PNAS RESEARCH ARTICLE | EVOLUTION

Genomic diversification of the specialized parasite of the fungus-growing ant symbiosis

Kirsten Gotting[a,b], Daniel S. May[a], Jeffrey Sosa-Calvo[c], Lily Khadempour[d], Charlotte B. Francoeur[a], Aileen Berasategui[e], Margaret W. Thairu[a], Shelby Sandstrom[a], Caitlin M. Carlson[a], Marc G. Chevrette[f,g], Mônica T. Pupo[h], Tim S. Bugni[i], Ted R. Schultz[c], J. Spencer Johnston[j], Nicole M. Gerardo[e], and Cameron R. Currie[a,k,1]

Edited by Joan Strassmann, Washington University in St. Louis, St. Louis, MO; received August 22, 2022; accepted October 20, 2022

Fungi shape the diversity of life. Characterizing the evolution of fungi is critical to understanding symbiotic associations across kingdoms. In this study, we investigate the genomic and metabolomic diversity of the genus *Escovopsis*, a specialized parasite of fungus-growing ant gardens. Based on 25 high-quality draft genomes, we show that *Escovopsis* forms a monophyletic group arising from a mycoparasitic fungal ancestor 61.82 million years ago (Mya). Across the evolutionary history of fungus-growing ants, the dates of origin of most clades of *Escovopsis* correspond to the dates of origin of the fungus-growing ants whose gardens they parasitize. We reveal that genome reduction, determined by both genomic sequencing and flow cytometry, is a consistent feature across the genus *Escovopsis,* largely occurring in coding regions, specifically in the form of gene loss and reductions in copy numbers of genes. All functional gene categories have reduced copy numbers, but resistance and virulence genes maintain functional diversity. Biosynthetic gene clusters (BGCs) contribute to phylogenetic differences among *Escovopsis* spp., and sister taxa in the Hypocreaceae. The phylogenetic patterns of co-diversification among BGCs are similarly exhibited across mass spectrometry analyses of the metabolomes of *Escovopsis* and their sister taxa. Taken together, our results indicate that *Escovopsis* spp. evolved unique genomic repertoires to specialize on the fungus-growing ant-microbe symbiosis.

attine | symbiosis | parasitism | *Escovopsis*

Significance

Fungi represent a kingdom of organisms with diverse interspecies associations and broad variability in genome content. The fungal genus *Escovopsis* is a parasite of the multipartite fungus-growing ant symbiosis. We present this genus as a eukaryotic model of genomic reduction and diversification over the 60-million-year-old history of the fungus-growing ant symbiosis. This genomic evolution represents an example of a eukaryotic genus evolving a reduced genomic toolkit while maintaining ancient host associations.

Fungi engage in complex interspecific interactions that have helped shape the biocomplexity and biodiversity of life on earth (1). In terrestrial ecosystems, fungi play critical roles as beneficial and antagonistic symbionts. Fungi occupied the terrestrial environment prior to the first land plants and animals (2–4), and played a key role in species radiations (5) in the other two kingdoms. Interactions between fungi and other organisms are mediated by small molecules, including those encoded by groupings of genes known as biosynthetic gene clusters (BGCs) (6, 7). These molecules are often targets of evolution due to their mediation of symbiotic interactions, and frequently facilitate symbiotic specialization (8, 9). While a number of studies focusing on plant–fungal symbioses have provided key insights into the macroevolutionary dynamics of fungi (5, 10–12), exploring the evolution of animal–fungal and fungal–fungal associations within ancient clades of fungal symbionts is necessary for understanding the processes underlying major transitions in these associations over time, and how these transitions impact genomic evolution.

The genus *Escovopsis* (Ascomycota, Pezizomycotina: Hypocreales: Hypocreaceae), a specialized mycoparasite of the fungi cultivated by fungus-growing ants, is part of an emerging model system for the study of fungal symbiosis. Fungus-growing ants (Formicidae: Myrmicinae: Attini: Attina) participate in a >60-My-old obligate nutritional symbiosis with specialized fungi (Basidomycota: Agaricomycetes: Agaricales: Agaricaceae or Pterulaceae), which they farm for food. Lineages of these ant-fungus mutualists have evolved diverse methods of fungal cultivation across a broad geographic range in the American tropics, with ants and fungal cultivars exhibiting high levels of symbiont fidelity (13, 14). *Escovopsis* has only been observed within the structural matrix of the gardens (Fig. 1*A*) where the ants maintain their fungal cultivar (15) and in the refuse piles where ants discard old garden material (15–17). No free-living form of *Escovopsis* has been identified and *Escovopsis* is likely horizontally transmitted between gardens (15). *Escovopsis* gets energy from necrotrophically degrading (18) and consuming the fungus-growing ants' cultivar (19), functioning as a mycoparasite. *Escovopsis* can destabilize ant colonies by completely overwhelming the fungus-growing ant gardens (Fig. 1 *A* and *B*) and producing insecticidal molecules that disrupt ant behaviors (20, 21). To defend against *Escovopsis,* fungus-growing ants weed and groom out hyphae and conidia (16, 22). Additionally, many fungus-growing ants associate with *Pseudonocardia* (Actinobacteria)

Author affiliations: [a]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706; [b]Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI 53706; [c]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560; [d]Department of Earth and Environmental Sciences, Rutgers University, Newark, NJ 07102; [e]Department of Biology, Emory University, Atlanta, GA 30322; [f]Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI 53705; [g]Department of Plant Pathology, University of Wisconsin-Madison, Madison, WI 53705; [h]School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP 14040-903, Brazil; [i]Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin-Madison, Madison, WI 53705; [j]Department of Entomology, Texas A&M University, College Station, TX 77843; and [k]David Braley Centre for Antibiotic Discovery, Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

Author contributions: K.G., D.S.M., J.S.-C., L.K., A.B., M.G.C., M.T.P., T.S.B., T.R.S., N.M.G., and C.R.C. designed research; K.G., D.S.M., J.S.-C., L.K., M.W.T., S.S., C.M.C., and J.S.J. performed research; K.G., D.S.M., J.S.-C., and J.S.J. analyzed data; and K.G., C.B.F., T.R.S., and C.R.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: ccurrie@mcmaster.ca.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2213096119/-/DCSupplemental.

Published December 12, 2022.

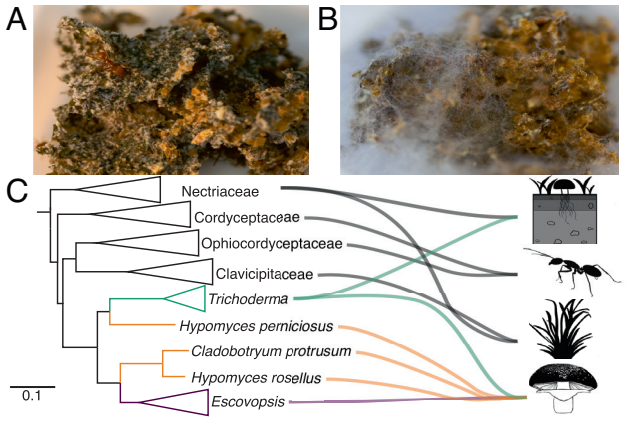

Fig. 1. *Escovopsis* parasites are descended from parasites of mushroom-forming fungi. (*A*) *Atta* fungus-growing ant with basidiomycete (Agaricales: Agaricaceae) fungus garden. Photograph by Caitlin Carlson. (*B*) Fungus Garden overgrown by *Escovopsis* parasite after experimental inoculation. Photograph by Caitlin Carlson. (*C*) The phylogenetic placement of *Escovopsis* within hypocrealean fungi and rooted to *S. cerevisiae*. Common host associations are indicated by the lines connected to the illustrations; in order from *Top* to *Bottom*: free-living, insects, plants, and mushroom-forming fungi. The colors indicate groupings of fungi used in this paper: green—*Trichoderma*, orange—*Hypomyces* and *Cladobotryum*, purple—*Escovopsis*. Illustrations by Kirsten Gotting.

that produces antifungal molecules that inhibit *Escovopsis* growth (23–25). Despite these defenses, *Escovopsis* parasitizes the gardens of the majority of fungus-growing ant lineages (15, 23, 26).

The fungus-growing ant symbiosis is characterized by phylogenetically concordant associations between the ants, their cultivars, and the specialized fungal parasites of the symbiosis in the genus *Escovopsis* (26). In total, nine species of *Escovopsis* have been described to associate with fungus-growing ant lineages across four of the five different ant agricultural strategies, and many more species, and possibly genera, are likely yet to be described (27). These strategies are based on the macroevolutionary ant-cultivar relationships described by Schultz and Brady (28) as follows: "Lower agriculture" encompasses the mutualism between the broadest range of attine ants and agaricaceous fungi in the genera *Leucocoprinus* and *Leucoagaricus* cultivated in filamentous forms. The *Escovopsis* species that associate with lower-agriculture ant lineages are *Escovopsis trichodermoides* and *Escovopsis kreiselii* (29, 30). Yeast agriculture is practiced by species within the *Cyphomyrmex rimosus* group that cultivates agaricaceous fungi in an atypical yeast-like form and is not known to be parasitized by any *Escovopsis* species (22). "Coral-fungus agriculture" is practiced by a subset of ants in the genus *Apterostigma* (the *pilosum* species group) that cultivate pterulaceous fungi in the genus *Myrmecopterula* (31). *Escovopsis clavatus* and *Escovopsis multiformis* associate with coral-fungus ant agriculture (32–34). "Higher agriculture" is practiced by attine ants that cultivate a distinct clade of agaricaceous fungi derived from lower-agriculture fungi that are parasitized by *Escovopsis moelleri, E. microspora, E. lentecrescens, E. weberi,* and *E. aspergilloides* (35–37). Within higher agriculture, there are three ant genera, *Amoimyrmex*, *Acromyrmex,* and *Atta,* known as leaf-cutter ants, most of which cultivate the single, highly derived fungal species *Leucoagaricus gongylophorus*. In summary, specific species of *Escovopsis* associate with specific ant-cultivar hosts.

Comparative genomic studies of fungi have provided valuable insights into macroevolutionary patterns in diverse fungal symbionts (38), such as those of plants (39–43) and lichens (44). Comparative genomic and metabolomic studies could likewise provide valuable insights into the coevolutionary dynamics of *Escovopsis*. A study of the *E. weberi* draft genome sequence revealed a genome that contained primary metabolism genes and exhibited reductions in several carbohydrate-active enzyme (CAZyme) gene families and genes associated with sexual reproduction (45). These gene family reductions are consistent with the obligate parasitic life history of *E. weberi*. In addition, the *E. weberi* genome revealed unique suites of BGCs, some of which are expressed during parasitic growth toward the cultivar (45). Further, *E. weberi* was shown to produce melinacidins capable of inhibiting *Pseudonocardia in vitro*, as well as shearinines capable of influencing ant behavior *in vivo* (20, 21). We hypothesize that the pattern of genome reduction exemplified in the *E. weberi* genome was inherited by all species of *Escovopsis* from their shared common ancestor, which originated more or less coincident with the origin of ant agriculture. Additionally, we hypothesize that *Escovopsis* lineages will exhibit genomic diversification and specialization across the different fungus-growing ant lineages practicing different agricultural strategies. Overall, we expect that *Escovopsis* genomes will share a similar suite of core genes that facilitate parasitism of fungus-growing ant gardens and that they will also possess unique subsets of non-core accessory genes that evolved subsequent to the evolutionary divergences of the fungus-growing ant lineages.

