Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA*

[3]*Center for the Study of Social Insects, UNESP – São Paulo State University, Rio Claro, SP, Brazil.*

Short title: Convergence in microbiomes from insects fungiculture

Keywords: Functional convergence. Host-microbiome association. Fungiculture. Attini. Macrotermitinae. Scolytinae. Metagenome. CAZy

Status: In prep.

Correspondence to:

Dr. Andre Rodrigues (andrer@rc.unesp.br)

Universidade Estadual Paulista, Departamento de Bioquímica e Microbiologia

Av. 24-A, n. 1515, Bela Vista, Rio Claro, SP, Brazil, Zipcode: 13.506-900.

1 Introduction

Host-associated microbial communities comprise complex ecological systems, shaped by several stochastic and selective forces.[1,2] Most of microbial functional traits are not phylogenetically exclusive, and different microorganisms could occupy similar functional niches within the host.[2,3] Thus, the assembly of functionally equivalent species within a niche could be an outcome of natural selection acting mainly on microbial functional traits.[2] Identifying the functional repertoire of a microbiome could potentially predict ecological parameters and evolutive processes shaping the microbiota composition.[2,4] Also, the functional capabilities of host-associated microbial communities provide valuable information on how the interactions among community members influence the host's physiology.[5-7] Interactive host-microbiota metabolism could be considered a key determinant of the host life style and fitness over an evolutionary time scale.[6,8,9] The dynamics of this metabolic communication also affects the diversity, composition, and functional capacity of the associated microbiota, which is shaped by a series of interactions throughout the host life, including diet, physicochemical variations, and maturation phases.[9]

Only by considering the microbiota as a component of ecological and physiological processes, the nutritional ecology of insects can be explained.[6,10] Particularly for herbivorous insects, which may feed on recalcitrant and indigestible plant tissues, with low nutritional content and toxic compounds, the association with symbiotic microorganisms mediates the use of resources otherwise non-accessible to the host.[6,11] Fungus-growing insects are a noticeable example of insect-microbial symbiosis for exploring recalcitrant plant biomass. The active maintenance of fungus crops, also known as fungiculture, evolved independently in tree insect lineages:[12] ants in the tribe Attini (Hymenoptera: Formicidae: Myrmicinae), which are strict to the New World, occurring from the south of Argentina to the south of United States;[13] termites in the subfamily Macrotermitinae (Isoptera: Termitidae), which occur in the Old-World tropics, mainly in Africa and Asia;[14] and beetles in the subfamily Scolytinae (Coleoptera: Curculionidae), which are found in wet tropical forests.[15] These "fungus farmers" exhibit specific methods for fungus cultivation, harvesting and propagation (Fig. 1).[12]

The fungal lignocellulose-degrading capacity has been fundamental for the evolutionary success of the fungus-growing insects symbiosis.[16-18] The so-called lower attines (Fig. 1A)

cultivate basidiomycete fungal symbionts (mostly of the tribe Leucocoprinae) in diverse substrates. These fungi are not truly domesticated and are able to sustain a free live existence.[12,19-21] Higher attines (Fig 1B) cultivate a fully and highly domesticated basidiomycete fungal symbiont in the genus *Leucoagaricus,* which is unable of a free-living existence and apparently has been clonally dispersed through vertical transmission.[12,22-24] The ambrosia beetle fungal symbionts are usually the ascomycete genera *Ambrosiella* and *Raffaelea,* which are cultivated on the walls of galleries excavated in the xylem of trees, and comprise the only nutritional source for larvae and adults (Fig 1C).[18,25-27] Although bark beetles cannot be considered true "fungus farmers", they are commonly associated with ascomycetes in the genera *Ophiostoma*, *Ceratocystiopsis*, *Grosmannia*, and *Ceratocystis.* Bark beetles main nutritional source is phloem, but the fungal association may be related to improving nutrients availability and detoxification of host-defense metabolites.[27,28] Plant biomass degradation in fungus-growing termites' colonies (Fig 1D) depends on the enzymatic activity of a basidiomycete fungal symbiont in the genus *Termitomyces,* in combination with workers' gut microbiota.[17,29] *Termitomyces* grows within fungus combs, which are structures built from termite feces containing pre-digested plant material (processed by the termite and by the enzymatic activity of the gut microbiota) and asexual spores of *Termitomyces*.[17,30-32]

Fungus-growing insects also harbor bacterial microbiotas of important physiological consequences for the host life style. For example, some components of the bacterial microbiota from higher Attini ants are able to fix atmospheric nitrogen,[33] and the community also exhibit the metabolic capacity for lignocellulose degradation, aminoacids and vitamins biosynthesis.[34,35] Bacteria found in the galleries of ambrosia beetles are considered secondary symbionts,[36,37] while bacteria associated with bark beetles are able to produce antimicrobial compounds[38] and can degrade terpenes, which are employed as chemical defense released by conifers as a response to bark beetles attacking behavior.[39,40] In termite fungiculture, besides antimicrobial-producing bacteria which could suppress antagonistic fungus of the symbiosis,[41] the workers gut microbiome and the fungus comb microbiome aid in the comb continuous lignocellulose decomposition.[17,32] The metabolic potential of the microbiota associated with fungus-growing insects suggests that the bacterial community could be a usual and determinant characteristic of fungiculture systems.[42]

Despite differences in insect host geographic distribution and life history, bacterial microbiomes associated with fungus-growing insects exhibit a highly similar taxonomic composition, indicating that these communities could be an example of convergence of host-microbiome association.[42] Convergent evolution manifests as the independent evolution of similar traits or phenotypes (as physiological or morphological features of a microorganism) in multiple lineages.[43-45] The convergence of phenotypes provides remarkable evidence that particular environmental conditions probably select for specific traits - which represents similar evolutionary solutions.[46] When referring to convergent evolution of microbiotas, we are considering that a set of similar functions and/or taxonomic composition may be found in the microbiota associated to different and distantly related insect hosts. Here, we explore whether convergence patterns may be found in the microbiota associated with fungiculture systems. For expanding the geographic distribution of the microbiomes available, we included the first Attini microbiomes from South America. Our choice for adding microbiomes from the lower attini *Mycocepurus goeldii* and the higher attini *Atta sexdens rubropilosa* was based on the wide distribution of these species in Brazil.[47,48] We first described the taxonomic composition and the functional capacity of the bacterial microbiota from fungus gardens of both ant species. Then, we included these microbiomes in a broad comparison of microbiota functional profile in several levels: between lower and higher attini, among fungus-growing insects, and between fungus-growing and non-fungus-growing insects. Comparative analysis of the microbiomes support the bacterial microbiota taking place in plant material decomposition. Metabolic parallels may be found among microbiomes from fungus-growing insects and other lignocellulose-feeding insects, particularly for pathways involved with the metabolism of carbohydrates, amino acids, aromatic compounds, and cofactors and vitamins. Nevertheless, notable differences in taxonomic composition of microbiomes from fungiculture and insects' gut not only emphasize taxonomic similarities in fungiculture systems,[42] but also reinforce different host-microbiome associations as functionally similar environments. As such, fungiculture systems' microbiota comprise a remarkable example of functionally convergent bacterial community.

2 Methods

2.1 Sampling

Fungus gardens from healthy colonies of *Atta sexdens rubropilosa* and *Mycocepurus goeldii* were collected from nests near Botucatu, São Paulo State, Brazil (S22º49.886'/W48º25.426' and S22º54.353'/W48º14.562', respectively), in July and October 2015, respectively. Both *A. sexdens rubropilosa* and *M. goeldii* colonies were sampled in shadowed and humid areas of eucalyptus cultivation, with approximately 3-10 meters of distance between colonies. Fungus gardens' top and bottom sections were sampled from two colonies of *A. sexdens rubropilosa*, and were combined for resulting 92.59 g. Because of the smaller size of *M. goeldii* fungus gardens, central and peripherical regions were sampled from 18 colonies and were combined for totalizing 50.58 g of fungus garden (Fig S1). Immediately after collection, samples were kept under controlled conditions (25 °C, in the dark) for subsequent preparations.

2.2 Bacterial enrichment, DNA extraction and sequencing

The bacterial fraction of total fungus garden samples was isolated through a modification of a previously described protocol (Fig S1).[34,35] Briefly, workers, larvae, and pupae were removed from the samples, and fungus garden were buffered in 1X PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) containing 0.1% Tween 80 and gently centrifuged (30 minutes at 40 xg). This mixture was incubated at room temperature for six days for fungus gardens of *A. sexdens rubropilosa* and for ten days for *M. goeldii* gardens. During this period, the fungus garden settled at the bottom of the tubes. To recover bacteria that still could be in the liquid phase, the buffer was carefully transferred to another tube, filtered and centrifuged (30 minutes at 2800 xg), then the resulting pellet was stored at 8 °C. The fungus garden was washed in fresh buffer, centrifuged (30 minutes at 40 xg) and incubated in the same conditions. The washing and incubation steps were repeated three times. Following these washing steps, the fungus garden was shaken for 3 minutes, filtered, and centrifuged for 30 minutes at 2800 xg. Then, the several pellets resulting from the same sample were joined. The presence of bacteria in the final pellet was confirmed through bright-field microscopy. DNA was subsequently extracted from the bacterial fraction using the PowerSoil DNA Isolation Kit (MoBio Laboratories), by adapting the fabricant protocol. Basically, instead of extracting DNA from a 0.25 g sample as suggested in the protocol, we extracted from a 0.40 g sample. We empirically verified this

adaptation resulting in DNA samples with higher quantity and quality from our bacterial samples. DNA was sequenced through Illumina HiSeq 2000, paired ending sequencing (100bp).

2.3 Assembling, annotation, and analysis of *Mycocepurus goeldii* and *Atta sexdens rubropilosa* metagenomes

Reads quality control and preprocessing were carried out in Solexa QA v3.1.5,[49] sorting the reads by quality (*phred cutoff*=13) and length (*length cutoff*=60). Preprocessing quality was checked in FastQC. Quality-controlled reads were assembled using default settings in MEGAHIT v1.0.6.[50] Quality of assembled contigs were verified in PRINSEQ.[51]

Quality-controlled contigs were uploaded to the Integrated Microbial Genomes (IMG/M ER) database for gene identification and annotation through the standard pipeline of IMG.[52] In summary, the identification of protein-coding genes is performed by combining the *ab initio* gene calling tolls: GeneMark,[53] Metagene,[54] Prodigal,[55] and FragGeneScan.[56] Protein-coding genes then are compared with Clusters of Orthologous Groups database (COG)[57] using RPS-BLAST, and Protein Families database (Pfam)[58] using HMMER3. Metagenome protein-coding genes are also associated with KEGG Orthology (KO) terms, KEGG Enzymes and EC numbers[59] through USEARCH.[60]

The abundance of protein coding-genes for each assigned function (COG, Pfam, KO, KEGG Enzymes) was standardized by the total number of protein-coding genes (i.e., we determined the proportion of protein-coding genes codifying a function by the total number of protein-coding genes). Statistical comparisons were performed based on Fisher's exact test for the standardized abundance of function-annotated genes, using STAMP v2.1.3.[61]

Genes were taxonomic assigned using the "Phylogenetic Distribution of Genes" comparative tool of IMG,[62] which estimates the phylogenetic composition of the metagenome by comparing (through RPS-BLAST) the best BLASTp hits with COG database. The taxonomic assignment was performed at bacterial genera (identity percentage > 60%). The genus-level taxonomic was further investigated by aligning contigs > 100Kbp with 16S rDNA sequences in the Ribosomal Database Project,[63] and through taxonomic assignment of contigs > 100Kbp using PhyloPhytiaS.[64]

Diversity indices were calculated using the standardized abundance of taxonomic-assigned protein-coding genes using PAST 3.[65] The Venn diagram was calculated in R v3.3.2.

2.4 Comparisons of the bacterial communities from insects' fungiculture

We included the two metagenomes annotated here in a broader comparison, using metagenomes publicly available in IMG/M database (Table S1 and Dataset_Methods, available at: https://www.dropbox.com/s/hb0hediysud5jm9/Dataset_Methods.xlsx?dl=0).

Microbiomes were compared for all annotated functions (COG, Pfam, KO, KEGG Enzymes, CAZy). Every comparison was performed with the standardized abundance of protein-coding genes assigned to a specific function. For comparing Carbohydrates-Active Enzymes (CAZymes)[66] from different fungus-growing insects' bacterial communities, CAZymes were designated for the protein-coding genes using the pipeline described in Takasuka *et al.*[67] Statistical comparisons were performed between lower and higher Attini microbiomes, among fungus-growing insects microbiomes, and between fungus-growing and non-fungus-growing insects microbiomes, through STAMP v2.1.3.[61]

Profiles of KEGG pathways (via KO terms) were determined in IMG/M, as the KEGG genes total dataset (standardized abundance of all genes classified by KEGG pathways), and as the subsets of pathways defined by higher levels of organization in KEGG pathways (standardized abundance of genes classified by each metabolic pathway: carbohydrates metabolism, energy metabolism, lipids metabolism, amino acids metabolism, aromatic compounds metabolism, cofactors and vitamins metabolism). To visualize similarities among samples, non-metric multidimensional scaling (NMDS) ordination based on Correlation distances was calculated for profiles of KEGG pathways, for each annotated function (COG, Pfam, KO, KEGG Enzymes, CAZy), and for protein coding genes taxonomic assigned to class level (using "Phylogenetic Distribution of Genes" comparative tool of IMG, identity percentage > 60%).

To determine significant differences between microbiomes from fungicultures (Attini ants, Macrotermitinae termites, and Scolytinae beetles) and insects' gut (termite, Passalidae beetle, and honey bee guts), similarity analysis (ANOSIM) based on 9999 permutations was calculated in PAST 3.[65]

KEGG Enzymes standardized abundance overrepresented in fungiculture's microbiome (Kruskal-Wallis test) were manually classified using the MetaCyc database.[68] Phylogenetic reconstruction from metagenomes was performed through an alignment-free *k*-mer based approach (*k*-mer length= 23), described in Fan *et al.*[69] The standardized relative abundances of bacterial classes were estimated using taxonomic assigned protein-coding genes (using "Phylogenetic Distribution of Genes" comparative tool of IMG, identity percentage > 60%).

