DISTRIBUTION OF CICADAS IN ILLINOIS WITH A FOCUS ON THE NATURAL HISTORY, POPULATION GENETICS, AND CONSERVATION OF *MEGATIBICEN DORSATUS* (HEMIPTERA:CICADIDAE)

BY

CATHERINE E. DANA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Entomology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Dr. Sam W. Heads, Chair and Director of Research
Dr. Mark A. Davis
Assistant Professor Marianne Alleyne
Professor Andrew V. Suarez
Professor Emeritus James B. Whitfield

ABSTRACT

This dissertation examines an array of risk factors affecting Illinois cicadas (Hemiptera: Cicadidae) with considerations for successful management of species of greatest conservation need. To start, I broadly focus on the cicada biodiversity inhabiting Illinois' habitat mosaic. In my first chapter, I reviewed the distribution and ecology of the cicada species of Illinois. New county and state records, including DNA barcoding to confirm species identity are included. On clifftop prairies in southwestern Illinois multiple populations of both *Neotibicen auriferus* and *Beameria venosa* were discovered, both new state records.

For my second chapter, I investigated the impact of habitat fragmentation on cicadas in Illinois by looking at the distribution and genetic diversity of the prairie cicada *Megatibicen dorsatus*. This species was historically in high quality prairies, but is contemporarily restricted to small remnant prairies scattered throughout the state. A double digest restriction associated DNA sequencing (ddRADSeq) library was created with DNA from 452 *M. dorsatus* individuals from across Illinois. This method revealed gene flow among remaining populations along railroad rights-of-way, but less connectivity to higher quality prairies like Loda Cemetery Prairie Nature Preserve. Development of methods like RADSeq for cicada genetics and conservation can be applied worldwide to this cosmopolitan group of insects. With large-scale changes in the landscape occurring globally, it is important to assess the impact and long-term trends that might be influencing cicada populations and their geographic distribution.

Finally, in my third chapter I explored the role antimicrobial, superhydrophobic surfaces present on the wings of cicadas might play on colonization by microbes. I examined the external microbiome of *M. dorsatus* to identify the microbial community that they are exposed to in the prairie. I compared the microbiome of *M. dorsatus* to another more common species of cicada

that lives in urban areas throughout the state, *Neotibicen pruinosus pruinosus*. I divided individuals into separate parts to test if less microbial diversity and different community structure was present on the forewings compared to the rest of the body. To ensure that these differences were a product of exposure to the environment, I also compared the microbiome of early- and late-season collected individuals. I found that the legs were a likely source point of microbes. Cicadas were also carriers of several plant pathogens, although more evidence is needed to see if they play a role in spreading infections among plants. Due to my methodology not being specific to the ectobiome, I discovered gut endosymbionts of both species and added more information to the phylogeny of these different endosymbionts (i.e., *Candidatus* Hodgkinia, *Candidatus* Sulcia muelleri, and Yeast-Like Symbionts).

I hope that these findings will help provide information that can be used for the conservation of cicadas on the Illinois landscape through insight into the lands in need of protection, species diversity and distribution, necessary corridors for gene flow, ongoing and potential threats, and pathogens. Cicadas provide an important ecosystem service of nutrient cycling from below to aboveground. Among insects, cicadas are considerably large bodied and thus provide a substantial resource to insectivorous species. Given the sparse number of remnant prairies in Illinois (less than 0.07% of original prairies) and how specialized several cicada species are to this environment, their conservation is at a crucial time.

ACKNOWLEDGMENTS

Thank you to the fellow graduate students and technicians over the years who put in long and often frustrating hours chasing after cicadas -- Bailey Clancy, Shannon Carlson, Marysa Cristofaro, Vanessa Gabel, Jocelyn Hedlund, Tyler Hedlund, Kate Johnson, Masato Keeley, Grace Lewis, Sara Merkelz, Natalie Mills, Jonathan Tetlie, Tom Schmeelk, Ivan Shilov, Samantha Davis, Jacob Tamarri, and Sara Wilson. Thank you to my friends for being there through messenger, through the pandemic, sleepless nights, and random thoughts. Allen, Sarah, and Brendan, my unofficial committee members, I couldn't have gotten through these past two plus years without you. Christina Silliman-Swanson – don't think that I didn't notice your subtle and not so subtle help these past few years. You are a dear friend and hopefully we can get back to some SciCom adventures soon.

Mark Davis, I couldn't have asked for a better lab space than the Collaborative Conservation Genetics Lab. It is truly a wonderful collaborative environment to work in and provided me with so many wonderful memories over the years and I'm looking forward to continuing working with you. Thank you Sam Heads for your guidance and support over these many years and for giving me the opportunity to fall in love with cicadas. Thank you as well to the rest of my committee members, Jim, Andy, and M for providing me with fantastic edits, feedback, and a roadmap on how to get this dissertation into shape.

Jordyn Chace, I have loved your company in the lab these past years as you've really helped me keep my sanity with your positivity and good conversation. Aron Katz has been a wonderful friend and colleague – I've learned so much troubleshooting with you over the years. Thank you as well to Dylan O'Hearn for helping to work through the ddRADSeq benchwork and

pipeline. Roberto Cucalon Tamayo provided much appreciated code and tremendous guidance for much of the population genetics work.

Thank you to various funding agencies, including, but not limited to: Friends of Nachusa Grasslands, U.S. Fish and Wildlife Service, Ross Memorial Fund, and the Friends of Nachusa Grasslands, Prairie Biotic Research, and the University of Illinois Department of Entomology. Thank you to all the heritage biologists, land managers, and other staff from both the Department of Natural Resources and the Illinois Natural History Survey. I also want to give a special thank you to Joanne and Mike Fricke for their help on their wonderfully managed preserve.

The greatest thanks go to my dear husband, Jared Thomas, without whom none of this would have been possible. Thank you to Julian for putting up with hot summer days in the prairie and noisy cicadas flying around our house. Thank you to Corinne for keeping me company through the many late hours of writing.

TABLE OF CONTENTS

CHAPTER 1: A REVIEW OF THE DIVERSITY, ECOLOGY, DISTRIBUTION, AND CONSERVATION STATUS OF CICADAS (HEMIPTERA: CICADIDAE) IN ILLINOIS

Abstract

Less than 0.07% of original prairies remain on the Illinois landscape, presenting a unique challenge for the conservation of organisms exclusively utilizing this habitat. Many species of cicada (Hemiptera: Cicadidae) rely on remnant midwestern prairies. By spending most of their lives as nymphs feeding on xylem underground, cicadas provide critical ecosystem service of nutrient cycling; bringing nutrients from below ground to above ground predators. Despite their ecological importance, much is still unknown about the conservation status, impact of land management, host plant utilization, and life cycle duration of many species of North American cicadas. Here, I review the diversity, ecology, and distribution of this ecologically important group of insects in Illinois based on extensive fieldwork and studies of historic collections. Twenty-two species in eight genera are reported from Illinois including 15 new county records and two new records of species previously unknown from the state. High-resolution photographs and identification notes are provided along with distribution maps for all species. Acoustic information for several species is also included.

Keywords

Auchenorrhyncha, bioacoustics, barcoding, Cicadomorpha, Cicadoidea

Introduction

Cicadas (Hemiptera: Cicadidae) are an often large-bodied family of insects that inhabit various biomes across the globe on every continent apart from Antarctica. There are over 3,300 species of cicadas worldwide (Dmitriev 2022), each with a unique song that has been described as the "best single character for species recognition" (Myers 1929, page 217). The oft-reported loudest cicada on record is the African shrill thorntree cicada, *Brevisana brevis* (Walker, 1850), with a song that can exceed 106 decibels (Villet 1987). However, an Illinois species, *Megatibicen pronotalis* (Davis, 1938) rivals *B. brevis*, with alarm calls that can peak at 108.9 decibels (Sanborn and Phillips 1995).

Male cicadas produce airborne sounds used in conjunction with mating displays (calling or courtship songs), interference buzzes between conspecific males, and as an alarm when handled or disturbed by potential predators (Simmons and Young 1978; Cooley and Marshall 2001). These sounds are produced by the buckling of their abdominal tymbals, and sound production is amplified via abdominal air sacs (Pringle 1954). Sound is modulated and amplified by behavioral positioning of the abdomen and contraction of the tensor muscles (Myers 1929; Hennig et al. 1994), tympanal openings, and the opercula (coverings) present in some species (Young and Bennet-Clark 1995). In many species, once they are within a certain distance, the female will begin to "wing-flick" (Myers 1929). The male and female will then continue this pattern of singing and wing flick call and response before mating (Cooley 2001; Cooley and Marshall 2001; Sueur and Aubin 2004). A related family, the Tettigarctidae or "hairy cicadas," as well as leafhoppers (Cicadellidae), treehoppers (Membracidae), spittlebugs (Cercopidae), and planthoppers (Fulgomorpha) have rudimentary tymbal structures similar to cicadas (Pringle 1957), but instead are known to use substrate borne vibrations for communication (Ossiannilsson

1949; Shaw 1976; Hunt 1993; Claridge et al. 1999). Indeed, there is evidence that some species of cicadas use substrate borne vibrations in addition to acoustic male choruses (Stölting et al. 2002).

Cicadas spend the majority of their lives underground (Beamer 1928), emerging as adults after anywhere from one to 21 years depending on the species (Lloyd and Dybas 1966; Dietrich 2003; supplementary material of Marshall et. al 2018). While underground, cicada nymphs feed solely on low-nutrient xylem fluids from plant roots (Cheung and Marshall 1973; White and Strehl 1978). Despite feeding on xylem for their entire nymphal lifespan, species like the periodical cicadas have the highest recorded biomass per area for a terrestrial animal (Dybas and Davis 1962). Due to the lack of specific amino acids in xylem, cicadas have developed a relationship with an interdependent complex of bacterial endosymbionts to provide cicadas with the missing amino acids (Van Leuven et al. 2014; Campbell et al. 2015; Łukasik et al. 2018). Multiple species of endosymbionts can exist in the cells of a single cicada individual and are provisioned into cicada eggs for transmission to the next generation (Van Leuven et al. 2014). These endosymbionts have been speciating within cicadas for millions of years and have all undergone genome reduction (Campbell et al. 2017).

Nymphs are not necessarily dependent on a particular plant species, be it dicot or monocot (Lloyd and White 1987), although cicada species may have preferences (Callaham et al. 2000). Instead, oviposition suitability of the host plant may be the more limiting factor (White et al. 1982). Little is known about the host plant specificity in North American cicadas, likely due to their subterranean life cycle and their cryptic egg clusters. Many species of the genus *Neotibicen* Hill & Moulds, 2015 are hardwood generalists (Beamer 1928; Hill et al. 2015;

Sanborn and Phillips 2013), but more specific host plant information is not available for most species.

Much of the Illinois distributional data for cicadas are dated (73% of databased cicadas in the Illinois Natural History Survey Insect Collection were collected prior to 1970), and given the small size and fragmented nature of remnant Illinois prairies, it is possible that for some cicada species, these habitat patches fall below the minimal fragment size to support a population and may exceed achievable dispersal distances. In many species, negative effects of habitat loss not only result in a loss of biodiversity, but a loss of genetic diversity, population abundance, and distribution (Fahrig 2003). Several well-known entomologists have accessioned or identified cicada specimens at the Illinois Natural History Museum Insect Collection, including, Thomas E. Moore, Richard D. Alexander, Raymond H. Beamer, Milton W. Sanderson, William T. Davis, Herbert H. Ross, and Chris Simon (McElrath 2022). Some of the earliest collectors of cicadas at the INHS include Charles A. Hart and Stephen A. Forbes, with specimens from as early as 1877. The Chicago Field Museum also houses representative Illinois cicada specimens, with a large focus on the periodical cicada collection of Henry S. Dybas. It should be noted that one of the original "state collection[s] of insects" was lost in the Chicago Fire of 1871 (Chicago Academy of Science 1871). This collection included type specimens from the first state entomologist of Illinois, Benjamin D. Walsh, who wrote one of the original descriptions of the periodical cicadas (Walsh and Riley 1868).

Midwestern cicadas vary considerably in their habitat requirements, ranging from species that thrive in urban environments (e.g., the dog-day midwestern cicada, *Neotibicen pruinosus pruinosus* (Say, 1825)) to habitat specialists found only in grasslands with sandy soils (e.g., *Diceroprocta vitripennis* (Say, 1830)) (Sanborn and Phillips 2013). Few species are found only

in high-quality prairies (e.g., *Okanagana balli* (Davis, 1919) and *Cicadettana calliope calliope* (Walker, 1850)) and many of these species have very sparse records in Illinois (e.g., *O. balli*, *Okanagana rimosa rimosa* (Say, 1830), and *Beameria venosa* (Uhler, 1888)).

Given their unique habitat requirements, cicadas can be used as indicator species for habitat quality. Cicadas are an important component of ecosystems as a food source that connects aboveground and belowground organisms and energy flow. Due to their often-significant body mass and potential large-scale emergences (*i.e. Magicicada* spp.), cicadas can provide a substantial food source for insectivorous birds (Karban 1982; Rosenberg et al. 1982; Luukkonen 1987; Pons 2020), snakes (Surface 1906; Fitch 1982; Smith et al. 2019), mammals (Hahus and Smith 1990; Storm and Whitaker 2007; Krohne et al. 1991) including bears (Soper et al. 1976), and various other predators (Riley 1892; Marlatt 1907; Myers 1929). Periodical cicadas serve as a resource pulse in their emergence year with a long interval between adult generations (for a review on the community dynamics of pulsed resources see Ostfeld and Keesing 2000). This pulsed resource is difficult for specialist predators to predict and often has delayed impacts. However, in the years following an emergence there is increased wood accumulation (Koenig and Liebhold 2003) likely as a result of an increase in microbial biomass and nitrogen availability resulting from dead and decaying adult cicadas (Yang 2004); thus, cicadas can play an important role in large scale nutrient cycling in ecosystems (Andersen 1994; Callaham et al. 2000; Callaham et al. 2003; Smith et al. 2006).

Despite being a dominant species of the ecological soundscape during the summer months (Pijanowski et al. 2011), literature on the taxonomy, natural history, and conservation status of cicadas in North America is lacking. Here, I review the diversity of cicadas in Illinois through the study of historical museum specimens, and new fieldwork and collecting efforts. Full

active season acoustic monitoring and analysis were performed at several sites across the state and helped to document species. DNA barcoding was also performed to confirm species records.

Materials and Methods

To document the species of cicadas that inhabit Illinois, I utilized multiple methods including field collection of individuals, audio recording and processing, high quality photographs of both specimens from my studies and from museum collections (Table A.1), COI barcoding, and utilization of publicly available distribution data. A list of sites visited by me or field technicians assisting in this project can be found in Appendix A (Table A.2).

Field Collection

Cicadas were caught using 15-inch sweep nets then placed in mesh cages or on ice before being taken back to the lab. If legs were removed in the field (for later DNA analysis), scissors were flame sterilized using 80-100% ethanol. Legs were placed dry in 1.5 mL tubes on ice and then transported back to the lab for long term dry storage at -20°C at the Illinois Natural History Survey.

Acoustic Data

Calls were recorded using Wildlife Acoustics Song Meters 4.0 (SM4) (Wildlife Acoustics Inc., Maynard, MA) Firmware Version 2.2.0 that were placed in natural areas throughout the state (Appendix Table A.2) and timed to record during daylight hours. SM4 recorders were mounted several feet above the ground on fence posts or trees using bike cable locks. Audio recorders were set to a sample rate of 44100 Hz on dual channel built-in internal microphones.

HOBO® Pendant dataloggers MX2202 (Onset Computer Corp., Bourne, MA) were mounted alongside audio recorders and logged ambient temperature and light intensity. Audio files recorded on SM4 recorders were then visualized on Kaleidoscope Pro Version 5.4.2 (Wildlife Acoustics Inc., Maynard, MA).

Specimen Imaging

Photographs of cicada specimens were taken using a Canon EOS 5D Mark III camera with a Canon EF 100mm f/2.8L macro IS USM lens (Canon Inc., Tokyo, Japan) mounted to a StackShot Automated Focus Stacking Macro Rail (Cognisys Inc., Traverse City, MI) on a copy stand. Resulting images were combined using the focus stacking program Helicon Focus version 7.6.6 (Helicon Soft, Kharkiv, Ukraine). Image background was removed, lighting corrected, and scale bars were added in Adobe Photoshop CC version 19.0 (Adobe Systems Inc., San Jose, CA).

DNA Extraction, PCR, and Sanger Sequencing

DNA was extracted from cicada legs using QIAGEN DNEasy® Blood and Tissue Kits (QIAGEN Inc., Germantown, MD). Several modifications were made to the QIAGEN protocol, including using sterile plastic pellet pestles (Thermo Fisher Scientific, Waltham, MA) to grind muscle tissue out of the chitinous exoskeleton of cicada legs as an additional step to maximize the amount of DNA extracted. Samples were incubated in Proteinase K at 56°C for 18 – 24 hours with additional vortexing during this period. Another additional step performed for maximum DNA yield was overnight DNA precipitation after the addition of chilled 100% ethanol. DNA was eluted from QIAGEN columns in 80 to 100 µl of warmed AE buffer. After elution DNA concentration was measured using the High Sensitivity dsDNA Invitrogen™ Qubit™ 3

Fluorometer system (Invitrogen, Thermo Fisher Scientific, Waltham, MA). In order to amplify Cytochrome c oxidase I (COI) region of the mitochondria for barcoding, I used modified methodology from Hill et al. (2015) to amplify the full region in two parts (Figure 1.1) using 2X GoTaq® DNA Polymerase Master Mix (Promega Corp., Madison, WI) and primer pairs found in Table 1.2. PCR products were Sanger sequenced by Eurofins Genomics using their "crude PCR products" PrePaid plate service (Eurofins Genomics, Louisville, KY) or by the Sanger Core Facility at the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign (Table 2). COI sequences were assembled in Geneious Prime® 2021.2.2 (https://www.geneious.com).

Distribution Data Sources and Maps

Localities listed in Table 1.1 and used for state county maps are based off several sources, including: (1) Personal specimens collected; (2) Illinois Natural History Survey (INHS) Insect Collection records and specimens; (3) Sanborn and Phillips (2013) locations; (4) bugguide.net with individual records verified; (5) Global Biodiversity Information Facility (GBIF) records which include iNaturalist citizen science records verified individually by image; and (6) additional records in the literature (Davis 1925; Cooley et al. 2009; Cooley et al. 2013; Liebhold et al. 2013; Tumlison 2013; Lee 2016). In cases where a nymph, teneral individual, or unclear photos were used for either bugguide.net or iNaturalist, the record was discarded. Low map resolution resulted in uncertainty surrounding county records from Sanborn and Phillips (2013) which required county and state records to be confirmed from other sources when possible. The Sanborn and Phillips (2013) map was overlaid with lowered opacity over a county

map in Adobe Illustrator CC Version 22.0.0 (Adobe Systems Incorporated, San Jose, CA) and counties were chosen based on the center of the circles provided.

Fine-scale, county-level maps were created in ArcGIS Pro 2.8.0 (Environmental Systems Research Institute or “Esri”, Redlands, CA) using the World Hillshade, World Terrain Base, and World Terrain Reference basemaps. "USA States" layer package from Esri_dm map data downloaded via ArcGIS Online along with county lines feature layer from IDOTAdmin. LiDAR derivatives data was acquired from the Illinois Geospatial Data Clearinghouse website (https://clearinghouse.isgs.illinois.edu/data/elevation/illinois-height-modernization-ilhmp) for Madison and Monroe counties. This data was added over the top of the basemap layers. Monroe and Madison counties shape outlines were selected, and the LiDAR overlays were clipped to the selected shapefiles followed by stroke and layer shading adjustment for ease of visibility when printed. Microsoft Excel spreadsheet data for GPS coordinates were imported into ArcGIS Pro to plot X and Y point values over the top of the LiDAR layers. Species of Illinois

Species are grouped below tribal classification. Specified taxonomy follows that of Marshall et al. (2018) based on a phylogeny reconstructed using both nuclear and mitochondrial genes as well as morphological characters. Through the course of this project, I documented 22 species of cicadas within the state of Illinois, including two new state records.

Periodical cicadas (Tribe Lamotialnini)

Periodical cicadas emerge early in the season, during the months of May and June, dependent on the region (Krohne et al. 1991; Miller 1997; Yang 2006). Their bodies are primarily black in color and can be identified to species based on the presence of orange accents on the ventral surface of the abdomen (presence / absence) and orange marking between the eye

and the wing attachment point (presence / absence). Most live specimens have red eyes, but some are more pink or brown in coloration and fade in preserved specimens. According to maps generated by Sanborn and Phillips (2013), data collected by Cooley (2021), and specimens in the Illinois Natural History Survey (INHS) Insect Collection, all seven species of periodical cicadas (*Magicicada* Davis, 1925) occur in Illinois. This includes both the 13-year periodical cicadas, *Magicicada neotredecim* Marshall & Cooley, 2000, *Ma. tredecassini* Alexander & Moore, 1962, *Ma. tredecula* Alexander & Moore, 1962, and *Ma. tredecim* (Walsh & Riley, 1868), as well as the 17-year periodical cicadas, *Ma. cassinii* (Fisher, 1852) (Figure 1.2), *Ma. septendecim* (Linnaeus, 1758) (Figure 1.3), and *Ma. septendecula* Alexander & Moore, 1962 (Figure 1.4). Periodical cicada broods are both geographically (largely parapatric) and temporally isolated from each other (Marlatt 1907; Simon 1988). Allochronic speciation has led to the formation of three species groups, or "cognate" species – e.g., *Ma. cassinii* is genetically similar and often indistinguishable morphologically and behaviorally from *Ma. tredecassini* (Marshall and Cooley 2000; Cooley et al. 2001; Sota et al. 2013; Simon et al. 2022).

The 17-year "Northern Illinois Brood" XIII emerge throughout the Chicago region, west into eastern Iowa, and into southern Wisconsin (Figure 1.5). This brood has been noted to emerge in significant numbers off-cycle, four years prior to the main emergence year (Dybas 1969; Cooley et al. 2016). It is possible that this brood may be splitting into two separate broods as off-cycle individuals ("stragglers") were observed mating and ovipositing in 2020 (Dana *pers. obs.*), four years before the 2024 emergence (Figure 1.6). Marlatt (1923) mapped Brood XIII as extending into Michigan. Smaller numbers of off-cycle individuals were observed in 2020 by CED at Jubilee College State Park in Peoria County, more consistent with the numbers expected of typical "stragglers." This brood seem to have contracted in range in more recent studies

(Simon 1988; Moore 2016). Other range contractions have eventually resulted in extinction (Marlatt 1923; Manter 1974) or near extinction (Cooley et al. 2004; Gilbert and Klass 2006; Dana *pers. obs.* 2018).

Another 17-year brood, the "Great Eastern Brood" X (Figure 1.5) is one of the largest *Magicicada* broods (geographically) that extends from far eastern Illinois (near Danville south to Marshall), south to northern Georgia and northeast as far as Long Island (Cooley et al. 2009; Dana, pers. obs. 2021). Like Brood XIII mentioned previously, a significant number of individuals were observed to emerge early in 2017 in Indianapolis, IN and Bloomington, IN, including individuals infected with the *Massospora* fungal pathogen (Figure 1.7; Dana *pers. obs.*).

The 17-year Brood III (Figure 1.5), known as the "Iowan Brood", is found in the western portion of Illinois and its distribution falls mostly on the west side of the Illinois River. Outside of Illinois, it stretches across both Iowa and Missouri (Cooley et al. 2013). A map created by Stannard (1975) showed a separate population of Brood III in central Illinois (Champaign, DeWitt, and Piatt counties). However, later studies (Lloyd et al. 1983; Cooley et al. 2013) argue that this population is a disjunct population of XXIII that co-emerged during the same year as Brood III in 1963 (when Stannard 1975 observed them).

During this study, no 13-year broods emerged in Illinois. However, the 13-year "Great Southern Brood" XIX (Figure 1.5) is due to emerge in the southern half of Illinois in 2024, the same year as the next emergence of the 17-year Brood XIII. While these two broods largely occupy different regions in Illinois, it is possible that individuals of the 13- and the 17-year broods may encounter each other given their large geographic range. Marshall and Cooley (2000) described a new species, *Magicicada neotredecim,* that has increased black and orange

striping on the ventral surface of the abdomen compared to *Ma. tredecim*. *Ma. neotredecim* occurs primarily in Southern Illinois and Missouri but has a smaller range in several other states (Table 1.1; Sanborn and Phillips 2013). The more salient character difference is that the male call has a higher pitch than *Ma. tredecim* and this difference is greater in areas (like Illinois) where the two species are sympatric (Marshall and Cooley 2000).

The "Mississippi Valley" Brood XXIII is another 13-year brood that occurs from southern Illinois down into Mississippi and Arkansas (Figure 1.5). This brood is scheduled to next emerge in Illinois in 2028. Stannard's map (1975) of Illinois broods showed a disjunct population of Brood III in Piatt, Champaign, and DeWitt counties, but this disjunct population belongs to Brood XXIII (Cooley et al. 2013). The confusion in mapping was a result of Brood III and Brood XXIII co-emerging in 1963; subsequent maps have since been revised (Kritsky and Meyer 1976; Lloyd et al. 1983).

Dog-day or annual cicadas (Tribe Cryptotympanini)

"Dog-day" cicadas are so-called due to their tendency to call during the hottest days of summer. In Tribe Cryptotympanini, this includes species in two genera in Illinois: *Megatibicen* Sanborn & Heath, 2016 and *Neotibicen* Hill et al. 2015. The life cycle of dog-day cicadas is commonly reported to be between two and six years (Cole 1954; Dietrich 2003). However, this has not been confirmed for any species of *Neotibicen* or *Megatibicen* (Beamer 1928; Heath 1978).

Of the typically larger *Megatibicen*, Illinois is known to be home to the following: *Me. auletes* (Germar, 1834) (Figure 1.8), *Me. dorsatus* (Say, 1825) (Figure 1.9), and *Me. pronotalis walkeri* Metcalf, 1955 (Figure 1.10). The ecological niches of these species are substantially

different. *M. auletes* is sometimes known as the "Northern dusk-singing cicada" or "giant oak cicada" as a result of its association with oak trees. This species is on average the largest of the *Megatibicen* species – primarily distinguishable by its size, pruinose underbelly, and marked pruinosity on the final three tergites (Davis 1922). *Me. auletes* has a patchy distribution throughout Illinois (Figure 1.11).

Sometimes called the "grand prairie cicada", *Me. dorsatus* (Figure 1.9) can be found in remnant or high-quality prairies, like cemetery prairies, nature preserves, or right-of-way prairies (Figure 1.12). Illinois is the only state east of the Mississippi where this species can be found, otherwise it has a more central United States distribution, stretching from Mexico to Colorado and into southern South Dakota (Table 1.1). A subspecies that calls later in the season, *Me. pronotalis walkeri* (Figure 1.10) or "Walker's cicada" has a wide central US range and appears to be associated with riparian plants (like willow, cottonwood, and birch) (Beamer 1925; Beamer 1928, Dana *pers. obs.*). It can be heard calling in the late summer and has been found emerging in central and southern Illinois in late July and early August (Figure 1.13, Dana *pers. obs.*).

Six species of *Neotibicen* occur in Illinois, including two distinct subspecies (Table 1.1). One of the most commonly encountered throughout the state is *Neotibicen pruinosus pruinosus* (Figure 1.14), the "scissor grinder cicada" or the "silver bellied cicada" (however, this common name is used for several species) and is known for its distinctive dusk chorus. It is well established in city parks and more developed areas throughout Illinois. A very similar species in appearance to *Neot. p. pruinosus, Neot. linnei* (Smith & Grossbeck, 1907) (Figure 1.14) has a distinctive call compared to *Neot. p. pruinosus* – much quicker in its repetition of the echeme, and also typically shorter in duration (not counting the wind up and cool down, lasting approximately 10 seconds). *Neot. linnei* is distinguished morphologically from *Neot. p.*

pruinosus by the sharp bend in the costal line of the leading edge of the forewing (Figure 1.14). There are several other characters that can help distinguish these species (Lee 2016), but these characters can vary throughout their range and male song is the most reliable character (Dana *pers. obs.*). *Neot. canicularis* (Harris, 1841) (Figure 1.14) is also very similar in appearance to *Neot. p. pruinosus* and *Neot. linnei* but is much smaller in size and has a distinct call from the other species. Rather than consisting of separate echemes, it instead has a long metallic whine. It is more common in the northern part of the state and its range stretches up into Canada (Table 1.1).

Neot. l. lyricen (De Geer, 1773) (Figure 1.15) is often referred to as the lyric cicada and can be identified by the distinctive black pronotal band which is typically green in other *Neotibicen* species. Some citizen science portals list *Neot. lyricen engelhardti* (Davis, 1910) as being present in Illinois, but I have not been able to confirm these records and the primary subspecies found in Illinois is *Neot. l. lyricen*. *Neot. tibicen tibicen* (Linnaeus, 1758) (Figure 1.16) is sometimes called the "swamp cicada" or "morning cicada." It can often be found in more riparian areas and at first glance can be confused with *Neot. l. lyricen* – it also possesses a black pronotal band and *Neot. l. lyricen* can have more dark coloration (much darker even in *Neot. lyricen engelhardti*). However, *Neot. t. tibicen* is ventrally covered in white pruinosus and has a more humpbacked appearance when viewed laterally. *Neot. l. lyricen*, in contrast, has a black stripe running down the central portion of the ventral surface of the abdomen.

Other species

In addition to those listed previously, Illinois is home to cicada species in four additional distinct tribes (Cicadettini, Fidicinini, Leptosaltrini, and Tibicinini). *Cicadettana calliope*

calliope (Tribe Cicadettini) (Figure 1.17) is a small, grass associated cicada found primarily in high quality or remnant prairies (Figure 1.18). The call of this subspecies is high in pitch (approximately 17 – 21 kHz), which can be just out of the range of hearing in older humans. Marshall and Hill (2017) removed this subspecies from the genus *Cicadetta* due to morphological and molecular characters that separate North American species from those present in Europe.

Diceroprocta vitripennis (Tribe Fidicinini) (Figure 1.19) is sometimes referred to as the "green winged cicada" and is a species known for being associated with scrub habitat. Although unpublished, acoustic surveys by Carl Strang place *D. vitripennis* in both Will and Kankakee counties in regions with woody plants and sandy soils (Strang 2022). There is a consistent annual population of *D. vitripennis* at Henry Allan Gleason Nature Preserve and nearby Sand Ridge State Forest (Dana *pers. obs.*), both sites known for sandy soils. Unlike *Neotibicen* spp. and many other groups of cicadas, this species lacks the "smokey Z", or infuscated crossveins, on its forewings.

Neocicada hieroglyphica hieroglyphica (Say, 1830) (Tribe Leptosaltriini) (Figure 1.20) is a species that I have been unable to capture in Illinois – although I have heard males calling frequently and on an annual basis from high in the trees at Fults Hill Prairie Nature Preserve in Monroe County during the months of June and July. Unlike the lack of infuscation mentioned in *D. vitripennis*, *Neoc. h. hieroglyphica* has additional infuscation beyond the "smokey-Z" that make this species easily identifiable when compared to other North American species.

Okanagana balli (Tribe Tibicinini) (Figure 1.21) is associated with high quality silt loam prairies (Betz and Lamp 1989; IWAP 2015) and males have been frequently observed from the base of plants rather than the highest point (Dana *pers. obs.*). There are at least two prairies in

Illinois where this species still occurs, though its cryptic habits, short emergence time (late June to early July), and easily startled nature make it difficult to locate (Figure 1.22). Species of *Okanagana* have what is considered to be a "proto-periodical" life cycle (Soper et al. 1976; Williams and Simon 1995), where individuals live longer than one-year and populations have greater abundance in some years in comparison to others. The last known record of *Okanagana rimosa rimosa* (Figure 1.23) in Illinois was prior to 1962 and if this subspecies still occurs in Illinois, based on historical records, it would likely be found in one of the more northern counties of Illinois (Figure 1.24). Although seemingly rare in Illinois, *O. r. rimosa* has a large range across the United States (Table 1.1).

New Records

New State Records

Beameria venosa (Uhler, 1888) (Figure 1.25)

Distribution in Illinois: Monroe County (Figure 1.26).

Voucher specimens:

INHS Insect Collection 382,292 – 382,295. 3♂1♀, USA, Illinois, "Fults Hill Prairie". 0°N, 0°W. 10-vii-1985. No collector information. Det. C.E. Dana 2021.

INHS 557,546 – 557,549. 3♂1♀, USA, Illinois, Monroe Co., Prairie du Rocher, "Fults Prairie", 38.155°, -90.187°. 14-vii-2007. R. Rakitov.

INHS 1,001,451 and 1,001,452 [Dana Coll. ID: BV180001M and BV180002F], 1♂1♀, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, "North hill prairie". 38.157686°, -90.191259°. Sweep net. 10-vii-2018. C.E. Dana.

INHS 1,001,453 and 1,001,454 [Dana Coll. ID: BV180006F and BV180007M], 1♀1♂, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, “South hill prairie.” 38.154774°, -90.185110°. Sweep net. 10-vii-2018. C.E. Dana.

INHS 1,001,455 and 1,001,456 [Dana Coll. ID: BV190063M and BV190064F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Eagle Prairie”. 38.286587°, -90.303072°. Sweep net. 19-vii-2019. C.E. Dana, G.M. Lewis, J.R. Tetlie.

INHS 1,001,457 [Dana Coll. ID: BV190075M], 1♂, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, “South hill prairie.” 38.154774°, -90.185110°. Sweep net. 12-vii-2019. C.E. Dana, N.R. Mills, G.M. Lewis.

INHS 1,001,458 and 1,001,459 [Dana Coll. ID: BV190077M and BV190078F], 1♂1♀, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, “North hill prairie”. 38.157686°, -90.191259°. Sweep net. 12-vii-2019. C.E. Dana, N.R. Mills, G.M. Lewis.

INHS 1,001,460 and 1,001,461 [Dana Coll. ID: BV200046M and BV200047F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Turkey Prairie”. 38.290969°, -90.296490°. Sweep net. 14-vii-2020. C.E. Dana, N.R. Mills, J. Fricke, M. Fricke.

INHS 1,001,462 and 1,001,463 [Dana Coll. ID: BV200094M and BV200095F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve,

"Long Prairie". 38.288285°, -90.296292°. Sweep net. 14-vii-2020. C.E. Dana, N.R. Mills, J. Fricke, M. Fricke.

INHS 1,001,464 and 1,001,465 [Dana Coll. ID: BV200107M and BV200108F], 1♂1♀, USA, Illinois, Monroe County, Valmeyer, Salt Lick Point Land and Water Reserve, "Newman Prairie". 38.303930°, -90.307731°. Sweep net. 15-vii-2020. C.E. Dana, N.R. Mills, J. Fricke.

Remarks: *B. venosa* (Tribe Fidicinini) was found in the Illinois Natural History Survey Insect Collection (INHS 557,546 – 557,549) but was not included in Sanborn and Phillips (2013) report. The localities listed above (Monroe County) all fall within the Northern Section of the Ozark Division of the 14 Natural Divisions in Illinois, as defined by Schwegman et al. (1973). According to records from Sanborn and Phillips (2013) this species has been recorded on the opposite side of the Mississippi River in Missouri.

Neotibicen auriferus (Say 1825) (Figure 1.27)

Distribution in Illinois: Madison Co., Monroe Co. (Figure 1.28)

Voucher specimens:

INHS 837,720 and 837,721 [Dana Coll. ID: NAur19001M and NAur19003F], 1♂1♀, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, "North hill prairie". 38.157686°, -90.191259°. Sweep net. 28-viii-2019. C.E. Dana. GenBank Accession OK626637.

INHS 837,722 and 837,723 [Dana Coll. ID: NAur19006M and NAur19007F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve,

“Eagle Prairie”. 38.286587°, -90.303072°. Sweep net. 7-ix-2019. C.E. Dana, J.R. Tetlie, C.J. Williams, N.R. Mills, J. Fricke. GenBank Accession OK626638.

INHS 837,724 and 837,725 [Dana Coll. ID: NAur19008M and NAur19009F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Turkey Prairie”. 38.290969°, -90.296490°. Sweep net. 7-ix-2019. C.E. Dana, J.R. Tetlie, C.J. Williams, N.R. Mills, J. Fricke. GenBank Accession OK626639; GenBank Accession OK626640.

INHS 837,726 and 837,727 [Dana Coll. ID: NAur19010M and NAur19011F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Long Prairie”. 38.288285°, -90.296292°. Sweep net. 7-ix-2019. C.E. Dana, J.R. Tetlie, C.J. Williams, N.R. Mills, J. Fricke.

INHS 837,728 and 837,729 [Dana Coll. ID: NAur20072M and NAur20073F], 1♂1♀, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, “North hill prairie”. 38.290969°, -90.296490°. Sweep net. 4-ix-2020. C.E. Dana, J.R. Tetlie.

INHS 837,730 and 837,731 [Dana Coll. ID: NAur20074M and NAur20075F], 1♂1♀, USA, Illinois, Monroe County, Prairie du Rocher, Fults Hill Prairie Nature Preserve, “South hill prairie.” 38.154774°, -90.185110°. Sweep net. 4-ix-2020. C.E. Dana, J.R. Tetlie.

INHS 1,001,465 and 1,001,466 [Dana Coll. ID: NAur20076M and NAur20077F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Long Prairie”. 38.288285°, -90.296292°. Sweep net. 5-ix-2020. C.E. Dana, J.R. Tetlie, J. Fricke.

INHS 1,001,467 and 1,001,468 [Dana Coll. ID: NAur20078M and NAur20079F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Turkey Prairie”. 38.290969°, -90.296490°. Sweep net. 5-ix-2020. C.E. Dana, J.R. Tetlie, J. Fricke.

INHS 1,001,469 and 1,001,470 [Dana Coll. ID: NAur20080M and NAur20081F], 1♂1♀, USA, Illinois, Monroe County, Waterloo, Illinois Ozarks Nature Preserve, “Eagle Prairie”. 38.286587°, -90.303072°. Sweep net. 5-ix-2020. C.E. Dana, J.R. Tetlie, J. Fricke.

INHS 1,001,471 [Dana Coll. ID: NAur20082M], 1♂, USA, Illinois, Madison County, Godfrey Township, John M. Olin Nature Preserve. 38.915828°, -90.225474°. Sweep net. 11-ix-2020. C.E. Dana.

INHS 1,001,472 [Dana Coll. ID: NAur210007M], 1♂, USA, Illinois, Monroe County, Valmeyer, Salt Lick Point Land and Water Reserve, “Boy Scout Prairie”. 38.308095°, -90.303985°. Sweep net. 12-ix-2020. M.J. Thomas, G.M. Lewis, J.R. Tetlie, J. Fricke.

INHS 1,001,473 [Dana Coll. ID: NAur210008M], 1♂, USA, Illinois, Monroe County, Valmeyer, Salt Lick Point Land and Water Reserve, “Newman Prairie”. 38.303930°, -90.307731°. Sweep net. 12-ix-2020. M.J. Thomas, G.M. Lewis, J.R. Tetlie, J. Fricke.

INHS 1,001,474 [Dana Coll. ID: NAur210066M], 1♂, USA, Illinois, Madison County, Godfrey Township, John M. Olin Nature Preserve. 38.915828°, -90.225474°. Sweep net. 10-ix-2021. M.J. Thomas, G.M. Lewis, J.R. Tetlie.

Additional material examined:

INHS 837,721 [Dana Coll. ID: NAur19003F] GenBank Accession OK626637 1467 bp COI – 99.46% identity to Accession KR674194.1 (*Neotibicen auriferus* isolate 07.US.KS.MDP.01); Next best match 98.56% identity to Accession KR674222.1 (*Neotibicen davisi harnedi* isolate 08.MS.STX.01).

INHS 837,722 [Dana Coll. ID: NAur19006M] GenBank Accession OK626638 779 bp COI – 99.61% identity to Accession KR674194.1 (*Neotibicen auriferus* isolate 07.US.KS.MDP.01); Next best match 98.97% identity to Accession KR674222.1 (*Neotibicen davisi harnedi* isolate 08.MS.STX.01).

INHS 837,724 [Dana Coll. ID: NAur19008M] GenBank Accession OK626639 1440 bp COI – 99.39% identity to Accession KR674194.1 (*Neotibicen auriferus* isolate 07.US.KS.MDP.01); Next best match 98.57% identity to Accession KR674222.1 (*Neotibicen davisi harnedi* isolate 08.MS.STX.01).

INHS 837,725 [Dana Coll. ID: NAur19009F] GenBank Accession OK626640 1470 bp COI – 99.40% identity to Accession KR674194.1 (*Neotibicen auriferus* isolate 07.US.KS.MDP.01); Next best match 98.51% identity to Accession KR674222.1 (*Neotibicen davisi harnedi* isolate 08.MS.STX.01).

Comparative material examined from other states:

INHS 1,001,475 [Dana Coll. ID: NAur200083F], 1♀, USA, Missouri, Maries County, Vichy, Spring Creek Gap Conservation Area. 38.143826°, -91.810103°. Sweep net. 15-ix-2020. C.E. Dana.

Remarks: *Neotibicen auriferus* (Tribe Cryptotympanini) is distinct from other Illinois species within the genus as it oviposits in grass stems rather than in shrubs or trees

(Beamer 1925). This behavior was observed at the Illinois Ozarks Nature Preserve in 2019 (Figure 1.30). This species may have been misidentified in previous studies in other hill prairies of Illinois (Wallner 2011) due to its similar size, general appearance, and male chorus to *Neot. canicularis* (Figure 1.14; Figure 1.29). Figure 1.29 shows a spectrogram from Fults Hill Prairie Nature Preserve ("North Prairie"), recorded 5-ix-2019, recorded the same year INHS 837,720 and 837,721 were collected.

New County Records

Okanagana balli Davis 1919 (Figure 1.21)

Distribution in Illinois (new county record in bold): Cook Co., Kane Co, **McLean Co.**, Ogle Co., Will Co. (Figure 1.22)

Voucher specimens:

INHS 1,001,476 and 1,001,477 (Dana Coll. ID: OK210004M and OK210005F), 1♂1♀, USA, Illinois, McLean Co., Chenoa, Weston Cemetery Prairie Nature Preserve. 40.746701°, -88.614504°. Sweep net. 25-vi-2021. M. Keeley, S. Merkelz, M. Cristofaro.

Additional material examined:

Chicago Field Museum 4188481, 1♂, USA, Illinois, Will Co., 3 mi S of Monee, relict prairie along Illinois Central RR. 9-vii-1961. H.S Dybas.

Chicago Field Museum 4188482, 1♂, USA, Illinois, Cook Co., Kensington Railroad Tracks. 29-vi-1978. R.W. Hamilton.

Remarks: There is also a known population of this species at James Woodworth Prairie Preserve (42.0598188°, -87.841581°) in Glenview, Cook County, Illinois. This site was

visited by our group on July 2, 2016 and male cicadas were observed calling and photo-documented, but not collected due to lack of permissions.

Cicadettana calliope calliope (Walker, 1850) (Figure 1.17)

Distribution in Illinois: (new county records in bold): Champaign Co., Clay Co., Effingham Co., **Fayette Co.,** Ford Co., Grundy Co., Henry Co., Iroquois Co., **Madison Co.**, Marion Co., Mason Co., **McLean Co.**, Monroe Co., **Morgan Co.**, Ogle Co., Perry Co., Sangamon Co., **Vermillion Co.** (Figure 1.18)

Voucher specimens:

INHS 1,001,478 and 1,001,479 [Dana Coll. ID: CC190045M and CC190066F], 1♂1♀, USA, Illinois, Fayette Co., Farina, Right-Of-Way (ROW) Prairie along Route 37 at County Rd. 1900 E crossroad ("Tract 5" of 12 Mile Prairie). 38.819829°, -88.788420°. Sweep net. 2-vii-2019. C.E. Dana, G.M. Lewis, S. Carlson.

INHS 1,001,480 [Dana Coll. ID: CC190105M], 1♂, USA, Illinois, Fayette Co., LaClede Township, Right-Of-Way (ROW) Prairie along Route 37 at County Rd. 700N crossroad ("Tract 4" of 12 Mile Prairie). 38.842177°, -88.760195°. Sweep net. 10-vii-2019. C.E. Dana, G.M. Lewis, S.M. Wilson.

INHS 1,001,481 and 1,001,482 [Dana Coll. ID: CC200086F and CC200087M], 1♀1♂, USA, Illinois, McLean Co., Chenoa, Weston Cemetery Prairie Nature Preserve. 40.746701°, -88.614504°. Sweep net. 29-vi-2020. C.E. Dana, N.R. Mills.

INHS 1,001,483 and 1,001,484 [Dana Coll. ID: CC210021M and CC210022F], 1♂1♀, USA, Illinois, Madison Co., Godfrey, John M. Olin Nature Preserve. 38.915828°, -90.225474°. Sweep net. 10-vi-2021. J.R. Tetlie, M. Keeley, S. Merkelz, G.M. Lewis.

INHS 1,001,485 and 1,001,486 [Dana Coll. ID: CC210036M and CC210037F], 1♂1♀, USA, Illinois, Morgan Co., Meredosia, Meredosia Hill Prairie Nature Preserve. 39.853305°, -90.465657°. Sweep net. 23-vi-2021. J.R. Tetlie, M. Keeley.

INHS 1,001,487 [Dana Coll. ID: CC200123M], 1♂, USA, Illinois, Vermillion Co., Rankin, Pellsville Cemetery Prairie. 40.461108°, -87.923331°. Sweep net. 10-vii-2020. C.E. Dana, J.R. Tetlie.

INHS 1,001,488 [Dana Coll. ID: CC210064M], 1♂, USA, Illinois, Vermillion Co., Rankin, Pellsville Cemetery Prairie. 40.461108°, -87.923331°. Sweep net. 28-vi-2021. G.M. Lewis, M. Keeley.

Additional material examined:

INHS 1,001,489 and 1,001,490 [Dana Coll. ID: CC210014M and CC210015F], 1♂1♀, USA, Illinois, Monroe Co., Prairie du Rocher, Fults Hill Prairie Nature Preserve, “North hill prairie”. 38.290969°, -90.296490°. Sweep net. 11-vi-2021. J.R. Tetlie, M. Keeley, S. Merkelz, G.M. Lewis.

Remarks: Specimens from Monroe County are included in “additional material examined” as they have not been reported in the county after 1957 (INHS Specimen 555916-555919) and my specimens increase the range within the county to include an additional clifftop prairies. Archived notes from the collector, Milton Sanderson, indicate

that they found *C. calliope calliope* at a clifftop prairie at what is now “White Rock Nature Preserve” based on locality notes.

Megatibicen auletes (Germar, 1834) (Figure 1.8)

Distribution in Illinois (new county records in bold) (Figure 1.11): Bond Co., Champaign Co., Coles Co., Cook Co., Cumberland Co., Franklin Co., Iroquois Co., Jefferson Co., Kankakee Co., La Salle Co., **Lee Co.**, Macoupin Co., Madison Co., Massac Co., **Marion Co.,** Mason Co., Morgan Co., Peoria Co., Pope Co., Saint Clair Co., Union Co., Washington Co., Will Co., Williamson Co., Woodford Co.

Voucher specimens:

INHS 1,001,491 [Dana Coll. ID: MAu190003F], 1♀, USA, Illinois, Lee Co., Nachusa Township, Nachusa Grasslands preserve, “Tellabs”. 41.893917°, -89.376566°. Found under oak tree beneath poison ivy. 23-vii-2019. C.E. Dana, S.M. Wilson, S. Carlson.

INHS 1,001,492 [Dana Coll. ID: MAu190001M], 1♂, USA, Illinois, Marion Co., Salem, Bryan Memorial Park. 38.637800°, -88.946921°. Collected emerging nymph and let eclose in cage. 31-vii-2019. C.E. Dana, S.M. Wilson, G.M. Lewis.

Megatibicen dorsatus (Say, 1825) (Figure 1.12)

Distribution in Illinois (new county records in bold) (Figure 1.11): Champaign Co., Christian Co., Clinton Co., Effingham Co., Fayette Co., Ford Co., Franklin Co., Hancock Co., Iroquois Co., Jasper Co., Marion Co., Mason Co., **Vermilion Co.,** Washington Co.

Voucher specimens:

Dana Coll. ID: MDor20209M, 1♂, USA, Illinois, Vermilion Co., Rankin, Pellsville Cemetery Prairie. 40.461108°, -87.923330°. Caught by net near the border with Herschel Workman Pheasant Area, but very few individuals calling. 18-viii-2020. C.E. Dana.

Megatibicen pronotalis walkeri Metcalf, 1955 (Figure 1.10)

Distribution in Illinois (new county records in bold) (Figure 1.13): Adams Co., **Champaign Co.**, Clay Co., Clinton Co., Cook Co., Hancock Co., Hardin Co., Jackson Co., Jersey Co., Kankakee Co., Madison Co., **Marion Co.**, Mason Co., Montgomery Co., Peoria Co., Perry Co., Pike Co., Pope Co., Pulaski Co., Randolph Co., Rock Island Co., Saint Clair Co., Saline Co., Shelby Co., Union Co., Washington Co., White Co., **Woodford Co.**

Voucher specimens:

INHS [Dana Coll. ID: MPro180001F], 1♀, USA, Illinois, Marion Co., Omega Township, Stephen A Forbes State Recreation Area, Rocky Point Beach. 38.71515543°, -88.75247666°. Sweep net. 2-viii-2018. C.E. Dana.

INHS 1,001,495 [Dana Coll. ID: MPro190003M], 1♂, USA, Illinois, Champaign Co., Champaign, Legacy Ave. 40.152140°, -88.283143°. Sweep net. 1-x-2019. T. McElrath.

INHS 1,001,496 [Dana Coll. ID: MPro190002M], 1♂, USA, Illinois, Woodford Co., Low Point, Woodford Fish and Wildlife Habitat Area, Jenkins Marsh. 40.879762°, - -89.45621931°. Sweep net. 19-ix-2019. C.E. Dana.

Additional material examined:

INHS 1,001,494 [Dana Coll. ID: MPro170001M], 1♂, USA, Illinois, Union Co., Dongola, Dongola Gas Station. 37.3686004°, -89.15747569°. Caught at gas station lights in the evening. 20-viii-2017. C.E. Dana. (Figure 1.10)

Discussion

Illinois is unique, as it is the only state east of the Mississippi where several western cicada species can be found; of note, it is the only state where *Diceroprocta vitripennis* and *Megatibicen dorsatus* occur east of the Mississippi (Table 1.1). Two new state records for *Neotibicen auriferus* and *Beameria venosa* were included in this study, found solely in clifftop prairies in southwestern Illinois (Figure 1.26; Figure 1.28). These clifftop prairies are also home to other animals commonly found across the Mississippi in the Ozarks, like the scorpion *Centruroides vittatus* (Shelley and Sissom 1995) and the Eastern Coachwhip Snake (Smith 1961). The localities for *B. venosa* and *Neot. auriferus* all fall within the Northern Section of the Ozark Division of the 14 Natural Divisions in Illinois, as defined by Schwegman et al. (1973). *B. venosa* is listed on the Missouri State Wildlife Action Plan as a vulnerable (S3) as well as a characteristic species for grassland habitats. The same wildlife action plan says that characteristic species are "indicative of the diversity and health of the wildlife characteristic of a specific habitat type, are ideal for monitoring management effectiveness and overall community health" (MWAP 2015, p. 4). These clifftop habitats are also uniquely threatened. Like many other hill prairies in Southern Illinois, Fults Hill Prairie Nature Preserve has lost a large percentage of prairie habitat acreage to woody encroachment as a result of a lack of regular fire management (Jones and Bowles 2013).