Here, we investigate the macroevolutionary history of *Escovopsis* mycoparasites, within the iconic fungus-growing ant symbiosis. Specifically, we sequenced and assembled 18 high-quality draft genomes from *Escovopsis,* which, combined with the seven publicly available genomes, represent strains isolated across the phylogenetic diversity of fungus-growing ants. We use maximum likelihood (ML) and coalescent-based phylogenomic inference to construct a species tree of *Escovopsis* to ascertain its phylogenetic position within the Hypocreales. We date the origin of species lineages across *Escovopsis* spp. using the program MCMCTREE (46) as implemented in the Phylogenetic Analysis for Maximum LIkelihood (PAML) package for phylogenetic dating analysis and estimate genome sizes across all species using k-mer estimation and propidium iodide flow cytometry. We then estimate the effect of genome reduction on lineage-specific functional gene diversity and determine how this relates to BGCs underlying the production of small molecules. Finally, we use high-resolution electrospray ionization-mass spectrometry (HRESI-LC-MS) to conduct untargeted metabolomic characterization of *Escovopsis* across representative strains to determine variation and diversification of small molecule production. Taken together, our results reconstruct the evolutionary histories of different species of *Escovopsis* and their diversification across associated fungus-growing ant lineages from genomes to metabolomes. We observe gene loss across the monophyletic group of *Escovopsis* species that has implications for niche specialization parallel to major evolutionary events following the origin of the fungus-growing ant symbiosis.

Results and Discussion

Macroevolutionary Dynamics of *Escovopsis*. To examine the macroevolutionary dynamics of the garden pathogen *Escovopsis*, we sequenced, assembled, and annotated 18 draft genomes of *Escovopsis* species. The 18 strains, together with seven previously sequenced and publicly available *Escovopsis* genomes, span the phylogenetic diversity of fungus-growing ants, including four representatives from lower-attine agriculture, six from *Apterostigma* coral agriculture, seven from higher-attine agriculture, and eight from leaf-cutter ant agriculture (Dataset S1). The 25 *Escovopsis* genomes have an average Benchmarking Universal Single-Copy Orthologs (BUSCO) completeness of 97.8% for the ascomycete lineage dataset, an average N50 of 1.04 Mbp, and an average assembly length of 29.7 Mbp. Using the alignment-independent k-mer method GenomeScope (47), we estimated the average genome size to be 24.7 Mbp, with an average of 6,665 protein-coding genes (Dataset S1).

Next, we reconstructed the phylogeny of the 25 *Escovopsis* genomes together with 77 publicly available genomes of outgroup fungal species from the order Hypocreales. In total, we aligned 1,706 single-copy orthologs from the BUSCO ascomycete lineage dataset, resulting in a concatenated alignment composed of 2,882,422 base pairs. Our by-gene partitioned ML analyses resulted in strong statistical bootstrap support for the phylogeny, with most nodes exceeding 95% bootstrap support (*SI Appendix*, Fig. S1*A*). Additionally, we constructed a coalescent-based species tree using gene trees from the same 1,706 single-copy orthologs with most nodes on the primary topology exceeding a quartet score of 0.9 (Dataset S2 and *SI Appendix*, Fig. S1*B*).

Our phylogenomic analyses show that the strains of *Escovopsis* used in our study all form a monophyletic group, which is supported by both ML and coalescent-based analyses (*SI Appendix*, Fig. S1). Further, we find that the evolutionary placement of the genus *Escovopsis* is within the family Hypocreaceae (Fig. 1), and that the sister clade is composed of *Cladobotryum protrusum* (GCA_004303015.1) and *Hypomyces rosellus* (GCA_011799845.1), two mycoparasites of agaricaceous fungi. *Escovopsis* forms a monophyletic group of five clades: i) one clade (the sister of all the rest) parasitizing lower agriculture, ii) two clades parasitizing coral-fungus agriculture, iii) one clade parasitizing non–leaf-cutting higher agriculture, and, in the most apical clade, iv) two sister clades, one parasitizing non-leaf-cutting higher agriculture and one parasitizing leaf-cutter agriculture (Fig. 2, colored boxes). One clade of *Escovopsis* parasitizing lower agriculture (quartet support 0.99) is composed of four strains: two from a currently undescribed *Escovopsis* species (ICBG2046 and ICBG2048), isolated from a *Cyphomyrmex muelleri* ant garden from Panama and a *Mycetophylax* cf. *faunulus* ant garden from Guyana, respectively, and two strains of *E. kreiselii* (ICBG2047 and ICBG2049), both isolated from *Myrmicocrypta ednaella* ant gardens from Panama. Of the two clades of *Escovopsis* coral-fungus agriculture parasites, the first clade (quartet support 0.99) consists of two strains of an undescribed *Escovopsis* species (ICBG712 and ICBG721) that were isolated from *Apterostigma urichii* ant host gardens from Brazil and the second (quartet support 0.93) consists of *E. multiformis* strains (ICBG1065 and ICBG1075) isolated from *Apterostigma bruchi* gardens from Brazil and *E. clavatus* strains (ICBG1054 and ICBG726) isolated from *Apterostigma urichii* gardens from Brazil. Of the two clades of higher-agriculture-parasitizing *Escovopsis*, the first (quartet support 0.84) consists of three *E. aspergilloides* strains (ICBG730, ICBG710, and NIGD00000000) and the second (quartet support 0.99) consists of three *E. moelleri* strains (ICBG751, ICBG1096, and ICBG733). All of the higher-agriculture-associated *Escovopsis* strains were isolated from *Paratrachymyrmex diversus* gardens from Brazil except for NIGD00000000, which was isolated from a *Paratrachymyrmex cornetzi* fungus garden from Panama. The clade of *Escovopsis weberi* leaf-cutter parasites (quartet support 0.82) consists of nine strains (ICBG731, ICBG736, ICBG742, LGSR00000000, NIGB 00000000, NIGC00000000, NQYQ00000000, NQYR0000 0000, NQYS00000000). Leaf-cutter-associated strains were isolated from *Acromyrmex hystrix* from Brazil (ICBG742), *Acromyrmex echinatior* from Panama (NIGC00000000, NQYQ00000000, NQYR00000000), *Atta cephalotes* from Panama (LGSR00000000), *Atta colombica* from Panama (NQYS00000000), and *P. diversus* from Brazil (ICBG736), with two strains that do not have host identifications (ICBG731 and NIGB00000000) from Brazil and Panama.

The temporal sequence of evolutionary divergences in *Escovopsis* revealed by our ML analyses mirrors the evolutionary history of fungus-growing ants. The most early-diverging *Escovopsis* lineage (Fig. 2, *Left*) contains species associated with lower agriculture, whereas the next two diverging lineages contain parasites of coral-fungus agriculture. *Escovopsis* spp. associated with higher and leaf-cutter agriculture are the most derived in the genus (Fig. 2). Our coalescent-based gene tree supports most of these major evolutionary transitions; however, it is fundamentally conflicted about the position of the coral-fungus-parasitizing clade containing ICBG712 and ICBG721. The primary topology, supported by 719 genes, reconstructs ICBG712 and ICBG721 as the most early-diverging *Escovopsis* clade, whereas the secondary topology, supported by 706 genes, is consistent with our ML tree (i.e., with the lower-agriculture clade as the sister to the remainder of the genus) (*SI Appendix*, Fig. S1 *A* and *B* and Dataset S2). As an independent topology test, we reconstructed the phylogeny of the mitochondrial genome. All major nodes of the mitochondrial genome tree are supported by greater than 98% bootstrap support and support the same major topology as that of the primary coalescent-based gene tree by placing the clade consisting of ICBG712 and ICBG721 (coral agriculture) as the most early-diverging *Escovopsis* clade (*SI Appendix*, Fig. S1 *B* and *C*). The clade consisting of ICBG712 and ICBG721 positioned as the most early-diverging *Escovopsis* clade is also supported by the elongation factor 1-alpha gene tree (*SI Appendix*, Fig. S1*D*) and the mitochondrial genome tree (*SI Appendix*, Fig. S1*C*).

The divergence dates resulting from our genome-scale *Escovopsis* and fungus-growing ant divergence-dating analyses reveal that the dates of origin of the majority of *Escovopsis* clades coincide with the dates of origin of their fungus-growing ant hosts. For example, *Escovopsis* likely originated from 72.69 (stem; 47.74 to 97.99 Mya highest posterior density (HPD)) to 61.82 (crown; 40.45 to 83.81 Mya HPD) Mya, whereas fungus-growing ants likely originated from 72.89 (stem; 63.68 to 81.92 Mya HPD) to 68.04 (crown; 59.35 to 76.93 Mya HPD) Mya. The ancestral fungus-growing ant practiced lower agriculture (13, 28) and the earliest-diverging lineage of *Escovopsis* parasitizes lower agriculture, suggesting that most or all *Escovopsis* species are descended from a parasite of lower agriculture. *Escovopsis* parasites of higher ant agriculture arose from 40.85 (stem; 26.06 to 55.82 Mya HPD) Mya to 30.75 (crown; 19.47 to 42.83 HPD) Mya, which closely corresponds to the origins of higher ant agriculture from 34.78 (stem; 29.08 to 40.57 Mya HPD) to 28.71 (crown; 24.13 to 33.37 Mya HPD) Mya. Most recently, *Escovopsis* parasites of leaf-cutting ants arose 17.83 (stem; 10.72 to 25.52 HPD) to 11.9 (crown; 6.85 to 17.2 HPD) Mya, whereas leaf-cutting ants arose from 20.0 (stem; 16.49 to 23.61 Mya HPD) to 18.94 (crown; 15.47 to 22.39 Mya HPD) Mya. Correspondence between the dates of origin of coral-fungus-cultivating ants and their *Escovopsis* parasites is more complex. The clade of coral-fungus-cultivating *Apterostigma* ant species originated from 24.95 (stem; 20.08 to 29.97 Mya) to 18.67 (crown; 14.76 to 22.66 Mya HPD) Mya. The younger of the two clades of coral-fungus-parasitizing *Escovopsis*, containing *E. multiformis* and *E. clavatus*, originated from 40.85 (stem; 26.06 to 55.82 Mya HPD) to 25.96 (crown; 15.1 to 37.62 Mya HPD) Mya, coinciding with the HPD ant stem age. However, the older clade of coral-fungus-parasitizing *Escovopsis,* containing ICBG712 and ICBG721, originated from 54.13 (stem; 34.99 to 73.53 Mya HPD) to 1.01 (crown; 0.48 to 1.62 Mya HPD) Mya, suggesting that, paradoxically, it originated prior to the origin of coral-fungus agriculture. An alternative and arguably more likely hypothesis, however, is that the ancestor of the coral-fungus-associated clade (ICBG712 and ICBG721) (Fig. 2) originated as a parasite of lower ant agriculture but that, coincident with or following the origin of coral-fungus agriculture 20 to 30 Mya, its descendant lineage host-switched to become a parasite of coral-fungus agriculture. This hypothesis of host-switching from lower to coral-fungus agriculture is also consistent with the alternative topologies produced by the coalescent,