3 Results

3.1 Metagenomic sequencing

Sequencing of the bacterial community obtained from *Mycocepurus goeldii* fungus garden yielded 5.4 Gbp of raw sequence data (53 329 142 reads, Q30= 91.224%). The community from *Atta sexdens rubropilosa* fungus garden resulted in 6.7 Gbp of raw data (66 381 084 reads, Q30= 85.958%). Reads of each library were assembled into metagenomes consisting of 249-364 Mbp of sequence data. Assembled contigs comprised sequences with good quality and length (Table 1).

3.2 Bacterial community from *Mycocepurus goeldii* and *Atta sexdens rubropilosa* fungiculture: structure and function

The bacterial community predicted by BLASTp analysis revealed most of the protein-coding genes to belong to the phyla Actinobacteria, Bacterioidetes, Firmicutes and Proteobacteria (Fig S2). These phyla were further explored at genus level, revealing that *Pseudomonas, Pantoea, Rhizobium, Enterobacter, Achromobacter, Stenotrophomonas* and *Serratia* were the most abundant genera in the bacterial community of *A. sexdens rubropilosa* fungus garden. *Pseudomonas* was also the most abundant genus in the bacterial community of *M. goeldii* fungus garden, followed by *Dysgonomonas, Bacteroides, Parabacteroides, Prevotella, Comamonas* and *Burkholderia* (Fig. 2A). Most of the genera represented in Fig. 2 were further confirmed in the bacterial community through alignment of the metagenomes with a 16S rDNA database (Fig. S3) and taxonomic assignment of the sequences (Fig. S4).

M. goeldii fungus gardens' bacterial community is richer and slightly more diverse, sharing about 55% of the estimated genera with the bacterial community of *A. sexdens rubropilosa* fungus garden. On the other hand, approximately 82% of estimated genera of the *A.*

sexdens rubropilosa community are shared with the *M. goeldii* community. Both bacterial communities are considered highly uneven (Fig. 2B).

The standardized abundance of predicted genes annotated with COG, Pfam, KO and KEGG Enzymes were compared between the two communities (Dataset S1). We classified the annotated functions through KEGG pathways (via KO terms). Pathways related to carbohydrate metabolism, nucleotide metabolism, amino acid metabolism and membrane transport were relatively well represented (>1.5% KO annotated genes) in both metagenomes (Fig 3). By comparing KO annotated genes, we observed that standardized abundance of several functions differed significantly between the two communities (Fisher's exact test with Bonferroni correction, P <0.05), especially those involved in carbohydrate metabolism, nucleotide metabolism, amino acid metabolism, genetic information process, membrane transport, cell motility and cell cycle (Fig 3).

3.3 Bacterial communities from insect fungiculture

3.3.1 *Lower and higher Attini: bacterial communities are functionally similar*

The samples annotated in this study were grouped with other bacterial metagenomes derived from attini fungus gardens available on the IMG/M database (Table S1 and Dataset_Methods). The standardized abundance of genes annotated with COG, Pfam, KO, KEGG Enzymes, and CAZy were compared between bacterial communities from lower attini (*M. goeldii*, *Ap. dentigerum*, and *C. longiscapus*) and higher attini (*Trachymyrmex*, *Ac. echinator*, *At. colombica*, *At. cephalotes*, *At. sexdens rubropilosa*) fungus gardens. Through these comparisons no differences were found (White's non-parametric t-test with Bonferroni correction, P>0.05, Dataset S2). Although there are differences when comparing the communities as individual samples (Fig 3), these differences are not significant when expanding the comparison to a larger group, implying that bacterial communities in attini fungus gardens tend to be functionally similar.

3.3.2 *Convergence patterns: functional and taxonomic similarities of microbiomes from different fungiculture systems*

We attempted to verify whether the functional profile of bacterial communities from each fungus-growing system (Attini ants, Macrotermitinae termites and Scolytinae beetles) has a specific functional profile, intrinsic of each fungicultural environment.

By comparing the members of CAZy families, we observed no differences between the bacterial communities from attini ants' and termites' fungus gardens (White's non-parametric t-test with Bonferroni correction, $P>0.05$, Dataset S3), and between communities from termites' fungus gardens and beetles' fungus gallery (White's non-parametric t-test with Bonferroni correction, $P>0.05$, Dataset S3). On the other hand, when comparing communities from attini ants' fungus gardens and beetles' fungus gallery, there are differences in the standardized abundance of the glycoside hydrolases GH75 (chitosanase), GH98 (galactosidase, xylanase), GH119 (α-amylase), the glycoside transferases GT23 (fucosyltransferase), GT29 (sialyltransferase), GT34 (galactosyltransferase), GT70 (glucuronosyltransferase), GT84 (glucan synthase), GT87 (mannosyltransferase), GT89 (arabinofuranosyltransferase), the carbohydrate esterase CE15 (methylesterase), the auxiliary activities AA3 (cellobiose dehydrogenase), AA5 (oxidase), and the carbohydrate-binding modules CBM4 (binding to xylan, glucan and amorphous cellulose), CBM16 (binding to cellulose and glucomannan), CBM23 (mannan-binding function; White's non-parametric t-test with Bonferroni correction, $P<0.05$, Dataset S3). Although showing differences in some CAZy functions, it is worthy to note the similarities in the pattern of CAZy families' distribution (Fig 1E). This similar pattern could insinuate that, by codifying a similar CAZy profile, the bacterial community from different fungicultures has the potential to metabolize a similar range of carbohydrates.

To estimate the lignocellulose influence in shaping the functional similarity found in fungiculture systems, we compared microbiomes from fungus-growing insects with microbiomes from environments with high lignocellulose content, as the wood-feeding termites *Nasutitermes corniger* and *Microcerotermes parvus* guts, the litter-feeding *Cornitermes* sp. gut, and the Passalidae beetle gut (which also feed on recalcitrant plant biomass). For these analysis, we also included microbiomes from environments not enriched in lignocellulose content, as the honey bee gut, the soil-feeding termite *Cubitermes ugandensis* gut, the humus-feeding termites *Termes hospes* and *Neocapritermes taracua* guts. In general, in ordinations of function abundances and

taxonomic-assigned protein-coding genes abundances, microbiomes from insect fungiculture formed a cluster relatively distinct from insects' gut microbiomes, although *M. goeldii* and *At. colombica* (dump top and bottom) microbiomes tended to set closer to honey bee gut and Passalidae beetle gallery than to other fungicultures' communities (Fig 4). However, analysis of similarities (ANOSIM) indicated that insect fungiculture microbiomes and insects' gut microbiomes could share a large degree of their functional profile (Fig. 4).

Microbiomes from insect fungiculture and insects' gut sharing a large degree of their functional profile (Fig 4) bring into question whether these bacterial communities potentially metabolize a similar range of nutrients, and as consequence, have similar role in plant biomass decomposition. For both fungiculture and insects' gut microbiomes, genes classified through KEGG pathways are predominantly involved in carbohydrates and amino acids metabolism. While the abundance of genes involved in carbohydrates metabolism is slightly higher in insects' gut microbiome, genes involved in the metabolism of aromatic compounds tend to be more abundant in fungus-growing insects' microbiome (Fig 5A). In ordinations of the entire set of KEGG pathways, microbiomes from insect fungiculture grouped independently from insects' gut microbiomes, and the two groups are relatively distinguishable (Fig 5B). At lower levels of the KEGG hierarchy, microbiomes from fungiculture and insects' gut differ most by lipids metabolism (Fig 5D) and energy metabolism (Fig 5F). The microbiomes were most similar for pathways involved with carbohydrates metabolism (Fig 5C) and relatively similar regarding pathways involved with cofactors and vitamins (Fig 5H), amino acids (Fig 5G), and aromatic compounds metabolisms (Fig 5E). Ordinations coupled with analysis of similarities revealed that communities from environments enriched in lignocellulose contents tend to be metabolically similar, especially for carbohydrates, amino acids, and aromatic compounds metabolic pathways.

Molecules and nutrients that could be differentially metabolized for fungiculture bacterial communities were further investigated by comparing the standardized abundance of genes annotated with KEGG Enzymes between fungiculture systems (metagenomes from Attini ants, Macrotermitinae termites, and Scolytinae beetles) and non-fungiculture systems (insects' gut and soil metagenomes). Compared with metagenomes from non-fungiculture systems, many KEGG Enzymes categories were overrepresented in fungiculture's microbiome (Kruskal-Wallis test with Bonferroni correction, $P<0.05$), including: carbohydrates assimilation and degradation; glycan, sugar, and sugar derivatives degradation, assimilation, and transportation; proteins and

aminoacids degradation and assimilation; lipids degradation and assimilation; aromatic compounds degradation and assimilation; energetic metabolism and nucleosides, nucleotides, and polynucleotides metabolism (Fig S4).

Attempting to determine whether the tendency of functional similarity (Fig. 4 and Fig. 5) has phylogenetic underpinnings, we used an alignment-free *k*-mer based approach for phylogenetic reconstruction from metagenomes and compared the class-level microbiota composition. Alignment-free phylogenetic inferences displayed fungiculture grouping separately from non-fungiculture microbiomes, and within non-fungiculture, insects' gut microbiomes formed a cluster distinct from soil microbiomes (Fig. 6). Furthermore, the relative abundance of bacterial classes among the inferred clusters differed remarkably (Fig 6, Dataset S6). In fungiculture microbiomes, Gamma-proteobacteria was the most abundant class (21.8 - 67.3%), followed by Beta-proteobacteria (0.35 - 17.1%), Alpha-proteobacteria (0.42 - 13.6%), Bacteroidia (0 - 24.4%), and Bacilli (0.019 - 8.35%). It is interesting to note the distinct class-level profile of *At. colombica* dump (top and bottom) microbiomes: the most abundant class was Alpha-proteobacteria (9.36 - 9.6%), and then Actinobacteria (8.37 - 8.67%), Gamma-proteobacteria (2.64 - 6.92%), and Beta-proteobacteria (3.17 - 6.28%). In soil microbiomes, Alpha-proteobacteria (4.69 - 17.4%), Actinobacteria (3.8 - 5.3%), Acidobacteriia (0.71 - 7.8%), Beta-proteobacteria (2.01 - 2.83%) were the most abundant classes. In insect gut microbiomes, the most abundant classes were Bacteroidia (0 - 35.4%), Clostridia (0.01 - 22.1%), Gamma-proteobacteria (0.02 - 21.7%), and Bacilli (0.16 - 17.7%).

4 Discussion

Fungus-growing insects (ants, termites, and beetles) are intriguing hosts for studying the interactions with bacterial microbiota and their functions. These insects independently evolved a symbiotic association with fungi that, by metabolizing recalcitrant plant biomass, produce nutrients available to their host.[12,42] Here, by comparing the functional profile of microbiomes associated with diverse fungus-growing insects, we found conspicuous functional and taxonomic similarities evidencing functional convergence in bacterial communities. The convergence seems to be driven by lignocellulose intake and the metabolic activity of the symbiotic fungus, suggesting interactions between the bacterial microbiota and the cultivated fungi for nutrient cycling in fungus-growing systems. Our results give further evidence for bacterial communities

associated with fungiculture participating in plant biomass deconstruction, possibly complementing the metabolism of the symbiotic fungus.

By analyzing the bacterial microbiomes from *M. goeldii* and *At. sexdens rubropilosa*, our results reiterate the abundant presence of *Pseudomonas, Pantoea, Enterobacter,* and *Stenotrophomonas* (Fig. 2). These genera were previously reported as the most abundant in Attini ants' fungus gardens,[34,35,70] and apparently, are important to the community structure.[42] It was surprising to find out that the genus *Dysgonomonas* (which was reported in environments with high lignocellulosic content)[71-73] seems to be highly abundant in *M. goeldii* microbiome (Fig. 2). Besides, some of the genera relatively abundant found in both metagenomes as *Pseudomonas, Burkholderia, Achromobacter, Comamonas, Sphingobium,* and *Acinetobacter,* have shown the ability to break down lignin components.[74-77]

Bacterial communities from fungus-growing systems have a highly similar functional profile (Fig. 1E, Fig. 4, and Fig. 5), which could reflect similarities in niche conditions of the fungicultural environment. Host-niche conditions are determined by interaction of many factors, which in turn influence bacterial communities structure and function. For mammalian hosts, even though the community composition is widely defined by host diet,[78,79] a considerable variation is correlated with host phylogeny.[78,80] Thus, diet and host phylogeny differently impact the dynamics defining the animal gut bacterial community.[81,82] Insect gut microbial communities can be influenced by diet, pH, environmental factors, life stage and specificity to the host. Although host diet and taxonomy are variables that most contribute to gut community composition, the influence of these factors is not alike across insects guilds.[81] For hymenopteran and termite gut microbial community structure, host phylogeny seems to have a stronger influence, while diet has higher impact in structuring microbial communities from insects which feed on lignocellulose-derived material.[81] Therefore, given the central role that lignocellulose has in fungiculture systems,[12] it seems plausible to assume that the high lignocellulosic content in these environments could shape the community structure (Fig. 2) and drive the functional convergence to some extent.[83,84]

Host-associated microbiomes are structurally and functionally adaptable to shifts on the available nutritional resources.[85-87] The human gut microbiota show a remarkable difference between plant-based (fiber-rich) and animal-based diets.[86] A plant-based diet (rich in starch, fiber, plant polysaccharides, and nonanimal proteins) enriched the microbiota for Bacterioidetes

and depleted for Firmicutes, and exclusively exhibited the genera *Prevotella, Xylanibacter* (Bacterioidetes) and *Treponema* (Spirochaetes) when compared to low-fiber diet. The particular composition of these microbiomes could represent a response to long-term high fiber intake, as a way to optimize the energetic metabolism derived from plant polysaccharides.[85] In fact, humanized mice that fed on diet depleted in complex carbohydrates had quick and evident reduction in microbiota diversity, impacting the dynamics of host-microbiota interactions.[88] Also for insect gut microbiota, diet variability and complexity influence the composition and diversity of microbiomes.[83] In lignocellulosic diets varying in content, fiber-richer diets resulted in dominance of Firmicutes in cockroach gut microbiota,[89] and in the prevalence of Spirochaetes in cotton weevil gut microbiota.[83] Functionally flexible microbial communities that could adapt the functional profile in response to diet possibly have enlarged host dietary flexibility.[86] The impact of short-term diet shifts in the microbiota structure suggest that a long-term plant biomass intake could favor a bacterial community that optimally utilize plant material. In some way, the microbiota could have coevolved with the host and its diet, optimizing the energetic metabolism derived from plant polysaccharides which are indigestible to the host.[85] For that reason, the intake of lignocellulosic substrates in a long-term scale could be a selective force determining the bacterial community structure and functional repertoire in fungiculture systems.