There are several potential explanations for the presence of more western species in Illinois, despite the geographic barrier of the Mississippi River. One possibility for why these more western species occur in Illinois is a result of the Holocene Hypsithermic Interval, having colonized during this warmer period after the last known glaciation event. Approximately 9000 ya there was a general warming across the globe followed by a dry period (from ~8000 ya to 2000 ya) that may have promoted invasion of more western species across the barrier Mississippi River, similar to shifts in other regions of North America during this time period (Kaul et al. 1988; Purdue 1989; Ratcliffe and Hammond 2002; Wolverton 2005). Species that spread across the Mississippi may still remain in Illinois as a result of microclimates maintained in habitats like hill prairies, but more work is needed on the genetic timescale and investigation into rates of migration into some of these isolated populations.

Potential for further work

Given the limited time and large geographic scale of Illinois, the western and northern regions of the state were not surveyed in great detail during this study. Other cicada species may have exploited the Holocene Hypisthermal Interval of dry conditions and moved eastward into Illinois, and further acoustic and manual surveys should be conducted in Western and Northern Illinois. This would likely help illuminate the distribution of the more cryptic *Okanagana* species in Illinois. It is likely that the range of *O. balli* is greater than our current knowledge and unpublished reports suggest that the range is greater in the Chicago region (Strang 2022). The last known sighting of *O. r. rimosa* in Illinois was prior to 1962, which could be due to it being extirpated from the state or from minimal efforts at locating it.

Phylogenetic and taxonomic work has been done to resolve the relationships of North American species previously united under *Tibicen* Latreille 1825, however, the authors (Hill et al. 2015) note that there was poor phylogenetic resolution within the clade containing *Neot. canicularis*, *Neot. pruinosus,* and *Neot. lyricen* due to difficulties encountered with mtDNA. Heath (1978) created a key those cicada genera whose distribution includes all of North America, north of Mexico, but only in the past decade has a key to the species of *Neotibicen* been published (Lee 2016) and other genera lack similar keys. Available keys often only address regional species (e.g., Froeschner 1952; Alexander et al. 1972) and sometimes use characters, like wing curvature, that early authors found were unreliable (Beamer 1928). Often the most reliable way to identify a species is by its call (Boulard 2006), but this cannot be applied to museum specimens and even when collecting, it can be unclear which male is calling. Furthermore, it is difficult to induce some species of males to call in captivity.

Ongoing threats to cicada populations

In addition to loss in significant biodiversity in the Anthropocene, there is also concern over the loss of insect biomass, known colloquially as "Insect Armageddon" (Sorg et al. 2013; Hallmann et al. 2017). This loss of biomass could have large-scale impacts on higher trophic levels. Pons (2020) reviewed the diet of Paleartic birds, the role that cicadas play as a food source, and decline of insectivorous birds. North American birds have undergone widespread decline over the past half century, with grassland birds showing the largest decline — approximately 53% in population and 74% of species since the 1970s (Rosenberg et al. 2019). As an example, Illinois is home to the Greater Prairie Chicken, *Tympanuchus cupido pinnatus*, which have seen an incredible decline in their population and genetic diversity over the past few

decades (Bouzat et al. 2009; Mussmann et al. 2017). This species is known to be partly insectivorous, but little is known about its specific diet and how it has changed with the changing landscape.

It is possible that treating ash trees for emerald ash borer (*Agrilus planipennis* Fairmaire, 1888) may impact cicada populations in urban areas, although to what extent has not been determined. Many insecticides recommended for use against emerald ash borer are systemic insecticides applied by trunk injection or soil drenching (Herms et al. 2019) and typically have high activity against piercing and sucking insects like cicadas (Horowitz and Ishaaya 2004; Hahn et al. 2011). Unlike other hemipterans, few studies have been done on plant pathogen transmission resulting from cicadas feeding, and those that exist show little to no evidence of cicadas acting as vectors (Cornara et al. 2020). Although there is some evidence that cicadas can act as pests by restricting above ground tree growth (Karban 1980), most impacts of cicadas on trees appears superficial (White and Sedcole 1993; Cook and Holt 2002) and other studies suggest a need to document impact on production (Saljoqi et al. 2010). Above ground damage from ovipositing *Magicicada* females can be easily controlled by bagging young trees (Ahern et al. 2005) which is more effective than insecticide treatments in reducing damage (Miller 1997). Covering young trees with muslin has been the recommendation since the early 20th century in the case of apple orchards (Marlatt 1907; Herrick 1925).

Soil compaction may also impact cicada diversity as a result of soil hardness, and this impact may be mitigated in some species by better burrowing abilities (Moriyama and Numata 2015). This may be driving some of the distribution patterns seen in Illinois and warrants future research.

Recommendations for listing and management

Given the fragmented landscape and unknown impacts of land management practices, the Wildlife Conservation and Restoration Program and the Illinois State Wildlife Grant Program have created an Illinois Wildlife Action Plan which lists three cicadas as Species in Greatest Conservation Need (SGCN) or places them on the watchlist – *D. vitripennis*, *O. balli*, and *Me. dorsatus* (IWAP 2015, pages 49, 265, 273). Given their highly restricted range and rarity, I propose adding *Neot. auriferus* and *B. venosa* as SGCN species, as well (Table 1.3). I also found that *O. balli* has an extremely restricted distribution (Figure 1.22) and should be considered for elevation from the Illinois Watch List to threatened in Illinois. While *Me. auletes* is found in multiple counties throughout Illinois, and I was able to add more information on distribution with this study (Figure 1.11), more information is needed on population sizes given their primary association with oak trees and patchy distribution within surburban areas (Sanborn and Phillips 2013; Dana *pers. obs.*). Despite my efforts, I was unable to sample *Neoc. h. hieroglyphica* over the years of this study. Acoustic recordings of *Neoc. h. hieroglyphica* were obtained, but a larger effort to locate individuals to better assess their status is needed. Therefore, I suggest placing *Neoc. h. hieroglyphica* on the Illinois Watch List. Given that *O. r. rimosa* was not observed throughout the study, nor has it been reported in any known collection since prior to 1962, more information is needed. Given timing and logistical constraints, I was unable to sample to any large degree in Northern Illinois. Thus, I recommend placement of *O. r. rimosa* on the Illinois Watch List to determine if this species is indeed extirpated within Illinois.

In terms of general management recommendations for cicada conservation, land managers should do patchy burning, especially in smaller prairies like cemetery prairies where there may be no refuge. Additionally, any mowing efforts should consider oviposition phenology

for a given species, as I have observed mowing in areas either during or directly after cicadas were observed ovipositing on plants in the area (both in grass parking lots and along roadsides). With these changes, I hope that conservation of cicadas, as ecologically important insects, will be ensured for many years to come.

Acknowledgments

Funding was provided by the State Wildlife Grant Program through the U.S. Fish and Wildlife Service and administered by the Illinois Department of Natural Resources to help address conservation needs identified in the Illinois Wildlife Action Plan. Additional funding was provided by the Friends of Nachusa Grasslands, the Ross Memorial Fund (Herbert Holdsworth Ross Award), and Prairie Biotic Research. Thank you to the Illinois Preserves Nature Commission, Illinois Department of Transportation, Illinois Department of Natural Resources, Grand Prairie Friends, Nachusa Grasslands, and the Nature Conservancy for their site permissions and assistance. Thank you to field technicians Masato Keeley, Ivan Shilov, Sara Merkelz, Samantha Davis, Kate Johnson, Jocelyn Hedlund, Vanessa Gabel, Maryssa Cristofaro, Bailey Clancy, and Shannon Carlson. Thank you to Grace Lewis for help with photography of specimens. Many site managers and heritage biologists helped to make much of the field work possible, including, but not limited to Joann Fricke, Mike Fricke, and Terry Esker. Thank you to Mark Davis (Illinois Natural History Survey) for access to the Collaborative Conservation Genetics Laboratory for DNA extractions and PCRs.

Tables

Table 1.1. Cicada species of Illinois. Bolded states are new records.

Species name	Common name	Known Range in North America
Beameria venosa (Uhler, 1888)		AZ, AR, CO, IA, **IL**, KS, MO, NE, NM, OK, TX, UT, Mexico (Baja California Sur, Chihuahua, Coahuila, Durango, Nuevo Leon, Sonora, Tamaulipas, Veracruz)
Cicadettana calliope calliope (Walker, 1850)	Small grass cicada	AL, AR, CO, FL, GA, IL, IN, IA, KS, KY, LA, MD, MS, MO, NE, NC, OH, OK, SC, SD, TN, TX, VA
Diceroprocta vitripennis (Say, 1830)	Green-winged scrub cicada	AL, AR, IL, IN, KS, KY, LA, MI, MS, MO, NE, OK, TN, TX, WI
Magicicada cassinii (Fisher, 1852)*	Cassin's periodical cicada (17-year)	GA, IA, IL, IN, KS, KY, MD, MO, NC, NE, NJ, NY, OH, OK, PA, TN, TX, VA, WI, WV
Magicicada neotredecim Marshall & Cooley, 2000	(13-year)	AR, IA, IL, IN, KY, MO, TN
Magicicada septendecim (Linnaeus, 1758)	Linnaeus' periodical cicada (17-year)	CT, DC, DE, GA, IA, IL, IN, KS, KY, MA, MD, MI, MO, NC, NE, NJ, NY, OH, PA, RI, SC, TN, VA, WI, WV
Magicicada septendecula Alexander & Moore, 1962	The little (17-year) cicada	GA, IA, IL, IN, KS, KY, MO, NC, NJ, NY, OH, PA, TN, VA, WV
Magicicada tredecassini Alexander & Moore, 1962	Cassin's (13-year) cicada	AL, AR, GA, IA, IL, IN, KY, MD, MO, MS, NC, OK, SC, TN, VA
Magicicada tredecim (Walsh & Riley, 1868)	Riley's (13-year) cicada	AL, AR, GA, IL, IN, KY, LA, MD, MO, MS, NC, OK, SC, TN, VA
Magicicada tredecula Alexander & Moore, 1962	(13-year)	AL, AR, GA, IA, IL, IN, KY, LA, MO, MS, NC, OK, SC, TN, VA
Megatibicen auletes (Germar, 1834)	Northern dusk singing cicada	AL, AR, CT, DC, DE, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, MI, MO, MS, NC, NE, NJ, NY, OH, OK, PA, SC, TN, TX, VA, WI, WV, Canada (Ontario)
Megatibicen dorsatus (Say, 1825)	Grand prairie cicada	AR, CO, ID, IL, IA, KS, MO, MT, NE, NM, OK, SD, TX, WY
Megatibicen pronotalis walkeri Metcalf, 1955	Walker's cicada	AL, AR, FL, GA, IA, IL, IN, KS, KY, LA, MD, MI, MN, MS, MO, NE, NC, ND, OH, OK, SD, TN, TX, VA, WV, WI, WY

Table 1.1. (continued).

Species name	Common name*	Known Range in North America
Neocicada hieroglyphica hieroglyphica (Say, 1830)	Hieroglyphic cicada	AL, AR, DE, FL, GA, IL, IN, KS, KY, LA, MD, MS, MO, NJ, NY, NC, OH, OK, SC, TN, TX, VA
Neotibicen auriferus (Say, 1825)	Plain's dog day cicada	AR, **IL**, KS, MO, NE, NM, OK, TX
Neotibicen canicularis (Harris, 1841)	Dog-day cicada	AR, CT, DC, IL, IN, IA, KS, ME, MD, MA, MI, MN, MO, NE, NH, NJ, NY, NC, ND, OH, PA, PE, RI, SC, SD, TN, VT, VA, WV, WI, Canada (Manitoba, New Brunswick, Nova Scotia, Ontario, Quebec)
Neotibicen linnei (Smith & Grossbeck, 1907)	Linne's cicada	AL, AR, CT, DE, DC, FL, GA, IL, IN, IA, KS, KY, LA, ME, MD, MA, MI, MN, MS, MO, NE, NJ, NY, NC, OH, ON, PA, SC, TN, VT, VA, WV, WI
Neotibicen lyricen lyricen (De Geer, 1773)	Lyric cicada	AL, AR, CT, DE, DC, FL, GA, IL, IN, IA, KS, KY, LA, MD, MA, MI, MS, MO, NE, NH, NJ, NY, NC, OH, OK, ON, PA, RI, SC, TN, TX, VA, WV, WI
Neotibicen pruinosus pruinosus (Say, 1825)	Scissor grinder cicada	AL, AR, CO, FL, GA, IL, IN, IA, KS, KY, LA, MA, MD, MI, MN, MS, MO, NE, NJ, NY, NC, OH, OK, PA, SC, SD, TN, TX, VA, WV, WI
Neotibicen tibicen tibicen (Linnaeus, 1758)	Swamp cicada	AL, AR, CT, DE, DC, FL, GA, IL, IN, IA, KS, KY, LA, MD, MA, MI, MS, MO, NE, NJ, NY, NC, OH, OK, PA, RI, SC, SD, TN, TX, VT, VA, WV, WI
Okanagana balli Davis, 1919	Ball's prairie cicada	IL, IA, KS, MN, NE, ND, SD, WI
Okanagana rimosa rimosa (Say, 1830)**	Say's cicada	CA, CT, ID, IL, IN, IA, ME, MD, MA, MI, MN, MT, NV, NH, NJ, NY, ND, OH, OR, PA, SD, UT, VT, VA, WA, WI, WY, Canada (Alberta, British Columbia, Manitoba, New Brunswick, Ontario, Quebec)

*In the literature, the species is often spelled *Magicicada cassini*, the correct Latin form, however, the misspelling has never been formally corrected using International Code of Zoological Nomenclature rules. It was, however, addressed in Alexander and Moore (1962).
**Last known record of *Okanagana rimosa rimosa* in Illinois was prior to 1962.

Table 1.2. Primers used for sequencing cytochrome c oxidase subunit I (COI).

Name	Sequence (5' to 3')	Length	Source
C1-J-1490	GGTCAACAAATCATAAAGATATTGG	25	Folmer et al. 1994; Hill et al. 2015
TibCOIRev	CCTCTTTCYTGHGTAATAATGTRTG	25	Hill et al. 2015
C1-J-2195	TTGATTTTTTGGTCATCCAGAAGT	24	Simon et al. 1994; Hill et al 2015
TL2-N-3014	TCCAATGCACTAATCTGCCATATTA	25	Simon et al. 1994; Hill et al. 2015

Table 1.3. Recommended changes to Illinois Wildlife Action Plan (IWAP) Species Greatest Conservation Need (SGCN). Older nomenclature is indicated in the "Synonym" column when appropriate (i.e. when the species name has changed). Current species names are based on Hill et al. (2015), Sanborn and Heath (2016), and Marshall and Hill (2017).

Current Species Name	Synonym	Common Name	Current Status in IL	Proposed Status
Megatibicen auletes	*Tibicen auletes*	Northern Dusk-Singing Cicada	NA*	Illinois Watch List
Megatibicen dorsatus	*Tibicen dorsatus*	Giant Grassland Cicada	Illinois SGCN	Illinois SGCN
Neotibicen auriferus			NA	Illinois SGCN
Cicadettana calliope calliope	*Cicadetta calliope*	Southern Grass Cicada	NA	Illinois Watch List
Okanagana balli	-	NA	Illinois Watch List	Threatened
Okanagana rimosa rimosa	-	Say's Cicada	NA	Illinois Watch List
Diceroprocta vitripennis	-	Green Winged Cicada	Illinois Watch List	Illinois Watch List
Neocicada hieroglyphica hieroglyphica	-	Hieroglyphic Cicada	NA	Illinois Watch List
Beameria venosa	-	Concealed-tymbal cicada	NA**	Illinois SGCN

*Megatibicen auletes is on the Connecticut SGCN list (CWAP 2015).
**Beameria venosa is on the Missouri SGCN list.

Figures

Figure 1.1. Cytochrome c oxidase subunit I (COI) region of a generalized mitochondrial genome (based on Accession MG737764.1 - *Cicadettana calliope calliope* (Walker, 1850) isolate CICCAL mitochondrion, partial genome) showing the position of primers used to amplify for sequencing. Note the overlap between the two regions allowing for a longer read by aligning shorter reads. Image designed using Geneious Prime version 2021.2 created by Biomatters.

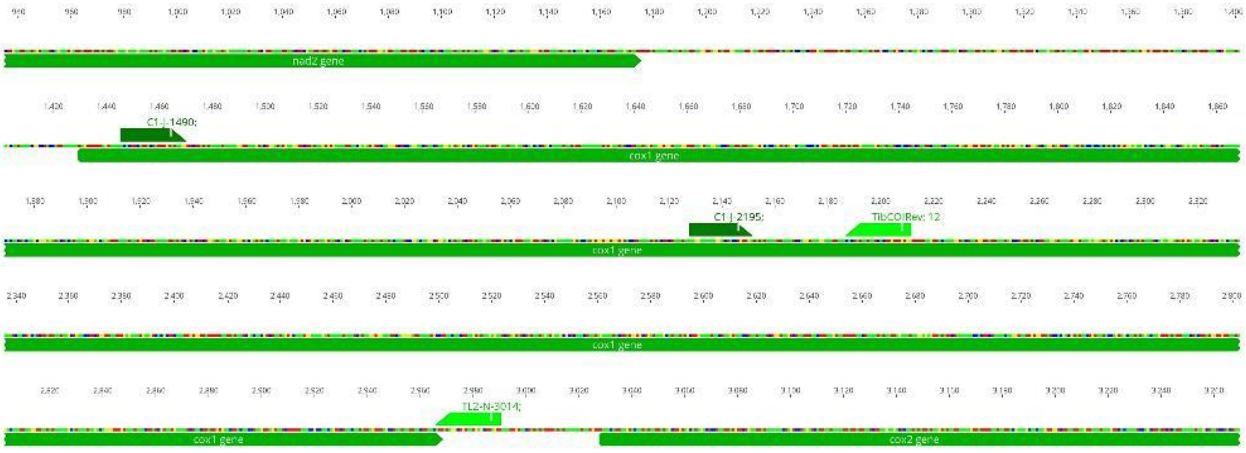


Figure 1.2. 17-year periodical cicada, *Magicicada cassinii* (Fisher, 1852), (Dana Collection MC190093M) ♂, Brood VIII. Settler's Cabin Park, parking lot west of Settler's Cabin Park Wave Pool, USA: PA: Allegheny County: Pittsburgh. 40.433928°, -80.154878°. 9.vi.2019. C.E. Dana, M.J. Thomas, J.R. Dana.

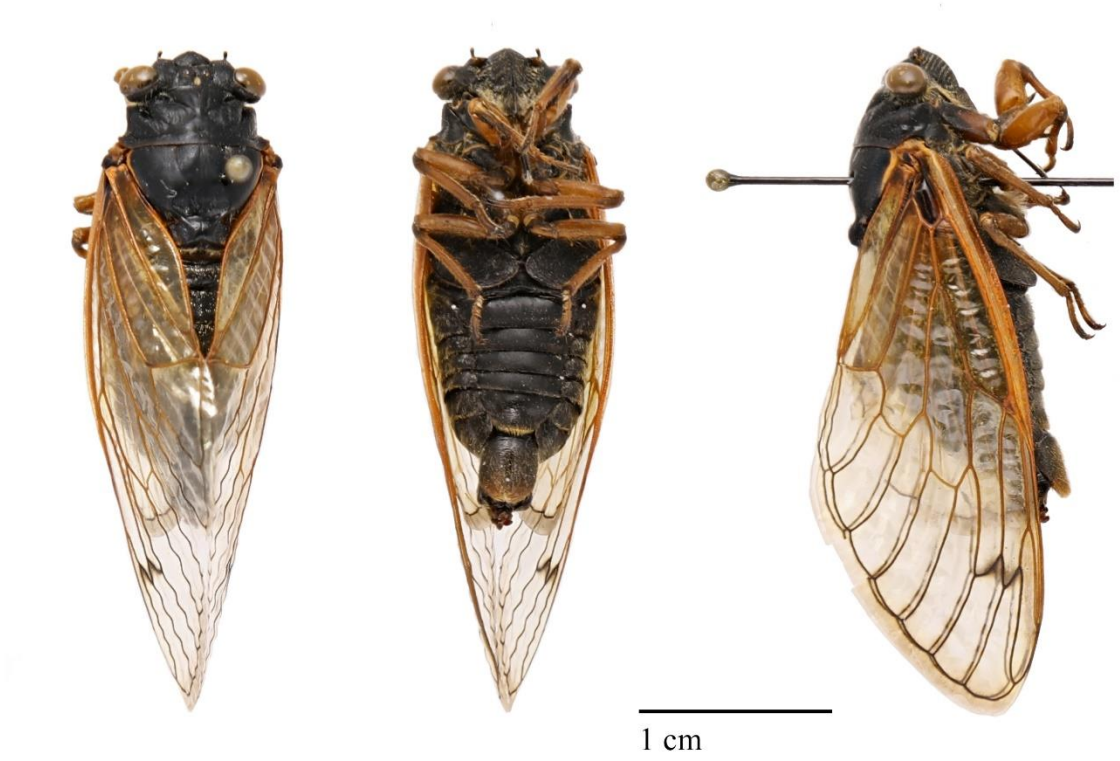


Figure 1.3. 17-year periodical cicada, *Magicicada septendecim* (Linnaeus, 1758), (Dana Collection MC190077M) ♂, Brood VIII. Penn State Beaver Campus Athletic Fields, USA: PA: Beaver County: Monaca. 40.681083 -80.296715. 8-vi-2019. C.E. Dana, M.J. Thomas, J.R. Dana.

Figure 1.4. 17-year periodical cicada, *Magicicada septendecula* (Alexander & Moore, 1962), (Dana Collection MC210432M) ♂, Brood X, USA: IL: Vermillion County: Oakwood. 630 E Rd. Collected on Solter property with permission. 40.08984°, -87.822605°. 11.vi.2021. C.E. Dana.

Figure 1.5. 17-year periodical cicada brood distribution by Illinois county.

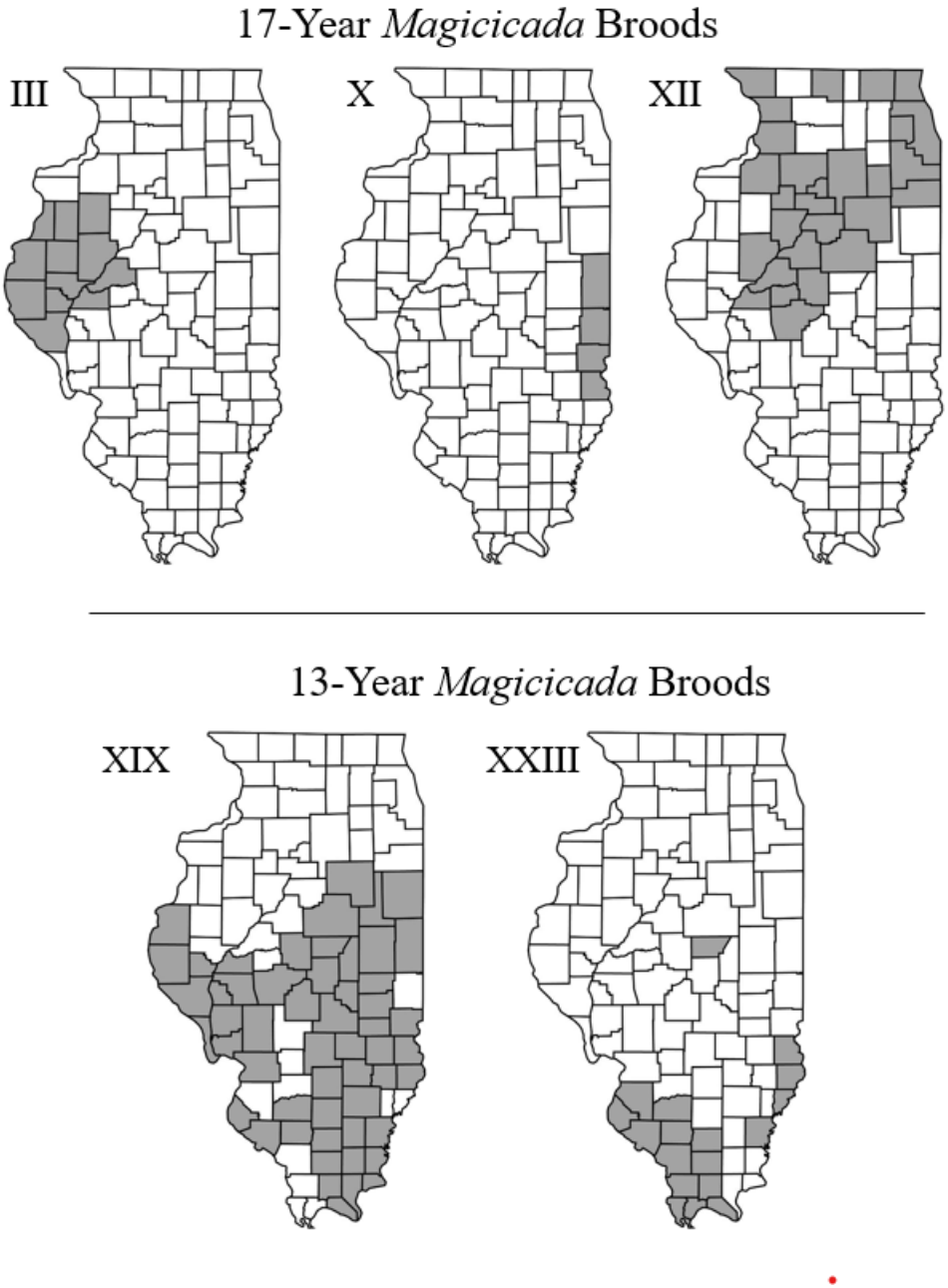


Figure 1.6. Images of Brood XIII. Magicicada cassinii (Fisher, 1852) mating and ovipositing. **a.** Mating *M. cassinii pair,* USA: IL: DuPage County: Hinsdale. Burns Field Park 41.808606°, -87.936069°. C.E. Dana. **b.** Female *M. cassinii* individual ovipositing in branch, USA: IL: Cook County: Westchester. Bemis Woods. 41.824603°, -87.914633°. C.E. Dana.

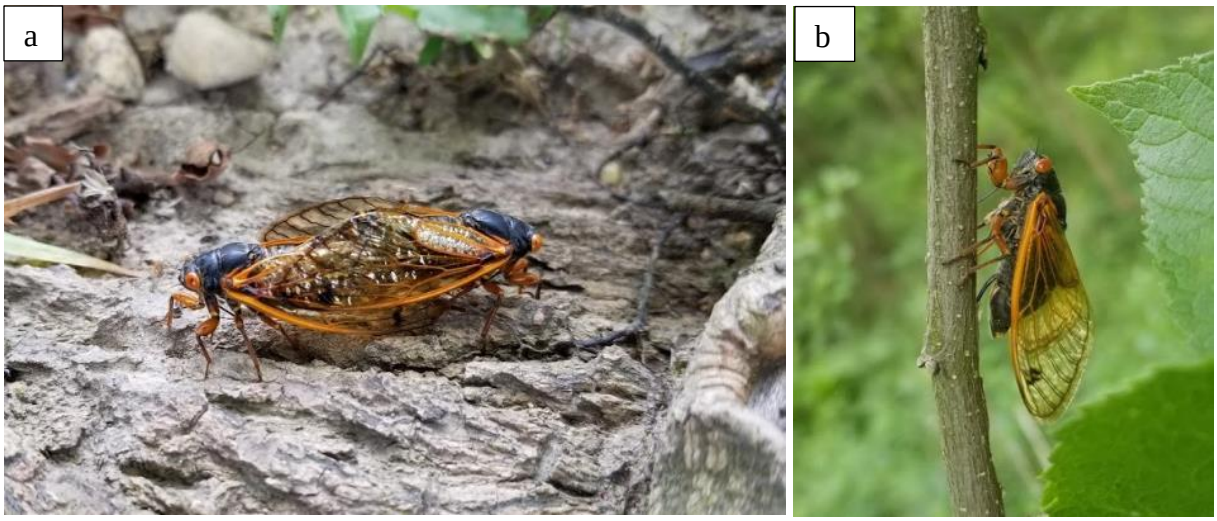


Figure 1.7. Off-cycle emergence of Brood X *Magicicada* spp., USA: IN: Monroe County: Bloomington. Indiana University Bloomington on East 7th Street. 39.168683°, -86.515171°. 23.v.2017. C.E. Dana.

Figure 1.8. *Megatibicen auletes* (Germar, 1834). ♂ (left): (Dana Collection MAu1901M), (INHS Insect Collection 1,001,492), USA: IL: Marion County: Salem. Bryan Memorial Park. 38.637800°, -88.946921°. 31-vii-2019. C.E. Dana. ♀ (right): (Dana Collection MAu1801F), USA: IL: Marion County: Kinmundy. Stephen A. Forbes State Recreation Area. 38.715155°, -88.752477°. 2-viii-2018. C.E. Dana.

Figure 1.9. *Megatibicen dorsatus* (Say, 1825), (Dana Collection NDor170028M) ♂, USA: IL: Iroquois County: Buckley. Right-of-Way Prairie South of Buckley. 40.58152°, -88.044949°. 10-viii-2017. C.E. Dana.

Figure 1.10. *Megatibicen pronotalis walkeri* (Metcalf, 1955), (Dana Collection MPro170001M), (INHS Insect Collection 1,001,494) ♂, USA: IL: Union County: Dongola. Dongola Gas Station. 37.368600°, -89.157475°. Caught at gas station lights in the evening. 20-viii-2017. C.E. Dana.

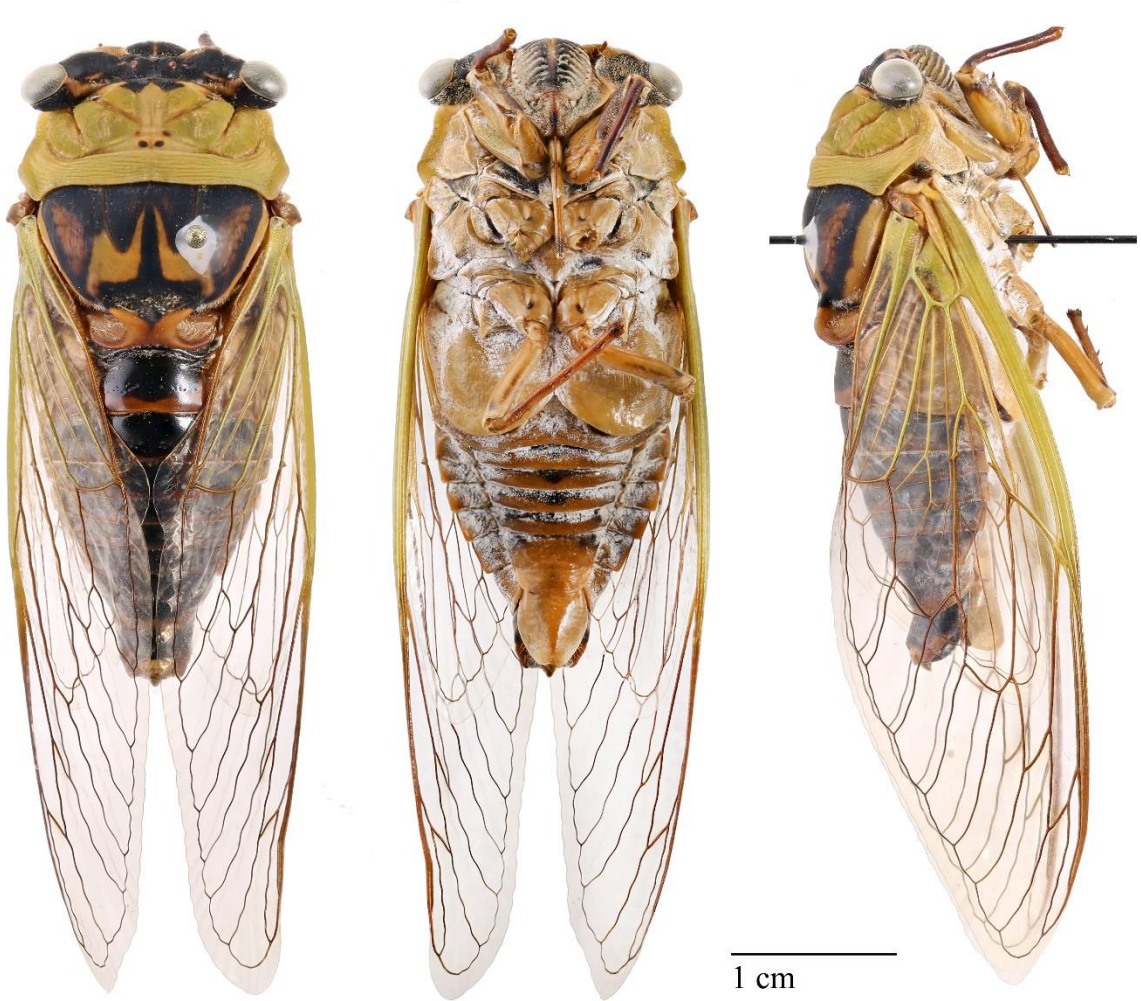


Figure 1.11. County map showing distribution of *Megatibicen auletes* (Germar, 1834) in Illinois. Grey counties are those based on previous records and black counties are those newly added based on specimens collected during the duration of this study.

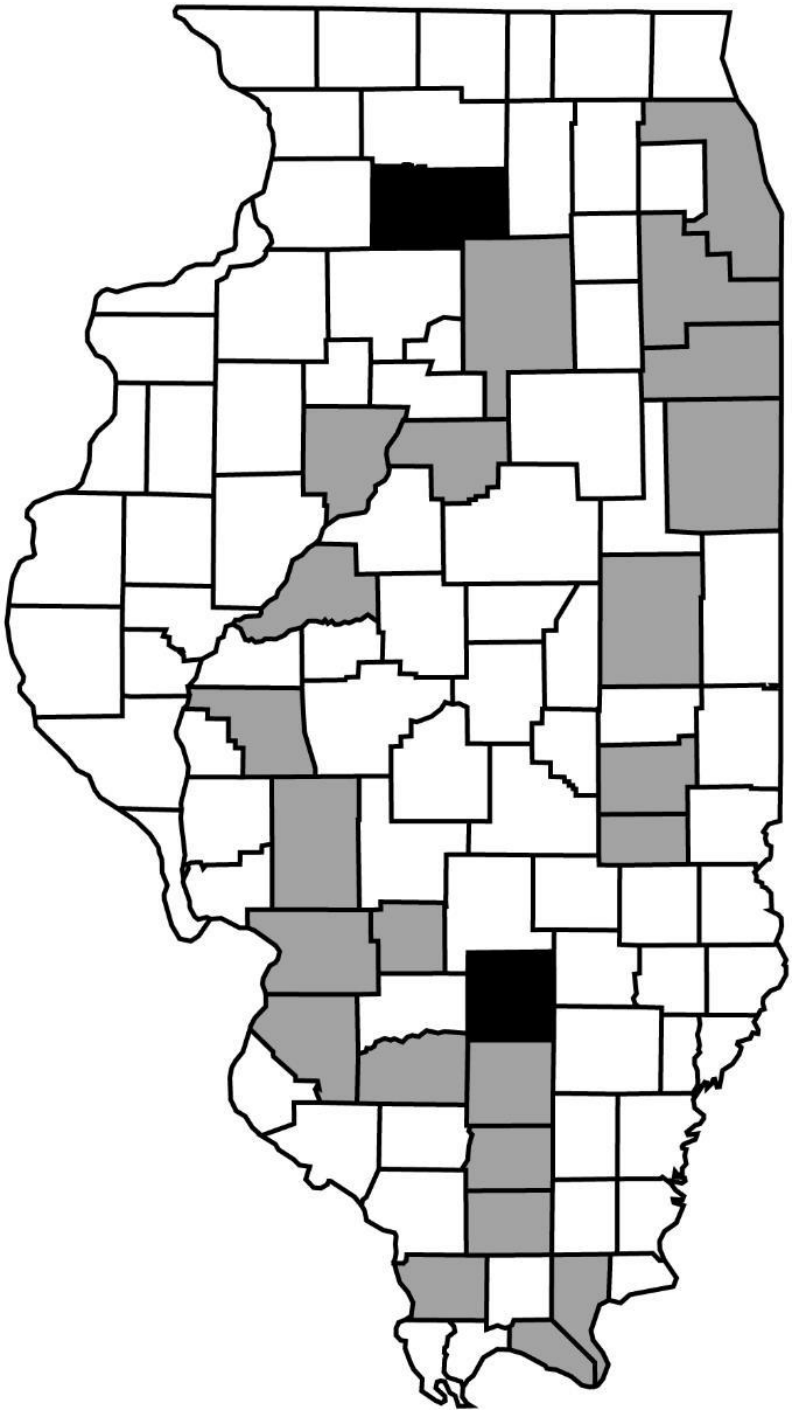


Figure 1.12. County map showing distribution of *Megatibicen dorsatus* (Say, 1825) in Illinois. Grey counties are those based on previous records and black counties are those newly added based on specimens collected during the duration of this study.

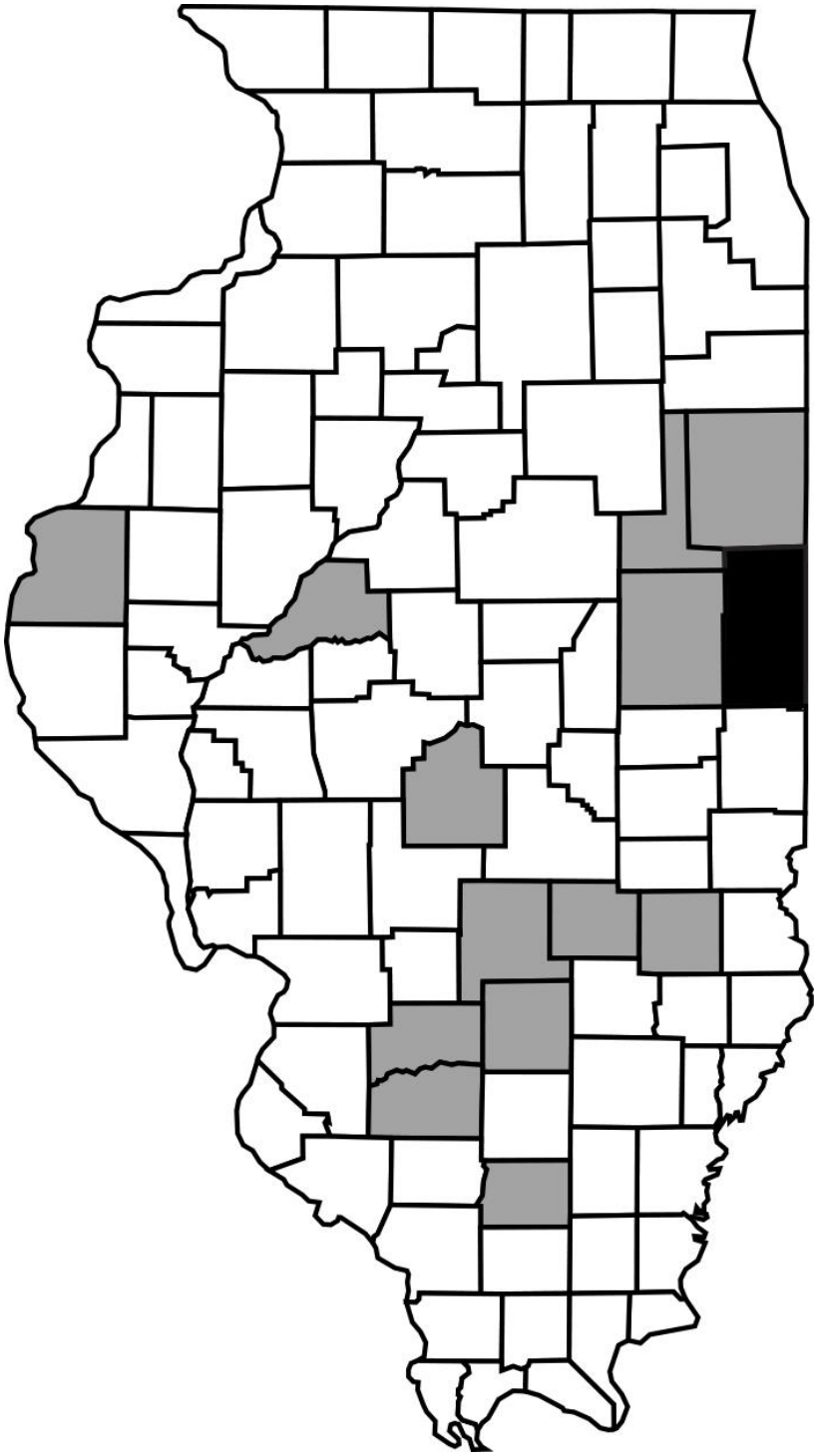


Figure 1.13. County map showing distribution of *Megatibicen pronotalis walkeri* (Metcalf, 1955) in Illinois. Grey counties are those based on previous records and black counties are those newly added based on specimens collected during the duration of this study.

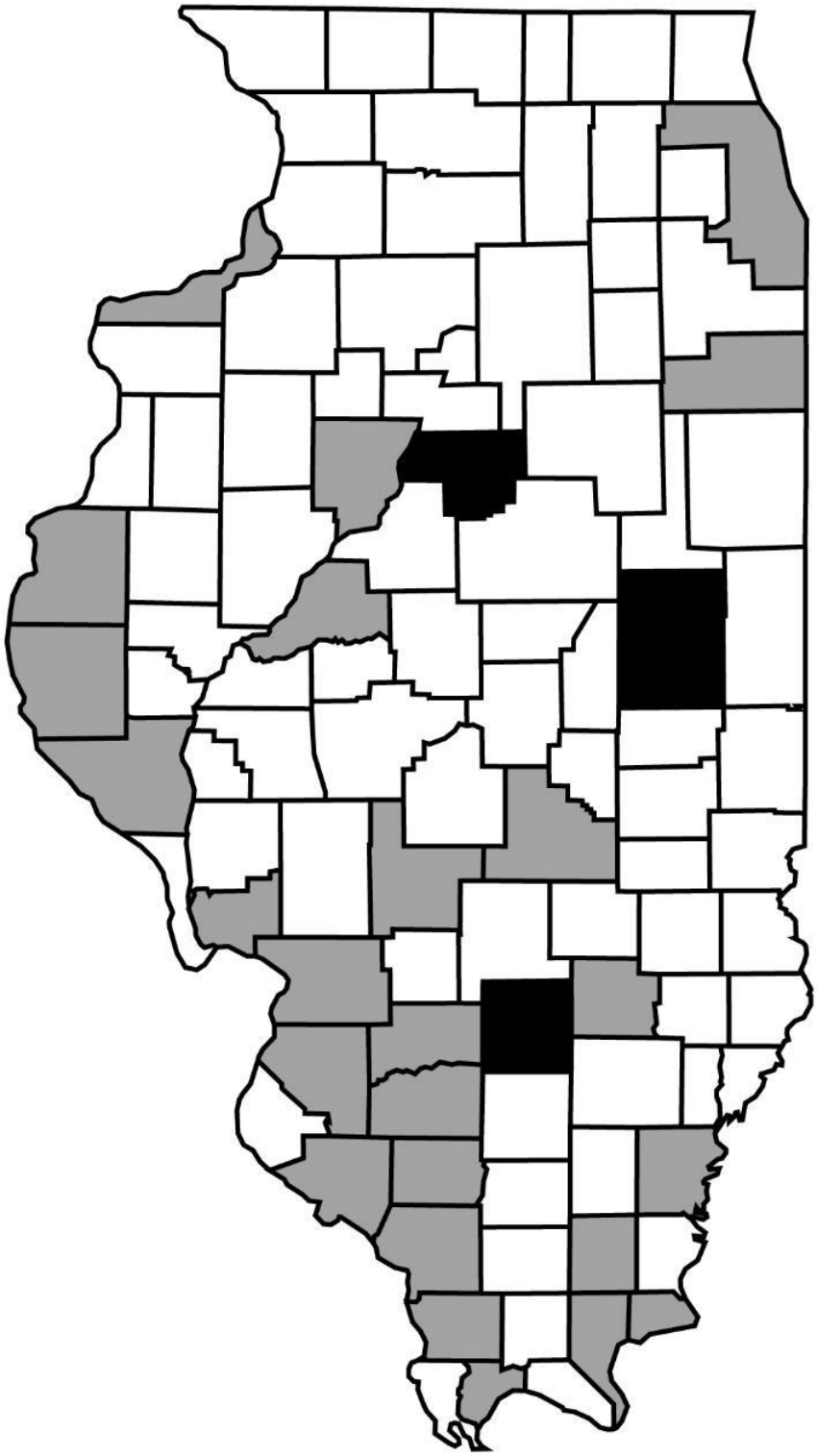


Figure 1.14. *Neotibicen linnei* (Smith & Grossbeck, 1907), *Neotibicen pruinosus pruinosus* (Say, 1825), and *Neotibicen canicularis* (Harris, 1841) with specimen information.

INHS ## 674,049
Neotibicen linnei
Peoria IL

INHS ## 674,338
Neotibicen pruinosus
Champaign IL

INHS ## 669,162
Neotibicen canicularis
Algonquin IL

1 cm

Figure 1.15. *Neotibicen lyricen lyricen* (De Geer, 1773), (Dana Collection NLyr1704M) ♂, USA: FL: Alachua County: Gainesville. Alfred A. Ring Park. 29.671720°, -82.347323°. 12-vi-2017. T. Hedlund.

Figure 1.16. *Neotibicen tibicen tibicen* (Linnaeus, 1758). (Dana Collection NTT200010M) ♂, USA: IL: Fayette County: La Clede Township. 12 Mile Prairie Tract 4. 38.842177°, -88.7602°. 13-viii-2020. C.E. Dana.

Figure 1.17. *Cicadettana calliope calliope* (Walker, 1850), (Dana Collection CC21036M), (INHS Insect Collection 1,001,485) ♂. USA: IL: Morgan County: Meredosia. Meredosia Hill Prairie Nature Preserve. 39.85330486°, -90.46565664°. 23-vi-2021. J.R. Tetlie, M. Keeley.

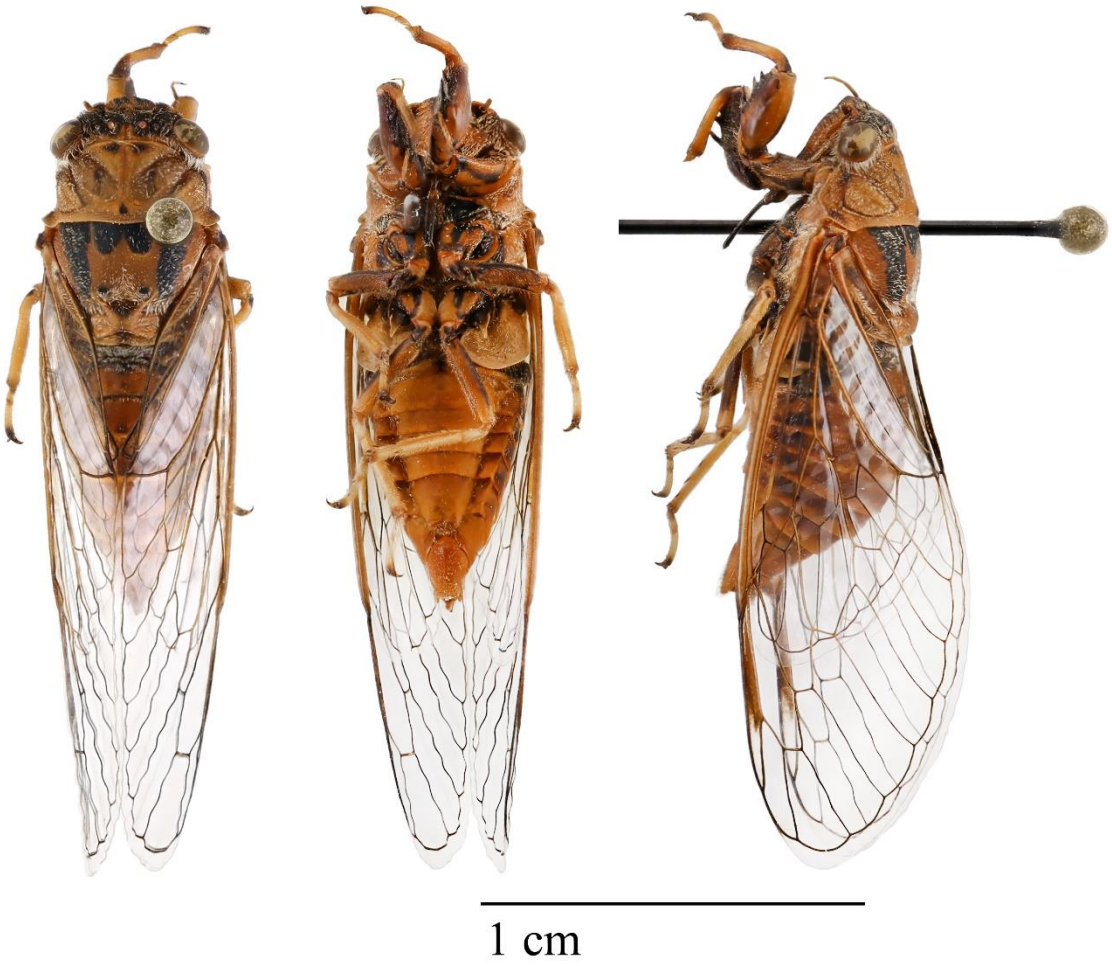


Figure 1.18. County map showing the distribution of *Cicadettana calliope calliope* (Walker, 1850) in Illinois. Grey counties are those based on previous records and black counties are those newly added based on specimens collected during the duration of this study.

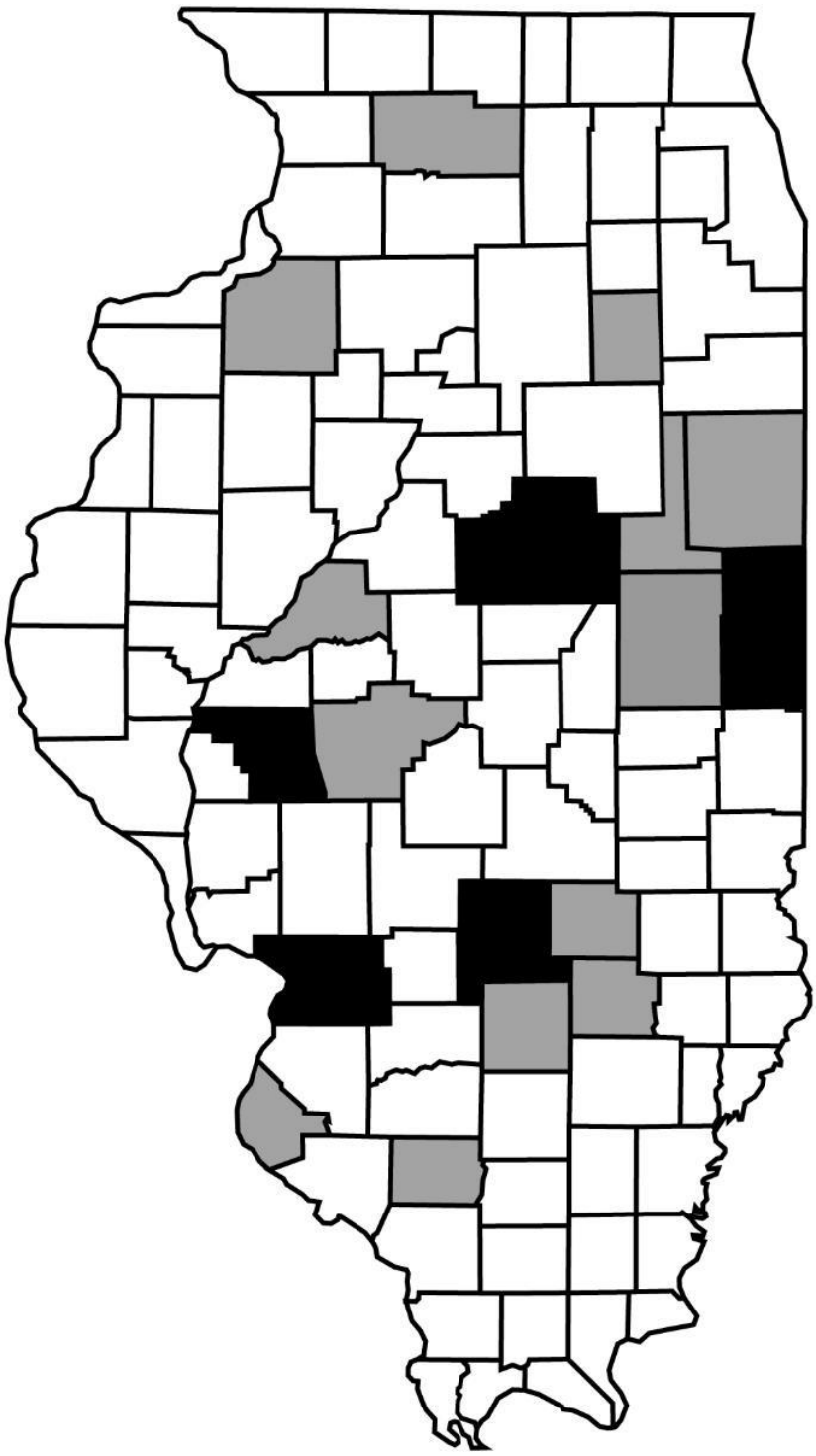


Figure 1.19. *Diceroprocta vitripennis* (Say, 1830), (Dana Collection DV200005M) ♂, USA: IL: Mason County: Topeka. Henry Allen Gleason Nature Preserve. 40.378118°, -89.92856°. 24-vii-2020. C.E. Dana.

