Summer 8-31-2021

Foe to Friend: Parallel Domestication of Ophiocordyceps from Fungal Parasite to Beneficial Symbiont in Cicadas

Jason Vailionis
University of Connecticut - Storrs, jason.vailionis@uconn.edu

Eric RL Gordon
University of Connecticut - Storrs, egord003@ucr.edu

Chris Simon
University of Connecticut - Storrs, chris.simon@uconn.edu

Foe to Friend: Parallel Domestication of *Ophiocordyceps* from Fungal Parasite to Beneficial Symbiont in Cicadas

Submitted to the department of Ecology and Evolutionary Biology in partial fulfillment of the requirements for a Bachelor of Science as a University Scholar

University of Connecticut

Sept 2021

Jason Vailionis

Advisor: Dr. Chris Simon

Additional Committee Members: Drs. Paul Lewis, Jonathan Klassen, and Rachel O'Neill

Abstract

Nutritional symbioses are integral to the survival and diversity of many insects. The majority of herbivorous insects in the order Hemiptera possess stable, inherited symbionts that produce essential amino acids and vitamins. However, instability has been observed in cicadas, with one bacterial symbiont, *Hodgkinia cicadicola*, being repeatedly replaced by a new fungal symbiont, *Ophiocordyceps*. The fungal symbionts are thought to be derived from parasitic *Ophiocordyceps* species, but little is known about these parasitic ancestors or how the transition from parasite to mutualist occurs. We used a combination of targeted amplified genes and metagenomic sequencing to investigate the evolution of endosymbiotic *Ophiocordyceps* across 25 species of cicadas in the tribe Cryptotympanini. At least four parallel instances of *Ophiocordyceps* domestication were found in the studied group, arising from a single monophyletic clade of cicada-parasitic *Ophiocordyceps* with only one having been known previously. The genome of a symbiotic *Ophiocordyceps* strain from the cicada *Megatibicen auletes* has been sequenced and annotated, paving the way for future comparative analyses between symbiotic and parasitic *Ophiocordyceps*.

INTRODUCTION

Multicellular life evolved from a world of microbes via symbiosis. That shared evolutionary history is responsible for the complex communication and interdependence that microbes, animals, and plants are now known to form. Next-generation sequencing has unveiled a world of unculturable microbes that form symbiotic relationships with multicellular hosts, ranging from commensal to parasitic to mutualistic. Host-dependent microbes face unique evolutionary pressures and combat host control and immunity, while microbe-dependent multicellular hosts must either reacquire their symbionts every generation or rely on inherited microorganisms whose genomes often degrade over time. Despite these presumed challenges, symbiotic microbes are abundant in nature and have had a marked impact on eukaryote evolution.

Hemipteran Endosymbiosis

Endosymbiosis is any form of symbiotic relationship in which one organism lives inside another. Endosymbiosis is an important source of biological diversity, with many organisms harboring endosymbiotic microbes inside of them. Some of the most prolific are insects, many of which have internal bacteria living in specialized organs or compartments within the gut. Endosymbionts are often required when insects have highly specialized diets on nutritionally-limited substrates Microbes can supplement their hosts' nutritional needs with their expanded metabolism. Insects of the order Hemiptera have an especially high prevalence of endosymbionts. In the mid 20th century, Buchner and his student Müller described the various symbionts of 405 different plant bugs (Hemiptera: Auchenorrhyncha) using light microscopy and

noted the diversity of symbiont identities (Müller 1940, 1962; Buchner 1965). Though not all hemipterans possess endosymbionts, those that are dietary specialists on plant sap require endosymbionts to produce essential amino acids and vitamins. In these hemipteran insects, endosymbionts are typically housed in a specialized "bacteriome" organ in the lower abdomen. Here, provisioned with nutrients by the host (Ankrah et al. 2018), the symbionts grow to a regulated abundance. Symbionts are inoculated into the eggs and thus passed down vertically from mother to offspring. These properties of endo-cellularity, vertical transmission and host control dictate the evolution of endosymbiotic microbes (McCutcheon et al. 2019).

Properties of endosymbionts: genome reduction, and patterns of symbiont replacement and acquisition

Endosymbiotic, vertically transmitted microbes, which spend their entire life cycle within a host, follow a convergent trajectory of genome evolution that begins after they are initially acquired. A free-living bacterium typically has a genome size of 5-10Mb that encodes all functions needed to survive, extract nutrients from its environment, and compete with other microorganisms (Ochman and Davalos 2006). Once the bacterium is incorporated into the host, however, it will never leave and it no longer requires most of its genes. This reduced selection pressure on the symbiont allows deleterious or non-functional mutations to accumulate (Moran 1996). Bacterial populations within the bacteriome have a small effective population size, and a population bottleneck occurs each time they are transmitted from mother to offspring (Moran 1996). These factors result in increased genetic drift, which contributes to the fixation of deleterious mutations. In eukaryotes, small population size tends to lead to genome expansion via accumulation of mobile elements (Lynch 2006), but in bacteria, which have a stronger deletion bias, genome size reduction is typically observed (Mira et al. 2001). Quickly after acquisition, bacterial genomes lose genes until the tiny remaining genome (150-500kb) encodes only essential cellular activities and functions needed for host survival, e.g., amino acid synthesis (McCutcheon and Moran 2012). Additional factors contributing to gene loss include the loss of DNA repair genes and the development of mutator phenotypes, which often occur early after symbiont acquisition by the host (McCutcheon et al. 2019). As endosymbiont genomes become smaller, they converge functionally. They almost always display a strong AT bias, though there are exceptions. Bacteria are thought to have a mutational bias towards AT-rich genomes (Hershberg and Petrov 2010) which is normally counteracted by selection for GC-rich genomes and/or by GC-biased gene conversion, which occurs during recombination (Lassalle et al. 2015). Since endosymbiotic bacteria reproduce asexually and have reduced selection pressure, the AT-mutation bias prevails. Aggressive gene loss leads to small, streamlined genomes that encode only functions necessary for replication and host survival. Endosymbionts reach a point of extreme genomic stability once their genomes are small and mostly protein-coding (Silva et al.

2003). Though this form is stable and can last over 100 million years, it does not seem to persist indefinitely. Deleterious mutations continue to accumulate due to a lack of recombination.

When symbionts lose essential functions like amino acid production, the host must accommodate by horizontal gene transfer, acquiring an additional symbiont, or replacing the old symbiont entirely; hosts that are not successful at this addition or replacement during symbiont degradation will go extinct. Examples of each of these strategies exist in hemipteran insects. For example, the genomes of *Sulcia* endosymbionts from a spittlebug, leafhopper, and a cicada show very high levels of sequence conservation except for the complete and precise loss of the tryptophan biosynthetic pathway in spittlebug-*Sulcia*, a function which is encoded in the spittlebug's other endosymbiont *Zinderia* (McCutcheon and Moran 2010). Presumably this gene loss occurred after the split between cercopids and the others, and it was concurrent with the acquisition of *Zinderia*. Further variation exists within the cercopids, where *Zinderia* has been replaced by a *Sodalis*-like bacterium in a single cercopid tribe (Koga et al. 2013). Additionally, a species of spittlebug with all three symbionts exists, demonstrating the ability of hosts to augment their symbiont capacity by obtaining new symbionts (Koga et al. 2013). Mealybugs have a unique symbiotic relationship in which one gammaproteobacterial endosymbiont lives inside the cytoplasm of another endosymbiont, *Tremblaya.* In the mealybug *Planococcus citri,* the genes required for gammaproteobacterial symbiont *Moranella* to synthesize peptidoglycan are distributed between *Moranella* and horizontally-transferred genes in the host (Bublitz et al. 2019). Like the cercopids, mealybugs have replaced their gammaproteobacterial symbiont multiple times, so mealybug lineages today have gammaproteobacterial symbionts of different ages and origins (Husnik and McCutcheon 2016). These instances of symbiont replacement and horizontal gene transfer represent a successful mechanism for the host to counteract the degrading genomes of its symbiont.

By observing the identity of symbionts throughout Hemiptera, it is clear that symbionts have been repeatedly acquired, replaced, or lost, and that the symbionts of any given lineage are an ever-changing assemblage of 1-3 partners. Acquisition of new symbionts represents an opportunity for ecological adaptation (Sudakaran et al. 2017), even though it can arise out of the seemingly maladaptive process of symbiont genome degradation. The mechanism by which a new symbiont is acquired is uncertain. New symbionts must evade the host immune system and be properly localized within the correct tissue for transmission to offspring. Insect-vectored pathogens or insect parasites represent likely sources of new symbionts due to their ability to invade the host and initiate close intracellular contact over long evolutionary times (McCutcheon et al. 2019). The period of symbiont acquisition remains relatively unexplored because most systems in nature are in the long-lasting stable symbiotic phase. However, cicadas are in an ongoing phase of symbiont transition, and represent a unique opportunity to study this phenomenon.

The ancestral endosymbionts of Cicadidae; genome fragmentation in *Hodgkinia* and transition to *Ophiocordyceps*

Cicadas ancestrally possess two bacterial endosymbionts, *Candidatus* Sulcia muelleri (Bacteroidetes; referred to as *Sulcia*) and *Candidatus* Hodgkinia cicadicola (Alphaproteobacteria; referred to as *Hodgkinia*) (McCutcheon et al. 2009a). *Sulcia* produces 8 essential amino acids: leucine, valine, isoleucine, threonine, lysine, arginine, tryptophan, and phenylalanine, while *Hodgkinia* produces the two remaining essential amino acids: histidine, and methionine, as well as vitamin B_{12} (McCutcheon and Moran 2010). *Sulcia* was acquired in an ancestor of all auchennorrhynchan insects and has remained as a symbiotic partner for at least 260 million years (Moran et al. 2005). *Sulcia* is still present in most auchennorrhyncan lineages but in all known cases it can only synthesize 8 amino acids or fewer, implying the existence of another symbiont in the ancestor of Auchennorrhyncha (Bennett and Moran 2013). *Hodgkinia* is unique to cicadas, with an age of at least 100 million years (Simon et al. 2019). *Hodgkinia* has one of the smallest known bacterial genomes at around 150 kb while the *Sulcia* genome is around 250 kb (McCutcheon et al. 2009a, 2009b; McCutcheon and Moran 2012). Interestingly, *Hodgkinia* appears to be in the final stage of symbiont evolution, and is displaying unique and chaotic genome expansion not seen in any other symbionts. The first sequenced *Hodgkinia* genome from *Diceroprocta semicincta* had a single circular genome of 144 kb, but subsequent sequencing in a different cicada, *Tettigades undata,* revealed that *Hodgkinia* had fragmented into two interdependent *Hodgkinia* species inhabiting the same bacteriome (Van Leuven et al. 2014). These two strains lost genes or developed pseudogenes in a complementary pattern so that at least one copy of each ancestral *Hodgkinia* gene remained functional between the two. The overall coding capacity of the *Hodgkinia* genomes remained the same as the single, complete *Hodgkinia* genome, but with almost twice the genome size distributed between two genomes. Van Leuven et al. hypothesized that this lineage splitting could be a nonadaptive product of endosymbiont evolution where multiple inactivating mutations arise within a bacteriome, are masked due to *Hodgkinia*'s polyploidy, reach abundance due to the cicadas' long life cycle, and become fixed due to population bottleneck when passed to offspring. Since then, *Hodgkinia* lineage splitting has been found to be much more widespread. Further sequencing of *Tettigades* species' *Hodgkinia* symbionts revealed that multiple independent lineage splits have occurred during the past four million years of evolution in this group, resulting in closely related cicada species with 1-6 *Hodgkinia* species and varied genomic organization among them (Łukasik et al. 2018). All original coding genes were retained between the *Hodgkinia* species but the relative abundance of many genes varies. In *Magicicada*, the cicadas with the longest life cycle from egg to adult, tens of *Hodgkinia* lineages exist in a single cicada with high variance in size, number of lineages, and gene copy number between *Magicicada* species (Campbell et al. 2015, 2017). The genomes can total over 1 mb (Campbell et al. 2015), an almost 10-fold increase in nucleotides

from the ancestral *Hodgkinia* genome, but each individual *Hodgkinia* genome is still characterized by genome reduction.