Lignocellulose intake selecting for a microbiota which optimally explore plant biomass could explain microbiomes from insect fungiculture and herbivorous insects' gut sharing a large degree of their functional profile (Fig 4). This indicates that communities from environments enriched in lignocellulose contents could be metabolically similar, especially for carbohydrates, amino acids, and aromatic compounds metabolic pathways (Fig 5). Lignocellulose-enriched environments could represent a similar set of available nutritional resources, that in turn could reflect similarities in the microbiota functional profile. A considerable amount of carbon and energy are derived from the metabolism of carbohydrates. The "microbiota-accessible carbohydrates" include those that the host's organism cannot directly degrade, but are metabolized the host microbiota.[90] These carbohydrates have a wide range of chemical composition, solubility, and complexity, representing vast possibilities of ecological niches for those components of the microbiota which codify specific enzymes for carbohydrates degradation, as glycoside hydrolases and polysaccharide lyases (Fig 1E).[66,90] Similarities in carbohydrate metabolism capacity of microbiomes from insect fungiculture and insects' gut (Fig

4D and Fig 5C) suggest that these different herbivorous systems have a similar intake of carbohydrates accessible to the bacterial community.

Besides carbohydrates metabolism, microbiomes from insect fungiculture and insects' gut are functionally similar for amino acids, cofactors, vitamins, and aromatic compounds (Fig 5), suggesting that the microbiota from these different herbivorous systems potentially synthesize nutrients, digest recalcitrant plant polymers, and neutralize plant toxins.[11,35] It is interesting to point out the similarity in the metabolism of aromatic compounds (Fig 5E and Fig S5). This could indicate that the microbiota takes part on the degradation of lignin-derived components[91] and detoxification of compounds employed in plant chemical defense,[92] both fundamental for allowing the use of plant tissues as food sources.

Despite the notable functional similarity, microbiomes from insect fungiculture and herbivorous insects' gut are structurally different at bacterial class-level composition, possibly reflecting the microbiota phylogenetic distribution (Fig 6). Also, it is worthy to notice the remarkable functional and taxonomic similarities in microbiomes associated with fungus-growing insects. The functional and taxonomic similarities become even more noticeable when considering the wide geographic distribution of the microbiomes analyzed, which we further expanded in this study. These evidences not only reinforce the taxonomic convergence,[42] but also point to functional convergence in fungiculture associated microbiomes. These convergent microbiomes are possibly an outcome of several selective evolutionary forces, mainly the lignocellulose intake and the metabolic activity of the symbiotic fungus. The breakdown of lignocellulose components was demonstrated for the fungal symbiont of fungus-growing ants,[16,34,93-95] fungus-growing-termites,[17,96,97] ambrosia beetles,[18] and bark beetles.[28,98] The fungal symbiont metabolic activity over plant biomass could provide specific sets of available nutritional resources, determining metabolic niches for the bacterial community. Moreover, we found further evidence for bacterial communities participating in plant biomass deconstruction (Fig 1E, Fig 4 and Fig 5), possibly integrating and/or complementing the symbiont fungus metabolism.[34,35,92] Convergent interactions[99] between the bacterial microbiota and the cultivated fungi could contribute for a maximized lignocellulose degradation, suggesting that fungicultures comprise an efficient system for nutrient cycling.

Acknowledgements

We would like to thank to FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) for providing scholarships to MOB (n. 2015/16415-0 and n. 2016/02767-5). We are grateful to Prof. Dr. Luiz C. Forti and Dr. Nilson S. Nagamoto for helping with field sampling, to Sérgio Kakazu for helping with DNA purification, and Dra. Milene Ferro for the initial quality control of the metagenomes sequenced and annotated in this study, and for analysis suggestions. We also thank to Prof. Dr. Odair C. Bueno for providing lab colonies for establishing the enrichment protocol.

5 References

1. Nemergut, D.R. e*t al.* Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* **77**(3), **342-356** (2013).

2. Shafquat, A., Joice, R., Simmons, S. & Huttenhower, C. Functional and phylogenetic assembly of microbial communities in the human microbiome. *Trends Microbiol*. **22**(5), 261–266 (2014).

3. Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S. & Thomas, T. Bacterial community assembly based on functional genes rather than species *Proc of the Natl Acad Sci USA* **108**(34), 14288-14293 (2011).

4. Raes, J., Letunic, I., Yamada, T., Jensen, L. J. & Bork, P. Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Mol Syst Biol*. **7**: 473; doi: 10.1038/msb.2011.6 (2011).

5. Kinross, J. M., Darzi, A.W. & Nicholson, J. K. Gut microbiome-host interactions in health and disease. *Genome Med* **3**(14) (2011).

6. Engel, P. & Moran, N. A. The gut microbiota of insects – diversity in structure and function. *FEMS Microbiol Rev* **37**, 699–735 (2013).

7. Nakayama, J. *et al.* Diversity in gut bacterial community of school-age children in Asia. *Sci Rep* **5**, 8397; doi:10.1038/srep08397 (2015).

8. Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R. & Gordon, J. I. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* **6,** 776-788 (2008a).

9. Nicholson, J.K. *et al.* Host-Gut Microbiota Metabolic Interactions. *Science* **336**, 1262-1267 (2012).

10. Douglas, A. E. The microbial dimension in insect nutritional ecology. *Funct Ecol* **23**, 38-47 (2009).

11. Hansen, A.K. & Moran, N.A. The impact of microbial symbionts on host plant utilization by herbivorous insects. *Mol. Ecol.* **23,** 1473–1496 (2013).

12. Mueller, U.G., Gerardo, N.M., Aanen, D.K., Six, D.L. & Schultz TR. The evolution of agriculture in insects. *Annu Rev Ecol Evol Syst* **36,** 563– 595 (2005).

13. Mayhé-Nunes, A. J. & Jaffé, K. On the biogeography of Attine (Hymenoptera: Formicidae). *Ecotropicos* **11**(1), 45-54 (1998).

14. Nobre, T., Rouland-Lefèvre, C. & Aanen, D. K. Comparative biology of fungus cultivation in termites and ants. In *Biology of Termites: A Modern Synthesis*. Bignell, D. (2011)

15. Jordal, B. H. & and Cognato, C. Molecular phylogeny of bark and ambrosia beetles reveals multiple origins of fungus farming during periods of global warming. *BMC Evol Biol* **12**:133, doi**:** 10.1186/1471-2148-12-133 (2012).

16. Aylward, F.O. *et al. Leucoagaricus gongylophorus* produces diverse enzymes for the degradation of recalcitrant plant polymers in leaf-cutter ant fungus gardens. *Appl Environ Microbiol* **79**(12),3770-3778 (2013).

17. Poulsen, M. *et al.* Complementary symbiont contributions to plant decomposition in a fungus-farming termite. *Proc Natl Acad Sci USA* **111,** 14500-14505 (2014).

18. De Fine Lich, H.H. & Biedermann, P. H. W. Patterns of functional enzyme activity in fungus farming ambrosia beetles. *Front Zool* **9**, 13; doi: 10.1186/1742-9994-9-13 (2012).

19. Chapela, I., Rehner, S.A., Schultz, T. R. & Mueller, U.G. Evolutionary history of the symbiosis between fungus-growing ants and their fungi. *Science* **266**(5191), 1691-1694 (1994).

20. Mueller, U.G., Rehner, S.A. & Schultz, T.R. The evolution of agriculture in ants. *Science* **281** (5385), 2034-2038 (1998).

21. Shik, J.Z. *et al.* Nutrition mediates the expression of cultivar-farmer conflict in a fungus-growing ant. *Proc of the Natl Acad Sci USA* **113**(36), 10121–10126 (2016).

22. Schultz, T.R. & Brady, S.G. Major evolutionary transitions in ant agriculture. *Proc of the Natl Acad Sci USA* **105**(14), 5435-54401 (2008).

23. De Fine Licht, H.H., Boomsma, J.J. & Tunlid, A. Symbiotic adaptations in the fungal cultivar of leaf-cutting ants. *Nat Commun* **5**, 5675; doi:10.1038/ncomms6675 (2014).

24. Nygaard, S. *et al.* Reciprocal genomic evolution in the ant–fungus agricultural symbiosis *Nat Commun* **7,** 12233; doi:10.1038/ncomms12233 (2016).

25. Beaver, R. A. Insect-fungus relationships in the bark and ambrosia beetles. In *Insect-Fungus Interactions* (ed. Wilding, N., Collins, N.M., Hammond, P.M. & Webber, J.F) 121–143 (Academic Press: London, UK, 1989).

26. Farrell, B. D. *et al.* The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution* **55**, 2011–2027 (2001).

27. Six, D. L. Ecological and evolutionary determinants of bark beetle-fungus symbioses. *Insects* **3,** 339-366 (2012).

28. DiGuistini, S. *et al.* Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont *Grosmannia clavigera*, a lodgepole pine pathogen. *Proc Natl Acad Sci USA* **108**(6) 2504–2509 (2011).

29. Liu, N. *et al.* Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). *PLoS One* **8,** e69184 (2013).

30. Nobre, T. & Aanen, D. K. Fungiculture or Termite Husbandry? The Ruminant Hypothesis. *Insects* **3**(1), 307-323 (2012).

31. Poulsen, M. Towards an integrated understanding of the consequences of fungus domestication on the fungus-growing termite gut microbiota. *Environ Microbiol* **17**(8), 2562–2572 (2015).

32. Li, H. Lignocellulose pretreatment in a fungus-cultivating termite. *Proc Natl Acad Sci USA*; www.pnas.org/cgi/doi/10.1073/pnas.1618360114 (2017).

33. Pinto-Tomás, A. A. *et al.* Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. *Science* **326,** 1120-1123 (2009).

34. Suen, G. *et al.* An insect herbivore microbiome with high plant biomass degrading capacity. *PLoS Genet* **6**(9), e1001129 (2010).

35. Aylward, F. O. *et al.* Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. *ISME J* **6,** 1688-1701 (2012).

36. Haanstad, J.O. & Norris, D. M. Microbial symbiotes of the ambrosia beetle *Xyletorinus politus*. *Microb Ecol* **11**, 267–276 (1985).

37. Grubbs, K. J. *et al.* The complete genome sequence of *Streptomyces* cf. *griseus* (XyelbKG-1), an Ambrosia beetle associated Actinomycete. *J Bacteriol* **193**, 2890–2891 (2011).

38. Scott, J.J. *et al.* Bacterial protection of beetle-fungus mutualism. *Science* **322**(5898), 63 (2008).

39. Boone, C. K. Bacteria associated with a tree-killing insect reduce concentrations of plant defense compounds. *J Chem Ecol* **39**,1003–1006 (2013).

40. Xu, L-T., Lu, M. & Sun, J-H. Invasive bark beetle-associated microbes degrade a host defensive monoterpene. *Insect Sci* **23**, 183–190 (2016).

41. Um, S., Fraimout, A., Sapountzis, P., Oh, D-C. & Poulsen, M. The fungus-growing termite *Macrotermes natalensis* harbors bacillaene-producing *Bacillus* sp. that inhibit potentially antagonistic fungi. *Sci Rep* **3**, 3250; doi: 10.1038/srep03250 (2013).

42. Aylward, F. O. *et al.* Convergent bacterial microbiotas in the fungal agricultural systems of insects. *mBio* **5**(6), e02077-14; doi:10.1128/mBio.02077-14 (2014).

43. Stayton, C. T. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* **69**(8), 2140-2153 (2015).

44. Arbuckle, K. & Speed, M. P. Analysing Convergent Evolution: A Practical Guide to Methods. In *Evolutionary Biology: Convergent Evolution, Evolution of Complex Traits, Concepts and Methods*. Pontarotti, P. (2016).

45. Martiny, J. B. H., Jones, S. E., Lennon, J. T. Martiny, A. C. Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**(6261), aac9323 (2015).

46. Stern, D. L. The genetic causes of convergent evolution. *Nat Rev Genet* **14**(11), 751-64 (2013).

47. Rabeling, C., Verhaagh, M. & Engels, W. Comparative study of nest architecture and colony structure of the fungus-growing ants, *Mycocepurus goeldii* and *M. smithii*. *J Insect Sci* **7**:40; insectscience.org/7.40 (2007).

48. Zanetti, R. *et al.* An overview of integrated management of leaf-cutting ants (Hymenoptera: Formicidae) in Brazilian forest plantations. *Forests* **5**(3), 439-454 (2014).

49. Cox, M. P., Peterson, D. A. & Biggs, P. J. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BCM Bioinformatics* **11**, 485; doi: 10.1186/1471-2105-11-485 (2010).

50. Li, D., Liu, C-M., Luo, R., Sadakane, K. & Lam, T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**(10), 1674-1676 (2015).

51. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**(6), 863-864 (2011).

52. Markowitz, V.M *et al.* IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res* **42**, D568–D573; doi:10.1093/nar/gkt919 (2014).

53. Zhu, W., Lomsadze, A. & Borodovsky, M. *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Res* **38**, e132; doi: doi.org/10.1093/nar/gkq275 (2010).

54. Noguchi, H., Park, J. & Takagi, T. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res* **34,** 5623–5630 (2006).

55. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**(119); **doi:** 10.1186/1471-2105-11-119 (2010).

56. Rho, M., Tang, H. & Ye, Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* **38**, e191; doi: 10.1093/nar/gkq747 (2010).

57. Tatusov, R. L. *et al.* The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**(41); **doi:** 10.1186/1471-2105-4-41 (2003).

58. Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res* **40**, D290–D301; doi: 10.1093/nar/gkr1065 (2012).

59. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. &Tanabe, M. KEGG for integration and interpretation of large scale molecular data sets. *Nucleic Acids Res* **40,** D109–D114; doi: doi.org/10.1093/nar/gkr988 (2012).

60. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

61. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**(21), 3123-3124 (2014).