Figure 1.20. *Neocicada hieroglyphica hieroglyphica* (Say, 1830), (Dana Collection NH17001M) ♂. USA: MO: Shannon County: Fremont. Peck Ranch Conservation Area. 37.041452°, -91.163398°. 18-vii-2020. J.R. Tetlie.

Figure 1.21. *Okanagana balli* (Davis, 1919), (Dana Collection DOK200001M) ♂, USA:.IL: McLean County: Chenoa. Weston Cemetery Prairie Nature Preserve. 40.746767°, -88.614269°. 7-vii-2020. C.E. Dana.

Figure 1.22. Known distribution of *Okanagana balli* (Davis, 1919). Grey counties are those based on previous records and black counties are those newly added based on specimens collected during the duration of this study.

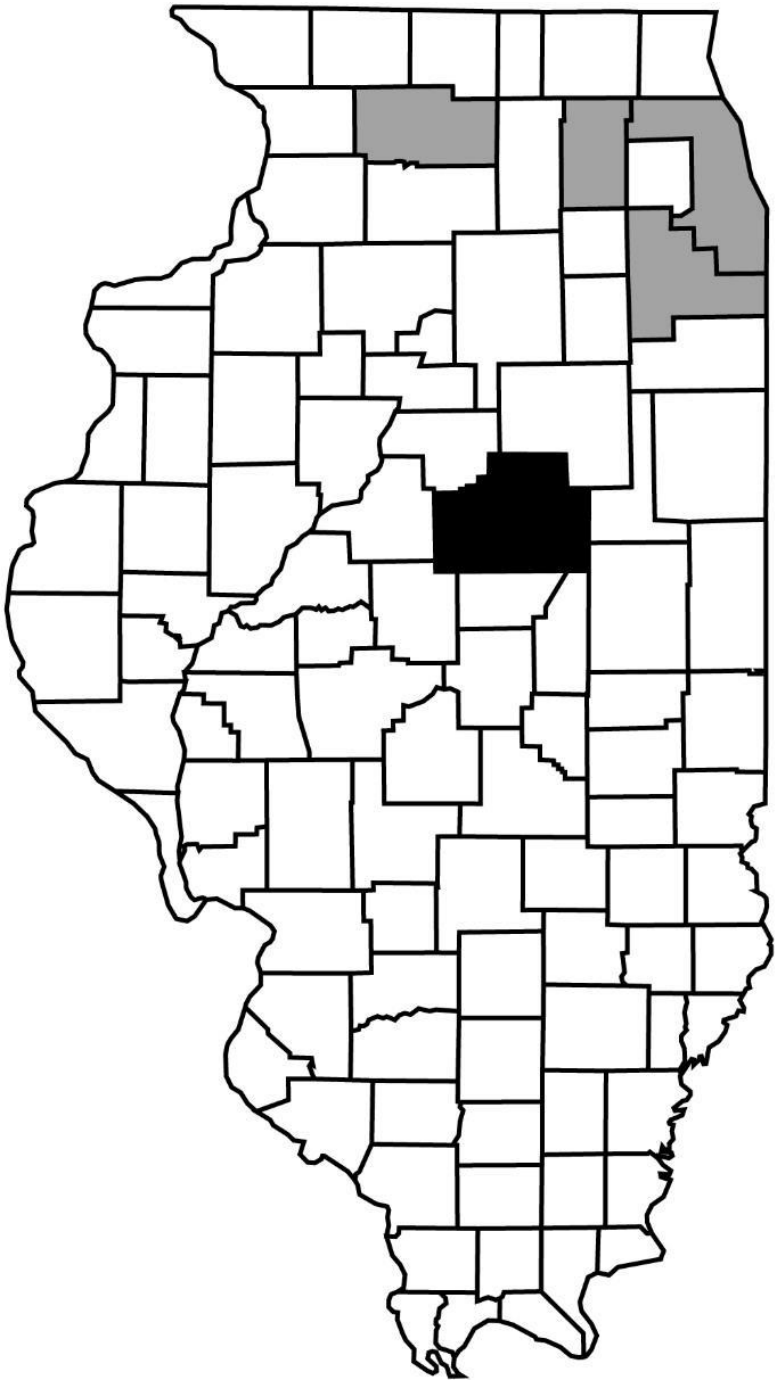


Figure 1.23. *Okanagana rimosa rimosa* (Say, 1830), (Field Museum Specimen 418846) ♂, USA:MI: Marquette County. 1-x-1956. H.S. Dybas.

Figure 1.24. Known historical distribution of *Okanagana rimosa rimosa* (Say, 1830), grey counties are those based on previous records.

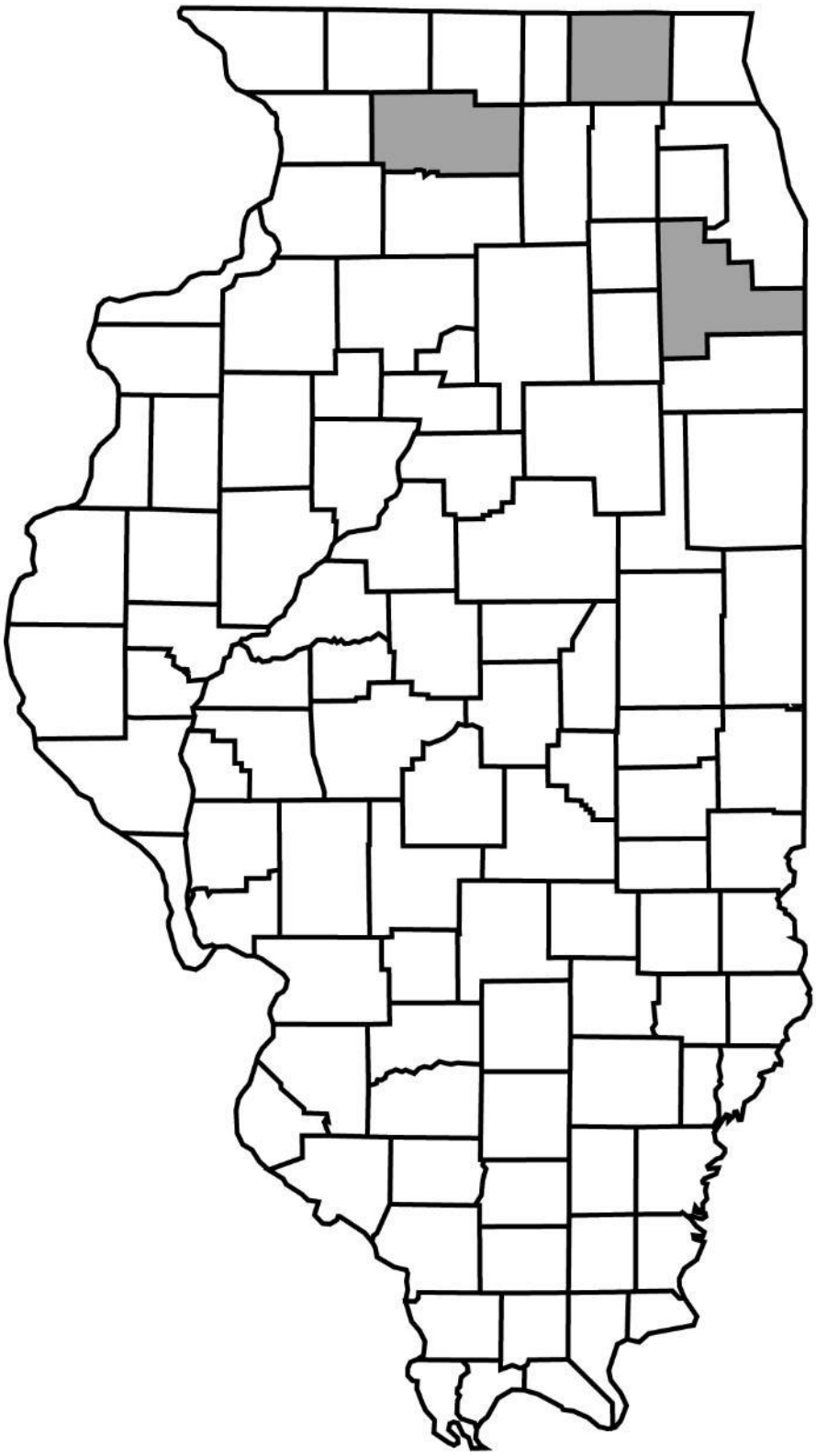


Figure 1.25. *Beameria venosa* (Uhler, 1888), (Dana Collection BV200107M) ♂, USA: IL: Monroe County: Valmeyer. Salt Lick Land and Water Reserve, “Newman Prairie”. 38.30252961°, -90.3087846°. 15.vii.2020. C.E. Dana.

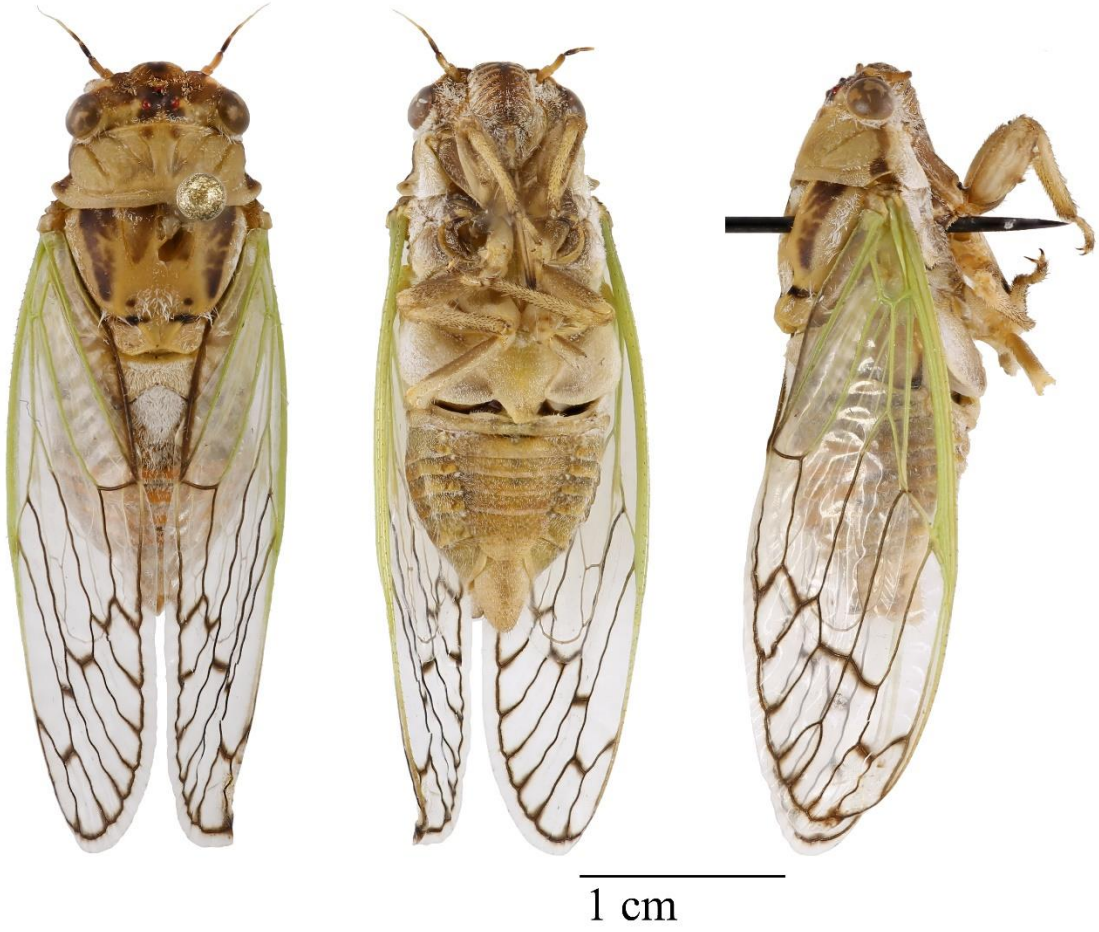


Figure 1.26. Distribution and study sites (hill prairies) of *Beameria venosa* (Uhler, 1888) in Monroe County, IL

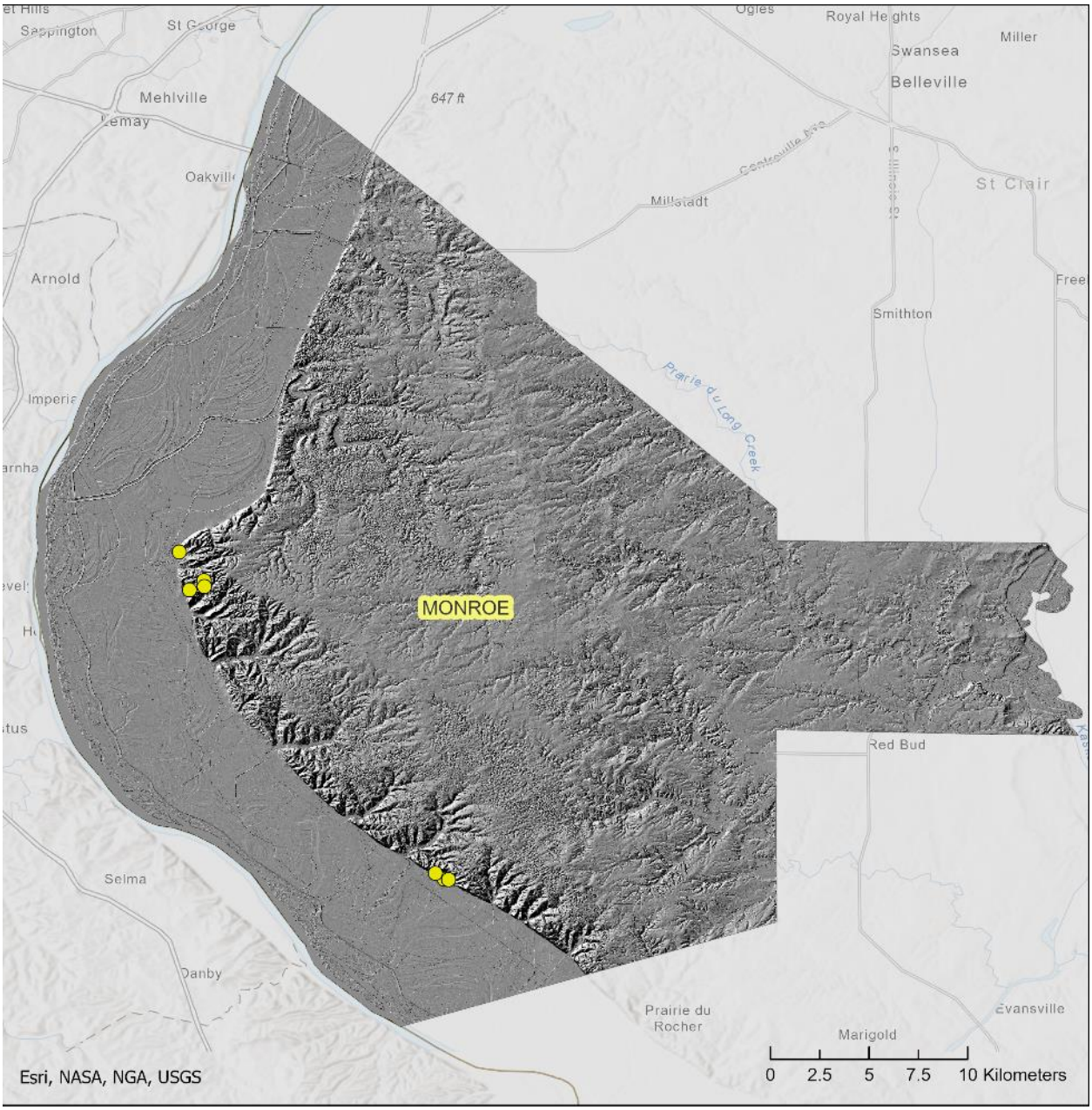


Figure 1.27. Image of *Neotibicen auriferus* (Say, 1825), (Dana Collection NAur210007M), USA: IL: Monroe County: Valmeyer. Salt Lick Land and Water Reserve, “Boyscout Prairie”. 38.30809486°, -90.30398494°. 12.ix.2021. M.J. Thomas, G.M. Lewis, J.R. Tetlie.

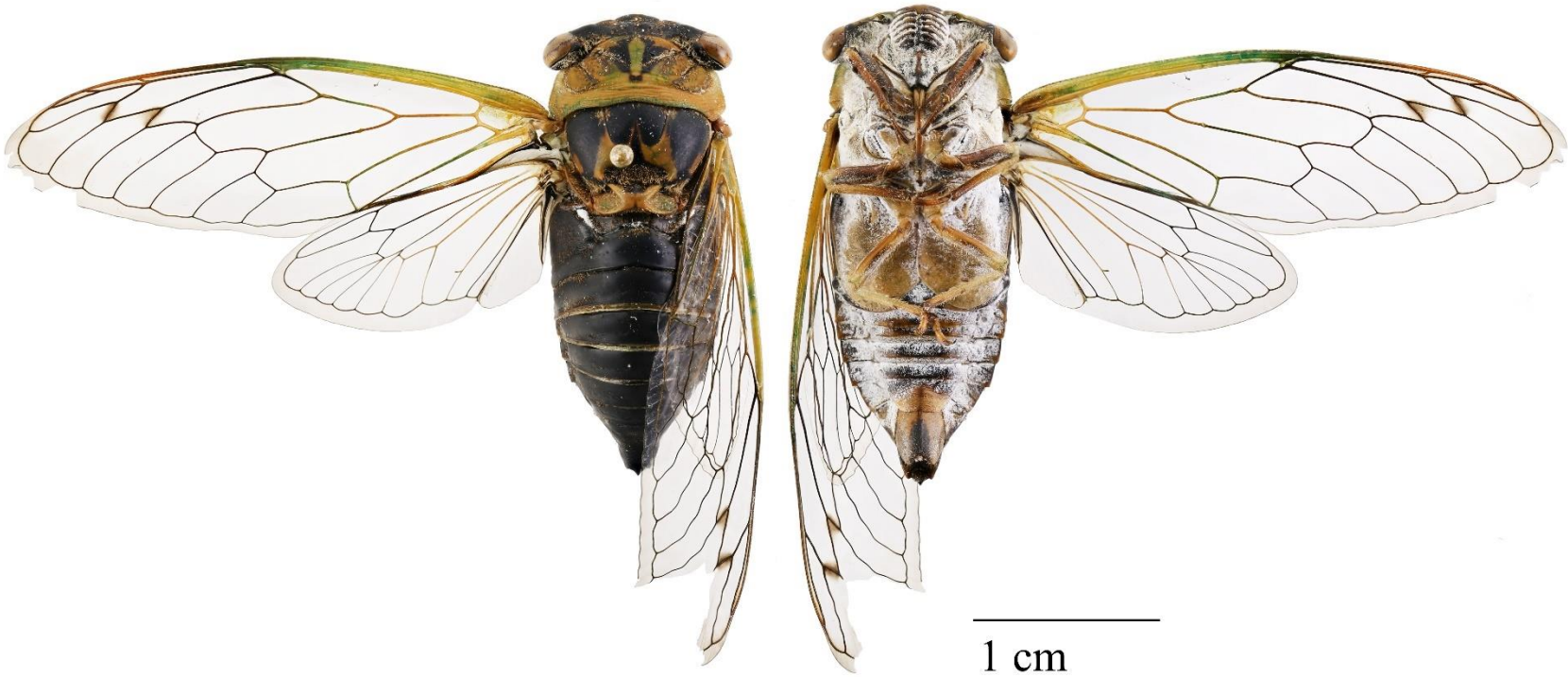


Figure 1.28. Distribution and study sites (hill prairies) of *Neotibicen auriferus* (Say, 1825).

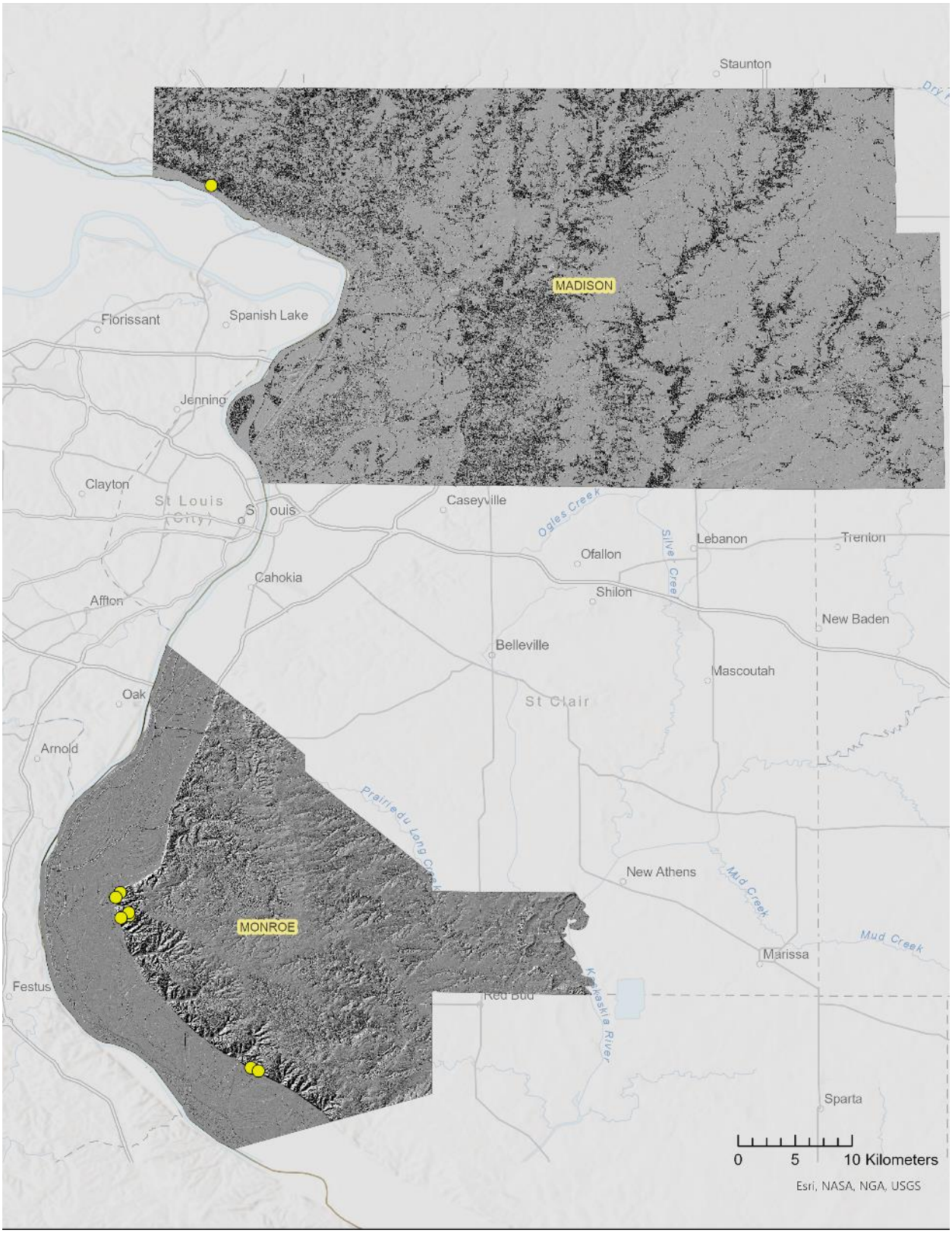


Figure 1.29. Spectrogram of male calls from (a) *Neotibicen auriferus* (Say, 1825) recorded at Fults Hill Prairie Nature Preserve in Monroe Co. (38.1576773°, -90.191327°) on 5-ix-2019 at 11:33 am and (b) *Neotibicen canicularis* recorded at Nachusa Grasslands in Lee Co. (41.878441°, -89.360444°) 23-vii-2019 at 8:41 am. Amplitude is shown in the upper window and spectrogram on the lower window. Produced using Kaleidoscope 5.4.2 software.

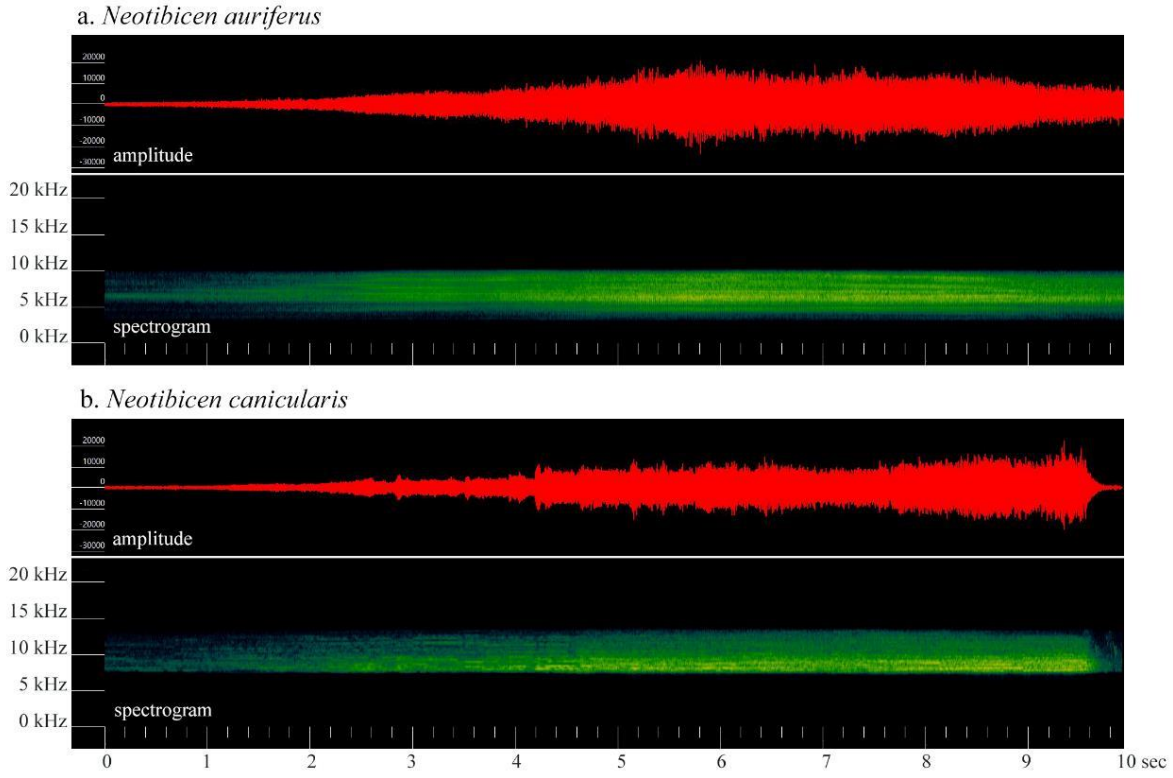


Figure 1.30. Image of *Neotibicen auriferus* (Say, 1825), female next to oviposition scar left by same female on grass stem, USA: IL: Monroe County: Valmeyer. Illinois Ozarks Nature Preserve, “Eagle Prairie”. 38.286749°, -90.3030262°. 7.ix.2019. C.E. Dana.

References

Ahern, R.G., Frank, S.D., and M.J. Raupp. 2005. Comparison of exclusion and imidacloprid for reduction of oviposition damage to young trees by periodical cicadas (Hemiptera: Cicadidae). Journal of Economic Entomology 98(6): 2133–2136.

Alexander, R.D., and T.E. Moore. 1962. The evolutionary relationships of 17-year and 13-year cicadas, and three new species (Homoptera, Cicadidae, Magicicada). Miscellaneous Publications. Museum of Zoology, University of Michigan 121: 1–59.

Alexander, R.D., Pace, A.E., and D. Otte. 1972. The singing insects of Michigan. Great Lakes Entomologist 5: 33–69.

Andersen, D.C. 1994. Are cicadas (*Diceroprocta apache*) both a" keystone" and a "critical-link" species in lower Colorado River riparian communities? The Southwestern Naturalist 39(1): 26–33.

Beamer, R.H. 1925. Notes on the oviposition of some Kansas cicadas. Annals of the Entomological Society of America 18(4): 479–482.

Beamer, R.H. 1928. Studies on the biology of Kansas Cicadidae. University of Kansas Science Bulletin 18: 155–263.

Betz, R.F., and H.F. Lamp. 1989. Species composition of old settler silt-loam prairies. Prairie Pioneers: Ecology, History and Culture: Proceedings of the Eleventh North American Prairie Conference, Lincoln, Nebraska.

Boulard, M. 2006. Acoustic signals, diversity and behaviour of cicadas (Cicadae, Hemiptera). In: Insect sounds and communication, physiology, behaviour, ecology and evolution (Drosopoulos, S. and M. F. Claridge Eds.). Boca Raton: Taylor and Francis. pp. 331–350.

Bouzat, J.L., Johnson, J.A., Toepfer, J.E., Simpson, S.A., Esker, T.L., and R.L. Westemeier. 2009. Beyond the beneficial effects of translocations as an effective tool for the genetic restoration of isolated populations. Conservation Genetics 10(1): 191–201.

Callaham, M.A., Blair, J.M., Todd, T.C., Kitchen, D.J., and M.R. Whiles. 2003. Macroinvertebrates in North American tallgrass prairie soils: effects of fire, mowing, and fertilization on density and biomass. Soil Biology and Biochemistry 35(8): 1079–1093.

Callaham, M.A., Whiles, M.R., Meyer, C.K., Brock, B.L., and R.E. Charlton. 2000. Feeding ecology and emergence production of annual cicadas (Homoptera: Cicadidae) in tallgrass prairie. Oecologia 123(4): 535–542.

Campbell, M.A., Łukasik, P., Simon, C., and J.P. McCutcheon. 2017. Idiosyncratic genome degradation in a bacterial endosymbiont of periodical cicadas. Current Biology 27(22): 3568–3575.

Campbell, M.A., Van Leuven, J.T., Meister, R.C., Carey, K.M., Simon, C., and J.P. McCutcheon. 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. Proceedings of the National Academy of Sciences 112(33): 10192–10199.

Cheung, W.W.K., and A.T. Marshall. 1973. Water and ion regulation in cicadas in relation to xylem feeding. Journal of Insect Physiology 19(9): 1801–1816.

Chicago Academy of Science. 1871. To the friends and correspondents of the Chicago Academy of Sciences. Chicago, IL. Library of Congress. https://www.loc.gov/item/rbpe.01802000/.

Claridge, M.F., Morgan, J.C., and M. S. Moulds. 1999. Substrate-transmitted acoustic signals of the primitive cicada, *Tettigarcta crinita* Distant (Hemiptera Cicadoidea, Tettigarctidae). Journal of Natural History 33(12): 1831–1834.

Cole, L.C. 1954. The population consequences of life history phenomena. The Quarterly Review of Biology 29(2): 103–137.

Cook, W.M., and R.D. Holt. 2002. Periodical cicada (*Magicicada cassini*) oviposition damage: visually impressive yet dynamically irrelevant. The American Midland Naturalist 147(2): 214–224.

Cooley, J.R. 2021. *Magicicada* broods and information. University of Connecticut. https://cicadas.uconn.edu/ [Accessed 10/8/2021]

Cooley, J.R. 2001. Long-range acoustical signals, phonotaxis, and risk in the sexual pair-forming behaviors of *Okanagana canadensis* and *O. rimosa* (Hemiptera: Cicadidae). Annals of the Entomological Society of America 94(5): 755–760.

Cooley, J.R., Kritsky, G., Edwards, M.J., Zyla, J.D., Marshall, D.C., Hill, K.B.R., Knauss, R., and C. Simon. 2009. The distribution of periodical cicada brood X in 2004. American Entomologist 55(2): 106–112.

Cooley, J.R., Kritsky, G., Marshall, D.C., Hill, K.B.R., Bunker, G., Neckermann, M.L., Yoshimura, J., Cooley, J.E., and C. Simon. 2016. A GIS-based map of periodical cicada Brood XIII in 2007, with notes on adjacent populations of Broods III and X (Hemiptera: *Magicicada* spp.). Bulletin of the Entomological Society of America 62(4): 241–246.

Cooley, J.R., and D.C. Marshall. 2001. Sexual signaling in periodical cicadas, *Magicicada* spp. (Hemiptera: Cicadidae). Behaviour 138(7): 827–855.

Cooley, J.R., Marshall, D.C., Richards, A.F., Alexander, R.D., Irwin, M.D., Coelho, J.R., and C. Simon. 2013. The distribution of periodical cicada Brood III in 1997, with special emphasis on Illinois (Hemiptera: *Magicicada* spp.). American Entomologist 59(1): 9–14.

Cooley, J.R., Marshall, D.C., and C. Simon. 2004. The historical contraction of periodical cicada Brood VII (Hemiptera: Cicadidae: *Magicicada*). Journal of the New York Entomological Society 112(2): 198–204.

Cooley, J.R., Simon, C., Marshall, D.C., Slon, K., and C. Ehrhardt. 2001. Allochronic speciation, secondary contact, and reproductive character displacement in periodical cicadas (Hemiptera: *Magicicada* spp.): genetic, morphological, and behavioural evidence. Molecular Ecology 10(3): 661–671.

Cornara, D., Marra, M., Tedone, B., Cavalieri, V., Porcelli, F., Fereres, A., Purcell, A., and M. Saponari. 2020. No evidence for cicadas' implication in *Xylella fastidiosa* epidemiology. Entomologia Generalis 20: 125–132.

CWAP. 2015. Connecticut's Wildlife Action Plan – 2015 Revision. https://portal.ct.gov/DEEP/Wildlife/CT-Wildlife-Action-Plan/CT-WAP-Current-Status [Accessed: 11/25/22]

Davis, W.T. 1910. Observations on *Cicada pruinosa* and a description of a new species. Entomological News 21: 457–458.

Davis, W.T. 1919. Cicadas of the genera *Okanagana*, *Tibicinoides* and *Okanagodes*, with descriptions of several new species. Journal of the New York Entomological Society 27: 179–223.

Davis, W.T. 1922. An annotated list of the cicadas of Virginia with description of a new species. Journal of the New York Entomological Society 30(1): 36–52.

Davis, W.T. 1925. *Cicada tibicen*, a South American species, with records and descriptions of North American cicadas. Journal of the New York Entomological Society 33(1): 35–51

Davis, W.T. 1938. New North American cicadas with notes on described species. Journal of the New York Entomological Society 46(3): 291–311.

De Geer, C. 1773. Des cigales exotiques de la quatrieme famille. Cinquième mémoire. Des cigales. Mémoires pour servir à l'histoire des Insects 3: 212–222.

Dietrich, C.H. 2003. Auchenorrhyncha (Cicadas, Spittlebugs, Leafhoppers, Treehoppers, and Planthoppers). Pp. 66–74. *In* Resh, V. H., and R. Cardé (Eds.) Encyclopedia of Insects. Elsevier Science, San Diego, CA.

Dmitriev, D.A. 2022. 3i World Auchenorrhyncha Database (D.A. Dmitriev, S. McKamey, A. Sanborn, D.M. Takiya, and J. Zahniser, Eds.; Aug 2022). Species File Group. http://dmitriev.speciesfile.org/3i.asp [Accessed 10/2/2022]

Dybas, H.S. 1969. The 17-year cicada: A four-year "mistake"? Bulletin Field Museum Natural History 40: 10–12.

Dybas, H.S., and D.D. Davis. 1962. A population census of seventeen-year periodical cicadas (Homoptera: Cicadidae: *Magicicada*). Ecology 43(3): 432–444.

Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34(1): 487–515.

Fairmaire, M.L. 1888. Coléoptères de l'intérieur de la Chine. Annales de la Société Entomologique de Belgique 32: 7–46.

Fisher, J.C. 1852. On a new species of cicada. Proceedings of the Academy of Natural Sciences of Philadelphia 5: 272–273.

Fitch, H.S. 1982. Resources of a snake community in prairie-woodland habitat of northeastern Kansas. Wildlife Research Report 13: 83–97.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondria cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–297.

Froeschner, R.C. 1952. A synopsis of the Cicadidae of Missouri (Homoptera). Journal of the New York Entomological Society 60(1): 1–14.

Germar, E.F. 1834. Observations sur plusieurs espèces du genre *Cicada*, Latr. Revue Entomologique 2: 65–66.

Gilbert, C., and C. Klass. 2006. Decrease in geographic range of the finger lakes brood (Brood VII) of the periodical cicada (Hemiptera: Cicadidae: *Magicicada* spp.). Entomologica Americana 114(1): 78–85.

Hahn, J., Herms, D.A., and D.G. McCullough. 2011. Frequently asked questions regarding potential side effects of systemic insecticides used to control emerald ash borer. University of Minnesota, Michigan State University, and The Ohio State University Extension Emerald Ash Borer Fact Sheet. 4 pp. Accessed online: 10/12/2020. https://extension.entm.purdue.edu/EAB/PDF/PotentialSideEffectsofEABInsecticidesFAQ.pdf

Hahus, S.C., and K.G. Smith. 1990. Food habits of *Blarina*, *Peromyscus*, and *Microtus* in relation to an emergence of periodical cicadas (*Magicicada*). Journal of Mammalogy 71(2): 249–252.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., and D. Goulson. 2017. More than 75% decline over 27 years in total flying insect biomass in protected areas. PloS One 12: p.e0185809. DOI: 10.1371/journal.pone.0185809

Harris, T.W. 1841. A report on the insects of Massachusetts, injurious to vegetation. Folsom, Wells, and Thurston. Cambridge, MA. Pg 175–178.

Heath, M.S. 1978. Genera of American Cicadas North of Mexico (PhD Thesis). University of Florida. DOI: 10.5962/bhl.title.42291.

Hennig, R.M., Weber, T., Moore, T.E., Huber, F., Kleindienst, H.-U., and A.V. Popov. 1994. Function of the tensor muscle in the cicada *Tibicen linnei*. Journal of Experimental Biology 187(1): 33–44.

Herms, D.A., McCullough, D.G., Smitley, D.R., Sadof, C., Miller, F.D., and W. Cranshaw. 2019. Insecticide options for protecting ash trees from emerald ash borer. North Central IPM Center Bulletin. 3rd Edition. 16 pp. Accessed online: 10/12/2020. http://www.emeraldashborer.info/documents/multistate_eab_insecticide_fact_sheet.pdf

Herrick, G.W. 1925. Manual of Injurious Insects. Henry Holt and Company, New York, USA.

Hill, K.B.R., Marshall, D.C., Moulds, M.S., and C. Simon. 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.

Horowitz, A.R., and I. Ishaaya (Eds.). 2004. Insect pest management: field and protected crops. Springer-Verlag Berlin Heidelberg.

Hunt, R.E. 1993. Role of vibrational signals in mating behavior of *Spissistilus festinus* (Homoptera: Membracidae). Annals of the Entomological Society of America 86(3): 356–361.

IWAP. 2015. 2015 Implementation Guide to the Illinois Wildlife Action Plan. https://www2.illinois.gov/dnr/conservation/IWAP/Documents/IWAPImplementationGuide2015.pdf [Accessed: 10/18/22]

Jones, M. and M. Bowles. 2013. Tree ring analysis of eastern red cedar reveals fire history of Fults Hill Prairie Nature Preserve, and its relationship to climate and loss of prairie vegetation. Morton Arboretum Publications, Lisle, Illinois.

Karban, R. 1980. Periodical cicada nymphs impose periodical oak tree wood accumulation. Nature 287(5780): 326–327.

Karban, R. 1982. Increased reproductive success at high densities and predator satiation for periodical cicadas. Ecology 63(2): 321–328.

Kaul, R.B., Kantak, G.E., and S.P Churchill. 1988. The Niobrara River Valley, a postglacial migration corridor and refugium of forest plants and animals in the grasslands of central North America. The Botanical Review 54(1): 44–81.

Koenig, W.D., and A.M. Liebhold. 2003. Regional impacts of periodical cicadas on oak radial increment. Canadian Journal of Forest Research 33(6): 1084–1089.

Kritsky, G., and R.A. Meyer. 1976. The emergence of the periodical cicada (Brood XXIll) in Illinois in 1976. Transactions of the Illinois State Academy of Science 69: 196–197.

Krohne, D.T., Couillard, T.J., and J.C. Riddle. 1991. Population responses of *Peromyscus leucopus* and *Blarina brevicauda* to emergence of periodical cicadas. American Midland Naturalist 126(2): 317–321.

Latreille, L.A. 1825. Familles naturelles du règne animal, exposées succinctement et dans un ordre analytique, avec lindication de leurs genres. 1–570. J.B. Baillière, Libraire (Paris, France).

Lee, Y.J. 2016. Description of three new genera, *Paratibicen*, *Megatibicen*, and *Ameritibicen*, of Cryptotympanini (Hemiptera: Cicadidae) and a key to their species. Journal of Asia-Pacific Biodiversity 9(4): 448–454.

Liebhold, A.M., Bohne, M.J., and R.L. Lilja. 2013. Active periodical cicada broods of the United States. USDA Forest Service Northern Research Station, Northeastern Area State and Private Forestry. https://www.fs.fed.us/foresthealth/docs/CicadaBroodStaticMap.pdf [Accessed: 11/1/2021]

Linnaeus, C. 1758. II. Hemiptera. Systema naturae: per regna tria naturae, secundum classes, ordines, genera, species cum characteribus, differentiis, synonymis, locis. Editio decima, reformata. L. Salvii. Stockholmiae 1: 1–824.

Lloyd, M., and H.S. Dybas. 1966. The periodical cicada problem. II. Evolution. Evolution 20(4): 466–505.

Lloyd, M., Kritsky, G., and C. Simon. 1983. A simple mendelian model for 13- and 17- year life cycles of periodical cicadas, with historical evidence of hybridization between them. Evolution 37: 1162–1180.

Lloyd, M., and J. White. 1987. Xylem feeding by periodical cicada nymphs on pine and grass roots, with novel suggestions for pest control in conifer plantations and orchards. The Ohio Journal of Science 87(3): 50–54.

Łukasik, P., Nazario, K., Van Leuven, J.T., Campbell, M.A., Meyer, M., Michalik, A., Pessacq, P., Simon, C., Veloso, C., and J.P. McCutcheon. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of cicadas. Proceedings of the National Academy of Sciences 115(2): E226–E235.

Luukkonen, D.R. 1987. Status and breeding ecology of the Loggerhead Shrike in Virginia. Doctoral dissertation, Virginia Tech.

Manter, J.A. 1974. Brood XI of the periodical cicada seems doomed. Memoirs, Connecticut Entomological Society 1974: 100–101.

Marlatt, C.L. 1907. The periodical cicada. Bulletin of the United States Department of Agriculture, Bureau of Entomology 71: 1–181.

Marlatt, C.L. 1923. The periodical cicada. U.S. Department of Agriculture, Bureau of Entomology Bulletin 71: 1–183.

Marshall, D.C., and J.R. Cooley. 2000. Reproductive character displacement and speciation in periodical cicadas, with description of a new species, 13-year *Magicicada neotredecim*. Evolution 54(4): 1313–1325.

Marshall, D.C., and K.B.R. Hill. 2017. A new genus for North American *Cicadetta* species (Hemiptera: Cicadidae). Zootaxa 4306(4): 537–550.

Marshall, D.C., Moulds, M., Hill, K.B.R., Price, B.W., Wade, E.J., Owen, C.L., Goemans, G., Marathe, K., Sarkar, V., Cooley, J.R., Sanborn, A.F., Kunte, K., Villet, M.H., and C. Simon. 2018. A molecular phylogeny of the cicadas (Hemiptera: Cicadidae) with a review of tribe and subfamily classification. Zootaxa 4424(1): 1–64.

McElrath, T. 2022. Illinois Natural History Survey Insect Collection. Illinois Natural History Survey. Occurrence dataset https://doi.org/10.15468/eol0pe accessed via GBIF.org on 2022-10-04.

Metcalf, Z.P. 1955. New names in the Homoptera. Journal of the Washington Academy of Sciences 45(8): 262–267.

Miller, F.D. 1997. Effects and control of periodical cicada *Magicicada septendecim* and *Magicicada cassini* oviposition injury on urban forest trees. Journal of Arboriculture 23: 225–232.

Moore, T.E. 2016. 17-Year cicadas in Michigan. Newsletter of the Michigan Entomological Society 61: 32–34.

Moriyama, M., and H. Numata. 2015. Urban soil compaction reduces cicada diversity. Zoological Letters 1(19): 1–9. DOI: 10.1186/s40851-015-0022-3.

Mussmann, S.M., Douglas, M.R., Anthonysamy, W.J.B., Davis, M.A., Simpson, S.A., Louis, W., and M.E. Douglas. 2017. Genetic rescue, the greater prairie chicken and the problem of conservation reliance in the Anthropocene. Royal Society Open Science 4(2): 160736.

MWAP. 2015. Missouri Wildlife Action Plan. Missouri Department of Conservation. https://short.mdc.mo.gov/ZqN [Accessed: 10/18/22]

Myers, J.G. 1929. Insect singers: a natural history of the cicadas. London: Routledge.

Ossiannilsson, F. 1949. Insect drummers: A study on the morphology and function of the sound-producing organ of Swedish Homoptera, Auchenorrhyncha. Opuscula Entomologica Supplementa 10: 1–146.

Ostfeld, R.S., and F. Keesing. 2000. Pulsed resources and community dynamics of consumers in terrestrial ecosystems. Trends in Ecology and Evolution 15(6): 232–237.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H., and N. Pieretti. 2011. Soundscape ecology: the science of sound in the landscape. BioScience 61(3): 203–216.

Pons, P. 2020. True cicadas (Cicadidae) as prey for the birds of the Western Palearctic: a review. Avian Research 11: 1–9.

Pringle, J.W.S. 1954. A physiological analysis of cicada song. Journal of Experimental Biology 31(4): 525–560.

Pringle, J.W.S. 1957. The structure and evolution of the organs of sound-production in cicadas. Proceedings of the Linnean Society of London 167(2): 144–159.

Purdue, J.R. 1989. Changes during the Holocene in the size of white-tailed deer (*Odocoileus virginianus*) from central Illinois. Quaternary Research 32(3): 307–316.

Ratcliffe, B.C., and P.C. Hammond. 2002. Insects and the native vegetation of Nebraska. Transactions of the Nebraska Academy of Sciences 28: 29–47.

Riley, C.V. 1892. The larger digger wasp. Insect Life 4: 248–252.

Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., and P.P. Marra. 2019. Decline of the North American avifauna. Science 366(6461): 120–124.

Rosenberg, K.V., Ohmart, R.D., and B.W. Anderson. 1982. Community organization of riparian breeding birds: response to an annual resource peak. The Auk 99: 260–274.

Saljoqi, A.-U.-R., Rahimi, M.S., Rehman, S., and I.A. Khan. 2010. Population density of cicada, *Tibicen* spp. (Homoptera: Cicadidae) in grape vineyards in Northern Plains of Afghanistan. Sarhad Journal of Agriculture 26: 75–78.

Sanborn, A.F., and M.S. Heath. 2016. *Megatibicen* n. gen., a new North American cicada genus (Hemiptera: Cicadidae: Cicadinae: Cryptotympanini). Zootaxa 4168(3): 577–582.

Sanborn, A.F., and P.K. Phillips. 1995. Scaling of sound pressure level and body size in cicadas (Homoptera: Cicadidae; Tibicinidae). Annals of the Entomological Society of America 88(4): 479–484. https://doi.org/10.1093/aesa/88.4.479

Sanborn, A.F., and P.K. Phillips. 2013. Biogeography of the cicadas (Hemiptera: Cicadidae) of North America, north of Mexico. Diversity 5(2): 166–239.

Say, T. 1825. Descriptions of new hemipterous insects, collected in the expedition to the Rocky Mountains, performed by the order of Mr. Calhoun, Secretary of War, under command of Major Long. Journal of the Natural Academy of Sciences of Philadelphia 4(2): 307–345.

Say, T. 1830. Descriptions of new North American hemipterous insects, belonging to the first family of the section Homoptera of Latreille. Journal of the Academy of Natural Sciences of Philadelphia 6: 235–244.

Schwegman, J.E., Fell, G.B., Hutchison, M.D., Paulson, G., Shepard, W.M., and J. White. 1973. Comprehensive plan for the Illinois Nature Preserves System. Part 2. The Natural Divisions of Illinois. Illinois Nature Preserves Commission (Rockford, IL). 32 pp. http://hdl.handle.net/2142/17165

Shaw, K.C. 1976. Sounds and associated behavior of *Agallia constricta* and *Agalliopsis novella* (Homoptera: Auchenorrhyncha: Cicadellidae). Journal of the Kansas Entomological Society 49(1): 1–17.

Shelley, R.M., and W.D. Sissom. 1995. Distributions of the scorpions *Centruroides vittatus* (Say) and *Centruroides hentzi* (Banks) in the United States and Mexico (Scorpiones, Buthidae). The Journal of Arachnology 23(2): 100–110.

Simmons, P., and D. Young. 1978. The tymbal mechanism and song patterns of the bladder cicada, *Cystosoma saundersii*. Journal of Experimental Biology 76(1): 27–45.

Simon, C. 1988. Evolution of 13- and 17-year periodical cicadas (Homoptera: Cicadidae: *Magicicada*). Bulletin of the Entomological Society of America 34(4): 163–176.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved PCR primers. Annals of the Entomological Society of America 87: 651–701.

Simon, C., Cooley, J.R., Karban, R., and T. Sota. 2022. Advances in the evolution and ecology of 13- and 17-year periodical cicadas. Annual Review of Entomology 67: 457–482.

Smith, C.F., Schuett, G.W., Reiserer, R.S., Dana, C.E., Collyer, M.L., and M.A. Davis. 2019. Drought-induced suppression of female fecundity in a capital breeder. Scientific Reports 9(1): 1–13.

Smith, D.M., Kelly, J.F., and D.M. Finch. 2006. Cicada emergence in southwestern riparian forest: influences of wildfire and vegetation composition. Ecological Applications 16(4): 1608–1618.

Smith, J.B., and J.A. Grossbeck. 1907. Studies in certain cicada species. Entomological News and Proceedings of the Entomological Section of the Academy of Natural Sciences Philadelphia 18: 116–129.

Smith, P.W. 1961. The amphibians and reptiles of Illinois. Illinois Natural History Survey Bulletin 28(1): 1–298.

Soper, R.S., Delyzer, A.J., and L.F. Smith. 1976. The genus *Massospora* entomopathogenic for cicadas. Part. II. Biology of *Massospora levispora* and its host *Okanagana rimosa*, with notes on *Massospora cicadina* on the periodical cicadas. Annals of the Entomological Society of America 69(1): 89–95.