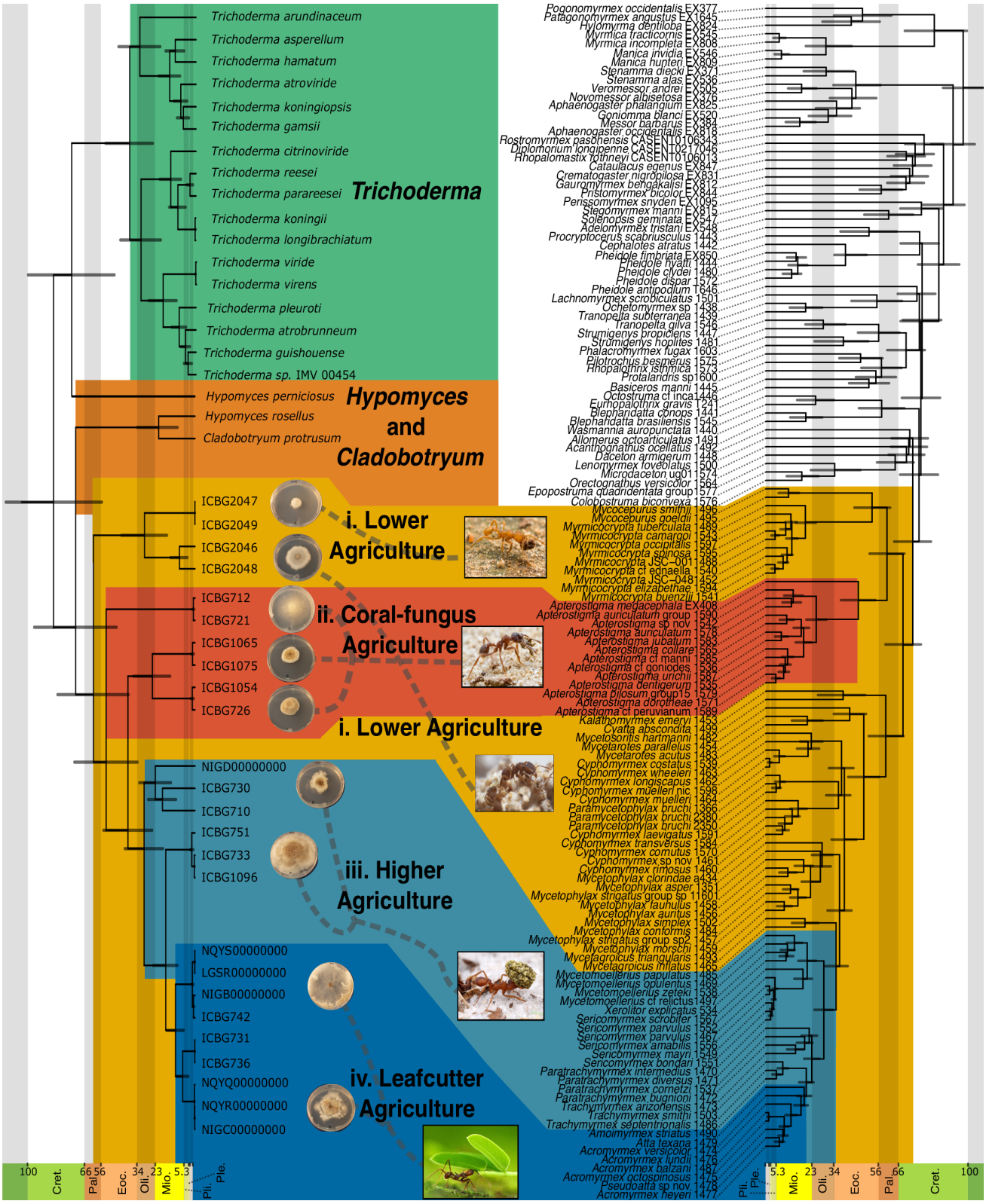

Fig. 2. *Escovopsis* originated simultaneously with the origin of fungus-growing ants and specialized parasitic lineages of *Escovopsis* correspond to specialized host lineages of fungus-growing ants. The phylogenetic relationships and the divergence dating analysis of *Escovopsis* genomes are consistent with the dating of the ant agricultural systems they parasitize. *Left*: Time tree of hypocrealean fungi, including *Escovopsis* (i. lower: yellow, ii. coral-fungus: red, and iii. higher, iv. leafcutter: blue boxes), its sister clade (orange box), and *Trichoderma* (green box). *Right*: Time tree of fungus-growing ants and non-fungus-growing ant outgroups. *Center Left*: Culture plates of *Escovopsis* strains after seven days of growth on potato dextrose agar. *Center Right*: Images of ants corresponding to selected ant agricultural systems. The dashed gray lines connecting *Escovopsis* fungal strains to ant images indicate parasite–host associations. Vertical gray bars indicate bins across the time scale. Photographs of *Escovopsis* by Kirsten Gotting, photographs of *Trachymyrmex, Apterostigma, Myrmicocrypta*, and *Atta* ants photographs by Alex Wild, used by permission, and photograph of *Mycetophylax asper* ant by Don Parsons.

mitochondrial genome, and single-gene elongation factor 1-alpha analyses discussed above (*SI Appendix*, Fig. S1).

These results indicate that *Escovopsis* is a monophyletic group within the family Hypocreaceae. The species in the sister group of *Escovopsis*, represented in our phylogeny by *Hypomyces rosellus* and *Cladobotryum protrusum*, are, like *Escovopsis*, mycoparasites of agaricaceous fungi, as is *Hypomyces perniciosus*, an early-diverging member of the sister clade of *H. rosellus, C. protrusum,* and *Escovopsis* (Figs. 1 and 2). The molecular dating analysis indicates that *Escovopsis* parasitism of agaricaceous fungi may have coincided with the inception of the fungus-growing ant-microbe symbiosis, and subsequently differentiated in concordance with diversification of the symbiosis after the K-Pg extinction event. The origin of the main clades of *Escovopsis* corresponds to the origins of the major ant agricultural systems. This temporal correspondence of evolutionary histories supports significant co-diversification and coevolution in the fungus-growing ant agricultural symbiosis. Future studies integrating more fossil evidence and more *Escovopsis* genomes will further resolve the evolutionary history of *Escovopsis* and its correlations with fungus-growing ant evolution. Additional sampling of *Hypomyces* and *Cladobotryum* genomes is necessary to better resolve the phylogenetic position of the *Escovopsis* clade consisting of ICBG712 and ICBG721, which is currently confounded by gene tree discordance.

Genome Reduction in *Escovopsis*. To test the hypothesis that a reduced genome is an ancestral feature of *Escovopsis* genomes, we estimated genome sizes using k-mer-based methods and assembly lengths, as well as examined the relative contributions of coding sequence (CDS), introns, and repeat content to genome sizes (Dataset S1). *Escovopsis* genome sizes, as estimated with k-mer-based methods, range from 19 to 28.8 Mbp, which is substantially smaller than the genome size estimates for *Cladobotryum protrusum* (39.6 Mbp), *Hypomyces perniciosus* (44.6 Mbp), and *Hypomyces rosellus* (39 Mbp). Further, k-mer estimates of *Trichoderma* spp. genome sizes range from 31.28 to 39.18 Mbp (*SI Appendix*, Fig. S2*A*). *Escovopsis* genomes exhibit substantial reductions in CDS DNA, ranging between 9.4 Mbp and 10.9 Mbp. In comparison, *Hypomyces perniciosus*, *Hypomyces rosellus*, and *Cladobotryum protrusum* genomes have 13.6 Mbp, 15.2 Mbp, and 15.2 Mbp, respectively, while *Trichoderma* spp. range from 13.5 to 17.5 Mbp (*SI Appendix*, Fig. S2*C*). Additionally, *Escovopsis* spp. have between 10,734 and 13,420 introns in CDS regions (*SI Appendix*, Fig. S2*E*), while *Hypomyces perniciosus, Hypomyces rosellus*, and *Cladobotryum protrusum* have 17,690, 16,633, and 18,581 introns in CDS regions, respectively. Comparatively, *Trichoderma* spp. have between 15,546 and 23,690 introns in CDS regions. While *Escovopsis* genomic features are smaller for both metrics, mean CDS lengths of *Escovopsis* genes are on average 1,561 base pairs, which is only 3% smaller than *Hypomyces perniciosus, Hypomyces rosellus*, and *Cladobotryum protrusum*, and is larger than *Trichoderma* spp. whose genes are on average 1,427 base pairs (*SI Appendix*, Fig. S2*D*). Repeat content as estimated with k-mer frequency spectra across genome assemblies for *Escovopsis* genomes is highly variable (*SI Appendix*, Fig. S2*F*) ranging from 0.6 to 6.5 Mbp, like the variability among *Hypomyces rosellus* (0.98 Mbp), *Cladobotryum protrusum* (1.5 Mbp), and *Hypomyces perniciosus* (11.5 Mbp). The distributions of these estimates are broader than the distributions of repeat content in *Trichoderma* genomes (1.3 Mbp), which is also smaller than the average for *Escovopsis* (1.5 Mbp). Additionally, we estimated repeat content with k-mer frequency spectra across Illumina reads. Estimates for *Escovopsis* ranged from 0.38 to 1.68 Mbp, which is smaller than the range estimated across genomic reads from *Trichoderma* which ranged from 0.9 to 3.8 Mbp (*SI Appendix*, Fig. S2*G*). The differences in repeat content as estimated from assemblies and sequenced reads could reflect technical variation in genome assembly, with the estimates from the Illumina reads likely reflecting values closer to biological reality.

To further establish genome reduction in *Escovopsis* genomes, we conducted propidium iodide flow cytometry. We estimate genome sizes of 31.71 Mbp for ICBG2048, a strain of *Escovopsis* from lower agriculture, 21.94 Mbp for ICBG1065 and 26.52 for ICBG1075, strains of *Escovopsis multiformis*, 19.7 Mbp for ICBG1054, a strain of *E. clavatus,* 30.34 Mbp for ICBG730, a strain of *Escovopsis* from *P. diversus* higher agriculture, and 24.43 Mbp for ICBG736, a strain of *Escovopsis weberi* isolated from leaf-cutter agriculture. This range of genome sizes from 19.7 to 31.71 Mbp is like our k-mer-based predictions, which ranged from 19 to 28.8 Mbp. In particular, the genome size for ICBG1065 predicted with both methods varied by only 2.95 at 18.99 Mbp with k-mers and 21.94 Mbp with flow cytometry.

To understand how *Escovopsis* genomes differ across lineages, we examined shared orthologues and gene loss between *Escovopsis* and three outgroup genera, *Cladobotryum*, *Hypomyces*, and *Trichoderma*, using clade-specific pangenomes. Genes were included if they were present in all *Hypomyces* and *Cladobotryum* genomes and in at least 95% of all genomes for *Escovopsis* and *Trichoderma* genomes. In total, 8,783 orthologous genes were considered in this analysis (Fig. 3*A*). 1,943 genes are completely lost in *Escovopsis* (*SI Appendix*, Fig. S3*A*). We tested these genes for enrichment of gene ontology (GO) terms to determine if there were specific molecular functions, cellular components, or biological processes that predominated across these lost genes. These genes were enriched for GO terms such as "regulation of transcription, DNA-templated," "transmembrane transport," "regulation of metabolic process," "RNA biosynthetic process," "transcription, DNA-templated," and "oxidation-reduction process" (*SI Appendix*, Fig. S3*B*). Within *Escovopsis* genomes, 6,849 orthologues are shared, of which 6,573 are also shared with at least one of the outgroup taxa (*SI Appendix*, Fig. S3*A*). There are 729 orthologues that are shared between *Trichoderma, Hypomyces,* and *Cladobotryum* genomes that are not present in *Escovopsis*. Additionally, there are 605 orthologues that are only present in the genus *Trichoderma* and 600 only present in *Hypomyces* and *Cladobotryum*. Finally, *Escovopsis* shares 486 genes with *Hypomyces* and *Cladobotryum* and 133 genes with *Trichoderma* (*SI Appendix*, Fig. S3*A*).

While sequencing-dependent genome size estimates tend to underestimate genome sizes, the genome size differences are uneven across the genus *Escovopsis*, with some species exhibiting larger, lineage-specific reductions. This result may be partially due to changes in gene size. When considering all annotated genes, *Escovopsis* genes are on average longer and have fewer introns compared to other fungi in the Hypocreaceae. Longer genes may be a result of loss of intergenic regions and gene fusion. In addition to structural changes in gene sizes, certain genes were lost in *Escovopsis*, in which GO enrichment tests indicated were metabolism and membrane transport genes, as well as genes related to transcription. The reduction in genes related to transmembrane transport could be a result of specialization of this parasite, resulting in the loss of genes related to generalized interspecies signaling. Additionally, many transcription factors are multi-copy genes, where a loss of some copies may not sacrifice the function (*SI Appendix*, Fig. S3*B*). Overall, the patterns of gene loss suggest non-uniform reductive evolution of *Escovopsis* lineages (*SI Appendix*, Fig. S3*A*).