62. Markowitz, V. M. *et al.* IMG/M-HMP: A Metagenome comparative analysis system for the Human Microbiome Project. *PLoS One* **7**(7), e40151; doi: 10.1371/journal.pone.0040151 (2012).

63. Cole, J.R. *et al.* Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**, D633–D642; doi:10.1093/nar/gkt1244 (2014)

64. Patil, K. R., Roune, L. & MChardy, A.C. The PhyloPythiaS Web Server for Taxonomic Assignment of Metagenome Sequences. *PLoS ONE* **7**(6), e38581; doi: 10.1371/journal.pone.0038581 (2012).

65. Hammer, Ř., Harper, D. A. T. & Ryan, P. D. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol Electronica* **4**(1); http://palaeo-electronica.org/2001_1/past/issue1_01.htm (2001).

66. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**, D233–D238; DOI:10.1093/nar/gkn663 (2009).

67. Takasuka TE, Book AJ, Lewin GR, Currie CR, Fox BG. Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces. Sci Rep* **3**, 1030; doi: 10.1038/srep01030 (2013).

68. Caspi, R. *et al*. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **38,** D473–D479; doi:10.1093/nar/gkp875 (2010).

69. Fan H, Ives AR, Surget-Groba Y, Cannon CH. An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics* **16,** 522; doi: 10.1186/s12864-015-1647-5 (2015).

70. Kellner, K., Ishak, H. D., Linksvayer, T. A. & Mueller, U. G. Bacterial community composition and diversity in an ancestral ant fungus symbiosis. *FEMS Microbiol Ecol* **91,** 7; doi: 10.1093/femsec/fiv073 (2015).

71. Yang, Y-J. *et al. Dysgonomonas macrotermitis* sp. nov., isolated from the hindgut of a fungus-growing termite. *Int J Syst Evol Microbiol* **64**, 2956–2961 (2014).

72. Pramono, A. K., Sakamoto, M., Iino, T., Hongoh, Y. & Ohkuma, M. *Dysgonomonas termitidis* sp. nov., isolated from the gut of the subterranean termite *Reticulitermes speratus Int J Syst Evol Microbiol* **65**, 681–685 (2015).

73. Duang, J., Liang, J., Wang, Y., Du, W. & Wang, D. Kraft lignin biodegradation by *Dysgonomonas* sp. WJDL-Y1, a newly anaerobic bacterial strain isolated from sludge of a pulp and paper mill. *J Microbiol Biotechnol* **26**(10),1765-1773 (2016).

74. Pieper, D. H. Aerobic degradation of polychlorinated biphenyls. *Appl Microbiol Biotechnol* **67**, 170-191 (2005).

75. Masai E, Katayama Y, Fukuda M. Genetic and biochemical investigations on bacterial catabolic pathways for lignin derived aromatic compounds. *Biosci Biotechnol Biochem* **71**, 1-15 (2007).

76. Ahmad, M., *et al.* Development of novel assays for lignin degradation: comparative analysis of bacterial and fungal lignin degraders. *Mol BioSyst* **6**, 815-821 (2010).

77. Bugg, T. D. H., Ahmad, M., Hardiman, E.M. & Singh, R. The emerging role for bacteria in lignin degradation and bio-product formation. *Curr Opin Biotechnol* **22**, 394-400 (2011).

78. Ley, R. E. *et al.* Evolution of mammals and their gut microbes. *Science* **320**,1647–1651 (2008b).

79. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332,** 970–974 (2011).

80. Zhu, L., Wu, Q., Dai, J., Zhang, S. & Wei, F. Evidence of cellulose metabolism by the giant panda gut microbiome. *Proc. Natl Acad. Sci. USA* **108**, 17714–17719 (2011).

81. Colman, D. R., Toolson, E. C. & Takacs-Vesbach, C. D. Do diet and taxonomy influence insect gut bacterial communities? *Mol Ecol* **21**(20), 5124-37. doi: 10.1111/j.1365-294X.2012.05752.x. (2012).

82. Sanders, J. G. *et al.* Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. *Nat Commun* **6**, 8285; doi: 10.1038/ncomms9285 (2015)

83. Guerrero, B. E. *et al.* Effect of different lignocellulosic diets on bacterial microbiota and hydrolytic enzyme activities in the gut of the cotton boll weevil (*Anthonomus grandis*). *Front Microbiol* **7**, 2093; doi: 10.3389/fmicb.2016.02093 (2016).

84. Wong, M. T. *et al.* Substrate-driven convergence of the microbial community in lignocellulose-amended enrichments of gut microflora from the Canadian Beaver (*Castor canadensis*) and North American Moose (*Alces americanus*). *Front Microbiol* **7,** 961; doi: 10.3389/fmicb.2016.00961 (2016).

85. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107**(33), 14691-14696 (2010).

86. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* **505**(7484), 559-63: doi: 10.1038/nature12820 (2014a).

87. David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome Biol* **15**, R89; doi: 10.1186/gb-2014-15-7-r89 (2014b).

88. Sonnenburg, E. D. *et al.* Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529,** 212–215; doi:10.1038/nature16504 (2016).

89. Bertino-Grimaldi, D. *et al.* Bacterial community composition shifts in the gut of *Periplaneta americana* fed on different lignocellulosic materials. *SpringerPlus* **2,** 609; doi: 10.1186/2193-1801-2-609 (2013).

90. Sonnenburg, E. D. & Sonnenburg, J. L. Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab* **20**(5),779-86; doi: 10.1016/j.cmet.2014.07.003 (2014).

91. Scully, E. D. *et al.* Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One* **8**(9), e73827; doi: 10.1371/journal.pone.0073827 (2013).

92. Adams, A. S. *et al.* Mountain pine beetles colonizing historical and naïve host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. *Appl Environ Microbiol* **79**(11), 3468-75; doi: 10.1128/AEM.00068-13 (2013).

93. Nagamoto, N.S. *et al.* Microscopic evidence supports the hypothesis of high cellulose degradation capacity by the symbiotic fungus of leaf-cutting ants. *J Biol Res-Thessalon* **16,** 308 – 312 (2011).

94. Aylward, F.O. *et al.* Enrichment and broad representation of plant biomass-degrading enzymes in the specialized hyphal swellings of *Leucoagaricus gongylophorus,* the fungal symbiont of leaf-cutter ants. *PLoS One* **10**(8), e0134752; doi: 10.1371/journal.pone.0134752 (2015).

95. Khadempour, L. *et al.* The fungal cultivar of leaf-cutter ants produces specific enzymes in response to different plant substrates. *Mol Ecol* **25**,5795–5805 (2016).

96. Hyodo, F., Inoue, T., Azuma, J. I., Tayasu, I. & Abe, T. Role of the mutualistic fungus in lignin degradation in the fungus-growing termite *Macrotermes gilvus* (Isoptera; Macrotermitinae). *Soil Biol Biochem* **32,** 653-658 (2000).

97. Hyodo, F. *et al.* Differential role of symbiotic fungi in lignin degradation and food provision for fungus-growing termites (Macrotermitinae: Isoptera). *Funct Ecol* **17,**186–193 (2003).

98. Valiev, A., Ogel, Z. B. & Klepzig, K. D. Analysis of cellulase and polyphenol oxidase production by southern pine beetle associated fungi. *Symbiosis* **49**, 37-42 (2009).

99. Bittleston LS, Pierce NE, Ellison AM, Pringle A. Convergence in multispecies interactions. *Trends Ecol Evol* **31**(4), 269-280 (2016).

100. Ghurye, J. S., Cepeda-Espinoza, V. & Pop, M. Metagenomic Assembly: Overview, Challenges and Applications. **Yale J Biol Med 89**: 353-362 (2016).

101. Misof, B. *et al.* Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**(6210),763-767 (2014).

102. Rossmassler, K. *et al.* Metagenomic analysis of the microbiota in the highly compartmented hindguts of six wood- or soil-feeding higher termites. *Microbiome* **3,** 56 doi: 10.1186/s40168-015-0118-1 (2015).

103. Vargas-Asensio, G. *et al.* Uncovering the cultivable microbial diversity of costa rican beetles and its ability to break down plant cell wall components. *PLoS One* **9**(11); e113303. doi: 10.1371/journal.pone.0113303 (2014).

104. Engel, P., Martinson, V. G. & Moran, N. A. Functional diversity within the simple gut microbiota of the honey bee. *Proc. Natl Acad. Sci. USA* **109,** 11002–11007 (2012).

105. VanInsberghe, D. *et al.* Non-symbiotic Bradyrhizobium ecotypes dominate North American forest soils. *ISME J* **9,** 2435–2441; doi:10.1038/ismej.2015.54 (2015).

106. Anantharaman, K. *et al.* Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* **7,** 13219; doi: 10.1038/ncomms13219 (2016).

107. White, R. A., III. *et al.* Moleculo long-read sequencing facilitates assembly and genomic binning from complex soil metagenomes. *mSystems* **1**(3), e00045-16; doi:10.1128/mSystems.00045-16 (2016).

Figure legends

Fig 1 – Fungus-growing insects. A schematic illustration of insects fungiculture systems. A) Lower attines (illustration based on *Mycocepurus goeldii*) forage for flower parts, seeds, wood fragments, plant debris, arthropods feces and carcasses to cultivate basidiomycete fungal symbionts.[12,19,20,22] B) Higher attines (illustration based on *Atta cephalotes*) cultivate a fungal symbiont which produces swollen hyphae, containing nutrients, named gongylidia, from which all ant larvae and most of adult ants feed on.[12,22-24] C) Ambrosia beetles (illustration based on *Xyleborus affinis*) cultivate their fungal symbiont within tunnel systems – also known as galleries – in the xylem of trees. The fungal symbiont is the sole nutritional source for larvae and adults.[18,25,27] For bark beetles, the main nutritional source is phloem, and the benefits from the association with fungi are not completely clear.[27] D) In fungus-growing termites' colonies (illustration based on *Odontotermes formosanus*), younger workers ingest plant substrate which is fragmented and mixed with *Termitomyces* spores, resulting in the substrate known as fresh fungus combs. *Termitomyces* hyphae grow within the comb depleting plant biomass, ultimately becoming the mature comb which is consumed by older workers.[30,31] E) Composition of CAZymes in microbiomes from fungus-growing insects. Simplified phylogeny of insect orders based on the reported by Misof *et al*.[101] Bars show the distribution in standardized abundance of CAZy family members across different fungiculture systems. Labeled in green: fungus growing termites fungiculture. Labeled in light blue: lower Attini fungiculture. Labeled in dark blue: higher Attini fungiculture. Labeled in brown: ambrosia beetle fungiculture. Labeled in orange: bark beetle fungiculture. GH: glycoside hydrolases; GT: glycoside transferases; AA: auxiliary activities; CE: carbohydrate esterase, CBM: carbohydrate-binding modules, PL: polysaccharide lyases.

Fig 2 – Taxonomic structure of bacterial communities from *Mycocepurus goeldii* and *Atta sexdens rubropilosa* fungus gardens. A) Bacterial genera were predicted through taxonomic assignment of protein coding genes of both metagenomes, and were used to (B) estimate diversity indices and the number of genera shared (Venn diagram).

Fig 3 – Functional profile of bacterial communities from *Mycocepurus goeldii* and *Atta sexdens rubropilosa* fungus gardens. Functions were grouped per their known role in metabolic interactions and reactions. Functions in black indicate no significant differences between communities (Fisher's exact test with Bonferroni correction, $P > 0.05$). Functions in red text indicate significant differences (Fisher's exact test with Bonferroni correction, $P < 0.05$).

Fig. 4 – Functional similarity of bacterial communities from insects fungiculture and insects gut. Non-metric multidimensional scaling (NMDS) ordination was based on Correlation distances, using the standardized abundance of protein-coding genes annotated with: A) Pfam; B) COG; C) KO, and D) CAZy. E) The class-level microbiota composition ordination was

estimated using the standardized abundance of protein-coding genes taxonomic-assigned to Class level. All ANOSIM *P* values were significant ($P<0.05$, corrected *P* values). *R* values are usually between 0 and 1; lower values mean that the two groups evaluated are not significantly different, while higher values represent significant differences between the two groups. Samples indicated with an arrow were assembled and annotated in this study.

Fig. 5 – Microbiomes from fungiculture systems and insects gut exhibit similarities in several metabolic pathways. A) Standardized abundance of genes classified through KEGG pathways as related to carbohydrates metabolism, amino acids metabolism, energy metabolism, lipids metabolism, cofactors and vitamins metabolism, and aromatic compounds metabolism. Non-metric multidimensional scaling (NMDS) ordination was based on Correlation distances for: B) All KEGG pathways; C) Carbohydrates metabolism; D) Lipids metabolism; E) Aromatic compounds metabolism; F) Energy metabolism; G) Amino acids metabolism; H) Cofactors and vitamins metabolism. All ANOSIM *P* values were significant ($P<0.05$, corrected *P* values). *R* values fall usually between 0 and 1; lower values mean that the two groups evaluated are not significantly different, while higher values represent significant differences between the two groups. Samples indicated with an arrow were assembled and annotated in this study.

Fig. 6 – Phylogeny inference and class-level microbiota composition. Phylogenetic reconstructions were carried out through alignment-free *k*-mer based approach (*k*-mer length= 23). The standardized relative abundances of bacterial classes were estimated using taxonomic assigned protein-coding genes. Colors in the phylogenetic tree represent the groups: lower Attini (light blue), higher Attini (dark blue), fungus growing termite (light green), ambrosia and bark beetles (orange), soil (gray), Passalidae beetle gut (brown), termite gut (dark green), honey bee gut (purple).

Table and Figures

Table 1 – Summary of assembling and annotation statistics of the community metagenomes from fungus gardens of two Attini ants.

