Received: 26 July 2017 | Revised: 5 October 2017 | Accepted: 11 October 2017

DOI: 10.1111/mec.14394

SPECIAL ISSUE: THE HOST-ASSOCIATED MICROBIOME: PATTERN, PROCESS AND FUNCTION

WILEY MOLECULAR ECOLOGY

Know your farmer: Ancient origins and multiple independent domestications of ambrosia beetle fungal cultivars

Dan Vanderpool[1] | Ryan R. Bracewell[2] | John P. McCutcheon[1]

[1]Division of Biological Sciences, University of Montana, Missoula, MT, USA

[2]Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, MT, USA

Correspondence
Dan Vanderpool, Division of Biological Sciences, University of Montana, Missoula, MT, USA.
Email: daniel1.vanderpool@umontana.edu

Present address
Ryan R. Bracewell, Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA.

Funding information
NSF GRFP; NSF GROW, Grant/Award Number: DGE-1313190; University of Montana Institute on Ecosystems Graduate Fellowship, Grant/Award Number: IAA-1443108; EPSCoR, Grant/Award Number: EPS-1101342; USDA AFRI NIFA, Grant/Award Number: 2013-67011-21113

Abstract

Bark and ambrosia beetles are highly specialized weevils (Curculionidae) that have established diverse symbioses with fungi, most often from the order Ophiostomatales (Ascomycota, Sordariomycetes). The two types of beetles are distinguished by their feeding habits and intimacy of interactions with their symbiotic fungi. The tree tissue diet of bark beetles is facilitated by fungi, while ambrosia beetles feed solely on fungi that they farm. The farming life history strategy requires domestication of a fungus, which the beetles consume as their sole food source. Ambrosia beetles in the subfamily Platypodinae originated in the mid-Cretaceous (119–88 Ma) and are the oldest known group of farming insects. However, attempts to resolve phylogenetic relationships and the timing of domestication events for fungal cultivars have been largely inconclusive. We sequenced the genomes of 12 ambrosia beetle fungal cultivars and bark beetle associates, including the devastating laurel wilt pathogen, *Raffaelea lauricola*, to estimate a robust phylogeny of the Ophiostomatales. We find evidence for contemporaneous diversification of the beetles and their associated fungi, followed by three independent domestication events of the ambrosia fungi genus *Raffaelea*. We estimate the first domestication of an Ophiostomatales fungus occurred ~86 Ma, 25 million years earlier than prior estimates and in close agreement with the estimated age of farming in the Platypodinae (96 Ma). Comparisons of the timing of fungal domestication events with the timing of beetle radiations support the hypothesis that the first large beetle radiations may have spread domesticated "ambrosia" fungi to other fungi-associated beetle groups, perhaps facilitating the evolution of new farming lineages.

KEYWORDS
ectosymbiosis, genomics, insect agriculture, life history evolution, symbiosis

1 | INTRODUCTION

Fungal farming by insects evolved independently in at least three insect orders: Hymenoptera (ants), Blattodea (termites) and Coleoptera (beetles) (Mueller & Gerardo, 2002). The most well studied of these insect fungiculture systems evolved 61–57 million years ago (Ma) in the Attine ants (Branstetter et al., 2017) which include the leaf-cutters of Central and South America (Currie et al., 2003; De Fine Licht et al., 2013; Mueller, Rehner, & Schultz, 1998; Suen et al., 2011). The evolution of agriculture has resulted in colony sizes numbering in the millions for leaf-cutter ants, making them the dominant herbivore in the New World Tropics (Holldobler & Wilson, 1990). Insect fungiculture also evolved with a single origin 40–25 Ma (Roberts et al., 2016) in the Macrotermitinae termites, which subsequently diversified and spread through Africa and Asia (Mueller & Gerardo, 2002; Aanen & Boomsma, 2005). The third insect–fungal agricultural symbiosis is between ambrosia beetles of the weevil family Curculionidae and (primarily) "Ophiostomatoid" fungi of the

orders Ophiostomatales and Microascales. Ambrosia beetles make up the most speciose and diverse assemblages of both insects and cultivars with an estimated 3,400 species of beetles (Farrell et al., 2001) and a probably comparable number of domesticated fungal species (de Beer & Wingfield, 2013). Within Curculionidae, agriculture evolved multiple times including a minimum of 14 independent transitions to fungal farming in the subfamily Scolytinae (Hulcr & Stelinski, 2017). In contrast, a single transition to fungal farming occurred in the subfamily Platypodinae with all but two of the ~1,400 species practicing fungiculture. The wholly ambrosial "core" Platypodinae (consisting of the tribes Tesserocerini and Platypodini) has a maximal stem age estimated at 119 Ma and a minimum crown age of 88 Ma, making them the oldest known insect farmers (Jordal, 2015).

Scolytinae is comprised of ~4,000 species of bark beetle and ~2,000 species of ambrosia beetle (Farrell et al., 2001; Jordal & Cognato, 2012), the most important distinction between the two beetle types being their primary food source. Both bark and ambrosia beetles are wood-borers that spend the majority of their adult lives under the bark or in the sapwood of dead or dying trees. Bark beetles, however, feed primarily on the plant tissue of these trees, often including the relatively nutrient-rich phloem (i.e., phloeophagy or phloeomycetophagy). A tiny minority of bark beetle species which attack and kill living trees (~1% of 6,000 Scolytines) (Kirkendall, Biedermann, & Jordal, 2014) have gained notoriety for landscape-scale forest devastation throughout North America (Raffa et al., 2008; Tsui et al., 2011). In contrast, ambrosia beetles colonize dead and dying trees by boring past the bark into the sapwood and cultivating fungal "gardens" that breakdown dead tissue and concentrate nutrients for the beetles' consumption (i.e., xylomycetophagy or mycetophagy) (Kasson et al., 2013). "Ambrosia" is a general term used to reference the fungi farmed by ambrosia beetles, and it is well recognized that like the beetles, these fungi are multiply derived within several fungal orders (Cassar & Blackwell, 1996; de Beer & Wingfield, 2013).

Both bark and ambrosia beetles form obligate symbioses with fungi. Morphological adaptations to this lifestyle evolved in many species in the form of exoskeletal fungal "pouches" called mycangia (Francke-Grosmann, 1956) which ensure faithful transmission of symbiont fungal spores. The fungi have reciprocally adapted to insect dispersal by evolving sticky spores on the tips of fruiting structures that act to facilitate attachment to their beetle vector (Malloch & Blackwell, 1993) and by fulfilling beetle nutrition by producing thickened, nutrient-rich conidia ("ambrosial growth") (Neger, 1908). In their association with bark beetles, Ophiostomatoid fungi serve as nutritional symbionts (Bentz & Six, 2006; Six & Klepzig, 2004), terpenoid detoxifiers (Wang et al., 2012, 2014) and pheromone producers (Blomquist et al., 2010). From these mutualisms, more intimate symbioses evolved in the form of domestication and cultivation by groups of farming ambrosia beetles (Batra, 1963; Farrell et al., 2001; Jordal, Normark, & Farrell, 2000). Ambrosia fungi are fully domesticated, vertically transmitted and serve as the sole food source for their insect caretakers. While domestication results in a loss of independence, these fungi are compensated with dispersal to new environments and protection from competing fungi.

Many fungi within the order Ophiostomatales (Ascomycota, Sordariomycetes) are involved in mutualisms with bark and ambrosia beetles. Genomic studies of these fungi have primarily focused on tree-killing bark beetle associates due to the economic impact of these forest pests during outbreaks (DiGuistini et al., 2009). At the time this study began, there were 12 published Ophiostomatales fungal genomes, nine were bark beetle associates, two were human or animal pathogens, and one was an ambrosia beetle fungal cultivar. In 2002, an invasion of the ambrosia beetle *Xyleborus glabratus* to the southeast United States caused alarm when it was recognized to be the vector of *Raffaelea lauricola*, the fungal pathogen responsible for the deadly tree disease known as "laurel wilt" (Fraedrich et al., 2008). Hundreds of millions of Redbay trees throughout the region have succumbed to the disease with other tree species in the *Lauracea* family, including the avocado *Persea americana*, showing signs of susceptibility (Ploetz et al., 2012). This outbreak has prompted a number of studies attempting to describe the biology and mechanism of pathogenicity for *R. lauricola* (Harrington & Fraedrich, 2010; Inch & Ploetz, 2011; Ploetz et al., 2012).

Past efforts to definitively reconstruct evolutionary relationships among Ophiostomatales genera have been inconclusive. The inferred phylogenies were often poorly supported, or the relationships among taxa were unstable across studies, with results varying depending on the taxa or loci sampled, and on the phylogenetic methods employed (Bateman et al., 2017; Dreaden et al., 2014; Massoumi Alamouti, Tsui, & Breuil, 2009; Musvuugwa et al., 2015; Taerum et al., 2013). Phylogenetic uncertainty makes it difficult to address long-standing questions regarding the timing and number of domestication events for ambrosial cultivars in the Ophiostomatales (Dreaden et al., 2014). We wanted to address uncertainties regarding the taxonomy, systematics, number of domestication events and timing of these events in the beetle-associated Ophiostomatales. We used both Illumina and Pacific Biosciences technologies to sequence and annotate 11 Ophiostomatales and one Microascales genomes and transcriptomes. This data set provides the first annotated draft genome for the laurel wilt pathogen, *R. lauricola*, and doubles the number of Ophiostomatales genomes available for comparative genomics. We combined these data with existing genomes from the fungal order and performed a phylogenomic analysis with 18 ingroup taxa and four outgroups. Extensive topology hypothesis testing was performed to address any lingering uncertainty in the resulting whole-genome tree as well as to test hypotheses regarding the monophyly of *Raffaelea*. We used the resulting whole-genome topology as a constraint tree in a phylogenetic analysis with fewer loci but with expanded taxon sampling. This strategy resulted in a well-supported phylogeny for the major clades within the Ophiostomatales as well as confident ancestral state reconstruction (ASR) for life history habit. Additionally, comprehensive molecular dating analyses, bolstered by fossil evidence, were performed revealing the timing of both the origin of Ophiostomatales and multiple fungal domestication events.

2 | METHODS

2.1 | Culturing and library preparation

2.1.1 | Illumina genome sequencing

Fungal cultures were grown at room temperature in 25 ml malt extract broth (17 g malt extract/litre) for 1–5 weeks. Tubes were centrifuged at >4,500 g to remove excess liquid, decanted, and the pellet frozen and ground under liquid nitrogen. Powdered fungal extract was immediately split into several 1.5-ml tubes for nucleic acid extraction. Genomic DNA was isolated using the Mo Bio PowerSoil ® DNA Isolation Kit (Mo Bio Laboratories Inc.)

For genome sequencing, PCR-free multiplexed sequencing libraries were created using the Illumina TruSeq kit with an average insert size of 180 bp. The libraries were pooled in equimolar ratios and sequenced using 101-bp PE sequencing on an Illumina HiSeq 2000. Mate-pair libraries were constructed using the Illumina Nextera Mate Pair kit and size selected to ~4 kb using Pippin Prep (Sage Science). Fragments were circularized and sheared to ~800 bps using a Covaris S220 sonicator. Libraries were pooled in equimolar ratios and sequenced both as 101 bp PE reads on an Illumina HiSeq 2000 and as 259 bp PE reads on an Illumina MiSeq. Quality filtering and trimming of mate-pair Illumina reads were performed using NextClip (Leggett, Clavijo, Clissold, Clark, & Caccamo, 2013).

2.1.2 | Pacific Biosciences genome sequencing

Fungal cultures were grown at room temperature in 25-ml liquid malt extract medium for 3–8 weeks. Tubes containing cultured fungal tissue were centrifuged and decanted to remove excess liquid, and then, fungal tissue was pressed between filter paper discs to further remove liquid. Samples were frozen at −80°C and then transferred to a 4.5-L Labconco benchtop freeze dryer system and dried for >4 hr. Freeze-dried samples were lyophilized using two steel beads per tube with 1–2, 20-s cycles, oscillating at 30 Hz. It is notoriously difficult to extract ultrapure DNA from Ophiostomatoid fungi. To obtain the more than 10 μg of ultrapure, high molecular weight DNA required for Pacific Biosciences (PacBio) SMRT sequencing, we modified the Qiagen Genomic Tip 500 g extraction kit protocol as follows. Lyophilized tissue was added to 12 ml of Y1 buffer containing 1,000 units of lyticase and mixed thoroughly. Each sample was incubated approximately 2 hr at 30°C before it was pelleted and resuspended in 15 ml of buffer G2 containing 30 μl of RNase A (100 mg/ml). 1 ml of Proteinase K stock solution was added, and the buffer was incubated at 50°C for 8–12 hr. Finally, the remaining cellular debris was pelleted and the supernatant retained for processing as outlined in Part II of the Genomic-tip protocol for isolation of genomic DNA. This process was repeated until ~75 μg of HMW DNA was obtained. This HMW DNA was subsequently purified using a MoBio PowerClean Pro DNA Cleanup Kit. While this step ensures ultrapure DNA for PacBio library construction, it reduced total yield to ~15 μg. The integrity and final concentration of genomic DNA were determined using pulse-field gel electrophoresis, fluorescence and absorbance spectrophotometry.

Sequencing libraries were prepared using a SMRTbell Template Prep Kit 1.0 with BluePippin size selection for fragments >20 kb. The resulting libraries were sequenced on four single-molecule real-time (SMRT) cells using P6 version 2 chemistry and reagents by SciLife labs in Uppsala, Sweden.

2.2 | Transcriptome library preparation and sequencing

Total RNA was isolated using the Ambion Ribopure Yeast Kit (Life Technologies). RNA quality and concentration were assessed using either a Bioanalyzer 2100 with an Agilent RNA 6000 Nano Kit (Agilent Technologies) or an Agilent Tape Station 2100. Libraries were prepared from 4 μg total RNA using the Illumina TruSeq Stranded RNA Kit following the manufacturer's protocols. The resulting cDNA libraries were pooled in equimolar ratios and sequenced using an Illumina Hiseq 2500 with either 100-bp or 150-bp PE reads. Illumina reads were trimmed and quality filtered using Trimmomatic version 0.30 (Lohse et al., 2012) with the HEADCROP option enforced for RNA-seq reads.