While many of the *Escovopsis* genomes share high levels of sequence similarity, as measured by average nucleotide identity

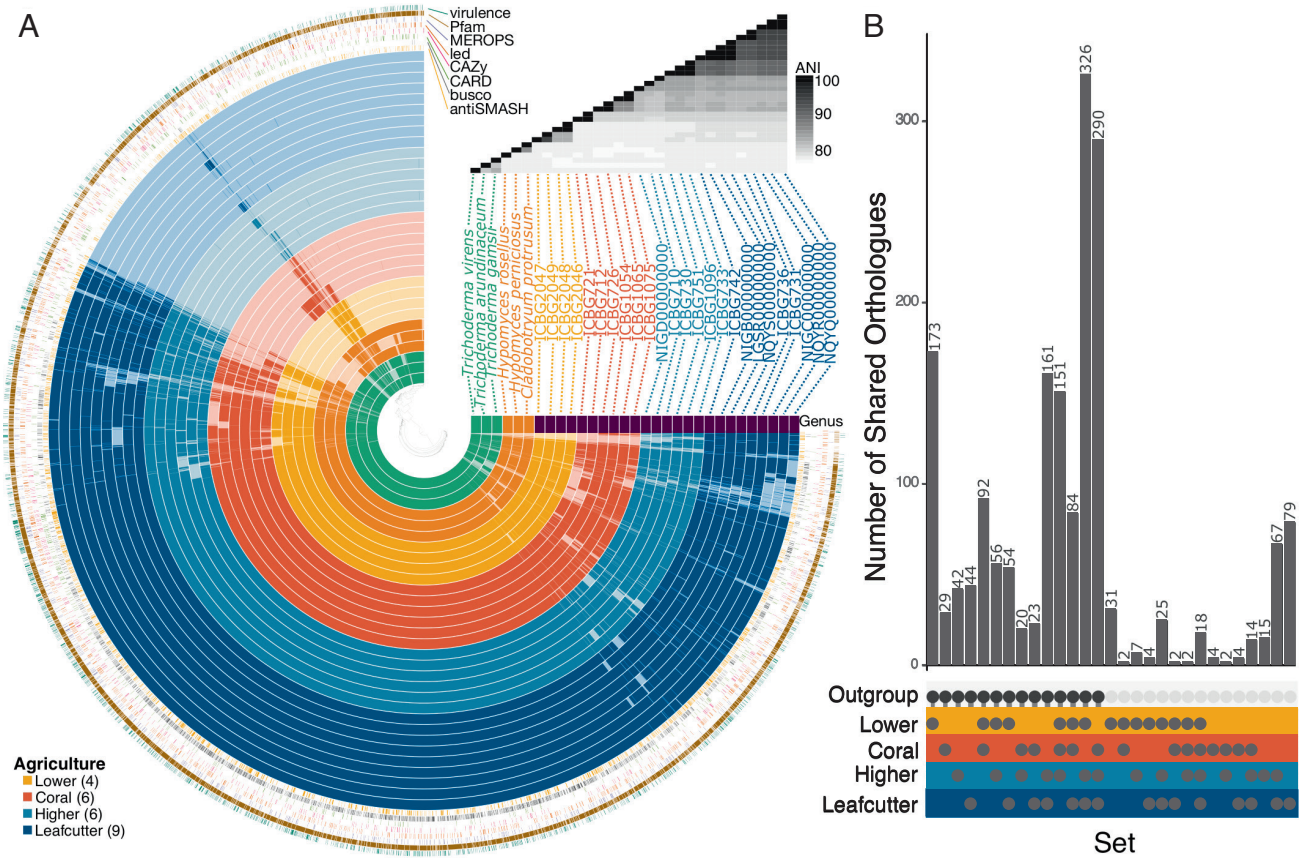

Fig. 3. (*A*) Anvi'o interactive visualization of presence-absence matrix of 8,783 orthologues among depicted genomes, with the rows along the radial axis indicating each genome and the columns indicating the presence (saturated) or absence (unsaturated) of each ortholog. Functional and informative annotations of orthologs are placed in the outer rings, in order from the outermost ring, number of genomes with each orthologue, presence in Fig. 3*B*, orthologs with a virulence annotation, orthologs with a PFAM annotation, orthologs with a MEROPS annotation, orthologs with a LED annotation, orthologs with a CAZyme annotation, orthologs with a CARD annotation, orthologs in an antiSMASH-annotated BGC. The legend describes the colors associated with each genome, indicating the genus each genome originates from, and for *Escovopsis* spp. the fungus-growing ant agriculture origin. The upper right-hand corner indicates pairwise average nucleotide identity (ANI) between genomes for this set of 8,783 orthologues. (*B*) UpSet (48,49) plot of shared orthologues content between sets that have greater than or equal to 40 orthologues. Shared orthologues among *Escovopsis* genomes show distinct evolutionary trajectories. These genes are not enriched for any specific GO term, but are annotated to various functional categories as shown in *B*.

(Fig. 3 *A*, *Upper Right* corner), the highest sequence similarity is found among species of *Escovopsis* parasitizing the same ant agricultural system. 276 genes are unique to *Escovopsis* and have different patterns of distribution corresponding with host association, i.e., with the parasitized ant agricultural systems. The distribution of these 276 orthologues across ant agriculture shows the distinct trajectories that lineages of *Escovopsis* have taken during their evolution (Fig. 3*B*). We observe that *Escovopsis* from lower, coral, higher, and leaf-cutter fungal agriculture have 31, 4, 15, and 79 lineage-specific orthologous genes, respectively. None of these groups of genes were enriched for any GO terms but do have functional domain annotations, as noted in the outer rings of Fig. 3*A*. Taken together, these data suggest that the selective pressures unique to different ant agricultural systems have played a role in shaping the genomic diversity seen in extant lineages of *Escovopsis*.

To determine the influence of reduced genome size on protein-coding gene families, we analyzed orthologous genes among and within species of *Escovopsis* by comparing the mean number of gene copies in each orthologous group in *Escovopsis* against the number in other fungal outgroups in the Hypocreaceae. We examined genes characterized as fungal virulence factors from the database of fungal virulence factors: DFVF (50), lipase-encoding genes from the Lipase Engineering Database: LED (51), peptidases from the MEROPS peptidase database (52), CAZymes from dbCAN2 for automated CAZyme annotation (53), antimicrobial resistance genes from the Comprehensive Antibiotic Resistance Database: CARD (54), and BGCs using antiSMASH (55) with transmembrane domains or extracellular locations (see *Methods* for more detail on these designations). This analysis shows that *Escovopsis* spp. genomes are reduced in gene copies of 1,479 orthologues, whereas, in contrast, gains of gene copies occur in only 56 orthologues. CAZymes and resistance genes have the highest loss-to-gain ratios, 6.60 and 6.63, respectively, followed by peptidase genes (2.94), lipase genes (1.99), and virulence genes (1.72) (Fig. 4*A*). *Escovopsis* spp. also have a reduction in BGCs, with a loss-to-gain ratio of 4.62 and with eight clusters gained and 37 lost.

To examine how this observed reduction of copy number impacts the diversity of gene functions present in *Escovopsis*, we calculated entropy for the same protein-coding gene families. We hypothesize that a uniform reduction in copy number would uniformly reduce the amount of diversity, as measured by entropy, observed in different gene classes. We expect that the overall gene number will be reduced in every category, while the evenness of different functions may be increased or decreased based on whether there is selection on those functions. As a control, we used annotated BUSCOs, which have similar diversity levels across all genomes (*SI Appendix*, Fig. S2*F*). Lipase genes, peptidase genes, CAZyme genes, and BGCs supported our hypothesis of uniform diversity reduction, but resistance and virulence genes increased entropy relative to the outgroup taxa (Fig. 4*B* and *SI Appendix*, Fig. S4). We looked at whether the

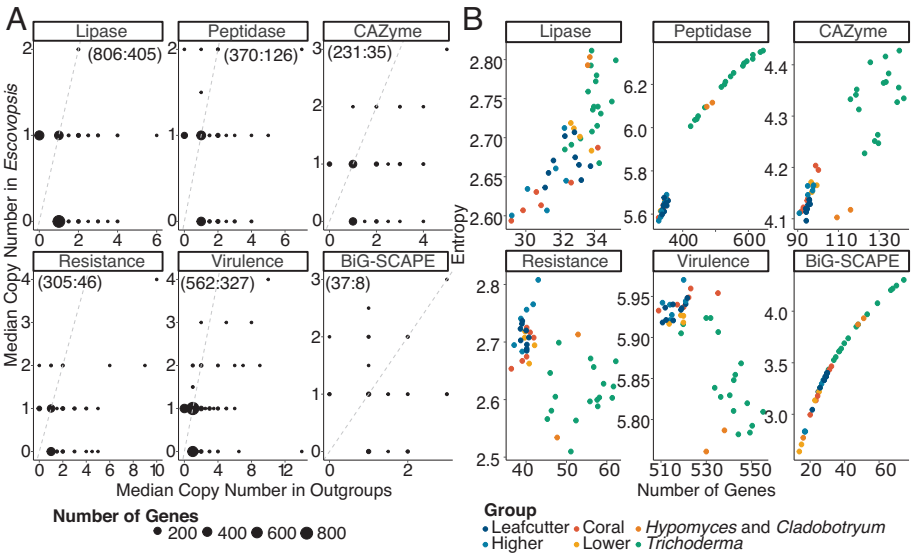

Fig. 4. *Escovopsis* spp. exhibit reduced copy numbers in all functions, while maintaining diversity in resistance and virulence gene functions. (*A*) For lipase genes, peptidase genes, CAZyme genes, resistance genes, virulence genes, and BGCs identified as similar by BiG-SCAPE, the mean copy number of each gene in *Escovopsis* (y-axis) is compared to the mean copy number in the outgroups (x-axis). The dotted gray line indicates a 1:1 copy ratio, with the size of the circle indicating the number of genes with a given ratio. The numbers in inset parenthesis indicate the total number of gene copies per functional group that are lost and gained. (*B*) The same gene categories as in (*A*) are analyzed for functional diversity of genes as measured by entropy on the y-axis versus the number of genes for each category in each genome on the x-axis. Colors indicate the grouping of the genome either as a part of ant-agriculture or genus-level groupings.

diversity increases in select gene categories and genomes were due to increased genetic evenness or richness and determined that virulence and resistance genes have increased evenness relative to the outgroup taxa (*SI Appendix*, Fig. S4 *A* and *E*). CAZymes had increased evenness, but the overall decrease in gene richness resulted in lower entropy values (*SI Appendix*, Fig. S4 *D* and *G*). Lipases and peptidase decreased both in richness and evenness of genes (*SI Appendix*, Fig. S4 *B* and *C*). BGCs were the only group of loci examined that exhibited decreased richness but not evenness of annotated clusters (*SI Appendix*, Fig. S4*G*).

Our results regarding copy number reduction and diversity support the hypothesis that *Escovopsis* genomes likely initially evolved a reduced suite of genes, which subsequently diversified across ant agriculture as different *Escovopsis* lineages specialized on the differing ecologies of the ant-cultivar agricultural strategies. This genomic diversification is exhibited through signaling molecules encoded by BGCs, as well as CAZymes, lipases, and peptidases for breaking down host material, resistance genes for confronting antibiotic pressure from ant-fungus-garden bacterial symbionts, and virulence genes for detecting and interacting with hosts. The case for reductive evolution followed by diversification is further supported by differing patterns of functional diversity for categories of protein-coding genes. The increase in entropy in protein-coding genes for resistance and virulence gene categories indicates that, although these categories exhibit gene copy reduction categories in *Escovopsis* genomes, functional genetic diversity has been maintained. Range restriction of *Escovopsis* to the fungus-growing ant-microbe symbiosis likely relaxed constraint on genes required for a generalist lifestyle, resulting in purifying selection and genome-wide reduction. While this reduction occurred in all classes of genes, certain virulence and resistance genes that likely confer pathogenic advantages to *Escovopsis* maintained functional diversity.