	M. goeldii	*A. sexdens rubropilosa*
Number of contigs	267 959	139 603
Size of assembled data (Mbp)	364.1	249.6
Mean contig length (bp)[1]	1358.97	1788.20
Largest contig (kbp) [1]	1040.6	716.7
Mean GC content	49.45 ± 12.58 %	57.53 ± 8.14 %
N50 contig size (bp) [1]	4 324	5 495
Contigs > 100 kbp	229	153
Number of annotated sequences	151 313	82 170
Protein coding genes	395 806	269 449
COG annotation (%)	62.62	75.17
Pfam annotation (%)	67.82	76.45
KO annotation (%)	44.03	54.32
KEGG Enzymes annotation (%)	23.60	26.73

[1]Quality and length of assembled contigs are considered good according to described by Ghurye *et al*.[100]

Fig. 1

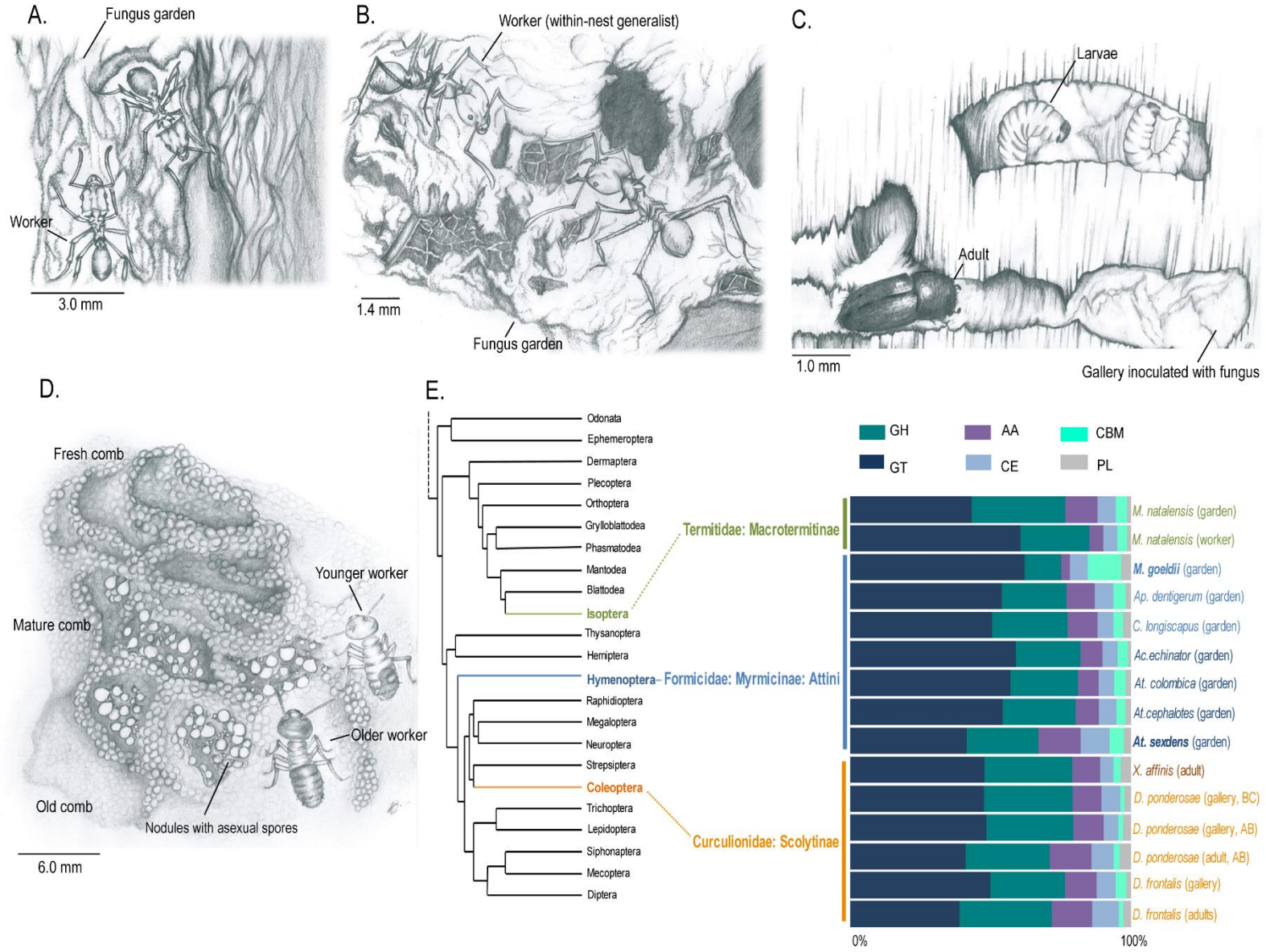

A.
Fungus garden
Worker
3.0 mm
B.
Worker (within-nest generalist)
1.4 mm
Fungus garden
C.
Larvae
Adult
1.0 mm
Gallery inoculated with fungus
D.
Fresh comb
Younger worker
Mature comb
Older worker
Old comb
Nodules with asexual spores
6.0 mm
E.
Odonata
Ephemeroptera
Dermaptera
Plecoptera
Orthoptera
Grylloblattodea
Phasmatodea
Mantodea
Blattodea
Isoptera
Thysanoptera
Hemiptera
Hymenoptera
Raphidioptera
Megaloptera
Neuroptera
Strepsiptera
Coleoptera
Trichoptera
Lepidoptera
Siphonaptera
Mecoptera
Diptera
Termitidae: Macrotermitinae
Formicidae: Myrmicinae: Attini
Curculionidae: Scolytinae
GH
AA
CBM
GT
CE
PL
M. natalensis (garden)
M. natalensis (worker)
M. goeldii (garden)
Ap. dentigerum (garden)
C. longiscapus (garden)
Ac.echinator (garden)
At. colombica (garden)
At.cephalotes (garden)
At. sexdens (garden)
X. affinis (adult)
D. ponderosae (gallery, BC)
D. ponderosae (gallery, AB)
D. ponderosae (adult, AB)
D. frontalis (gallery)
D. frontalis (adults)
0%
100%

Fig. 2

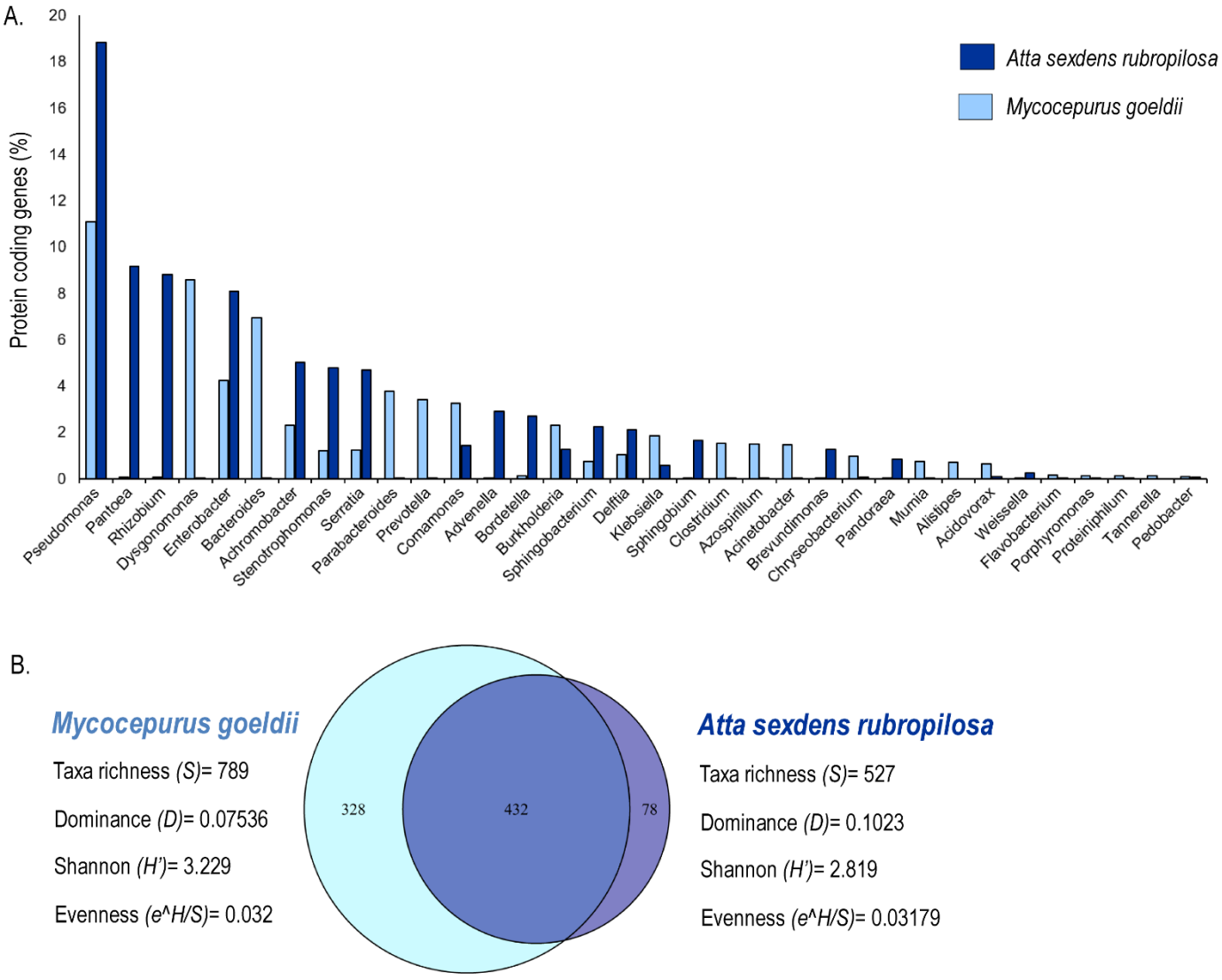

A.
Protein coding genes (%)
20
18
16
14
12
10
8
6
4
2
0
Atta sexdens rubropilosa
Mycocepurus goeldii
Pseudomonas
Pantoea
Rhizobium
Dysgonomonas
Enterobacter
Bacteroides
Achromobacter
Stenotrophomonas
Serratia
Parabacteroides
Prevotella
Comamonas
Advenella
Bordetella
Burkholderia
Sphingobacterium
Delftia
Klebsiella
Sphingobium
Clostridium
Azospirillum
Acinetobacter
Brevundimonas
Chryseobacterium
Pandoraea
Mumia
Alistipes
Acidovorax
Weissella
Flavobacterium
Porphyromonas
Proteiniphilum
Tannerella
Pedobacter
B.
Mycocepurus goeldii
Taxa richness (S)= 789
Dominance (D)= 0.07536
Shannon (H')= 3.229
Evenness (e^H/S)= 0.032
328
432
78
Atta sexdens rubropilosa
Taxa richness (S)= 527
Dominance (D)= 0.1023
Shannon (H')= 2.819
Evenness (e^H/S)= 0.03179