Because of the varied gene dosage and because cicadas must transfer more bacteria to each offspring to ensure all *Hodgkinia* copies are present, *Hodgkinia* lineage splitting is thought to be maladaptive for the cicada and can lead to a feedback loop worsening genome degradation (Van Leuven et al. 2014; Campbell et al. 2017; Łukasik et al. 2018). If this trait is maladaptive, it may explain the recent discovery that some cicadas have acquired a new fungal symbiont, *Ophiocordyceps*, and subsequently lost *Hodgkinia* (Matsuura et al. 2018). Fifteen of twenty-four sampled Japanese cicadas from multiple cicada lineages lacked *Hodgkinia* and possessed *Ophiocordyceps* symbionts, consistent with at least three independent acquisitions of *Ophiocordyceps*. These independent acquisitions are marked by structural diversity in tissue localization (Matsuura et al. 2018; Wang et al. 2021). The transition in symbionts could be partially explained by *Hodgkinia* genome fragmentation, but unpublished work from our lab (Eric RL Gordon) does not find a clear relationship between increased fragmentation in extant sister groups and *Hodgkinia* replacement. *Ophiocordyceps* has also emerged in hemipteran insects that never possessed a *Hodgkinia*-like symbiont. Yeast-like symbionts, often confirmed as *Ophiocordyceps*, have been noted in multiple Hemipteran groups - both Sternorrhyncha (aphids mealybugs, scales and relatives) and Auchenorrhyncha (leafhoppers, planthoppers, cicadas, spittlebugs and relatives) but not in Heteroptera. They have been reported in leafhoppers of the subfamilies Ledrinae (Nishino et al. 2016) and Deltocephalinae (Kobiałka et al. 2018), delphacid planthoppers (Noda 1977; Hongoh and Ishikawa 2000), and cerataphidine aphids (Hongoh and Ishikawa 2000; Vogel and Moran 2013). *Ophiocordyceps*-allied fungi have been found in the cytoplasm of fat body cells in kermesid scales (Podsiadło et al. 2018), in *Kerria lacca*, a kerriid scale (Vashishtha et al. 2011), and as a prevalent microbiome member in coccid scales (Gomez-Polo et al. 2017), representing a putative symbiont in those organisms. *Ophiocordyceps* appears uniquely capable of forming mutualistic symbioses compared to other fungal insect pathogens. The tight internal relationship that *Ophiocordyceps* parasites form with their host might facilitate the transition to host-beneficial endosymbionts, especially if a pre-existing symbiont is undergoing maladaptive genome degradation.

Ophiocordyceps **taxonomy, the nature of the parasite**

Ophiocordyceps, the newly acquired symbiont, is derived from a naturally occurring fungal parasite which kills cicada nymphs before they emerge from the ground, sprouting a fruiting body from their carcass and releasing spores. In cicadas that have *Ophiocordyceps* as a symbiont, the fungi never develop hyphae but persist inside the cicada in a yeast-like form (Matsuura et al. 2018). The genus *Ophiocordyceps* contains around 140 described species of entomoparasitic fungi (Ainsworth 2008). Ophiocordycipitaceae fungi were previously included in the genus *Cordyceps* within the family Clavicipitaceae. This group has been reclassified into

three families, Ophiocordycipitaceae, Clavicipitaceae, and Cordycipitaceae, with Ophiocordycipitaceae and Clavicipitaceae as sister groups (Sung et al. 2007). Further naming confusion comes from the tendency of single fungi to be given two names, one for their anamorphic (asexual) and one for their teleomorphic (sexual) phase, since these phases can have very distinct morphologies. *Ophiocordyceps* can correspond to anamorphic fungi under the names *Hirustella, Hymenostilbe, Paraisaria,* or *Synglíocladium*. Other cicada-infecting fungi, *Beauveria* and *Isaria*, are anamorphs of *Cordyceps* fungi. Together, these three families comprise one clade of Hypocreales, an ascomycete fungal order. The insect-parasitic lifestyle has evolved independently in *Ophiocordycipitaceae, Cordycipitaceae,* and some *Clavicipitaceae* (Wang et al. 2016).

The most well-known members of *Ophiocordyceps* are the "zombie-ant" fungi (species complex *Ophiocordyceps unilateralis*), which manipulate the behavior of their hosts to increase transmission of fungal spores and evade the social defenses of ants. Infected ants leave their colony, climb up the stems of plants above foraging trails and clasp leaf blades with their jaws. Then, the fruiting body emerges and disperses its spores onto the foraging ants below. The zombie-ant *Ophicordyceps* are species specific to their hosts, likely due to the complex nature of their host manipulation (de Bekker et al. 2017). However, other *Ophiocordyceps* species show limited or no behavioral manipulation, and less host specificity. Since the evolution of an entomoparasitic lifestyle in the ancestor of Ophiocordycipitaceae, the fungi have gone through multiple order-level host switches and behavioral manipulation has evolved at least once (Hughes et al. 2016; Araújo and Hughes 2019). Today, *Ophiocordyceps* are known to infect at least nine insect orders, as well as trap-door spiders (Hughes et al. 2016). The most common hosts of *Ophiocordyceps* are in the orders Hemiptera, Hymenoptera, Coleoptera, and Lepidoptera. In addition to switching between hosts, *Ophiocordyceps* can specialize on either immature or adult forms of their host. In cicadas, *Ophiocordyceps* fruiting bodies have only been found on nymphs. Coleoptera-infecting *Ophiocordyceps* comprise multiple different clades which specialize on either adult or larvae (but never both).

Ophiocordyceps have evolved to parasitize cicadas numerous times, but all currently known cicada symbiont species appear to have arisen from one monophyletic clade of *Ophiocordyceps* parasites including *O. sobolifera, O. longissima,* and *O. yakusimensis* (Matsuura et al. 2018). Due to lack of sampling of *Ophiocordyceps* parasites outside of Japan, it is likely that there is undiscovered diversity of *Ophiocordyceps* parasites that has contributed to the evolution of *Ophiocordyceps* symbionts. Recent data from our lab and collaborators' supports over ten *Ophiocordyceps* acquisitions in phylogenetically and geographically distinct groups of cicadas (Haji et al. 2021). Documenting the ongoing *Ophiocordyceps* acquisition throughout Cicadidae provides a fruitful opportunity to better understand the process of symbiont domestication.

Dog-Day cicadas as a study system for *Ophiocordyceps* acquisition; Goals

The common Dog-Day cicadas of North America comprise two genera, *Neotibicen* and *Megatibicen,* within the tribe Cryptotympanini. They are especially well-known for producing loud mating calls, which are the sound of summer across much of the United States. This group is of interest for the study of symbiont transition because of its diversity, geographic distribution, and relationship to other *Ophiocordyceps*-possessing cicadas. *Neotibicen* and *Megatibicen* together are sister to the large genus *Cryptotympana*, which is exclusive to Asia (Hill et al. 2015). *Neotibicen* and *Megatibicen* are largely native to central and eastern North America although some extend into Mexico and the western US) (Sanborn and Phillips 2013; Hill et al. 2015). Two *Cryptotympana* species are confirmed to have *Ophiocordyceps (Matsuura et al. 2018)* but we have only recently discovered that the Dog-Day cicadas also possess this fungal symbiont. Other cryptotympanine genera inhabit western North America including *Cacama,* for which evidence from this thesis suggests may have an *Ophiocordyceps* symbiont*,* and *Hadoa,* which does not. Sequencing of *Ophiocordyceps* symbionts of the Dog-Day cicadas presented here supports multiple domestications of genetically different *Ophiocordyceps* in the clade, including a possible instance of a free-living *Ophiocordyceps* replacing a symbiotic *Ophiocordyceps*. It is clear that this early stage of symbiont acquisition is dynamic rather than a single acquisition event in evolutionary history. The goal of my work was to decode the order of symbiont acquisition events within the Dog-Day cicadas and relatives using genomic data, determine the parasitic ancestors for each acquisition event, and begin to understand the genetic changes that occur during the very first phase of symbiont domestication. Additionally, I aimed to understand whether eukaryotic endosymbionts undergo the same patterns of genome evolution as bacterial symbionts. Some evidence from existing *Ophiocordyceps* symbionts in Hemiptera suggests that eukaryotic symbionts do not undergo genome size reduction, but rather gene loss is balanced by proliferation of mobile elements and introns (Vogel and Moran 2013; Fan et al. 2015). Genomic sequencing of symbionts of Dog-Day cicadas will allow us to address this gap in knowledge.

METHODS

Sampling

A total of 43 specimens were used for this project, including 9 species of *Neotibicen*, 8 *Megatibicen*, 4 *Cacama*, 2 *Hadoa*, and 2 *Raiateana*. Female specimens were prioritized since *Ophiocordyceps* DNA could potentially be amplified from the ovaries. Species were chosen with the intent of broadly sampling target genera, particularly *Neotibicen* and *Megatibicen*. The samples were collected by various collectors between 1997 and 2020 (Table 1) and stored in ethanol at -80 °C until use, except for the samples of *Raiateana kuruduadua, Neotibicen canicularis,* and *Megetibicen cultriformis*, which were pinned (see Table 1).

Sample processing

Cicada samples were dissected in phosphate buffered saline. Before each dissection, all dissection tools (dishes, forceps, pins, scissors) were sterilized in 10% bleach followed by a 95% ethanol wash, then placed in a UV crosslinker for two minutes. The abdominal tissue was separated into three sections: guts and fat bodies, bacteriomes, and ovaries (if present). Each was stored in 95% ethanol until DNA extraction. Before DNA extraction, the samples were homogenized using sterilized microcentrifuge pestles. Twelve samples (denoted with 'LN2' in Table 1) were frozen with liquid nitrogen immediately before being homogenized in order to improve lysis of fungal cells. For all samples, DNA extraction was performed using the DNeasy Blood and Tissue Kit with the following modifications: 1) during step 1a of the quick-start protocol, incubation was performed for 24 hours, and 2) during steps 8 and 9 (elution), 60 µl of Buffer AE was used instead of 200 µl. The three pinned specimens were extracted using a modification of the quick-start protocol. Abdomens were removed and holes were poked in the abdomen using sterilized forceps. The abdomen was then submerged in a mixture with 9 parts buffer ATL : 1 part proteinase K. After incubating overnight at 56°C, the lysate was pipetted out of the mixture and then treated normally through the protocol from step 2. DNA extracts were quantified using the Qubit dsDNA BR assay.