Sorg, M., Schwan, H., Stenmans, W., and A. Muller. 2013. Ermittlung der Biomassen flugaktiver Insekten im Naturschutzgebiet Orbroicher Bruch mit Malaise Fallen in den Jahren 1989 und 2013. Mitteilungen aus dem Entomologischen Verein Krefeld: 1–5.

Sota, T., Yamamoto, S., Cooley, J.R., Hill, K.B.R., Simon, C., and J. Yoshimura. 2013. Independent divergence of 13- and 17-y life cycles among three periodical cicada lineages. Proceedings of the National Academy of Sciences 110(17): 6919–6924.

Stannard, L.J. 1975. The distribution of periodical cicadas in Illinois. Illinois Natural History Survey Biological Notes 91: 3–12.

Stölting, H., Moore, T.E., and R. Lakes-Harlan. 2002. Substrate vibrations during acoustic signalling in the cicada *Okanagana rimosa*. Journal of Insect Science 2(2): 1–7.

Storm, J.L., and J.O. Whitaker, Jr. 2007. Food habits of mammals during an emergence of 17-year cicadas (Hemiptera: Cicadidae: *Magicicada* spp.). Proceedings of the Indiana Academy of Science 116(2): 196–199.

Strang, C.A. 2022. Singing insects of the Chicago region: A guide to crickets, katydids, grasshoppers and cicadas with audible displays. https://natureinquiries.wordpress.com/ [Accessed: 9/26/2022]

Sueur, J., and T. Aubin. 2004. Acoustic signals in cicada courtship behaviour (order Hemiptera, genus *Tibicina*). Journal of Zoology 262(3): 217–224.

Surface, H.A. 1906. The serpents of Pennsylvania. Pennsylvania Department of Agriculture Monthly Bulletin of the Division of Zoology 4: 113–201.

Tumlison, R. 2013. First record of the aridland cicada, *Beameria venosa*, in Arkansas. The Southwestern Naturalist 58(3): 384–386.

Uhler, P.R. 1888. Preliminary survey of the Cicadaea of the United States, Antilles, and Mexico. Entomologica Americana 4(5): 81–85.

Van Leuven, J.T., Meister, R.C., Simon, C., and J.P. McCutcheon. 2014. Sympatric speciation in a bacterial endosymbiont results in two genomes with the functionality of one. Cell 158(6): 1270–1280.

Villet, M. 1987. Sound pressure levels of some African cicadas (Homoptera: Cicadoidea). Journal of the Entomological Society of Southern Africa 50(2): 269–273.

Walker, F. 1850. List of the specimens of homopterous insects in the collection of the British Museum. Order of Trustees. London, England. Part 1: 1–260.

Wallner, A.M. 2011. Evaluating North American tallgrass prairie quality using the Auchenorrhyncha quality index (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Walsh, B.D., and C.V. Riley. 1868. The periodical cicada. American Entomologist 1(4): 63–72.

White, E.G., and J.R. Sedcole. 1993. A study of the abundance and patchiness of cicada nymphs (Homoptera: Tibicinidae) in a New Zealand subalpine shrub-grassland. New Zealand Journal of Zoology 20(1): 38–51.

White, J., and C.E. Strehl. 1978. Xylem feeding by periodical cicada nymphs on tree roots. Ecological Entomology 3(4): 323–327.

White, J., Lloyd, M., and R. Karban. 1982. Why don't periodical cicadas normally live in coniferous forests? Environmental Entomology 11(2): 475–482.

Williams, K.S., and C. Simon. 1995. The ecology, behavior, and evolution of periodical cicadas. Annual Review of Entomology 40(1): 269–295.

Wolverton, S. 2005. The effects of the hypsithermal on prehistoric foraging efficiency in Missouri. American Antiquity 70(1): 91–106.

Yang, L.H. 2004. Periodical cicadas as resource pulses in North American forests. Science 306(5701): 1565–1567.

Yang, L.H. 2006. Periodical cicadas use light for oviposition site selection. Proceedings of the Royal Society of London B: Biological Sciences 273(1604): 2993–3000.

Young, D., and H. Bennet-Clark. 1995. The role of the tymbal in cicada sound production. The Journal of Experimental Biology 198(4): 1001–1020.

CHAPTER 2: LOW HABITAT QUALITY RAILROAD RIGHTS-OF-WAY ACT AS RESERVOIRS FOR GENETIC DIVERSITY IN PRAIRIE-ASSOCIATED CICADAS

Abstract

The Illinois landscape is dominated by agriculture and has very little original prairie remaining. Right-of-way prairies that occur along railroads represent a largely unexplored refuge for rare or threatened animals and plants, like the prairie-associated cicada *Megatibicen dorsatus* (Hemiptera: Cicadidae). I non-lethally collected *M. dorsatus* DNA samples from both railroad rights-of-way and high quality nature preserves throughout Illinois. Despite this species having a large genome, I was able to successfully create a double digest restriction-site associated DNA sequence library for 452 individuals with average 19.8x coverage. My results indicate that there is some contiguity along the railroad rights-of-way, but that populations at nearby high quality nature preserves do not always have significant gene flow occurring with surrounding areas. There is some isolation by distance occurring, but there are also other factors impeding gene flow. I discuss potential hypotheses for inconsistent gene flow, including agriculture, edge effects, and historical railway continuity.

Keywords:

Cemetery prairies, corridors, grasslands, habitat fragmentation, restoration, population genetics

Introduction

For many endangered ecosystems, habitat loss and fragmentation has restricted remaining areas to small patches that are often isolated or that follow features of the landscape. The ability of these remnants to support viable populations may depend on connectivity among patches via corridors that facilitate dispersal and gene flow. For example, rights-of-way (ROWs) that follow features such as railroad tracks can cover as much as 2.5% of a region's land area (Huijiser and Clevenger 2006). These linear habitats can serve as corridors and even refugia for native species of plants but there is less information on how animals navigate and utilize these spaces (Leston and Koper 2016; Reed and Schwarzmeier 1975). However, due to their high edge to interior ratios, many of these relicts have been degraded via invasive species, development, and neglect (Bolin et al. 1988).

Given changes in utilized modes of transportation, many railroad lines have become abandoned as well, halting the management of invasive and woody plants along unused tracks (Hyman and Manley 1977). What is truly unique about some of these relict prairies is the absence of substantial soil disturbance through tilling or other anthropogenic means. Soil disturbance can have unexpected bimodal consequences as plant species richness can be increased with disturbance, especially in prairies, but larger disturbance events can reduce species diversity (Collins and Barber 1985; Hughes et al. 2007). Fires provide disturbance regimes necessary for the maintenance of prairies and prevent woody encroachment. In recent years management using fire has all but ceased along railways and been replaced by herbicide treatment (sometimes targeted), mowing, and tilling (Harrington and Leach 1989; Terry 2018). Fires can help prevent woody encroachment into prairies. Prior to 1980, most wildfires in Wisconsin were started by railways (Harrington and Leach 1989) with some exceptions (Wolf

2004), and many wildfires in Illinois were also started by locomotives (Jones and Bowles 2013; Jones and Bowles 2016) as well as other human activities (McClain et al. 2021).

Before European settlement, over 60% of Illinois was covered by prairie (Anderson 1970; Robertson and Schwartz 1994). Now over 75% of Illinois' total land is under agriculture and very little native habitat remains (Iverson 1988). Other threats to natural areas in the Midwest include mowing, spraying, and degradation of roadside and hedgerow habitats for vector and pest control; if maintained properly, these areas might otherwise serve as corridors (IWAP 2015). Conversion of natural areas to agriculture over the past 150 years, subsequent cropland abandonment, and fire suppression (Ramankutty and Foley 1999) have resulted in a landscape with less than 0.07% of original or remnant prairie (White 1978). In fact, over 83% of the prairies identified by the Illinois Natural Areas Inventory are less than 10 acres in size (White 1978; Robertson and Schwartz 1994). All of these factors make Illinois a uniquely fragmented landscape with limited natural corridors for gene flow.

Megatibicen dorsatus (Say, 1825)[1] commonly called the "bush cicada" or the "grand western cicada", is found in prairies throughout the central region of North America. It has a loud, characteristic call that has been described as a "tractor-like rattle" (Hill et al. 2015). Beamer (1928) even reported that some female cicadas had been attracted to the sound of a running tractor and landed on the cab. Towards the western range in Texas, Oklahoma, Nebraska, Kansas, and Colorado, *Me. dorsatus* can co-occur with the cryptic species *Megatibicen tremulus* (Cole 2008) and can be distinguished by song. Unlike most other cicadas

[1] Recent synonyms for *Megatibicen dorsatus* (Say 1825) include *Tibicen dorsatus, Ameritibicen dorsatus,* and *Neotibicen dorsatus*. *Ameritibicen* Lee 2016 was erected the same year as *Megatibicen* Sanborn and Heath 2016, but based on the International Commission on Zoological Nomenclature (ICZN) principal of priority, *Megatibicen* was retained (Sanborn and Heath 2017).

in the group previously united under *Neotibicen* that oviposit and inhabit trees, both coniferous and deciduous, *Me. dorsatus* is associated with more herbaceous plants and shrubs (Froeschner 1952; Hill et al. 2015; Sanborn and Philips 2013). I have observed it ovipositing in gray-headed coneflower, *Ratibida pinnata* (Asteraceae), fragrant sumac, *Rhus aromatica* (Anacardiaceae), and mulberry, *Morus rubra* (Moraceae) (Dana *pers. obs.*). Sanborn and Phillips (2013) list a variety of other associated plants across its rangeland including sagebrush, *Artemesia* spp. (Asteraceae), and squawbush sumac, *Rhus trilobata* (Anacardiaceae). There are also records of association with goldenrod, *Solidago* (Asteraceae) (Froeschner 1952), and a spurious record of oviposition in elm, *Ulmas* (Ulmaceae) (Hoffmann 1942). As a result of these plant associations, this species is thought to be an intact prairie specialist and thus of conservation concern, given substantial losses of this unique habitat type.

The grand western cicada's prairie habitat is severely threatened in Illinois and its known range is steadily decreasing in area such that many populations are no longer contiguous and have likely been disconnected for several decades. The primary barrier to gene flow in Illinois is row-crop agriculture. In the last half-decade, populations were discovered along railroad rights-of-way (ROW) prairies in Champaign, Ford, and Iroquois counties (near Loda Cemetery Prairie Nature Preserve and Prospect Cemetery Nature Preserve) and Jasper/Marion counties (near Prairie Ridge State Natural Area). To assess the potential for species to recolonize restored prairies and other natural areas, I identified the current geographical distribution of the habitat specialist *M. dorsatus* and evaluated the potential of corridors to facilitate dispersal by looking at the distribution of polymorphisms across the genome using double digest restriction-site associated DNA sequencing (ddRADSeq).

Materials and Methods

Collection methods

Megatibicen dorsatus individuals were collected by insect net at prairie locations throughout the state (Figure 2.1; Appendix Table B.1) starting in 2015 and during the months of July and August. A voucher pair was taken whole to document populations and will be accessioned into long term collections at the Illinois Natural History Survey (INHS). Given the large body size and desire to minimize impact on the population, one hind leg was sufficient for DNA needs. Individual legs were clipped across the femur using flame-sterilized scissors before being released at the same prairie. Legs were placed dry in 1.5 mL centrifuge tubes and kept on ice until they could be stored at -20°C long term at the INHS. Attempts were made to collect at least 12 individuals of equal male:female ratio from a site, but this was not always possible given the difficulty in collecting. Three sites were chosen for annual collections: a right-of-way prairie south of the town of Paxton Illinois (40.429590°, -88.108767°), Loda Cemetery Prairie Nature Preserve (40.527315°, -88.075882°), and Henry Allan Gleason Nature Preserve (40.379807°, -89.929928°). Sites were chosen based on records in the INHS Insect Collection, word of mouth, and audio surveys of railroad prairies.

DNA extraction

DNA was extracted from cicada legs using QIAGEN DNEasy® Blood and Tissue Kits (QIAGEN Inc., Germantown, MD) with several modifications to the kit methods. Half of a leg clipping was used for each extraction, cut in half using a sterile razor blade on a fresh piece of parafilm. I used sterile plastic pellet pestles (Thermo Fisher Scientific, Waltham, MA) to grind muscle tissue out of the chitinous exoskeleton of cicada legs as an additional step to maximize

the amount of DNA extracted. Samples were incubated in ATL Buffer and Proteinase K at 56°C for 18–24 hours with additional vortexing during this period and overnight DNA precipitation at 4°C after the addition of chilled 100% ethanol. DNA was eluted from QIAGEN columns in 80 to 100 µl of warmed AE buffer. DNA concentration was measured using the High Sensitivity dsDNA Invitrogen™ Qubit™ 3 Fluorometer system (Invitrogen, Thermo Fisher Scientific, Waltham, MA). If the concentration of the DNA was measured to be lower than 13 ng/µl, DNA was concentrated using sodium acetate and isopropanol precipitation and resuspension. The overall amount of DNA needed for the next steps was 250 ng for each specimen.

In silico testing

To determine the number of samples that can be run on an Illumina flowcell, an estimate of genome size is needed to determine how many unique fragments will occur in an individual. Size ranges were chosen based on estimates from FRAGMATIC (Chafin et al. 2018) using MspI and PstI enzymes and genome assemblies for two cicada species *Magicicada septendecim* (Linnaeus, 1758) (NCBI GCA_011326945.1) and *Ma. septendecula* Alexander & Moore, 1962 (NCBI GCA_011763675.1). The genome size estimates for these two species assemblies are 1.6 Gbps, much smaller than the estimates for cicada species more closely related to *Me. dorsatus*. The genome size of *Neotibicen lyricen* (De Geer, 1773) was estimated between 6.7 and 7 Gbps, while the more closely related *Megatibicen resh* (Haldeman, 1852) was smaller at 5.2 to 5.7 Gbps (Hanrahan and Johnston 2011). To account for this large potential difference in genome size and avoid loss in coverage, the number of estimated fragments (using *Magicicada* reference genomes) was multiplied by a factor of 3. With an estimated 3.5 billion reads available from sequencing, along with 452 samples, a size range of 300 to 600 bp was chosen for size selection

of fragments (rounded up to 370 to 700 bp to account for adapters). More details on the calculations can be found in the Appendix (Table B.2).

ddRADSeq library preparation

Library preparation methods were modified from the Clark et al. (2014) and an OpenWetWare protocol (2017). 250 ng of genomic DNA per sample was used in library preparation. Genomic DNA was digested for three hours at 37°C using the enzymes PstI and MspI and heat inactivated at 80°C for 20 minutes. PstI and MspI adapters were ligated to fragmented DNA for two hours at 25°C and T4 ligase enzymes (New England Biolabs) were heat inactivated at 65°C for 20 minutes. For both enzyme heat inactivations, thermocyclers were stepped down gradually back to 20°C, no more than 3°C decrease every 90 seconds. Samples were then pooled by plate, cleaned, and then concentrated using the QIAquick® PCR Purification kit (QIAGEN Inc., Germantown, MD). Fragment size was selected using 1.5% Agarose Gel Cassettes with internal standards on a BluePippin machine (Sage Science, Inc., Beverly, MA) in the size range 370 to 700 bp (without adapters the range would be 300 to 600 bp). After size selection fragments were amplified using Illumina indexing primers (Appendix Table B.4). Four PCRs were run in parallel and pooled to avoid possible amplicon bias (98°C 30 seconds; 15 cycles of 98°C 10 seconds, 65°C 30 seconds, 72°C 30 seconds; 72°C 5 minutes) using Kappa Hi-Fi DNA Polymerase Master Mix (Roche Sequencing and Life Science, Wilmington, MA). Pooled PCR products were cleaned once more using QIAQuick PCR Cleanup kits. To remove primer-dimer products a secondary size selection was run on the BluePippin and the increase in size range due to base pairs added from primers was accounted for (new range 370 to 835 bp). Plates were combined in equal ratios based on DNA concentration and the

individual sample number per plate. Adapter sequences (including barcodes) and Illumina sequence primers can be found in the Appendix (Table B.3).

Sequencing

To ensure even coverage across 96-well plates, a MiSeq titration run was completed prior to more in depth sequencing. Based on this run, additional amounts of plate ngs were spiked-in to the pool and this newly created, secondary pool was sequenced in greater depth. Quality control in the form of fragment size assessment using a Fragment Analyzer (Advanced Analytical Technologies, Inc.) ensured that any primer sequences were removed. Sequencing was provided by the Carver Center (University of Illinois at Urbana-Champaign) using the Illumina NovaSeq 6000 S4 flowcell producing paired 150 bp reads.

Sequence processing and analysis

Due to the use of variable length MspI adapters (Appendix B.3), base pairs needed to be trimmed prior to use of the Stacks pipeline. Reverse reads were run through cutadapt 2.10 to remove any excess base pairs from the 5' end (i.e., any that were prior to the CGG cut site) from variable length adapters used in library preparation (Martin 2011). To identify common loci across samples for population genetics analysis I used Stacks v2.54 (Catchen et al. 2011; Catchen et al. 2013). In order to optimize the *de novo* assembly of loci in the Stacks program I selected the parameters based on trial runs using 15 randomly selected individuals to maximize the number of SNPs and loci available for further analysis (Paris et al. 2017; Rochette and Catchen 2017). Further filtering was done using VCFTools 0.1.16 (Danecek et al. 2011) and

based on poor coverage (more than 50% missing data), an additional set of individuals was removed (Puritz et al. 2014a; Puritz et al. 2014b).

The maximum number of loci possible was output as a STRUCTURE file. Given the large size of the dataset, FastSTRUCTURE was used to get an idea of the best value for k (Raj et al. 2014). STRUCTURE 2.3.4 (Settings: Admixture Model, Length of Burnin Period: 1,000, Number of MCMC Reps after Burnin: 10,000) to determine k (clusters/populations) and better visualize data. Statistics were visualized using STRUCTURE HARVESTER (Earl and VonHoldt 2012). After output from Stacks, analysis using STRUCTURE, and output to vcf results results were visualized using R 4.2.1 in RStudio (2022) using multiple packages, including vcfR 1.13.0 (Knaus and Grünwald 2016; Knaus and Grünwald 2017) and poppr 2.9.3 (Kamvar et al. 2014; Kamvar et al. 2015). Pairwise F_{ST} values were calculated using hierfstat 0.5-11 (Weir and Cockerham 1984). A Mantel test was performed in R Studio using adegenet 2.1.8 (Jombart 2008; Jombart and Ahmed 2011).

Examples of code for multiple parts of this pipeline used can be found in Appendix B.

Results

Adult *Megatibicen dorsatus* were found at many previously reported locations (McElrath 2022) and several new locations (e.g., Eldon Hazlet State Recreation Area, Carlyle, IL) (Appendix Table B.1). Beamer (1928) reported that the emergence and activity of *Me. dorsatus* in Kansas was between the end of June to the end of September. However, in Illinois, my earliest date of collection was July 17th and the latest was September 16th. The furthest distance between collection localities was 240 km between railroad ROW prairie “South of Buckley” on U.S. Route 45 and Eldon Hazlet State Recreation Area (Figure 2.2). Complications arising from

permits, weather, and scheduling resulted in missing an annual collection at one site (Loda Cemetery Prairie Nature Preserve) and lower sample size at another (Table 2.1). Some individuals were also removed due to low coverage or other quality reasons (e.g., poor metadata quality).

The prepared library was sequenced at the Carver Center (University of Illinois at Urbana-Champaign) using Illumina NovaSeq 6000 S4 flowcell to produce paired 150 bp reads. This NovaSeq run was shared with another project and produced a total of 5.72 billion reads, of which 3.94 billion were of the *Megatibicen dorsatus* library. After processing using STACKS (removing low quality reads and assigning reads to individual IDs based on barcodes), an average of 8.50 million reads (95% CI: 8,139,063 – 8,869,307) were retained for each of the 452 individuals (Appendix Table B.1). Individuals with less than one million reads were discarded (n = 8). *de novo map* (denovo_map.pl) ideal parameters were used on the remaining individuals (m = 3, M = 2, n = 3) according to a set of trials run on a random subset of individuals. Of the resulting 2,625,687 loci, the effective coverage had a mean of 19.8x (σ = 6.5, min = 5.2x, max = 46.3). The gstacks pipeline results were visualized in R to show effective coverage and other metrics (Figure 2.3). After filtering with VCF tools to remove individuals with greater than 50% missing data, 413 individuals were retained for analysis.

fastSTRUCTURE revealed that when testing between 2 and 30 the value for K to maximize marginal likelihood was 4 (Figure 2.4) and that the best value to explain structure in data was 13 (Figure 2.5). STRUCTURE plots were generated (Figure 2.7; Figure 2.15) and best K was chosen based on STRUCTURE HARVESTER (Table 2.2; Figure 2.6). The best K according to the Evano method was 6, contrary to what was chosen by fastSTRUCTURE.

Principal components analysis (PCA) plots showed nonoverlapping ellipses (99% samples) among several populations. Both Loda Cemetery Prairie Nature Preserve and Henry Allan Gleason Nature Preserve had very distinctive ellipses (Figure 2.8 and Figure 2.9). The circle that overlapped slightly with Loda was from a rest stop prairie (Main Line Station Rest Stop Northbound on the 57 Highway) that was 2.8 km away, although this was a small sample size due to minimal population size at the rest stop prairie. It appears there is some structure from geographic locality by looking at latitude and longitude (Figure 2.10 and Figure 2.11).

Isolation by distance (IBD) was tested using a Mantel test and the observation value was 0.253568 based on 999 replicates ($p = 0.085$) (Figure 2.12) indicating that our observation did not significantly show an effect of distance on population structure, although there was a positive correlation in the overall relationship between geographic distance and genetic distance when viewed as a scatterplot (Figure 2.13; Figure 2.14).

Pairwise F_{ST} values calculated using hierfstat 0.5-11 (Weir and Cockerham 1984) can be seen in Table 2.3 for sampled locations. The sample comparison with the highest pairwise value was LODA-GLEASON at a F_{ST} value of 0.1045 and separated by a distance of 158 km. The sample comparison with the lowest pairwise value was TRACT5-NKIN with a F_{ST} value of 0.0049 and separated by a distance of 4.8 km. Not all populations separated by under 10 km had similar pairwise values however, as LODA-BUCKLEY have a pairwise F_{ST} value of 0.0795 and are separated by 6.5 km. Pairwise F_{ST} values were also calculated for a k of 6 (Table 2.4) and the groups can be visualized, along with their structure plots in Figure 2.15. The lowest F_{ST} value of 0.0162 is seen between Groups R (light blue, including the 12-Mile Tracts, ELDON, and assorted other locations in Southern Illinois and V (green, the southern half of the US-45 samples, including LUDLOW and SPAX, as well as PROSPECT). LODA again shows higher

divergence from other populations, especially compared again to GLEASON (F_{ST} = 0.1034) and RANKIN (F_{ST} = 0.1005).

Discussion

Across 24 locations in Illinois, my analysis revealed six populations: (1) Loda Cemetery Prairie Nature Preserve (LODA), (2) Henry Allan Gleason Nature Preserve (GLEASON), (3) South of Buckley Canadian National railroad right-of-way prairie along US-45 (BUCKLEY), (4) right-of-way prairies also along US-45 but south of Paxton (SPAX and LUDLOW), (5) Rankin Union Pacific railroad right-of-way prairie (RANKIN), and (6) southern Illinois sites including Eldon Hazlet State Recreation Area (ELDON) and right-of-way prairies along IL-37 called "12 Mile Prairie" (Figure 2.15). Two locations in particular stood out as distinctive from the rest of the samples, LODA and GLEASON, both in PCA and STRUCTURE Plots (Figure 2.4; Figure 2.5; Figure 2.7; Figure 2.8; Figure 2.9; Figure 2.15). GLEASON is a distinct habitat from the others sampled: a sand prairie with sand dunes, large bushes of *Rhus aromatica* (fragrant sumac), and *Opuntia humifusa* (prickly pear cactus) (Hart and Gleason 1907; McClain et al. 2005). LODA on the other hand is a heavily managed black soil prairie that is part of the Grand Prairie Section of the natural divisions of Illinois (Schwegman 1997) (Figure 2.16). LODA is 6.5 km from BUCKLEY, 9.3 km from Prospect Cemetery Prairie Nature Preserve (PROSPECT), and 7 km from SPAX. GLEASON is 19.5 km from Long Branch Sand Prairie Nature Preserve (LONGBRANCH), 196 km from ELDON, 158 km from LODA, and 194 km Tract 1 of 12 Mile Prairie (TRACT1) – all substantial distances compared to LODA and surrounding populations.

As discussed in Chapter 1, Illinois lies at the extreme east edge of the range of *Megatibicen dorsatus* distributional range within the central United States. Indeed, the RANKIN

site is the furthest east site that I have been able to locate *Me. dorsatus,* and BUCKLEY is the furthest north site. If these are indeed the furthest north and east within Illinois and perhaps across the entire range, these localities represent the outermost edge. Perhaps the reason that they are so distinct is because these populations are on the outmost range of the distribution. This effect has been seen in other systems before (Assis et al. 2013). Eckert et al. (2008) propose that as you reach the range edges, genetic drift plays more of a role than gene flow in population genetic structuring.

Railroad ROWs along US-45 and 12 Mile Prairie (i.e., IL-37) are on a Canadian National (CN) line and are ultimately connected (Figure 2.17), although there is a large amount of development along this corridor that likely blocks gene flow. The BUCKLEY population is quite distinct, despite only being 6.5 km from LODA and 17.2 km from the US45 South of Paxton (SPAX) ROW population. The RANKIN ROW prairie also comes out as distinct and is not part of the same rail system as the other ROW samples; it is instead along a Union Pacific railroad line. Given issues getting permissions and lack of continuity, we only have collections from directly along the county roads. I also attempted to add to this sample by collecting at nearby Pellsville Cemetery Prairie and connecting Herschel Pheasant Habitat Area, but the cicada populations were not well established in these prairies. The few samples that we did collect from Pellsville showed highest level of similarity to the populations along the southern half of US-45, not to nearby Rankin ROW.

Over the years of this study there has been increased woody and invasive encroachment at the "south of Paxton" railroad prairie – such that a large portion of the prairie is now shaded by sumac. I also noted that during the removal of some of the trees and other foliage at 12 Mile Prairie land managers (IDOT or Canadian National) have done substantial damage to the soil

surface which could lead to colonization by invasive plant species and disrupt the native prairie plant species assemblages. The precise impacts of fire and soil disturbance are unknown for many cicada species; although there is evidence that soil compaction decreases diversity and abundance in urban areas (Moriyama and Numata 2015) and fire can be beneficial or detrimental to cicadas depending on the species (Smith et al. 2006; Callaham et al. 2002; Callaham et al. 2003; Pons 2015). Fire is no longer used to manage railroad prairies.

Cicadas represent an important large-bodied foodsource for a variety of predators that help nutrient cycling from belowground to above ground ecosystems. During this study I observed several animals preying on live, adult cicadas, including the non-native Chinese mantis, *Tenodera sinensis*, and the eastern kingbird, *Tyrannus tyrannus*. Other records of this species acting as a food source include predation by cicada killer wasps, *Sphecius* spp. (Holliday et al. 2009) and tachinid flies (Stucky 2015; Stucky 2016). These right-of-way lands likely act as corridors for many other species, but research is very limited on organisms that inhabit these systems.

Future studies will include this dataset and *Me. dorsatus* samples collected in 2021 and 2022 at the same sites. We also have a similar DNA repository of *Cicadettana calliope calliope* (Walker, 1850), the small grass cicada, collected at many of the same locations and some additional sites where *Me. dorsatus* was not present. Comparison of these two species would likely reveal insight into the differences in their connectivity on the landscape. *Cicadettana c. calliope* was found at more prairies than *Me. dorsatus* over the course of this study, even appearing in some restored prairies (unlike *Me. dorsatus*). However, it is a small and cryptic species, with ultrasonic calls that could only be heard by ~50% of the field technicians involved

in this project. It is possible its range might be more contiguous on the landscape, but genetic data might help better resolve this question.

As we look towards future management of disturbed prairies, we should consider that while plants and animals can be artificially reintroduced into a disturbed ecosystem, there is a loss in genetic diversity in doing so and that there may be below ground impacts that are not easily monitored or may have long term consequences. Insects may not always be able to recolonize newly restored areas, especially if agricultural fields separate natural areas. Contiguity of habitat on a fragmented landscape is vital for the conservation of cicadas. Given my findings, management of these railroad right-of-way prairies is vital for the conservation of this species within Illinois.

Acknowledgments

Thank you to the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign, especially Chris Wright, Leslie Benson, and Alvaro Hernandez for their guidance and service in the sequencing of this library. The development of this library was based on the Sacks protocol available on wetlab and with guidance from Nick LaBonte and Lindsay V. Clark. Thank you to Mark Davis (Illinois Natural History Survey) for assistance and access to the Collaborative Conservation Genetics Laboratory. Thank you to the Illinois Nature Preserves Commission, the Illinois Department of Natural Resources, and Illinois Department of Transportation for site permissions and guidance. Thank you to my field technicians that have helped collect cicadas over the years, including Thomas M. Schmeelk, Jonathan R. Tetlie, Kate Johnson, Tyler Hedlund, Vanessa Gabel, Shannon Carlson, Bailey Clancy, and Jenny Chlipala.

Tables

Table 2.1. Number of *Megatibicen dorsatus* specimens and their respective locality by year. Counts represent n after filtering for individuals with low reads or coverage. Sites under US45 right-of-way (ROW) or 12 Mile ROW railroad prairie headings are ordered from North to South.

															US45 ROW				12 MILE ROW						
		Gleason NP	Long Branch	Eldon Hazlet	Ballard NC	Rankin	Pellsville Cem	Loda CPNP	Soloner Tract	Forbes	Jasper	Donnelly	Prospect CPNP	Mainline	S Buckley	N Loda	S Paxton	S Ludlow	Tract 1	Tract 2	Oriole	Tract 3	Tract 4	Tract 5	N Kinmundy
	n =	**58**	**3**	**20**	**12**	**28**	**2**	**52**	**2**	**1**	**2**	**2**	**20**	**4**	**19**	**1**	**66**	**12**	**8**	**22**	**1**	**24**	**23**	**23**	**9**
Year Collected	**2015**	12			2			11	2	1	2	2					12								
	2016	6			10			11								1	12								
	2017	4				7											10								
	2018	12	2			11		12									12								9
	2019	12		8		10		11					12	4	10		11	8	1	10	1	12	12	11	
	2020	12	1	12			2	7					8		9		9	4	7	12		12	11	12	

Table 2.2. STRUCTURE HARVESTER Evanno table output showing values calculated by STRUCTURE based on K ("natural populations") and the number of replicates for each value of K.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	1	-3597776.400000	0.000000	—	—	—
2	3	-3548055.466667	3175.798938	49720.933333	7793.066667	2.453892
3	3	-3506127.600000	2432.787724	41927.866667	18182.266667	7.473840
4	3	-3482382.000000	1126.615893	23745.600000	11169.875000	9.914537
5	4	-3469806.275000	1368.706098	12575.725000	5182.616667	3.786508
6	3	-3462413.166667	3634.683377	7393.108333	296262.091667	81.509739
7	2	-3751282.150000	177216.586425	-288868.983333	1360914.166667	7.679384
8	2	-5401065.300000	739196.842552	-1649783.150000	—	—

Table 2.3. FST values between locations with 8 or more samples. Cells are shaded from the lowest F_{ST} value (white) to highest (dark green). Higher values represent greater difference in allele frequencies, or genetic distance, between locations.

	BNC	BUCKLEY	ELDON	FORBES	GLEASON	LODA	LUDLOW	NKIN	PROSPECT	RANKIN	SPAX	TRACT1	TRACT2	TRACT3	TRACT4	TRACT5
BNC	NA															
BUCKLEY	0.0619	NA														
ELDON	0.0213	0.0468	NA													
FORBES	0.0250	0.0691	0.0144	NA												
GLEASON	0.0661	0.0850	0.0503	0.0702	NA											
LODA	0.0797	0.0795	0.0667	0.1008	0.1045	NA										
LUDLOW	0.0456	0.0578	0.0319	0.0395	0.0739	0.0774	NA									
NKIN	0.0175	0.0486	0.0088	0.0161	0.0548	0.0717	0.0314	NA								
PROSPECT	0.0420	0.0519	0.0292	0.0473	0.0682	0.0725	0.0353	0.0314	NA							
RANKIN	0.0715	0.0863	0.0583	0.0779	0.0965	0.1012	0.0566	0.0615	0.0636	NA						
SPAX	0.0297	0.0412	0.0187	0.0258	0.0583	0.0579	0.0215	0.0192	0.0155	0.0482	NA					
TRACT1	0.0194	0.0512	0.0138	0.0114	0.0606	0.0745	0.0367	0.0083	0.0324	0.0641	0.0228	NA				
TRACT2	0.0195	0.0521	0.0121	0.0159	0.0572	0.0741	0.0338	0.0071	0.0332	0.0611	0.0230	0.0102	NA			
TRACT3	0.0215	0.0528	0.0150	0.0142	0.0611	0.0744	0.0369	0.0099	0.0355	0.0640	0.0251	0.0108	0.0119	NA		
TRACT4	0.0216	0.0541	0.0147	0.0143	0.0609	0.0756	0.0384	0.0097	0.0354	0.0650	0.0253	0.0131	0.0118	0.0125	NA	
TRACT5	0.0180	0.0480	0.0100	0.0124	0.0558	0.0722	0.0331	0.0049	0.0313	0.0612	0.0218	0.0094	0.0080	0.0106	0.0104	NA

Table 2.4. F_{ST} values between structure groups (k = 6). Groups match those in Figure 2.15: R (light blue, 12-mile group, Eldon Hazlet, assorted southern Illinois locations), S (red, Rankin group), T (yellow, Gleason group), U (lilac, Loda group), V (green, southern half US-45 group), and W (dark blue, US-45 Buckley group). Values are colored such that the darkest green represents the highest F_{ST} value and light green the lowest.

	R	S	T	U	V	W
R	NA					
S	0.0547	NA				
T	0.0500	0.0968	NA			
U	0.0626	0.1005	0.1034	NA		
V	0.0162	0.0480	0.0566	0.0562	NA	
W	0.0404	0.0829	0.0813	0.0741	0.0365	NA

Figures

Figure 2.1. Example of railroad right-of-way prairie showing prairie habitat adjacent to mowed region. In Illinois, these prairies are also often surrounded by agricultural fields, like corn and soy. Image is of 12 Mile Prairie which is located south of Effingham Illinois and stretching along IL-37 towards Salem Illinois. These prairies are named "Tracts" 1 through 5 in this study.

Figure 2.2. Partial map of Illinois showing collection locations of *Megatibicen dorsatus* (red dots). Loda Cemetery Prairie Nature Preserve (Loda CPNP), Prospect Cemetery Nature Preserve (Prospect CPNP), Ballard Nature Center (Ballard NC) and Henry A. Gleason Nature Preserve are indicated on the map as well. Railroad rights-of-way (ROWs) that follow US 45 are indicated in dashed lines north of Champaign. Railroad ROWs that follow IL-37 (12 Mile Prairie) prairies are indicated in dashed lines south of Effingham.

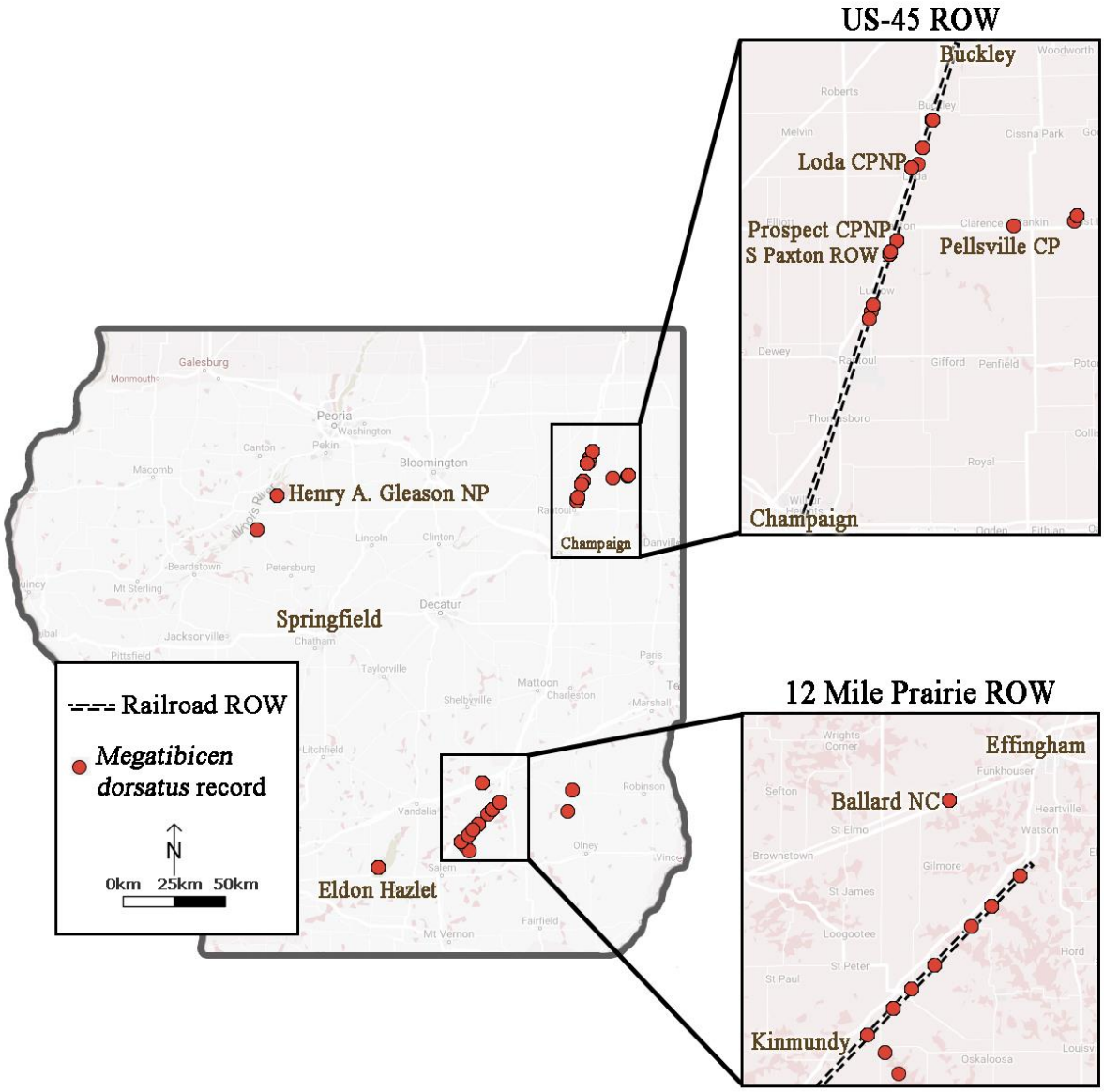


Figure 2.3. Effective coverage and loci per sample distributions for all samples after the gstacks portion of the Stacks pipeline.

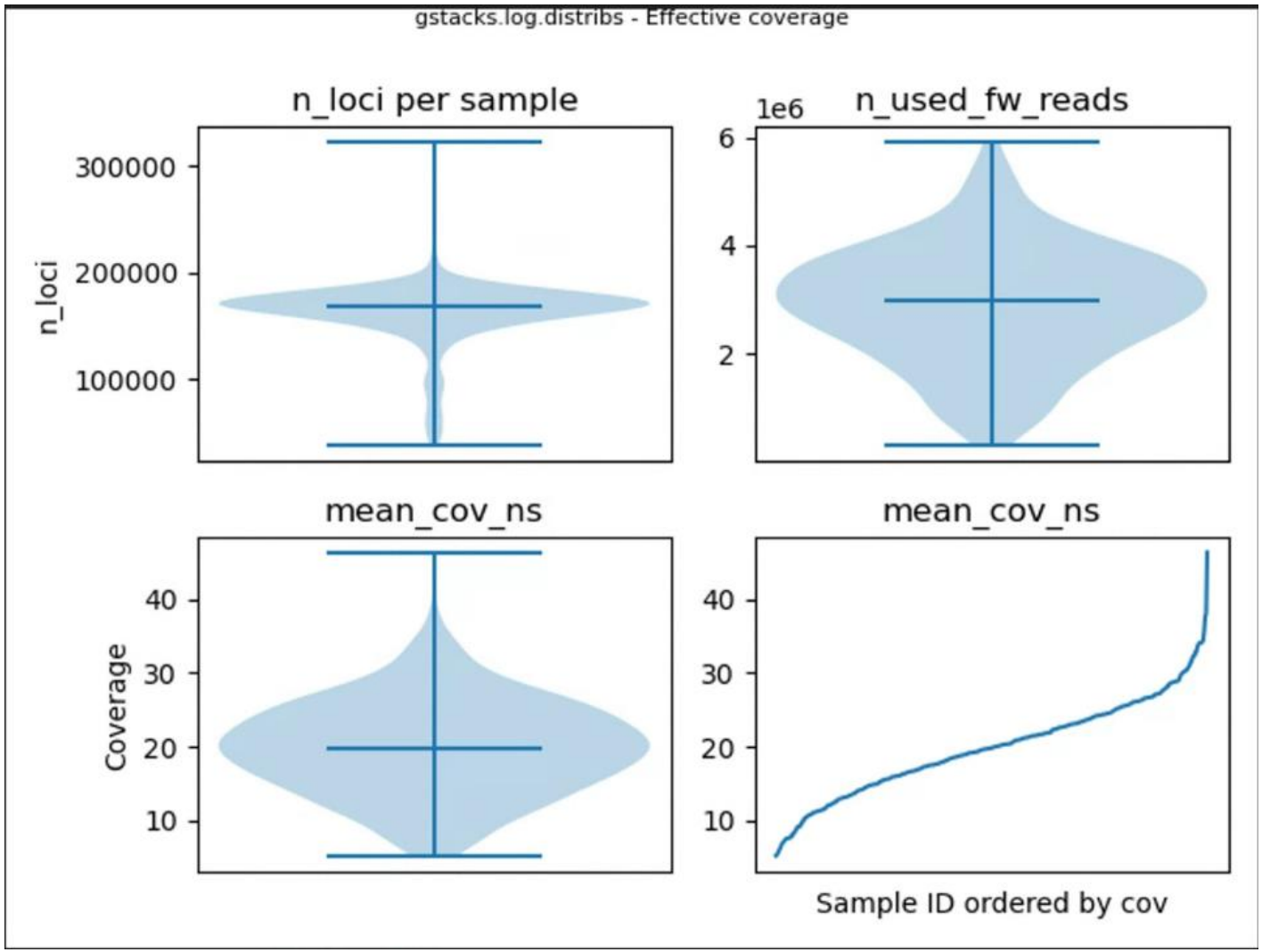


Figure 2.4. fastSTRUCTURE output generated in R and ordered by site for values of k between 3 and 6. According to FastStructure the -best value to maximize marginal likelihood was K=4.

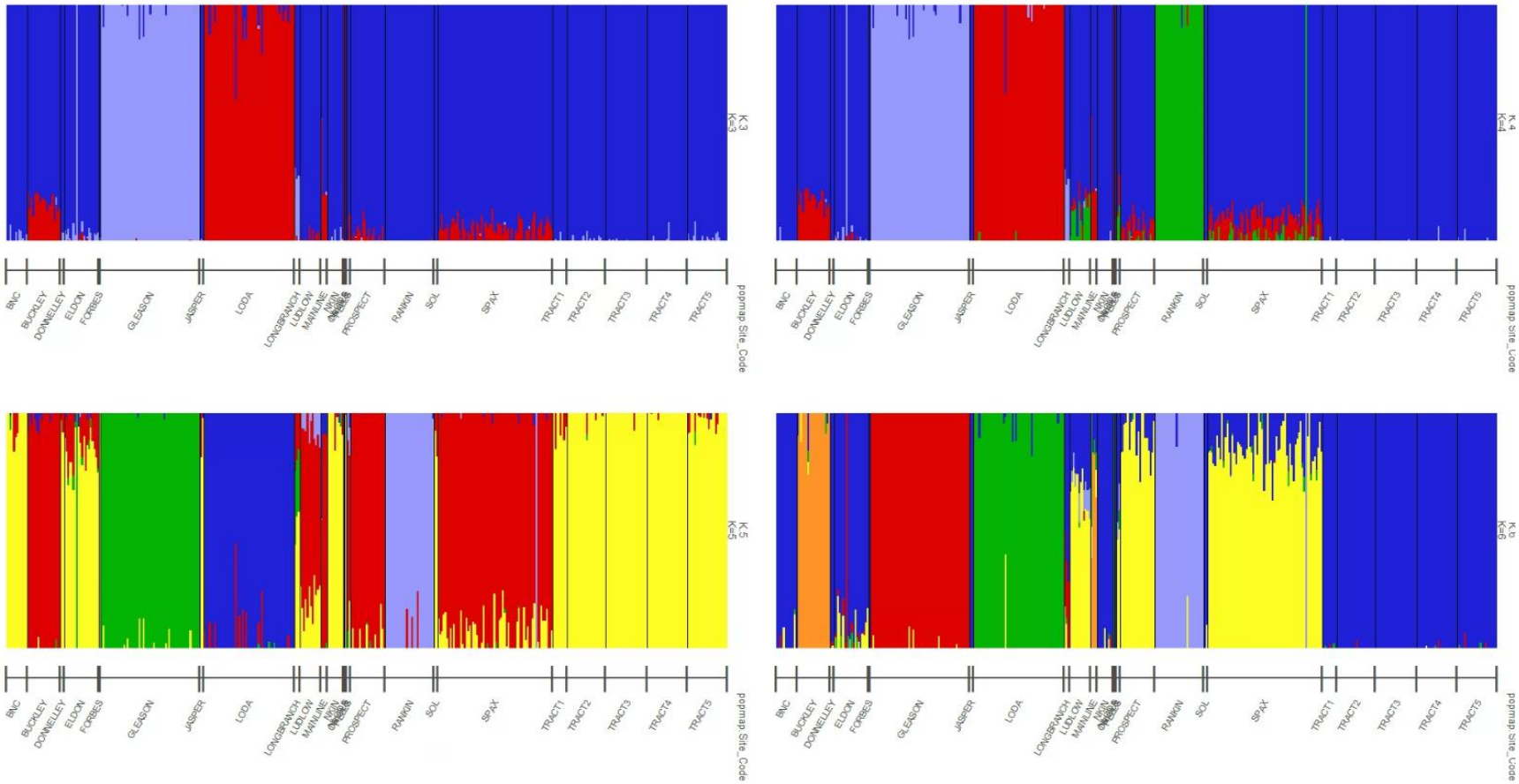


Figure 2.5. fastSTRUCTURE output generated in R and ordered by site for values of k between 11 and 14.

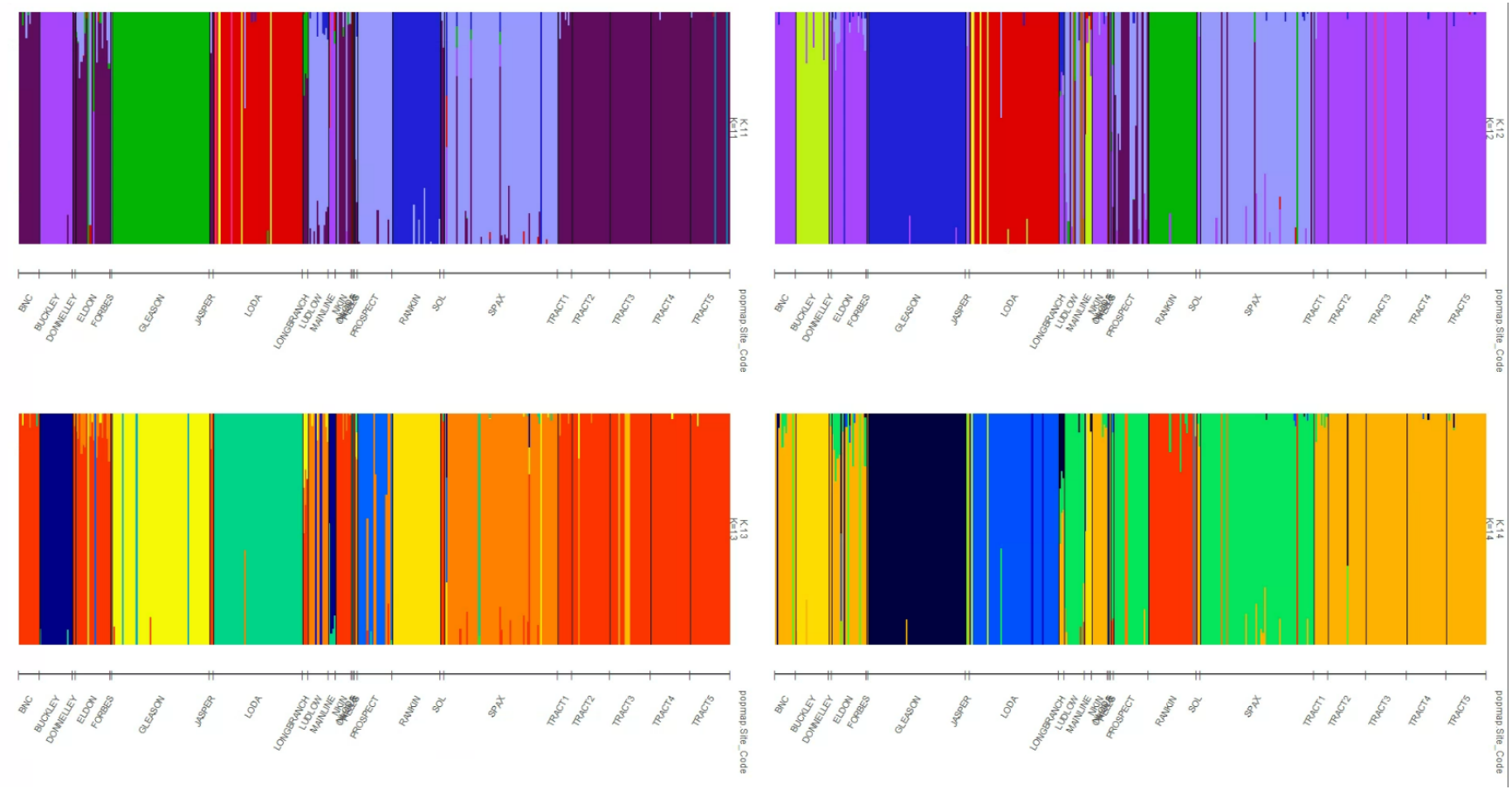


Figure 2.6. STRUCTURE HARVESTER figures for utilizing the Evanno (2005) method. (a) L(K) (mean +/- SD) (b) Rate of change of the likelihood distribution (mean); (c) Absolute value of the 2nd order rate of change of the likelihood distribution (mean); (d) DeltaK = mean(L"(K))/sd(L(K))

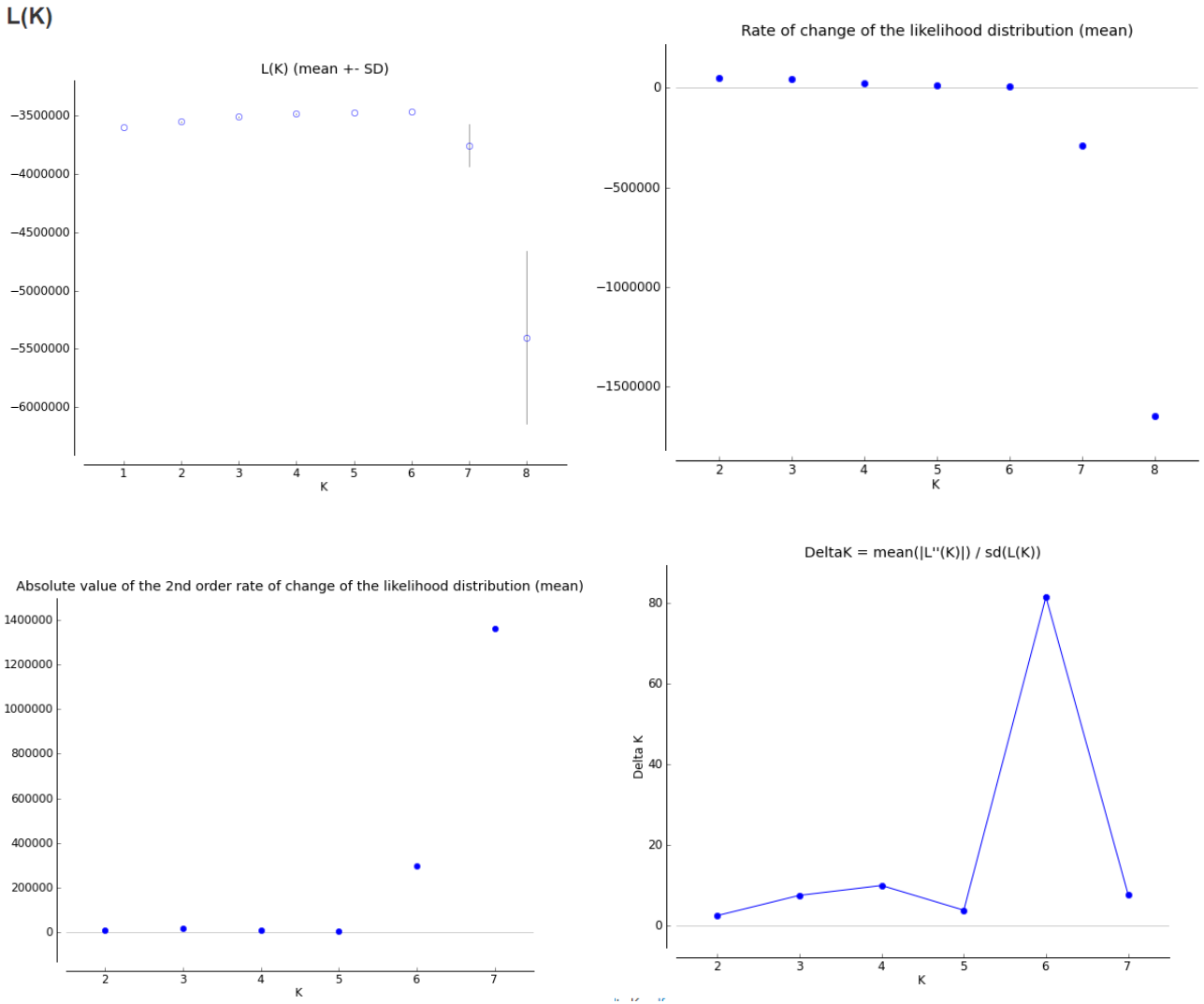
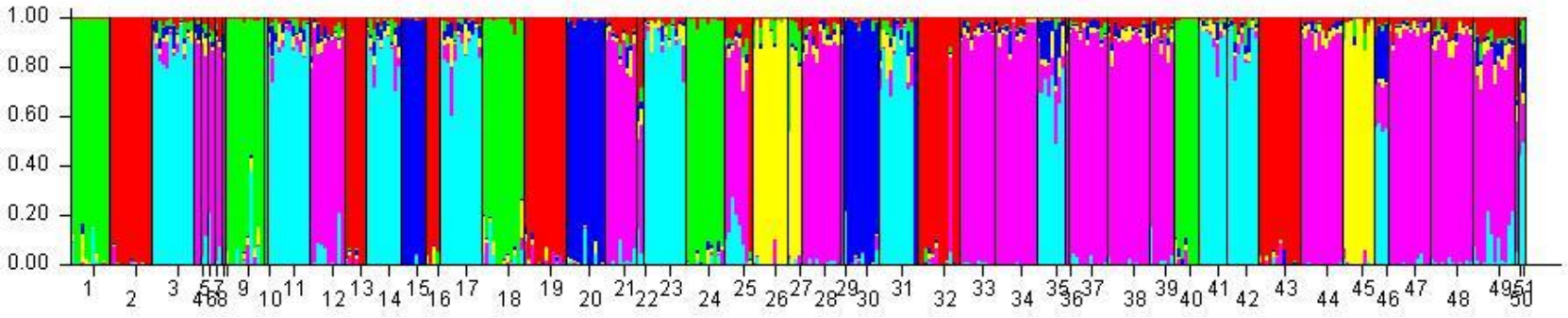


Figure 2.7. STRUCTURE Model with k = 6 based on STRUCTURE HARVESTER output and sorted by site and year population map.