2.3 | Genome assembly and annotation

2.3.1 | Illumina sequenced libraries

Preliminary assemblies of paired-end reads were performed with Velvet (Zerbino & Birney, 2008) and AllpathsLG (Butler et al., 2008) with similar results. Final assemblies, which included mate-pair data, were performed with AllPathsLG. RNA-seq reads were assembled using Trinity (Grabherr et al., 2011; Haas et al., 2013) with the jaccardclip option enforced for fungal genomes.

Protein annotation of draft assemblies was performed using the MAKER 2.3.36 pipeline (Altschul, Gish, Miller, Myers, & Lipman, 1990; Camacho et al., 2009; Holt & Yandell, 2011; Slater & Birney, 2005). Repetitive elements were masked by first generating classified repeat libraries for each assembled genome using RepeatModeler (Bao, 2002; Benson, 1999; Smit, Hubley, & Green, 1996; Price, Jones, & Pevzner, 2005; Smit & Hubley, 2008). This repetitive sequence library was combined with a MAKER provided transposable element library and was used in RepeatMasker 3.0 (Smit et al., 1996) for masking within the MAKER pipeline. Both protein and RNA-seq data were used as evidence for generating "hints" for ab initio gene predictors. We used the Trinity assembly and the published transcriptome from *Grosmannia clavigera* (DiGuistini et al., 2011) as transcript evidence for each species. Whole proteomes for *G. clavigera* (DiGuistini et al., 2009), *Ophiostoma piceae* (Haridas et al., 2013) and *Magnaporthe grisea* (Dean et al., 2005) were provided as protein evidence. Iterative rounds of training were used to produce hidden Markov models for the ab initio gene predictors SNAP (Korf, 2004) and Augustus (Stanke, Schöffmann, Morgenstern, & Waack, 2006) while self-training was used for GeneMark-ES (Ter-Hovhannisyan,

Lomsadze, Chernoff, & Borodovsky, 2008). Ab initio gene predictors are notorious for overcalling genes (Larsen & Krogh, 2003; Skovgaard, Jensen, Brunak, Ussery, & Krogh, 2001), we set MAKER to only call gene models which had RNA-seq or protein evidence (AED < 1) supporting them (keep_preds = 0). The remaining ab initio gene calls were then scanned for protein family domains (PFAM) using InterProScan5 (IPRscan) (Jones et al., 2014). Predicted proteins lacking functional evidence but that contained at least one PFAM domain were promoted to full gene models using scripts included with MAKER.

Functional annotation of the final protein sets was accomplished by first scanning each proteome using the command line version of InterProScan5. The ipr2gff3 script bundled with MAKER was used to update the MAKER-generated gff3 with the resulting IPRscan information. BLASTP was then used to search each protein set against the UniProt Swiss-Prot database. The best blast hit for each protein was used as the closest gene I.D. for the query protein. The "maker_functional_fasta" and "maker_functional_gff" scripts bundled with MAKER were used to update the respective FASTA and gff3 files with BLASTP results. Finally, genome completeness was assessed using BUSCO 2.0 (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015) on the annotated protein set.

2.3.2 | PacBio sequenced libraries

Genome assembly was performed using the hierarchical genome assembly process (HGAP) (Chin et al., 2013) as implemented in SMRT Portal. Genome-guided transcriptome assembly was performed using Trinity with the jaccardclip option enforced. Protein annotation of HGAP assemblies was performed using the MAKER3 with EVidence Modeler (Haas et al., 2008) pipeline as described above with the following modifications. Proteins from all annotated Ophiostomatales genomes, including the six proteomes annotated from Illumina assemblies in this study, were used as evidence for ab initio gene predictors. Genome-guided Trinity assemblies from all species submitted for PacBio sequencing were used as EST evidence.

At the time this analysis began, there were several Ophiostomatales genomes for which only nucleotide sequence was available. To include these genomes in our analysis, we first had to perform protein annotations using MAKER. *Ophiostoma novo-ulmi* (Forgetta et al., 2013) was annotated as described for the Illumina-sequenced genomes above using MAKER2. There has recently been a functional annotation of *O. novo-ulmi* published (Comeau et al., 2015); however, since our annotation was completed before this publication, we used our version in all downstream analyses. Functional annotation for *Graphilbum fragrans* (B. D. Wingfield et al., 2015) and *Leptographium procerum* (van der Nest et al., 2014) was performed with MAKER3 as described for the PacBio sequenced genomes above.

To determine the number of ribosomal operons present in our PacBio sequenced genomes, we performed iterative BLAST searches, first using rDNA sequences obtained from GenBank as the initial query. Once a full-length operon was identified for each species, it was extracted and used as a query in a second round of BLAST searches to determine the number of partial and full-length operons present in each genome.

2.4 | Phylogenomic analysis

Phylogenomic analyses for eighteen Opiostomatales genomes (11 from this study) and four outgroup taxa were performed. Single-copy, orthologous, protein sequences were identified using OrthoMCL (Li, 2003; van Dongen, 2000) using various inflation parameters ranging from 1.5 to 8. Ortholog sequences were aligned using the L-INSI-I algorithm in MAFFT version 7.271 (Katoh & Standley, 2013) and poorly aligned regions were trimmed using GBlocks (Castresana, 2000; Talavera & Castresana, 2007). Final alignments were analysed using supertree and supermatrix approaches. For the supertree approach, a best fitting evolutionary model of amino acid substitution was selected for each aligned and trimmed amino acid sequence using automatic model selection with free rate heterogeneity incorporated in IQTree 1.5.3 (Nguyen, Schmidt, von Haeseler, & Minh, 2015). Maximum-likelihood (Felsenstein, 1981) phylogenies with 1000 UltraFast bootstrap replicates (Minh, Nguyen, & von Haeseler, 2013) were generated for each locus. We used CONSENSE 3.695 (Felsenstein, 2005) to build consensus bootstrap trees for each loci with all poorly supported nodes (<95% UF bootstrap support) collapsed. IC values were then calculated for an extended majority rule consensus (eMRC) tree using only the remaining, well-supported nodes with RAxML 8.2.9 (Salichos & Rokas, 2013; Salichos, Stamatakis, & Rokas, 2014).

For the supermatrix approach, the same loci used to generate the eMRC tree were concatenated into a single alignment. A maximum-likelihood phylogeny with 1,000 UltraFast (UF) bootstrap replicates was estimated using IQTree 1.5.3 with a posterior mean site frequency (PSMF) model (Wang, Minh, Susko, & Roger, 2017) and 60 mixture classes. This analysis was repeated three times to ensure convergence in topologies. IC values were calculated for the supermatrix topology using the well-supported bootstrap trees from the supertree analysis. The resulting phylogenies were viewed and manipulated using DENDROSCOPE 3.2.10 (Huson & Scornavacca, 2012) and EvolView 2 (He et al., 2016).

2.5 | Choosing among alternative phylogenetic hypotheses

The topologies obtained in the supermatrix and supertree phylogenomic analyses were very similar with the exception of the position of the root node. To differentiate among these alternatives and to compare alternative hypotheses around nodes with low support values, we used the approximately unbiased (AU) test (Shimodaira, 2002) implemented in IQTree1.5.3. Additionally, we used the AU test to compare alternative topologies regarding the monophyly of *Raffaelea* (see Fig. S1). We estimated the likelihood of 12 different topologies under the PSMF model with 60 rate classes using the concatenated amino acid data set. A parsimony starting tree was used to estimate the site frequencies for the 60 rate classes, and 10,000 RELL replicates were performed for each tested tree.

2.6 | Molecular dating

We used PhyloBayes 3.3 (Lartillot, Lepage, & Blanquart, 2009) with site-specific evolutionary rates modelled using nonparametric infinite mixtures (CAT-GTR) to generate an ultrametric tree from our whole-genome concatenated tree. A diffuse gamma prior with the mean equal to the standard deviation was used for the root of the tree. A log-normal autocorrelated relaxed clock model (Thorne, Kishino, & Painter, 1998) with a Dirichlet prior on divergence times was used for dating the tree.

We employed an alternative dating strategy to take advantage of the existing fungal fossil record outside the Ophiostomatales. First, 73 (75 including two outgroup species) whole proteomes spanning Basidiomycetes and Ascomycetes (including a subsample of available genomes in the Ophiostomatales) were clustered using OrthoMCL (Li, 2003; van Dongen, 2000). The resulting 562 single-copy orthologs present in at least 75% of species were aligned and trimmed using MAFFT version 7.271 and Gblocks as above. A maximum-likelihood phylogeny was estimated for the group using RAxML 8.2.4 with GTR+G. To date the resulting phylogeny, six different fossils spanning Basidiomycete/Ascomycete split (Dikaryons) were used as minimum constraints on nodes. Two of these were in the Basidiomycetes and correspond to a minimum age of 90 My for the Agaricales (Hibbett, Grimaldi, & Donoghue, 1997) and a minimum age of 360 My (Stubblefield, Taylor, & Beck, 1985) for the Basidiomycota in general. The remaining four fossils were all of Ascomycete origin. The oldest of these, *Paleopyrenomycites devonicus*, was used to constrain the minimum age of the Pezizomycotina to 395 My (Taylor & Berbee, 2006; Taylor, Hass, Kerp, Krings, & Hanlin, 2005). The second Ascomycete calibration was a minimum age of 136 My for the origin of the Diaporthales (Bronson, Klymiuk, Stockey, & Tomescu, 2013). The third Ascomycete calibration was a minimum age of 99 My for the origin of the genus *Ophiocordyceps* (Sung, Poinar, & Spatafora, 2008). Finally, 50 My was used as the minimum age for the origin of the genus *Aspergillus* (Dorfelt & Schmidt, 2005).

To determine how taxon sampling and fossil calibrations affected dates within the Ophiostomatales, the 75 taxon data set was pared down to 54 whole Ascomycete genomes and the four ascomycete fossil calibrations described above. The Ascomycete taxon set was independently dated using three different amino acid alignments of 60 loci (24,404 amino acids), 40 loci (12,552 amino acids) and 30 loci (11,790 amino acids). To further investigate the effects of fossil constraints and taxon sampling, the 54 taxon Ascomycete data set was pared down to 35 Sordariomycetes (38 with three outgroups) and the *Ophiocordyceps* and *Aspergillus* fossil calibrations from above. The 35 Sordariomycete taxon set was dated using the same 40 locus alignment used in the 75 Dikaryons and the 54 Ascomycete above. Finally, in order to date the 18 whole Ophiostomatales genomes (22 with four outgroup taxa), we used the 95% credibility intervals from our prior dating analyses to constrain two nodes within the Ophiostomatales. The root node was constrained to a maximum of 165 My and a minimum of 56 My. The node corresponding to the most recent common ancestor of *R. lauricola* and *G. clavigera* was constrained to 150 and 46 My. A single fossil calibration was used in the outgroup to constrain the origin of the Diaporthales (Bronson et al., 2013) to a minimum age of 136 My (see Figure 1). The 22 taxon tree was then dated using two independent amino acid alignments of 30 loci (14,763 aa) and 40 loci (18,991 aa). For each alignment, the Markov chain was run for 15,000 steps with two independent runs to ensure convergence. 5,000 steps were discarded as burn-in with each chain sampled every 10 steps for a total of 1,000 samples used to estimate divergence times for each run.

2.7 | Phylogenetic analysis and ancestral state reconstruction

We used the resulting topology from our concatenated phylogenomic analysis as a constraint tree in a second phylogenetic analysis, which included a more comprehensive sampling of Ophiostomatales species. Where possible, we used published GenBank (Clark, Karsch-Mizrachi, Lipman, Ostell, & Sayers, 2016) sequences for ribosomal small subunit (SSU), ribosomal large subunit (LSU), beta-tubulin (tubB), elongation factor-1 alpha (EF1A) and RNA polymerase II (RPB2) genes for many Ophiostomatales. When genomes were available, we extracted full-length sequences from the contigs therein. For species without all gene regions in GenBank, the regions that were available were used. Accession numbers for GenBank sequences are listed in Table S1. Seventy-one ingroup and 15 outgroup (86 total) taxa were included, the majority of which overlapped with taxa in Dreaden et al. (2014). To survey for appropriate outgroup taxa, rDNA sequences from several *Raffaelea* species were compared to GenBank using BLASTN (Camacho et al., 2009) while excluding Ophiostomatales. The resulting closest hits were used as outgroups as well as those species for which genomic data are available and were used in our phylogenomic analyses. Sequences were aligned using MAFFT version 7.271 (Katoh & Standley, 2013) and visually adjusted using AliView (Larsson, 2014) and Mesquite 3.2 (Liu, Endara, & Burleigh, 2015). PartitionFinder 1.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012) was used to find the optimal combination of Eleven data partitions (SSU, LSU, 1st, 2nd and 3rd codon positions of each protein coding gene) resulting in eight partitions modelled with GTR+G using eight gamma rate categories (Tavaré, 1986; Yang, 1993). A maximum-likelihood phylogeny was estimated using RAxML 8.2.9 with 100 bootstrap replicates. The resulting bootstrap trees were filtered in RAxML (-J MR_DROP flag) to eliminate "rogue taxa" (Aberer, Krompass, & Stamatakis, 2013; Pattengale, Aberer, Swenson, Stamatakis, & Moret, 2011) that could mislead the resulting phylogenetic and ASR analyses. After the removal of two taxa (*Afroraffaelea ambrosiae* CBS141678 and *Raffaelea ellipticospora*, 86 taxa topology included as Fig. S3), the maximum-likelihood analysis was repeated with 200 bootstrap replicates.