Target Gene Amplification

The bacterial *16S* rRNA gene was amplified via PCR using the 27F 1492R universal bacterial primer pair to screen for *Sulcia* in bacteriome tissue. Amplified DNA was sequenced in both directions with Sanger sequencing via Eurofins Genomics, and the two ends were assembled using de novo assembly tool within Geneious Prime v2019.1.3. The identity of each sequence was determined using BLASTN against the NCBI Nucleotide database (Johnson et al. 2008) if close or identical sequences were found. Next, bacteriome samples were screened for *Ophiocordyceps* by amplifying and sequencing the RNA polymerase II largest subunit (*RPB1*) gene. In samples that were positive for *Ophiocordyceps*, two additional genes were sequenced: *β*-tubulin (*β-tub*) and elongation factor 1 alpha (*EF-1α*). The 18S rRNA gene was also amplified in several samples but found to coamplify the insect homolog resulting in unusable chromatographs and was discontinued. A list of all primers used is available in Table 2.

Sanger data processing, alignment, and tree building

Sanger chromatograms for *RPB1*, *EF-1α*, and *β-tub* were analyzed in Geneious. First, beginning and ending regions were trimmed based on chromatogram quality using the 'Trim Ends' tool. Trimmed regions were manually inspected and modified to maximize the number of usable bases. Next, forward and reverse sequences were aligned, and bases with disagreements between the two sequences were manually inspected and either fixed or coded as the corresponding ambiguous base. The full gene sequences were extracted and aligned with each

other for each gene using the MAFFT Multiple Alignment v1.3.7 plugin (Katoh and Standley 2013) in Geneious. Each alignment was inspected individually to exclude the possibility of cross contamination between two otherwise unrelated samples with close or identical sequences. The individual gene alignments were read into SequenceMatrix v1.8 (Vaidya et al. 2011), which was used to create a concatenated alignment in PHYLIP format. A concatenation approach was chosen due to the low number of genes and since the taxon sampling for each gene was inconsistent. Phylogenetic trees were built using the concatenated alignments with RAxML v8.2.12 (Stamatakis 2014) on the CIPRES Science Gateway v3.3 (Miller et al. 2011). RAxML was run with the GTRCAT model and with each gene partitioned separately. Bootstraps were obtained with the rapid bootstrapping method (Stamatakis et al. 2008) with 100 iterations. Trees were visualized using a combination of FigTree v1.4.4 and the ggtree R package v3.0.4 (Yu et al. 2017).

For the Hypocreales tree (Figure 3), additional sequences were collected from 55 species of related fungi from Genbank, including 29 Ophiocordycipitaceae parasites, 19 Ophiocordycipitaceae symbionts, 5 Cordycipitaceae, and 2 Clavicipitaceae (see Table 3). For these taxa, data was collected for up to five genes: *RPB1*, *RPB2*, *EF-1α*, *18S*, and *28S* rRNA. Species were chosen randomly with the intention of getting a broad sampling of *Ophiocordyceps* and some more distant relatives throughout the order Hypocreales, with preference for species that have data for at least four of our five target genes. Sequences were aligned with our new data and manually trimmed in Geneious to minimize missing data at the ends of each gene. Tree building was performed as described in the previous paragraph.

Genus	Species	Specimen code (project)	Specimen code (internal)	District	Collection date	Lat	Long	Sex	Metagenome seq?	Collector	Description of location
Hadoa	townsendii	Vtib1	12.US.NM.TOC.09	Sierra Co., NM	15 Jun 2012	33.129083	-107.1177	f	n	Kathy Hill and Dave Marshall	Approx. 7.5 mi E of Rt. 179 on Rt. 51, E. of Truth or Consequences
Cacama	valvata	Vtib2	12.US.NM.SER.09	San Miguel Co., NM	15 Jun 2012	35.3438	-105.3099	m	n	Kathy Hill and Dave Marshall	CR B28A 4.4 mi. S. of Serafina
Hadoa	duryi	Vtib3	12.US.NM.SER.17	San Miguel Co., NM	15 Jun 2012	35.3438	-105.3099	f	n	Kathy Hill and Dave Marshall	CR B28A 4.4 mi. S. of Serafina
Megatibicen	auletes	Vtib4	N/A	Raleigh, NC	19 Sep 2011	UNKNOWN	UNKNOWN	f	n	Kathy Hill, Dave Marshall, W.H. Reynolds	
Megatibicen	dorsatus	Vtib5	N/A	Pana, IL	14 Aug 2014	UNKNOWN	UNKNOWN	m	n	Jake Readnour	Anderson Prarie Park off W 9th Street
Neotibicen	tibicen	Vtib6	12.US.NC.GRL.05	Wake Co., NC	12 Aug 2012	36.216533	-79.112966	f	n	W.H. Reynolds	Garner cr very nr. the Garner-s. Raleigh city limit either along US 50/70 and/or 401. Near Lake Benson.
Megatibicen	pronotalis	Vtib7	N/A	Kirksville, IL	17 Aug 2014	UNKNOWN	UNKNOWN	m	n	Jake Readnour	Willbarn Creek rec area 1175 N outside of Kirksville, IL
Neotibicen	linnei	Vtib8	12.US.NC.GRL.04	Wake Co., NC	31 Aug 2012	36.216533	-79.112966	f	n	W.H. Reynolds	Garner cr very nr. the Garner-s. Raleigh city limit either along US 50/70 and/or 401. Near Lake Benson.
Neotibicen	lyricen	Vtib9	03.US.VA.CHA.03	Albemarle Co., VA	17 Aug 2003	38.010133	-78.476133	m	n	Kathy Hill, Dave Marshall	Albemarle Co. E side of Charlottesille, SE Cnr I64 and Rt20 JCT
Neotibicen	davisi	Vtib10	N/A	Wake Co., NC	15 Aug 2009	UNKNOWN	UNKNOWN	f	n	W.H. Reynolds	Hwy NC 50/70 split (BP station, coll. @ lights)
Neotibicen	winnemanna	Vtib11	N/A	Blount Co., TN	12 Aug 1997	UNKNOWN	UNKNOWN	f	n	Allen Sanborn	Maryville
Megatibicen	resh	Vtib12	14.US.OK.EWP.02	Cleveland Co., OK	15 Aug 2014	35.351383	-97.567366	f	n	Robbie Sanders	Earlywine Park at SW 119th and May Ave in SW Oklahoma City
Megatibicen	figuratus	Vtib13	14.US.OK.RSC.03	Le Flore Co., OK	15 Aug 2014	34.63675	-94.606783	f	n	Robbie Sanders	Rattlesnake Creek Hwy 63, 8 mi W of Arkansas border
Neotibicen	auriferus	Vtib14	14.US.OK.NBR.02	Beaver Co., OK	15 Aug 2014	36.823133	-100.701966	f	n	Robbie Sanders	N1240 Rd, N side of Beaver River, W of Beaver River Wildlife Management Area
Megatibicen	dealbatus	Vtib15	14.US.OK.BDP.02	Beaver Co., OK	15 Aug 2014	36.839033	-100.509566	f	n	Robbie Sanders	Beaver Dunes Park Campground N of Beaver
Neotibicen	latifasciatus	Vtib16	11.US.FL.SSE.03	Marion Co., FL	23 Sep 2011	29.2084	-81.9459	f	n	K. Hill, D. Marshall	Rt. 40 7 mi. E of Silver Springs
Neotibicen	superbus	Vtib17	11.US.OK.HHE.02	Latimer Co., OK	17 Jun 2011	34.811366	-95.464683	m	n	Dave Marshall, Kathy Hill	Latimer Co. 3.2mi E of Pittsbug Co. line E of Harts Horn, ~20mi ESE of McAlester. Hwy 1/Rt63, ~5 mi ESE of US270
Megatibicen	harenosus	Vtib18	N/A	Chaves Co., NM	11-12 Sept 2009	33.4077	-103.8666	m	n	J. A. Cole, K. T. Eldredge, W. Chatfield-Taylor	Mescalero Dunes OHV Area
Cacama	valvata	Vtib19	12.US.NM.SER.08	San Miguel Co., NM	15 Jun 2012	35.3438	-105.3099	m	n	Kathy Hill and Dave Marshall	CR B28A 4.4 mi. S. of Serafina
Cacama	californica	Vtib20	N/A	Riverside Co, CA	26 Jun 2005	33.5856	-118.4570	m	n	Jeffrey A. Cole	Pinyon Flat Cpgr. 14 mi SW of Palm Desert on SR74
Cacama	californica	Vtib21	N/A	Riverside Co, CA	27 Jun 2005	33.5856	-118.4570	m	n	Jeffrey A. Cole	Pinyon Flat Cpgr. 14 mi SW of Palm Desert on SR75
Neotibicen	canicularis	Vtib22	N/A	Tolland Co., CT	14 Sep 2016	41.8155	-72.2663	f	n	Diler Haji	Storrs Campus
Raiateana	kuruduadua	Vtib23	03.FJ.WE.WEI.01	Vitu Levu	27 Jan 2003	-17.56	177.947	m	n	K. Hill, D. Marshall	S of Korovou
Megatibicen	cultriformis	Vtib24	07.US.AZ.PDL.03	Santa Cruz Co., AZ	02 Sep 2007	31.400483	-111.0911	f	n	Kathy Hill and Dave Marshall	SW edge of Lake Rena Blanca. Lower Thumb Rock picnic area
Raiateana	knowlesi	Vtib25	N/A	Namosi Province, Fiji	2017	-18.161033	-178.029983	f	n	Nunia Thomas-Moko	Viti Levu; Garrick Reserve
Raiateana	knowlesi	Vtib26	N/A	Namosi Province, Fiji	09 Oct 2017	-18.158317	-178.133117	f	n	Nunia Thomas-Moko	Viti Levu; Garrick Reserve
Raiateana	knowlesi	Vtib27	N/A	Namosi Province, Fiji	09 Oct 2017	-18.158317	-178.133117	f	n	Nunia Thomas-Moko	Viti Levu; Garrick Reserve
Raiateana	knowlesi	Vtib28	N/A	Namosi Province, Fiji	2017	UNKNOWN	UNKNOWN	f	n	Nunia Thomas-Moko	Viti Levu; Garrick Reserve
Cacama	moorei	Vtib29	N/A	Tucson, AZ	Jun 04 2019	32.3106	-110.8206	f	n	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	Sabino Canyon Recreation Area on trails near visitor center
Cacama	moorei	Vtib30	N/A	Tucson, AZ	Jun 04 2019	32.3106	-110.8206	f	n	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	Sabino Canyon Recreation Area on trails near visitor center
Cacama	moorei	Vtib31	N/A	Tucson, AZ	Jun 04 2019	32.2155	-110.7397	m	n	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	E. Tinnrod St off Freeman Rd
Neotibicen	linnei	LN2_1	12.US.NC.GRL.03	Wake Co., NC	31 Aug 2012	36.216533	-79.112966	m	n	W.H. Reynolds	Garner cr very nr. the Garner-s. Raleigh city limit either along US 50/70 and/or 401. Near Lake Benson.
Neotibicen	latifasciatus	LN2_2	11.US.FL.SSE.01	Marion Co., FL	23 Sep 2011	29.2084	-81.9459	m	y	K. Hill, D. Marshall	Rt. 40 7 mi. E of Silver Springs
Neotibicen	davisi	LN2_3	06.US.NC.BWC.01	Carteret Co., NC	4 Sep 2006	34.694	-76.826	m	y	Kathy Hill and Dave Marshall	W of Atlantic Beach
Neotibicen	superbus	LN2_4	N/A	Dimmit Co., TX	Jun 07 2019	28.5502	-99.7531	m	y	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	Picnic area on TX85 E of Carrizo Springs
Megatibicen	auletes	LN2_5	14.US.OK.RSC.02	Le Flore Co., OK	15 Aug 2014	34.63675	-94.606783	f	y	Robbie Sanders	Rattlesnake Creek Hwy 63, 8 mi W of Arkansas border
Megatibicen	dorsatus	LN2_6	14.US.OK.SRW.02	Bryan Co. OK	15 Aug 2014	34.002333	-96.403433	f	n	Robbie Sanders	Field SE of intersection of Radio Rd and W University Blvd. Durant OK
Cacama	californica	LN2_7	N/A	Riverside Co., CA	Jun 16 2020	33.5800	-116.45167	f	y	Jeffrey A. Cole	Ribbonwood Equestrian Campground, Santa Rosa Mountains
Cacama	crepitans	LN2_8	N/A	Orange Co., CA	Jul 22 2020	33.65194	-117.56806	m	y	Jeffrey A. Cole	Bellview Trail West, off Heritage Drive, Robison Ranch
Cacama	valvata	LN2_9	N/A	San Bernardino Co., CA	Jun 18 2020	35.03889	-115.39889	m	n	Jeffrey A. Cole	Wildhorse Canyon Road off Black Canyon Road, Mojave National Preserve
Cacama	moorei	LN2_10	N/A	San Bernardino Co., CA	Jun 19 2020	35.03889	-115.39889	m	n	Jeffrey A. Cole	Wildhorse Canyon Road off Black Canyon Road, Mojave National Preserve
Cacama	moorei	LN2_11	N/A	Tucson, AZ	Jun 04 2019	32.3106	-110.8206	m	n	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	Sabino Canyon Recreation Area on trails near visitor center
Cacama	moorei	LN2_12	N/A	Tucson, AZ	Jun 04 2019	32.3106	-110.8206	f	n	Eric RL Gordon, Allegra Bargnesi, Jason Vailionis	Sabino Canyon Recreation Area on trails near visitor center