Fig. 3

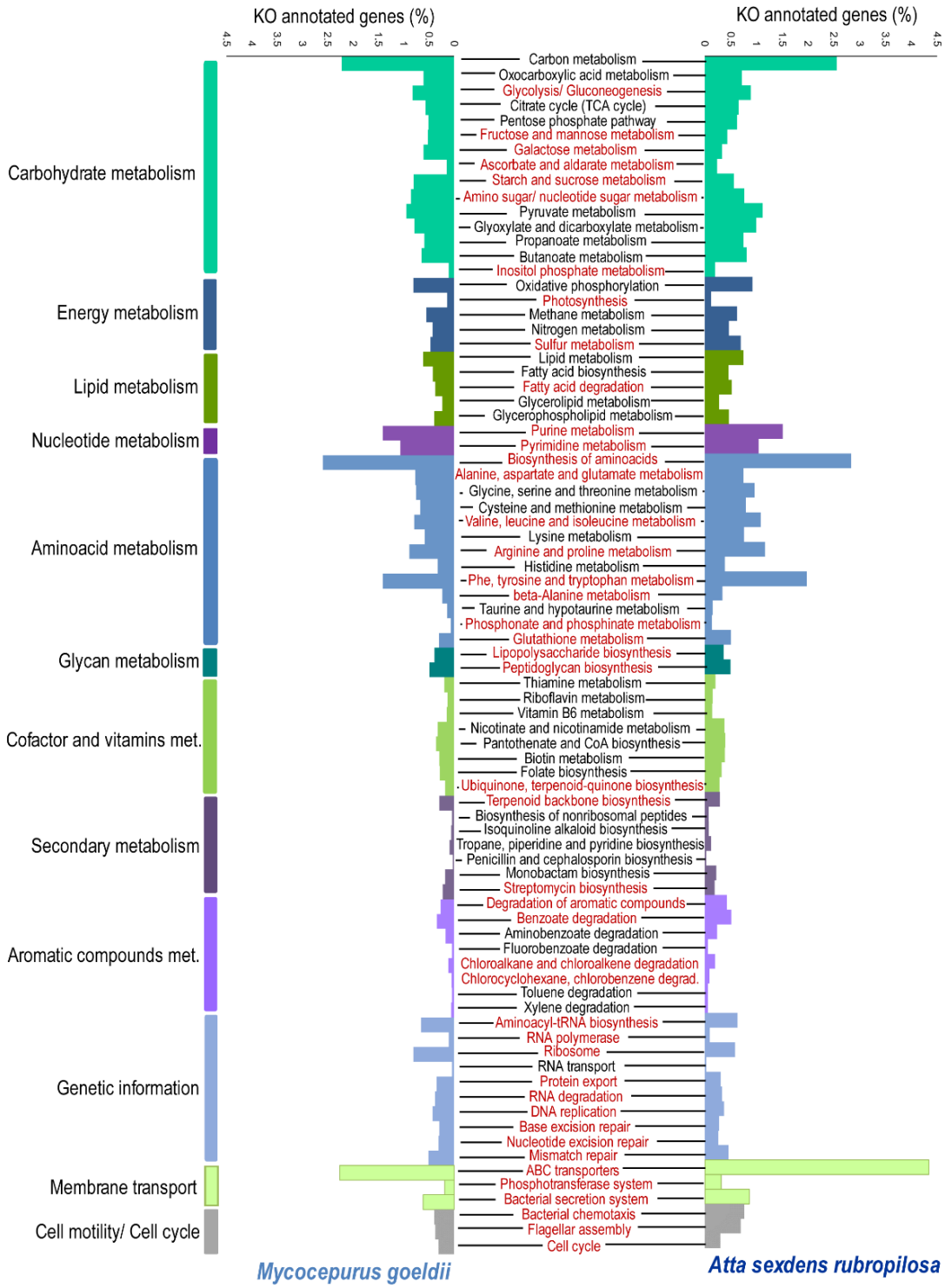

KO annotated genes (%)
KO annotated genes (%)
4.5 4 3.5 3 2.5 2 1.5 1 0.5 0
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Carbohydrate metabolism
Energy metabolism
Lipid metabolism
Nucleotide metabolism
Aminoacid metabolism
Glycan metabolism
Cofactor and vitamins met.
Secondary metabolism
Aromatic compounds met.
Genetic information
Membrane transport
Cell motility/ Cell cycle
Carbon metabolism
Oxocarboxylic acid metabolism
Glycolysis/ Gluconeogenesis
Citrate cycle (TCA cycle)
Pentose phosphate pathway
Fructose and mannose metabolism
Galactose metabolism
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Amino sugar/ nucleotide sugar metabolism
Pyruvate metabolism
Glyoxylate and dicarboxylate metabolism
Propanoate metabolism
Butanoate metabolism
Inositol phosphate metabolism
Oxidative phosphorylation
Photosynthesis
Methane metabolism
Nitrogen metabolism
Sulfur metabolism
Lipid metabolism
Fatty acid biosynthesis
Fatty acid degradation
Glycerolipid metabolism
Glycerophospholipid metabolism
Purine metabolism
Pyrimidine metabolism
Biosynthesis of aminoacids
Alanine, aspartate and glutamate metabolism
Glycine, serine and threonine metabolism
Cysteine and methionine metabolism
Valine, leucine and isoleucine metabolism
Lysine metabolism
Arginine and proline metabolism
Histidine metabolism
Phe, tyrosine and tryptophan metabolism
beta-Alanine metabolism
Taurine and hypotaurine metabolism
Phosphonate and phosphinate metabolism
Glutathione metabolism
Lipopolysaccharide biosynthesis
Peptidoglycan biosynthesis
Thiamine metabolism
Riboflavin metabolism
Vitamin B6 metabolism
Nicotinate and nicotinamide metabolism
Pantothenate and CoA biosynthesis
Biotin metabolism
Folate biosynthesis
Ubiquinone, terpenoid-quinone biosynthesis
Terpenoid backbone biosynthesis
Biosynthesis of nonribosomal peptides
Isoquinoline alkaloid biosynthesis
Tropane, piperidine and pyridine biosynthesis
Penicillin and cephalosporin biosynthesis
Monobactam biosynthesis
Streptomycin biosynthesis
Degradation of aromatic compounds
Benzoate degradation
Aminobenzoate degradation
Fluorobenzoate degradation
Chloroalkane and chloroalkene degradation
Chlorocyclohexane, chlorobenzene degrad.
Toluene degradation
Xylene degradation
Aminoacyl-tRNA biosynthesis
RNA polymerase
Ribosome
RNA transport
Protein export
RNA degradation
DNA replication
Base excision repair
Nucleotide excision repair
Mismatch repair
ABC transporters
Phosphotransferase system
Bacterial secretion system
Bacterial chemotaxis
Flagellar assembly
Cell cycle
Mycocepurus goeldii
Atta sexdens rubropilosa

Fig. 4

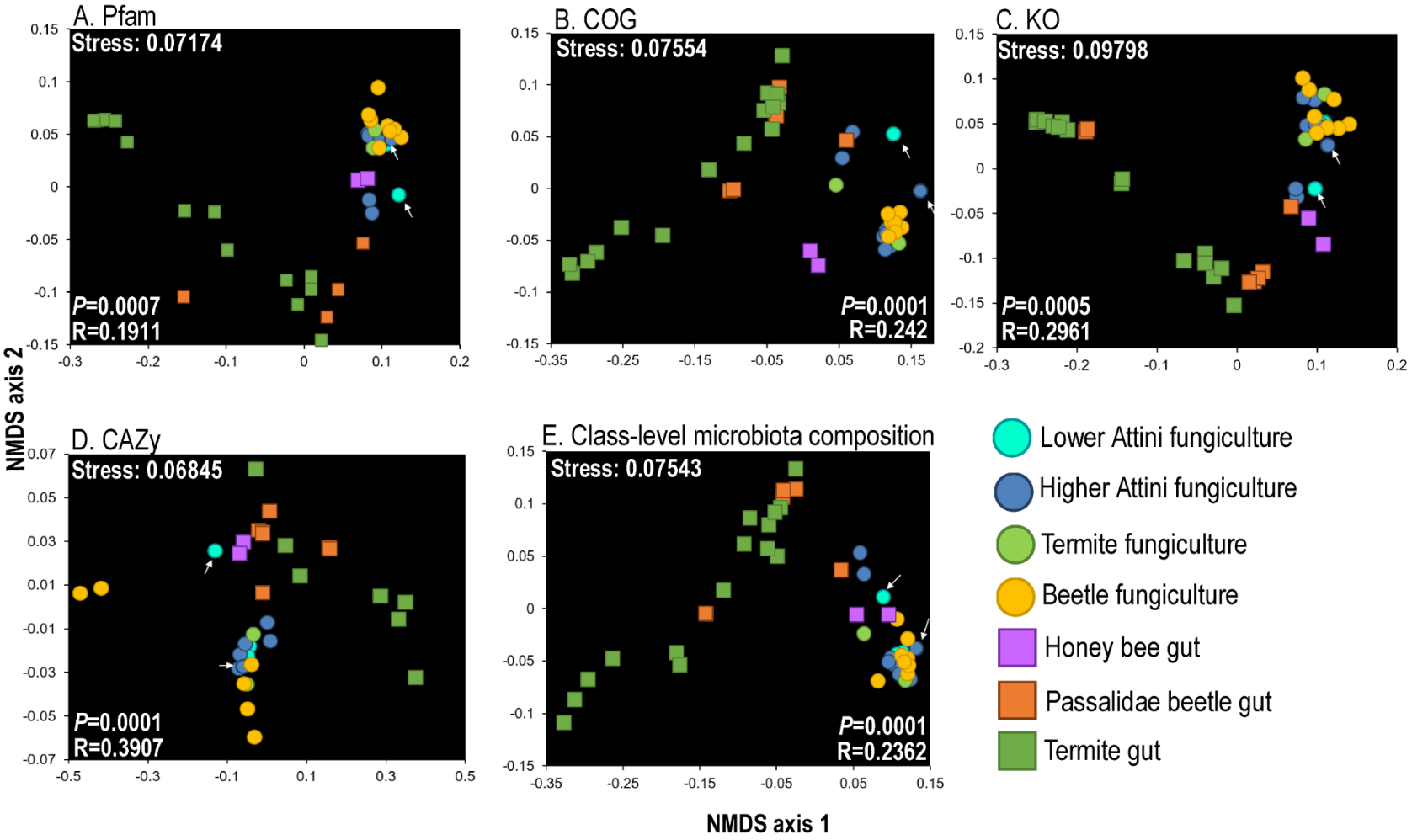

A. Pfam
Stress: 0.07174
P=0.0007
R=0.1911
B. COG
Stress: 0.07554
P=0.0001
R=0.242
C. KO
Stress: 0.09798
P=0.0005
R=0.2961
D. CAZy
Stress: 0.06845
P=0.0001
R=0.3907
E. Class-level microbiota composition
Stress: 0.07543
P=0.0001
R=0.2362
NMDS axis 2
NMDS axis 1
Lower Attini fungiculture
Higher Attini fungiculture
Termite fungiculture
Beetle fungiculture
Honey bee gut
Passalidae beetle gut
Termite gut

Fig. 5

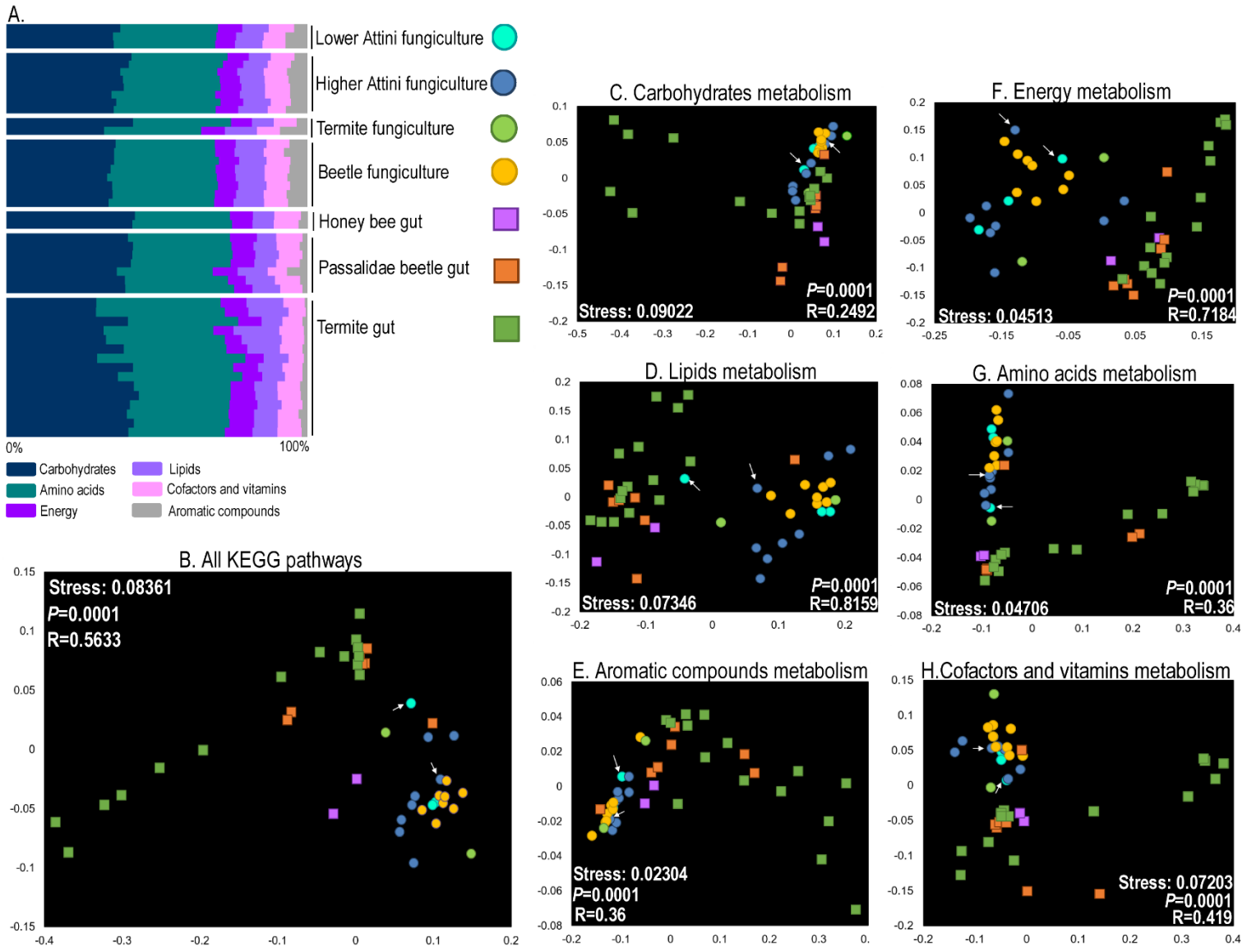
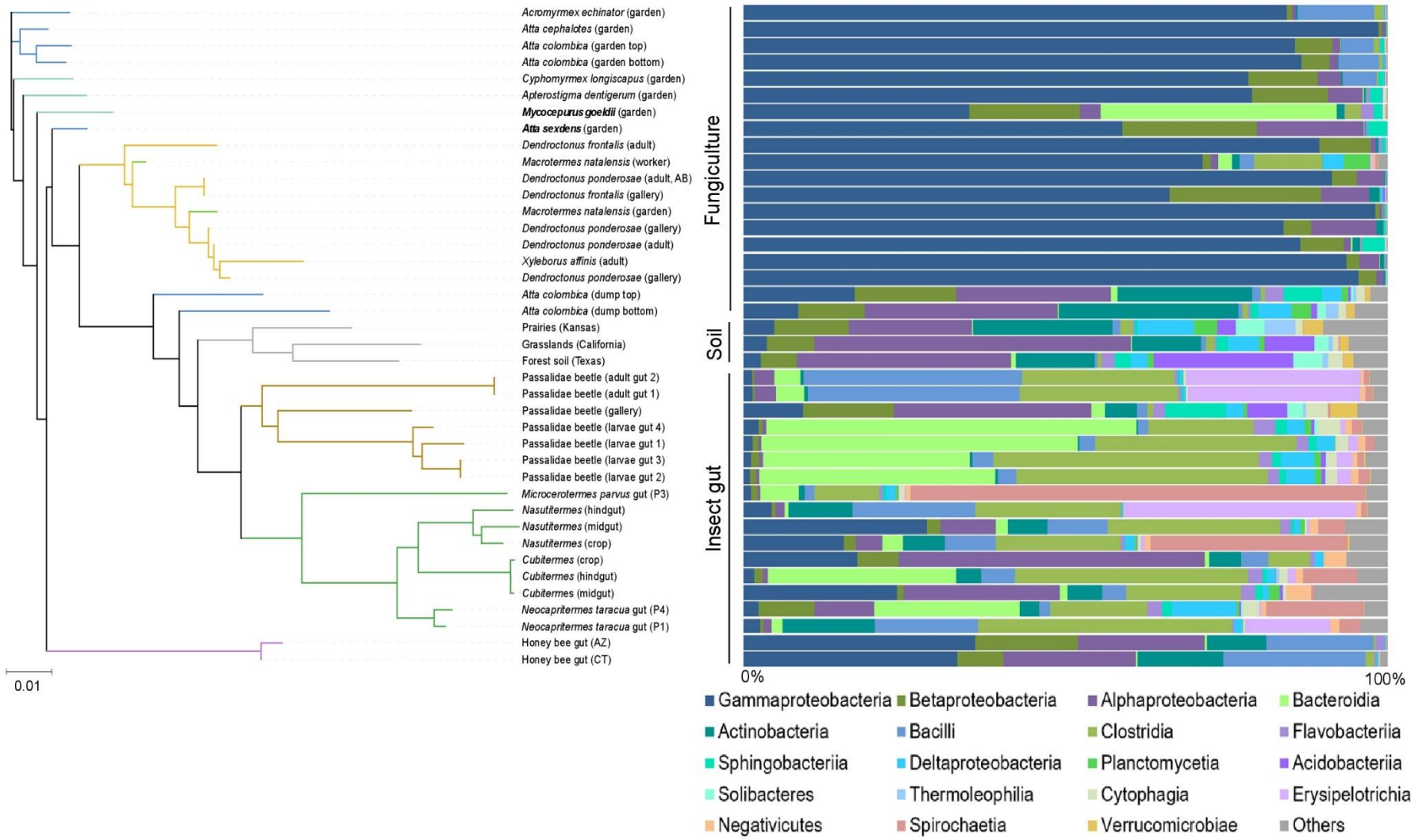