We used Mesquite to reconstruct ancestral character states for fungal farming on the 84 taxon constrained tree with three different methods: parsimony, likelihood (Pagel, 1999) and stochastic character mapping (Huelsenbeck, Nielsen, & Bollback, 2003; Nielsen, 2002). With parsimony, we reconstructed ancestral states using models that

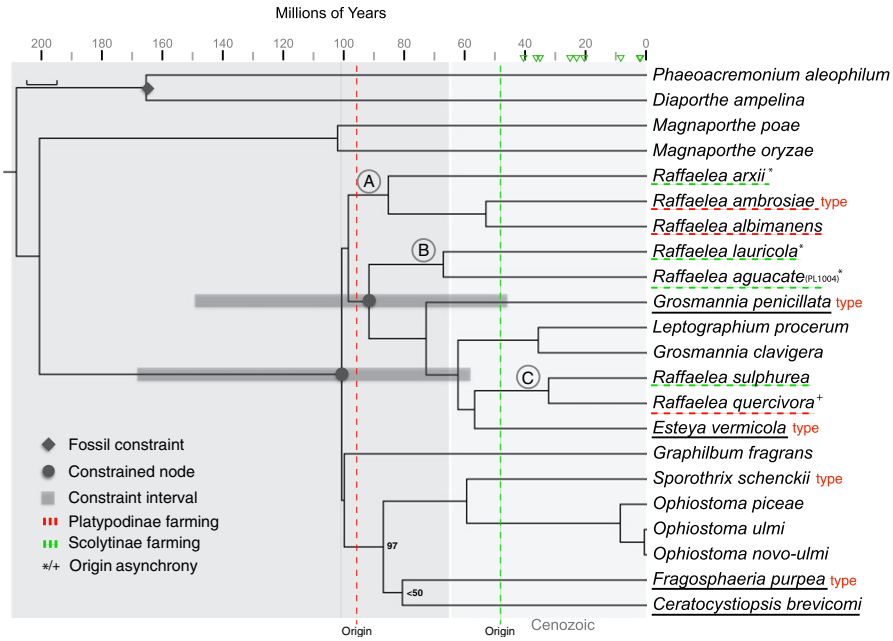

FIGURE 1 The timing of domestication events in Ophiostomatales are indicated by A, B and C. The origin of farming in Platypodinae (red dotted line) is from Jordal (2015) while the timing for the known origins in Scolytinae (green line and triangles) is from Jordal and Cognato (2012) and Gohli et al. (2017). Underlined taxa are genomes sequenced in this study, colours indicate farming beetle associate (Platypodinae/Scolytinae); however, many *Raffaelea* species are promiscuous and associate with more than one beetle. *Indicates an asynchrony in the origin of farming and domestication, that is, farming in the ancestor of the beetle associated with this fungus evolved *after* the ancestor of this fungus was domesticated. +Indicates an asynchrony in origin where the farming in the ancestor of the beetle host evolved *before* the ancestor of this fungus was domesticated. Maximum-likelihood phylogeny estimated from 978,795 aligned amino acids (2,207 single-copy, concatenated and trimmed loci) using IQTree 1.5.3 with a posterior mean site frequency model and 60 mixture classes. A single fossil constraint was implemented in the outgroup species while upper and lower bounds derived from multiple fossil constrained analyses (see Methods) were placed on the ingroup node and the crown node encompassing the *Grosmannia/Leptographium/Esteya/Raffaelea* clade (grey bars). Underlined species represent genomes sequenced in this study. Type species for the genus are indicated in red. Nodes receiving less than maximum bootstrap support are indicated with their respective values [Colour figure can be viewed at wileyonlinelibrary.com]

included unordered parsimony, 2:1 forward-biased stepmatrix and a 1:2 reversal-biased step matrix. In estimating the likelihood reconstruction of ancestral states, we used an asymmetrical Markov k-state 2-parameter model with root state frequencies equal to the equilibrium rate. Rates were optimized from the data by first estimating a single rate under the Mk1 model (Lewis, 2001) with a starting value of 1.0. Next, we estimated both a forward and reverse rate under three scenarios: symmetric forward and reverse rates with starting values of 1.0 and 1.0, a forward-biased scenario with 1.0 and 0.1 starting values and a reverse biased scenarios with 0.1 and 1.0 as starting values. Rates were optimized under each condition and the strategy resulting in the highest likelihood was used in the reconstruction (option 3 in the Mesquite optimization settings). This same model was used in the likelihood stochastic character mapping approach.

3 | RESULTS

3.1 | Ophiostomatales genomes vary twofold in size

The results from our sequencing effort for both Illumina and PacBio technologies are summarized in Tables 1, 2 and S2. The PacBio assemblies produced significantly higher median N50 values with significantly fewer scaffolds than the Illumina assemblies (*t* test $p = .0006$, $p = .024$, respectively). Genome assembly sizes (Table 1) varied within the Ophiostomatales by as much as twofold; however, assembly sizes are often lower than actual genome sizes, especially for short-read assemblies due to the collapsing of repeat regions. Nevertheless, the number of annotated genes correlates with assembly size regardless of the sequencing technology used (Fig. S2). BUSCO (Simão et al., 2015) statistics of genome completeness generated using 290 fungal-specific BUSCOS were not different (*t* test $p = .44$) regardless of sequencing technology used (Table 1). This is an indication that the evidence-based annotation pipeline used here is robust across a range of genome assembly qualities.

For each PacBio assembly, there were multiple large contigs that appear to be whole chromosomes with telomere repeat sequences at the start and end of the contig. We used the number of telomere repeat sequences, as well as contig size and coverage (to ensure repeats were not spurious) to estimate the number of chromosomes for each PacBio sequenced species (Table 2). Additionally, we used BLASTN coupled with coverage to identify the mitochondrial genome in these same species (Table 2).

Genome annotation produced a broad range in the number of annotated proteins within the Ophiostomatales (Table 1). We annotated previously sequenced (but not annotated) genomes from *Graphilbum fragrans* (8,878 protein coding genes) (Wingfield et al., 2015),

TABLE 1 Assembly and annotation statistics for 12 fungal genomes sequenced this study

Species	N50 (mb)	Assembly size (mb)	Max scaffold (mb)	Number scaffolds	Per cent complete[a]	Filtered gene models
Ambrosiella xylebori	1.86	27.1	4.32	38	94.1%	6,503
Ceratocystiopsis brevicomi	1.5	20.6	3.49	198	95.2%	6,327
Raffaelea ambrosiae	3.53	40.7	6.22	69	100%	9,913
Raffaelea arxii	2.2	36.7	5.56	123	98.9%	10,816
Raffaelea lauricola	2.87	34.6	4.91	206	98.2%	9,553
Raffaelea aguacate	0.46	35.9	2.39	414	97.9%	10,194
Raffaelea sulphurea	0.55	23.8	2.47	157	99.6%	7,774
Esteya vermicola[b]	4.98	33.97	7.11	29	96.9%	8,012
Fragosphaeria purpea	5.1	34.75	8.11	25	99.0%	8,493
Grosmannia penicillata	3.49	27.68	4.75	43	99.3%	7,284
Raffaelea albimanens	5.27	39.04	6.77	11	97.9%	9,715
Raffaelea quercivora	3.69	26.41	5.2	23	99.3%	8,003

Species in bold indicate those sequenced with PacBio long read technology, all others sequenced with Illumina short insert libraries and scaffolded with mate-pair libraries.
[a]Denotes fraction of 290 complete fungal BUSCOS present in annotated protein set (does not include fragments).
[b]Denotes estimated statistics for this species after filtering for contaminant.

L. procerum (8,430 protein coding genes) (van der Nest et al., 2014) and *O. novo-ulmi* (8,314 protein coding genes) (Forgetta et al., 2013). Strain information and accession numbers for the corresponding sequencing projects are given for each annotated genome in Table S3. Protein and transcript FASTA files as well as gff3 files containing functional annotation are available at Dryad https://doi.org/10.5061/dryad.tk569.

3.2 | A resolved phylogeny for the Ophiostomatales

After alignment, trimming and concatenation of 2,207 single-copy orthologs present in 18 Ophiostomatales genomes and four outgroup taxa, our concatenated alignment was 978,795 amino acids long. The use of whole-genome protein alignments to resolve the Ophiostomatales phylogeny resulted in slightly different topologies for the supertree and supermatrix approaches. In the supermatrix topology, estimated under a more complex substitution model (Figure 1), the tree is rooted on the internal branch separating *Graphilbum*, *Sporothrix schenckii*, *Fragosphaeria purpea*, *Ceratocystiopsis brevicomi* and *Ophiostoma* genera from the *Raffaelea*, *Grosmannia*, *Esteya vermicola* and *Leptographium* genera. *Graphilbum* (Gf), *Fragosphaeria* (Fp) and *Ceratocystiopsis* (Cb) are all genera represented by a single taxon in this analysis resulting in long terminal branches. Additionally, the available outgroup genomes are comprised of distantly related taxa. The topological differences between the supertree (Fig. S1) and supermatrix (Figure 1) analyses are characterized by whether the long terminal branches of Gf, Fp and Cb associate with the long outgroup branches. The supertree estimated using the relatively simple models available in the automatic model tests implemented in IQTree 1.5.3, roots the tree on the long branch leading to *Gf* (Fig. S1), indicating a potential long-branch attraction issue (Felsenstein, 1978).

The supermatrix topology had maximal bootstrap support for all but two nodes. However, it is well established that concatenation will often result in artificially inflated support values that mask uncertainty in the data. Internode certainty values (IC) estimated using topologies from the 2,207 supertree loci indicate conflict among individual loci with respect to the supermatrix topology (data not shown). We collapsed poorly supported nodes (<95% UF bootstrap support) in the 2,207 loci trees used to calculate IC values for the supermatrix tree. While this strategy improved average IC values for the tree, low IC values still persisted for several nodes (data not shown). Low IC values are indicative of either different evolutionary histories among loci or incorrectly resolved trees due to low or noisy phylogenetic signal (Salichos & Rokas, 2013). However, in this case, differences in the position of the root between the supertree and supermatrix were driven entirely by six of the 2,207 topologies recovered from our supertree analysis. Additionally, both supertree and supermatrix topologies split the ambrosial genus *Raffaelea* into three distinct clades, making the genus polyphyletic. We used the AU test to explore the differences between the supertree and supermatrix topologies, evaluate nodes that were uncertain and to test hypotheses of *Raffaelea* monophyly. The 12 topologies compared in the AU test and their associated likelihood values can be seen in Fig. S1. The supermatrix topology was significantly better than the other topologies tested. IQTree 1.5.3 automatically runs a suite of topology tests in the course of performing the AU test, and the supermatrix tree was preferred among all topologies tested here. Results for all tests are reported in Table S4.

3.3 | Ancient origins and multiple domestication events within the Ophiostomatales

Results of the molecular dating analysis can be seen in Figure 1 and Table 3. The order Ophiostomatales originated ~101 Ma during a rapid radiation that gave rise to two major groups, the *Ophiostoma/Sporothrix/Fragosphaeria/Certatocystiopsis* group and the *Grosmannia/*

TABLE 2 Ribosomal operon counts and estimated number of chromosomes identified for genomes sequenced with PacBio

Species	Complete operons	Partial operons	Average[b] % identity (*SD*)	Telomere repeats	Estimate the number of chromosomes	mtGenome size (kb)
Esteya vermicola[a]	14	3	99.29 (0.30)	12	6	NA[a]
Fragosphaeria purpea	24	4	99.52 (0.81)	10	5	58
Grosmannia penicillata	33	9	99.60 (0.52)	16	8	200
Raffaelea albimanens	9	1	99.91 (0.05)	12	6	139
Raffaelea quercivora	29	2	99.79 (0.28)	12	6	152

Operons were considered full length if the BLASTN match was ~80% of the full-length query sequence. Mitochondrial scaffolds were determined using BLAST coupled with coverage statistics (all identified contigs had ~20× higher coverage).
[a]Indicates low-level contaminant present in the sequenced culture. Only scaffolds that could be unambiguously identified were used in the final analyses; therefore, this count may be an underestimate of operon number.
[b]Indicates average identity to a single query sequence from the same species.

Raffaelea/Leptographium/Esteya group. The ambrosial genus *Raffaelea* appears to have at least three different origins, the earliest of which arose ~86 Ma (crown node age, Figure 1A, Table 3). This clade contains the type species for the genus, *Raffaelea ambrosiae*, as well as *Raffaelea albimanens and Raffaelea arxii* (henceforth RA clade). The second origin of ambrosia fungi occurred ~67 Ma, corresponding to the crown node for the split of *R. lauricola* + *Raffaelea aguacate* (referred to as the *R. lauricola* complex in prior studies, henceforth referred to as the RL clade) (Figure 1B, Table 3). The final origin of fungal domestication within the Ophiostomatales occurred ~33 Ma (Figure 1C, Table 3) when the clade containing *Raffaelea sulphurea and Raffaelea quercivora* split from the lineage that became the nematophagous species *Esteya vermicola*.

The results of the 2,207 loci supertree analysis show similar support for a three domestications hypothesis. Short single-copy ortholog sequences are unlikely to accurately resolve very deep divergences so we collapsed all nodes that were not strongly supported (<95% UFbootstrap) in each of the 2,207 maximum-likelihood trees. There were 117 loci that strongly supported the *Leptographium/Grosmannia/Esteya/Rsulphurea* clade (LGER) grouping with the next most frequent conflicting topology (16) supporting a *R. sulphurea* + *Leptographium/Grosmannia* sister group relationship. Twenty-nine loci strongly support RL+LGER as sister to one another with the next most frequent topology (6) being a RA + RL sister group relationship.