Lineage-Specific Differences among *Escovopsis* spp. Are Characterized by Genome Reduction and Gene Content Diversification. To identify evolutionary changes associated with the origin, life history, and diversification differentiating the garden pathogen *Escovopsis* from closely related Hypocreaceae, we examined the genomic-predicted putative secretome composition as characterized by fungal virulence factors, lipase-encoding genes, peptidases, carbohydrate-active enzymes, and antimicrobial resistance genes with either transmembrane domains or extracellular locations (see *Methods* for more detail on these designations), as well as BGCs. We then chemically characterized the metabolome for a subset of strains using HRESI-LC-MS. This suite of secreted proteins, secondary metabolites, and molecules likely mediates host interactions in symbiosis. Fig. 5*A* shows broad patterns of representation of secretome annotations across the genomes, with a broad swath shared, lost, and specific to *Escovopsis*. Additionally, we found a diversity of chemical potential across all *Escovopsis* lineages, and we found that molecules stratify *Escovopsis* lineages similarly to BGCs when analyzed with PCA (Fig. 5 *B* and *C*). With this in mind, we consider these groups of annotations and molecules with respect to genome reduction, genome retention, and genome diversification in comparison with outgroup taxa from *Trichoderma, Hypomyces,* and *Cladobotryum*. With respect to genome reduction, we identified 72 secretome annotations across 90% of *Escovopsis* species, which is smaller than the predicted secretomes in *Trichoderma* (129 across 90% of *Trichoderma* species), *Hypomyces* (169 across *Hypomyces* genomes), and *Cladobotryum protrusum* (272). Eighty-one annotations are lost across all *Escovopsis* genomes but are present in genomes of *Trichoderma, Hypomyces,* and *Cladobotryum*. Of note are 16 CAZymes annotated as GH1, GH27, GH13_1, GH5_15, GH5_31, AA12, GH37, GH105, GH24, GH39, GH43_14, GH32, GH43_34, AA14, PL20, and PL8_4. With respect to our metabolomics analysis, five molecular families are present in *Cladobotryum protrusum*, *Hypomyces perniciosus,* and *Hypomyces rosellus* that are not present in *Escovopsis* (Dataset S4).

Next, we examined gene annotations and molecules that are shared and retained in *Escovopsis* from the ancestral state. Among the secretome, 37 annotations are retained across all genomes surveyed (Fig. 4*A*). Seventeen of these are CAZymes, 16 are lipases, 11 are virulence genes, and one is a resistance gene (Dataset S4). Several are predicted to be involved in chitin degradation (ids:

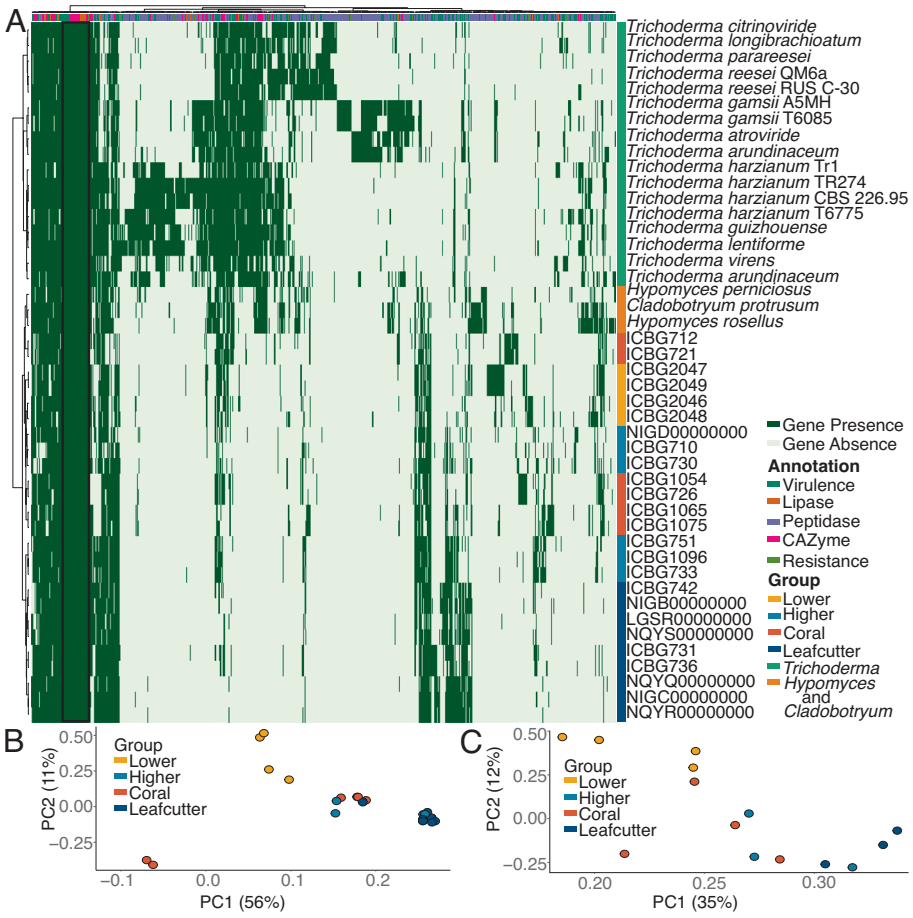

Fig. 5. Garden pathogen genomes have specialized genes related to host-signaling and specialized metabolites. (*A*) Heatmap of the secretome genes present in *Escovopsis* and outgroups. The columns represent individual genes and their presence (dark green) and absence (light green) across genomes in the rows. The black box on the *Left* side indicates genes present in all genomes analyzed. (*B*) PCA of BGC presence/absence as defined by BiG-SCAPE clustering per strain across ant agriculture. (*C*) PCA of the presence/absence of molecular networking subclusters as identified with GNPS.

GH16, GH18, Q8J1Y3_BEABA, O59928_HYPVI, AA11, AA7). Additionally, two (ids: abH36, CE5) are predicted to degrade cutin. All *Escovopsis* genomes share a putative type one polyketide synthase (PKS) BGC. Additionally, four BGCs are present in the majority of *Escovopsis* genomes and are putatively characterized as "Nonribosomal Peptide Synthetase" (NRPS), "Terpene," or "Other Ketide Synthase" (Other KS) (*SI Appendix*, Fig. S5*A*). Our metabolomic analyses identified 79 molecular families present across all *Escovopsis* metabolomes, none of which have natural product representatives in the Global Natural Product Social Molecular Networking Database (GNPS) (56, 57). Additionally, known Epipolythiodiketopiperazines are present in all *Escovopsis* strains except for ICBG1075, as well as in *Cladobotryum protrusum*, *Hypomyces perniciosus,* and *Hypomyces rosellus*.

Lastly, we considered genomic diversification among compared genera. We observe 148 secretome annotations that are lineage-specific to *Escovopsis*, defined as less than 10% presence in *Trichoderma* genomes, and a greater than 30% presence in at least one *Escovopsis* agricultural lineage (Dataset S4). Four of these proteins are predicted to be involved in resistance to tetracycline antibiotics and are present in all but coral agriculture-associated *Escovopsis*. Additionally, there are three CAZymes, two lipases, 82 peptidases, and 56 virulence factors that fall into this category. Eight of these proteins are predicted to be involved in degrading plant tissue (ids: Q9C1R1_FUSOX, Q00350_COCCA, PLYB_COLGL, MER0000339, MER0019139, MER0011060, CUTI2_FUSSO). Additionally, several are annotated as subtilisin-like serine protease, which play a role in fungal virulence (ids: SUB4A_COCP7, SUB7A_COCP7, SUB6_ARTOC, MER1073654). There are 82 BGCs across ant agriculture, with only three matching clusters present in the MIBiG (58) repository. The known clusters are identified as follows. BGC0001583 is an emodin BGC from *Escovopsis weberi* and is present in all higher and leaf-cutter agriculture *Escovopsis*. BGC0001777 is a shearinine D BGC and is present in all but two strains (ICBG751 and NIGD00000000) of higher and leaf-cutter agriculture *Escovopsis* as well as in coral agriculture (ICBG1065). Our metabolomics results additionally confirmed the presence of shearinines in all higher and leaf-cutter strains surveyed, as well as in ICBG1054 and ICBG726 from coral agriculture *Escovopsis*. Lastly, among identified BGCs, BGC0001585: melinacidin IV BGC is present in all but two strains of leaf-cutter agriculture (NIGB0000000 and NQYS00000000) as well as in all strains of higher agriculture except for the clade formed by ICBG730, ICBG710, and NIGD00000000. Eleven other unidentified clusters have mixed lineage specificity across agricultural types (*SI Appendix*, Fig. S5*A*). Twenty-four BGCs are only found in *Escovopsis* strains ICBG712 and ICBG721. Fourteen BGCs are only found among lower-agriculture-associated *Escovopsis* strains. Six BGCs are only found in *Escovopsis* strains ICBG726 and ICBG1054, and three BGCs are only found in *Escovopsis* strains ICBG1065 and ICBG1075. The remaining clusters from the original 82 have mixed patterns of lineage specificity (*SI Appendix*, Fig. S5*A*). Our metabolomics results identified eight molecular families specific to leaf-cutter agriculture *Escovopsis*, 15

molecular families specific to higher-agriculture *Escovopsis,* 62 molecular families specific to coral-agriculture *Escovopsis*, and 49 molecular families specific to lower-agriculture *Escovopsis*. Fifteen molecular families are shared between leaf-cutter, higher, and coral agriculture *Escovopsis* (including shearinines), 12 molecular families are shared between higher, coral, and lower-agriculture *Escovopsis*, 11 molecular families are shared across higher and coral agriculture *Escovopsis*, and 25 molecular families are shared between coral and lower *Escovopsis*.

The secretome, BGCs, and metabolomics highlight the ancient evolutionary history of these organisms, originating from a free-living common ancestor, and the genetic basis of a range of life history strategies. The reduced genomes of these specialized fungus-garden mycoparasites likely highlight the need to preserve chitin-degrading genes in the core secretome, whereas the functional capacity for degradation of plant tissue has been reduced across all lineages of *Escovopsis* with the loss of GH1, GH13_1, GH5_15, GH105, GH39, GH5_31, GH43_14, GH32, AA14, and GH43_34, which play roles in the degradation of xylan, cellulose, amylose, and fructose. One exception to this trend is proteins involved in cutin degradation, which could indicate that *Escovopsis* encounters these polymers, a key component of plant cuticles, during their life cycle. While most *Escovopsis* species had similar evolutionary patterns in their genomes and metabolomes, ICBG712 and ICBG721, isolated from the *Apterostigma* coral-fungus symbiosis, had differences in all functional suites of genes as well as metabolites, which is notable given that these *Escovopsis* are associated with gardens containing cultivars distantly related to those of the other agricultural systems. These results indicate that these genes likely persist in *Escovopsis* lineages as part of a co-evolutionary dynamic involving fungus-growing ant-microbe behavioral and antimicrobial defenses. Additionally, the trend toward greater across-lineage functional diversity in virulence and resistance genes is underscored by the presence of tetracycline resistance genes in all but one clade of *Escovopsis* and the presence of subtilisin-like serine proteases that have unorganized specificity across lineages. Lastly, the evolutionary origin of shearinines and melinacidin IV, two BGCs that are not found in *Hypomyces*, *Cladobotryum,* or *Trichoderma* or in lower attine-associated *Escovopsis,* represent genomic diversification known to have negative consequences for both leaf-cutter ants and their associated *Pseudonocardia* bacteria (20). Taken together, these data suggest that *Escovopsis* spp. evolved a diverse and streamlined molecular toolkit to adapt to the unique selective pressures applied by the fungus-growing ant niche. Additionally, these results demonstrate that *Escovopsis* has diversified along with its ant-cultivar hosts. Future research will unveil how these genes are causative for various traits that facilitate the parasitism of *Escovopsis*.