A.
Lower Attini fungiculture
Higher Attini fungiculture
Termite fungiculture
Beetle fungiculture
Honey bee gut
Passalidae beetle gut
Termite gut
0%
100%
Carbohydrates
Lipids
Amino acids
Cofactors and vitamins
Energy
Aromatic compounds
B. All KEGG pathways
Stress: 0.08361
P=0.0001
R=0.5633
C. Carbohydrates metabolism
Stress: 0.09022
P=0.0001
R=0.2492
D. Lipids metabolism
Stress: 0.07346
P=0.0001
R=0.8159
E. Aromatic compounds metabolism
Stress: 0.02304
P=0.0001
R=0.36
F. Energy metabolism
Stress: 0.04513
P=0.0001
R=0.7184
G. Amino acids metabolism
Stress: 0.04706
P=0.0001
R=0.36
H.Cofactors and vitamins metabolism
Stress: 0.07203
P=0.0001
R=0.419

Fig. 6


Acromyrmex echinator (garden)
Atta cephalotes (garden)
Atta colombica (garden top)
Atta colombica (garden bottom)
Cyphomyrmex longiscapus (garden)
Apterostigma dentigerum (garden)
Mycocepurus goeldii (garden)
Atta sexdens (garden)
Dendroctonus frontalis (adult)
Macrotermes natalensis (worker)
Dendroctonus ponderosae (adult, AB)
Dendroctonus frontalis (gallery)
Macrotermes natalensis (garden)
Dendroctonus ponderosae (gallery)
Dendroctonus ponderosae (adult)
Xyleborus affinis (adult)
Dendroctonus ponderosae (gallery)
Atta colombica (dump top)
Atta colombica (dump bottom)
Prairies (Kansas)
Grasslands (California)
Forest soil (Texas)
Passalidae beetle (adult gut 2)
Passalidae beetle (adult gut 1)
Passalidae beetle (gallery)
Passalidae beetle (larvae gut 4)
Passalidae beetle (larvae gut 1)
Passalidae beetle (larvae gut 3)
Passalidae beetle (larvae gut 2)
Microcerotermes parvus gut (P3)
Nasutitermes (hindgut)
Nasutitermes (midgut)
Nasutitermes (crop)
Cubitermes (crop)
Cubitermes (hindgut)
Cubitermes (midgut)
Neocapritermes taracua gut (P4)
Neocapritermes taracua gut (P1)
Honey bee gut (AZ)
Honey bee gut (CT)
0.01
Fungiculture
Soil
Insect gut
0%
100%
Gammaproteobacteria
Betaproteobacteria
Alphaproteobacteria
Bacteroidia
Actinobacteria
Bacilli
Clostridia
Flavobacteriia
Sphingobacteriia
Deltaproteobacteria
Planctomycetia
Acidobacteriia
Solibacteres
Thermoleophilia
Cytophagia
Erysipelotrichia
Negativicutes
Spirochaetia
Verrucomicrobiae
Others

Supplemental material

Table S1- Microbiomes utilized in functional and taxonomic comparisons.

Classification	Genome Name / Sample Name	IMG Genome ID	Reference study
Lower Attini	*Apterostigma dentigerum* (fungus garden)	2029527003	Aylward *et al.*[42]
Lower Attini	*Mycocepurus goeldii* (fungus garden)	3300009856	This study
Lower Attini	*Cyphomyrmex longiscapus* (fungus garden)	2030936005	Aylward *et al.*[42]
Higher Attini	*Trachymyrmex* (fungus garden)	2084038018	Aylward *et al.*[42]
Higher Attini	*Acromyrmex echinatior* (fungus garden)	2035918000	Aylward *et al.*[42]
Higher Attini	*Atta colombica* (fungus garden top)	2029527005	Aylward *et al.*[42]
Higher Attini	*Atta colombica* (fungus garden bottom)	2029527006	Aylward *et al.*[42]
Higher Attini	*Atta colombica* (dump bottom)	2040502000	Aylward *et al.*[42]
Higher Attini	*Atta colombica* (dump top)	2038011000	Aylward *et al.*[42]
Higher Attini	*Atta cephalotes* (fungus garden)	2029527004	Aylward *et al.*[42]
Higher Attini	*Atta sexdens rubropilosa* (fungus garden)	3300009944	This study
Fungus-growing termites	*Macrotermes natalensis* (fungus garden)	2065487014	Aylward *et al.*[42]
Fungus-growing termites	*Macrotermes natalensis* (worker)	2065487013	Aylward *et al.*[42]
Ambrosia beetle	*Xyleborus affinis* (adult)	2043231000	Aylward *et al.*[42]
Ambrosia beetle	*Xyleborus affinis* (gallery)	2084038008	Aylward *et al.*[42]
Bark beetle	*Dendroctonus ponderosae* (gallery)	2029527007	Aylward *et al.*[42]
Bark beetle	*Dendroctonus ponderosae* (adult)	2032320009	Aylward *et al.*[42]
Bark beetle	*Dendroctonus ponderosae* (adult)	2035918003	Aylward *et al.*[42]
Bark beetle	*Dendroctonus frontalis* (adult)	2044078006	Aylward *et al.*[42]
Bark beetle	*Dendroctonus frontalis* (gallery)	2044078007	Aylward *et al.*[42]
Bark beetle	*Dendroctonus ponderosae* (gallery)	2032320008	Aylward *et al.*[42]
Termite gut	*Cubitermes ugandensis* foregut (crop)	3300001474	Rossmassler *et al.*[102]
Termite gut	*Cubitermes ugandensis* midgut	3300001468	Rossmassler *et al.*[102]

Termite gut	*Cubitermes ugandensis* hindgut (P1 segment)	3300002185	Rossmassler *et al.*[102]
Termite gut	Nasutitermes corniger foregut (crop)	3300001542	Rossmassler *et al.*[102]
Termite gut	*Nasutitermes corniger* (midgut)	3300001466	Rossmassler *et al.*[102]
Termite gut	*Nasutitermes corniger* hindgut (P1 segment)	3300001541	Rossmassler *et al.*[102]
Termite gut	*Neocapritermes taracua* hindgut (P1 segment)	3300002501	Rossmassler *et al.*[102]
Termite gut	*Cornitermes* sp. hindgut (P1 segment)	3300002552	Rossmassler *et al.*[102]
Termite gut	*Termes hospes* hindgut (P1 segment)	3300002508	Rossmassler *et al.*[102]
Termite gut	*Microcerotermes parvus* hindgut (P3 segment)	3300002449	Rossmassler *et al.*[102]
Termite gut	*Termes hospes* hindgut (P3 segment)	3300006226	Rossmassler *et al.*[102]
Termite gut	*Neocapritermes taracua* hindgut (P3 segment)	3300006045	Rossmassler *et al.*[102]
Termite gut	*Neocapritermes taracua* hindgut (P4 segment)	3300002504	Rossmassler *et al.*[102]
Termite gut	*Cornitermes* sp. hindgut (P4 segment)	3300002834	Rossmassler *et al.*[102]
Termite gut	*Termes hospes* hindgut (P4 segment, Th196)	3300002462	Rossmassler *et al.*[102]
Passalidae beetle gut	Passalidae beetle (adult gut 2)	3300000097	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (larvae gut 3)	3300000114	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (larvae gut 1)	3300000062	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (larvae gut 2)	3300000839	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (larvae gut 4)	3300000838	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (gallery)	3300000036	Vargas-Asensio *et al.*[103]
Passalidae beetle gut	Passalidae beetle (adult gut 1)	3300000836	Vargas-Asensio *et al.*[103]
Honey bee gut	Honey bee gut (AZ)	3300005721	Engel *et al.*[104]
Honey bee gut	Honey bee gut (CT)	3300000333	Engel *et al.*[104]
Soil	Forest soil (Texas)	3300001143	VanInsberghe *et al.*[105]
Soil	Grasslands (California)	3300002886	Anantharaman *et al.*[106]
Soil	Prairies (Kansas)	2067725004	White *et al.*[107]

Fig S1 – Bacterial enrichment protocol

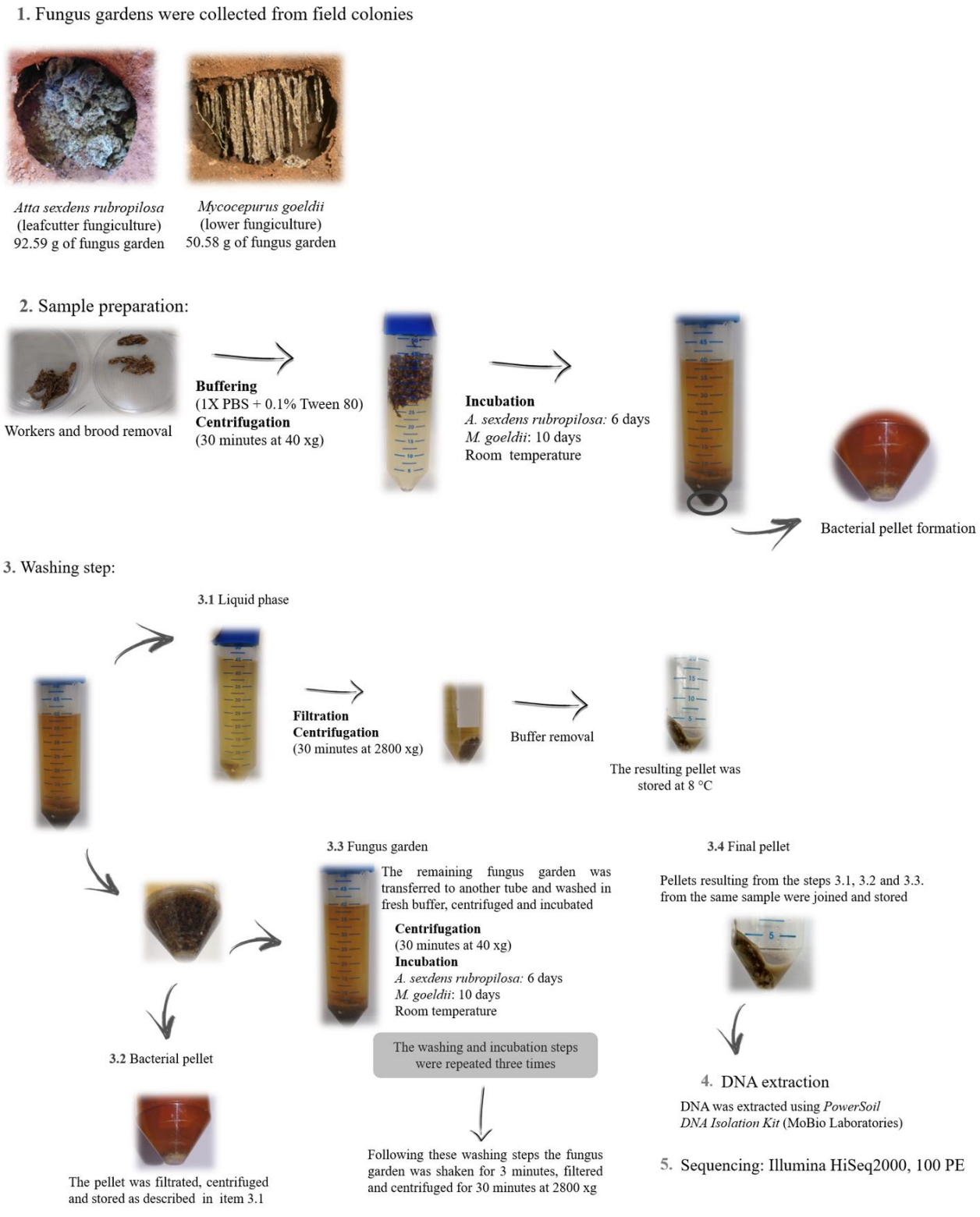


Fig S2- Most abundant phyla in the bacterial community. Normalized abundance of taxonomic-assigned protein coding genes. Differences between *M. goeldii* and *At. sexdens* estimated by Fisher's exact test ($P<0.05$).