TABLE 3 Dates for major splits within the Ophiostomatales

Node	Mean minimum (crown) age	Mean maximum (stem) age
Ophiostomatales origin	101 ± 25	199 ± 49
R. arxii + *R. ambrosiae* + *R. albimanens*	86 ± 22	99 ± 25
(*R. laur* + *R. aguacate*) + (*Gros.* + *Lept.* + *Esteya* + *Raff*)	92 ± 23	99 ± 25
R. lauricola + *R. aguacate*	67 ± 17	92 ± 23
Esteya + *R. sulphurea* + *R. quercivora*	58 ± 15	63 ± 17
R. sulphurea + *R. quercivora*	33 ± 10	58 ± 15
(*Fragosphaeria* + *Cerato*) + (*Ophiostoma* + *Sporothrix*)	88 ± 22	100 ± 25

Three origins of ambrosia fungi (grey rows). The minimum age for the split of *R. arxii* + *R. ambrosiae* + *R. albimanens* is older than any prior estimate for the evolution of the farming habit in any insect and corresponds closely with the origin of Platypodinae.

Once rogue taxa were filtered from the ingroup, 69 taxa remained in our phylogenetic analysis. The resulting phylogenetic tree had a well-resolved backbone with bootstrap support >70% for the majority of nodes (Figure 2). Our use of a constraint tree resulted in strong support for the root of the tree on the internal branch between the *Ophiostoma* et al. clade and the *Raffaelea* et al. clade. This is highly relevant because the primary shortcoming of all previous analyses of the group was their inability to resolve backbone nodes with any degree of confidence.

Ancestral state reconstruction for fungal domestication events largely resulted in the same conclusion for each methodology employed. Unordered parsimony reconstructed 2–3 domestication events in five steps (not including the *Ambrosiella* sp. outgroups) with the ancestral state at the base of the RA and RL clades being equivocal. Weighted parsimony with a 2:1 backward bias reconstructs only two domestication events, with the most parsimonious reconstruction requiring nine steps. Weighted parsimony with a 2:1 forward bias resulted in three unambiguous domestication events, requiring only five steps. Maximum-likelihood reconstructed three domestication events, although two were not statistically significant. The asymmetric two-parameter model resulted in the best log likelihood of −22.75451379 marginal probability reconstruction. Transition rates were asymmetric with a twofold faster forward rate of 0.3482 and a backward rate of 0.1719. Proportional likelihoods at the ancestral node leading to RA were noncultivar fungi = 0.8050, cultivar = 0.1950. For the ancestral node leading to RL noncultivar fungi = 0.8041, cultivar = 0.1959. For the ancestral node leading to *R. sulphurea* noncultivar fungi = 0.9490, cultivar = 0.051. The majority of stochastic mapping realizations resulted in the tree seen in Figure 2, which is in agreement with both the likelihood reconstruction and the two most parsimonious ASRs.

4 | DISCUSSION

4.1 | Raffaelea monophyly and fungal domestication

Previous studies have disagreed about whether the ambrosia fungi genus *Raffaelea* is monophyletic (Farrell et al., 2001; Harrington,

Aghayeva, & Fraedrich, 2010; Massoumi Alamouti et al., 2009; Ploetz, Hulcr, Wingfield, & de Beer, 2013) or polyphyletic (de Beer & Wingfield, 2013; Dreaden et al., 2014; Musvuugwa et al., 2015; Simmons et al., 2016). The rapid radiation at the base of the Ophiostomatales, compounded by limited data, made it impossible for these studies to confidently resolve deep nodes among major clades within the group. This makes conclusions regarding ASR, taxonomic classification and comparative genomics as robust as the underlying phylogeny on which they are based. In the studies where *Raffaelea* is polyphyletic, the clade containing *R. sulphurea* (henceforth referred to as RS) is more closely related to the *Leptographium/Grosmannia* species. Our phylogenomic analysis indicates that the RL clade shares a more recent common ancestor with the LGER group than with the RA clade which contains the *Raffaelea* type species, *R. ambrosiae* (Figure 1). Furthermore, the crown node for RA is ~20 million years older than that of RL, although the RL stem node slightly overlaps with the RA crown node (Table 3). Maximum-likelihood ASR indicates there were three independent domestication events among taxa currently assigned to *Raffaelea*. de Beer and Wingfield (2013) suggested the RL complex might not share a most recent common ancestor with *Raffaelea* sensu stricto and stated: "*. . .their generic status and position in the Ophiostomatales should await further investigation, including more gene regions, before any new combinations are made.*"-pg. 34. With the inclusion of ~2,200 more gene regions, the analysis presented here fulfils these requirements and validates their viewpoint. Heeding this advice, we elected to exclude *Afroraffaelea ambrosiae* from our primary conclusions on domestication as there was no support for its correct position within the Ophiostamatales (similar to when the species was first described (Bateman et al., 2017)). It is possible that the rapid radiation at the base of the Ophiostomatales resulted in a mis-rooting in our supermatrix analysis, one that would unite the RA and RL clades with increased taxon sampling and a different data set. While we cannot completely exclude this possibility, the RL+LGER relationship was recovered in our 75 whole-genome Dikarya analysis (562 loci), our 54 whole-genome Ascomycete analysis (293 loci), and our 36 whole-genome Sordariomycete analysis (375 loci) (data not shown). Additionally, a three domestications hypothesis was supported in our supertree analysis and in our topology tests (see Results above). It is worth noting, that to facilitate primer design, included in the online supplementary of this publication is ~6,500 bp of aligned CDS from 22 fungal species spanning >200 My of evolution. Newly collected CDS data can be aligned to this matrix and coupled with the phylogenomic constraint tree (Figure 1) for quick placement of additional taxa. The alignment and the Newick formatted constraint tree from Figure 1 are available on Dryad (https://doi.org/10.5061/dryad.tk569).

The challenges associated with accurate ASR are well established (Ekman, Andersen, & Wedin, 2008; Goldberg & Igić, 2008; Griffith et al., 2015; Litsios & Salamin, 2012; Wright, Lyons, Brandley, & Hillis, 2015). Issues such as phylogenetic uncertainty, taxon sampling and transition rates of a character state can influence the reconstruction. To ameliorate these issues as much as possible, we took three different approaches. First, we expanded taxon sampling to include a total of 69 ingroup taxa and 15 outgroup taxa. Increased taxon sampling has been demonstrated to improve phylogenetic accuracy (Hillis, 1998; Philippe et al., 2011; Zwickl & Hillis, 2002) which in turn can increase the accuracy of ASR. Second, we used the supermatrix topology from our phylogenomic analysis as a constraint tree in our phylogenetic analysis. This serves to stabilize previously unresolved backbone nodes in the tree as well as the position of the root. Third, we reconstructed ancestral character states using three different methodologies (following the suggestion of Ekman et al., 2008). Maximum-likelihood, unordered parsimony, and forward-biased, weighted parsimony all reconstructed three domestication events for species currently in *Raffaelea*. Only one possible reversal from farming back to bark or phloem feeding has ever been detected in ambrosia beetles (Hulcr & Stelinski, 2017). Similarly in the ambrosia fungi, there have been no detected cases, of which we are aware, of cultivars reverting to a free-living state. Due to the fungi's dependence on the beetle for protection from competitors and desiccation, transportation to new substrates, as well as very limited sexual reproduction in the fungi (Musvuugwa et al., 2015), reversals seem unlikely (Farrell et al., 2001). Our models still allowed reversals, which makes our estimates conservative. More realistic models would make reversals very unlikely or exclude this possibility altogether, likely resulting in the reconstruction of even more origins of domesticated cultivars.

4.2 | PacBio sequencing produces chromosome-length scaffolds in small fungal genomes

A useful aspect of this study is the direct comparison of different genome sequencing approaches applied to multiple, small eukaryotic genomes of similar size, gene density and repeat content. While slightly more expensive per base, PacBio sequencing of small fungal genomes is the best approach in several respects. First, the hands-on time required for generating large amounts of high molecular weight DNA required for PacBio library preparation was far less than the time invested for building multiple PE and mate-pair libraries. Additionally, as the PacBio sequencing was performed on four SMRT cells for each genome, the average coverage we generated was higher than necessary for accurate genome assembly, and one could get by with less. All together, this suggests that PacBio genomes produce a lower cost per genome compared with Illumina sequencing that includes mate-paired library construction and sequencing, at least for typical Sordariomycete genomes.

One of the issues associated with genome assembly using Illumina short reads is their difficulty in spanning and resolving large repetitive regions. Our short-read assembled genomes produced no assembled ribosomal operons because of this limitation. We were able to overcome this to some extent by reassembling these genomes using Velvet (Zerbino & Birney, 2008) with a low coverage cut-off of ~300×. This resulted in assemblies for the mitochondrial genome and a ribosomal operon for most species; however, this approach does not produce ordered and orientated ribosomal

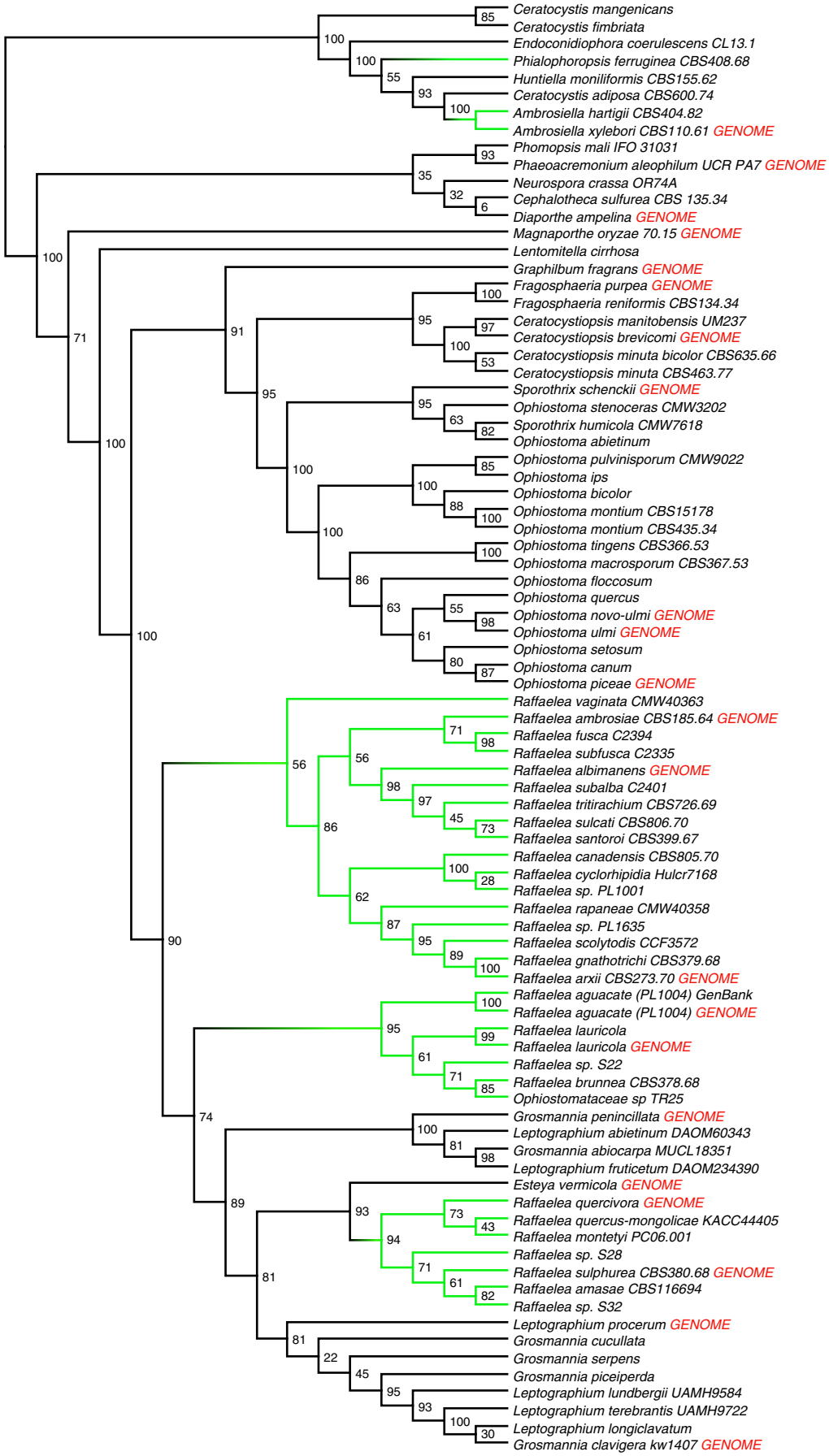
Ceratocystis mangenicans
Ceratocystis fimbriata
Endoconidiophora coerulescens CL13.1
Phialophoropsis ferruginea CBS408.68
Huntiella moniliformis CBS155.62
Ceratocystis adiposa CBS600.74
Ambrosiella hartigii CBS404.82
Ambrosiella xylebori CBS110.61 GENOME
Phomopsis mali IFO 31031
Phaeoacremonium aleophilum UCR PA7 GENOME
Neurospora crassa OR74A
Cephalotheca sulfurea CBS 135.34
Diaporthe ampelina GENOME
Magnaporthe oryzae 70.15 GENOME
Lentomitella cirrhosa
Graphilbum fragrans GENOME
Fragosphaeria purpea GENOME
Fragosphaeria reniformis CBS134.34
Ceratocystiopsis manitobensis UM237
Ceratocystiopsis brevicomi GENOME
Ceratocystiopsis minuta bicolor CBS635.66
Ceratocystiopsis minuta CBS463.77
Sporothrix schenckii GENOME
Ophiostoma stenoceras CMW3202
Sporothrix humicola CMW7618
Ophiostoma abietinum
Ophiostoma pulvinisporum CMW9022
Ophiostoma ips
Ophiostoma bicolor
Ophiostoma montium CBS15178
Ophiostoma montium CBS435.34
Ophiostoma tingens CBS366.53
Ophiostoma macrosporum CBS367.53
Ophiostoma floccosum
Ophiostoma quercus
Ophiostoma novo-ulmi GENOME
Ophiostoma ulmi GENOME
Ophiostoma setosum
Ophiostoma canum
Ophiostoma piceae GENOME
Raffaelea vaginata CMW40363
Raffaelea ambrosiae CBS185.64 GENOME
Raffaelea fusca C2394
Raffaelea subfusca C2335
Raffaelea albimanens GENOME
Raffaelea subalba C2401
Raffaelea tritirachium CBS726.69
Raffaelea sulcati CBS806.70
Raffaelea santoroi CBS399.67
Raffaelea canadensis CBS805.70
Raffaelea cyclorhipidia Hulcr7168
Raffaelea sp. PL1001
Raffaelea rapaneae CMW40358
Raffaelea sp. PL1635
Raffaelea scolytodis CCF3572
Raffaelea gnathotrichi CBS379.68
Raffaelea arxii CBS273.70 GENOME
Raffaelea aguacate (PL1004) GenBank
Raffaelea aguacate (PL1004) GENOME
Raffaelea lauricola
Raffaelea lauricola GENOME
Raffaelea sp. S22
Raffaelea brunnea CBS378.68
Ophiostomataceae sp TR25
Grosmannia penincillata GENOME
Leptographium abietinum DAOM60343
Grosmannia abiocarpa MUCL18351
Leptographium fruticetum DAOM234390
Esteya vermicola GENOME
Raffaelea quercivora GENOME
Raffaelea quercus-mongolicae KACC44405
Raffaelea montetyi PC06.001
Raffaelea sp. S28
Raffaelea sulphurea CBS380.68 GENOME
Raffaelea amasae CBS116694
Raffaelea sp. S32
Leptographium procerum GENOME
Grosmannia cucullata
Grosmannia serpens
Grosmannia piceiperda
Leptographium lundbergii UAMH9584
Leptographium terebrantis UAMH9722
Leptographium longiclavatum
Grosmannia clavigera kw1407 GENOME