Conclusions

Our study reveals the distinct trajectories that *Escovopsis* spp. have undergone in their evolutionary histories, resulting in small genomes that encode diverse metabolites. These differences across species correspond to differences in fungus-growing ant agriculture. Our co-phylogenetic results suggest a significant "coevolutionary arms race" dynamic between *Escovopsis* and the ant-fungal cultivars, in which fungal cultivars continually diversify (resulting in the major ant-cultivar agricultural systems), perhaps to avoid parasitism, and in which *Escovopsis* continually adapts and evolves to parasitize these new systems. Additionally, the observed gene loss across *Escovopsis* species serves as further evidence that this genus likely originated as mycoparasites of agaricaceous cultivar ancestors that then diversified alongside the evolutionary history of ant-cultivar agriculture. This genomic diversification likely resulted in directional selection on niche-maintenance genes and in relaxed selection on genes ancillary to the ant-cultivar symbiotic ecology. Future research will aim to disentangle the relationship between species of *Escovopsis* and transitions to different agricultural strategies employed by the ants as well as different cultivar types.

Materials and Methods

Sample Collection, Isolation, DNA Extraction, and Genome Assembly. We collected samples from Brazil in 2017, and from Panama in 1999 (Dataset S1). Research and collection of biological samples in Brazil were authorized by SISBIO (46555-6) and CNPq (010936/2014-9). Strains were isolated and maintained on potato dextrose agar as in Currie et al. (15). For genomic DNA preparation, strains were grown up on OXOID malt extract agar (CM59) 8.33 g, agar 7.5 g, 500 mL distilled water with a cellophane overlay. Tissue was scraped off the plates with a scalpel and stored at −80 °C. DNA was prepared by grinding fungal tissue in liquid nitrogen and isolating with a CTAB-based protocol. Genomic DNA was sequenced at the University of Wisconsin-Madison Biotechnology Center in July of 2017 and January of 2019 with 2X300 base pair Illumina MiSeq. Raw Illumina genomic reads were filtered for bacterial and mitochondrial contamination by aligning to contaminant genomes using bwa mem (59) and discarding aligned reads using pysam (https://github.com/pysam-developers/pysam) (60). Reads were corrected using Musket (61) with default settings. Paired-end reads were then merged using FLASh (62) with default settings. Merged and unmerged reads from FLASh were used as input for genome assembly with SPAdes v3.11.1 (63) with default settings. Genomes that were particularly fragmented were scaffolded using RagOO (64) with a closely related (defined as greater than or equal to 95% average nucleotide identity) genome as a reference. Mitochondrial genomes were assembled using NOVOplasty (65), or downloaded from the NCBI from the following accessions: *Hypomyces aurantius, Trichoderma atroviride* (MN125601.1), *Trichoderma gamsii* (KU687109.1), *Trichoderma asperellum* (NC_037075.1), *Trichoderma hamatum* (NC_036144.1), *Trichoderma lixii* (MT495248.1), and *Trichoderma reesei* (AF447590.1). Species of *Escovopsis* were identified using a phylogenetic species concept by grouping species based on similarity to sequences of elongation factor 1-alpha of published species (*SI Appendix*, Fig. S1*D*).

Genome Size Estimation. Reads used for genome assembly were quality and adapter trimmed using fastp (66) using default parameters for single end reads, and the "--detect_adapter_for_pe" flag for paired end reads. Remaining reads were k-mer-counted using jellyfish (67) and a k-mer size of 31. These counts were made into a histogram with KAT "kat hist" (68). Genome sizes were estimated using genomescope2 (69), using the mean read length as reported by fastp, the k-mer size of 31, and the k-mer frequency spectra as produced by jellyfish. Genome sizes were only reported for genomes that had sufficient reads to meet 25X coverage of the genome. Furthermore, the k-mer frequency spectra for any genome with at least 25X coverage was manually inspected for an approximate Poisson distribution shape. If the spectra did not fit a Poisson distribution, the estimated genome size from genomescope2 was not reported. The computer program "agat_sp_statistics.pl" to tabulate genomic features is shown in *SI Appendix*, Fig. S2 (70). KAT "kat sect" with options -E and -F was used to estimate repetitive content in genome assemblies. As a second measure of nuclear genome size estimation, we performed propidium iodide flow cytometry after Bennett et al. (71). Chopping tissue with a fresh razor blade in Galbraith buffer supplemented with agarose was used to release nuclei from fungi. *Drosophila melanogaster* (1C = 175 Mbp) was used as an external standard to anchor the estimates. After filtration through 40 μm nylon, the DNA in the nuclei of sample and standard were stained with 1 mg/mL concentration of Propidium Iodide for 2 h in the cold and dark, with red PI fluorescence scored using a CytoFLEX (BeckmanCoulter) flow cytometer. The 1C amount of DNA in each sample was estimated as the ratio (mean 2C red PI fluorescence of the sample/mean red PI fluorescence of the standard) times 175 Mbp.

Gene Prediction and Protein Annotation. The Trinity-v2.6.5 de novo transcriptome (72) assembler with default parameters was used to assemble an initial transcriptome from RNA-seq data for *Escovopsis* (Bioproject: PRJNA253870). Augustus-v3.3 (73) was used to train and optimize initial gene models for

Escovopsis based on the first genome annotation (45) and the assembled Trinity transcriptome. Repeat families were predicted for each genome using RepeatModeler-v1.0.11 (74). Genes were predicted for each genome using MAKER-v2.31.9 (75), with transcript evidence from the assembled transcriptome, proteomes for publicly available *Trichoderma* assemblies, and the UniProt (76) database of curated proteins. After an initial run with MAKER, BUSCO- v4.1.1 (77–79) was used to assess the completeness and quality of the annotation. Then, hmm-assembler.pl from Semi-HMM-based Nucleic Acid Parser (SNAP) (80) was used to retrain on high confidence genome models used as input for a second run with MAKER. This process was repeated iteratively until the annotation reached greater than or equal to 90% completeness, or the quality of the annotation as measured by BUSCO degraded.

InterProScan (81) was used to annotate proteins with Pfam, TMHMM, SignalP, and GO terms. CAZymes (carbohydrate-active enzymes) were annotated to protein sequences using dbcan (53). Lipases, peptidases, and virulence factors were annotated by aligning protein sequences with DIAMOND to the LED (51), MEROPS (52), and DFVF (50), respectively, and taking the top hit for each protein according to bitscore. Resistance genes were annotated using the RGI software distributed by CARD (54). Subcellular locations of proteins were annotated using TargetP and WoLF PSORT (82). Orthologous protein families were annotated with OrthoFinder (83).

To annotate BGCs, antiSMASH (55) was run twice using the options "--taxon fungi --full-hmmer --smcogs --subclusterblast --knownclusterblast" for both runs, in addition to the following. The first run used just the genome sequence file as input, and the second run used the genome sequence file, and the maker-annotated gene models as an input with the "--gff3" option. BiG-SCAPE (84) with the option "--include_singletons --mix --mibig" was used to find clusters of similar BGCs among genomes and clusters from the MIBiG repository (58).

Secretome Annotation. To define the secretome of each genome, we annotated the presence of lipases, proteinases, carbohydrate-active enzymes, resistance genes, and fungal virulence genes among the protein-coding genes for each genome. These annotated proteins were considered part of the secretome if they contain either a signal peptide as annotated with SignalP via InterProScan, one transmembrane domain as annotated with TMHMM via InterproScan, or a subcellular location outside of the cell as annotated with TargetP and WoIF PSORT.

Phylogenetic Reconstruction. Representative RefSeq genomes of Hypocreales fungi were used to reconstruct phylogenetic relationships, along with *Saccharomyces cerevisiae*, *Candida albicans,* and *Saitoella complicata* as outgroups (Dataset S3). BUSCOv4.1.1 ran with the ascomycota_odb10 lineage database generated single-copy orthologues from each genome. The 1,706 shared single-copy orthologs were used as input for building the species tree. Sequences were aligned individually using MAFFT v7.475 (85,86) with default parameters, then trimmed using trimAL v1.2rev59 (87) with the "-automated1" flag. Output alignments from trimAL were used as input to IQ-TREE2 (88–90) with parameters "-bb 1000 -alrt 1000" to construct individual gene trees, using ModelFinder to select the best evolutionary model for each gene (89). These alignments and evolutionary models were used to reconstruct the species tree with concatenation-based methods, whereas the individual gene trees were used to reconstruct the species tree with coalescent-based methods. Methods for both approaches are as follows. The concatenation-based tree was constructed from the set of 1,706 BUSCO gene alignments from MAFFT and trimAL. The alignment was partitioned by gene, using the best evolutionary model as predicted with ModelFinder. We reconstructed the tree using IQ-TREE2 (88–91) with parameters "-bb 1000 -alrt 1000." The resulting tree was rooted to *S. cerevisiae* and *C. albicans* in FigTree (92). The multi-allele coalescent species tree was constructed from the 1,706 gene trees which were used as input for ASTRAL v5.7.7 (93, 94) with option "-t 32." Mitochondrial genome trees were reconstructed similarly to the individual gene trees using full genome alignments.

Ant phylogenomic analysis used ultraconserved elements (UCEs) markers and is based on a modified version of the alignments of Li et al. (95) and Hanisch et al. (96). The dataset consisted of 133 taxa (57 outgroup Myrmicinae taxa and 76 fungus-growing ant taxa), 558,222 nucleotide characters, representing 942 UCE loci. We employed the Sliding Window Site Characteristics based on Entropy (97) approach to first split the UCE loci into its three regions (a core and two flanking regions). The subsets identified by the SWSC-EN algorithm were then used as input to identify the best partitioning scheme using ModelFinder (89) as implemented in IQ-TREE multicore v2.0.6 (90). For the merging step, we used the -m MF+MERGE command, the fast relaxed -rcluster algorithm (98), and compared the top 10% of the resulting partition schemes using the corrected Akaike Information Criterion (AICc). The evaluated models were restricted to those implemented in RaxML by using the command -mset raxml. The best-fit partitioning scheme consisted of 853 subsets. ML analysis was performed in IQ-TREE multicore v2.0.6 with ultrafast bootstrap (88) and SH-like approximation likelihood ratio test (99) set at 1,000 replicates, with other settings set at default values.

Divergence Dating. We inferred divergence dates of *Escovopsis* and the fungus-growing ants using the Bayesian program MCMCTREE (100) implemented as part of the package PAML (46) which uses the approximate-likelihood approach of (101). For both analyses, we input the alignments and the ML topologies generated for each dataset (see above). For *Escovopsis,* three calibrations nodes were used for phylogenetic dating analyses that were retrieved from the TimeTree database (102) and previously published dates across the Ascomycota (103). The calibration nodes used were the origin of the Saccharomycotina (304 to 590 Mya), the divergence between *Cordyceps tenuipes* (GCA_003025305.1) and *Torrubiella hemipterigena* (GCA_000825705.1) 50.2 Mya, and the *Trichoderma* root 20 Mya. The calibrations were employed as follows: i) the root was modeled as the uniform distribution B(3.04, 5.09), i.e., of 304 to 509 Mya; ii) the most recent common ancestor (MRCA) of GCA_007896495_1 and GCA003012105_1 was modeled as the Cauchy distribution L(0.20, 0.1, 1.0, 0.025), and iii) the MRCA of GCA_003025305_1 and GCA_000825705_1 was modeled as the Cauchy distribution L(0.502, 0.1, 1.0, 0.025). For the fungus-growing ants, we employed information available from two fossils as well as from two published studies (secondary calibrations) to calibrate our analysis. The two fossils were: i) A *Pheidole* species from Florissant Formation (34 Ma) (104,105) used to calibrate the crown node of the genus *Pheidole* and modeled as the Cauchy distribution L(0.34, 0.05, 0.085, 1e-300) and ii) the species *Mycetomoellerius primaevus* from Dominican Amber (15 Ma) (106) used to calibrate the MRCA of *Mycetomoellerius* and *Acromyrmex* (i.e., the higher Attina), and modeled as the uniform distribution B(0.15, 0.35), i.e., of 15 to 35 Mya. The two secondary calibrations were: i) the root (or crown age of the subfamily Myrmicinae) [95% CI 110.1 to 87.1 Mya, median 98.6 Mya; (13,105)], modeled as the skew-normal distribution SN(0.986, 0.06, 0). ii) The crown age of the tribe Crematogastrini [95% CI 93.704 to 66.132 Mya, median 78.55 Mya; (107)] modeled as the skew-normal distribution SN(0.7855, 0.08, 0).