Table 1 - Sample data for cicadas used in this study

Gene Name	Forward primer name	Forward primer Sequence	Reverse primer name	Reverse primer sequence	Annealing temperature used	Citation
16S	27F	AGAGTTTGATCCTGGCTCAG	1492R	GGTTACCTTGTTACGACTT	56°C	Heuer et al. 1997
18S	Hyp325F	GTTTGGGTAGTGGCCAAAC	Hyp760R	CCTGCCTGGAGCACTCT	52°C	Matsuura et al. 2018
B-tub	Ophi_Btub44448F	CGYGAGGAGTTCCYGACCG	Ophi_Btub5243R	CRTCCTGGTAYTGCTGGTACTC	60°C	Vanderpool, D. unpublished
EF1a	Ef1a-OphF1	CCGGCCACCGTGACTTCAT	Ef1a-Oph2218R 1	ATGACACCGACGGCRMCGGTYTG	62°C	Matsuura et al. 2018
RPB1	RPB1-OhpF1	GTCTWCCACCCGGGCTT	RPB1-OhpR1	GGCRATGTCGTTGTCCAT	60°C	Matsuura et al. 2018

Table 2 - Primer sets used in this study

					Accession number					
Family	Symb/parasite	Revised name	Genus	Species	18S	28S	EF1a	RPB1	RPB2	Btub
Clavicipitaceae	Parasite	Claviceps purpurea	Claviceps	purpurea	(No data)	(No data)	MT430207	AY489648	MT184193	(No data)
Clavicipitaceae	Parasite	Metarhizium album	Metarhizium	album	DQ522560	DQ518775	(No data)	DQ522398	DQ522452	(No data)
Cordycipitaceae	Parasite	Beauveria caledonica	Beauveria	caledonica	AY245650	MN523535	(No data)	EF469086	HQ880962	(No data)
Cordycipitaceae	Parasite	Cordyceps brongniartii	Cordyceps	brongniartii	AY282741	MH872067	(No data)	JN992350	HQ880926	(No data)
Cordycipitaceae	Parasite	Cordyceps militaris	Cordyceps	militaris	CP023322	MH878247	(No data)	JN992492	CP023324	(No data)
Cordycipitaceae	Parasite	Cordyceps takaomontana	Cordyceps	takaomontana	AB044631	MF416545	(No data)	MF416646	EF495079	(No data)
Cordycipitaceae	Parasite	Cordyceps tenuipes	Cordyceps	tenuipes	AB086205	MH878312	(No data)	MH885443	MH521877	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps agriotidis	Cordyceps	agriota	AY245651	DQ518754	(No data)	DQ522368	DQ522418	(No data)
Ophiocordycipitaceae	Parasite	Tolypocladium capitatum	Cordyceps	capitata	AB027318	U57086	(No data)	JN992471	DQ522421	(No data)
Ophiocordycipitaceae	Parasite	Tolypocladium inegoensis	Cordyceps	inegoensis	AB027322	DQ118741	(No data)	DQ127243	(No data)	(No data)
Ophiocordycipitaceae	Parasite	Tolypocladium ophioglossoides	Cordyceps	ophioglossoides	AY489691	U47827	(No data)	JN992470	AB968562	(No data)
Ophiocordycipitaceae	Parasite	Polycephalomyces paradoxa	Cordyceps	paradoxa	AB027323	JN941411	(No data)	JN992465	EF495092	(No data)
Ophiocordycipitaceae	Parasite	Hirsutella rhossiliensis	Hirsutella	rhossiliensis	KM652081	MH872887	(No data)	KM652046	(No data)	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps annulata	Ophiocordyceps	annulata	KJ878915	KJ878881	KJ878962	KJ878995	(No data)	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps araracuarensis	Ophiocordyceps	araracuarensis	KC610788	NG_060292	KC610738	KF658665	KC610716	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps arborescens	Ophiocordyceps	arborescens	NG_064858	NG_060238	(No data)	(No data)	AB968534	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps brunneirubra	Ophiocordyceps	brunneirubra	(No data)	MH753688	(No data)	MK751466	MK751468	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps clavata	Ophiocordyceps	clavata	JN941727	MH879601	(No data)	JN992461	AB968548	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps coccidiicola	Ophiocordyceps	coccidiicola	AB031195	AB968419	(No data)	(No data)	AB968545	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps formosana	Ophiocordyceps	formosana	KJ878908	(No data)	(No data)	KJ878988	KR052522	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps gracilis	Ophiocordyceps	gracilis	MH697671	EF468810	(No data)	MH697678	MH697682	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps heteropoda	Ophiocordyceps	heteropoda	AY489690	EF468812	AB968596	JN992451	AB968555	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps humbertii	Ophiocordyceps	humbertii	MK874748	MK875536	(No data)	MK863829	(No data)	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps kniphofioides	Ophiocordyceps	kniphofioides	MF416620	MK875538	(No data)	KX713712	KC610717	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps konnoana	Ophiocordyceps	konnoana	AB031192	(No data)	(No data)	EF468862	EF468916	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps longissima	Ophiocordyceps	longissima	AB968394	MG031219	AB968584	MG196183	AB968546	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps nutans	Ophiocordyceps	nutans	JN941712	MH668079	(No data)	JN992446	EF495090	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps prolifica	Ophiocordyceps	prolifica	JN941708	JN941434	(No data)	JN992442	AB972957	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps pseudorhizoidea	Ophiocordyceps	pseudorhizoidea	(No data)	(No data)	(No data)	MK751469	MK214089	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps rhizoidea	Ophiocordyceps	rhizoidea	EF468969	MH753674	(No data)	EF468873	EF468923	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps sinensis	Ophiocordyceps	sinensis	JX968028	KU239985	JX968018	MK984587	EF495094	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps sobolifera	Ophiocordyceps	sobolifera	EF468972	MH141338	AB968590	KJ879013	AB968551	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps tricentri	Ophiocordyceps	tricentri	AB968393	MH668081	(No data)	(No data)	AB968554	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps unilateralis	Ophiocordyceps	unilateralis	DQ522554	MK752812	(No data)	KX713730	DQ522436	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps yakusimensis	Ophiocordyceps	yakusimensis	AB044632	KJ878902	(No data)	KJ879018	KJ878953	(No data)
Ophiocordycipitaceae	Parasite	Tolypocladium cylindrosporum	Tolypocladium	cylindrosporum	MK984565	MK984577	(No data)	MK984584	MK984573	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cryptotympana atrata	Cryptotympana	atrata	(No data)	(No data)	(No data)	LC370819	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cryptotympana facialis	Cryptotympana	facialis	(No data)	(No data)	LC370821	LC370819	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Euterpnosia chibensis	Euterpnosia	chibensis	(No data)	(No data)	LC370893	LC370891	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Euterpnosia okinawana	Euterpnosia	okinawana	(No data)	(No data)	LC370899	LC370897	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Graptopsaltria bimaculata	Graptopsaltria	bimaculata	(No data)	(No data)	LC370864	LC370862	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Graptopsaltria nigrofuscata	Graptopsaltria	nigrofuscata	(No data)	(No data)	LC370849	LC370847	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Hamiltonaphis styraci	Hamiltonaphis	styraci	D55719	MF536295	(No data)	(No data)	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Hyalessa maculaticollis	Hyalessa	maculaticollis	(No data)	(No data)	LC370988	LC370986	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Meimuna iwasakii	Meimuna	iwasakii	(No data)	(No data)	LC370978	LC370976	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Meimuna kuroiwae	Meimuna	kuroiwae	(No data)	(No data)	LC370965	LC370963	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Meimuna opalifera	Meimuna	opalifera	(No data)	(No data)	LC371011	LC371009	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Meimuna oshimensis	Meimuna	oshimensis	(No data)	(No data)	LC370978	LC370943	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Mogannia minuta	Mogannia	minuta	(No data)	(No data)	LC371004	LC371002	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Nilaparvata lugens	Nilaparvata	lugens	AF267233	XR_005574003	(No data)	(No data)	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Pochonia suchlasporia	Pochonia	suchlasporia	KM096334	MH878400	(No data)	KJ398601	LN714696	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Sogatella furcifera	Sogatella	furcifera	JF773150	AF267237	(No data)	(No data)	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Tanna japonensis	Tanna	japonensis	(No data)	(No data)	LC370913	LC370911	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Terpnosia nigricosta	Terpnosia	nigricosta	(No data)	(No data)	LC370886	LC370884	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Terpnosia vacua	Terpnosia	vacua	(No data)	(No data)	LC370871	LC370869	(No data)	(No data)

Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen superbus	Neotibicen	superbus	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen davisi	Neotibicen	davisi	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen auriferus	Neotibicen	auriferus	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen tibicen	Neotibicen	tibicen	(No data)	(No data)	To be submitted	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Pagiphora annulata	Pagiphora	annulata	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Tettigettalna josei	Tettigettalna	josei	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Scieroptera sanaoensis	Scieroptera	sanaoensis	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen harenosus	Megatibicen	harenosus	(No data)	(No data)	To be submitted	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen pronotalis	Megatibicen	pronotalis	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen dorsatus	Megatibicen	dorsatus	(No data)	(No data)	To be submitted	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen cultriformis	Megatibicen	cultriformis	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen canicularis	Neotibicen	canicularis	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen figuratus	Megatibicen	figuratus	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen dealbatus	Megatibicen	dealbatus	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen latifasciatus	Neotibicen	latifasciatus	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen linnei	Neotibicen	linnei	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen lyricen	Neotibicen	lyricen	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Neotibicen winnemanna	Neotibicen	winnemanna	(No data)	(No data)	(No data)	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen auletes	Megatibicen	auletes	(No data)	(No data)	To be submitted	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Megatibicen resh	Megatibicen	resh	(No data)	(No data)	(No data)	To be submitted	(No data)	To be submitted
Ophiocordycipitaceae	Symbiont	(Symb. of) Raiateana kuruduadua	Raiateana	kuruduadua	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Huechys parvula	Huechys	parvula	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Parasite	Ophiocordyceps sp. Arizona	Ophiocordyceps	sp. Arizona	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cacama crepitans	Cacama	crepitans	(No data)	(No data)	To be submitted	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cacama moorei	Cacama	moorei	(No data)	(No data)	To be submitted	(No data)	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cacama californica	Cacama	californica	(No data)	(No data)	To be submitted	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Cacama valvata	Cacama	valvata	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Caledopsalta judithae	Caledopsalta	judithae	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Kikihia scutellaris	Kikihia	scutellaris	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Kikihia muta	Kikihia	muta	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Kikihia peninsularis	Kikihia	peninsularis	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)
Ophiocordycipitaceae	Symbiont	(Symb. of) Kikihia cauta	Kikihia	cauta	(No data)	(No data)	(No data)	To be submitted	(No data)	(No data)

Table 3 - Accession numbers for related fungi included in the expanded phylogeny. Sequences that were obtained in this study but not yet uploaded to GenBank are marked 'To be submitted'.

Metagenome sequencing, processing, and assembly

Based on the DNA quantifications and *RPB1* band intensity, five liquid-nitrogen-ground samples were chosen and sequenced to 200 million 150 bp paired-end reads, on an Illumina HiSeq X by Admera Health after library preparation with an insert size of 350 bp. The chosen samples were bacteriomes of: *Neotibicen latifasciatus, Neotibicen davisi, Neotibicen superbus, Megatibicen auletes, Cacama californica,* and *Cacama crepitans*. Processing of raw reads began with the clumpify.sh script from BBMap v37.41 (Bushnell 2014) with the dedupe parameter enabled, which removes duplicate sequences in the raw reads (likely to be optical or PCR duplicates). Next, Trimmomatic v0.36 (Bolger et al. 2014) was used to trim low quality bases and Illumina adaptors using the following parameters: ILLUMINACLIP:$ADAPTER_PATH:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:50. Paired-end reads were merged with bbmerge.sh from BBMap and the unmerged and unpaired reads were manually added to a single file with the rest of the merged reads. Assembly was performed with a special configuration of SPAdes 3.13.1 (Bankevich et al. 2012) capable of high kmer assembly with the following parameters: --only-assembler -k 21,33,55,77,99,127,151,189,229. The majority of contigs in the assemblies were expected to be of host (cicada) DNA. To filter the fungal contigs from them, each contig was aligned using BLASTN to the genome of a closely related *Ophiocordyceps* species cultured

from the cicada *Meimuna opalifera* (Matsuura et al. 2018). Contigs with strong matches (e-value < 1e-100) were considered to be of fungal origin in our samples. After selecting the fungal contigs, only the *Megatibicen auletes* sample had a relatively complete genome, and the rest of the analysis focused on this one sample. Completeness of the genome was assessed with BUSCO 5.0.0 (Simão et al. 2015) using the Hypocreales lineage set. BUSCO uses gene prediction and protein alignment software to search for conserved, single-copy orthologs in a genome assembly, and the proportion of these orthologs that are present is used as a proxy for genome completeness. Since the BLAST-based method of classifying *Ophiocordyceps* contigs could potentially miss novel sequences or parts of the genome that were lost in the reference organism, an alternative binning approach was also employed using MetaBat2 v2.12.1 (Kang et al. 2019). MetaBat2 output 24 bins from the *M. auletes* assembly and of those, bin11 perfectly matched the fungal contigs retrieved by the BLASTN method. The lack of novel sequences found in the *Ophiocordyceps* MetaBat2 bin suggests that any unidentified fungal contigs in our assembly must be less than 1500 bp (since that is the minimum sequence size that was specified for MetaBat2) if they exist.

Genome annotation

Annotation of the *M. auletes*-sourced *Ophiocordyceps* genome was performed using the MAKER v2.31.9 pipeline (Cantarel et al. 2008). Repeat masking was performed with RepeatMasker v4.1.0 (Smit et al. 2015), using RepeatModeler (Smit and Hubley 2015) to generate a novel repeat database for our organism, which was combined with fungal repeats from the Dfam database (Storer et al. 2021). A combination of protein and transcriptome data were used by MAKER to inform gene models. The full dataset included 9 proteomes: *Beauveria bassiana* PRJNA225503 (Xiao et al. 2012), *Claviceps purpurea* PRJEA76493 (Schardl et al. 2013), *Cordyceps militaris* PRJNA225510 (Zheng et al. 2011), *Fusarium graminearum* PRJNA235346 *(Gardiner et al. 2014)*, *Metarhizium album* PRJNA72731 (Hu et al. 2014), *Ophiocordyceps australis* SAMN07142923 (de Bekker et al. 2017), *Ophiocordyceps polyrhachis-furcata* PRJNA200756 *(Wichadakul et al. 2015)*, *Ophiocordyceps sinensis* PRJNA608258 (Shu et al. 2020), *Ophiocordyceps unilateralis* PRJNA280567 (de Bekker et al. 2017) and RNAseq data from the closely related *Ophiocordyceps* symbiont of *M. opalifera* SAMN08222404 (Matsuura et al. 2018). Running MAKER with the parameter correct_est_fusion=1, was crucial to prevent gene fusion caused by the transcriptome data, since the genome is so small and gene-dense that the UTRs of the transcripts caused the formation of false mega-proteins in the annotation. The reference proteomes were used as evidence for generating “hints” for the *ab initio* gene predictor SNAP v2013-02-16 (Korf 2004). Two additional gene predictors were used but were not trained on protein evidence: AUGUSTUS v3.3.3 (Stanke and Waack 2003), using the *Fusarium graminearum* model, and GeneMark-ES v4.68 (Ter-Hovhannisyan et al. 2008), which was self-trained on the genome alone using the --fungus option. Functional annotation was completed using BLASTP against the Swiss-Prot

database (UniProt Consortium 2021), and InterProScan (Jones et al. 2014) to assign functional identifications to predicted proteins.

RESULTS

Phylogenetics of "tibicen group" *Ophiocordyceps*

Phylogenetic reconstruction of the symbiotic *Ophiocordyceps* of Cryptotympanini indicated three independent symbiont clades (Figure 1). Group 1 (green) includes the symbionts of all sampled *Megatibicen* cicadas, but only half of the sampled *Neotibicen.* Group 2 (blue) includes the remaining *Neotibicen* symbionts, as well as symbionts of the genus *Cryptotympana* (sequenced in Matsuura et al. 2018). The split in *Neotibicen* symbiont identity corresponds with a phylogenetic split at the base of *Neotibicen*. Group 3 (orange) includes only symbionts of *Cacama*. Given that *Ophiocordyceps* symbionts must be vertically inherited once they are acquired by a cicada, the mismatch between cicada and *Ophiocordyceps* phylogenies indicates that multiple acquisitions of *Ophiocordyceps* must have occurred. Three permutations of symbiont domestication events could explain the observed phylogenetic mismatch, requiring at least four independent acquisitions of *Ophiocordyceps*, but our current data cannot indicate which is most likely (Figure 2).

In our tree of Hypocreales (Figure 3), most cicada symbionts fall in a single monophyletic clade, along with three described parasitic species: *O. longissima, O. yakusimensis,* and *O. sobolifera*. We hypothesize that these three species and close relatives that remain undescribed are potential sources of domesticated symbionts. However, the symbionts of *Kikihia* and *Caledodopsalta* cicadas, which occur in New Zealand and New Caledonia respectively, fall in a different clade more closely related to the cicada parasite *O. araracuarensis* along with other *Ophiocordyceps* species not known to infect cicadas, though this relationship is not well supported. Given that *O. araracuarensis* is only found in the Amazon of Colombia (Sanjuan et al. 2015) it is likely that a currently undescribed parasite was the progenitor of symbionts in these cicadas.

Genome sequencing of the *Megatibicen auletes* symbiont

The final *M. auletes* genome assembly contained 1,607,034 contigs (2 Gb), of which 838 were determined to be of *Ophiocordyceps* origin with a total size of 22,237,293 bp. The genome size is slightly smaller than other *Ophiocordyceps* fungi, both parasitic and symbiotic, though GC content is slightly higher (Table 4). In total, 91.5% of BUSCO orthologs were found complete in our assembly, with another 3% present but fragmented. Repeat sequences constituted 4.39% of the genome, less than any close relatives, and 6,790 proteins were predicted in the annotated genome, with 6,426 being over 100 amino acids long. The total number of proteins is

slightly lower than in other related symbiont genomes, but is subject to change once training of the gene predictors is complete.