1 2015_LODA
2 2015_GLEASON
3 2015_SPAX
4 2015_BNC
5 2015_SOL
6 2015_JASPER
7 2015_DONNELLEY
8 2015_FORBES
9 2016_LODA
10 2016_NLODA
11 2016_SPAX
12 2016_BNC
13 2016_GLEASON
14 2017_SPAX
15 2017_RANKIN
16 2017_GLEASON
17 2018_SPAX
18 2018_LODA
19 2018_GLEASON
20 2018_RANKIN
21 2018_NKIN
22 2018_LONGBRANCH
23 2019_PROSPECT
24 2019_LODA
25 2019_ELDON
26 2019_BUCKLEY
27 2019_MAINLINE
28 2019_TRACT5
29 2019_ORIOLE
30 2019_RANKIN
31 2019_SPAX
32 2019_GLEASON
33 2019_TRACT2
34 2019_TRACT3
35 2019_LUDLOW
36 2019_TRACT1
37 2019_TRACT4
38 2020_TRACT2
39 2020_TRACT1
40 2020_LODA
41 2020_PROSPECT
42 2020_SPAX
43 2020_GLEASON
44 2020_TRACT3
45 2020_BUCKLEY
46 2020_LUDLOW
47 2020_TRACT4
48 2020_TRACT5
49 2020_ELDON
50 2020_LONGBRANCH
51 2020_PELLS

Figure 2.8. PCA (Principal components analysis) plot showing the different populations with ellipses surrounding 99% of members. Groups include every site with all years grouped together. Note that not all groups contained enough points to create an ellipse.

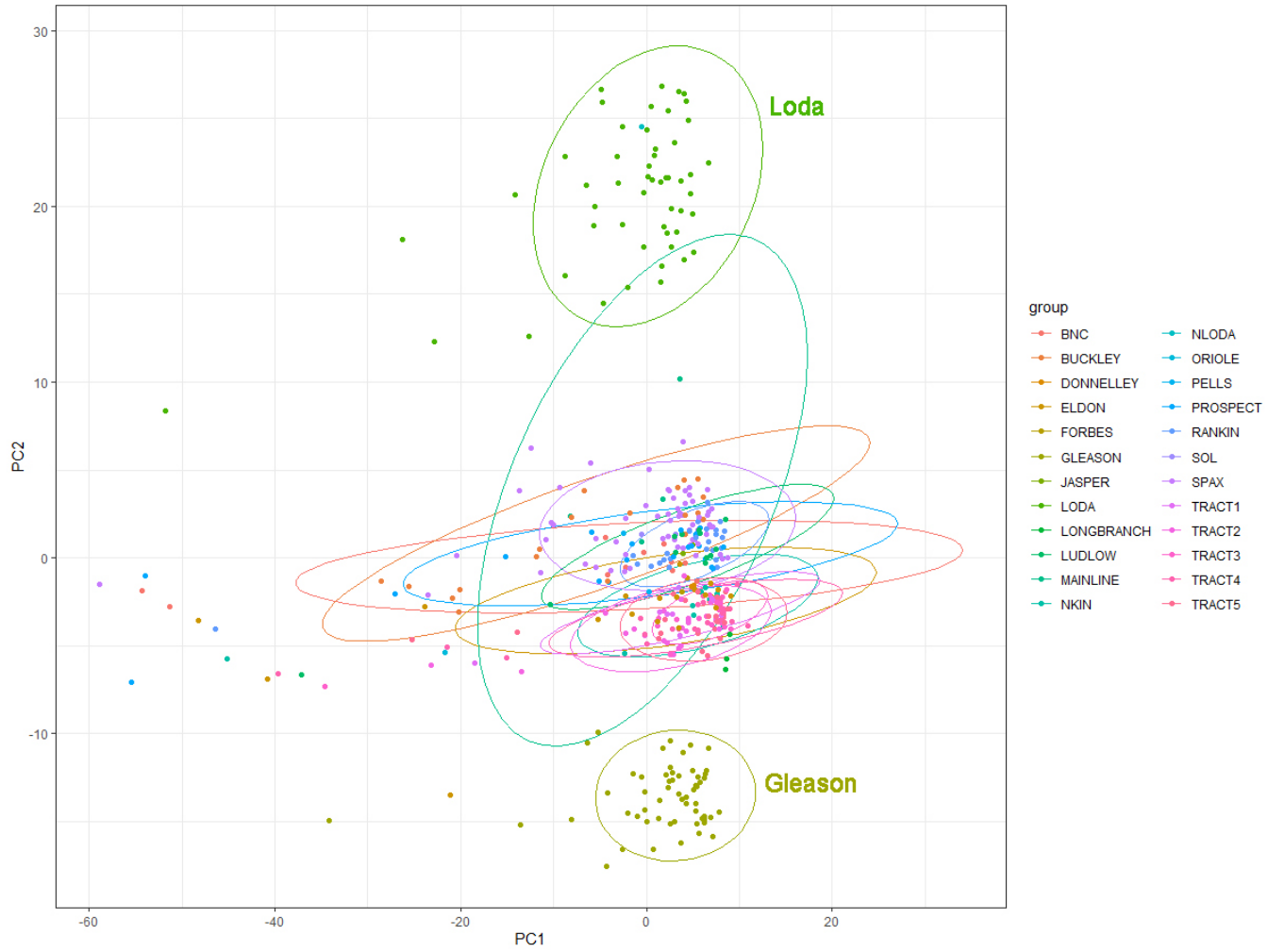


Figure 2.9. Principal component analysis (PCA) plot showing different populations with railroad prairies grouped by location IL37 (12 Mile Prairie, including Tract 1, Tract 2, Tract 3, Tract 4, Tract 5, Oriole, and North of Kinmundy) and US-45 (Ludlow, North of Loda, Buckley, and South of Paxton). Ellipses surround 99% of members of a group and not all groups contained enough points to create an ellipse.

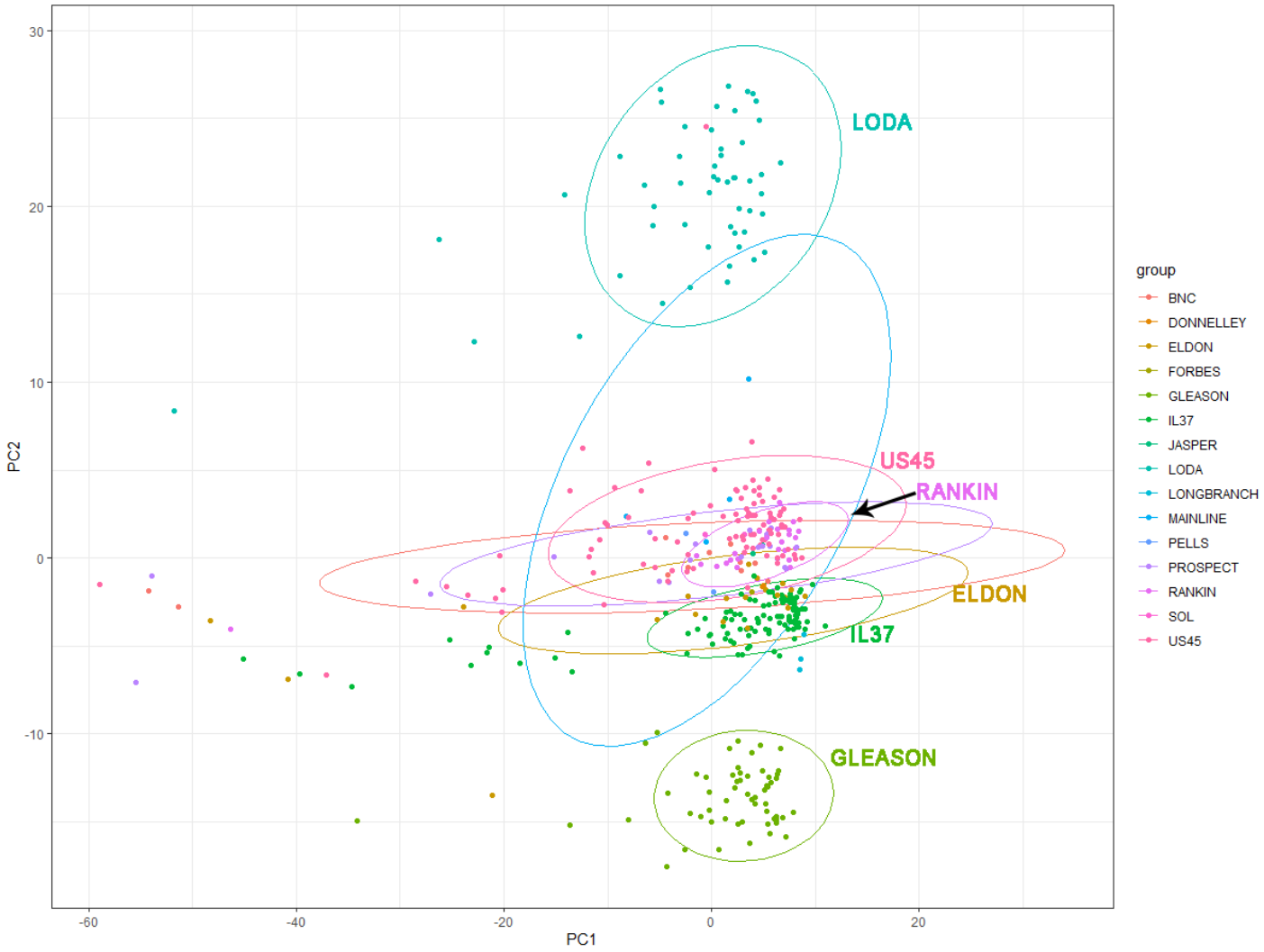


Figure 2.10. Principal component analysis (PCA) plots showing individuals by Latitude.

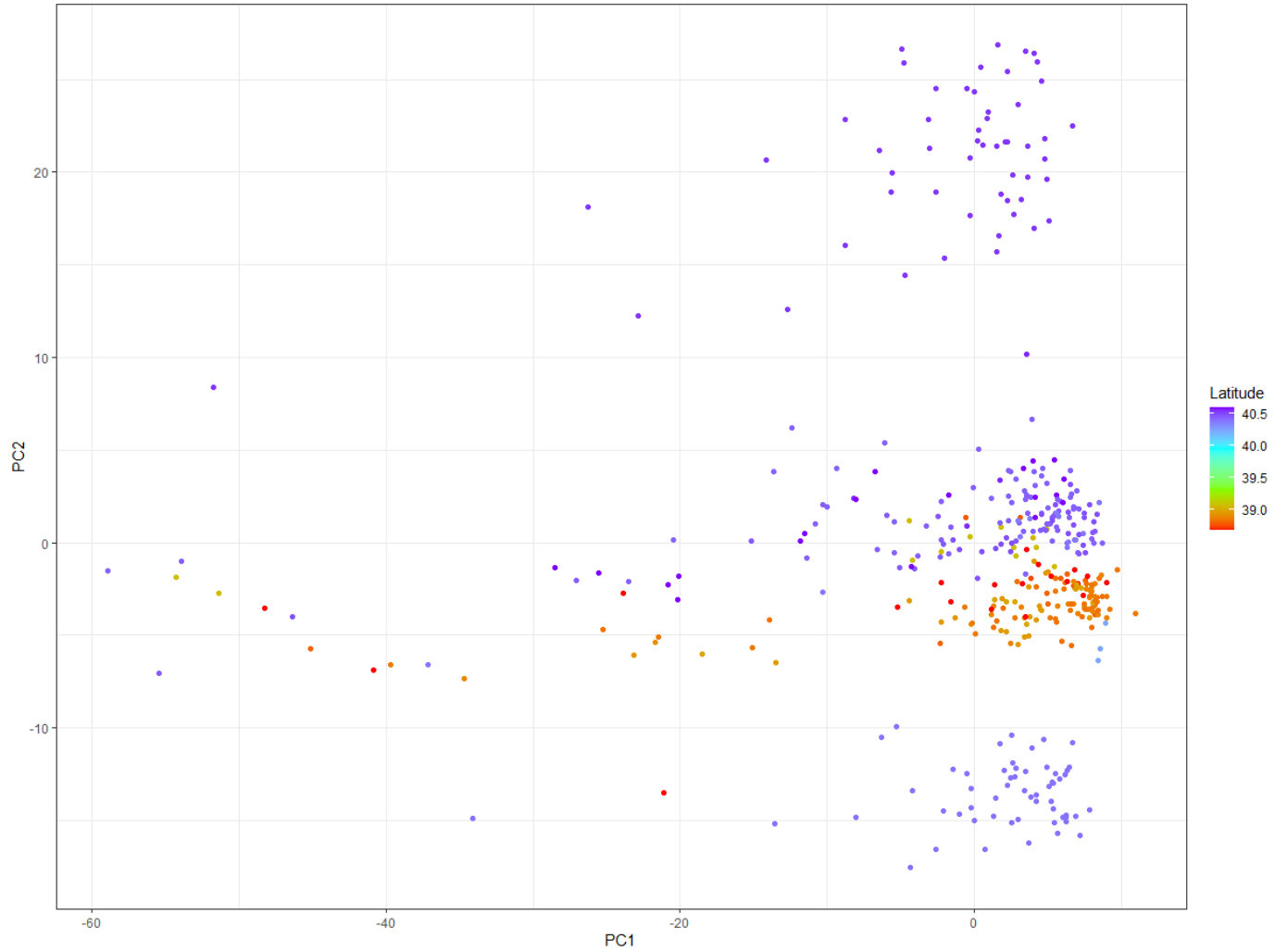


Figure 2.11. Principal component analysis (PCA) plot showing individuals colored by longitude.

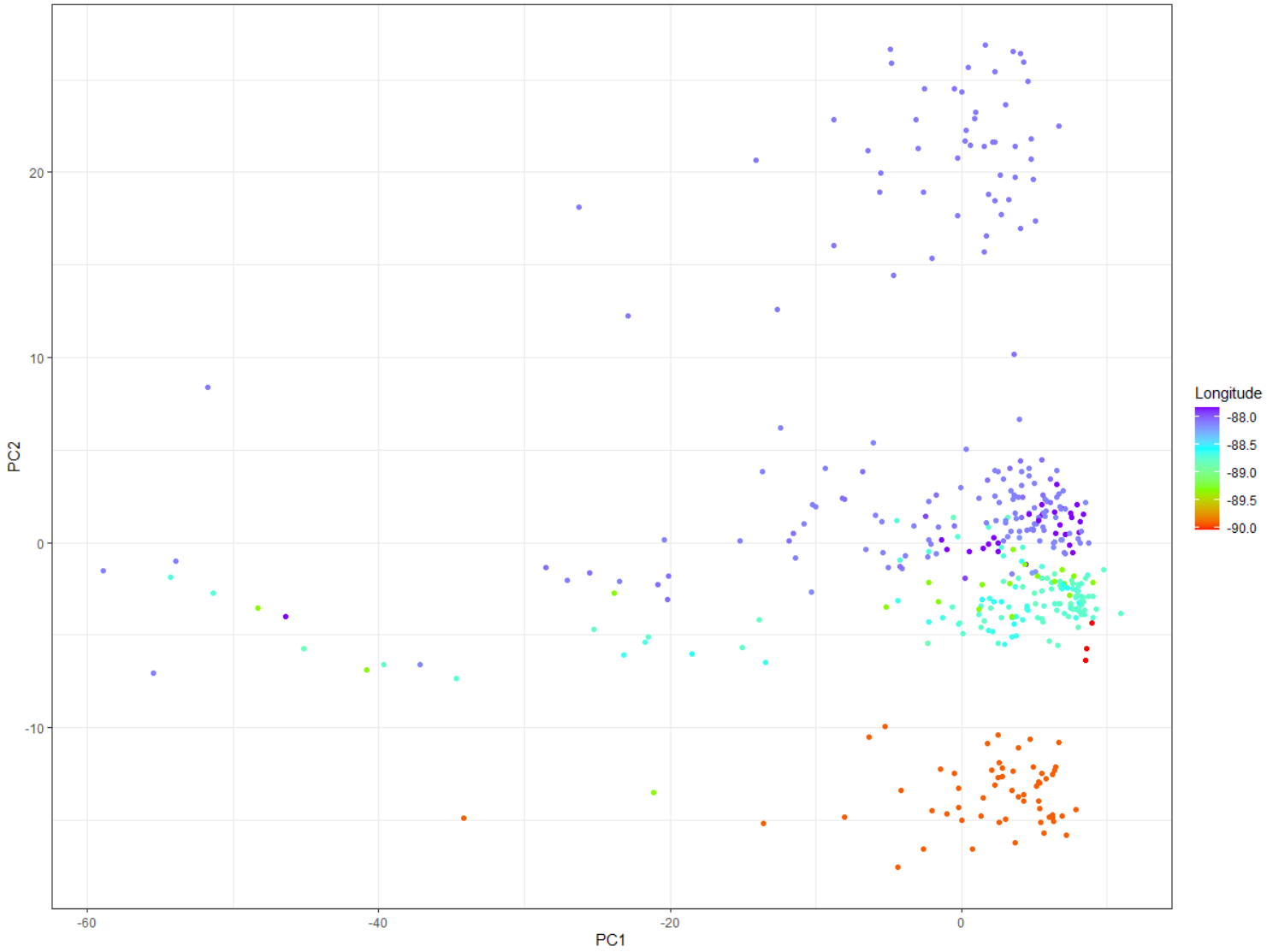


Figure 2.12. Mantel test created using a matrix of geographic distances (based on latitude and longitude) and genetic distances. Observed value can be seen at 0.253568. p = 0.085

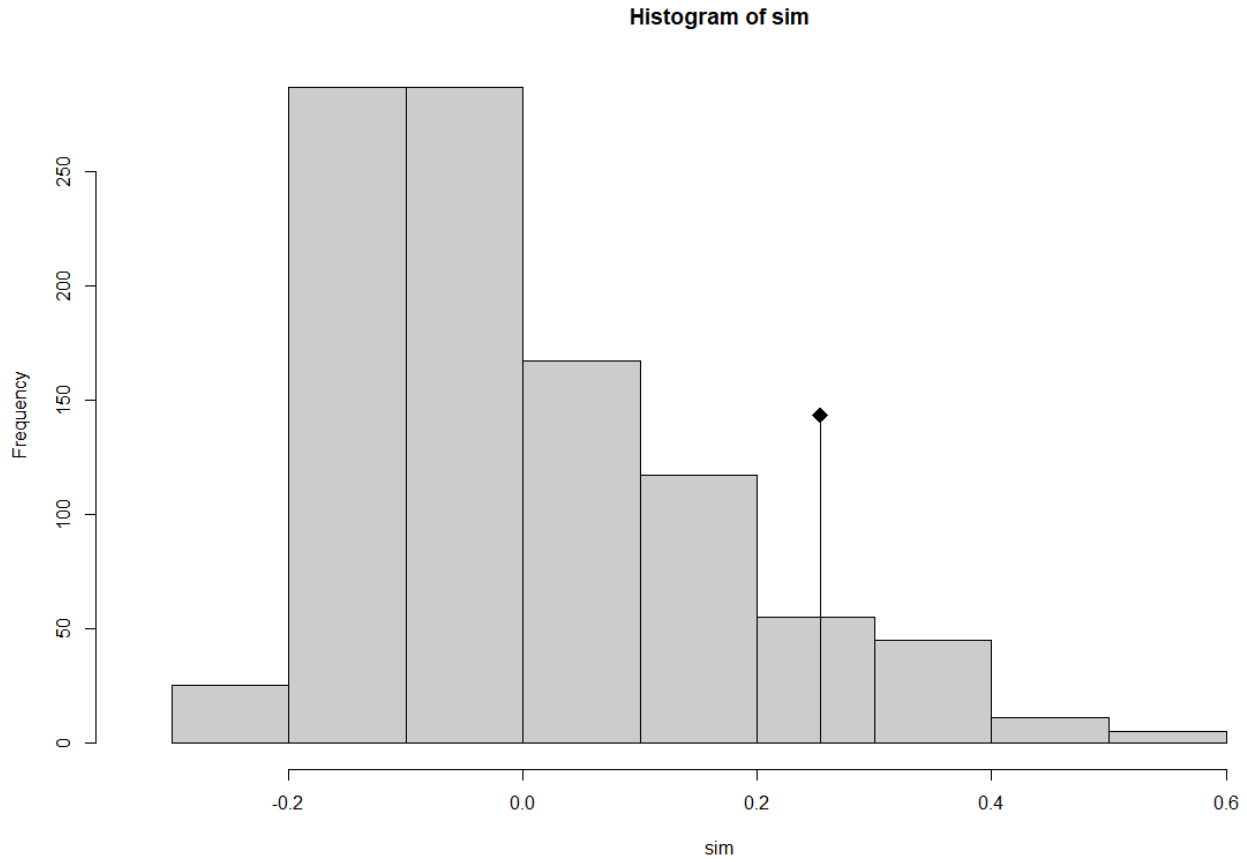


Figure 2.13. Isolation by distance scatter plot calculated adegenet 2.1.8 with a trend line to best fit the data.

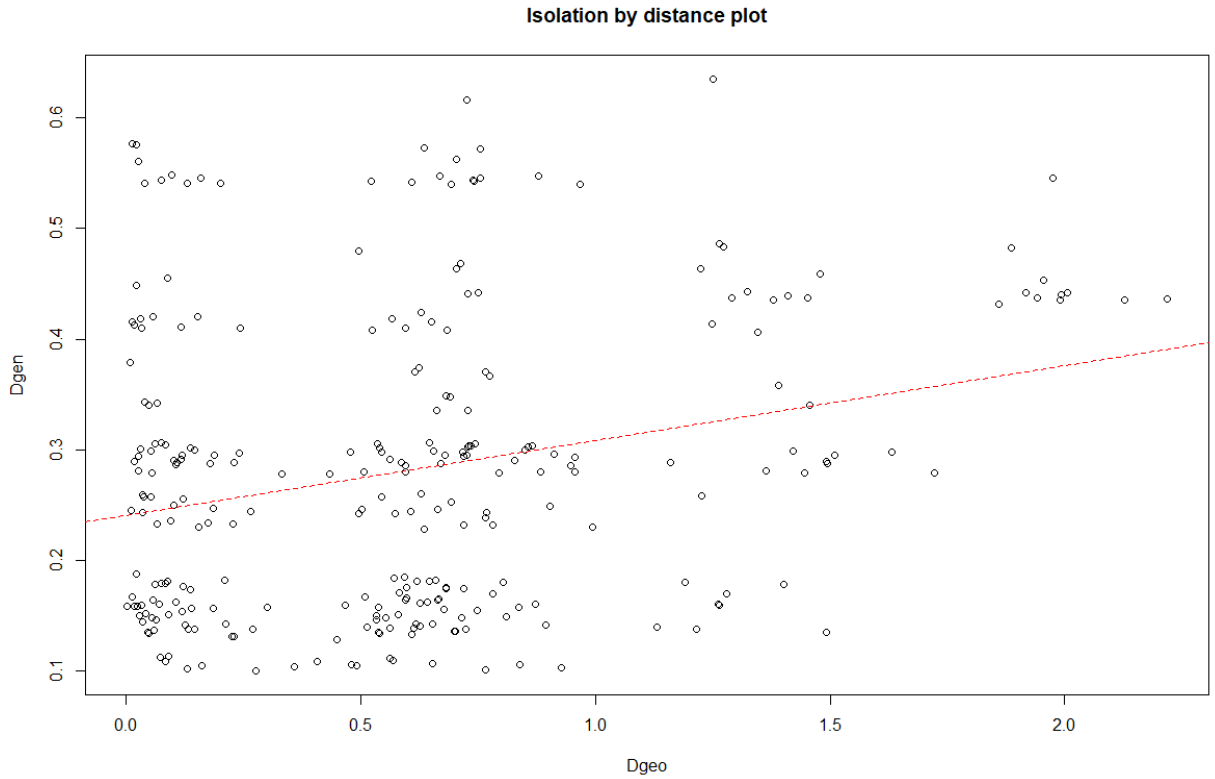


Figure 2.14. Scatter plot using euclidean distance between points and pairwise F_{ST} matrix to illustrate isolation by distance. Trend line was added to best fit data and a slight positive correlation can be seen.

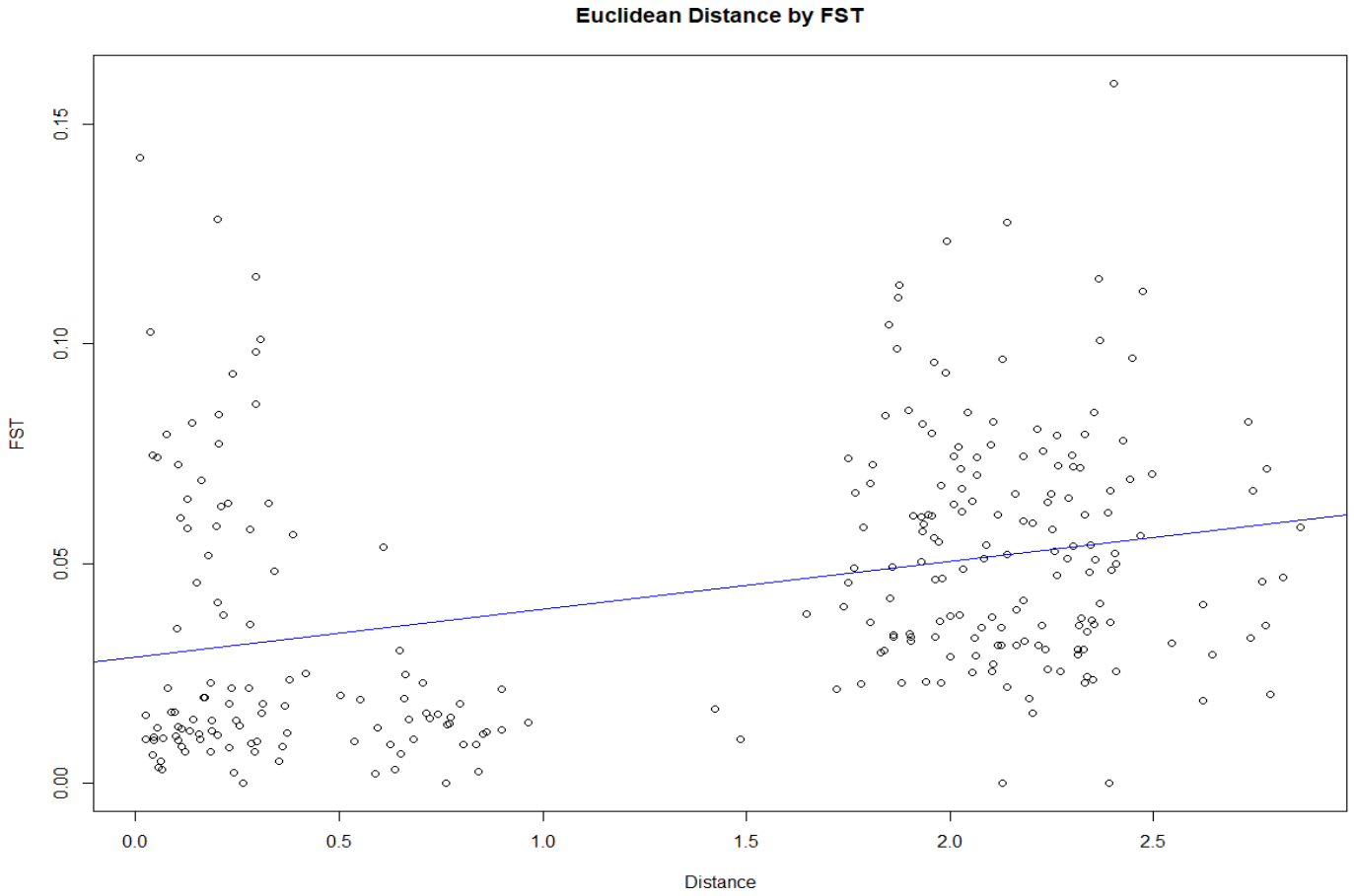


Figure 2.15. Map of *Megatibicen dorsatus* sampling locations and structure (k=6) plots by location. Colors on map indicate the population that the location best fits within (highest proportion). Colors correspond to groups from Table 2.4: R (light blue, 12-mile group), S (red, Rankin group), T (yellow, Gleason group), U (lilac, Loda group), V (green, southern half US-45 group), and W (dark blue, US-45 Buckley group).

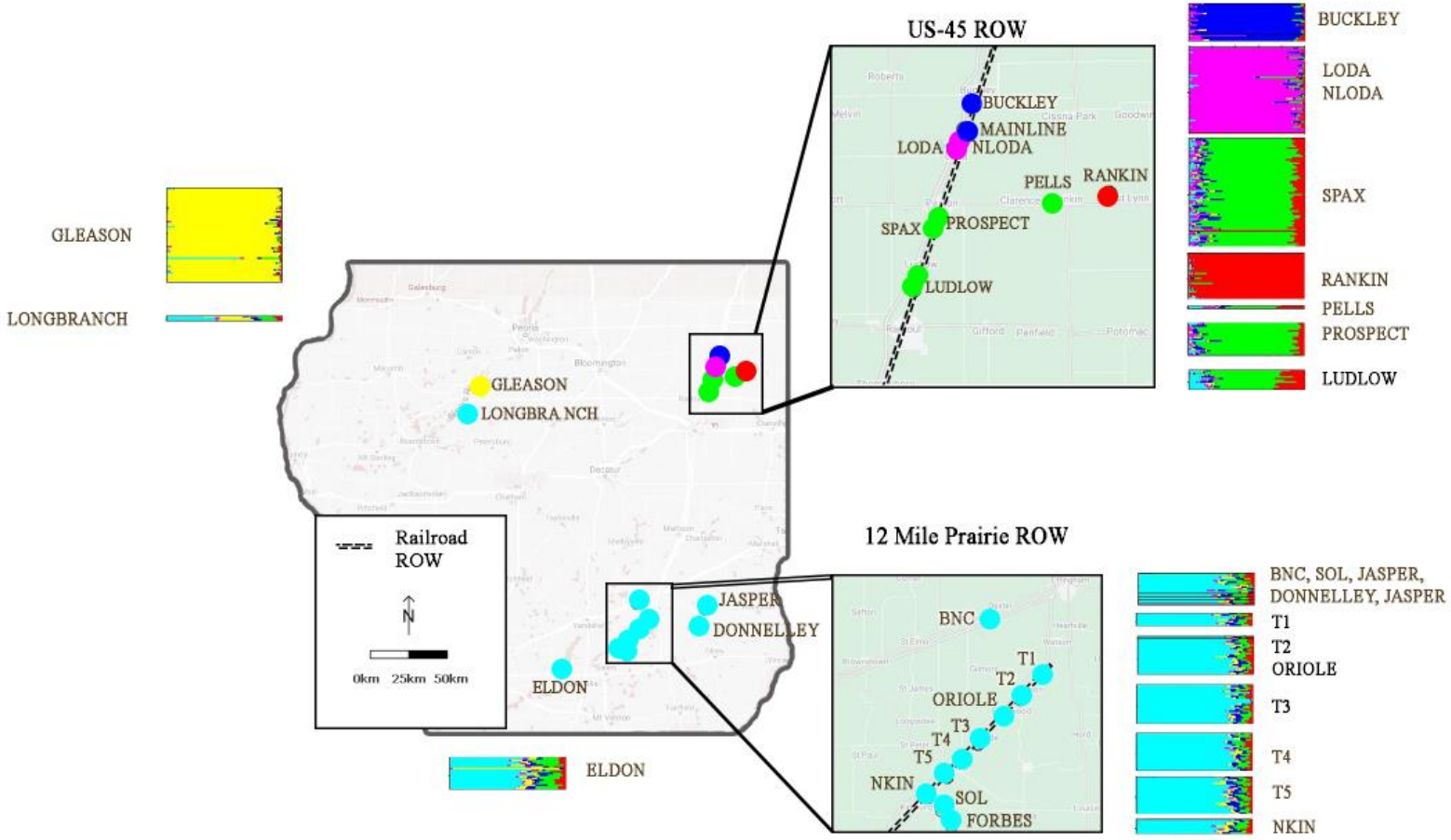


Figure 2.16. Overlay of population assignments (k = 6) on Schwegman's Natural Divisions of Illinois (Schegman 1997). 12 Mile Prairie ROW points fall within the Southern Till Plain Division, US-45 ROW points in Grand Prairie Division, and Gleason and Longbranch within the Illinois River and Mississippi River Sand Areas Division.

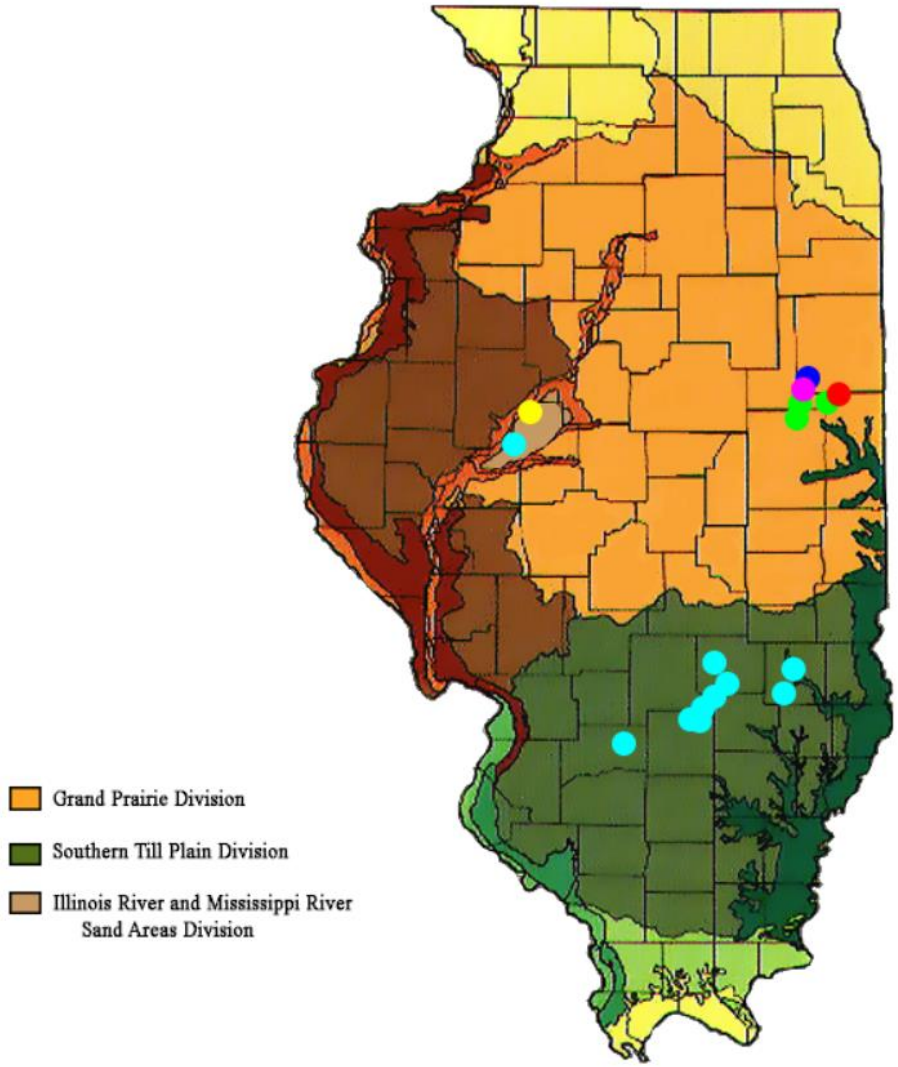
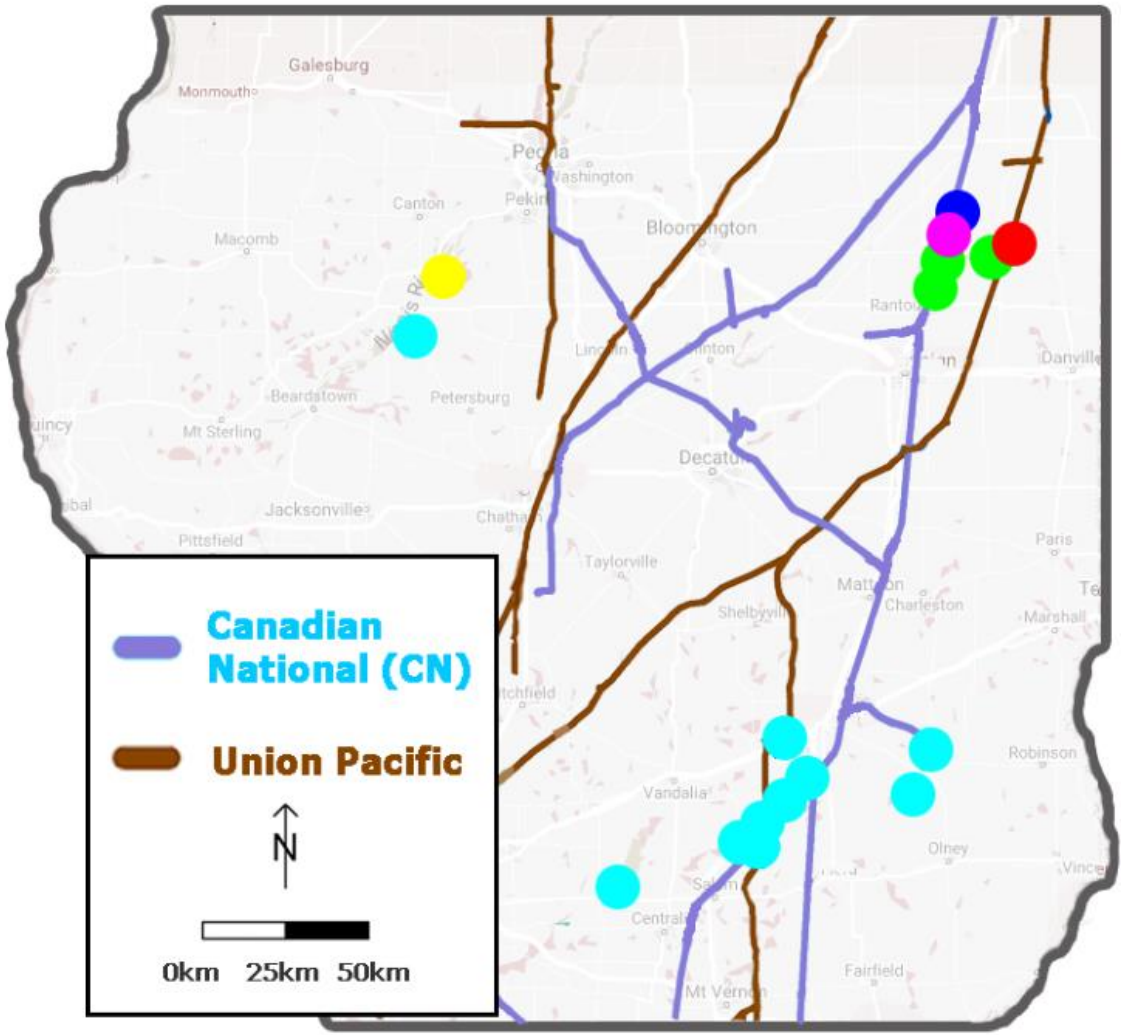


Figure 2.17. Overlay of STRUCTURE population assignments (k = 6) on the Canadian National and Union Pacific rail lines in Illinois.



References

Alexander, R.D., and T.E. Moore. 1962. The evolutionary relationships of 17-year and 13-year cicadas, and three new species (Homoptera, Cicadidae, Magicicada). Miscellaneous Publications. Museum of Zoology, University of Michigan 121: 1–59.

Anderson, R.C. 1970. Prairies in the prairie state. Transactions of the Illinois State Academy of Science 63(2): 214–221.

Assis, J., Castilho Coelho, N., Alberto, F., Valero, M., Raimondi, P., Reed, D., and E.A. Serrão. 2013. High and distinct range-edge genetic diversity despite local bottlenecks. PloS ONE 8(7): e68646. https://doi.org/10.1371/journal.pone.0068646

Beamer, R.H. 1928. Studies on the biology of Kansas Cicadidae. University of Kansas Science Bulletin 18: 155–263.

Bolin, K.E., Albrecht, N.J., and R.L. Jacobson. 1988. Identification, preservation, and management of Minnesota roadside prairie communities. Transportation Research Record 1279: 79–85.

Callaham, M.A., Whiles, M.R., and J.M. Blair. 2002. Annual fire, mowing and fertilization effects on two cicada species (Homoptera: Cicadidae) in tallgrass prairie. The American Midland Naturalist 148(1): 90–101.

Callaham, M.A., Blair, J.M., Todd, T.C., Kitchen, D.J., and M.R. Whiles. 2003. Macroinvertebrates in North American tallgrass prairie soils: effects of fire, mowing, and fertilization on density and biomass. Soil Biology and Biochemistry 35(8): 1079–1093.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., and J.H. Postlethwait. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. G3: Genes, Genomes, Genetics 1(3): 171–182.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and W.A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22(11): 3124–3140.

Chafin, T.K., Martin, B.T., Mussmann, S.M., Douglas, M.R., and M.E. Douglas. 2018. FRAGMATIC: in silico locus prediction and its utility in optimizing ddRADseq projects. Conservation Genetics Resources 10(3): 325–328.

Clark, L.V., Brummer, J.E., Głowacka, K., Hall, M.C., Heo, K., Peng, J., Yamada, T., Yoo, J.H., Yu, C.Y., Zhao, H., and S.P. Long. 2014. A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. Annals of Botany 114(1): 97–107.

Cole, J.A. 2008. A new cryptic species of cicada resembling *Tibicen dorsatus* revealed by calling song (Hemiptera: Auchenorrhyncha: Cicadidae). Annals of the Entomological Society of America 101(5): 815–823.

Collins, S.L., and S.C. Barber. 1986. Effects of disturbance on diversity in mixed-grass prairie. Vegetatio 64(2): 87–94.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., and G. McVean. 2011. The variant call format and VCFtools. Bioinformatics 27(15): 2156–2158.

De Geer, C. 1773. Des cigales exotiques de la quatrieme famille. Cinquième mémoire. Des cigales. Mémoires pour servir à l'histoire des Insects 3: 212–222.

Earl, D.A., and B.M. VonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4(2): 359–361.

Eckert, C.G., Samis, K.E., and S.C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Molecular Ecology 17(5): 1170-1188.

Evanno, G., Regnaut, S., and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14(8): 2611–2620.

Froeschner, R.C. 1952. A synopsis of the Cicadidae of Missouri (Homoptera). Journal of the New York Entomological Society 60(1): 1–14.

Haldeman, S.S. 1852. Appendix C. Insects. In: Stansbury, H. (Ed.) Exploration and survey of the valley of the Great Salt Lake of Utah, including a reconnoissance of a new route through the Rocky Mountains. pp. 367–379. Lippincott, Grambo, & Co.: Philadelphia, PA. https://doi.org/10.5962/bhl.title.51702

Hanrahan, S.J., and J.S. Johnston. 2011. New genome size estimates of 134 species of arthropods. Chromosome Research 19(6): 809–823.

Harrington, J.A., and M. Leach. 1989. Impact of railroad management and abandonment on prairie relicts. In T.B. Bragg and J. Stubbendeick (eds.). Prairie Pioneers: Ecology, History and Culture: Proceedings of the Fifth Midwest Prairie Conference. Lincoln, Nebraska.

Hart, C.A., and H.A. Gleason. 1907. On the biology of the sand areas of Illinois. Illinois State Laboratory of Natural History 7(7): 137–272.

Hill, K.B., Marshall, D.C., Moulds, M.S., and C. Simon. 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.

Hoffmann, C.H. 1942. Annotated list of elm insects in the United States. US Department of Agriculture. Miscellaneous Publication 466.

Holliday, C.W., Hastings, J.M., and J.R. Coelho. 2009. Cicada prey of New World cicada killers, *Sphecius* spp. (Dahlbom, 1843) (Hymenoptera: Crabronidae). Entomological News 120(1): 1–17.

Hughes, A.R., Byrnes, J.E., Kimbro, D.L., and J.J. Stachowicz. 2007. Reciprocal relationships and potential feedbacks between biodiversity and disturbance. Ecology Letters 10(9): 849–864.

Huijser M.P., and A.P. Clevenger. 2006. Habitat and corridor function of rights-of-way. In: Davenport J., Davenport J.L. (eds) The Ecology of Transportation: Managing Mobility for the Environment. Environmental Pollution, vol 10. Springer, Dordrecht. https://doi.org/10.1007/1-4020-4504-2_11

Hyman, W., and S. Manley. 1977. Vegetation and wildlife impacts of railroad branch line abandonment. Proceedings of the 18th Annual Transportation Research Forum 18(1): 454–462.

Iverson, L.R. 1988. Land-use changes in Illinois, USA: The influence of landscape attributes on current and historic land use. Landscape Ecology 2(1): 45–61.

IWAP. 2015. 2015 Implementation Guide to the Illinois Wildlife Action Plan (Revision 4-18-2016). https://www.dnr.illinois.gov/conservation/IWAP/Documents/FinalDraft2015_FINAL_Revision%204-18-16.pdf

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21): 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, M., and M. Bowles. 2013. Tree ring analysis of eastern red cedar reveals fire history of Fults Hill Prairie Nature Preserve, and its relationship to climate and loss of prairie vegetation. Morton Arboretum Publications, Lisle, Illinois.

Jones, M.D., and M.L. Bowles. 2016. Eastern redcedar dendrochronology links hill prairie decline with decoupling from climatic control of fire regime and reduced fire frequency. The Journal of the Torrey Botanical Society 143(3): 239–253.

Kamvar, Z.N., Brooks, J.C., and N.J. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in Genetics 6(208): 1–10. https://doi.org/10.3389/fgene.2015.00208

Kamvar, Z.N., Tabima, J.F., and N.J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281. https://doi.org/10.7717/peerj.281

Knaus, B.J., and N.J. Grünwald. 2016. VcfR: An R package to manipulate and visualize VCF format data. bioRxiv Press 041277. https://dx.doi.org/10.1101/041277

Knaus, B.J., and N.J. Grünwald. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources 17(1): 44–53. https://dx.doi.org/10.1111/1755-0998.12549

Lee, Y.J. 2016. Description of three new genera, *Paratibicen*, *Megatibicen*, and *Ameritibicen*, of Cryptotympanini (Hemiptera: Cicadidae) and a key to their species. Journal of Asia-Pacific Biodiversity 9(4): 448–454.

Leston, L., and N. Koper. 2016. Urban rights-of-way as reservoirs for tall-grass prairie plants and butterflies. Environmental Management 57(3): 543–557.

Linnaeus, C. 1758. II. Hemiptera. Systema naturae: per regna tria naturae, secundum classes, ordines, genera, species cum characteribus, differentiis, synonymis, locis. Editio decima, reformata. L. Salvii. Stockholmiae 1: 1–824.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17(1): 10–12. https://doi.org/10.14806/ej.17.1.200

McClain, W.E., Phillippe, L.R., and J.E. Ebinger. 2005. Floristic assessment of the Henry Allan Gleason Nature Preserve, Mason County, Illinois. Castanea 70(2): 146–154.

McClain, W.E., Ruffner, C.M., Ebinger, J.E., and G. Spyreas. 2021. Patterns of Anthropogenic Fire within the Midwestern Tallgrass Prairie 1673–1905: Evidence from Written Accounts. Natural Areas Journal 41(4): 283–300.