FIGURE 2 Maximum-likelihood phylogeny estimated for 84 taxa with many of the available Ophiostomatales sequences from GenBank (see Table S1) including SSU, LSU, β-tubulin, EF1-α, rpb2 (when available). RAxML 8.2.9 was used to estimate the tree using eight separate data partitions modelled using GTR+G and the whole-genome tree in Figure 1 as a constraint tree. Stochastic character mapping for the origin of domestication (green) was based on marginal probability reconstruction with an asymmetric two-parameter model with both rates estimated. Forward rate: 0.34821476, reverse rate: 0.17193707, −log Likelihood: 22.75451379 [Colour figure can be viewed at wileyonlinelibrary.com]

operons in a given genome, nor does it produce an accurate operon count. In contrast, our PacBio assemblies allowed us to ascertain both the placement within a scaffold and accurate copy numbers for most ribosomal operons (Table 2).

4.3 | Fungal farming begets fungal farming

With an origin in the mid-cretaceous (119–88 Ma), the Platypodinae is the oldest known group of fungal farming insects (Jordal, 2015; Jordal & Cognato, 2012; McKenna, Sequeira, Marvaldi, & Farrell, 2009). Their origin and initial diversification coincides with the diversification of other advanced weevil groups in the Curculionidae, including the bark and ambrosia beetles of Scolytinae. McKenna et al. (2009) hypothesize that the massive diversification of Curculionidae during this time was driven by the evolution and diversification of the "core" eudicots, a group comprising ~70% of all angiosperms (Friis, Pedersen, & Crane, 2016). Lacking representatives of all nonfarming Platypodine beetles, Jordal (2015) was unable to precisely date the origin of farming behaviour in the Platypodinae. However, one can infer it occurred with the emergence of the wholly farming "core" Platypodinae consisting primarily of two tribes, the Tesserocerini and Platypodini. In this study, we demonstrate that the beetle-associated order Ophiostomatales arose ~101 Ma, a time that coincides with the initial period of diversification in Platypodinae and Scolytinae weevils (Jordal, 2015; Jordal & Cognato, 2012; Jordal, Sequeira, & Cognato, 2011; McKenna et al., 2009). The short internode branches at the base of the Ophiostomatales tree (Figure 1) are indicative of a rapid radiation, resulting in two primary fungal groups, one containing the genera *Ophiostoma/Sporothrix/Fragosphaeria/Certatocystiopsis/Graphilbum* and the other containing the genera *Grosmannia/Raffaelea/Leptographium/Esteya*. Closely following this initial radiation, ~86 Ma (crown node), the first fungal cultivars arose with the clade containing the *Raffaelea* type species *R. ambrosiae + R. arxii* and *R. albimanens* (Figure 1A, Table 3). While slightly younger, the crown node age is similar to both the diversification of the core eudicots and Platypodinae. Additionally, for sparsely sampled clades such as this, it is unlikely the oldest lineages representative of that clade will be sampled, resulting in consistent underestimate for the age of the crown node. This same bias is likely to be true for the estimated ages of the Platypodinae (Jordal, 2015) and the Scolytinae (Gohli et al., 2017) though less pronounced as estimates for both were based on more extensive taxon sampling. This bias makes it likely that the actual age of the first fungal cultivars is slightly older than 86 million years making them contemporaries of Tesserocerini, the first farming beetle tribe which arose 96 Ma (Figure 3, 95% credibility 88.5–103.4) (Jordal, 2015).

It seems reasonably certain that the Tesserocerini domesticated the first fungi ~86 Ma, because fungal farming did not evolve in the Scolytinae (and the Platypodini did not yet exist) until ~48–60 Ma (Figure 3) (Farrell et al., 2001; Gohli et al., 2017; Jordal, 2015; Jordal & Cognato, 2012). The reason for this lag in the diversification of farming beetle lineages has been puzzled over since it first became apparent (Jordal, 2015; Jordal & Cognato, 2012; Jordal et al., 2011; McKenna et al., 2009). Jordal (2015) speculated that unfavourable conditions for the fungi contributed to low diversification rates and that when temperatures increased ~60 Ma with the onset of the Paleocene-Eocene temperature maximum (PETM) (Zachos, Pagani, Sloan, Thomas, & Billups, 2001), a corresponding increase in farming lineages is observed. Additionally, Jordal points out that favourable conditions during this time likely contributed to the diversification of other taxonomic groups, as well as the evolution of agriculture in ants. While increased global temperature could have played a part in widespread taxonomic diversification, Jordal cautions that other factors such as ecological release brought on by the K-T boundary extinction event (Labandeira, Johnson, & Wilf, 2002) could not be dismissed.

Another possible explanation is that the presence of fungal cultivars facilitates the evolution of the farming habit in the beetles. In addition to inhabiting the same environment and substrate as ambrosia beetles, bark beetles maintain facultative and obligate nutritional symbioses with fungi (Klepzig & Six, 2004). The introduction of ambrosia fungi to their environment would make it a relatively short evolutionary step to fungal farming behaviour. This idea was proposed as a potential model of transition to insect agriculture between beetles and ants (Sanchez-Pena, 2005), but was dismissed as unlikely due to inaccurate (at the time) dating of farming in the beetles and fungal domestication events (Mueller, Gerardo, Aanen, Six, & Schultz, 2005). While it is still unlikely to have occurred between beetles and ants, this model is more parsimonious for some of the independent origins of farming within ambrosia beetles. Several models have been proposed for the transition to agriculture in insects (reviewed in Mueller et al., 2005) but the initial transition to farming in the Tesserocerini likely occurred via a "transmission first" model (Mueller, Schultz, Currie, & Adams, 2001). In this scenario fungi associated with beetle host plants adapt to become dispersed by the beetle. Over time beetles would become dependent on these fungi through nutrient supplementation during passive feeding (much like what is observed in bark beetles today). Eventually cultivation evolves via behavioural/genetic elements that promote increased fungal production capable of sustaining large groups, for example, burrowing, tending, cropping (Biedermann, Klepzig, & Taborsky, 2009; Biedermann & Rohlfs, 2017). After evolving a farming habit

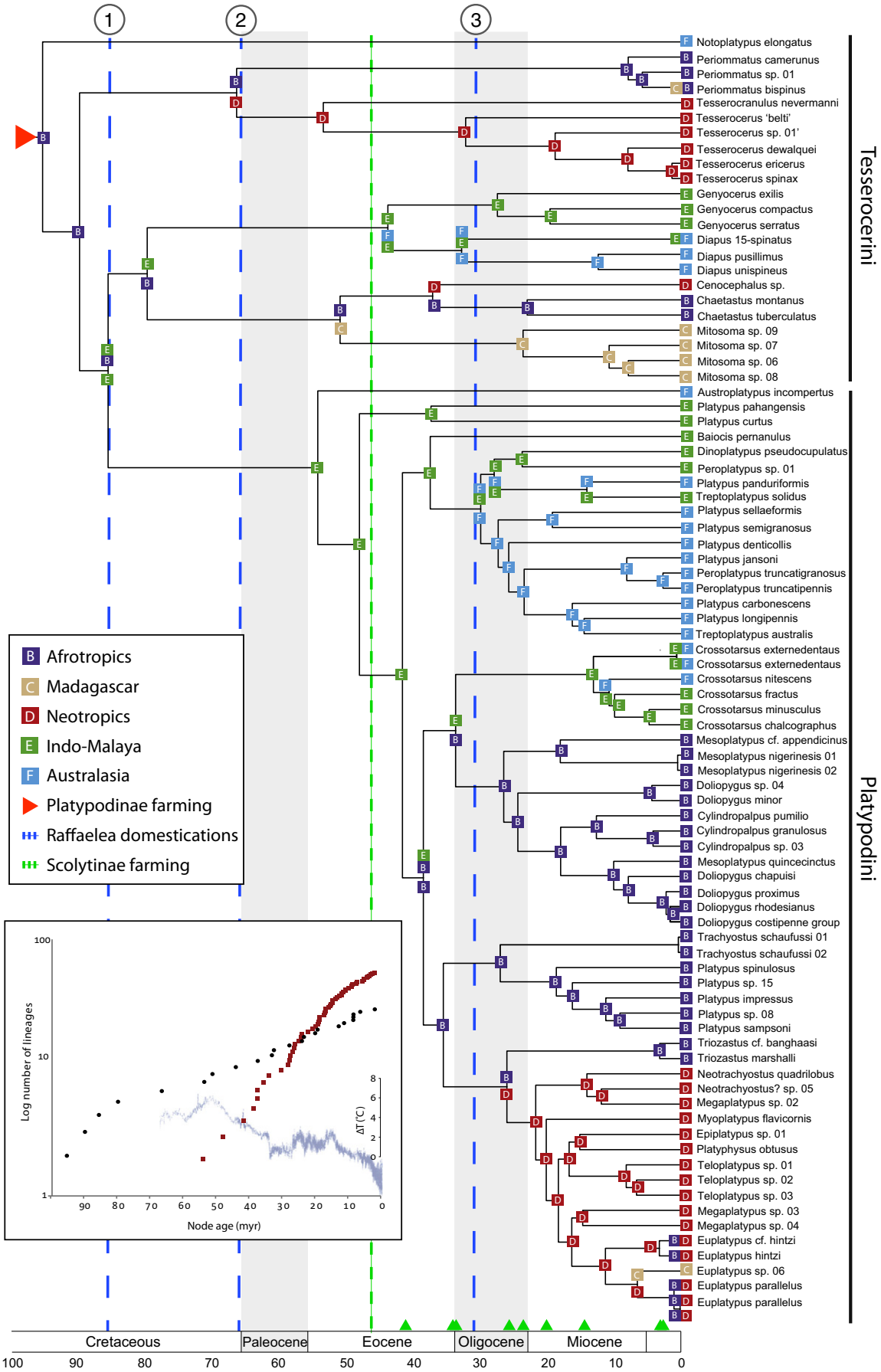
1
2
3
Notoplatypus elongatus
Periommatus camerunus
Periommatus sp. 01
Periommatus bispinus
Tesserocranulus nevermanni
Tesserocerus 'belti'
Tesserocerus sp. 01'
Tesserocerus dewalquei
Tesserocerus ericerus
Tesserocerus spinax
Genyocerus exilis
Genyocerus compactus
Genyocerus serratus
Diapus 15-spinatus
Diapus pusillimus
Diapus unispineus
Cenocephalus sp.
Chaetastus montanus
Chaetastus tuberculatus
Mitosoma sp. 09
Mitosoma sp. 07
Mitosoma sp. 06
Mitosoma sp. 08
Austroplatypus incompertus
Platypus pahangensis
Platypus curtus
Baiocis pernanulus
Dinoplatypus pseudocupulatus
Peroplatypus sp. 01
Platypus panduriformis
Treptoplatypus solidus
Platypus sellaeformis
Platypus semigranosus
Platypus denticollis
Platypus jansoni
Peroplatypus truncatigranosus
Peroplatypus truncatipennis
Platypus carbonescens
Platypus longipennis
Treptoplatypus australis
Crossotarsus externedentaus
Crossotarsus externedentaus
Crossotarsus nitescens
Crossotarsus fractus
Crossotarsus minusculus
Crossotarsus chalcographus
Mesoplatypus cf. appendicinus
Mesoplatypus nigerinesis 01
Mesoplatypus nigerinesis 02
Doliopygus sp. 04
Doliopygus minor
Cylindropalpus pumilio
Cylindropalpus granulosus
Cylindropalpus sp. 03
Mesoplatypus quincecinctus
Doliopygus chapuisi
Doliopygus proximus
Doliopygus rhodesianus
Doliopygus costipenne group
Trachyostus schaufussi 01
Trachyostus schaufussi 02
Platypus spinulosus
Platypus sp. 15
Platypus impressus
Platypus sp. 08
Platypus sampsoni
Triozastus cf. banghaasi
Triozastus marshalli
Neotrachyostus quadrilobus
Neotrachyostus? sp. 05
Megaplatypus sp. 02
Myoplatypus flavicornis
Epiplatypus sp. 01
Platyphysus obtusus
Teloplatypus sp. 01
Teloplatypus sp. 02
Teloplatypus sp. 03
Megaplatypus sp. 03
Megaplatypus sp. 04
Euplatypus cf. hintzi
Euplatypus hintzi
Euplatypus sp. 06
Euplatypus parallelus
Euplatypus parallelus
Tesserocerini
Platypodini
B Afrotropics
C Madagascar
D Neotropics
E Indo-Malaya
F Australasia
Platypodinae farming
Raffaelea domestications
Scolytinae farming
Log number of lineages
100
10
1
90 80 70 60 50 40 30 20 10 0
Node age (myr)
ΔT (°C)
8 6 4 2 0
Cretaceous
Paleocene
Eocene
Oligocene
Miocene
100 90 80 70 60 50 40 30 20 10 0