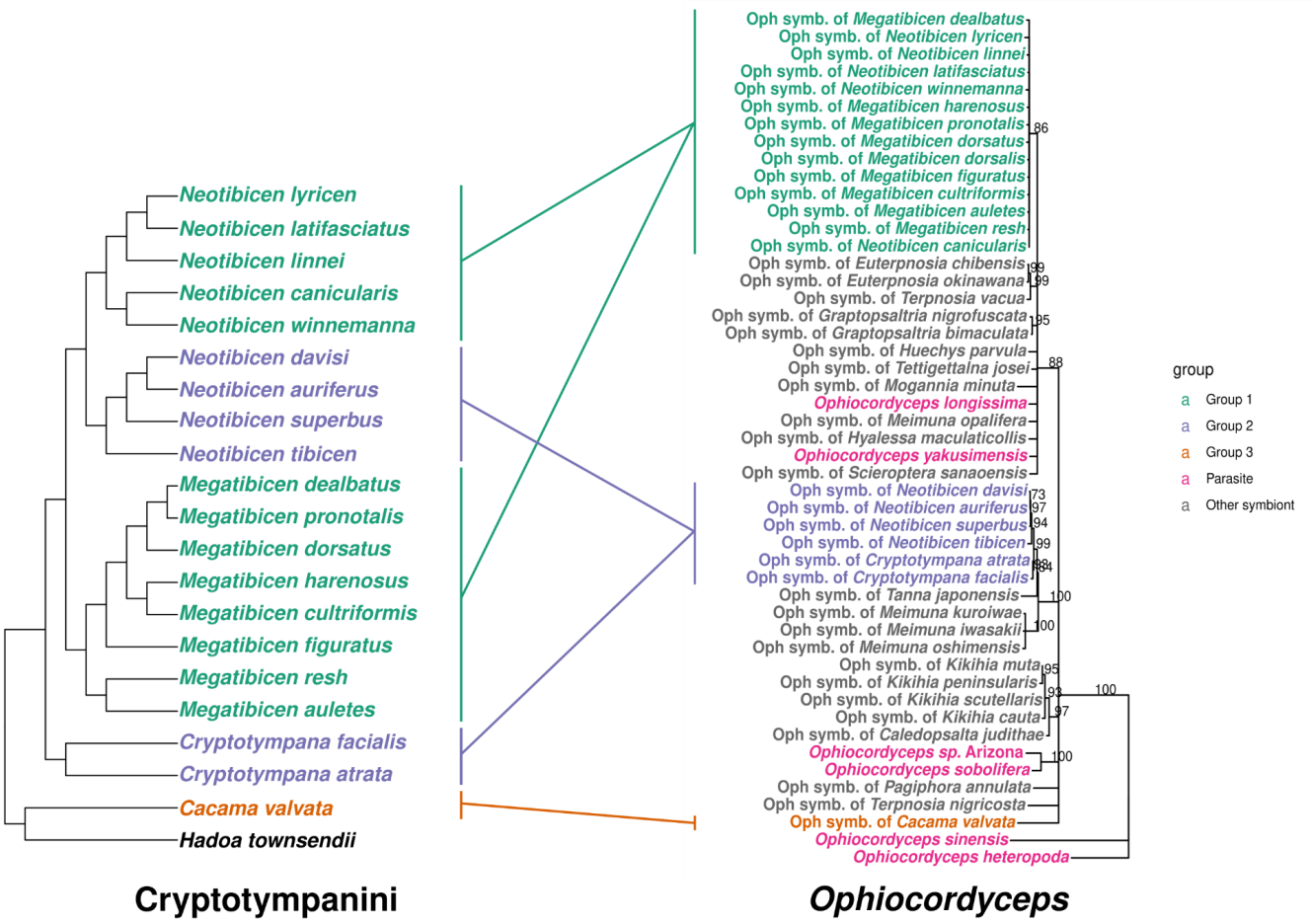

Figure 1 - Cophylogeny of cicadas (left) and *Ophiocordyceps* (right). The cicada phylogeny is modified from Hill et al. 2015 to include only those taxa used in this study and does not show bootstrap values, but each shown clade is well resolved in the original tree, except the blue Neotibicen clade which is present but poorly supported. They used *EF-1α* and mitochondrial *COI* sequences to generate the tree. Colors refer to monophyletic clades of symbionts that were derived from the sampled cicadas, except for the red group which refers to parasitic species of *Ophiocordyceps*. The colors on the hosts indicate that they possess a symbiont in the corresponding symbiont clade of that color. Grey *Ophiocordyceps* are symbionts of cicadas that are not in Cryptotympanini. The cicada *Hadoa townsendii* is a cryptotympanine cicada and was sampled for this study, but it lacked any *Ophiocordyceps* symbiont.

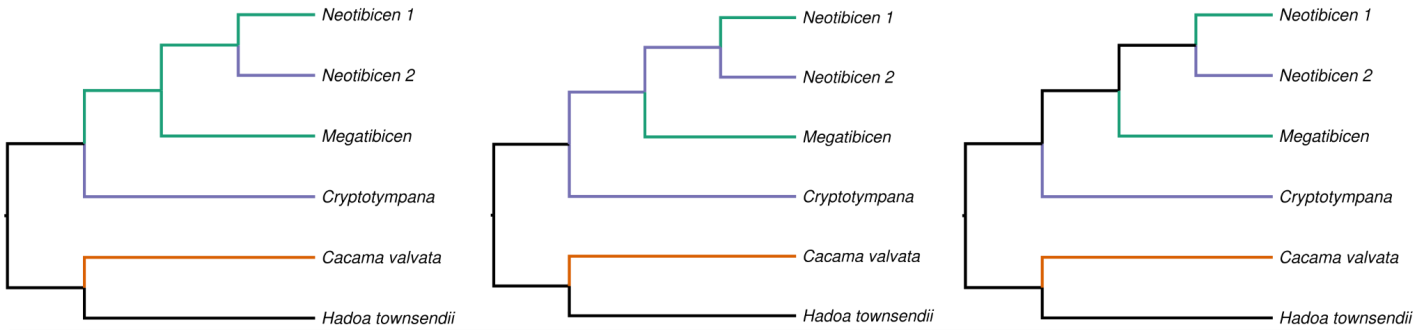

Figure 2 - Potential *Ophiocordyceps* domestication scenarios in Cryptotympanini. Colored branches indicate that *Ophiocordyceps* from that group are domesticated and inherited by cicadas along that branch, based on the color scheme defined in Figure 1. Assuming one *Ophiocordyceps* domestication at each color change on the tree, these scenarios suggest that *Ophiocordyceps* has been domesticated either four or five times in Cryptotympanini.

Fungus	(Symb. of) Megatibicen auletes	(Symb. of) Meimuna opalifera	(Symb. of) Cerataphis brasiliensis	(Symb. of) Nilaparvata lugens	Ophiocordyceps unilateralis	Cordyceps militaris
Symb/parasite	Symbiont	Symbiont	Symbiont	Symbiont	Parasite	Parasite
BUSCO score	91.50%	96.00%	89.80%	79.40%	94.10%	96.00%
Repeat %	4.39%	10.88%	7.18%	13.80%	7.61%	9.26%
Genome size (bp)	22,237,293	25,102,947	25,465,455	26,809,569	23,913,698	32,268,578
Number of scaffolds	838	31	1,181	578	2,536	32
Longest scaffold (bp)	270,671	5,998,088	233,904	873,274	167,401	6,263,243
N50 (bp)	48,194	4,912,384	39,063	310,669	27,475	4,551,492
GC content	0.62943	0.6019	0.56396	0.55288	0.55922	0.51418
# proteins	6,790	7,013	7,155	6,960	7,831	9,651
Accession number	To be deposited	PRJNA427071	PRJNA169168	PRJNA177649	PRJNA280567	PRJNA225510
Citation	N/A	Matsuura et al. 2018	Vogel et al. 2013	Xue et al. 2014	de Bekker et al. 2017	Zheng et al. 2011

Table 4 - Comparison of genome features between our genome (*Megatibicen auletes*) and other related genomes.

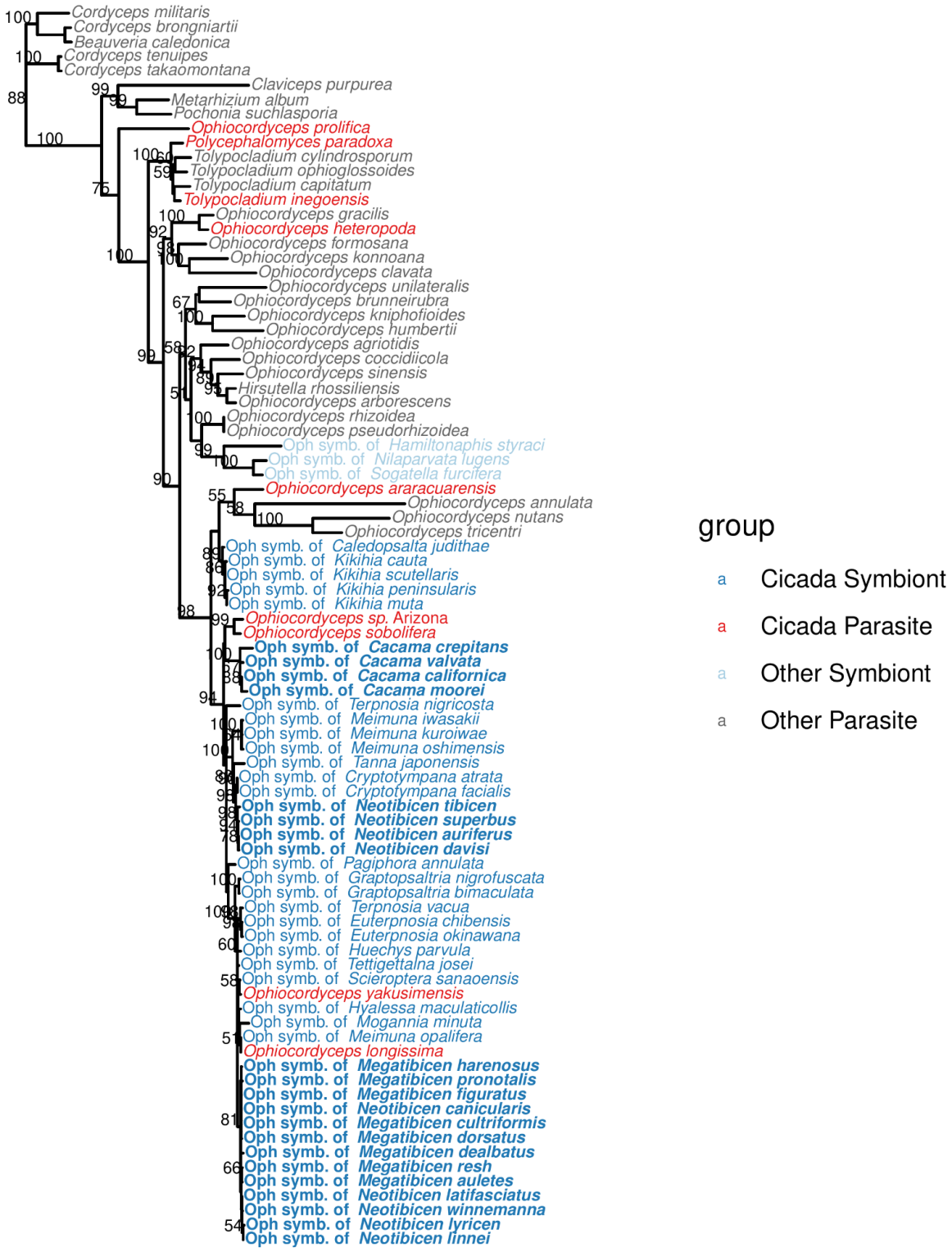

Figure 3 - Expanded RAxML phylogeny of fungi from the families Ophiocordycipitaceae, Cordycipitaceae, and Clavicipitaceae generated using five genes: *RPB1*, *RPB2*, *EF-1α*, *18S*, and *28S* rRNA. Species known to infect

cicadas as parasites are red, while beneficial symbionts are blue. Species sequenced as part of this study are in bold lettering. Nodes with less than 50% bootstrap support are not labeled

DISCUSSION

Phylogenetics of "tibicen group" *Ophiocordyceps*

We report evidence for multiple domestication events within the tribe Cryptotympanini, some potentially being quite recent (ca. 12 Ma to present; Hill et al. 2015) (Figs 1 & 2). Based on this result, along with observations of *Ophiocordyceps* in other phylogenetically independent groups of cicadas (Matsuura et al. 2018), cicada-*Ophiocordyceps* represents a case of parallel evolution which has occurred at least 10 times. In all known instances where *Ophiocordyceps* has been domesticated by cicadas, the former bacterial symbiont *Hodgkinia* has been lost. It is possible that the recurrent genome fragmentation of *Hodgkinia* is maladaptive, favoring the loss of *Hodgkinia* and domestication of a new symbiont. However, the reason why *Ophiocordyceps* of one particular clade is always the new symbiont, rather than some other fungus or microbe, is not clear. *Ophiocordyceps* symbionts have appeared in multiple families of Hemipteran insects and their adaptations to enter hosts and evade immunity likely help with their transition to mutualistic symbiosis. The specific mechanism that allows transition from parasite to mutualist is not known (i.e., is it a genomic change on the part of the host or symbiont, or just a change in gene expression from either? Is it the host exerting control over *Ophiocordyceps* inside its body or is it beneficial for *Ophiocordyceps* to keep its host alive?).