MCMCTREE analyses of both *Escovopsis* and fungus-growing ants used the independent-rates clock model and the GTR+G4 substitution model. For *Escovopsis*, we conducted six independent MCMCTREE runs, each consisting of 210 million generations, and each with the following settings: sampfreq = 3,000, nsample = 70,000, and burnin = 21,000,000. For the fungus-growing ants, we conducted four independent MCMCTREE runs, each consisting of 50 million generations, and each with the following settings: sampfreq = 1,000, nsample = 50,000, and burnin = 5,000,000. We assessed run convergence and stationarity by examining the resulting mcmc.txt files in Tracer v1.7.1 (108) using the criterion of ESS values higher than 500. Analyses were conducted on the Smithsonian High Performance Cluster (SI/HPC), Smithsonian Institution (https://doi.org/10.25572/SIHPC).

Genomic Comparisons. Lists of genes were tested for GO term enrichment using topGO. Significantly enriched, defined as having a Benjamini–Hochberg adjusted p value less than or equal to 0.01 GO terms were then tested for semantic similarity using REVIGO with the options "similarity: Medium (0.7), numbers associated with GO terms: *P*-values, database with GO term sizes: whole UniProt, semantic similarity measure: SimRel." We annotated orthologous genes with OrthoFinder (83) and used these to define the presence and copy number of genes across genomes and agriculture as visualized with UpSetR (48,49). Orthogroups were considered in our analysis if they were present in the pan-genome of each phylogenetic grouping, i.e., in 95% of genomes for groups with more than four representative genomes, and in all genomes for groups with four or fewer representative genomes. To characterize differences in functional genomic content, we calculated the Shannon entropy of genes for functional gene classes using the R package vegan (109). We made plots using ggplot2, cowplot, and pheatmap (110–113) and made extensive use of the tidyverse suite of R packages for data analysis (114). Lastly, Anvi'o interactive was used to visualize orthologous genes among genomes and the functional annotations of the orthologues (115).

Metabolomics. *Escovopsis* strains were grown on potato dextrose agar for 3 wk. The agar plates were cut into pieces and frozen at −80° C for 6 h. The cell material and agar were then extracted with ethyl acetate with excess anhydrous sodium sulfate to remove water overnight. The ethyl acetate was filtered and dried in vacuo. The cell material and agar were extracted a second time with methanol overnight and subsequently filtered and dried in vacuo. The two extracts were combined and analyzed by HRESI-LC-MS for the detection of small molecules. *Escovopsis* extracts were analyzed on a ThermoScientific Q-Exactive quadrupole orbitrap mass spectrometer coupled to a Dionex UPLC system. The UPLC method was 5% methanol for 0.5 min followed by a gradient from 5% methanol to 97% methanol over 16 min. 97% methanol was held for a 2-min wash before switching back to 5% methanol over 0.5 min and re-equilibrating at 5% methanol for 1 min. The flow rate was kept consistent at 0.35 mL/min. The method was run on a Phenomenex Kinetex XB-C18 column with dimensions 2.1 × 100 mm and 2.6-micron particle size. The mass spectrometer scanned from 200 to 2,000 m/z in positive mode and ion fragmentation was achieved using normalized collision energy of 30%, 35%, and 40%. The profile data were manually inspected and filtered using MzMINE2 (116). Additionally, an aligned feature table and quantification table were created for feature-based molecular networking using MzMINE2. Feature-based molecular networking was performed through GNPS (56,57) with a precursor ion mass tolerance of 0.05 Da, a fragment ion mass tolerance of 0.05 Da, a minimum cosine score of 0.7, and a minimum matched fragment ions of six. The GNPS spectral libraries were searched for matches to submitted MS/MS spectra with a cosine score threshold of 0.7 and minimum of six matched peaks to the library spectra. The resulting network file was visualized and analyzed using Cytoscape v3.8.0 (117).

Data, Materials, and Software Availability. Raw sequences associated with this paper have been deposited with links to Bioproject accession number PRJNA719863 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/), specifically, raw reads are available under the following SRA ids: SRR14176643, SRR14176642, SRR14176633, SRR14176632, SRR14176631, SRR14176630, SRR14176629, SRR14176628, SRR14176627, SRR14176626, SRR14176641, SRR14176640, SRR14176639, SRR14176638, SRR14176637, SRR14176636, SRR14176635, SRR14176634. All code is available at https://github.com/kirstengott/EscoEvoGenomics and Figshare at https://doi.org/10.6084/m9.figshare.21624261 (118), with datasets available through FigShare for antiSMASH annotation (10.6084/m9.figshare.16985137)(119), BiG-SCAPE (10.6084/m9.figshare.16985092) (120), anvi'o (10.6084/m9.figshare.16985119) (121), k-mer-based genome size estimates (10.6084/m9.figshare.16985122) (122), and phylogenetic tree analyses (10.6084/m9.figshare.16985125) (123).

ACKNOWLEDGMENTS. We are grateful to Reed Stubbendieck and Donny Hoang for comments on the manuscript. We thank Heidi Horn, Camila Carlos, Weilan Gomes de Paixão Melo, and Quimi Vidaurre Montoya for their work on the ICBG collections. We acknowledge the Programming for Biology course at Cold Spring Harbor and the course instructors Sofia M. Robb and Simon Prochnik, as well as Eric J. Ross and Jessen V. Bredeson for bioinformatic advice. This material is based on work supported in part by NIH grant U19 TW009872, NIH grant U19 AI142720, and NSF grant DEB-1927155, and São Paulo Research Foundation (FAPESP) grant #2013/50954-0. T.R.S. and J.S.C. were supported by NSF award DEB 1927161. A.B. was supported by the German Research Foundation by Germany's Excellence Strategy-EXC2124-390838134, projects 09.030 and BE 6922/1-1.

1. M. Blackwell, F. E. Vega, Lives within lives: Hidden fungal biodiversity and the importance of conservation. *Fungal Ecol.* **35**, 127–134 (2018).
2. T. Gan *et al.*, Cryptic terrestrial fungus-like fossils of the early Ediacaran period. *Nat. Commun.* **12**, 641 (2021).
3. J. W. Taylor, M. L. Berbee, Dating divergences in the fungal tree of life: Review and new analyses. *Mycologia* **98**, 838–849 (2006).
4. D. Y. Wang, S. Kumar, S. B. Hedges, Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. Biol. Sci.* **266**, 163–171 (1999).
5. F. Lutzoni *et al.*, Contemporaneous radiations of fungi and plants linked to symbiosis. *Nat. Commun.* **9**, 5451 (2018).
6. N. P. Keller, Translating biosynthetic gene clusters into fungal armor and weaponry. *Nat. Chem. Biol.* **11**, 671–677 (2015).
7. M. T. Robey, L. K. Caesar, M. T. Drott, N. P. Keller, N. L. Kelleher, An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2020230118 (2021).
8. E. A. Stahl, J. G. Bishop, Plant-pathogen arms races at the molecular level. *Curr. Opin. Plant Biol.* **3**, 299–304 (2000).
9. L. Noda-Garcia, D. S. Tawfik, Enzyme evolution in natural products biosynthesis: Target- or diversity-oriented? *Curr. Opin. Chem. Biol.* **59**, 147–154 (2020).
10. K. Clay, J. Holah, Fungal endophyte symbiosis and plant diversity in successional fields. *Science* **285**, 1742–1745 (1999).
11. M. G. A. van der Heijden, S. de Bruin, L. Luckerhoff, R. S. P. van Logtestijn, K. Schlaeppi, A widespread plant-fungal-bacterial symbiosis promotes plant biodiversity, plant nutrition and seedling recruitment. *ISME J.* **10**, 389–399 (2016).
12. J. D. Hoeksema *et al.*, Evolutionary history of plant hosts and fungal symbionts predicts the strength of mycorrhizal mutualism. *Commun. Biol.* **1**, 116 (2018).
13. M. G. Branstetter *et al.*, Dry habitats were crucibles of domestication in the evolution of agriculture in ants. *Proc. R. Soc. B Biol. Sci.* **284**, 20170095 (2017).
14. T. R. Schultz *et al.*, The most relictual fungus-farming ant species cultivates the most recently evolved and highly domesticated fungal symbiont species. *Am. Nat.* **185**, 693–703 (2015).
15. C. R. Currie, U. G. Mueller, D. Malloch, The agricultural pathology of ant fungus gardens. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7998–8002 (1999).
16. C. R. Currie, A. E. Stuart, Weeding and grooming of pathogens in agriculture by ants. *Proc. R. Soc. B Biol. Sci.* **268**, 1033–1039 (2001).
17. A. N. M. Bot, C. R. Currie, A. G. Hart, J. J. Boomsma, Waste management in leaf-cutting ants. *Ethol. Ecol. Evol.* **13**, 225–237 (2001).
18. H. T. Reynolds, C. R. Currie, Pathogenicity of *Escovopsis weberi*: The parasite of the attine ant-microbe symbiosis directly consumes the ant-cultivated fungus. *Mycologia* **96**, 955–959 (2004).
19. S. A. Steffan *et al.*, Microbes are trophic analogs of animals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15119–15124 (2015).
20. D. Heine *et al.*, Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat. Commun.* **9**, 1–11 (2018).
21. B. Dhodary, M. Schilg, R. Wirth, D. Spiteller, Secondary metabolites from *Escovopsis weberi* and their role in attacking the garden fungus of leaf-cutting ants. *Chemistry* **24**, 4445–4452 (2018).
22. C. R. Currie, A community of ants, fungi, and bacteria: A multilateral approach to studying symbiosis. *Annu. Rev. Microbiol.* **55**, 357–380 (2001).
23. C. R. Currie, J. A. Scottt, R. C. Summerbell, D. Malloch, Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature* **398**, 701–704 (1999).
24. D. C. Oh, M. Poulsen, C. R. Currie, J. Clardy, Dentigerumycin: A bacterial mediator of an ant-fungus symbiosis. *Nat. Chem. Biol.* **5**, 391–393 (2009).
25. E. B. Van Arnam *et al.*, Selvamicin, an atypical antifungal polyene from two alternative genomic contexts. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12940–12945 (2016).
26. C. R. Currie *et al.*, Ancient tripartite coevolution in the attine ant-microbe symbiosis. *Science* **299**, 386–388 (2003).
27. Q. V. Montoya, M. J. S. Martiarena, R. Bizarria Jr., N. M. Gerardo, A. Rodrigues, Fungi inhabiting attine ant colonies: Reassessment of the genus *Escovopsis* and description of *Luteomyces* and *Sympodiorosea* gens. nov. *IMA Fungus* **12**, 1–18 (2021).
28. T. R. Schultz, S. G. Brady, Major evolutionary transitions in ant agriculture. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5435–5440 (2008).
29. V. E. Masiulionis, M. N. Cabello, K. A. Seifert, A. Rodrigues, F. C. Pagnocca, *Escovopsis trichodermoides* sp. nov., isolated from a nest of the lower attine ant *Mycocepurus goeldii*. *Antonie van Leeuwenhoek. Int. J. General Mol. Microbiol.* **107**, 731–740 (2015).
30. L. A. Meirelles, Q. V. Montoya, S. E. Solomon, A. Rodrigues, New light on the systematics of fungi associated with attine ant gardens and the description of *Escovopsis kreiselii* sp. nov. *PLoS One* **10**, e0112067 (2015).
31. C. A. Leal-Dutra *et al.*, Reclassification of Pterulaceae Corner (Basidiomycota: Agaricales) introducing the ant-associated genus *Myrmecopterula* gen. nov., *Phaeopterula* Henn. and the corticioid Radulomycetaceae fam. nov. *IMA Fungus* **11**, 2 (2020).
32. Q. V. Montoya, M. J. Sutta Martiarena, D. A. Polezel, S. Kakazu, A. Rodrigues, More pieces to a huge puzzle: Two new *Escovopsis* species from fungus gardens of attine ants. *MycoKeys* **46**, 97–118 (2019).
33. E. J. Caldera, M. G. Chevrette, B. R. McDonald, C. R. Currie, Local adaptation of bacterial symbionts within a geographic mosaic of antibiotic coevolution. *Appl. Environ. Microbiol.* **85**, e01580–19 (2019), 10.1128/aem.01580-19.
34. Y. Christopher *et al.*, Disease management in two sympatric *Apterostigma* fungus-growing ants for controlling the parasitic fungus *Escovopsis*. *Ecol. Evol.* **11**, 6041–6052 (2021), 10.1002/ece3.7379.
35. J. O. Augustin *et al.*, Yet more "weeds" in the garden: Fungal novelties from nests of leaf-cutting ants. *PLoS One* **8**, e82265 (2013).
36. K. A. Seifert, R. A. Samson, I. H. Chapela, *Escovopsis aspergilloides*, a rediscovered hyphomycete from leaf-cutting ant nests. *Mycologia* **87**, 407 (2007).
37. J. Muchovej, T. Della Lucia, Escovopsis, a new genus from leaf cutting ant nests to replace *Phialocladus* nomen invalidum. *Mycotaxon* **37**, 191–195 (1990).
38. J. E. Stajich, Fungal genomes and insights into the evolution of the kingdom. *Microbiol. Spectr.* **5**, 1–25 (2017).
39. R. A. Ohm *et al.*, Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen dothideomycetes fungi. *PLoS Pathog.* **8**, e1003037 (2012).
40. L. S. Bittleston, N. E. Pierce, A. M. Ellison, A. Pringle, Convergence in multispecies interactions. *Trends Ecol. Evol.* **31**, 269–280 (2016).
41. A. Kohler *et al.*, Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat. Genet.* **47**, 410–415 (2015).
42. B. E. Wolfe, R. E. Tulloss, A. Pringle, The irreversible loss of a decomposition pathway marks the single origin of an ectomycorrhizal symbiosis. *PLoS One* **7**, e39597 (2012).
43. R. Riley *et al.*, Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9923–9928 (2014).