FIGURE 3 A dated phylogeny estimated from five genes for the Platypodinae ambrosia beetles. Three domestication events of ambrosia fungi currently in *Raffaelea* are labelled 1, 2 and 3 (blue lines). The earliest origin of farming in Scolytinae is indicated by the green dotted line, with subsequent times of origin indicated with green triangles. The timing for the origins of farming behaviour in Scolytine beetles is from Jordal and Cognato (2012) and Gohli et al. (2017). All known instances of farming in Scolytinae evolved after the radiation of Platypodinae. This figure is modified with permission from B.H. Jordal. Inset: lineage-through-time plots for Tesserocerini (black dots) and Platypodini (red squares), with Zachos et al.'s (2001) reconstruction of global average temperature changes plotted in blue. Ancestral areas reconstructed based on the dispersal–extinction–cladogenesis model with dispersal probabilities modelled according to Morley (2003) and Clayton, Soltis, & Soltis, (2009) (see Jordal, 2015 for details) [Colour figure can be viewed at wileyonlinelibrary.com]

~96 Ma, the Tesserocerini maintained a relatively modest diversification rate with only 1–2 lineages existing outside of their point of origin in the Indo-Malaya and Afrotropics (Jordal, 2015). Perhaps brought on by the onset of the PETM (Figure 3 inset) ~55 Ma the Tesserocerini, and especially the Platypodini, experienced massive diversification events, radiating throughout Australasia, Madagascar and the Neotropics (Figure 3). This radiation would have likely permeated the existing tropical forests of the time, exposing many fungus-associated bark beetles to fully domesticated fungal lineages. The difficulty of simultaneously evolving a farming habit and domesticating a fungal lineage would have been circumvented, explaining why shortly after the initial radiation 4 of the 10 farming lineages (for which timing of origin is known) in Scolytinae appeared (Jordal, 2015; Jordal & Cognato, 2012) (see Figure 3). Additionally, while increased temperatures during the PETM may have served as the impetus for beetle radiations and facilitated fungal colonization, fewer than half of Scolytine farming lineages arose near the end of this period and the remaining arose after temperatures cooled to near pre-PETM values (Figure 3 inset, and see Jordal & Cognato, 2012 and Gohli et al., 2017 for details on timing). By comparison, farming in ants and termites has evolved only once despite their undoubtedly intimate contact with various fungi in their environment. In stark contrast to ambrosia beetles (excluding *Austroplatypus* and perhaps some *Ambrosiophilus* species) (Kasson et al., 2016), farming ants and termites are eusocial and form large, well-defended, colonies to the exclusion of other insect species. This social structure likely impedes the comingling of cultivars and behaviour in a common substrate with other nonfarming, mycophagous species (Mueller et al., 2005). This requires nonfarmers to domesticate fungi "from scratch" in order for a new instance to arise.

If the presence of fungal cultivars enables nonfarming bark beetles to transition to fungal farming, we might expect to see evidence of this in beetle-fungal association patterns observed today. We would expect to see an increase in the rate of evolution of beetle farming lineages with a corresponding increase (perhaps with a time lag) in the number of domesticated fungal lineages. An increase in the rate of evolution of beetle farmers is clear: in the first 50 million years of fungal farming, 1 lineage evolved. In the 50 My following PETM radiations, at least 14 lineages evolved. As existing farmers domesticated new fungal lineages (Kasson et al., 2013, 2016; Mayers et al., 2015), an asynchrony between the age of the beetle's farming habit and the time of domestication for fungal lineages is observed. An example of this asynchrony is the Scolytine genus *Gnathotrichus* which did not evolve a farming habit until 16–34 Ma (95% credibility crown node, Gohli et al., 2017) despite its association with *R. gnathotrichi*, a very close relative of *R. arxii*, a lineage whose ancestor was domesticated ~86 Ma (Figure 1A). The reciprocal scenario is observed in the association of *Platypus quercivorus* and *R. quercivora*, where farming evolved in the *Platypus* ancestor ~96 Ma, but the *R. quercivora* ancestor was domesticated just ~33 Ma (Figure 1C, Table 3). A similar situation appears to have occurred in the Microascales cultivars where existing farming lineages appear to have independently domesticated at least three new fungal lineages (Mayers et al., 2015). Studies to confirm the exact nature of beetle-fungi domestications are confounded by the fact that beetle-fungi fidelity is (usually) limited, and beetles of the same species can carry multiple or different symbionts (Batra, 1966; Biedermann, Klepzig, Taborsky, & Six, 2013; Harrington et al., 2010; Kostovcik et al., 2014). It is possible there were two domesticated fungal lineages when the Platypodinae radiated. The timing for the origin of the first cultivar coincides with the origin of the farming tribe Tesserocini ~86 Ma and includes the ancestor of the type species *R. ambrosiae*. The second fungal cultivar to emerge ~67 Ma was the ancestor to RL, and it appears this timing largely overlaps with *very* long stem subtending the Platypodini radiation (Figure 3). Whether the ancestor to RL was domesticated by the Platypodini prior to their radiation is difficult to ascertain given the beetles' propensity for cultivar switching and cosmopolitan distribution. Further interrogation of this question will require exhaustive sampling and dating of all domesticated fungal lineages from four Ascomycete orders (and 3 known Basidiomycete species) (Kolarik & Kirkendall, 2010; Li et al., 2015). Comparisons of cultivar age to dated phylogenies of the Platypodinae and Scolytinae farmers (Gohli et al., 2017; Jordal, 2015; Jordal & Cognato, 2012) are required to determine whether (i) the farming habit evolved first (fungi newly domesticated), (ii) The cultivar arose first (beetle recently evolved farming habit) or (iii) both farming and cultivar arose contemporaneously (farming evolved "from scratch").

5 | CONCLUSIONS

Coincidence in the timing and diversification of Curculionidae and Ophiostomatales fungi suggest codiversification, perhaps driven by the radiation of the "core" eudicots. It is possible these fungi initially inhabited the same substrate as the beetle and were transient hitchhikers during beetle dispersal. Eventually this association evolved into a more intimate symbiosis similar to what we see today among bark beetles and fungi. Shortly after the initial radiation ~96 Ma, the

ancestor to Platypodinae ambrosia beetles made an evolutionary advance when it domesticated a fungus to serve as its sole food source. Despite a robust symbiosis with fungi related to Ophiostomatales cultivars, it would be another 50 million years before the sister group to Platypodinae, Scolytinae, evolved a farming habit. We propose that this lag could potentially be caused by the difficulty of simultaneously evolving a farming behaviour and domesticating a fungal lineage with characteristics adequate to serve as a dependable, sole food source. It is clear that after the Platypodinae radiations occurred they were quickly followed by five independent Scolytinae transitions to the farming habit with five more (that have been dated) occurring over the next ~35 million years. Additionally, with the increased number of farming beetle lineages, we expect an increased rate of new fungal lineage domestications. Without knowing the timing of all domesticated fungal lineages, it is impossible to say for certain whether the positive feedback loop suggested here actually occurred.

Doubtless there are ambrosia beetles other than Platypodinae in which farming evolved "from scratch" in the past 100 million years. This study highlights the importance of having a well-resolved, dated phylogeny for both the beetle (Gohli et al., 2017; Jordal, 2015) and their fungal cultivars for determining whether this is the exception rather than the rule for the evolution of new farming lineages. Future studies that include dating the origins of all fungal cultivars in both Ascomycete and Basidiomycete lineages will help clarify how the evolution of farming lineages proceeds.

ACKNOWLEDGEMENTS

The authors thank Jason Smith for contributing fungal strains for sequencing. Diana Six for general advice on fungal sampling and culture. DV would like to thank Peter Biedermann for extensive guidance on ambrosia beetle collecting and biology as well as helpful comments on this manuscript and Gernot Kunz for kindly providing photographs of ambrosia beetles. DV was supported by grants from NSF GRFP and NSF GROW travel award #DGE-1313190 and by the University of Montana Institute on Ecosystems Graduate Fellowship and Travel Enhancement award #IAA-1443108. DV and JPM were supported by the NSF Experimental Program to Stimulate Competitive Research (EPSCoR) Track-1 EPS-1101342 (INSTEP 3), and RB was supported by a USDA AFRI NIFA (grant no. 2013-67011-21113). "Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation."

DATA ACCESSIBILITY

Genome/transcriptome sequencing and assembly project BioProject: PRJNA395605; GenBank accessions PCDD00000000–PCDD00000000; NCBI SRA Sample accessions: SAMN07414043–SAMN07414054.

The resulting annotation files including gff3 file, protein FASTA files and transcript FASTA files for all taxa annotated this study are available at Dryad https://doi.org/10.5061/dryad.tk569.

Supermatrix alignment, supertree alignments, 86 taxon five gene alignment with defined data partitions file available at: Dryad https://doi.org/10.5061/dryad.tk569.

AUTHOR CONTRIBUTIONS

D.V. and J.M. conceived the study. D.V. cultured, extracted and constructed PE and MP libraries for all species except for *C. brevicomi*. R.B. cultured, extracted and constructed PE libraries for *C. brevicomi*. D.V. performed bioinformatic and phylogenetic analyses. D.V. wrote the manuscript with contributions from J.M. and R.B. All authors read and approved the final manuscript.

ORCID

Dan Vanderpool http://orcid.org/0000-0002-6856-5636

REFERENCES

Aanen, D. K., & Boomsma, J. J. (2005). Evolutionary dynamics of the mutualistic symbiosis between fungus-growing termites and Termitomyces Fungi. In F. E. Vega & M. Blackwell (Eds.), *Insect-fungal associations* (pp. 191–210). New York, NY: Oxford University Press.

Aberer, A. J., Krompass, D., & Stamatakis, A. (2013). Pruning rogue taxa improves phylogenetic accuracy: An efficient algorithm and webservice. *Systematic Biology*, *62*(1), 162–166. https://doi.org/10.1093/sysbio/sys078

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bao, Z. (2002). Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Research*, *12*(8), 1269–1276. https://doi.org/10.1101/gr.88502

Bateman, C., Huang, Y.-T., Simmons, D. R., Kasson, M. T., Stanley, E. L., & Hulcr, J. (2017). Ambrosia beetle *Premnobius cavipennis* (Scolytinae: Ipini) carries highly divergent ascomycotan ambrosia fungus, *Afroraffaelea ambrosiae* gen. nov. et sp. nov. (Ophiostomatales). *Fungal Ecology*, *25*, 41–49. https://doi.org/10.1016/j.funeco.2016.10.008

Batra, L. R. (1963). Ecology of ambrosia fungi and their dissemination by beetles. *Transactions of the Kansas Academy of Science (1903-)*, *66*(2), 213–236.

Batra, L. R. (1966). Ambrosia fungi: Extent of specificity to ambrosia beetles. *Science*, *153*(3732), 193–195. https://doi.org/10.1126/science.153.3732.193

Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, *27*(2), 573.

Bentz, B. J., & Six, D. L. (2006). Ergosterol content of fungi associated with *Dendroctonus ponderosae* and *Dendroctonus rufipennis* (Coleoptera: Curculionidae, Scolytinae). *Annals of the Entomological Society of America*, *99*(2), 189–194. https://doi.org/10.1603/0013-8746(2006)099[0189:ecofaw]2.0.co;2

Biedermann, P. H. W., Klepzig, K. D., & Taborsky, M. (2009). Fungus cultivation by ambrosia beetles: Behavior and laboratory breeding success in three xyleborine species. *Environmental Entomology*, *38*(4), 1096–1105.

Biedermann, P. H. W., Klepzig, K. D., Taborsky, M., & Six, D. L. (2013). Abundance and dynamics of filamentous fungi in the complex ambrosia gardens of the primitively eusocial beetle *Xyleborinus saxesenii* Ratzeburg (Coleoptera: Curculionidae, Scolytinae). *FEMS Microbiology Ecology*, *83*(3), 711–723. https://doi.org/10.1111/1574-6941.12026

Biedermann, P. H., & Rohlfs, M. (2017). Evolutionary feedbacks between insect sociality and microbial management. *Current Opinion in Insect Science*, *22*, 92–100. https://doi.org/10.1016/j.cois.2017.06.003

Blomquist, G. J., Figueroa-Teran, R., Aw, M., Song, M., Gorzalski, A., Abbott, N. L., … Tittiger, C. (2010). Pheromone production in bark beetles. *Insect Biochemistry and Molecular Biology*, *40*(10), 699–712. https://doi.org/10.1016/j.ibmb.2010.07.013

Branstetter, M. G., Ješovnik, A., Sosa-Calvo, J., Lloyd, M. W., Faircloth, B. C., Brady, S. G., & Schultz, T. R. (2017). Dry habitats were crucibles of domestication in the evolution of agriculture in ants. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1852), 20170095. https://doi.org/10.1098/rspb.2017.0095

Bronson, A. W., Klymiuk, A. A., Stockey, R. A., & Tomescu, A. M. F. (2013). A perithecial sordariomycete (Ascomycota, *Diaporthales*) from the lower cretaceous of Vancouver Island, British Columbia, Canada. *International Journal of Plant Sciences*, *174*, 278–292. https://doi.org/10.1086/668227

Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I. A., Belmonte, M. K., Lander, E. S., … Jaffe, D. B. (2008). ALLPATHS: De novo assembly of whole-genome shotgun microreads. *Genome Research*, *18*(5), 810–820. https://doi.org/10.1101/gr.7337908

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. BioMed Central Ltd. https://doi.org/10.1186/1471-2105-10-421

Cassar, S., & Blackwell, M. (1996). Convergent origins of ambrosia fungi. *Mycologia*, *88*(4), 596. https://doi.org/10.2307/3761153

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*(4), 540–552.

Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., … Korlach, J. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods*, *10*(6), 563–569. https://doi.org/10.1038/nmeth.2474

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2016). GenBank. *Nucleic Acids Research*, *44*(D1), D67–D72. https://doi.org/10.1093/nar/gkv1276

Clayton, J. W., Soltis, P. S., & Soltis, D. E. (2009). Recent long-distance dispersal overshadows ancient biogeographical patterns in a pantropical angiosperm family (Simaroubaceae, Sapindales). *Systematic Biology*, *58*, 395–410.

Comeau, A. M., Dufour, J., Bouvet, G. F., Jacobi, V., Nigg, M., Henrissat, B., … Bernier, L. (2015). Functional annotation of the *Ophiostoma novo-ulmi* genome: Insights into the phytopathogenicity of the fungal agent of Dutch elm disease. *Genome Biology and Evolution*, *7*(2), 410–430. https://doi.org/10.1093/gbe/evu281

Currie, C. R., Wong, B., Stuart, A. E., Schultz, T. R., Rehner, S. A., Mueller, U. G., … Straus, N. A. (2003). Ancient tripartite coevolution in the attine ant-microbe symbiosis. *Science*, *299*(5605), 386–388. https://doi.org/10.1126/science.1078155

de Beer, Z. W., & Wingfield, M. J. (2013). Emerging lineages in the Ophiostomatales. In K. A. Seifert, deBeer Z. W. & M. J. Wingfield (Eds.), *Ophiostomatoid fungi* (pp. 21–46). Utrecht, The Netherlands: CBS-KNAW Fungal Biodiversity Centre.

De Fine Licht, H. H., Schiøtt, M., Rogowska-Wrzesinska, A., Nygaard, S., Roepstorff, P., & Boomsma, J. J. (2013). Laccase detoxification mediates the nutritional alliance between leaf-cutting ants and fungus-garden symbionts. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(2), 583–587. https://doi.org/10.1073/pnas.1212709110

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., … Birren, B. W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature*, *434*(7036), 980–986. https://doi.org/10.1038/nature03449

DiGuistini, S., Liao, N. Y., Platt, D., Robertson, G., Seidel, M., Chan, S. K., … Jones, S. J. (2009). De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biology*, *10*(9), R94. https://doi.org/10.1186/gb-2009-10-9-r94

DiGuistini, S., Wang, Y., Liao, N. Y., Taylor, G., Tanguay, P., Feau, N., … Breuil, C. (2011). Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont *Grosmannia clavigera*, a lodgepole pine pathogen. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(6), 2504–2509. https://doi.org/10.1073/pnas.1011289108

Dorfelt, H., & Schmidt, A. R. (2005). A fossil *Aspergillus* from Baltic amber. *Mycological Research*, *109*(08), 956–960. https://doi.org/10.1017/S0953756205003497

Dreaden, T. J., Davis, J. M., de Beer, Z. W., Ploetz, R. C., Soltis, P. S., Wingfield, M. J., & Smith, J. A. (2014). Phylogeny of ambrosia beetle symbionts in the genus *Raffaelea*. *Fungal Biology*, *118*(12), 970–978. https://doi.org/10.1016/j.funbio.2014.09.001

Ekman, S., Andersen, H. L., & Wedin, M. (2008). The limitations of ancestral state reconstruction and the evolution of the ascus in the Lecanorales (lichenized Ascomycota). *Systematic Biology*, *57*(1), 141–156. https://doi.org/10.1080/10635150801910451

Farrell, B. D., Sequeira, A. S., O'Meara, B. C., Normark, B. B., Chung, J. H., & Jordal, B. H. (2001). The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution*, *55*(10), 2011–2027.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, *27*(4), 401–409.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*, 368–376.

Felsenstein, J. (2005). *PHYLIP (Phylogeny Inference Package) version 3.6*. Seattle, WA: Distributed by the author. Retrieved from http://evolution.genetics.washington.edu/phylip/faq.html#citation

Forgetta, V., Leveque, G., Dias, J., Grove, D., Lyons, R., Genik, S., … Dewar, K. (2013). Sequencing of the Dutch elm disease fungus genome using the Roche/454 GS-FLX Titanium system in a comparison of multiple genomics core facilities. *Journal of Biomolecular Techniques: JBT*, *24*(1), 39–49. https://doi.org/10.7171/jbt.12-2401-005

Fraedrich, S. W., Harrington, T. C., Rabaglia, R. J., Ulyshen, M. D., Mayfield, A. E., Hanula, J. L., … Miller, D. R. (2008). A fungal symbiont of the redbay ambrosia beetle causes a lethal wilt in redbay and other Lauraceae in the Southeastern United States. *Plant Disease*, *92*(2), 215–224. https://doi.org/10.1094/PDIS-92-2-0215

Francke-Grosmann, H. (1956). Hautdrusen als Trager der Pilzsymbiose bei Ambrosia-Kafern. *Zeitschrift Fur Morphologie Und Okologie Der Tiere*, *45*(3), 275–308. https://doi.org/10.1007/BF00430256

Friis, E. M., Pedersen, K. R., & Crane, P. R. (2016). The emergence of core eudicots: New floral evidence from the earliest Late Cretaceous. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1845), 20161325. https://doi.org/10.1098/rspb.2016.1325

Gohli, J., Kirkendall, L. R., Smith, S. M., Cognato, A. I., Hulcr, J., & Jordal, B. H. (2017). Biological factors contributing to bark and ambrosia beetle species diversification. *Evolution*, *71*, 1258–1272. https://doi.org/10.1111/evo.13219

Goldberg, E. E., & Igić, B. (2008). On phylogenetic tests of irreversible evolution. *Evolution*, *62*(11), 2727–2741. https://doi.org/10.1111/j.1558-5646.2008.00505.x

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Griffith, O. W., Blackburn, D. G., Brandley, M. C., Van Dyke, J. U., Whittington, C. M., & Thompson, M. B. (2015). Ancestral state reconstructions require biological evidence to test evolutionary hypotheses: A case study examining the evolution of reproductive mode in

squamate reptiles. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *324*(6), 493–503. https://doi.org/10.1002/jez.b.22614

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., … Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, *9*(1), R7. https://doi.org/10.1186/gb-2008-9-1-r7

Haridas, S., Wang, Y., Lim, L., Alamouti, S. M., Jackman, S., Docking, R., … Breuil, C. (2013). The genome and transcriptome of the pine saprophyte *Ophiostoma piceae*, and a comparison with the bark beetle-associated pine pathogen *Grosmannia clavigera*. *BMC Genomics*, *14*(1), 373. https://doi.org/10.1186/1471-2164-14-373

Harrington, T. C., Aghayeva, D. N., & Fraedrich, S. W. (2010). New combinations in *Raffaelea*, *Ambrosiella*, and *Hyalorhinocladiella*, and four new species from the redbay ambrosia beetle, *Xyleborus glabratus*. *Mycotaxon*, *111*(1), 337–361. https://doi.org/10.5248/111.337

Harrington, T. C., & Fraedrich, S. W. (2010). Quantification of propagules of the laurel wilt fungus and other mycangial fungi from the redbay ambrosia beetle, *Xyleborus glabratus*. *Phytopathology*, *100*(10), 1118–1123. https://doi.org/10.1094/PHYTO-01-10-0032

He, Z., Zhang, H., Gao, S., Lercher, M. J., Chen, W.-H., & Hu, S. (2016). Evolview v2: An online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Research*, *44*, W236–W241. https://doi.org/10.1093/nar/gkw370

Hibbett, D., Grimaldi, D., & Donoghue, M. (1997). Fossil mushrooms from Miocene and Cretaceous ambers and the evolution of Homobasidiomycetes. *American Journal of Botany*, *84*(7), 981.

Hillis, D. M. (1998). Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology*, *47*, 3–8. https://doi.org/10.2307/2585228

Holldobler, B., & Wilson, E. O. (1990). *The ants*. Cambridge, MA: The Belknap Press.

Holt, C., & Yandell, M. (2011). MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, *12*(1), 491. https://doi.org/10.1101/gr.403602

Huelsenbeck, J. P., Nielsen, R., & Bollback, J. P. (2003). Stochastic mapping of morphological characters. *Systematic Biology*, *52*(2), 131–158.

Hulcr, J., & Stelinski, L. L. (2017). The ambrosia symbiosis: From evolutionary ecology to practical management. *Annual Review of Entomology*, *62*, 285–303. https://doi.org/10.1146/annurev-ento-031616-035105

Huson, D. H., & Scornavacca, C. (2012). DENDROSCOPE 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology*, *61*(6), 1061–1067. https://doi.org/10.1093/sysbio/sys062

Inch, S. A., & Ploetz, R. C. (2011). Impact of laurel wilt, caused by *Raffaelea lauricola*, on xylem function in avocado, *Persea americana*. *Forest Pathology*, *42*(3), 239–245. https://doi.org/10.1111/j.1439-0329.2011.00749.x

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., … Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jordal, B. H. (2015). Molecular phylogeny and biogeography of the weevil subfamily Platypodinae reveals evolutionarily conserved range patterns. *Molecular Phylogenetics and Evolution*, *92*, 294–307. https://doi.org/10.1016/j.ympev.2015.05.028

Jordal, B. H. B., & Cognato, A. I. A. (2012). Molecular phylogeny of bark and ambrosia beetles reveals multiple origins of fungus farming during periods of global warming. *BMC Evolutionary Biology*, *12*, 133. https://doi.org/10.1186/1471-2148-12-133

Jordal, B. H., Normark, B. B., & Farrell, B. D. (2000). Evolutionary radiation of an inbreeding haplodiploid beetle lineage (Curculionidae, Scolytinae). *Biological Journal of the Linnean Society*, *71*(3), 483–499. https://doi.org/10.1111/j.1095-8312.2000.tb01270.x

Jordal, B. H., Sequeira, A. S., & Cognato, A. I. (2011). The age and phylogeny of wood boring weevils and the origin of subsociality. *Molecular Phylogenetics and Evolution*, *59*(3), 708–724. https://doi.org/10.1016/j.ympev.2011.03.016

Kasson, M. T., O'Donnell, K., Rooney, A. P., Sink, S., Ploetz, R. C., Ploetz, J. N., … Geiser, D. M. (2013). An inordinate fondness for *Fusarium*: Phylogenetic diversity of fusaria cultivated by ambrosia beetles in the genus *Euwallacea* on avocado and other plant hosts. *Fungal Genetics and Biology: FG & B*, *56*, 147–157. https://doi.org/10.1016/j.fgb.2013.04.004

Kasson, M. T., Wickert, K. L., Stauder, C. M., Macias, A. M., Berger, M. C., Simmons, D. R., … Hulcr, J. (2016). Mutualism with aggressive wood-degrading *Flavodon ambrosius* (Polyporales) facilitates niche expansion and communal social structure in *Ambrosiophilus* ambrosia beetles. *Fungal Ecology*, *23*, 86–96. https://doi.org/10.1016/j.funeco.2016.07.002

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kirkendall, L. R., Biedermann, P. H. W., & Jordal, B. H. (2014). Evolution and diversity of bark and ambrosia beetles. In F. E. Vega & R. W. Hofstetter (Eds.), *Bark beetles* (pp. 85–156). San Diego, CA: Academic Press.

Klepzig, K. D., & Six, D. L. (2004). Bark beetle-fungal symbiosis: Context dependency in complex associations. *Symbiosis*, *37*, 189–205.

Kolarik, M., & Kirkendall, L. R. (2010). Evidence for a new lineage of primary ambrosia fungi in *Geosmithia* Pitt (Ascomycota: Hypocreales). *Fungal Biology*, *114*(8), 676–689. https://doi.org/10.1016/j.funbio.2010.06.005

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, *5*(1), 59. https://doi.org/10.1186/1471-2105-5-59

Kostovcik, M., Bateman, C. C., Kolarik, M., Stelinski, L. L., Jordal, B. H., & Hulcr, J. (2014). The ambrosia symbiosis is specific in some species and promiscuous in others: Evidence from community pyrosequencing. *The ISME Journal*, *9*, 126–138. https://doi.org/10.1038/ismej.2014.115

Labandeira, C. C., Johnson, K. R., & Wilf, P. (2002). Impact of the terminal Cretaceous event on plant-insect associations. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(4), 2061–2066. https://doi.org/10.1073/pnas.042492999

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). Partitionfinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*(6), 1695–1701. https://doi.org/10.1093/molbev/mss020

Larsen, T. S., & Krogh, A. (2003). EasyGene – A prokaryotic gene finder that ranks ORFs by statistical significance. *BMC Bioinformatics*, *4*(1), 21. https://doi.org/10.1186/1471-2105-4-21

Larsson, A. (2014). AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, *30*(22), 3276–3278. https://doi.org/10.1093/bioinformatics/btu531

Lartillot, N., Lepage, T., & Blanquart, S. (2009). PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, *25*(17), 2286–2288. https://doi.org/10.1093/bioinformatics/btp368

Leggett, R. M., Clavijo, B. J., Clissold, L., Clark, M. D., & Caccamo, M. (2013). NextClip: An analysis and read preparation tool for Nextera Long Mate Pair libraries. *Bioinformatics*, *30*, 566–568. https://doi.org/10.1093/bioinformatics/btt702

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*(6), 913–925.