Based on our analysis (Fig. 3), it appears that the vast majority of cicada symbionts are derived from one monophyletic clade of *Ophiocordyceps* parasites. It is likely that there are undiscovered *Ophiocordyceps* parasites in this clade that are ancestors of the symbionts in Groups 1 and 2. If the symbionts in Groups 1 and 2 were each acquired from one of the described parasites present in our tree, it is expected that they would cluster with that parasite, or, if enough time has passed, diverge from one another. However they would not diverge from their source parasite in exactly the same way so as to produce clades like Groups 1 and 2 where symbionts from a non-monophyletic clade of cicadas form a monophyletic group of fungi. This could be explained if there is an unsampled fungal parasite which is genetically similar to either Group 1 or Group 2 and which was domesticated multiple times, giving the illusion of a single clade of symbiotic *Ophiocordyceps* which actually results from two domestications of genetically similar parasites. The use of only three conserved genes hinders our ability to differentiate between single monophyletic clades of cicada symbionts, or clades that include multiple domestications of the same parasite.

M. auletes genome and annotation

The newly sequenced genome and annotation of the *M. auletes* symbiont allow interesting insights into eukaryotic endosymbiont evolution by comparison to the closely related

symbiont of *Meimuna opalifera,* thought to be an extremely recently accommodated symbiont from a pathogenic ancestor. Compared to other *Ophiocordyceps* parasites and obligate endosymbionts, the *M. auletes* symbiont has the smallest genome, lowest repeat count, and lowest gene count (Table 4). These features could be the result of genome reduction, which is consistently observed in bacterial endosymbionts of Hemiptera. However, the features expected of a symbiont undergoing this characteristic genome reduction depend on the age of the association. The low repeat content is somewhat unexpected for a newly acquired endosymbiont. In the early stages of bacterial endosymbiont evolution, it is thought that repeat content will be high, as low population sizes and high genetic drift allow proliferation of mobile elements (Moran and Plague 2004). Gene loss occurs after the spread of these elements inactivates non-essential genes which are eventually lost. Further, eukaryotic genomes affected by small population size tend to undergo genome expansion rather than reduction, due to the buildup of repeat elements and introns (Lynch 2006). It has been hypothesized that the *Ophiocordyceps* symbionts of Hemiptera have maintained their genome sizes through a balance of the processes which cause genome expansion in eukaryotes with small population sizes and those that cause genome reduction in obligate endosymbionts (Vogel and Moran 2013; Fan et al. 2015). The small size, low repeat content, and low gene count of the *M. auletes* symbiont are not extreme, and could be consistent with this hypothesis. Without genomic sequences of recent parasitic ancestors of the symbiont, it is difficult to determine the age of its symbiotic association or the degree of genome size reduction, gene loss, and transposable element expansion. Based on phylogenetic analysis of cryptotympanine *Ophiocordyceps* (Figure 1), the *M. auletes* symbiont could be as old as the split between *Megatibicen*, *Neotibicen*, and the rest of the cryptotympanine cicadas. The other sequenced cicada symbiont from *Meimuna opalifera* was still culturable from a live cicada, and therefore likely to be a very recently acquired symbiont, while the *M. auletes* symbiont was unculturable using the same methods (unpublished result). Further analysis of the *M. auletes* symbiont will shed light on the evolutionary properties that have shaped its genome.

Conclusion

Symbiont replacement in cicadas is not a unique event but rather a parallel process affecting cicadas across the globe. Even in one tribe of cicadas, Cryptotympanini, there is evidence for at least four symbiont replacement events, and different lineages of symbionts can be found even within a single genus. There is some mechanism that allows *Ophiocordyceps* to continually transition from parasite to mutualist, even in more distant hemipteran insects, but it is not known if this mechanism is specific to *Ophiocordyceps* or if it could it happen with any specialized endoparasite. Many open questions still remain about this system; can cicadas with *Ophiocordyceps* symbionts still be affected by *Ophiocoryceps* parasites? Can different strains of *Ophiocordyceps* compete with each other as symbionts inside their host? How long is the transition from *Hodgkinia* to *Ophiocordyceps* and do the two ever coexist? Did other symbiont

replacement events involve parallel evolution like seen in cicadas or is it a result of the unique evolution of *Hodgkinia* or the unique life cycle of cicadas?

The completion of the *M. auletes* symbiont genome assembly will hopefully shed light on some of these questions, and will also help us understand the dynamics of genome evolution in eukaryotic endosymbionts. We plan to conduct more thorough analysis of this genome in the future, including investigations of evolutionary rate, gene loss, repeat content, introns, pseudogenes, mating type loci, gene duplication and rearrangements, and amino acid pathways. Though gene loss is relatively small in the symbiont genome sequenced compared to those of bacterial symbionts, we can examine the function of those genes and whether they have any impact on the former pathogenicity of the symbiont. We can also test whether other pathogenic genes are under relaxed purifying selection, accumulating mutations that may make them less functional than wild-type relatives. Specifically, the loss or extreme divergence of the mating type locus would allow us to infer that the symbiont has not undergone a sexual phase in many generations. As more *Ophiocordyceps* are sequenced at differing ages of symbiont evolution, the genome produced here will continue to be useful for understanding this unique and fascinating system.

Acknowledgements

Enormous thanks to my mentor Dr. Eric RL Gordon, who gave me oversight on literally every part of my research experience, taught me most of what I know, wrote half of my emails, forced me to like insects, and became one of my best friends throughout undergrad. I also want to thank my advisor, Dr. Chris Simon, for helping me and always pushing me to do more through deadlines and encouragement. I doubt I would have stayed on top of things without her help. Big shoutout to the other Simon lab members Mark Stukel, Allegra Bargnesi, Alexandra Porczak, Diler Haji, and Jefrin Thomas for always being there to discuss things and provide support, research related or not. Diler also taught me some R magic that I still use today.

Thanks to my University Scholar Committee members Rachel O'Neill, Jonathan Klassen, and Paul Lewis who gave feedback throughout my project, especially informing me of new methods and techniques that I was unfamiliar with.

I also want to acknowledge the many people who collected samples for this project - Jeffrey Cole, Dr. David Marshall, Kathy Hill, W.H. Reynolds, Robbie Sanders, Pam Rimer, Matt Kasson, Jake Readnour, Nunia Thomas-Moko, and Allen Sanborn, as well as Yu Matsuura who sent extracts from his *Ophiocordyceps* culture. This work could not have done without their efforts. Finally, I would like to thank Dr. Wu He, whom I consulted about the use of flow cytometry on this project.

Works Cited

Ainsworth G.C. 2008. Dictionary of the Fungi. CABI.

Ankrah N.Y.D., Chouaia B., Douglas A.E. 2018. The Cost of Metabolic Interactions in Symbioses between Insects and Bacteria with Reduced Genomes. MBio. 9.

Araújo J.P.M., Hughes D.P. 2019. Zombie-Ant Fungi Emerged from Non-manipulating, Beetle-Infecting Ancestors. Curr. Biol. 29:3735–3738.e2.

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Prjibelski A.D., Pyshkin A.V., Sirotkin A.V., Vyahhi N., Tesler G., Alekseyev M.A., Pevzner P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455–477.

de Bekker C., Ohm R.A., Evans H.C., Brachmann A., Hughes D.P. 2017. Ant-infecting Ophiocordyceps genomes reveal a high diversity of potential behavioral manipulation genes and a possible major role for enterotoxins. Sci. Rep. 7:1–13.

Bennett G.M., Moran N.A. 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a Phloem-feeding insect. Genome Biol. Evol. 5:1675–1688.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Bublitz D.C., Chadwick G.L., Magyar J.S., Sandoz K.M., Brooks D.M., Mesnage S., Ladinsky M.S., Garber A.I., Bjorkman P.J., Orphan V.J., McCutcheon J.P. 2019. Peptidoglycan Production by an Insect-Bacterial Mosaic. Cell. 179:703–712.e7.

Buchner P. 1965. Endosymbiosis of Animals with Plant Microorganisms. John Wiley and sons.

Bushnell B. 2014. BBMap: A fast, accurate, splice-aware aligner. .

Campbell M.A., Łukasik P., Simon C., McCutcheon J.P. 2017. Idiosyncratic Genome Degradation in a Bacterial Endosymbiont of Periodical Cicadas. Curr. Biol. 27:3568–3575.e3.

Campbell M.A., Van Leuven J.T., Meister R.C., Carey K.M., Simon C., McCutcheon J.P. 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. Proc. Natl. Acad. Sci. U. S. A. 112:10192–10199.

Cantarel B.L., Korf I., Robb S.M.C., Parra G., Ross E., Moore B., Holt C., Sánchez Alvarado A., Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18:188–196.

Fan H.-W., Noda H., Xie H.-Q., Suetsugu Y., Zhu Q.-H., Zhang C.-X. 2015. Genomic Analysis of an Ascomycete Fungus from the Rice Planthopper Reveals How It Adapts to an Endosymbiotic Lifestyle. Genome Biol. Evol. 7:2623–2634.

Gardiner D.M., Stiller J., Kazan K. 2014. Genome Sequence of Fusarium graminearum Isolate CS3005. Genome Announc. 2.

Gomez-Polo P., Ballinger M.J., Lalzar M., Malik A., Ben-Dov Y., Mozes-Daube N., Perlman S.J., Iasur-Kruh L., Chiel E. 2017. An exceptional family: Ophiocordyceps-allied fungus dominates the microbiome of soft scale insects (Hemiptera: Sternorrhyncha: Coccidae). Mol. Ecol. 26:5855–5868.

Haji D., Vailionis J., Stukel M., Gordon E., Lemmon E.M., Lemmon A.R., McCutcheon J.P., Simon C. 2021. Host-associated microbial diversity in New Zealand cicadas uncovers elevational structure and replacement of obligate bacterial endosymbionts by Ophiocordyceps fungal pathogens. bioRxiv.:2021.08.24.457591.

Hershberg R., Petrov D.A. 2010. Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet. 6:e1001115.

Heuer H., Krsek M., Baker P., Smalla K., Wellington E.M. 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. 63:3233–3241.

Hill K.B.R., Marshall D.C., Moulds M.S., Simon C. 2015. Molecular phylogenetics, diversification, and systematics of *Tibicen* Latreille 1825 and allied cicadas of the tribe Cryptotympanini, with three new genera and emphasis on species from the USA and Canada (Hemiptera: Auchenorrhyncha: Cicadidae). Zootaxa. 3985:219–251.

Hongoh Y., Ishikawa H. 2000. Evolutionary studies on uricases of fungal endosymbionts of aphids and planthoppers. J. Mol. Evol. 51:265–277.

Hughes D.P., Araújo J.P.M., Loreto R.G., Quevillon L., de Bekker C., Evans H.C. 2016. From So Simple a Beginning: The Evolution of Behavioral Manipulation by Fungi. In: Lovett B., St. Leger R.J., editors. Advances in Genetics. Academic Press. p. 437–469.