44. M. P. Nelsen, R. Lücking, C. Kevin Boyce, H. Thorsten Lumbsch, R. H. Ree, The macroevolutionary dynamics of symbiotic and phenotypic diversification in lichens. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 21495–21503 (2020), 10.1073/pnas.2001913117.
45. T. J. B. de Man *et al.*, Small genome of the fungus *Escovopsis weberi*, a specialized disease agent of ant agriculture. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3567–3572 (2016).
46. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
47. G. W. Vurture *et al.*, GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
48. A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, H. Pfister, UpSet: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).
49. N. Gehlenborg, Upsetr: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets. (2019) R package version 1.4.0. https://CRAN.R-project.org/package=UpSetR. Accessed 25 November 2022.
50. T. Lu, B. Yao, C. Zhang DFVF: database of fungal virulence factors. *Database*, bas032 (2012).
51. M. Fischer, J. Pleiss, The lipase engineering database: A navigation and analysis tool for protein families. *Nucleic Acids Res.* **31**, 319–321 (2003).
52. N. D. Rawlings, A. J. Barrett, R. Finn, Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **44**, D343–D350 (2016).
53. H. Zhang *et al.*, dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).
54. B. Jia *et al.*, CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).
55. K. Blin *et al.*, AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**, W36–W41 (2017).
56. M. Wang *et al.*, Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
57. L.-F. Nothias *et al.*, Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **17**, 905–908 (2020).
58. S. A. Kautsar *et al.*, MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **48**, D454–D458 (2020).
59. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
60. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009) https://doi.org/10.1093/bioinformatics/btp352.
61. Y. Liu, J. Schröder, B. Schmidt, Musket: A multistage k-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics* **29**, 308–315 (2013).
62. T. Magoč, S. L. Salzberg, FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
63. A. Bankevich *et al.*, SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
64. M. Alonge *et al.*, RaGOO: Fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* **20**, 224 (2019).
65. N. Dierckxsens, P. Mardulyn, G. Smits, NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**, e18 (2017).
66. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
67. G. Marçais, C. Kingsford, A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011) https://doi.org/10.1093/bioinformatics/btr011.
68. D. Mapleson, G. Garcia Accinelli, G. Kettleborough, J. Wright, B. J. Clavijo, KAT: A K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics* **33**, 574–576 (2017).
69. T. R. Ranallo-Benavidez, K. S. Jaron, M. C. Schatz, GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
70. J. Dainat. AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. Code version v0.8.0. Zenodo. https://www.doi.org/10.5281/zenodo.3552717. Accessed 25 November 2022.
71. M. D. Bennett, I. J. Leitch, H. J. Price, J. S. Johnston, Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be approximately 157 Mb and thus approximately 25% larger than the *Arabidopsis* genome initiative estimate of approximately 125 Mb. *Ann. Bot.* **91**, 547–557 (2003).
72. B. J. Haas *et al.*, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
73. M. Stanke, A. Tzvetkova, B. Morgenstern, AUGUSTUS at EGASP: Using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biol.* **7 Suppl 1**, S11.1-8 (2006).
74. A. F. A. Smit, R. Hubley, P. Green, RepeatModeler Open-1.0. 2008-2015. *Seattle, USA: Institute for Systems Biology.*, 2015. http://www.repeatmasker.org. Accessed 1 May 2018.
75. C. Holt, M. Yandell, MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 1–14 (2011).
76. B. Boeckmann *et al.*, The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**, 365–370 (2003).
77. R. M. Waterhouse *et al.*, BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2018).
78. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
79. M. Manni, M. R. Berkeley, M. Seppey, F. A. Simão, E. M. Zdobnov, BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
80. I. Korf, Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
81. P. Jones *et al.*, InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
82. P. Horton *et al.*, WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587 (2007).
83. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
84. J. C. Navarro-Muñoz *et al.*, A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* **16**, 60–68 (2020).
85. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
86. K. Katoh, K. Misawa, K.-I. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
87. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
88. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
89. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
90. B. Q. Minh *et al.*, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
91. O. Chernomor, A. von Haeseler, B. Q. Minh, Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **65**, 997–1008 (2016).
92. A. Rambaut. FigTree (http://tree.bio.ed.ac.uk/software/figtree/)(August 25, 2021).
93. M. Rabiee, E. Sayyari, S. Mirarab, Multi-allele species reconstruction using ASTRAL. *Mol. Phylogenet. Evol.* **130**, 286–296 (2019).
94. E. Sayyari, S. Mirarab, Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.* **33**, 1654–1668 (2016).
95. H. Li *et al.*, Convergent evolution of complex structures for ant-bacterial defensive symbiosis in fungus-farming ants. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10720–10725 (2018).
96. P. E. Hanisch, J. Sosa-Calvo, T. R. Schultz, The last piece of the puzzle? Phylogenetic position and natural history of the monotypic fungus-farming ant genus Paramycetophylax (Formicidae: Attini). *Insect Syst. Divers.* **6**, 1–17 (2022).
97. V. A. Tagliacollo, R. Lanfear, Estimating improved partitioning schemes for ultraconserved elements. *Mol. Biol. Evol.* **35**, 1798–1811 (2018).
98. R. Lanfear, P. B. Frandsen, A. M. Wright, T. Senfeld, B. Calcott, PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**, 772–773 (2016).
99. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
100. Z. Yang, B. Rannala, Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo Method. *Mol. Biol. Evol.* **14**, 717–724 (1997).
101. J. L. Thorne, H. Kishino, I. S. Painter, Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657 (1998).
102. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).
103. Shen, X.-X. *et al.*, (2020). Genome-scale phylogeny and contrasting modes of genome evolution in the fungal phylum Ascomycota. *Science Advances*, **6**. https://doi.org/10.1126/sciadv.abd0079.
104. F. M. Carpenter, The fossil ants of North America. *Bull. Mus. Comp. Zool.* **70**, 1–66 (1930).
105. P. S. Ward, S. G. Brady, B. L. Fisher, T. R. Schultz, The evolution of myrmicine ants: Phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). *Syst. Entomol.* **40**, 61–81 (2015).
106. C. Baroni Urbani, First description of fossil gardening ants (Amber Collection Stuttgart and Natural History Museum Basel; Hymenoptera: Formicidae. I: Attini). *Stutt. Beitr. Naturkd. B* **54**, 1–13 (1980).
107. B. B. Blaimer, P. S. Ward, T. R. Schultz, B. L. Fisher, S. G. Brady, Paleotropical diversification dominates the evolution of the hyperdiverse ant tribe Crematogastrini (Hymenoptera: Formicidae). *Insect. Syst. Divers.* **2**, 1–14 (2018).
108. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
109. Jari Oksanen, *et al.* (2020). vegan: Community Ecology Package. R packageversion 2.5-7. https://CRAN.R-project.org/package=vegan. Accessed 25 November 2022.
110. C. O. Wilke *et al.*, cowplot. R package version v1.1.1. Zenodo. https://doi.org/10.5281/zenodo.4411966. Accessed 25 November 2022.
111. H. Wickham, ggplot2: Elegant Graphics for Data Analysis. (Version v3.3.2) Springer-Verlag New York. ISBN 978-3-319-24277-4 (2016).
112. R. Kolde, pheatmap: Pretty HeatmapsR package version v1.0.12. CRAN. (https://cran.r-project.org/web/packages/pheatmap/index.html). Accessed 25 November 2022. (2019).
113. R. Suzuki, Y. Terada, H. Shimodaira, pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap ResamplingR package version v2.2-0. CRAN. (https://cran.r-project.org/web/packages/pvclust/index.html). Accessed 25 November 2022. (2019).
114. H. Wickham *et al.*, Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
115. A. M. Eren *et al.*, Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).
116. T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **11**, 395 (2010).
117. P. Shannon *et al.*, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
118. K. Gotting, EscoEvoGenomics. Figshare. Available at https:/doi.org/10.6084/m9.figshare.21624261.v1. Deposited 25 November 2022.
119. K. Gotting, antismash_2021_11_09.tar.gz. Figshare. Available at https:/doi.org/10.6084/m9.figshare.16985137.v1. Deposited 10 November 2021.
120. K. Gotting, bigscape_2021_11_09.tar.gz. Figshare. Available at https:/doi.org/10.6084/m9.figshare.16985092.v1. Deposited 10 November 2021.
121. K. Gotting, anvio_data_2021_11_09.tar.gz. Figshare. Available at https:/doi.org/10.6084/m9.figshare.16985119.v1. Deposited 10 November 2021.
122. K. Gotting, genome_size_kmer_2021_11_09.tar.gz. Figshare. Available at https:/doi.org/10.6084/m9.figshare.16985122.v1. Deposited 10 November 2021.
123. K. Gotting, trees_2021_11_09.tar.gz. Figshare. Available at https:/doi.org/10.6084/m9.figshare.16985125.v1. Deposited 10 November 2021.