Li, L. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, *13*(9), 2178–2189. https://doi.org/10.1101/gr.1224503

Li, Y., Simmons, D. R., Bateman, C. C., Short, D. P. G., Kasson, M. T., Rabaglia, R. J., & Hulcr, J. (2015). New fungus-insect symbiosis: Culturing, molecular, and histological methods determine saprophytic polyporales mutualists of ambrosiodmus ambrosia beetles. *PLoS ONE*, *10*(9), e0137689. https://doi.org/10.1371/journal.pone.0137689

Litsios, G., & Salamin, N. (2012). Effects of phylogenetic signal on ancestral state reconstruction. *Systematic Biology*, *61*(3), 533–538. https://doi.org/10.1093/sysbio/syr124

Liu, J., Endara, L., & Burleigh, J. G. (2015). MatrixConverter: Facilitating construction of phenomic character matrices. *Applications in Plant Sciences*, *3*(2), 1400088. https://doi.org/10.3732/apps.1400088

Lohse, M., Bolger, A. M., Nagel, A., Fernie, A. R., Lunn, J. E., Stitt, M., & Usadel, B. (2012). RobiNA: A user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, *40* (Web Server issue), W622–W627. https://doi.org/10.1093/nar/gks540

Malloch, D., & Blackwell, M. (1993). Dispersal biology of the ophiostomatoid fungi. In M. J. Wingfield, K. A. Seifert, & J. F. Webber (Eds.), *Ceratocystis and ophiostoma: Taxonomy, ecology and pathogenicity* (pp. 195–206). St. Paul, MN: American Phytopathological Society (APS).

Massoumi Alamouti, S., Tsui, C. K. M., & Breuil, C. (2009). Multigene phylogeny of filamentous ambrosia fungi associated with ambrosia and bark beetles. *Mycological Research*, *113*(Pt 8), 822–835. https://doi.org/10.1016/j.mycres.2009.03.003

Mayers, C. G., McNew, D. L., Harrington, T. C., Roeper, R. A., Fraedrich, S. W., Biedermann, P. H. W., … Reed, S. E. (2015). Three genera in the Ceratocystidaceae are the respective symbionts of three independent lineages of ambrosia beetles with large, complex mycangia. *Fungal Biology*, *119*(11), 1075–1092. https://doi.org/10.1016/j.funbio.2015.08.002

McKenna, D. D., Sequeira, A. S., Marvaldi, A. E., & Farrell, B. D. (2009). Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(17), 7083–7088. https://doi.org/10.1073/pnas.0810618106

Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, *30*(5), 1188–1195. https://doi.org/10.1093/molbev/mst024

Morley, R. J. (2003). Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics*, *6*, 5–20.

Mueller, U. G., & Gerardo, N. (2002). Fungus-farming insects: Multiple origins and diverse evolutionary histories. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15247–15249. https://doi.org/10.1073/pnas.242594799

Mueller, U. G., Gerardo, N. M., Aanen, D. K., Six, D. L., & Schultz, T. R. (2005). The evolution of agriculture in insects. *Annual Review of Ecology, Evolution, and Systematics*, *36*, 563–595.

Mueller, U. G., Rehner, S. A., & Schultz, T. R. (1998). The evolution of agriculture in ants. *Science*, *281*(5385), 2034–2038. https://doi.org/10.1126/science.281.5385.2034

Mueller, U. G., Schultz, T. R., Currie, C. R., & Adams, R. (2001). The origin of the attine ant-fungus mutualism. *Quarterly Review of Biology*, *76*, 169–197. https://doi.org/10.2307/2664003

Musvuugwa, T., de Beer, Z. W., Duong, T. A., Dreyer, L. L., Oberlander, K. C., & Roets, F. (2015). New species of Ophiostomatales from Scolytinae and Platypodinae beetles in the Cape Floristic Region, including the discovery of the sexual state of *Raffaelea*. *Antonie van Leeuwenhoek*, *108*(4), 933–950. https://doi.org/10.1007/s10482-015-0547-7

Neger, F. W. (1908). Ambrosiapilze. *Berichte Der Deutschen Botanischen Gesellschaft*, *26*, 735–754.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nielsen, R. (2002). Mapping mutations on phylogenies. *Systematic Biology*, *51*(5), 729–739. https://doi.org/10.1080/10635150290102393

Pagel, M. (1999). The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology*, *48*, 612–622. https://doi.org/10.2307/2585328

Pattengale, N. D., Aberer, A. J., Swenson, K. M., Stamatakis, A., & Moret, B. M. E. (2011). Uncovering hidden phylogenetic consensus in large data sets. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(4), 902–911. https://doi.org/10.1109/TCBB.2011.28

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biology*, *9*(3), e1000602. https://doi.org/10.1371/journal.pbio.1000602

Ploetz, R. C., Hulcr, J., Wingfield, M. J., & de Beer, Z. W. (2013). Destructive tree diseases associated with ambrosia and bark beetles: Black swan events in tree pathology? *Plant Disease*, *97*(7), 856–872. https://doi.org/10.1094/PDIS-01-13-0056-FE

Ploetz, R. C., Pérez Martínez, J. M., Smith, J. A., Hughes, M., Dreaden, T. J., Inch, S. A., & Fu, Y. (2012). Responses of avocado to laurel wilt, caused by *Raffaelea lauricola*. *Plant Pathology*, *61*(4), 801–808. https://doi.org/10.1111/j.1365-3059.2011.02564.x

Price, A. L., Jones, N. C., & Pevzner, P. A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics*, *21*(Suppl 1), i351–i358. https://doi.org/10.1093/bioinformatics/bti1018

Raffa, K. F., Aukema, B. H., Bentz, B. J., Carroll, A. L., Hicke, J. A., Turner, M. G., & Romme, W. H. (2008). Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. *BioScience*, *58*, 501–517.

Roberts, E. M., Todd, C. N., Aanen, D. K., Nobre, T., Hilbert-Wolf, H. L., O'Connor, P. M., … Stevens, N. J. (2016). Oligocene termite nests with in situ fungus gardens from the Rukwa Rift Basin, Tanzania, support a Paleogene African origin for insect agriculture. *PLoS ONE*, *11*(6), e0156847. https://doi.org/10.1371/journal.pone.0156847

Salichos, L., & Rokas, A. (2013). Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature*, *497*(7449), 327–331. https://doi.org/10.1038/nature12130

Salichos, L., Stamatakis, A., & Rokas, A. (2014). Novel information theory-based measures for quantifying incongruence among phylogenetic trees. *Molecular Biology and Evolution*, *31*, 1261–1271. https://doi.org/10.1093/molbev/msu061

Sanchez-Pena, S. R. (2005). New view on origin of attine ant–fungus mutualism: Exploitation of a preexisting insect–fungus symbiosis (Hymenoptera: Formicidae). *Annals of the Entomological Society of America*, *98*(2), 151–164. https://doi.org/10.1603/0013-8746(2005)098[0151:NVOOOA]2.0.CO;2

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, *51*(3), 492–508. https://doi.org/10.1080/10635150290069913

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simmons, D. R., de Beer, Z. W., Huang, Y.-T., Bateman, C., Campbell, A. S., Dreaden, T. J., … Hulcr, J. (2016). New *Raffaelea* species (Ophiostomatales) from the USA and Taiwan associated with ambrosia beetles and plant hosts. *IMA Fungus*, *7*(2), 265–273. https://doi.org/10.5598/imafungus.2016.07.02.06

Six, D. L., & Klepzig, K. D. (2004). Dendroctonus bark beetles as model systems for studies on symbiosis. *Symbiosis*, *37*(1), 207–232.

Skovgaard, M., Jensen, L., Brunak, S., Ussery, D., & Krogh, A. (2001). On the total number of genes and their length distribution in complete microbial genomes. *Trends in Genetics*, *17*(8), 425–428. https://doi.org/10.1016/S0168-9525(01)02372-1

Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, *6*(1), 31. https://doi.org/10.1186/1471-2105-6-31

Smit, A., & Hubley, R. (2008). *RepeatModeler 1.0.7*. Seattle. Retrieved from http://www.repeatmasker.org/RepeatModeler

Smit, A. F. A., Hubley, R., & Green, P. (1996). *RepeatMasker Open-3.0*. Seattle. Retrieved from http://www.repeatmasker.org

Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, *7*, 62. https://doi.org/10.1186/1471-2105-7-62

Stubblefield, S. P., Taylor, T. N., & Beck, C. B. (1985). Studies of Paleozoic fungi. IV. Wood-decaying fungi in Callixylon newberryi from the Upper Devonian. *American Journal of Botany*, *72*(11), 1765. https://doi.org/10.2307/2443734

Suen, G., Teiling, C., Li, L., Holt, C., Abouheif, E., Bornberg-Bauer, E., … Currie, C. R. (2011). The genome sequence of the leaf-cutter ant Atta cephalotes reveals insights into its obligate symbiotic lifestyle. *PLoS Genetics*, *7*(2), e1002007. https://doi.org/10.1371/journal.pgen.1002007

Sung, G.-H., Poinar, G. O., & Spatafora, J. W. (2008). The oldest fossil evidence of animal parasitism by fungi supports a Cretaceous diversification of fungal-arthropod symbioses. *Molecular Phylogenetics and Evolution*, *49*(2), 495–502. https://doi.org/10.1016/j.ympev.2008.08.028

Taerum, S. J., Duong, T. A., de Beer, Z. W., Gillette, N., Sun, J.-H., Owen, D. R., & Wingfield, M. J. (2013). Large shift in symbiont assemblage in the invasive red turpentine beetle. *PLoS ONE*, *8*(10), e78126. https://doi.org/10.1371/journal.pone.0078126

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. https://doi.org/10.1080/10635150701472164

Tavaré, S. (1986). Some probabilistic and statistical problems on the analysis of DNA sequences. *American Mathematical Society: Lectures on Mathematics in the Life Sciences*, *17*, 57–86.

Taylor, J. W., & Berbee, M. L. (2006). Dating divergences in the Fungal Tree of Life: Review and new analyses. *Mycologia*, *98*(6), 838–849.

Taylor, T. N., Hass, H., Kerp, H., Krings, M., & Hanlin, R. T. (2005). Perithecial ascomycetes from the 400 million year old Rhynie chert: An example of ancestral polymorphism. *Mycologia*, *97*(1), 269–285.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., & Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*, *18*(12), 1979–1990. https://doi.org/10.1101/gr.081612.108

Thorne, J. L., Kishino, H., & Painter, I. S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution*, *15*(12), 1647–1657.

Tsui, C. K. M., Roe, A. D., El-Kassaby, Y. A., Rice, A. V., Alamouti, S. M., Sperling, F. A. H., … Hamelin, R. C. (2011). Population structure and migration pattern of a conifer pathogen, *Grosmannia clavigera*, as influenced by its symbiont, the mountain pine beetle. *Molecular Ecology*, *21*(1), 71–86. https://doi.org/10.1111/j.1365-294X.2011.05366.x

van der Nest, M. A., Beirn, L. A., Crouch, J. A., Demers, J. E., de Beer, Z. W., De Vos, L., … Wingfield, B. D. (2014). IMA Genome-F 3: Draft genomes of *Amanita jacksonii*, *Ceratocystis albifundus*, *Fusarium circinatum*, *Huntiella omanensis*, *Leptographium procerum*, *Rutstroemia sydowiana*, and *Sclerotinia echinophila*. *IMA Fungus*, *5*(2), 473–486. https://doi.org/10.5598/imafungus.2014.05.02.11

van Dongen, S. (2000). *Graph clustering by flow simulation*. University of Utrecht. Retrieved from http://www.library.uu.nl/digiarchief/dip/diss/1895620/full.pdf

Wang, Y., Lim, L., DiGuistini, S., Robertson, G., Bohlmann, J., & Breuil, C. (2012). A specialized ABC efflux transporter GcABC-G1 confers monoterpene resistance to *Grosmannia clavigera*, a bark beetle-associated fungal pathogen of pine trees. *The New Phytologist*, *197*(3), 886–898. https://doi.org/10.1111/nph.12063

Wang, Y., Lim, L., Madilao, L., Lah, L., Bohlmann, J., & Breuil, C. (2014). Gene discovery for enzymes involved in limonene modification or utilization by the mountain pine beetle-associated pathogen *Grosmannia clavigera*. *Applied and Environmental Microbiology*, *80*(15), 4566–4576. https://doi.org/10.1128/AEM.00670-14

Wang, H.-C., Minh, B. Q., Susko, E., & Roger, A. J. (2017). Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology*. https://doi.org/10.1093/sysbio/syx068

Wingfield, B. D., Barnes, I., Wilhelm de Beer, Z., De Vos, L., Duong, T. A., Kanzi, A. M., … Wingfield, M. J. (2015). IMA Genome-F 5: Draft genome sequences of *Ceratocystis eucalypticola*, *Chrysoporthe cubensis*, *C. deuterocubensis*, *Davidsoniella virescens*, *Fusarium temperatum*, *Graphilbum fragrans*, *Penicillium nordicum*, *and Thielaviopsis musarum*. *IMA Fungus*, *6*(2), 493–506. https://doi.org/10.5598/imafungus.2015.06.02.13

Wright, A. M., Lyons, K. M., Brandley, M. C., & Hillis, D. M. (2015). Which came first: The lizard or the egg? Robustness in phylogenetic reconstruction of ancestral states. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *324*(6), 504–516. https://doi.org/10.1002/jez.b.22642

Yang, Z. (1993). Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution*, *10*(6), 1396–1401.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, *292*(5517), 686–693. https://doi.org/10.1126/science.1059412

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–829. https://doi.org/10.1101/gr.074492.107

Zwickl, D. J., & Hillis, D. M. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology*, *51*(4), 588–598.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Vanderpool D, Bracewell RR, McCutcheon JP. Know your farmer: Ancient origins and multiple independent domestications of ambrosia beetle fungal cultivars. *Mol Ecol*. 2018;27:2077–2094. https://doi.org/10.1111/mec.14394