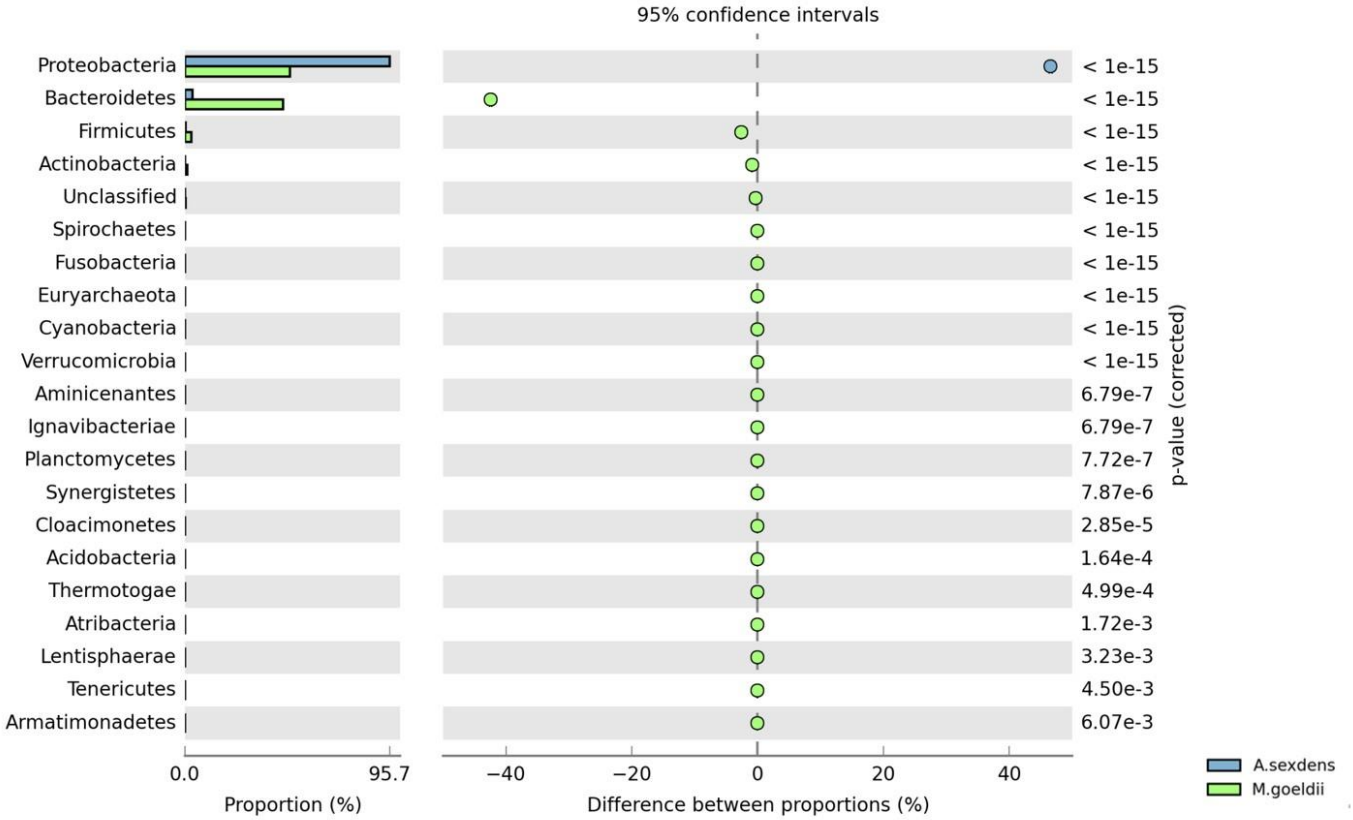


Fig S3- Genera represented in *M. goeldii* and *At. sexdens* microbiomes through alignment of the metagenomes with a 16S rDNA database

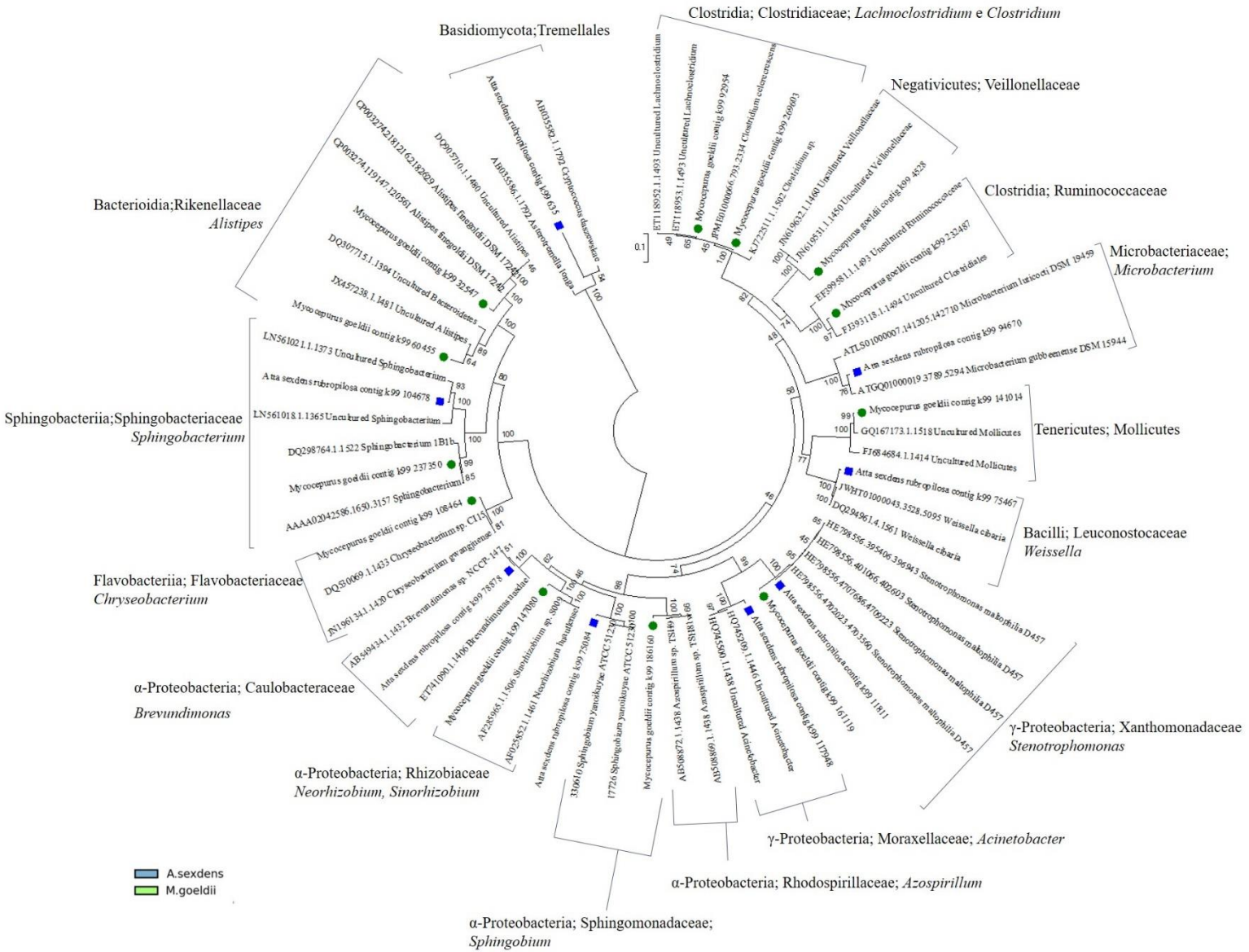


Fig S4- Taxonomic assignment of sequences > 100Kbp

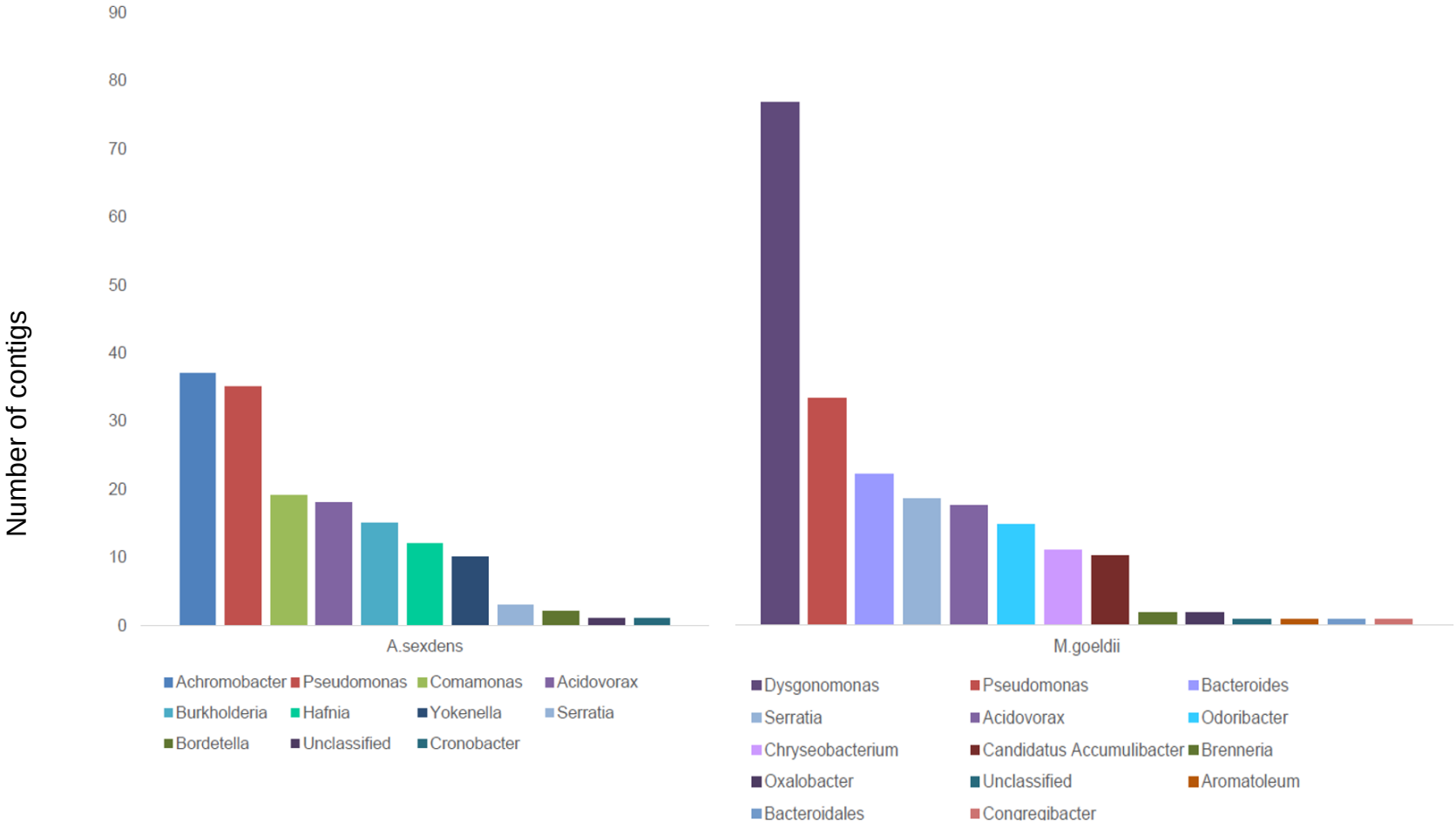


Fig S5- Supplemental Metagenomes from fungiculture and non-fungiculture systems with distinct enzymatic profiles. The normalized abundance of KEGG annotated genes were compared between fungiculture systems and non-fungiculture systems, and the enzymes with significant differences were classified according to its metabolic role/pathway through MetaCyc. Details about normalized abundances of each function within a specific metabolic pathway can be found at Dataset S5.


Number of annotated genes
(normalized values)
0.3
0
Mycocepurus goeldii (garden)
Apterostigma dentigerum (garden)
Cyphomyrmex longiscapus (garden)
Trachymyrmex (garden)
Acromyrmex echinator (garden)
Atta colombica (garden top)
Atta colombica (garden bottom)
Atta cephalotes (garden)
Atta sexdens (garden)
Atta colombica (dump bottom)
Dendroctonus frontalis (dump top)
Dendroctonus frontalis (adult)
Dendroctonus frontalis (gallery)
Dendroctonus ponderosae (adult, AB)
Dendroctonus ponderosae (gallery, AB)
Dendroctonus ponderosae (adult, BC)
Dendroctonus ponderosae (gallery, BC)
Xyleborus affinis (adult)
Xyleborus affinis (gallery)
Macrotermes natalensis (worker)
Macrotermes natalensis (garden)
Passalidae beetle (gallery)
Passalidae beetle (adult gut 1)
Passalidae beetle (adult gut 2)
Passalidae beetle (larvae gut 1)
Passalidae beetle (larvae gut 2)
Passalidae beetle (larvae gut 3)
Passalidae beetle (larvae gut 4)
Cubitermes (gut)
Cubitermes (crop)
Nasutitermes (midgut)
Nasutitermes (crop)
Nasutitermes (midgut)
Termite gut (gut)
Termite gut (P1)
Termite gut (P3)
Termite gut (P4)
Forest soil (Texas)
Grasslands (California)
Prairies (Kansas)
Alcohols and aldehydes degradation/assimilation
Carboxylates degradation/assimilation
Glycans, sugar and sugar derivatives degradation/ assimilation/transportation
Proteins and amino acids degradation/assimilation
Lipids degradation/assimilation
Aromatic compounds degradation/assimilation
Energetic metabolism
Polysaccharides biosynthesis
Amino acids biosynthesis
Lipids biosynthesis
Aromatic compounds biosynthesis
Cofactors, prosthetic groups, electron carriers, siderophores biosynthesis
Secondary metabolism
Nucleosides, nucleotides and polynucleotides metabolism
Amines and polyamines metabolism
Detoxification
Nitrogen metabolism
Inorganic compounds metabolism
Fungiculture
Non-fungiculture

Dataset S1- Normalized abundance of predicted genes annotated with COG, Pfam, KO and KEGG Enzymes were compared between the two communities. Available at: https://www.dropbox.com/s/a4adrezz4noe23d/Dataset_S1_Atta_Mycocepurus_Functions.xlsx?dl=0

Dataset S2- Normalized abundance of genes annotated with COG, Pfam, KO, KEGG Enzymes, and CAZy were compared between bacterial communities from lower attini and higher attini fungus gardens. Available at: https://www.dropbox.com/s/34mgvwruj7exik2/Dataset2_Lower_Higher_Attini_Functions.xlsx?dl=0

Dataset S3- Comparisons of the normalized abundance of CAZy annotated genes among fungus-growing insects. Available at: https://www.dropbox.com/s/7azpqxsm9y44k38/Dataset_S3_CAZy_families_distribution.xlsx?dl=0

Dataset S4 - KEGG Enzymes functions significantly different (Kruskal-Wallis test with Bonferroni correction, $P<0.05$) between fungiculture systems (metagenomes from Attini ants, Macrotermitinae termites and Scolytinae beetles) and non-fungiculture systems (insects' gut and soil metagenomes). Available at: https://www.dropbox.com/s/9s83n235ulluf7v/Dataset_S4_Enzymes_Significant_Difference.xlsx?dl=0

Dataset S5- KEGG Enzymes functions significantly different (Kruskal-Wallis test with Bonferroni correction, $P<0.05$) classified according to metabolic pathways in the MetaCyc database. Available at: https://www.dropbox.com/s/dtdg87fahcd8xa9/Dataset_S5_Enzymes_Classification.xlsx?dl=0

Dataset S6- Relative abundance (normalized data) of bacterial class in fungiculture and non-fungiculture metagenomes. Bacterial class were predicted from taxonomic assignment of protein-coding genes. Bacterial class with relative abundance <0.05% were designed as "others". Available at: https://www.dropbox.com/s/ii2z8ckx8d8r2fe/Dataset_S6_Class_Distribution_Percentage.xlsx?dl=0