Husnik F., McCutcheon J.P. 2016. Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. Proc. Natl. Acad. Sci. U. S. A. 113:E5416–24.

Hu X., Xiao G., Zheng P., Shang Y., Su Y., Zhang X., Liu X., Zhan S., St. Leger R.J., Wang C. 2014. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. Proc. Natl. Acad. Sci. U. S. A. 111:16796–16801.

Johnson M., Zaretskaya I., Raytselis Y., Merezhuk Y., McGinnis S., Madden T.L. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36:W5–W9.

Jones P., Binns D., Chang H.-Y., Fraser M., Li W., McAnulla C., McWilliam H., Maslen J.,

Mitchell A., Nuka G., Pesseat S., Quinn A.F., Sangrador-Vegas A., Scheremetjew M., Yong S.-Y., Lopez R., Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30:1236–1240.

Kang D.D., Li F., Kirton E., Thomas A., Egan R., An H., Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 7:e7359.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Kobiałka M., Michalik A., Walczak M., Szklarzewicz T. 2018. Dual "Bacterial-Fungal" Symbiosis in Deltocephalinae Leafhoppers (Insecta, Hemiptera, Cicadomorpha: Cicadellidae). Microb. Ecol. 75:771–782.

Koga R., Bennett G.M., Cryan J.R., Moran N.A. 2013. Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. Environ. Microbiol. 15:2073–2081.

Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics. 5:59.

Lassalle F., Périan S., Bataillon T., Nesme X., Duret L., Daubin V. 2015. GC-Content evolution in bacterial genomes: the biased gene conversion hypothesis expands. PLoS Genet. 11:e1004941.

Łukasik P., Nazario K., Van Leuven J.T., Campbell M.A., Meyer M., Michalik A., Pessacq P., Simon C., Veloso C., McCutcheon J.P. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of cicadas. Proc. Natl. Acad. Sci. U. S. A. 115:E226–E235.

Lynch M. 2006. The origins of eukaryotic gene structure. Mol. Biol. Evol. 23:450–468.

Matsuura Y., Moriyama M., Łukasik P., Vanderpool D., Tanahashi M., Meng X.-Y., McCutcheon J.P., Fukatsu T. 2018. Recurrent symbiont recruitment from fungal parasites in cicadas. Proc. Natl. Acad. Sci. U. S. A. 115:E5970–E5979.

McCutcheon J.P., Boyd B.M., Dale C. 2019. The Life of an Insect Endosymbiont from the Cradle to the Grave. Curr. Biol. 29:R485–R495.

McCutcheon J.P., McDonald B.R., Moran N.A. 2009a. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. PLoS Genet. 5:e1000565.

McCutcheon J.P., McDonald B.R., Moran N.A. 2009b. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. Proc. Natl. Acad. Sci. U. S. A. 106:15394–15399.

McCutcheon J.P., Moran N.A. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol. Evol. 2:708–718.

McCutcheon J.P., Moran N.A. 2012. Extreme genome reduction in symbiotic bacteria. Nat. Rev. Microbiol. 10:13–26.

Miller M.A., Pfeiffer W., Schwartz T. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses. Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery.:1–8.

Mira A., Ochman H., Moran N.A. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. 17:589–596.

Moran N.A. 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. U. S. A. 93:2873–2878.

Moran N.A., Plague G.R. 2004. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. 14:627–633.

Moran N.A., Tran P., Gerardo N.M. 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. Appl. Environ. Microbiol. 71:8802–8810.

Müller H.J. 1940. Die Symbiose der Fulgoroiden (Homoptera Cicadina). Zoologica. 98:1–220.

Müller H.J. 1962. Neuere vorstellungen über verbreitung und phylogenie der endosymbiosen der zikaden. Zeitschrift für Morphologie und Ökologie der Tiere. 51:190–210.

Nishino T., Tanahashi M., Lin C.-P., Koga R., Fukatsu T. 2016. Fungal and bacterial endosymbionts of eared leafhoppers of the subfamily Ledrinae (Hemiptera: Cicadellidae). Appl. Entomol. Zool. . 51:465–477.

Noda H. 1977. Histological and histochemical observation of intracellular yeastlike Symbiotes in the fat body of the smaller brown planthopper, Laodelphax striatellus (Homoptera : Delphacidae). Appl. Entomol. Zool. . 12:134–141.

Ochman H., Davalos L.M. 2006. The nature and dynamics of bacterial genomes. Science. 311:1730–1733.

Podsiadło E., Michalik K., Michalik A., Szklarzewicz T. 2018. Yeast-like microorganisms in the scale insect Kermes quercus (Insecta, Hemiptera, Coccomorpha: Kermesidae). Newly acquired symbionts? Arthropod Struct. Dev. 47:56–63.

Sanborn A.F., Phillips P.K. 2013. Biogeography of the Cicadas (Hemiptera: Cicadidae) of North America, North of Mexico. Diversity . 5:166–239.

Sanjuan T.I., Franco-Molano A.E., Kepler R.M., Spatafora J.W., Tabima J., Vasco-Palacios A.M., Restrepo S. 2015. Five new species of entomopathogenic fungi from the Amazon and evolution of neotropical Ophiocordyceps. Fungal Biol. 119:901–916.

Schardl C.L., Young C.A., Hesse U., Amyotte S.G., Andreeva K., Calie P.J., Fleetwood D.J., Haws D.C., Moore N., Oeser B., Panaccione D.G., Schweri K.K., Voisey C.R., Farman M.L., Jaromczyk J.W., Roe B.A., O'Sullivan D.M., Scott B., Tudzynski P., An Z., Arnaoudova E.G., Bullock C.T., Charlton N.D., Chen L., Cox M., Dinkins R.D., Florea S., Glenn A.E., Gordon A., Güldener U., Harris D.R., Hollin W., Jaromczyk J., Johnson R.D., Khan A.K., Leistner E., Leuchtmann A., Li C., Liu J., Liu J., Liu M., Mace W., Machado C., Nagabhyru P., Pan J., Schmid J., Sugawara K., Steiner U., Takach J.E., Tanaka E., Webb J.S., Wilson E.V., Wiseman J.L., Yoshida R., Zeng Z. 2013. Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the clavicipitaceae reveals dynamics of alkaloid loci. PLoS Genet. 9:e1003323.

Shu R., Zhang J., Meng Q., Zhang H., Zhou G., Li M., Wu P., Zhao Y., Chen C., Qin Q. 2020. A New High-Quality Draft Genome Assembly of the Chinese Cordyceps Ophiocordyceps sinensis. Genome Biol. Evol. 12:1074–1079.

Silva F.J., Latorre A., Moya A. 2003. Why are the genomes of endosymbiotic bacteria so stable? Trends Genet. 19:176–180.

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. 31:3210–3212.

Simon C., Gordon E.R.L., Moulds M.S., Cole J.A., Haji D., Lemmon A.R., Lemmon E.M., Kortyna M., Nazario K., Wade E.J., Meister R.C., Goemans G., Chiswell S.M., Pessacq P., Veloso C., McCutcheon J.P., Łukasik P. 2019. Off-target capture data, endosymbiont genes and morphology reveal a relict lineage that is sister to all other singing cicadas. Biol. J. Linn. Soc. Lond. 128:865–886.

Smit A.F.A., Hubley R. 2015. RepeatModeler Open-1.0. Available from http://www.repeatmasker.org.

Smit A.F.A., Hubley R., Green P. 2015. RepeatMasker Open-4.0. Available from http://www.repeatmasker.org.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Stamatakis A., Hoover P., Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst. Biol. 57:758–771.

Stanke M., Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics. 19 Suppl 2:ii215–25.

Storer J., Hubley R., Rosen J., Wheeler T.J., Smit A.F. 2021. The Dfam community resource of transposable element families, sequence models, and genome annotations. Mob. DNA. 12:2.

Sudakaran S., Kost C., Kaltenpoth M. 2017. Symbiont Acquisition and Replacement as a Source

of Ecological Innovation. Trends Microbiol. 25:375–390.

Sung G.-H., Hywel-Jones N.L., Sung J.-M., Luangsa-Ard J.J., Shrestha B., Spatafora J.W. 2007. Phylogenetic classification of Cordyceps and the clavicipitaceous fungi. Stud. Mycol. 57:5–59.

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 Res. 18:1979–1990.

UniProt Consortium. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49:D480–D489.

Vaidya G., Lohman D.J., Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics. 27:171–180.

Van Leuven J.T., Meister R.C., Simon C., McCutcheon J.P. 2014. Sympatric speciation in a bacterial endosymbiont results in two genomes with the functionality of one. Cell. 158:1270–1280.

Vashishtha A., Sharama K.K., Lakhanpaul S. 2011. Co-existence, phylogeny and putative role of Wolbachia and yeast-like symbiont (YLS) in Kerria lacca (Kerr). Curr. Microbiol. 63:206–212.

Vogel K.J., Moran N.A. 2013. Functional and evolutionary analysis of the genome of an obligate fungal symbiont. Genome Biol. Evol. 5:891–904.

Wang D., Huang Z., Billen J., Zhang G., He H., Wei C. 2021. Structural diversity of symbionts and related cellular mechanisms underlying vertical symbiont transmission in cicadas. Environ. Microbiol.

Wang J.B., St. Leger R.J., Wang C. 2016. Chapter Three - Advances in Genomics of Entomopathogenic Fungi. In: Lovett B., St. Leger R.J., editors. Advances in Genetics. Academic Press. p. 67–105.

Wichadakul D., Kobmoo N., Ingsriswang S., Tangphatsornruang S., Chantasingh D., Luangsa-ard J.J., Eurwilaichitr L. 2015. Insights from the genome of Ophiocordyceps polyrhachis-furcata to pathogenicity and host specificity in insect fungi. BMC Genomics. 16:881.

Xiao G., Ying S.-H., Zheng P., Wang Z.-L., Zhang S., Xie X.-Q., Shang Y., St. Leger R.J., Zhao G.-P., Wang C., Feng M.-G. 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in Beauveria bassiana. Sci. Rep. 2:1–10.

Xue J., Zhou X., Zhang C.-X., Yu L.-L., Fan H.-W., Wang Z., Xu H.-J., Xi Y., Zhu Z.-R., Zhou W.-W., Pan P.-L., Li B.-L., Colbourne J.K., Noda H., Suetsugu Y., Kobayashi T., Zheng Y.,

Liu S., Zhang R., Liu Y., Luo Y.-D., Fang D.-M., Chen Y., Zhan D.-L., Lv X.-D., Cai Y., Wang Z.-B., Huang H.-J., Cheng R.-L., Zhang X.-C., Lou Y.-H., Yu B., Zhuo J.-C., Ye Y.-X., Zhang W.-Q., Shen Z.-C., Yang H.-M., Wang J., Wang J., Bao Y.-Y., Cheng J.-A. 2014. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. Genome Biol. 15:521.

Yu G., Smith D.K., Zhu H., Guan Y., Lam T.T.-Y. 2017. Ggtree : An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8:28–36.

Zheng P., Xia Y., Xiao G., Xiong C., Hu X., Zhang S., Zheng H., Huang Y., Zhou Y., Wang S., Zhao G.-P., Liu X., St Leger R.J., Wang C. 2011. Genome sequence of the insect pathogenic fungus Cordyceps militaris, a valued traditional Chinese medicine. Genome Biol. 12:R116.