McElrath, T. 2022. Illinois Natural History Survey Insect Collection. Illinois Natural History Survey. Occurrence dataset https://doi.org/10.15468/eol0pe accessed via GBIF.org on 2022-10-04.

Moriyama, M., and H. Numata. 2015. Urban soil compaction reduces cicada diversity. Zoological Letters 1(19): 1–19. DOI: 10.1186/s40851-015-0022-3.

OpenWetWare. 2017. Sacks:RAD-seq. https://openwetware.org/mediawiki/index.php?title=Sacks:RAD-seq&oldid=1019322

Paris, J.R., Stevens, J.R., and J.M. Catchen. 2017. Lost in parameter space: a road map for Stacks. Methods in Ecology and Evolution 8(10): 1360–1373.

Puritz, J.B., Hollenbeck, C.M., and J.R. Gold. 2014a. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431 https://doi.org/10.7717/peerj.431

Puritz, J.B., Matz, M.V., Toonen, R.J., Weber, J.N., Bolnick, D.I., and C.E. Bird. 2014b. Demystifying the RAD fad. Molecular Ecology 23: 5937–5942. https://doi.org/10.1111/mec.12965

Pons, P. 2015. Delayed effects of fire and logging on cicada nymph abundance. Journal of Insect Conservation 19: 601–606. https://doi.org/10.1007/s10841-015-9781-6

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Raj, A., Stephens, M., and J.K. Pritchard. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 197(2): 573–589.

Ramankutty, N., and J.A. Foley. 1999. Estimating historical changes in global land cover: Croplands from 1700 to 1992. Global Biogeochemical Cycles 13(4): 997–1027.

Reed, D., and J. Schwarzmeier. 1975. The prairie corridor concept: possibilities for planning large scale preservation and restoration. In D.C. Glenn-Lewin and R.Q. Landers, Jr. (eds.). Proceedings of the Fifth Midwest Prairie Conference (Pages 158–165). Iowa State University. Ames, Iowa.

Robertson, K.R., and M.W. Schwartz. 1994. Prairies. In: The changing Illinois environment: critical trends. Technical report of the critical trends assessment project Volume 3: Ecological resources. Illinois Department of Energy and Natural Resources, Illinois Natural History Division, Champaign.

Rochette, N., and J. Catchen. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nature Protocols 12: 2640–2659.

RStudio Team. 2022. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/

Sanborn, A.F., and M.S. Heath. 2016. Megatibicen n. gen., a new North American cicada genus (Hemiptera: Cicadidae: Cicadinae: Cryptotympanini). Zootaxa 4168(3): 577–582.

Sanborn, A.F., and M.S. Heath. 2017. Priority and synonymy of some North American cicada genera (Hemiptera: Cicadidae: Cicadinae: Cryptotympanini). Zootaxa 4243(2): 377–382.

Sanborn, A.F., and P.K. Phillips. 2013. Biogeography of the cicadas (Hemiptera: Cicadidae) of North America, north of Mexico. Diversity 5(2): 166–239.

Say, Thomas. 1825. Descriptions of new hemipterous insects, collected in the expedition to the Rocky Mountains, performed by the order of Mr. Calhoun, Secretary of War, under command of Major Long. Journal of the Natural Academy of Sciences of Philadelphia 4(2): 307–345.

Smith, D.M., Kelly, J.F., and D.M. Finch. 2006. Cicada emergence in southwestern riparian forest: influences of wildfire and vegetation composition. Ecological Applications 16(4): 1608–1618.

Stucky, B.J. 2015. Infection behavior, life history, and host parasitism rates of *Emblemasoma erro* (Diptera: Sarcophagidae), an acoustically hunting parasitoid of the cicada *Tibicen dorsatus* (Hemiptera: Cicadidae). Zoological Studies 54(1): 1–17.

Stucky, B.J. 2016. Eavesdropping to find mates: the function of male hearing for a cicada-hunting parasitoid fly, *Emblemasoma erro* (Diptera: Sarcophagidae). Journal of Insect Science 16(1): 68; 1–7. https://doi.org/10.1093/jisesa/iew048

Schwegman, J. 1997. Illinois' natural divisions. In The Illinois Steward, Urbana, Illinois. 12 pp.

Terry, M.R. 2018. Vegetation management along roadside and railroad right-of-ways. Doctoral dissertation, University of Missouri—Columbia. https://hdl.handle.net/10355/66286

Walker, F. 1850. List of the specimens of homopterous insects in the collection of the British Museum. Order of Trustees. London, England. Part 1: 1–260.

Weir, B.S., and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6): 1358–1370.

White, J. 1978. Illinois natural areas inventory technical report. Department of Landscape Architecture, University of Illinois, Urbana-Champaign, and Natural Land Institute, Rockford, Illinois.

Wolf, J. 2004. A 200-year fire history in a remnant oak savanna in Southeastern Wisconsin. The American Midland Naturalist 152(2): 201–213.

CHAPTER 3: MICROBIOME DIVERSITY OF *MEGATIBICEN DORSATUS* AND *NEOTIBICEN PRUINOSUS PRUINOSUS* (HEMIPTERA: CICADIDAE) AND ENDOSYMBIONT COMMUNITY

Abstract

Cicadas have been a source of interest for the bioinspired design of novel material surfaces due to their superhydrophobic, anti-reflective, and anti-microbial wings. These functionalities are the result of arrays of nanopillars present on the surface of the wings. In order to elucidate why these nanopillars are present on the wings and not on other body parts, and their role in microbial community establishment, I explored the microbial diversity on different surfaces across the cicada body. I studied two cicada species that live in different ecosystems: prairies in the case of *Megatibicen dorsatus* and wooded areas (e.g., residential) in the case of *Neotibicen pruinosus pruinosus.* In addition, I explored how bacteria and fungi colonize these surfaces over time by collecting adult cicadas directly after emerging from the soil and several weeks after their assumed emergence. I found that legs are the likely exposure point to pathogenic microbes and that microbial communities likely establish quickly after emergence. Despite their antimicrobial properties, wings do not differ significantly in their microbial community from other body parts. Despite my efforts to focus on the ectobiome, this study also provides species specific insight into the cicada endosymbiont community. Similar to studies on other species of cicadas, I found evidence of yeast-like endosymbionts that replaced the endosymbiont *Candidatus* Hodgkinia in both *M. dorsatus* and *N. p. pruinosus.* These fungal endosymbionts were recruited from pathogenic fungi and still have high genetic similarity to *Ophiocordyceps* species.

Keywords

Beauveria, Hypocreales, ITS, metabarcoding, primary endosymbionts, YLS

Introduction

The wings of many cicadas, including *Megatibicen dorsatus* (Say, 1825) and *Neotibicen pruinosus pruinosus* (Say, 1825) are superhydrophobic, which aids in self-cleaning (Hasan et al. 2013; Oh et al. 2017). In some species, the nanostructures make a cicada wing superhydrophobic and bactericidal (Ivanova et al. 2012; Kelleher et al. 2016; Román-Kustas et al. 2020), thus providing cicadas with at least one less exposure point to entomopathogens. The wings of many species of cicada are also antireflective and transparent, reflecting as little as 1% of visible light (Han et al. 2016; Huang et al. 2015). Disruption of this transparent and antireflective surface could be detrimental to cicada fitness (i.e., through reduced camouflage due to dirt built-up and biofouling). The multi-functionality of the nano-structures present on cicada wing surfaces can lead to the guided bioinspired design of novel surfaces with similar functionalities (Oh et al. 2020).

In most insects, the exoskeleton is the first line of defense against pathogens (Brey et al. 1993; Klowden 2013). Indeed, most pathogens instead infect insects after being ingested (Mukherjeea and Vilcinskas 2018), rather than directly through the exoskeleton. Fungi can overcome the physical barrier of the integument by utilizing hydrolytic enzymes to degrade the cuticle and allow access to the hemolymph (Fan et al. 2007; Pereira et al. 2007). Given the way that cicadas cover their abdomens while they hold their wings tent-like over their body at rest, the self-cleaning and bactericidal properties of wing surfaces might provide an extra line of defense against entomopathogens. Microbes can build up on the surfaces of insect integuments

and are assumed to be primarily neutral or harmful to the host, but these microbes can also serve as a potential source of mutualistic interactions (Sen et al. 2009; Smith et al. 2021).

There are several known pathogens of cicadas, including the fungus *Massospora* Peck 1879 (Entomopthoraceae) (Soper 1963; Macias et al. 2020), the chalky bacteria *Luethyella okanaganae* (Lüthy and Soper 1969; Lüthy 1974; O'Neal et al. 2017), and the fungus *Metacordyceps* (Clavicipitaceae) (Li et al. 2010). Some of these pathogens can quickly kill the host after exposure. *Massospora* is an obligate fungal pathogen of cicadas (Macias et al. 2020; Soper 1974) that can fill the entirety of the abdomen (Soper et al. 1974; Soper et al. 1976) while still allowing the infected cicada to fly and infect others (Cooley et al. 2018; Soper 1963; Soper et al. 1976).

We know very little of the full range of microbiota that exists on the surfaces of cicada wings and bodies and how aforementioned surface-feature mediated superhydrophobicity might impact the colonization of the wing surface in contrast to the rest of the body. For this study I focused on two different cicada species, *Megatibicen dorsatus* and *Neotibicen pruinosus pruinosus*. These two species inhabit different habitats. *M. dorsatus* is a prairie-associated species that occurs across the central United States (Sanborn and Phillips 2013) and was collected in a disturbed railroad right of way prairie. *N. p. pruinosus*, commonly known as the scissor grinder cicada also occurs across a wide range of the United States but is associated with less threatened habitats and can be found in residential areas where it uses hardwood trees as a host plant.

Previous work done by myself and collaborators showed that the wings surfaces of both species have uniformly distributed conical nanopillars. The nanopillars of *M. dorsatus* are about 250 nm in height, whereas those present on *N. p. pruinosus* wings are taller at ~350-400 nm in

height (Oh, 2017). Comparison of the wettability (water contact angle behavior) revealed that both cicada species have superhydrophobic wings with droplets residing in the Cassie−Baxter state (apparent advancing contact angle >150°) (Oh 2017). We also showed that in *N. p. pruinosis*, the superhydrophobic, high aspect ratio pillars on the wings inhibited bacterial attachment much better than the less hydrophobic, lower aspect ratio pillars present on the wings of periodical *Magicicada* wings (Román-Kustas 2020), rendering *N. p. pruinosus* less prone to biofouling.

To determine the field relevance of the antimicrobial surface properties I studied the microbial communities present on *M. dorsatus* and *N. p. pruinosus* at two different time points: (1) immediately after emerging from the ground and eclosing to an adult and (2) several weeks after this emergence and thus after they have been exposed to their local microbiota. In addition, given the anti-microbial properties of the wing surface, I studied the microbial communities present on the wings as compared to different body segments and body parts, the exuviae, and the soil. This study is the first to determine if surfaces that have been shown to be antimicrobial and which have inspired novel engineered antimicrobial surfaces in fact shape microbial communities in the field.

Materials and Methods

Field collection

Megatibicen dorsatus and *Neotibicen pruinosus pruinosus* adults were collected at two different time points: (1) immediately, or as near as possible, after emergence from the ground and eclosing to their adult form and (2) late in the season (mid-August to September). *M. dorsatus* adults were collected at a railroad right-of-way (ROW) prairie south of the town of

Paxton, Illinois (40.4296°, -88.1091°). Both newly emerged and older, fully sclerotized adult *N. p. pruinosus* individuals were collected between Crystal Lake Park and a nearby residential area in Urbana, Illinois (40.1258°, -88.2080°). Two *Cicadettana calliope calliope* (Walker, 1850) were collected from the same site as *M. dorsatus* and one *Neotibicen lyricen lyricen* (De Geer, 1773) was collected from the same site as *N. p. pruinosus.* If adults had fully sclerotized by the time they were collected, they were placed in sterile 50 mL conical tubes at -20°C for at least 48 hours. Adults were collected by gloved hand and not by insect net to avoid contamination. Adult cicadas were dissected into parts using sterile razor blades and placed in enough ATL Buffer to cover until DNA was extracted: (1) head (1.5 mL microcentrifuge tube, 1 mL ATL Buffer), (2) forewings (15 mL conical tube, 5 mL ATL Buffer), (3) hindwings (1.5 mL microcentrifuge tube, 1 mL ATL Buffer), (4) body (15 mL conical tube, 5 mL ATL) (thorax and abdomen), and (5) legs (1.5 mL microcentrifuge tube, 1 mL ATL) (Figure 3.1). As mentioned above, the "body" for the purposes of this study is defined as only the thorax and the abdomen, minus the head, legs, and wings. If a molt was found that was associated with an individual, it was also sterile collected into a container with ATL buffer (15 mL conical tube, 5 mL ATL Buffer). Soil samples at or near the site of emergence (emergence hole/tunnel) were also taken and stored at -20C.

DNA Extraction

DNA was extracted from individual samples using the QIAGEN DNEasy® Blood and Tissue Kits (QIAGEN Inc., Germantown, MD). Individual samples consisted of the body parts described above. DNA extraction methods followed kit instructions with several modifications. Unlike in Chapter's 1 and 2 samples were not ground using plastic pestles, instead, body part samples were vortexed for several seconds, allowed to settle such that bubbles were no longer

present or could be avoided, inverted several times, and then 180 µl of the supernatant was transferred to a fresh tube. Both proteinase K digestion (56°C) and ethanol precipitation (4°C) were performed overnight (~18–24 hours). DNA was eluted in 100 µl pre-warmed AE Buffer. Due to the inhibitory presence of tannins in soil present on their surface, DNA extractions from cicada molts were run through Zymo OneStep PCR Inhibitor Removal Kit Columns (Zymo Research Corp., Irvine, CA). Soil samples were extracted using the QIAGEN DNEasy PowerSoil Pro Kit (QIAGEN Inc.). DNA concentration was quantified using the High Sensitivity dsDNA Invitrogen™ Qubit™ 3 Fluorometer system (Invitrogen, Thermo Fisher Scientific, Waltham, MA). Samples that came up "too low" to measure DNA concentration were still submitted for sequencing. Samples were plated across five 96 well plates for submission for sequencing and provided with DNA concentrations for titration. Multiple controls were included, including field controls (where ATL buffer was poured into the same collection tubes at the same site where adults were collected) and extraction controls (both for the Power Soil and Blood and Tissue kits).

Sequencing

Primers for multiple regions and focal taxa were used in the preparation of the library (Table 3.1) including primers for hypervariable regions of the 16S rRNA gene (V1-V3, V3-V4, and V4) and ITS (Internal Transcribed Spacer). A total of 223 samples were included in the library (Appendix Table C.1). The Fluidigm Access Array system (Fluidigm Corp., South San Francisco, CA) uses two primer sets to create a final amplicon that frames the region of interest: CS1 (5'-ACACTGACGACATGGTTCTACA-3') with CS2 (5'-TACGGTAGCAGAGACTTGGTCT-3') and Illumina i5 (5'-

AATGATACGGCGACCACCGAGATCT) with Illumina i7 (5'-CAAGCAGAAGACGGCATACGAGAT-XXXXXXXXXX-3', where the region denoted by XXXXXXXX is utilized for the index sequence). The prepared library (using Fluidigm) was sequenced at the Carver Center (University of Illinois at Urbana-Champaign) using the Illumina NovaSeq SP flowcell to produce paired-end reads.

dada2 and QIIME2 pipeline

The metadata table (Appendix Table C.1) was validated using Keemei (Rideout et al. 2016) to ensure that it met QIIME 2 formatting requirements. The length of the primer sequence was trimmed from both forward and reverse reads using Trimmomatic Version 0.39 (Bolger et al. 2014). Quality scores were examined using FastQC Version 0.11.8 (Andrews 2010) prior to further analysis in order to determine if further trimming was required before proceeding to subsequent steps in the pipeline. Data was imported into QIIME 2 Version 2021.4 (Bolyen et al. 2019; Hamady et al. 2008; Hamady and Knight 2009), which was used to demultiplex the data. Once the data were demultiplexed, sequences were denoised using dada2 to remove any chimeras and gene errors common with Illumina sequencing (Callahan et al. 2016) (Table 3.4). This also allows for the sorting of amplicons into amplicon sequence variants (ASVs). Example code with documentation can be found in Appendix C.

A trained classifier is needed to classify ASVs into taxonomic groups and can be set to different taxonomic units depending on desired level for operational taxonomic units (OTUs). For the ITS region, I used the publicly available UNITE general FASTA for Fungi Release 10-5-2021 (i.e., unite-ver8-seqs_99_10.05.2021.qza and unite-ver8-taxonomy_99_10.05.2021.qza) (Abarenkov et al. 2020; Bengtsson-Palme et al. 2013). For the Bacteria V3V4 dataset I utilized

the pre-formatted SILVA reference sequence and taxonomy files (Release 138.1) (Pruesse et al. 2007; Quast et al. 2013; Yilmaz et al. 2014) provided through the QIIME 2 website (Bokulich et al. 2021). Taxonomy assignments were made using the trained classifier and visualizations of taxonomy were created using QIIME 2 Viewer (Kaehler et al. 2019; Robeson et al. 2021; Rognes et al. 2016). Given the large amount of data, Shannon's entropy calculates the uncertainty of predicting the species in a sample to measure diversity and QIIME 2 was used to calculate this in ITS3-ITS4, but could not be calculated on bacterial V3-V4 due to the computational requirements needed (Shannon 1948).

Data Validation

Sequence identity was checked using the National Center for Biotechnology Information Basic Local Alignment Search Tool (NCBI BLAST) (Johnson et al. 2008). Furthermore, if any expected species (e.g., *Candidatus* Hodgkinia) were determined to not be present in a particular sample, the database (https://unite.ut.ee/ or https://www.arb-silva.de/search/) was checked to ensure that the missing organism was present. Relevant sequences in GenBank were also aligned to sequences in Geneious Prime® 2022.1.1. For some specific analyses, Geneious Prime was used to create local BLAST databases to search known accessions against ASVs present. In order to correct for contamination from extraction columns, buffers, lab consumables, PCR amplification, and user error, the number of reads for an ASV found in respective controls (i.e., buffer, extraction, field) was subtracted from the total number of reads from each sample for the identity provided.

Results

The prepared library (using Fluidigm) produced paired reads of 250 nucleotides in length for a total of 802 million reads. After sorting by primer 796 million reads remained, divided up amongst the primer pairs (PrimerSort mismatches allowed: 2) (Table 3.2). Sequence length included primers, so after trimming sequences were ~230bp in length, depending on the primer (Table 3.1). No hard quality trimming was needed as PHRED scores averaged above 30 for the full length of the sequences. After demultiplexing each ITS3-ITS4 sample had an average of 383,882 forward reads and reverse reads. The V3-V4 region was similar, with an average of 360,319 forward and reverse reads (Table 3.2). For the purposes of this study, focus is on these two regions, but future work will include a comparison to the other primer sets for better resolution of taxonomy and validation of results from QIIME 2 processing. After dada2 filtering, average sequence length of reads can be seen in Table 3.3.

Archaea

There was substantial bias in the feature count and sampling depth for Archaea in the few soil samples included in my analyses, reducing the quality of data in non-soil samples (Table 3.3). In the Archaea analysis, 1,764,036 (42.37%) features in 9 (3.86%) soil samples were retained at a sampling depth of 196,004. The sampling depth by sample type was much less skewed across fungi (ITS3-ITS4) and bacteria (V3-V4) features across my samples than it was with Archaea.

Bacteria (V3-V4)

Chloroplast V3-V4 reads made up anywhere from 0 reads to 86.8% of total reads per sample (body part) with the highest read counts being in the legs and hindwings of several cicadas. Similarly, mitochondrial V3-V4 reads made up anywhere from 0 to 79.7% of reads per sample with little pattern as to the location and more to do with the species – the non-focal *Cicadettana calliope calliope* samples had the largest proportion of their reads from mitochondrial DNA. These reads were removed from feature tables to get a better idea of read counts from bacteria.

Greater contamination was seen in bacterial reads using the V3-V4 primer set. In the field controls, most of the contamination was from *Escherichia-Shigella. Escherichia-Shigella* read counts were high across both species, time points, body part, and even in buffer controls. Given the similarity of these two genera it is not possible to know if *Escherichia coli* or *Shigella* spp. was amplified (Devanga Ragupathi et al. 2018). Contamination can be from DNA extraction kits as well as from the lab space being used, which can vary between studies (Salter et al. 2014; Weyrich et al. 2019). Given the low biomass samples, this contamination can have substantial impact on the diversity measures (Salter et al. 2014). Across all samples, 996 OTUs with 10 or more reads were identified.

After correcting for contamination in controls and pooling by sample type, but not body part, the top V3-V4 20 taxonomies were tabulated for early season *M. dorsatus* (MDE) (Table 3.11), late season *M. dorsatus* (MDL) (Table 3.12), early season *N. p. pruinosus* (NPE) (Table 3.13), late season *N. p. pruinosus* (NPL) (Table 3.14), soil collected near *N. p. pruinosus* emergence sites (Table 3.15), and soil collected near *M. dorsatus* emergence sites (Table 3.16).

Candidatus Sulcia muelleri (= *Ca.* Karelsulcia muelleri) was found in 25 of 26 body samples, including *M. dorsatus* (n=11), *N. p. pruinosus* (n=12), *C. c. calliope* (n=2), and *Neotibicen lyricen lyricen* (n=1). As mentioned previously, when referring to the "body" sample, I am referring to only the abdomen and thorax (the legs, head, and wings have all been removed). *Candidatus* Hodgkinia cicadicola did not initially get identified using the trained SILVA database in QIIME 2, nor were any samples found at the higher level of order Hyphomicrobiales that *Ca.* Hodgkinia belongs to. However, after importing the V3-V4 ASVs into Geneious and creating a Custom BLAST database over 100 ASVs were identified that were between 82.8 to 88 percent identity to the V3-V4 region of a *Ca.* H. cicadicola sequence available online (Accession: CP024987). This was necessary as the taxonomy assignments required a greater percent identity and *Ca.* Hodgkinia has lost large sections of its genome resulting in a loss of sequence stability (Campbell et al. 2015). From there the alignment showed the potential for at least two variants within *C. c. calliope* samples (Figure 3.14). Example ASV sequences can be found in Appendix C and showed highest percent identity to *Ca.* Hodgkinia from *Kosemia yezoensis* (tribe Cicadettini).

Fungi (ITS3-4)

Shannon's entropy measures showed a significant increase in fungal diversity between early and late collected *N. p. pruinosus* individuals, with the highest diversity observed in soil samples (Figure 3.3). Shannon's entropy incorporates both evenness and richness into one metric (Gauthier and Derome 2021; Kim et al. 2017). Percent of ITS reads were visualized using qiime2view (Bokulich et al. 2018; Bolyen et al. 2019; Callahan et al. 2016; Hamaday et al. 2008; Hamaday and Knight 2009; McDonald et al. 2012; McKinney 2010; Pedregosa et al. 2011) for

each set of samples pre-control corrections: forewing (Figure 3.4A), hindwing (3.5), head (Figure 3.6), legs (Figure 3.7), molt (Figure 3.8), pruinosity scraped from the body (Figure 3.8), and soil collected near *M. dorsatus* and *N. p. pruinosus* emergence sites (Figure 3.8). It should be noted that unknown and unidentified mean different things in the ITS tables. Unknown indicates that no resolution was found at this level and unidentified means that there was sequence data available in the UNITE database for that OTU, but it remains currently unresolved taxonomically. Field, buffer, and kit controls primarily consisted of ascomycete fungi in class Dothideomycetes, including *Cladosporium cladosporioides* (Capnodiales: Cladosporiaceae), *Sphaerulina* sp. (Capnodiales: Mycosphaerellaceae) and *Alternaria alternata* (Pleosporales: Pleosporaceae). Only one control, a field control, had greater than 20 reads for any ITS ASV. Both *C. cladosporoides* and *A. alternata* were present in most samples as well, making up anywhere from less than 0.05% to 97.9% of reads. After correcting for contamination in controls and pooling by sample type, but not body part, the top 20 taxonomies were tabulated for early season *M. dorsatus* (MDE) (Table 3.5), late season *M. dorsatus* (MDL) (Table 3.6), early season *N. p. pruinosus* (NPE) (Table 3.7), late season *N. p. pruinosus* (NPL) (Table 3.8), soil collected near *N. p. pruinosus* emergence sites (Table 3.9), and soil collected near *M. dorsatus* emergence sites (Table 3.10).

Ophiocordyceps cf. *longissima* (Kobayasi, 1963) (Ascomycota: Sordariomycetes: Hypocreales: Cordycipitaceae) was found in 27 cicada samples across all body parts and multiple species, including *M. dorsatus* and *N. p. pruinosus*. Most reads were found from "body" samples, including one early *N. p. pruinosus* individual (NPE 01) where 99.8% of reads were from *O.* cf. *longissima*. The next highest proportional sample, with 99.7% of reads, was from the

"body" sample of a late *M. dorsatus* (MDL 46) and 5 soil samples, with most reads being found on the body.

The pathogenic fungus, *Beauveria bassiana* (Ascomycota: Sordariomycetes: Hypocreales: Cordycipitaceae) was found on several individuals, but only substantially (>10 reads) on the legs and head of one early season *M. dorsatus* (samples MDE 64 and 65). In several ASVs it could not be identified to species, and only to genus (Table 3.9; Example sequence can be found in Appendix C).

Discussion

The impetus for this study was to determine if bacterial and fungal communities on cicadas differed both inter- and intraspecifically, as well as on different surfaces of the integument as a result of the antimicrobial structures present. *Neotibicen pruinosus pruinosus* has superhydrophobic forewings with uniform arrays of nanopillars that allow for these properties. *Megatibicen dorsatus* has similar spacing between nanopillars but with a shorter profile (i.e., height) (Oh et al. 2017). These species also inhabit different communities, with *M. dorsatus* residing in prairies and *N. p. pruinosus* inhabiting woodlands and residential areas (Sanborn and Phillips 2013; Dana *pers. obs*). These two species of cicada both have superhydrophobic wings with very similar wettability measures (Oh et al. 2017). In the field, I have observed that *M. dorsatus* wings tend to show greater degrees of wear, with reduced antireflective qualities (i.e., white, semi-opaque patches). As of now, the same has not been observed in *N. p. pruinosus*. This could be a result of the shorter profile of nanopillars, difference in behavior or habitat, a longer adult lifespan, or a combination of factors. This may also result in reduced antimicrobial properties in older *M. dorsatus* individuals.

My results indicate, at least in the case of *M. dorsatus*, that fungal diversity is established quickly after emergence (Figures 3.4 to 3.8). Unlike with eclosing *N. p. pruinosus* adults that can easily be found emerging on tree trunks in mowed residential parks, collecting *M. dorsatus* during or directly after eclosing was not possible due to habitat complexity limiting detectability. This short period of time where adults were inhabiting the prairie community resulted in the quick establishment of fungal diversity not seen in *N. p. pruinosus*. In *N. p. pruinosus* early collected samples had much lower fungal diversity (Figure 3.3).

Soil fungal diversity far exceeded that of any other sample. The average Shannon's entropy value was highest in legs in three of four sample types (not in early collected *N. p. pruinosus*) showing the legs might be the main potential exposure source for adults through contact with the soil (in comparison to wings, head, and body) (Figure 3.3). The low diversity in molt samples was almost certainly an artifact of incomplete inhibitor removal during sample processing based on my previous unpublished work on amplifying DNA from cicada molts. I found that just adding a small amount (2 ul) of molt extract was enough to inhibit a PCR reaction from another sample. The legs appear to be a large source of potential exposure to pathogens; for example, the generalist fungal entomopathogen, *Beauveria bassiana* (Ascomycota: Sordariomycetes: Hypocreales: Cordycipitaceae) was present on the legs of an early season *M. dorsatus* and in small amounts of reads in the soil (ASV sequence can be found in Appendix C).

Another entomopathogen present in our samples was *Metacordyceps chlamydosporia* which was previously found in South Korean cicadas (Kim et al. 2016). In my study, *M.* cf. *chlamydosporia* reads were found in many samples (n=29), including soil samples (associated with both *N. p. pruinosus* and *M. dorsatus*), head, wings, legs, and body of both species. The highest read count (118,792) was found on the molt of a *N. p. pruinosus* individual, suggesting a

potential infection point. However, the telomorph state of *M. chlamydosporia* is only known to grow on mollusks or nematodes (Kepler et al. 2012; Zare et al. 2001). For many of these identities or taxa only assigned to the genus level, it is possible that there are undescribed or unsequenced midwestern taxa. Additional sequence information and morphological study would be necessary for further identification and no fruiting bodies associated with cicadas have been found in the field thus far.

I did not identify any pathogens specifically known to infect cicadas. For example, I did not find any *Massospora* ASV's in my study, even though they are known cicada pathogens that comprise over a dozen species across North America, South America, Australia, and Asia (Ciferri et al. 1957; Macias et al. 2020; Soper 1974; Soper 1981). I expected to find *Massospora* in the soil samples because this is where cicadas are presumed to be infected by its resting spores and my collection site is within the range of the *Magicicada* host range map (although it is possible that this species has since become locally extirpated). Cicadas become exposed to *Massospora* when they construct their emergence tunnel to the soil surface to eclose into their adult form (Soper et al. 1976). Thereafter the fungus rapidly grows, filling the abdomen of the new adult (Soper et al. 1974; Soper et al. 1976). Future studies might replicate using soil samples collected from emergence holes of known carriers of *Massospora*, such as *Okanagana* spp. or *Magicicada* spp.

Plant pathogens

The fungal diversity found in my samples might illustrate the potential for transmission of plant pathogens due to hemipteran feeding mechanisms, but most research has been done on species other than cicadas. Indeed, there is little to no previous evidence of cicadas acting as

vectors (Cornara et al. 2020). While cicadas tend to be more specific on oviposition substrates, I have observed them feeding on a much wider breadth of herbaceous plants (e.g., sumac, mulberry, rose, elderberry, etc). I found evidence of several plant pathogens in some of my samples. For example, *Cercospora canescens* (Ascomycota: Dothideomycetes: Capnodiales: Mycosphaerellaceae) (Table 3.5; Table 3.8) was found on both species of cicadas and is a fungal plant pathogen that infects the leaves of several species of plants of economic importance (e.g., cowpea, Schneider 1973). Strawberry powdery mildew was found on the wings and legs of three different individuals with most reads from one early *M. dorsatus.* This fungal pathogen is an obligate parasite on strawberry leaves, flowers, and fruits and causes economic damage through reduced photosynthesis (Pertot et al. 2007). More research is needed to determine if these pathogens can be transmitted to the plant from cicada feeding damage or other contact. It also might indicate that these plants (strawberries) are present in the environment in which they were found and is a potential avenue to use cicadas, or other insects, as a passive methodology to sample the plant or plant pathogen community through environmental DNA (eDNA). Given that I have found that *M. dorsatus* can live for approximately a month in captivity (Dana *pers. obs.*), they might encounter a large array of plants during this time.

Endosymbiont communities

While my study intended to focus on the ecto-microbiome of cicadas, it revealed insight into the endosymbiont community of these cicada species as well. This was likely due to my methodology. I had hoped to optimize the amount of surface bacteria by submerging the sample in buffer after dissection, but this caused some endosymbionts to leak out from the alimentary canal. Most studies on the endosymbionts of cicadas dissect out the bacteriome, a specialized

organ found in the abdominal body cavity (Matsuura et al. 2018), but some of these endosymbionts have been found in a variety of other body tissues in cicadas (Zheng et al. 2017). Future studies might attempt swabbing the exterior of the cicada, with the caveat that this would also reduce DNA yield and introduce additional sources of contamination from handling.

Much like other sap feeding insects, cicadas have obligate intracellular bacterial symbionts that synthesize amino acids that are lacking in xylem, their primary food source (Buchner 1965). These species of bacteria include *Candidatus* Sulcia muelleri (Moran et al. 2005) (= *Candidatus* Karelsulcia muelleri[2]) (Bacteroidota: Baceroidia: Flavobacteriales: Blattabacteria) and *Candidatus* Hodgkinia cicadicola McCutcheon et al. 2009 (Proteobacteria: Alphaproteobacteria: Hyphomicrobiales) (Oren et al. 2020). *Candidatus* S. muelleri is found only in the suborder Auchenorrhyncha, of which cicadas are a member, and is thus a trait that provides evidence of monophyly of Auchenorrhyncha (Hu et al. 2022; Moran et al. 2005; Skinner et al. 2020; Takiya et al. 2006). In some species of cicadas, *Ca.* Hodgkinia has been lost and replaced with a fungal symbiont that is of entomopathogenic origin (Matsuura et al. 2018).

As expected from previous studies on auchenorrhynchans, my analyses found that the endosymbiont *Candidatus* Sulcia muelleri is present in all species tested, primarily in large reads in body samples (as previously mentioned, "body" refers solely to the abdomen and thorax sections). Indeed, the highest read counts were as high as 19.38% of total reads in the most teneral cicadas, early *Neotibicen pruinosus pruinosus*. Likely, the presence of the bacterial

[2] Oren (2017) created a list of Candidatus taxa, prokaryotes that have not been successfully cultured, to suggest corrected names that meet the requirements of the International Code of Nomenclature of Prokaryotes (ICNP). *Candidatus* Sulcia muelleri did not meet the ICNP requirements as the name already existed in the literature when it was named; thus, Oren suggested *Candidatus* Karelsulcia muelleri as it extended the tribute to the embryologist Karel Šulc who initially discovered the structures that house endosymbiotic bacteria, named bacteriomes (Moran et al. 2005). This is a well-known and still current issue in bacterial and archaeal taxonomy, along with the naming of species from solely genetic information, and indeed *Candidatus* taxa do not have priority according to the ICNP (Murray et al. 2020; Oren 2021). This is well outside the scope of this chapter, so for the purposes of stability and given the common usage in the endosymbiont literature I will use *Ca.* Sulcia muelleri in this study.

endosymbiont in my amplicons was from exposing the inside of the cicada to the buffers during the storage and extraction process, and also in the legs and wings as a result of contamination while cutting cicadas. *Candidatus* Sulcia muelleri is found in the bacteriomes of cicadas (Campbell et al. 2015; Łukasik et al. 2017), but there is evidence that it can be housed in other tissues as well (e.g., salivary glands, testes, ovaries) dependent on species (Zheng et al. 2015).

Candidatus Sulcia muelleri is not the only endosymbiont previously found in cicadas, as it does not provide all required amino acids to make up for the nutrient poor diet of xylem plant fluids. Indeed, in some species of cicadas, researchers have identified two species of *Candidatus* Hodgkinia that co-occur with *Ca.* Sulcia (Van Leuven et al. 2014). Whereas *Ca.* Sulcia has moderate genomic stability, there is often much less genomic stability in more recently acquired endosymbionts (Bennett et al. 2014) like *Ca.* Hodgkinia which would explain why my classifier did not originally identify it. The bacteria *Candidatus* Hodgkinia cicadicola has been described in numerous tribes in Cicadidae, including: Fidicinini (*Diceroprocta semicincta*) (Matsuura et al. 2018; McCutcheon et al. 2009), Lamotialnini (*Magicicada* spp.) (Campbell et al. 2015; McCutcheon et al. 2009), and Polyneurini (*Graptosaltria* spp.) (Matsuura et al. 2018). My study adds *Cicadettana calliope calliope* (tribe Cicadettini) to the groups of cicadas that house *Ca.* Hodgkinia. In fact, within the identified ASVs it appears that there is evidence of at least two variants within the samples (Figure 3.14; example sequences in Appendix C), but further sampling would be required to determine the amino acid synthesis machinery present in each variant. I did not find evidence of *Ca.* Hodgkinia in *M. dorsatus* or *N. p. pruinosus* even with more specific searches, thus indicating that another symbiont has likely taken its place in the production of the missing amino acids.

Multiple ASVs were identified as entomophagous fungi, particularly those in Clavicipitaceae (Ascomycota: Hypocreales). QIIME 2 identified pathogenic *Ophiocordyceps longissima* based on the UNITE database in both soil and cicada body part samples. This species of fungus has been reported only in Japan, China, and South Korea (Li et al. 2002; Kobayasi and Shimizu 1963; Lee and Oh 1998; Sung et al. 2011). However, Matsuura et al. (2018) discovered that phylogenetically similar fungi to pathogenic *Ophiocordyceps* spp. were actually bacterial endosymbionts from the bacteriomes of several species of cicadas. These "yeast-like" endosymbionts (YLS[3]) replaced *Ca.* Hodgkinia, which was likely lost due to further fragmentation of its genome, a common occurrence in endosymbionts. These YLSs help produce the amino acids that *Ca.* Sulcia muelleri does not. While the taxonomy differs among taxa, YLSs have been known to occur in a range of hemipteran hosts, including planthoppers (Suh et al. 2001) and aphids (Suh et al. 2001).

The YLSs that have been documented do not match their host phylogeny and went through multiple acquisition and replacement events (Matsuura et al. 2018; Wang et al. 2022). My finding of ASVs that identified with *Ophiocordyceps* in both *M. dorsatus* and *N. p. pruinosus*, is similar to findings of YLSs in sister taxa in *Cryptotympana* (Cicadidae: Cryptotympanini) and related missing *Ca.* Hodgkinia (Appendix Table C.2). However, the YLSs found in my study have higher percent relatedness to *Meimuna mongolica* (Cicadidae: Dundubiini) than to the YLS present in *Cryptotympana* species which might indicate another

[3] It should be noted that yeast-like endosymbionts (YLSs) does not refer to a specific clade, nor are they necessarily related to the prototypical *Saccharomyces cerevisiae* (Ascomycota: Saccharomycetes: Saccharomycetales: Saccharomycetaceae). Examples of other YLS, beyond those discussed in this paper, include the aphid *Hamiltonaphis styraci* and extracellular YLS (Ascomycota: Hypocreales: Clavipitaceae) (Fukatsu and Ishikawa 1996; Suh et al. 2001) and the drugstore beetle *Stegobium paniceum* and its respective YLS *Symbiotaphrina buchneri* (Ascomycota: Xylonomomycetes: Symbiotaphrinales: Symbiotaphrinaceae) (Baral et al. 2018; Gams and von Arx 1980).

loss and gain in the phylogeny (Matsuura et al. 2018; Wang et al. 2022). Further work would need to be done to confirm this, including sequencing of other regions from the genome of *M. dorsatus* and *N. p. pruinosus* YLSs. Further analyses would help determine the ancestry and relatedness of the YLS to *Ophiocordyceps* spp. to better understand its history of loss and gain in the phylogeny of the tribe Cryptotympanini.

Several other known symbionts were found in my samples, but I propose that they were incidental given their low read counts. For example, *Buchnera* was found on the legs of one late *N. p. pruinosus*, but in very small read count (20). *Buchnera* are obligate endosymbiont bacteria found in the bacteriocytes of aphids and likely were picked up from aphids in the environment (Buchner 1965; Sasaki and Ishikawa 1995). Another symbiont, *Candidatus* Hamiltonella, was found in large read count on the forewings of one early *M. dorsatus* and smaller read count on the legs of the same individual (MDE 12). Other reads found in my samples were under 10. *Candidatus* Hamiltonella is known for being both a defensive symbiont for aphids (Oliver et al. 2003; Oliver et al. 2008) as well as providing nutrition as a secondary endosymbiont for whiteflies (Su et al. 2014). One individual of early *N. p. pruinosus* had large read counts in the head, body, legs, and wings (67,786–417,965) of *Candidatus* Xiphenematobacter, which is a known endosymbiont of longidorid nematodes (Mobasseri et al. 2019). I did not include primer sets for eukaryotes to further elucidate if these reads might have come from a pathogenic nematode, as early trials showed that primarily cicada sequences would be returned from eukaryote primers.

The species complex *Pantoea-Erwinia* (Proteobacteria: Gammaproteobacteria: Enterobacterales: Erwinaceae) was found in high read counts across all sample types, although it was negligible in controls. This species complex is difficult to resolve and would require

multilocus sequence analyses to determine what species were present in my samples (Zhang and Qiu 2015). Zhou et al. (2015) also found *Pantoea-Erwinia* in the gut microbiome of another species of cicada, *Meimuna mongolica*, although their read counts were much lower proportionally. The role in the gut is uncertain, as *Pantoea* has been shown to be both pathogenic and beneficial in insects, dependent upon the host (for a review see Walterson and Stavrinides 2015).

Limitations of the current state of microbiome studies

Excitement about the results should be tempered by the knowledge that while large quantities of data can be obtained through large sequencing efforts such as this study, without culturing the bacteria or fungi it isn't possible to know the absolute identity of the species found. Knowledge of the sequence identified ASVs is only as good as the respective databases. If a bacterium or fungus was not cultured and subsequently sequenced then it cannot be properly identified. The UNITE database, for example, has been continuously updated since its inception in 2003, but there still remains a large portion of OTUs that cannot be identified even to phylum (Nilsson et al. 2019), as seen in this study (e.g., Table 3.7). Phukhamsasakda et al. (2022) examined the estimates of documented fungi species by looking at a set of genera, describing new species, and utilizing the UNITE database. They found that only examining the ITS region was insufficient to demarcate species, and that some species were only described using morphology (not ITS sequences). Available sequences also vary between online databases UNITE and PubMed, and a great number of species listed in Index Fungorum have no available sequence data. This can also reduce the ability to correctly assign taxonomy and is likely the

reason that many of the recovered ASVs have no resolution, even at the higher levels (i.e., class or even phylum).

During library preparation it is assumed that there will be some degree of bias due to amplification biases in the creation of the library (Degnan and Ochman 2012). While the Fluidigm system has many benefits compared to other methods, it does show decreased alpha diversity in comparison (Mallott et al. 2019). In addition, contaminants present in extractions can reduce the sequencing depth of microbes in lower abundance and even miss them completely (Laurence et al. 2014). This can also reduce the level of diversity estimated with methods like Shannon's diversity index, especially in studies like ours where small amounts of DNA are attempting to be amplified and sequenced. In particular this bias can be seen in our Archaea results where soil "reads" or number of sequences produced were primarily found in the soil samples, with very little depth in any cicada associated samples (Table 3.4; Figure 3.2c).

Finally, the presence of DNA on the wings does not tell us if the bacteria or fungus was alive at the time of collection. The wings might still be killing, or impaling, the potential pathogens on the nanopillars, but we are limited by our methodology. However, we might expect then that we would have smaller read counts for organisms that were not reproducing on the surface. This lends more credibility to our findings due to my focus on analysis of samples with higher read counts (>100).

Implications and future work

In this study, I found that legs are the likely source point of exposure to an external microbial community. This might support that wings protect the abdomen and thorax from exposure to entomopathogens, but we do see lower diversity of microbes on the heads as well.

We know that with greater sampling depth diversity plateaus at a final, correct value and this would be true for microbial studies as well. Shannon's entropy measures should account for this, but there is a possibility that contamination may be constraining our evaluation of the true level of diversity. More study is needed to elucidate if the bacteria and fungi on the wings are indeed alive and able to reproduce. While the findings on the endosymbiont community were unexpected due to the attempt of this study to focus on the ectobiome, it provides insight into the placement of the YLS as they relate to *Ophiocordyceps* spp. and the cicada phylogenetic tree. This will allow for more evidence of gains and losses of both *Ca.* Hodgkinia and YLS across a number of cicada tribes and species. Cicadas being a potential vector of plant pathogens was also unexpected and deserves more study to see if they are carrying these diseases across the larger landscape. How their uniquely large body size and subsequent increase in surface area plays into this dynamic is also an interesting avenue of study. There are many more sequences available for study in my dataset that will be part of a larger effort with microbiologists and phylogeneticists to further explore.

Tables

Table 3.1. Primers used in Fluidigm preparation of library. Given nucleotide changes in primers for sequencing optimization over the years, authors for primer are listed by precedence. Approximate product length does not include additional index or adapters used in amplification and sequencing but does include amplified region along with the locus specific primer sequences. The length may also be variable, especially in the case of the ITS3-ITS4 region.

Target	Primer Name	Primer Sequence (5' to 3')	Primer length (bp)	Approximate Product Length (bp)	Primer Reference(s)
Archaea	Arch349F	GYGCASCAGKCGMGAAW	17	457	Takai and Horikoshi 2000
	Arch806R	GGACTACVSGGGTATCTAAT	20		Takai and Horikoshi 2000
ITS3-ITS4	ITS3F	GCATCGATGAAGAACGCAGC	20	356	White et al. 1990
	ITS4R	TCCTCCGCTTATTGATATGC	20		White et al. 1990
16S V1-V3	F28-2-for (Gray28F)	GAGTTTGATCNTGGCTCAG	19	492	Lane 1991; Ishak et al. 2011
	R519-2-rev	GTNTTACNGCGGCKGCTG	18		Sørensen and Teske 2006; Ishak et al. 2011
16S V3-V4	V3-F357-N	CCTACGGGNGGCWGCAG	17	449	Herlemann et al. 2011
	V4 805R	GACTACHVGGGTATCTAATCC	21		Herlemann et al. 2011
16S V4	515F-Y	GTGYCAGCMGCCGCGGTAA	22	292	Turner et al. 1999; Caporaso et al. 2011; Quince et al. 2011; Parada et al. 2016
	806RB	GGACTACNVGGGTWTCTAAT	20		Caporaso et al. 2011; Apprill et al. 2015
16S V3-V5	F357	CCTACGGGAGGCAGCAG	17	570	Muyzer et al. 1993
	R926	CCGTCAATTCMTTTRAGT	18		Lane 1991

Table 3.2. Sequencing metrics after sorting by primer.

Primer Set	Direction	Number of Reads
Arch349F_Arch806R	Forward	6,203,232
Arch349F_Arch806R	Reverse	6,203,232
ITS3_ITS4	Forward	100,797,343
ITS3_ITS4	Reverse	100,797,343
V1_F28_V3_R519	Forward	67,451,478
V1_F28_V3_R519	Reverse	67,451,478
V3_F357_N_V4_R805	Forward	90,728,331
V3_F357_N_V4_R805	Reverse	90,728,331
V4_515F_New_V4_806R New	Forward	35,559,571
V4_515F_New_V4_806R_New	Reverse	35,559,571
V3_F357_V5_R926	Forward	97,386,130
V3_F357_V5_R926	Reverse	97,386,130
Total Reads:		**796,252,170**

Table 3.3. Average sequence length (bp) after dada2 filtering for archaea, fungi (ITS3-ITS4), and bacteria (V3-V4).

	Mean (bp)	Min	Max	Std Dev
Arch	379.7	233	450	61.3
ITS3-ITS4	325.1	230	447	37.7
V3-V4	417.0	233	450	12.1

Table 3.4. Frequencies of (a) features per sample and (b) reads per feature after dada2 filtering for archaea (Arch), fungi (ITS3-ITS4), and bacteria (V3-V4).

a)

	Min	1st quartile	Median	3rd quartile	Max	Mean
Arch	0	24	91	994	459,015	17,867
ITS3-ITS4	0	24,948	284,301	484,398	1,151,266	293,672
V3-V4	0	102,646	201,990	351,312	1,687,758	263,808

b)

	Min	1st quartile	Median	3rd quartile	Max	Mean
Arch	1	29	111	367	273,395	1,127
ITS3-ITS4	1	107	279	945	16,550,654	5,700
V3-V4	1	3	15	94	5,676,010	880

Table 3.5. Top 20 assigned taxonomies, operational taxonomic units (OTUs), for MDE (early *Megatibicen dorsatus*) ASVs in order of total proportion of reads, pooling all body parts and sample sets for the MDE group. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*	*cladosporioides*	36.83%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*	*alternata*	15.90%
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	*Filobasidium*	*oeirense*	3.40%
Unidentified						2.31%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Unidentified		1.94%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Septoria*		1.75%
Ascomycota	Dothideomycetes	Pleosporales				1.30%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*		1.30%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Cercospora*	*canescens*	1.27%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Neocladosporium*	*leucadendri*	1.23%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*		1.16%
Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	*Bulleromyces*	*albus*	1.12%
Ascomycota	Dothideomycetes	Dothideales	Aureobasidiaceae	*Aureobasidium*	*pullulans*	1.00%
Basidiomycota	Exobasidiomycetes	Entylomatales	Entylomatales fam Incertae sedis	*Tilletiopsis*	*washingtonensis*	0.91%
Ascomycota						0.91%
Ascomycota	Leotiomycetes	Erysiphales	Erysiphaceae	*Podosphaera*	*aphanis*	0.81%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Zymoseptoria*		0.72%
Ascomycota	Leotiomycetes	Erysiphales	Erysiphaceae	*Golovinomyces*		0.68%
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	*Metschnikowia*		0.67%
Unidentified						0.66%

Table 3.6. Top 20 assigned taxonomies for MDL (late *Megatibicen dorsatus*) ASVs to the species level in order of total proportion of reads, pooling all body parts and sample sets for the MDE group. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*	*cladosporioides*	38.10%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*	*alternata*	27.10%
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	*Filobasidium*	*oeirense*	3.51%
Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	*Bulleromyces*	*albus*	3.05%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*		2.71%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae			1.77%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*		1.58%
Basidiomycota	Spiculogloeomycetes	Spiculogloeales	Spiculogloeaceae	*Phyllozyma*	*linderae*	1.43%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	*Neoascochyta*	*rosicola*	1.34%
Unknown						1.29%
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	*Vishniacozyma*	*victoriae*	0.86%
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales fam Incertae sedis	*Sarocladium*	*strictum*	0.74%
Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	*Papiliotrema*	*aurea*	0.67%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	unidentified		0.64%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Zymoseptoria*		0.64%
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	*Sporobolomyces*	*roseus*	0.63%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae			0.62%
Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	*Papiliotrema*	*nemorosus*	0.61%
Ascomycota	Sordariomycetes	Hypocreales	Clavicipitaceae	*Metarhizium*		0.56%
Basidiomycota	Exobasidiomycetes	Entylomatales	Entylomatales fam Incertae sedis	*Tilletiopsis*	*washingtonensis*	0.49%

Table 3.7. Top 20 ITS reads for NPE samples (early collected *Neotibicen pruinosus pruinosus*) to the species level in order of total proportion of reads, pooling all body parts and sample sets. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Ascomycota	Sordariomycetes	Hypocreales	Ophiocordycipitaceae	*Ophiocordyceps*	*longissima*	26.53%
Unknown						24.52%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*	*cladosporioides*	13.16%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae			5.01%
Ascomycota	Pezizomycotina cls Incertae sedis	Pezizomycotina ord Incertae sedis	Pezizomycotina fam Incertae sedis	*Ciliophora*	unidentified	5.01%
Ascomycota	Eurotiomycetes	Chaetothyriales	unidentified	unidentified	unidentified	4.93%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*	*alternata*	4.12%
Ascomycota	Lecanoromycetes	Caliciales	Physciaceae	*Physcia*		3.86%
Basidiomycota	Exobasidiomycetes	Entylomatales	Entylomatales fam Incertae sedis	*Tilletiopsis*	*lilacina*	3.22%
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	*Filobasidium*	*oeirense*	2.85%
Ascomycota	Dothideomycetes	Dothideales	Dothideaceae	*Coniozyma*	*leucospermi*	1.64%
Ascomycota	Dothideomycetes	Pleosporales	unidentified	unidentified	unidentified	1.43%
Ascomycota	Dothideomycetes	Dothideales	Aureobasidiaceae	*Aureobasidium*	*pullulans*	1.37%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*		0.99%
Basidiomycota	Agaricomycetes	Cantharellales	Ceratobasidiaceae	*Thanatephorus*	*cucumeris*	0.71%
Chytridiomycota						0.26%
Ascomycota	Lecanoromycetes	Caliciales	Physciaceae	*Physcia*	*millegrana*	0.24%
Chytridiomycota	unidentified	unidentified	unidentified	unidentified	unidentified	0.07%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*		0.02%
Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	*Bulleromyces*	*albus*	0.01%

Table 3.8. Top 20 ITS reads for NPL samples (late collected *Neotibicen pruinosus pruinosus*) to the species level in order of total proportion of reads, pooling all body parts and sample sets. Percent reads is based on read count after correcting from contamination found present in associated controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*	*cladosporioides*	23.25%
Unknown						11.26%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae			8.04%
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*	*alternata*	6.91%
Ascomycota	Lecanoromycetes	Caliciales	Physciaceae	*Phaeophyscia*	*adiastola*	6.01%
Ascomycota	Dothideomycetes	Dothideales	Aureobasidiaceae	*Aureobasidium*	*pullulans*	4.05%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Cladosporium*		3.85%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Sphaerulina*		3.48%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Verrucocladosporium*	*dirinae*	2.99%
Ascomycota						1.76%
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	*Filobasidium*	*oeirense*	1.38%
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	*Epicoccum*	*nigrum*	1.38%
Ascomycota	Dothideomycetes	Capnodiales	Neodevriesiaceae	*Neodevriesia*	unidentified	1.12%
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	*Toxicocladosporium*	*strelitziae*	1.02%
Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae			0.92%
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	*Cercospora*	*canescens*	0.91%
Ascomycota	Lecanoromycetes	Caliciales	Physciaceae	*Hyperphyscia*	*adglutinata*	0.89%
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	*Engyodontium*	*album*	0.88%
Ascomycota	Lecanoromycetes	Caliciales	Physciaceae	*Physcia*	unidentified	0.87%
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	*Hannaella*	*luteola*	0.81%

Table 3.9. Top 20 ITS reads for soil collected near *Neotibicen pruinosus pruinosus* emergence sites. Percent reads is based on read count after correcting from contamination found present in associated controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Ascomycota	Pezizomycetes	Pezizales	Pyronemataceae	unidentified	unidentified	10.92%
Rozellomycota	Rozellomycotina cls Incertae sedis	GS11	unidentified	unidentified	unidentified	7.04%
Rozellomycota	unidentified	unidentified	unidentified	unidentified	unidentified	4.45%
Unknown						4.18%
Ascomycota	Eurotiomycetes	unidentified	unidentified	unidentified	unidentified	4.10%
Ascomycota	Sordariomycetes	Sordariales				4.00%
Mortierellomycota	Mortierellomycetes	Mortierellales	Mortierellaceae	*Mortierella*	*exigua*	3.11%
Ascomycota	Leotiomycetes	Helotiales	Helotiaceae	*Tetracladium*	unidentified	2.53%
Ascomycota	Sordariomycetes	Hypocreales	Bionectriaceae	*Clonostachys*	*rosea*	2.05%
Mortierellomycota	Mortierellomycetes	Mortierellales	Mortierellaceae	*Mortierella*	*zonata*	2.03%
Ascomycota	Sordariomycetes	Glomerellales	Plectosphaerellaceae	*Gibellulopsis*	*unidentified*	1.71%
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	*Exophiala*	*equina*	1.55%
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	*Cylindrocarpon*	unidentified	1.48%
Ascomycota	Sordariomycetes	Hypocreales	Bionectriaceae	*Bionectria*	*rossmaniae*	1.48%
Basidiomycota	Agaricomycetes	Agaricales	Lycoperdaceae	*Lycoperdon*	*pratense*	1.19%
Rozellomycota	Rozellomycotina cls Incertae sedis	GS10	unidentified	unidentified	unidentified	1.15%
Ascomycota	Sordariomycetes	Glomerellales	Plectosphaerellaceae	*Plectosphaerella*	*cucumerina*	1.14%
Basidiomycota	Agaricomycetes	Agaricales	Clavariaceae	*Clavaria*	unidentified	1.07%
Ascomycota	Leotiomycetes	Helotiales	Dermateaceae	*Blumeriella*	unidentified	1.04%
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	*Beauveria*		1.03%

Table 3.10. Top 20 ITS reads for soil collected near *Megatibicen dorsatus* emergence sites (MDE). Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Species	Proportion reads
Unknown						9.24%
Rozellomycota	unidentified	unidentified	unidentified	unidentified	unidentified	8.47%
Basidiomycota	Ustilaginomycetes	Urocystidales	Glomosporiaceae	*Thecaphora*	*frezzii*	4.53%
Ascomycota	Sordariomycetes	Sordariales				3.12%
Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	*Trichoderma*	*erinaceum*	2.75%
Rozellomycota	Rozellomycotina cls Incertae sedis	GS11	unidentified	unidentified	unidentified	2.66%
Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	*Xylodon*	*hyphodontinus*	2.59%
Basidiomycota	Agaricomycetes	Agaricales	Hygrophoraceae	*Hygrocybe*	unidentified	2.45%
Kickxellomycota	GS19	unidentified	unidentified	unidentified	unidentified	2.30%
Ascomycota						2.18%
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	*Fusarium*	*solani*	2.00%
Mortierellomycota	Mortierellomycetes	Mortierellales	Mortierellaceae	*Mortierella*	*zonata*	1.96%
Basidiomycota	Agaricomycetes	Agaricales	Stephanosporaceae	*Lindtneria*	*flava*	1.88%
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae			1.73%
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	*Schwanniomyces*	*occidentalis*	1.66%
Mortierellomycota	Mortierellomycetes	Mortierellales	Mortierellaceae	*Mortierella*	*exigua*	1.61%
Basidiomycota	Agaricomycetes	Phallales	Phallaceae	*Phallus*	*hadriani*	1.54%
Ascomycota	Dothideomycetes	Pleosporales				1.48%
Basidiomycota						1.47%
Ascomycota	Leotiomycetes	Helotiales	Helotiaceae	*Tetracladium*	unidentified	1.43%

Table 3.11. Top 20 V3-V4 reads for early collected Megatibicen dorsatus (MDE) summing all body parts, ordered by proportion of total reads the particular sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Proteobacteria	Gammaproteobacteria	Enterobacterales	Erwiniaceae	*Pantoea*	34.73%
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sphingomonas*	15.05%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Methylobacterium-Methylorubrum*	11.66%
Bacteroidota	Bacteroidia	Flavobacteriales	Blattabacteriaceae	Candidatus *Sulcia*	6.85%
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	6.82%
Proteobacteria	Gammaproteobacteria	Burkholderiales	Oxalobacteraceae	*Massilia*	2.09%
Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	*Curtobacterium*	1.68%
Bacteroidota	Bacteroidia	Cytophagales	Hymenobacteraceae	*Hymenobacter*	1.37%
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	1.29%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Aureimonas*	1.19%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Escherichia-Shigella*	0.98%
Patescibacteria	Parcubacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	0.94%
Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	*Staphylococcus*	0.84%
Bacteroidota	Bacteroidia	Flavobacteriales	Weeksellaceae	*Chryseobacterium*	0.83%
Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadales	*Saccharimonadales*	0.67%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*	0.67%
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	*Brevundimonas*	0.64%
Proteobacteria	Gammaproteobacteria	Enterobacterales			0.60%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Enterobacter*	0.57%
Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Roseomonas*	0.53%

Table 3.12. Top 20 V3-V4 reads for late collected *Megatibicen dorsatus* (MDL) summing all body parts, ordered by proportion of total reads the particular sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae		20.28%
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	18.91%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Erwiniaceae	*Pantoea*	12.43%
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sphingomonas*	10.74%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Enterobacter*	6.58%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Methylobacterium-Methylorubrum*	6.07%
Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	*Leifsonia*	5.26%
Bacteroidota	Bacteroidia	Flavobacteriales	Blattabacteriaceae	*Candidatus* Sulcia	3.27%
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	2.45%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Aureimonas*	1.85%
Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae		1.75%
Proteobacteria	Gammaproteobacteria	Enterobacterales			1.52%
Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	*Curtobacterium*	1.04%
Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Roseomonas*	0.74%
Bacteroidota	Bacteroidia	Cytophagales	Hymenobacteraceae	*Hymenobacter*	0.73%
Proteobacteria	Gammaproteobacteria	Burkholderiales	Oxalobacteraceae	*Massilia*	0.52%
Patescibacteria	Parcubacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	0.47%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae		0.46%
Firmicutes	Bacilli	Staphylococcales	Staphylococcaceae	*Staphylococcus*	0.41%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*	0.39%

Table 3.13. Top 20 V3-V4 reads for early collected Neotibicen pruinosus pruinosus summing all body parts, ordered by proportion of total reads the particular sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Bacteroidota	Bacteroidia	Flavobacteriales	Blattabacteriaceae	Candidatus *Sulcia*	19.38%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Enterobacter*	7.41%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Escherichia-Shigella*	6.78%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae		5.65%
Bacteroidota	Bacteroidia	Flavobacteriales	Crocinitomicaceae	*Fluviicola*	4.35%
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Thermomonas*	2.98%
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	2.45%
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae		2.14%
Actinobacteriota	Thermoleophilia	Gaiellales	Gaiellaceae	*Gaiella*	1.98%
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*	1.76%
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	*Niabella*	1.72%
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Pseudoxanthomonas*	1.59%
Verrucomicrobiota	Verrucomicrobiae	Verrucomicrobiales	Rubritaleaceae	*Luteolibacter*	1.39%
Patescibacteria	Parcubacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	*Candidatus* Nomurabacteria	1.31%
Planctomycetota	Planctomycetes	Pirellulales	Pirellulaceae	uncultured	1.29%
Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	*Arthrobacter*	1.26%
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sphingomonas*	1.23%
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	uncultured	uncultured	1.00%
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	uncultured	0.96%
Proteobacteria	Gammaproteobacteria	Burkholderiales	Oxalobacteraceae	*Massilia*	0.96%

Table 3.14. Top 20 V3-V4 reads for late collected *Neotibicen pruinosus pruinosus* (NPL) summing all body parts, ordered by proportion of total reads the particular sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Proteobacteria	Gammaproteobacteria	Enterobacterales	Erwiniaceae	*Pantoea*	23.88%
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Xiphinematobacteraceae	*Candidatus* Xiphinematobacter	16.61%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae		16.29%
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*	8.84%
Bacteroidota	Bacteroidia	Flavobacteriales	Blattabacteriaceae	*Candidatus* Sulcia	6.29%
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Escherichia-Shigella*	2.40%
Actinobacteriota	Actinobacteria	Pseudonocardiales	Pseudonocardiaceae	*Actinomycetospora*	1.96%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Methylobacterium-Methylorubrum*	1.96%
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Sphingomonas*	1.88%
Myxococcota	Polyangia	Polyangiales	Polyangiaceae		1.50%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae		1.17%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Methylocella*	0.97%
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	0.96%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	1174-901-12	0.78%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae		0.75%
Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Pseudarthrobacter	0.66%
Myxococcota	Polyangia	Polyangiales			0.64%
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	0.62%
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae		0.53%
Bacteroidota	Bacteroidia	Cytophagales	Hymenobacteraceae	*Hymenobacter*	0.50%

Table 3.15. Top 20 V3-V4 reads for soil collected near *Neotibicen pruinosus pruinosus* emergence sites, ordered by proportion of total reads across all sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	*Candidatus* Udaeobacter	6.88%
Actinobacteriota	Thermoleophilia	Gaiellales	uncultured	uncultured	5.87%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	uncultured	4.79%
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	uncultured	uncultured	4.04%
Actinobacteriota	Thermoleophilia	Solirubrobacterales	67-14	67-14	3.99%
Actinobacteriota	MB-A2-108	MB-A2-108	MB-A2-108	MB-A2-108	3.26%
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Xiphinematobacteraceae	*Candidatus* Xiphinematobacter	2.92%
Actinobacteriota	Thermoleophilia	Gaiellales	Gaiellaceae	*Gaiella*	2.89%
Actinobacteriota	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	2.18%
Planctomycetota	Planctomycetes	Gemmatales	Gemmataceae	uncultured	2.11%
Chloroflexi	KD4-96	KD4-96	KD4-96	KD4-96	2.06%
Actinobacteriota	Acidimicrobiia	IMCC26256	IMCC26256	IMCC26256	1.62%
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	Vicinamibacteraceae	Vicinamibacteraceae	1.49%
Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	1.48%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae		1.37%
Unknown					1.32%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	uncultured	1.21%
Methylomirabilota	Methylomirabilia	Rokubacteriales	Rokubacteriales	*Rokubacteriales*	1.04%
Actinobacteriota	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	*Solirubrobacter*	1.04%
Unknown					1.03%

Table 3.16. Top 20 V3-V4 reads for soil collected near Megatibicen dorsatus emergence sites, ordered by proportion of total reads across all sample type. Percent reads is based on read count after correcting from contamination found present in controls.

Phylum	Class	Order	Family	Genus	Proportion reads
Actinobacteriota	Thermoleophilia	Gaiellales	uncultured	uncultured	5.08%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	uncultured	4.94%
Actinobacteriota	Thermoleophilia	Solirubrobacterales	67-14	67-14	4.92%
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	uncultured	uncultured	3.36%
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	*Candidatus* Udaeobacter	3.23%
Planctomycetota	Planctomycetes	Gemmatales	Gemmataceae	uncultured	2.78%
Actinobacteriota	Thermoleophilia	Gaiellales	Gaiellaceae	*Gaiella*	2.72%
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Xiphinematobacteraceae	*Candidatus* Xiphinematobacter	2.22%
Actinobacteriota	MB-A2-108	MB-A2-108	MB-A2-108	MB-A2-108	1.79%
Chloroflexi	KD4-96	KD4-96	KD4-96	KD4-96	1.74%
Actinobacteriota	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	1.71%
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	Vicinamibacteraceae	*Vicinamibacteraceae*	1.56%
Actinobacteriota	Acidimicrobiia	IMCC26256	IMCC26256	IMCC26256	1.53%
Actinobacteriota	Acidimicrobiia	Microtrichales	uncultured	uncultured	1.48%
Methylomirabilota	Methylomirabilia	Rokubacteriales	Rokubacteriales	*Rokubacteriales*	1.34%
Unknown					1.31%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae		1.30%
Actinobacteriota	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	*Solirubrobacter*	1.13%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Microvirga*	1.08%
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	*Methyloligellaceae*	0.99%

Figures

Figure 3.1. *Neotibicen pruinosus pruinosus* adult showing dissected parts used in library creation: (1) head; (2) legs; (3) body; (4) hindwings; and (5) forewings. Not shown: molt.

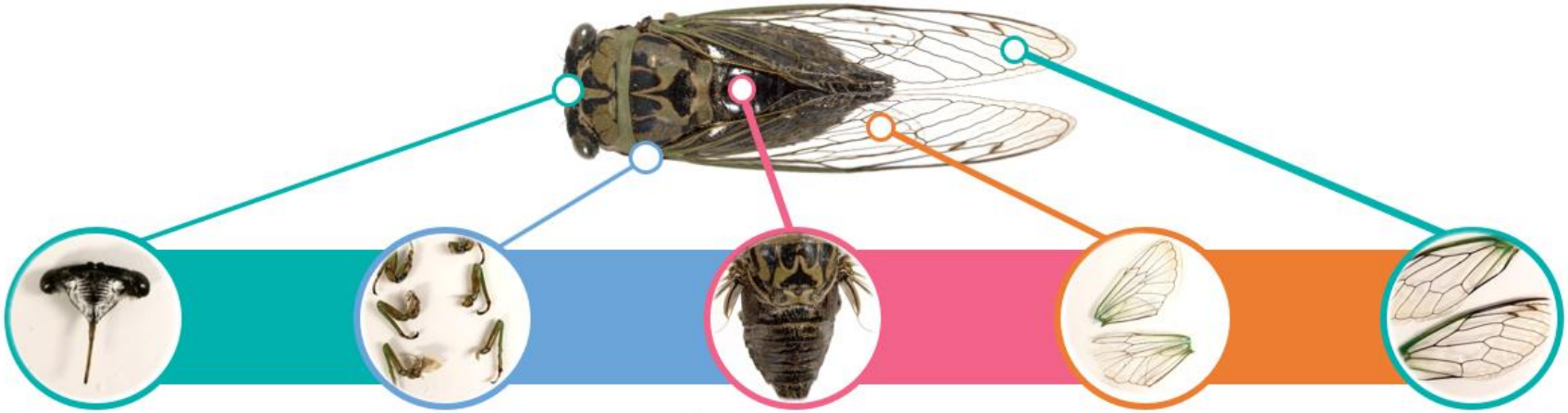
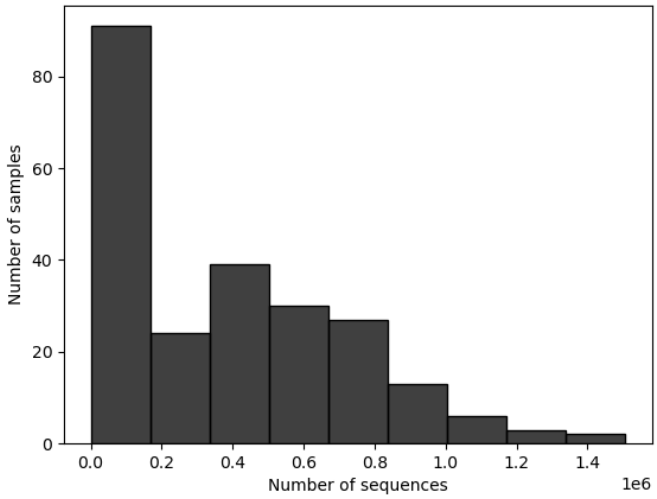
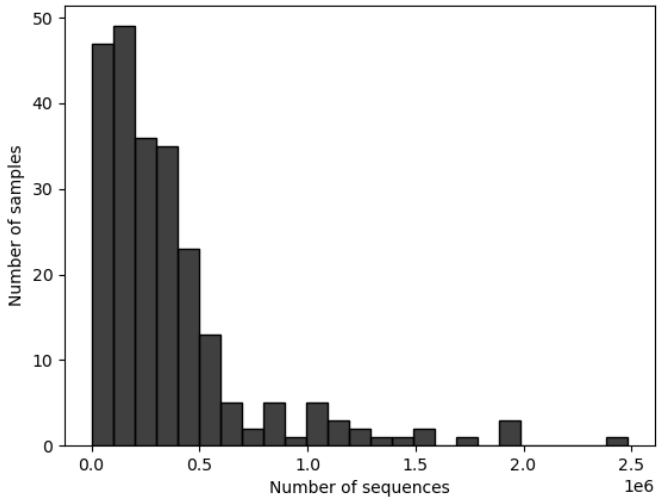


Figure 3.2. Distribution of (a) ITS, (b) V3V4 and (c) Archaea read counts per sample in paired reads.

a) ITS3-ITS4 reads



b) V3-V4 reads



(c) Archaea reads

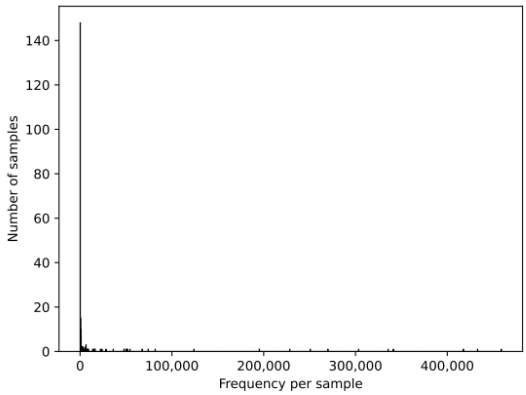


Figure 3.3 Alpha diversity visualized using Shannon entropy (Gauthier and Derome 2021). measurements of fungal ASVs (ITS) by category. The molt category is not representative of the diversity present due to presence of inhibitors not allowing for any PCR amplification in some samples, despite the presence of DNA.

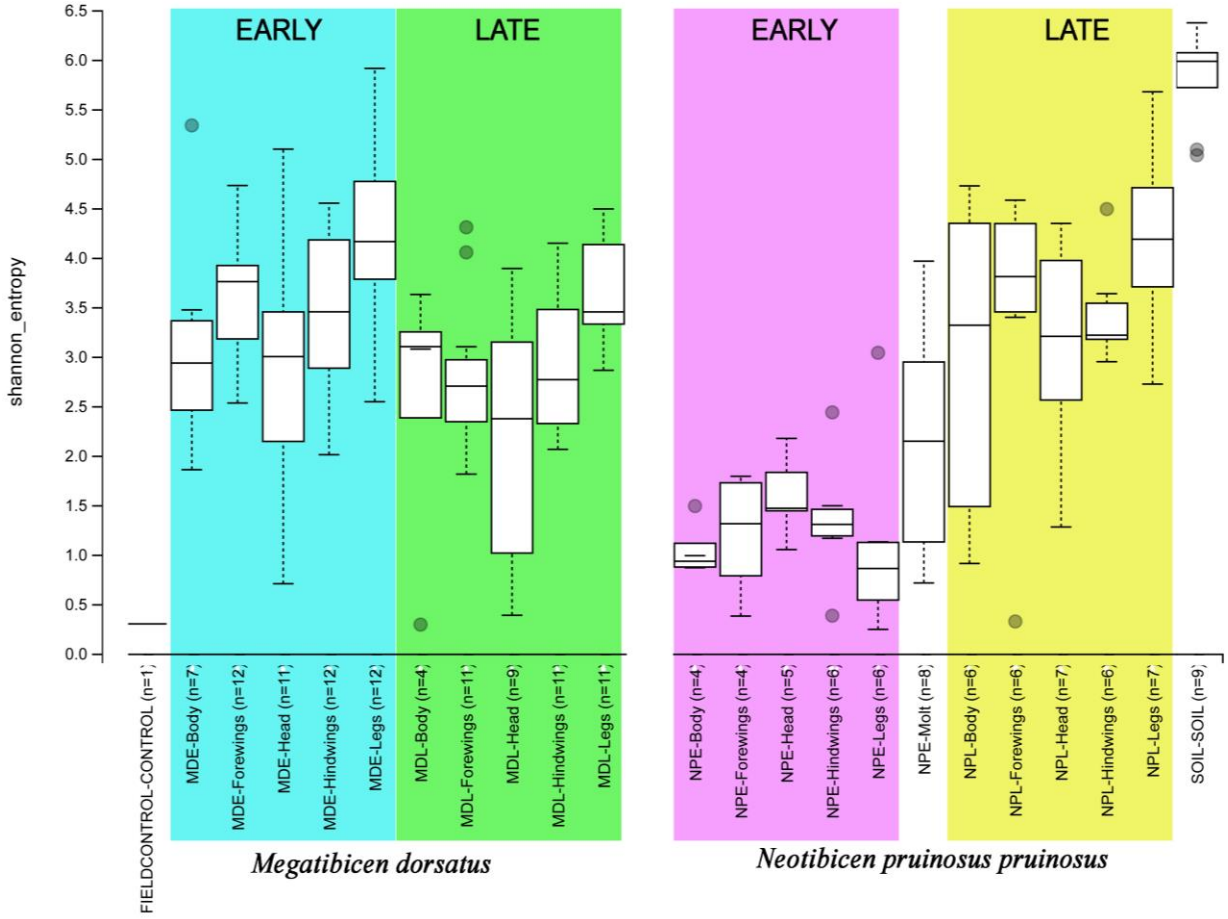


Figure 3.4.A. Forewing QIIME2 View generated species-level ITS taxonomic feature bar plot of unique fungal sequences (ASVs). Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.4.B.

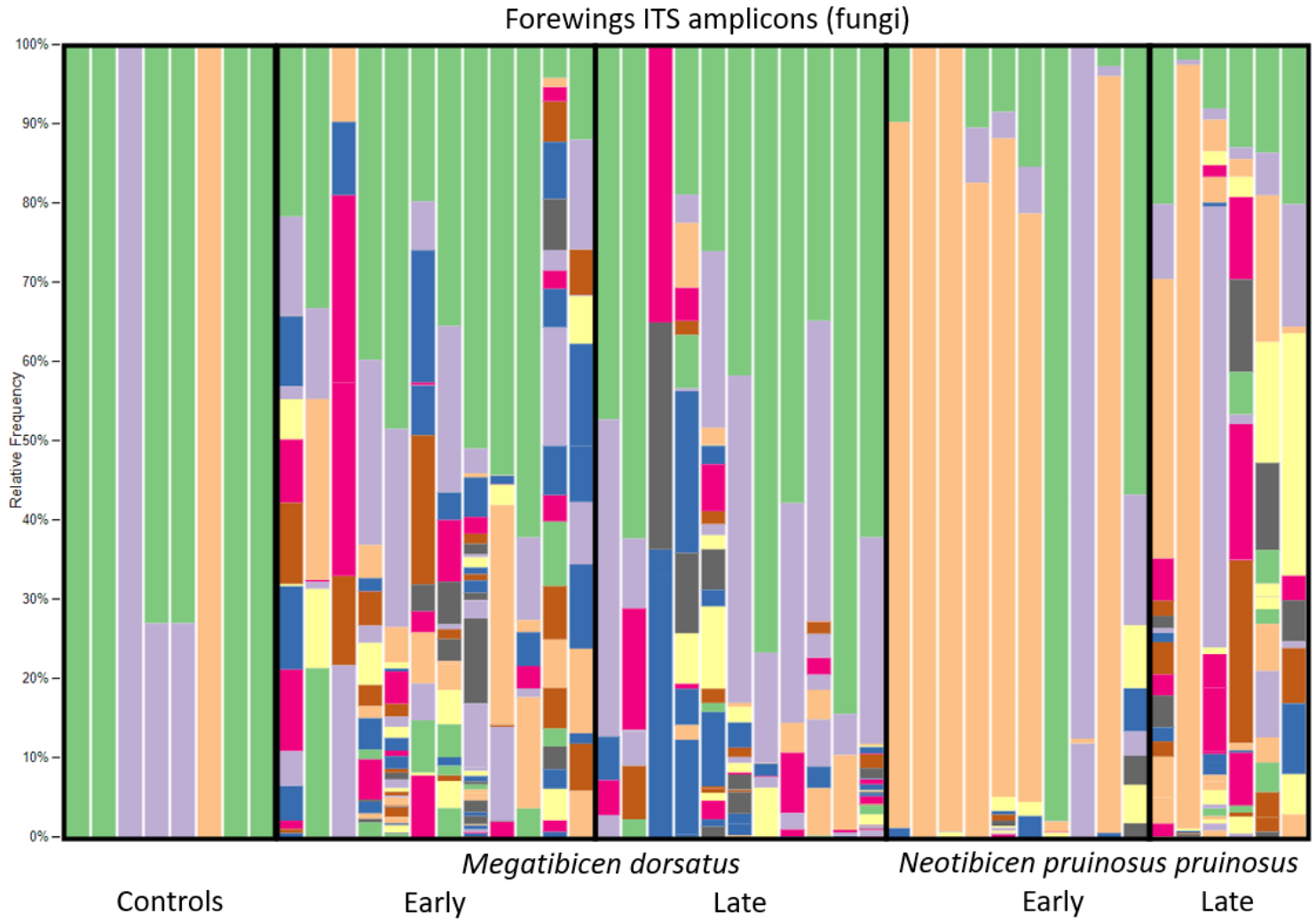


Figure 3.4.B. Legend of top 20 ASV identifications from ITS3-ITS4 amplified sequences. Full data set of assigned taxonomy assignment and read counts can be found as external file.

Figure 3.5. Hindwing QIIME2 View generated species-level ITS taxonomy of fungal sequences (ITS) by species and time points for samples. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.4.B.

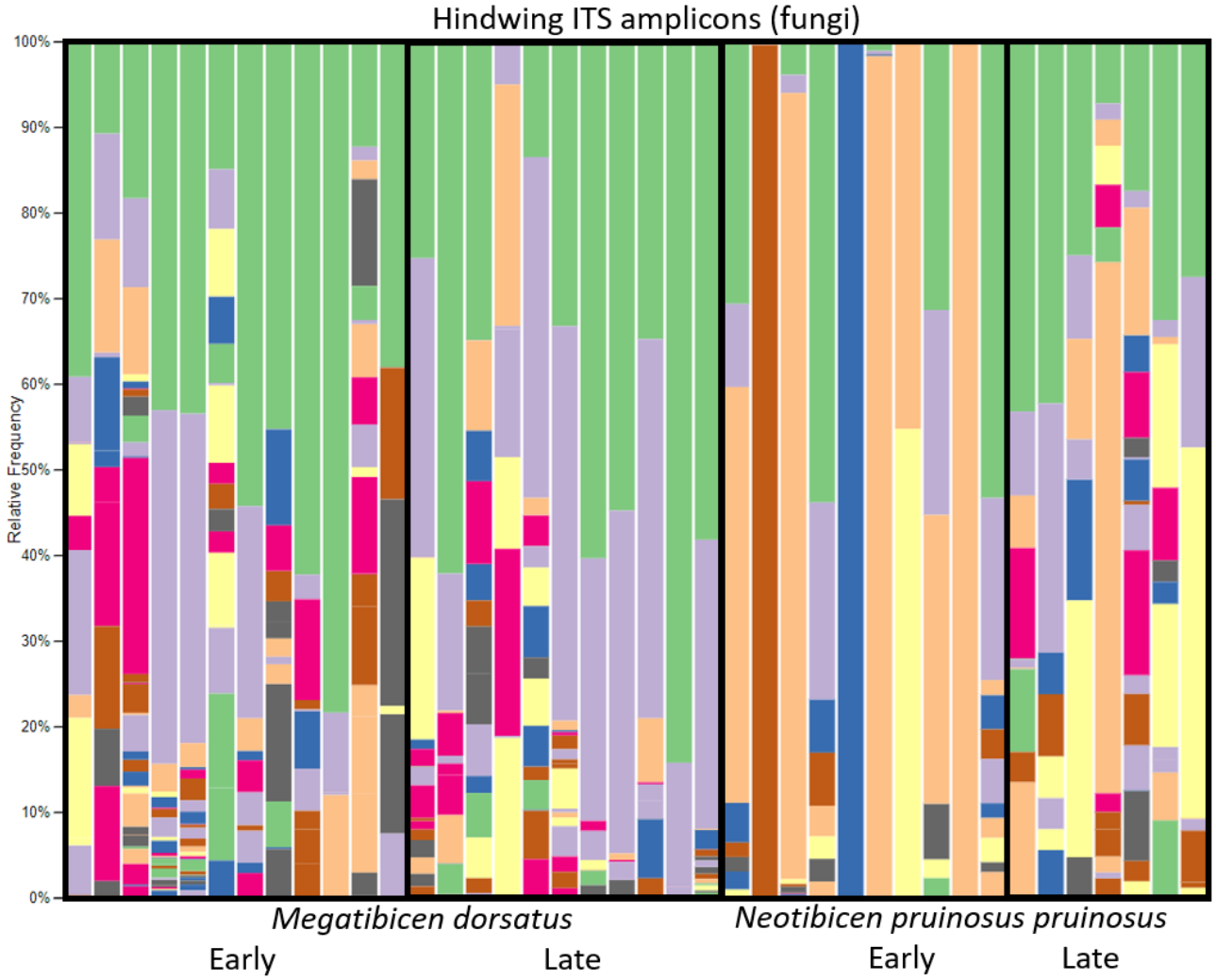


Figure 3.6. Head QIIME2 View generated species-level ITS taxonomic feature bar plot of fungal ASVs. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.4.B.

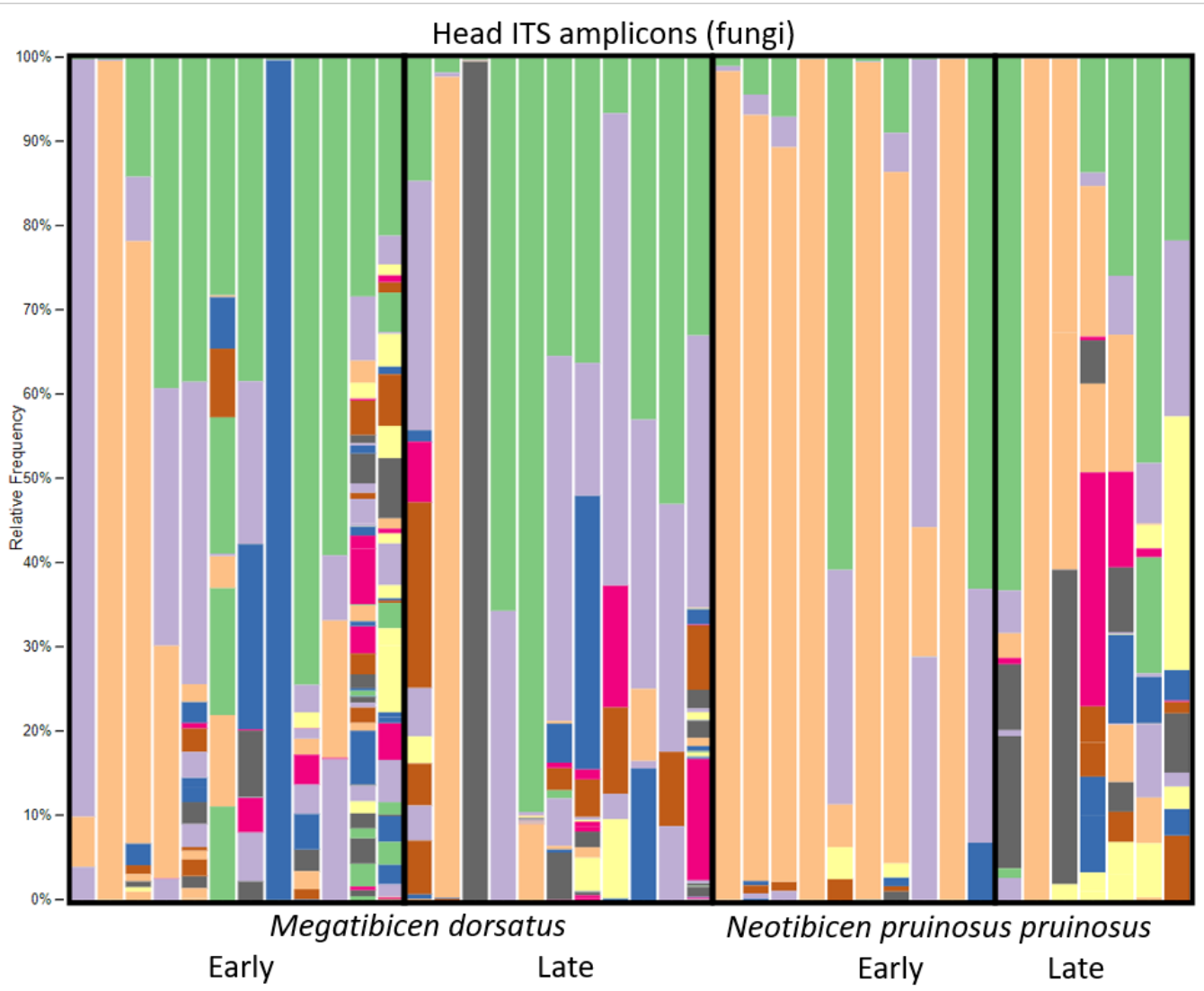


Figure 3.7. Leg QIIME2 View generated species level ITS assigned taxonomy of fungal sequences. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.4.B.

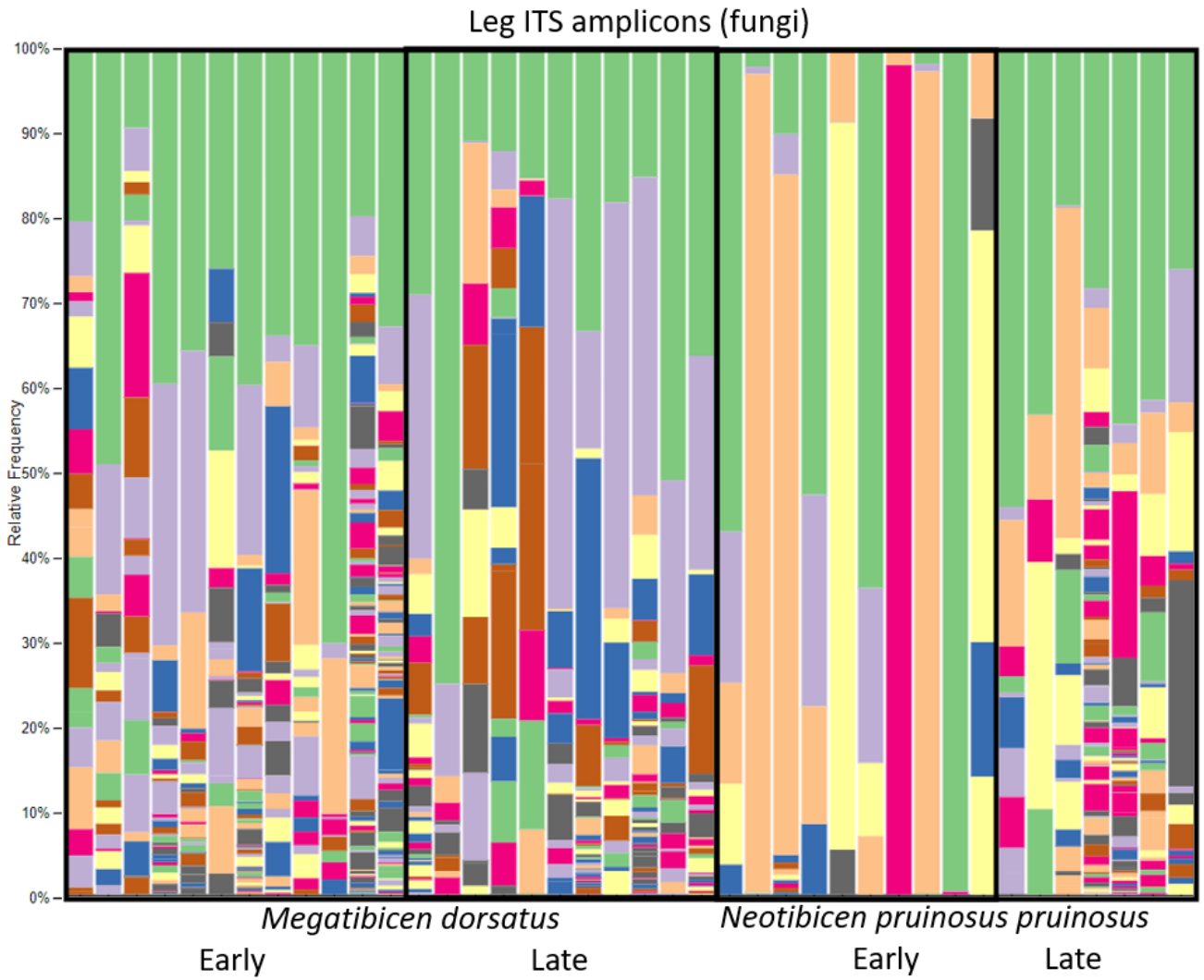


Figure 3.8. QIIME 2 View generated proportion of reads of molt, pruinosus, and soil ITS reads assigned fungal taxonomy using the UNITE classifier. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.4.B.

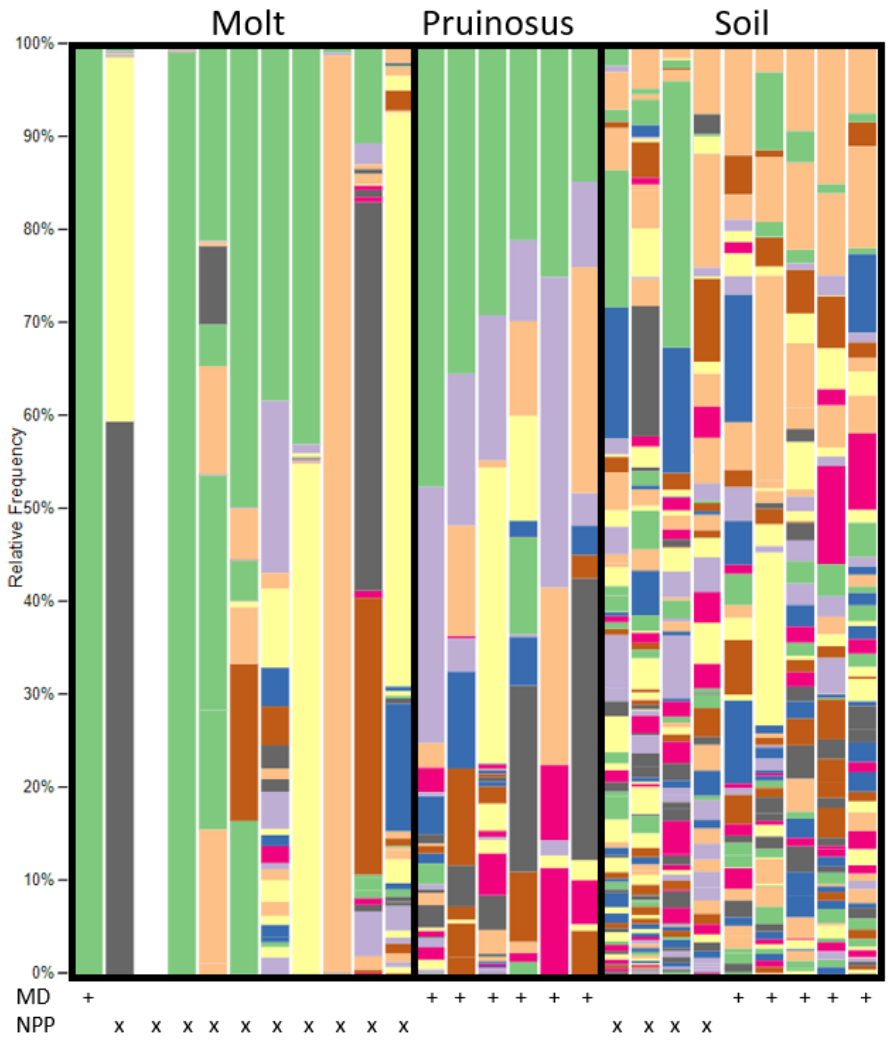


Figure 3.9.A. QIIME2 View generated forewing 16S V3-V4. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.9.B.

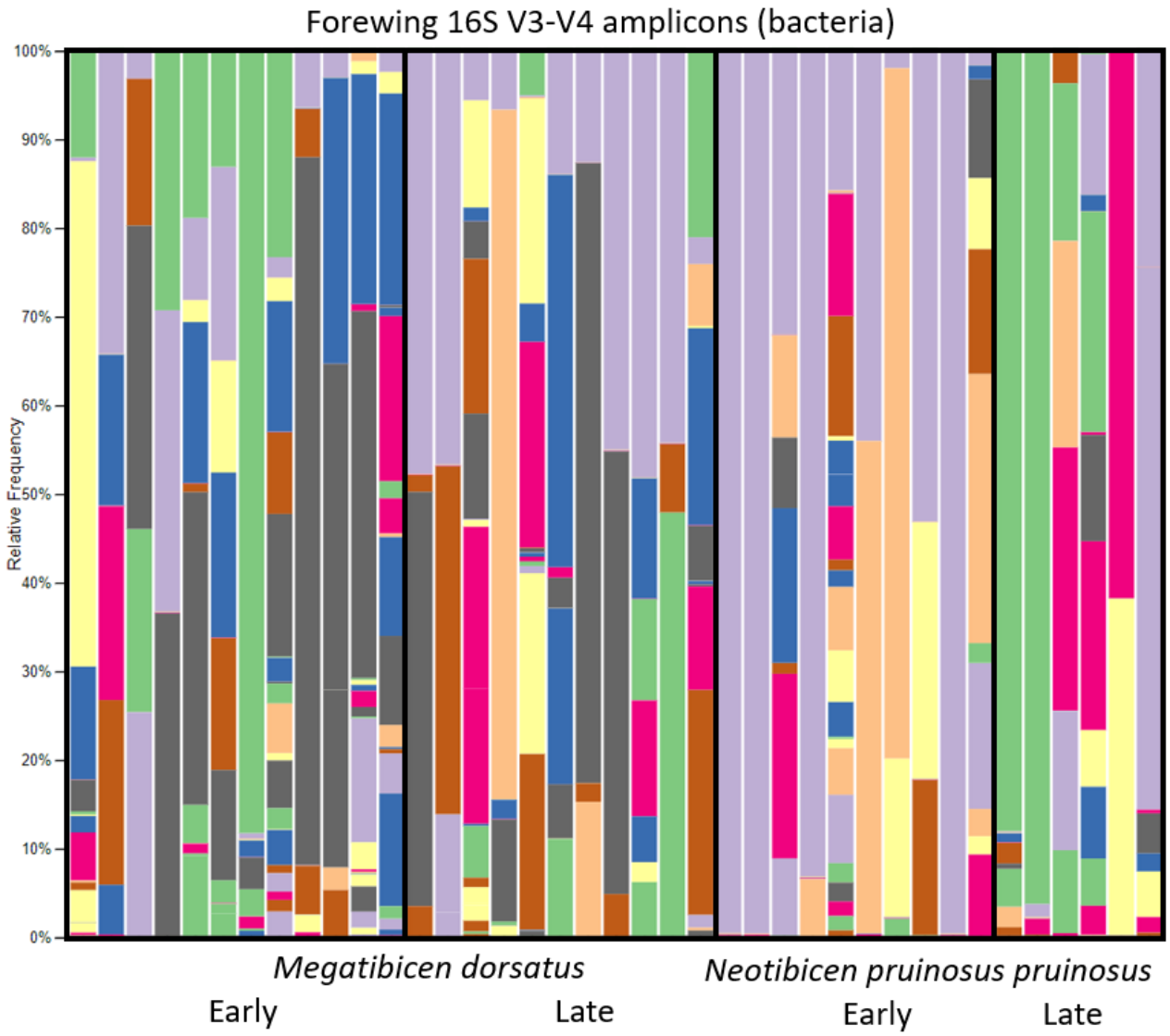


Figure 3.9.B. Legend of top 20 ASV identifications from 16S V3-V4 amplified sequences. Full data set of assigned taxonomy assignment and read counts can be found as external file.

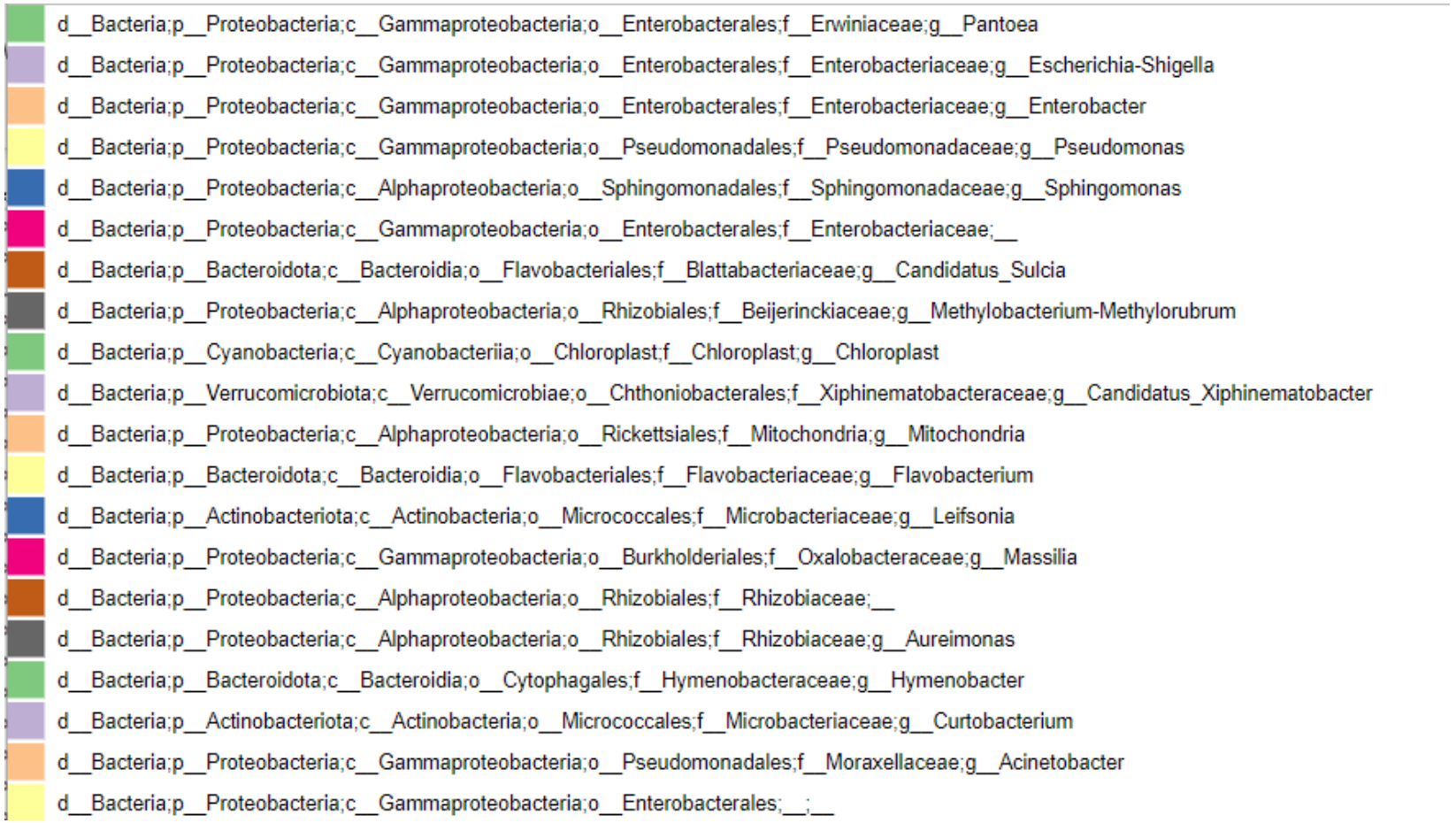


Figure 3.10. QIIME2 View generated proportion of head 16S V3-V4 ASVs. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.9.B.

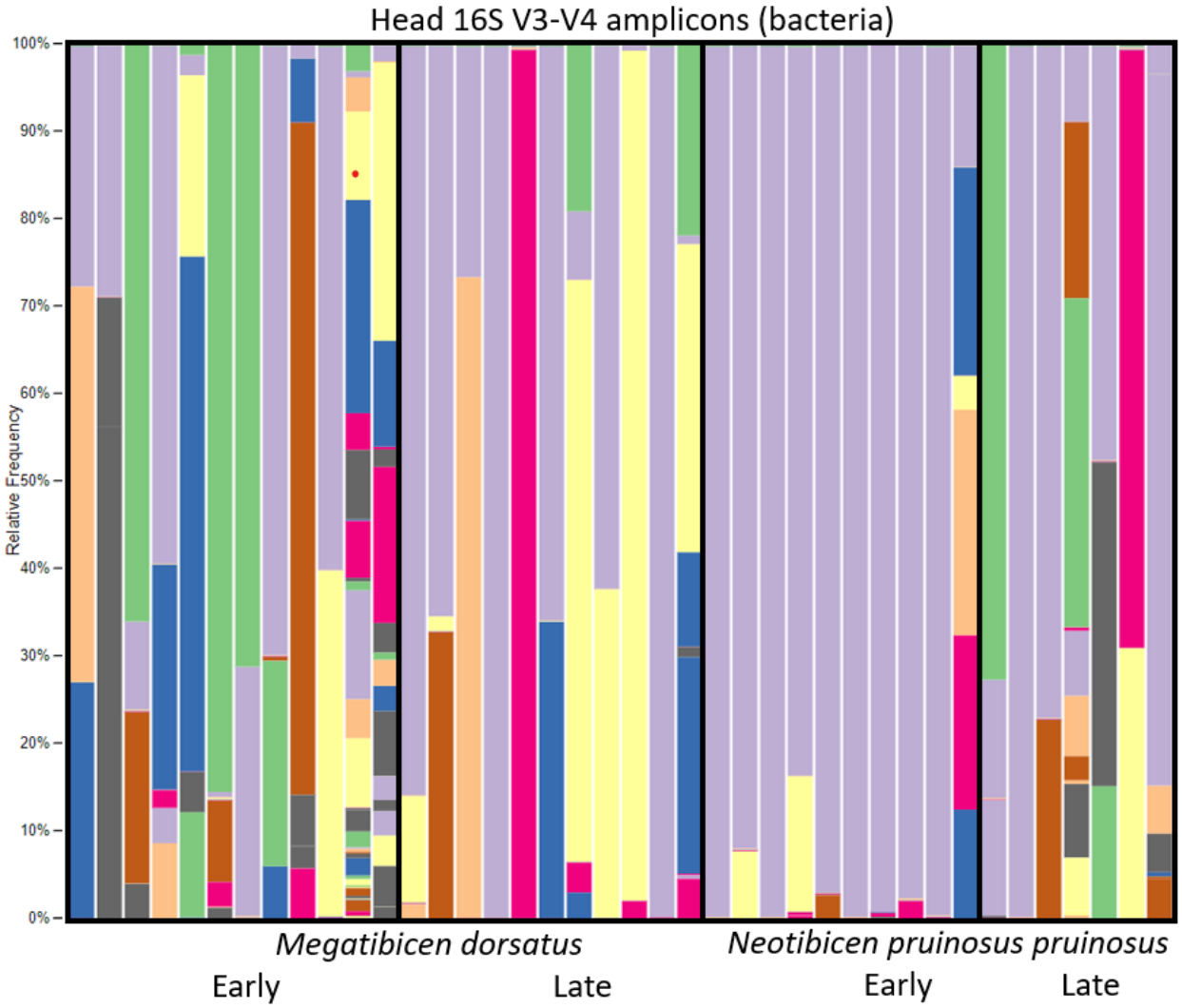


Figure 3.11. QIIME2 View generated proportion of hindwing 16S V3-V4 ASVs. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.9.B.

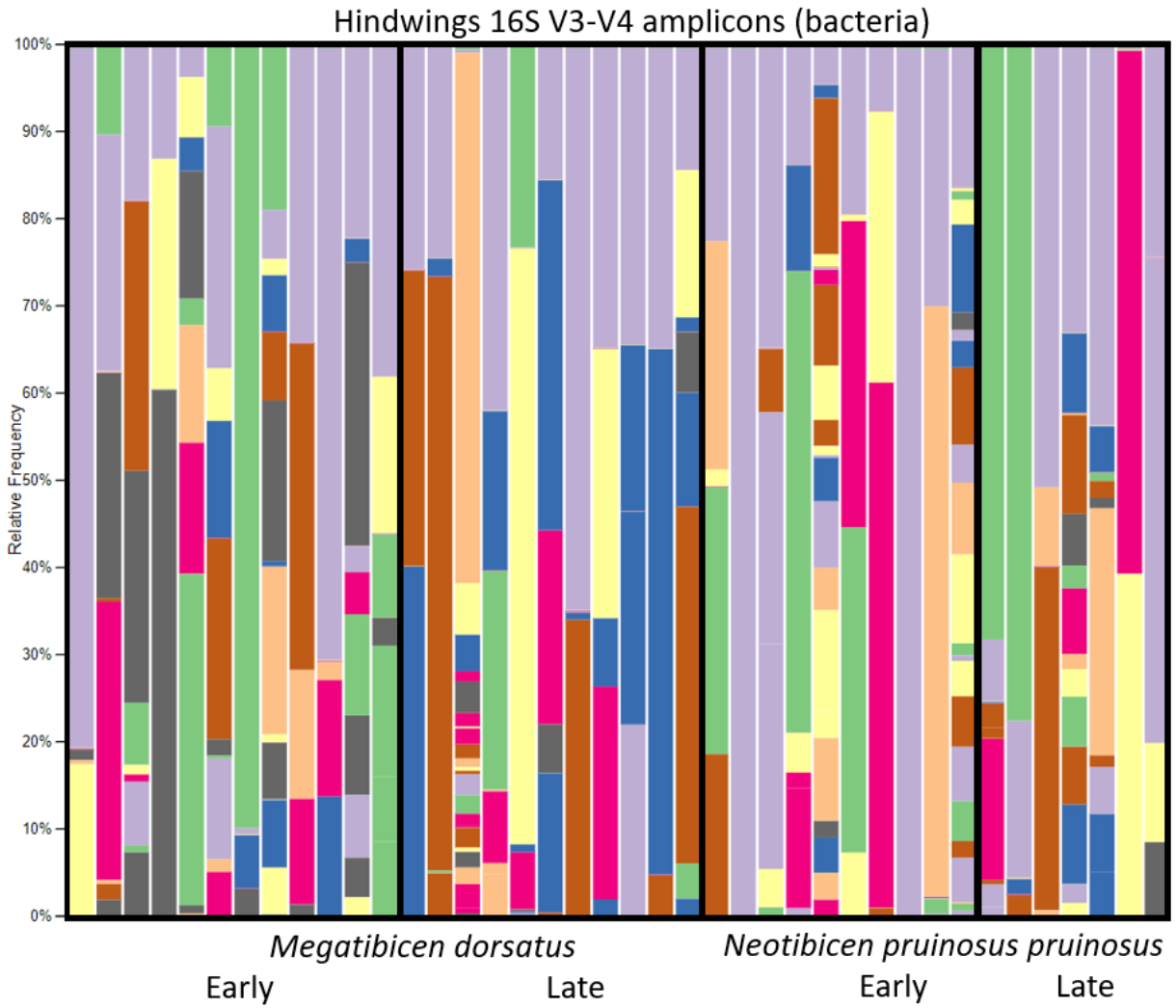


Figure 3.12. QIIME2 View generated proportion of leg 16S V3-V4 ASVs. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.9.B.

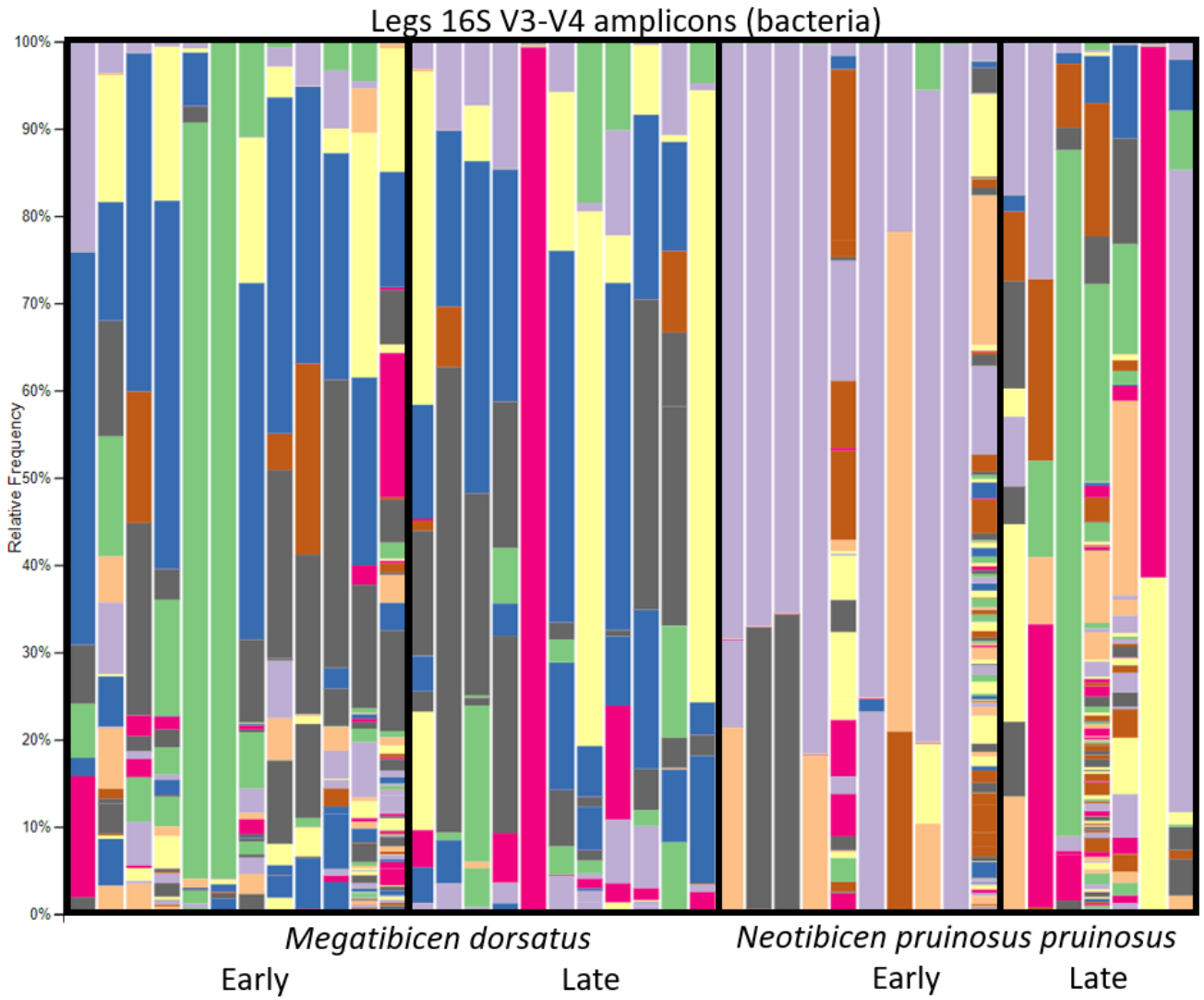


Figure 3.13. QIIME2 View generated proportion of molt, pruinosus (fuzz), and soil 16S V3-V4 ASVs. Full data set of assigned taxonomy assignment and read counts can be found as external supplemental file. Legend of top 20 assignments (associated from top down) can be seen in Figure 3.9.B. White assignments on molt samples indicate no reads were present, likely an artifact of PCR inhibitors (e.g., tannins).

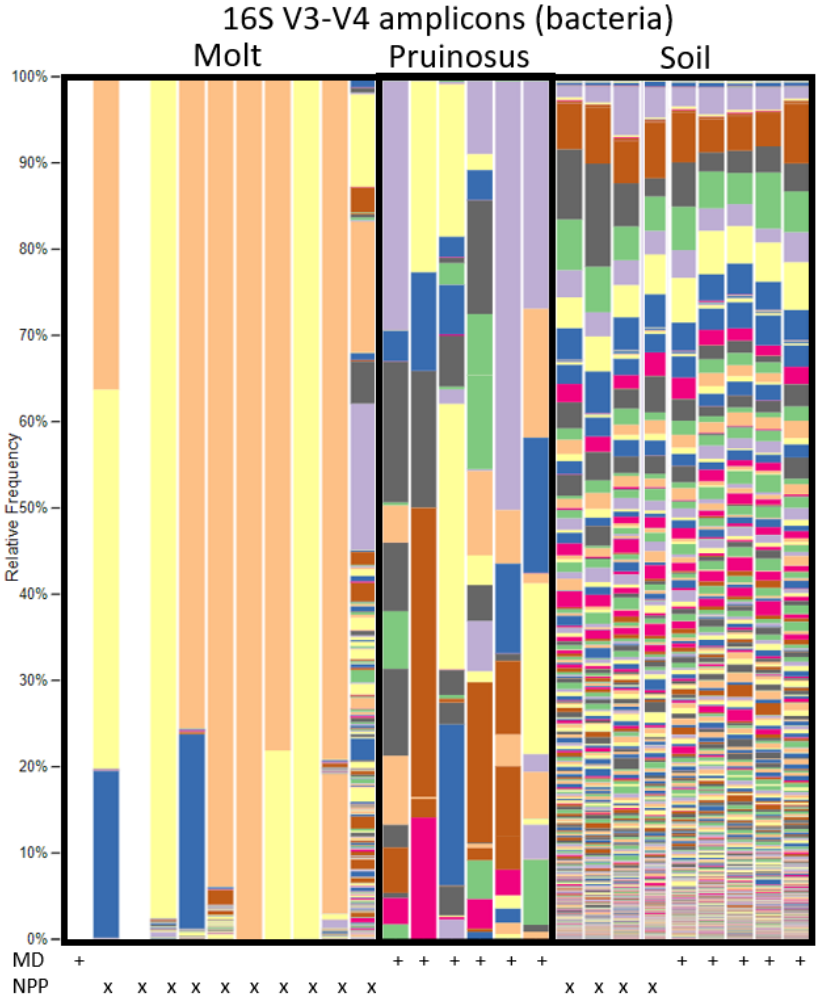


Figure 3.14. Geneious Prime ® 2022.1.1 generated alignment of bacterial *Candidatus* Hodgkinia ASVs found in our *Cicadettana calliope calliope* samples.

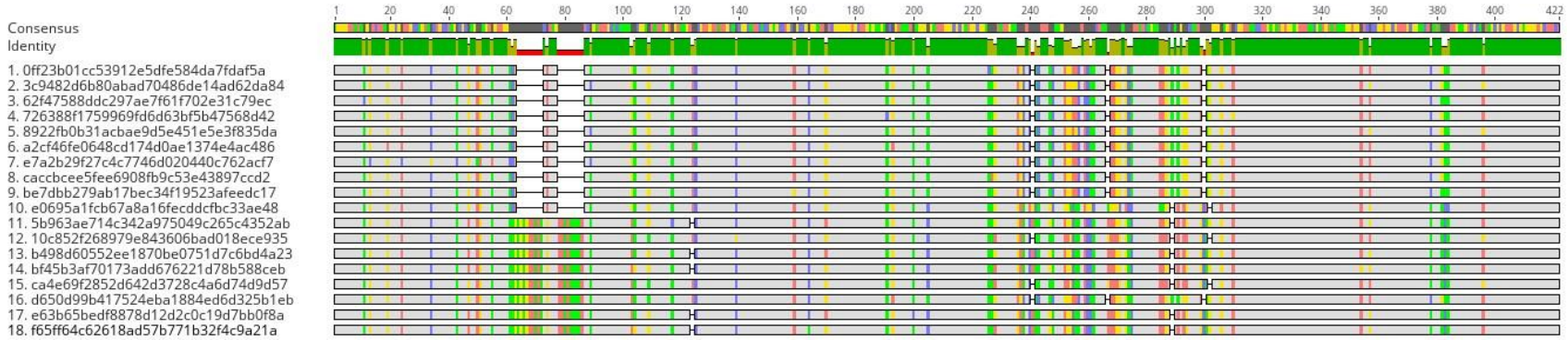
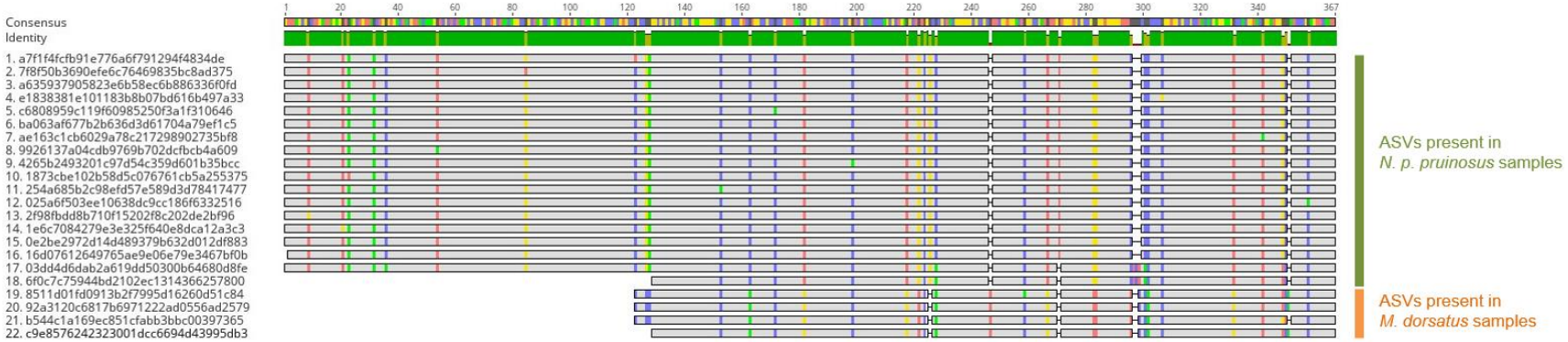


Figure 3.15. Geneious Prime ® 2022.1.1 generated alignment of fungal ITS3-4 ASVs that align closely with Yeast-Like Symbionts (YLS) and *Ophiocordyceps* spp.



Acknowledgments

This work was largely through collaboration and funding from support provided by the Army Basic Research Program with the United States Army Engineer Research and Development Center, Construction Engineering Research Laboratory (CERL) under Grant ARMY CESU W9132T-16-2-0011. Thank you to Mark Davis (Illinois Natural History Survey) for guidance and access to the Collaborative Conservation Genetics Laboratory. Thanks to the Roy J. Carver Biotechnology Center and Mark Band for guidance in library submission and sequencing. Thanks also go to the people that run the Biocluster at the Carl R. Woese Institute for Genomic Biology. Guidance on analyses was provided in part from workshops provided by the High-Performance Biological Computing group (HPCBio) at the Institute for Genomic Biology.

References

Abarenkov, K., Zirk, A., Piirmann, T., Pöhönen, R., Ivanov, F., Nilsson, R.H., and U. Kõljalg. 2020. UNITE QIIME release for Fungi [Data set]. UNITE Community. https://doi.org/10.15156/BIO/786385

Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Apprill, A., McNally, S., Parsons, R., and L. Weber. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology 75(2): 129–137.

Baral, H.O., Weber, E., Marson, G. and Quijada, L., 2018. A new connection between wood saprobism and beetle endosymbiosis: the rarely reported saprobic discomycete *Tromeropsis* is congeneric with the symbiotic yeast *Symbiotaphrina* (Symbiotaphrinales, Xylonomycetes) and two asexual morphs misplaced in *Hyphozyma*. Mycological Progress 17(1): 215–254.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., and A. Amend. 2013. ITSx: Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for use in environmental sequencing. Methods in Ecology and Evolution 4: 914–919.

Bennett, G.M., McCutcheon, J.P., MacDonald, B.R., Romanovicz, D., and N.A. Moran. 2014. Differential genome evolution between companion symbionts in an insect-bacterial symbiosis. mBio 5(5): e01697–14. https://doi.org/10.1128/mBio.01697-14

Bokulich, N.A., Kaehler, B.D., Rideout, J.R. Dillon, M., Bolyen, E., Knight, R., Huttley, G.A., and J.G. Caporaso. 2018. Optimizing taxonomic classification of marker-gene amplicon

sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6(90): 1–17. https://doi.org/10.1186/s40168-018-0470-z

Bokulich, N.A., Robeson, M., and M.R. Dillon. 2021. bokulich-lab/RESCRIPt. Zenodo. http://doi.org/10.5281/zenodo.3891931 [Accessed: 11/8/2021]

Bolger, A.M., Lohse, M., and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., and J.G. Caporaso. 2019.

Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37(8): 852–857. https://doi.org/10.1038/s41587-019-0209-9

Brey, P.T., Lee, W.J., Yamakawa, M., Koizumi, Y., Perrot, S., Francois, M., and M. Ashida. 1993. Role of the integument in insect immunity: epicuticular abrasion and induction of cecropin synthesis in cuticular epithelial cells. Proceedings of the National Academy of Sciences 90(13): 6275–6279.

Brumfield, K.D., Raupp, M.J., Haji, D., Simon, C., Graf, J., Cooley, J.R., Janton, S.T., Meister, R.C., Huq, A., Colwell, R.R., and N.A. Hasan. 2022. Gut microbiome insights from 16S rRNA analysis of 17-year periodical cicadas (Hemiptera: *Magicicada* spp.) Broods II, VI, and X. Scientific Reports 12: 16967. https://doi.org/10.1038/s41598-022-20527-7

Buchner, P. 1965. Endosymbiosis of animals with plant microorganisms. Rev. Eng. ed. [Translated by Bertha Mueller, with the collaboration of Francis H. Fockier]. New York: Interscience Publishers.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and S.P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13(7): 581–583.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., and R. Knight. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108 (Supplement 1): 4516–4522.

Ciferri, R., Machado, A.A., and A.F. Vital. 1957. A new species of the genus *Massospora* with an *Allomyces* species. Atti dell'Istituto Botanico della Universitá e Laboratorio Crittogamico di Pavia 5(14): 15–22.

Cooley, J.R., Marshall, D.C., and K.B.R. Hill. 2018. A specialized fungal parasite (*Massospora cicadina*) hijacks the sexual signals of periodical cicadas (Hemiptera: Cicadidae: *Magicicada*). Scientific Reports 8(1): 1–7.

Cornara, D., Marra, M., Tedone, B., Cavalieri, V., Porcelli, F., Fereres, A., Purcell, A., and M. Saponari. 2020. No evidence for cicadas' implication in *Xylella fastidiosa* epidemiology. Entomologia Generalis 20: 125–132.

De Geer, C. 1773. Des cigales exotiques de la quatrieme famille. Cinquième mémoire. Des cigales. Mémoires pour servir à l'histoire des Insects 3: 212–222.

Degnan, P.H., and H. Ochman. 2012. Illumina-based analysis of microbial community diversity. The ISME Journal 6(1): 183–194.

Devanga Ragupathi, N.K., Muthuirulandi Sethuvel, D.P., Inbanathan, F.Y., and B. Veeraraghavan. 2018. Accurate differentiation of *Escherichia coli* and *Shigella* serogroups: challenges and strategies. New Microbes and New Infections 21: 58–62. https://doi.org/10.1016/j.nmni.2017.09.003

Fan, Y., Fang, W., Guo, S., Pei, X., Zhang, Y., Xiao, Y., Li, D., Jin, K., Bidochka, M.J., and Y. Pei. 2007. Increased insect virulence in *Beauveria bassiana* strains overexpressing an engineered chitinase. Applied and Environmental Microbiology 73(1): 295–302. https://doi.org/10.1128/AEM.01974-06

Fukatsu, T., and H. Ishikawa. 1996. Phylogenetic position of yeast-like symbiont of *Hamiltonaphis styraci* (Homoptera, Aphididae) based on 18S rDNA sequence. Insect Biochemistry and Molecular Biology 26(4): 383–388.

Gams, W., and J.A. von Arx. 1980. Validation of *Symbiotaphrina* (imperfect yeasts). PERSOONIA: Molecular Phylogeny and Evolution of Fungi 10(4): 542–543.

Gauthier, J., and N. Derome. 2021. Evenness-richness scatter plots: a visual and insightful representation of Shannon Entropy measurements for ecological community analysis. Msphere 6(2): e01019–20. https://doi.org/10.1128/mSphere.01019-20

Haji, D., Vailionis, J., Stukel, M., Gordon, E., Lemmon, E.M., Lemmon, A.R., and C. Simon. 2022. Correlates of host-associated bacterial diversity in New Zealand cicadas and hybrids. Scientific Reports (Under Review). https://doi.org/10.21203/rs.3.rs-1875558/v1

Hamady, M., and R. Knight. 2009. Microbial community profiling for human microbiome projects: tools, techniques, and challenges. Genome Research 19(7): 1141–1152.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., and R. Knight. 2008. Error-correcting barcoded primers allow hundreds of samples to be pyrosequenced in multiplex. Nature Methods 5(3): 235–237.

Han, Z.W., Wang, Z., Feng, X.M., Li, B., Mu, Z.Z., Zhang, J.Q., Niu, S.C. and L.Q. Ren. 2016. Antireflective surface inspired from biology: a review. Biosurface and Biotribology 2(4): 137–150.

Hasan, J., Webb, H.K., Truong, V.K., Pogodin, S., Baulin, V.A., Watson, G.S., Watson, J.A., Crawford, R.J., and E.P. Ivanova. 2013. Selective bactericidal activity of nanopatterned superhydrophobic cicada *Psaltoda claripennis* wing surfaces. Applied Microbiology and Biotechnology 97(20): 9257–9262.

Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., and A.F. Andersson. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. The ISME Journal 5(10): 1571–1579.

Hu, Y., Dietrich, C.H., Skinner, R.K., and Y. Zhang. 2022. Phylogeny of Membracoidea (Hemiptera: Auchenorrhyncha) based on transcriptome data. Systematic Entomology: 1–14. https://doi.org/10.1111/syen.12563

Huang, Y.F., Jen, Y.J., Chen, L.C., Chen, K.H., and S. Chattopadhyay. 2015. Design for approaching cicada-wing reflectance in low-and high-index biomimetic nanostructures. ACS Nano 9(1): 301–311.

Ishak, H.D., Plowes, R., Sen, R. Kellner, K., Meyer, E., Estrada, D.A., Dowd, S.E., and U.G. Mueller. 2011. Bacterial diversity in *Solenopsis invicta* and *Solenopsis geminata* ant colonies characterized by 16S amplicon 454 pyrosequencing. Microbial Ecology 61: 821–831.

Ivanova, E.P., Hasan, J., Webb, H.K., Truong, V.K., Watson, G.S., Watson, J.A., Baulin, V.A., Pogodin, S., Wang, J.Y., Tobin, M.J., Löbbe, C., and R.J. Crawford. 2012. Natural bactericidal surfaces: mechanical rupture of *Pseudomonas aeruginosa* cells by cicada wings. Small 8: 2489–2494

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and T.L. Madden. 2008. NCBI BLAST: a better web interface. Nucleic Acids Research 36(Web Server issue): W5–W9. https://doi.org/10.1093/nar/gkn201

Kaehler, B.D., Bokulich, N.A., McDonald, D. Knight, R., Caporaso, J.G., and G.A. Huttley. 2019. Species abundance information improves sequence taxonomy classification accuracy. Nature Communications 10: 4643. https://doi.org/10.1038/s41467-019-12669-6

Kelleher, S.M., Habimana, O., Lawler, J., O'Reilly, B., Daniels, S., Casey, E. and A. Cowley. 2016. Cicada wing surface topography: an investigation into the bactericidal properties of nanostructural features. ACS Applied Materials & Interfaces 8(24): 14966–14974.

Kepler, R.M., Sung, G.H., Ban, S., Nakagiri, A., Chen, M.J., Huang, B., Li, Z.Z., and J.W. Spatafora. 2012. New teleomorph combinations in the entomopathogenic genus *Metacordyceps*. Mycologia 104(1): 182–197.

Kim, H.S., Adhikari, M., Yadav, D.R., Kim, S.W., Um, Y.H., Lee, H.B., and Y.S. Lee. 2016. First report of *Metacordyceps chlamydosporia* (*Cordyceps chlamydosporia*) isolated from soil in Korea. The Korean Journal of Mycology 44(1): 48–50.

Kobayasi, Y., and D. Shimizu. 1963. Monographic studies of *Cordyceps* 2. Group parasitic on Cicadidae. Bulletin of the National Science Museum 6(3): 286–314.

Lane, D.J. 1991. 16S/23S rRNA sequencing, p 115–174. *In* Stackebrandt E, Goodfellow M (ed), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, Chichester, England.

Laurence, M., Hatzis, C., and D.E. Brash. 2014. Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. PloS ONE 9(5): e97876. https://doi.org/10.1371/journal.pone.0097876

Lee, J.B., and D.-C. Oh. 1998. Higher fungi of Cheju-Do (1)-Unrecorded mushrooms. The Korean Journal of Mycology 26(4): 538–550.

Li, C.-R., Fan, M.-Z., Huang, B., Wang, S.-B., and Z-.Z. Li. 2002. The genus *Cordyceps* and its allies from Anhui I. Mycosystema 21(2): 167–171.

Li, C., Huang, B., Fan, M., Lin, Y., and Z. Li. 2010. *Metacordyceps guniujiangensis* and its *Metarhizium* anamorph: a new pathogen on cicada nymphs. Mycotaxon 111(1): 221–231.

Lüthy, P. 1974. *Corynebacterium okanaganae*, an entomopathogenic species of the Corynebacteriaceae. Canadian Journal of Microbiology 20(5): 791–794.

Lüthy, P., and R.S. Soper. 1969. Chalky disease, a bacterial septicemia of the cicada *Okanagana rimosa* Say. Journal of Invertebrate Pathology 14(2): 158–164.

Macias, A.M., Geiser, D.M., Stajich, J.E., Łukasik, P., Veloso, C., Bublitz, D.C., Berger, M.C., Boyce, G.R., Hodge, K., and M.T. Kasson. 2020. Evolutionary relationships among *Massospora* spp. (Entomophthorales), obligate pathogens of cicadas. Mycologia 112(6): 1060–1074. DOI: 10.1080/00275514.2020.1742033

Mallott, E.K., Malhi, R.S., and K.R. Amato. 2019. Assessing the comparability of different DNA extraction and amplification methods in gut microbial community profiling. Access Microbiology 1(7): 1–15.

Matsuura, Y., Moriyama, M., Łukasik, P., Vanderpool, D., Tanahashi, M., Meng, X.Y., McCutcheon, J.P., and T. Fukatsu. 2018. Recurrent symbiont recruitment from fungal parasites in cicadas. Proceedings of the National Academy of Sciences 115(26): E5970–E5979.

McCutcheon, J.P., McDonald, B.R., and N.A. Moran. 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. PLoS Genetics 5(7): e1000565. https://doi.org/10.1371/journal.pgen.1000565

McDonald, D., Clemente, J.C., Kuczynski, J., Rideout, J.R., Stombaugh, J., Wendel, D., Wilke, A., Huse, S., Hufnagle, J., Meyer, F., and R. Knight. 2012. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. Gigascience 1: 1–7. DOI: 10.1186/2047-217X-1-7

McKinney, W. 2010. Data structures for statistical computing in python. *In* Proceedings of the 9th Python in Science Conference 445(1): 51–56.

Moran, N.A., Tran, P., and N.M. Gerardo. 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. Applied and Environmental Microbiology 71(12): 8802–8810.

Mukherjee, K., and A. Vilcinskas. 2018. The entomopathogenic fungus *Metarhizium robertsii* communicates with the insect host *Galleria mellonella* during infection. Virulence 9(1): 402–413. https://doi.org/10.1080/21505594.2017.1405190

Murray, A.E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hugenholtz, P., Kämpfer, P., Konstantinidis, K.T., Lane, C.E., Papke, R.T., Parks, D.H., Rossello-Mora, R., Stott, M.B., Sutcliffe, I.C., Thrash, J.C., Venter, S.N., Whitman, W.B., Acinas, S.G., Amann, R.I., Anantharaman, K., Armengaud, J., Baker, B.J., Barco, R.A., Bode, H.B., Boyd, E.S., Brady, C.L., Carini, P., Chain, P.S.G., Colman, D.R., DeAngelis, K.M., de los Rios, M.A., Estrada-de los Santos, P., Dunlap, C.A., Eisen, J.A., Emerson, D., Ettema, T.J.G., Eveillard, D., Girguis, P.R., Hentschel, U., Hollibaugh, J.T., Hug, L.A., Inskeep, W.P., Ivanova, E.P., Klenk, H.-P., Li, W.-J., Lloyd, K.G., Löffler, F.E., Makhalanyane, T.P., Moser, D.P., Nunoura, T., Palmer, M., Parro, V., Pedrós-Alió, C., Probst, A.J., Smits, T.H.M., Steen, A.D., Steenkamp, E.T., Spang, A., Stewart, F.J., Tiedje, T.M., Vandamme, P., Wagner, M., Wang, F.-P., Yarza, P., Hedlund, B.P., and A.-L. Reysenbach. 2020. Roadmap for naming uncultivated Archaea and Bacteria. Nature Microbiology 5(8): 987–994.

Muyzer, G., De Waal, E.C., and A. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology 59(3): 695–700.

Nilsson, R.H., Larsson, K.H., Taylor, A., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., Saar, I., Kõljalg, U., and K. Abarenkov. 2019. The UNITE database for molecular identification of fungi: handling dark

taxa and parallel taxonomic classifications. Nucleic Acids Research 47(D1): D259–D264. https://doi.org/10.1093/nar/gky1022

Oh, J., Dana, C.E., Hong, S., Román, J.K., Jo, K.D., Hong, J.W., Nguyen, J., Cropek, D.M., Alleyne, M., and N. Miljkovic. 2017. Exploring the role of habitat on the wettability of cicada wings. ACS Applied Materials & Interfaces 9(32): 27173–27184.

Oh, J., Hoffman, J.B., Hong, S., Jo, K.D., Román-Kustas, J., Reed, J.H., Dana, C.E., Cropek, D.M., Alleyne, M. and N. Miljkovic. 2020. Dissolvable Template Nanoimprint Lithography: A Facile and Versatile Nanoscale Replication Technique. Nano Letters 20(10): 6989-6997.

Oliver, K.M., Campos, J., Moran, N.A., and M.S. Hunter. 2008. Population dynamics of defensive symbionts in aphids. Proceedings of the Royal Society B: Biological Sciences 275(1632): 293–299.

Oliver, K.M., Russell, J.A., Moran, N.A., and M.S. Hunter. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. Proceedings of the National Academy of Sciences 100(4): 1803–1807.

O'Neal, L., Schumann, P., Falsen, E., and P.A. Lawson. 2017. *Luethyella okanaganae* gen. nov., sp. nov., a novel genus and species of the family Microbacteriaceae isolated from the insect *Okanagana rimosa*. Current Microbiology 74(4): 419–424.

Oren, A. 2017. A plea for linguistic accuracy–also for Candidatus taxa. International Journal of Systematic and Evolutionary Microbiology 67(4): 1085–1094.

Oren, A. 2021. Nomenclature of prokaryotic '*Candidatus*' taxa: establishing order in the current chaos. New Microbes and New Infections 44(C): 100932. https://doi.org/10.1016/j.nmni.2021.100932

Oren, A., Garrity, G.M., Parker, C.T., Chuvochina, M., and M.E. Trujillo. 2020. Lists of names of prokaryotic *Candidatus* taxa. International Journal of Systematic and Evolutionary Microbiology 70(7): 3956–4042. https://doi.org/10.1099/ijsem.0.003789

Parada, A.E., Needham, D.M., and J.A. Fuhrman. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environmental Microbiology 18(5): 1403–1414.

Peck, C.H. 1879. Report of the Botanist. Thirty-First Annual Report on the New York State Museum of Natural History by the Regents of the University of the State of New York 42: 19–60.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and J. Vanderplas. 2011. Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research 12: 2825–2830.

Pereira, J., Noronha, E., Miller, R., and O. Franco. 2007. Novel insights in the use of hydrolytic enzymes secreted by fungi with biotechnological potential. Letters in Applied Microbiology 44: 573–581. https://doi.org/10.1111/j.1472-765X.2007.02151.x

Pertot, I., Fiamingo, F., Amsalem, L., Maymon, M., Freeman, S., Gobbin, D., and Y. Elad. 2007. Sensitivity of two *Podosphaera aphanis* populations to disease control agents. Journal of Plant Pathology 89(1): 85–96. http://www.jstor.org/stable/41998360

Phukhamsakda, C., Nilsson, R.H., Bhunjun, C.S., de Farias, A.R.G., Sun, Y.-R., Wijesinghe, S.N., Raza, M., Bao, D.-F., Lu, L., Tibpromma, S., Dong, W., Tennakoon, D.S., Tian, X.-G., Xiong, Y.-R, Karunarathna, S.C., Cai, L., Luo, Z.-L, Wang, Y., Manawasinghe, I.S., Camporesi, E., Kirk, P.M., Promputtha, I., Kuo, C.-H., Su, H.-Y., Doilom, M., Li, Y., Fu, Y.-P., and K.D. Hyde. 2022. The numbers of fungi: contributions from traditional taxonomic

studies and challenges of metabarcoding. Fungal Diversity 114: 327–386. https://doi.org/10.1007/s13225-022-00502-3

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and F.O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research 35(21): 7188–7196.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and F.O. Glöckner. 2013 The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41(D1): D590–D596.

Quince, C., Lanzen, A., Davenport, R.J., and P.J. Turnbaugh. 2011. Removing noise from pyrosequenced amplicons. BMC Bioinformatics 12(1): 38.

Rideout, J.R., Chase, J.H., Bolyen, E., Ackermann, G., González, A., Knight, R., and J.G. Caporaso. 2016. Keemei: cloud-based validation of tabular bioinformatics file formats in Google Sheets. GigaScience 5(1): 27. http://dx.doi.org/10.1186/s13742-016-0133-6

Robeson, M.S., O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., and N.A. Bokulich. 2021. RESCRIPt: Reproducible sequence taxonomy reference database management. PLoS Computational Biology 17(11): p.e1009581. https://doi.org/10.1371/journal.pcbi.1009581

Rognes, T., Flouri, T., Nichols, B., Quince, C., and F. Mahé. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4: p.e2584.

Román-Kustas, J., Hoffman, J.B., Reed, J.H., Gonsalves, A.E., Oh, J., Li, L., Hong, S., Jo, K.D., Dana, C.E., Miljkovic, N., and D.M. Cropek. 2020. Molecular and topographical organization: influence on cicada wing wettability and bactericidal properties. Advanced Materials Interfaces 7(10): 2000112.

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and A.W. Walker. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology 12(1): 1–12.

Sanborn, A.F., and Phillips, P.K. 2013. Biogeography of the cicadas (Hemiptera: Cicadidae) of North America, north of Mexico. Diversity 5(2): 166-239.

Sasaki, T., and H. Ishikawa. 1995. Production of essential amino acids from glutamate by mycetocyte symbionts of the pea aphid, *Acyrthosiphon pisum*. Journal of Insect Physiology 41(1): 41–46.

Say, T. 1825. Descriptions of new hemipterous insects, collected in the expedition to the Rocky Mountains, performed by the order of Mr. Calhoun, Secretary of War, under command of Major Long. Journal of the Natural Academy of Sciences of Philadelphia 4(2): 307–345.

Schneider, R.W. 1973. Epidemiology, yield-loss prediction, and control of *Cercospora* leaf spot of cowpea (*Vigna unguiculata*). Doctoral dissertation, University of Illinois at Urbana-Champaign. http://hdl.handle.net/2142/64219

Sen, R., Ishak, H.D., Estrada, D., Dowd, S.E., Hong, E., and U.G. Mueller. 2009. Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis* bacteria in nests of fungus-growing ants. Proceedings of the National Academy of Sciences 106(42): 17805–17810.

Shannon, C.E. 1948. A mathematical theory of communication. The Bell System Technical Journal 27(3): 379–423.

Skinner, R.K., Dietrich, C.H., Walden, K.K.O., Gordon, E., Sweet, A.D., Podsiadlowski, L., Petersen, M., Simon, C., Takiya, D.M., and K.P. Johnson. 2020. Phylogenomics of Auchenorrhyncha (Insecta: Hemiptera) using transcriptomes: examining controversial

relationships via degeneracy coding and interrogation of gene conflict. Systematic Entomology 45: 85–113. https://doi.org/10.1111/syen.12381

Smith, E.A., Anderson, K.E., Corby-Harris, V., McFrederick, Q.S., Parish, A.J., Rice, D.W., and I.L. Newton. 2021. Reclassification of seven honey bee symbiont strains as *Bombella api*s. International Journal of Systematic and Evolutionary Microbiology 71(9): 004950.

Soper, R.S. 1963. *Massospora levispora*, a new species of fungus pathogenic to the cicada, *Okanagana rimosa*. Canadian Journal of Botany 41(6): 875–878.

Soper, R.S. 1974. The genus *Massospora*, entomopathogenic for cicadas, part I, taxonomy of the genus. Mycotaxon 1: 13–40.

Soper, R.S. 1981. New cicada pathogens: *Massospora cicadettae* from Australia and *Massospora pahariae* from Afghanistan. Mycotaxon 13(1): 50–58.

Soper, R.S., Delyzer, A.J., and L.F. Smith. 1976. The genus *Massospora* entomopathogenic for cicadas. Part. II. Biology of *Massospora levispora* and its host *Okanagana rimosa*, with notes on *Massospora cicadina* on the periodical cicadas. Annals of the Entomological Society of America 69(1): 89–95.

Sørensen, K.B., and A. Teske. 2006. Stratified communities of active archaea in deep marine subsurface sediments. Applied and Environmental Microbiology 72(7): 4596–4603.

Su, Q., Xie, W., Wang, S., Wu, Q., Liu, B., Fang, Y., Xu, B., and Y. Zhang. 2014. The endosymbiont *Hamiltonella* increases the growth rate of its host *Bemisia tabaci* during periods of nutritional stress. PLoS ONE 9(2): e89002. doi:10.1371/journal.pone.0089002

Suh, S.O., Noda, H., and M. Blackwell. 2001. Insect symbiosis: derivation of yeast-like endosymbionts within an entomopathogenic filamentous lineage. Molecular Biology and Evolution 18(6): 995–1000.

Sung, G.H., Shrestha, B., Han, S.K., and J.-M. Sung. 2011. Growth and cultural characteristics of *Ophiocordyceps longissima* collected in Korea. Mycobiology 39(2): 85–91. DOI: 10.4489/MYCO.2011.39.2.085.

Takai, K., and K. Horikoshi. 2000. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. Applied and Environmental Microbiology 66(11): 5066–5072.

Takiya, D.M., Tran, P.L., Dietrich, C.H., and N.A. Moran. 2006. Co-cladogenesis spanning three phyla: leafhoppers (Insecta: Hemiptera: Cicadellidae) and their dual bacterial symbionts. Molecular Ecology 15(13): 4175–4191. https://doi.org/10.1111/j.1365-294X.2006.03071.x

Turner, S., Pryer, K.M., Miao, V.P.W, and J.D. Palmer. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence Analysis. Journal of Eukaryotic Microbiology 46(4): 327–338.

Van Leuven, J.T., Meister, R.C., Simon, C., and J.P. McCutcheon. 2014. Sympatric speciation in a bacterial endosymbiont results in two genomes with the functionality of one. Cell 158(6): 1270–1280.

Walker, F. 1850. List of the specimens of homopterous insects in the collection of the British Museum. Order of Trustees. London, England. Part 1: 1–260.

Walterson, A.M., and J. Stavrinides. 2015. *Pantoea*: insights into a highly versatile and diverse genus within the Enterobacteriaceae. FEMS Microbiology Reviews 39(6): 968–984.

Weyrich, L.S., Farrer, A.G., Eisenhofer, R., Arriola, L.A., Young, J., Selway, C.A., Handsley-Davis, M., Adler, C.J., Breen, J., and A. Cooper. 2019. Laboratory contamination over time during low-biomass sample analysis. Molecular Ecology Resources 19: 982–996. https://doi.org/10.1111/1755-0998.13011

White, T.J., Bruns, T.D., Lee, S.B., and J.W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* PCR Protocols: A Guide to Methods and Applications, Academic Press, Elsevier, 315–322.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., and F.O. Glöckner. 2014. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Research 42(DI): D643–D648. doi: 10.1093/nar/gkt1209

Zare, R., Gams, W., and H.C. Evans. 2001. A revision of *Verticillium* section *Prostrata*. V. The genus *Pochonia*, with notes on *Rotiferophthora*. Nova Hedwigia 73(1–2): 51–86.

Zhang, Y., and S. Qiu. 2015. Examining phylogenetic relationships of *Erwinia* and *Pantoea* species using whole genome sequence data. Antonie Van Leeuwenhoek 108(5): 1037–1046.

Zheng, Z., Wang, D., He, H., and C. Wei. 2017. Bacterial diversity of bacteriomes and organs of reproductive, digestive and excretory systems in two cicada species (Hemiptera: Cicadidae). PLoS ONE 12(4): e0175903. https://doi.org/10.1371/journal.pone.0175903

Zhou, W., Nan, X., Zheng, Z., Wei, C., and H. He. 2015. Analysis of inter-individual bacterial variation in gut of cicada *Meimuna mongolica* (Hemiptera: Cicadidae). Journal of Insect Science 15(1): 131. https://doi.org/10.1093/jisesa/iev113

APPENDIX A: CHAPTER 1 SUPPLEMENTARY MATERIALS

Table A.1. Specimens examined from museums including and in addition to those listed throughout the study.

Collection	Cat. Number	Genus	Species	Det. Label (Y/N)	Sex	Location Collected	Date Collected	Collector
INHS	382,292	*Beameria*	*venosa*	No	Male	Fults Hill Prairie	10.vii.1985	N/A
INHS	382,293	*Beameria*	*venosa*	No	Female	Fults Hill Prairie	10.vii.1985	N/A
INHS	382,294	*Beameria*	*venosa*	No	Male	Fults Hill Prairie	10.vii.1985	N/A
INHS	382,295	*Beameria*	*venosa*	No	Male	Fults Hill Prairie	10.vii.1985	N/A
INHS	557,546	*Beameria*	*venosa*	Yes	Male	Fults Prairie, Monroe Co., IL	14.viii.2007	O.R. Rakitov
INHS	557,547	*Beameria*	*venosa*	Yes	Male	Fults Prairie, Monroe Co., IL	14.viii.2007	O.R. Rakitov
INHS	557,548	*Beameria*	*venosa*	Yes	Male	Fults Prairie, Monroe Co., IL	14.viii.2007	O.R. Rakitov
INHS	557,549	*Beameria*	*venosa*	Yes	Female	Fults Prairie, Monroe Co., IL	14.viii.2007	O.R. Rakitov
INHS	667,048	*Oncotympana*	*maculaticollis*	Yes	Female	Port Arthur(TX?)	6-Aug-24	W.T.Davis
INHS	667,049	*Oncotympana*	*maculaticollis*	Yes	Female	Port Arthur(TX?)	6-Aug-24	W.T.Davis
INHS	669,162	*Tibicen*	*canicularis*	Yes		Algonquin, IL	7.viii.1907	W.A. Nason Col.
INHS	669,191	*Tibicen*	*davisi*	No	Male	Santa Rosa, N. Mex., Guadelupe County, 7/13/1961	7/13/1961	J.M. Kingsolver
INHS	669,195	*Tibicen*	*dealbatus*	No	Female	NM Otero Co,. Dog Canyon, 14 mi SSE Alamogordo, 6/22/1989, Malaise trap	6/22/1989	D.W.Webb/M.E. Irwin
INHS	674,041	*Tibicen*	*duryi*	No	Male	Pelahatchie Miss.	8/31/1921	Sweany Edwards
INHS	674,042	*Tibicen*	*figuratus*	Yes-W.T.M. Davis	Female	McCool, Miss.	ix/7/1921	Cecil Kennedy
INHS	674,049	*Tibicen*	*linnei*	Yes		Peoria, IL	3.viii.1938	F.F. Hasbrouck
INHS	674,187	*Tibicen*	*pronatalis*	No	Male	Hattiesburg, Miss.	ix-6-1921	M.L. Hemeter
INHS	674,198	*Tibicen*	*pronatalis*	Yes-T.E. Moore	Male	Belleville, IL. 8/27/1953, Sand	8/27/1953	Samuel F Moore
INHS	674,338	*Tibicen*	*pruinosus*	YES		Champaign, IL	3-viii-2001	C.H. Dietrich
INHS	674,344	*Tibicen*	*resonans*	No	Male	CAT 15 (?) Miss	9/7/1920	R.W Harned

Collection	Cat. Number	Genus	Species	Det. Label (Y/N)	Sex	Location Collected	Date Collected	Collector
INHS	674,347	*Tibicen*	*resonans*	Yes-T.E. Moore	Female	Sanford Fl 8/8/1939	1961, exact date unknown	R.H Beamer
INHS	674,348	*Tibicen*	*resh*	Yes-Davis	Male	Brownsville Tex. Dorner	unknown	Hald?
INHS	674,350	*Tibicen*	*resh*	Yes-T.E. Moore	Female	Tex.	1961, exact date unknown	Haldemann
INHS	744,393			No	Male	Douglas Co, Kansas near Lawrence	Sometime in the 80s	D. Yanega
INHS	744,428	*Neocicada*	*hieroglyphica*	No	Male	Mason Co., IL	22.vii.1967	Allison Roeske
INHS	744,439			No	Male	Sunken Gardens, St. Pete., Fla. 9/9/1965	9/9/1965	WVB?
INHS	744,478			No	Male	New Mexico Portales 7/15/1971	7/15/1971	W.H Bright
INHS	744,487	*Okanagana*		No		Lake Co., MN, off shore of Newton Lake	9.vii.1972	T.C Harsh et al
INHS	809,817			No	Male	Tennessee?	Sep-82	W.R.Rose
Field Museum	4,188,480			No		Relict prairie along IL central RR, 3 mi S of Monee, Will Co., IL	7/9/1961	H.S Dybas
Field Museum	4,188,481	*Okanagana*		No	Male	Relict prairie along IL central RR, 3 mi S of Monee, Will Co., IL	7/9/1961	H.S Dybas
Field Museum	4,188,482	*Okanagana*		No	Male	Kensington Railroad Tracks, Cook Co., IL	6/29/1978	R. W. Hamilton
Field Museum	4,188,483			No	Male	Kensington Railroad Tracks, Cook Co., IL	6/29/1978	R.W. Hamilton
Field Museum	4,188,486	*Okanagana*	*rimosa*	No	Male	Marquette Co., Michigan	10/1/1956	H.S Dybas
Field Museum	4,188,487	*Megatibicen*	*dorsatus*	No	Male	Mason Co., IL	8/18/2001	C. Grinter, D. Pollock, J. Louderman
Field Museum	4,188,488	*Neotibicen*	*auriferus?*	No	Male	Franklin County, Kansas (?)	1915, Exact Date Unknown	H.S Dybas
Field Museum	4,188,489			No	Female	Wayne Co., Ohio	1.vi.1965	H.S Dybas
Field Museum	4,188,490			No	Male	Wayne Co., Ohio	i.vi.1965	H.S Dybas

Table A.2. Sites visited over the years of this study and those that include audio recordings utilized in the search for cicada species.

Audio Data	Site Description	Latitude	Longitude	City	County
	Ballard Nature Center	39.060924	-88.704838	Altamont	Effingham
	Beadles Barrens Nature Preserve	38.352199	-88.126372	Ellery	Edwards
	Carl Becker Nature Preserve	41.016948	-87.540925	Pembroke Township	Kankakee
	Eldon Hazlet State Park	38.667732	-89.324938	Carlyle	Clinton
	Forest Glen Preserve	40.008753	-87.570736	Westville	Vermilion
	Fox Ridge State Park	39.401670	-88.141958	Hutton	Coles
YES	Fults Hill Prairie Nature Preserve	38.158146	-90.191337	Prairie du Rocher	Monroe
	Green River State Wildlife Management Area	41.635160	-89.506840	East Grove Township	Lee
YES	Henry A. Gleason Nature Preserve	40.378370	-89.927142	Topeka	Mason
	Herschel Workman Pheasant Area	40.461273	-87.920212	Butler Township	Vermilion
	Illinois Ozarks Nature Preserve	38.286584	-90.302765	Valmeyer	Monroe
	Iroquois County Conservation Area	40.989210	-87.578200	Beaverville Township	Iroquois
	Kennekuk Cove County Park	40.193378	-87.716952	Danville, IL	Vermilion
YES	Loda Cemetery Prairie Nature Preserve	40.527125	-88.076185	Loda	Iroquois
	Long Branch Sand Prairie Nature Preserve	40.227314	-90.053285	Havana	Mason
	Meredosia Hill Prairie Nature Preserve	39.856387	-90.464523	Arenzville	Morgan
	Merwin Savannah Nature Preserve	40.666075	-88.893600	Money Creek Township	McLean
YES	Nachusa Grasslands Nature Preserve	41.882542	-89.359092	Nachusa Township	Lee
	Olin Nature Preserve	38.916003	-90.224507	Godfrey Township	Madison
	Pellsville Cemetery Prairie	40.461001	-87.924490	Rankin	Vermilion
	Perdueville Habitat Area	40.405209	-88.213282	Paxton	Ford
	Prospect Cemetery Nature Preserve	40.444858	-88.097316	Paxton	Ford
	Rankin Right-of-Way Prairie	40.472686	-87.829045	Butler Township	Vermilion
YES	Revis Hill Prairie Nature Preserve	40.152862	-89.852320	Easton	Mason
	Richardson Wildlife Foundation	41.721890	-89.181650	West Brooklyn	Lee
	Ridgetop Hill Prairie Nature Preserve	40.652414	-89.161154	Secor	Woodford
	Russel M. Duffin Woods Nature Preserve	40.000328	-87.535810	Danville, IL	Vermilion
	Salt Lick Point Land and Water Reserve	38.303980	-90.307747	Valmeyer	Monroe

Audio Data	Site Description	Latitude	Longitude	City	County
	Sam Parr State Park	39.027065	-88.129087	Wade Township	Jasper
	Sand Prairie Scrub Oak Nature Preserve	40.190246	-90.074421	Bath Township	Mason
	Sand Ridge State Forest	40.391279	-89.872003	Forest City	Mason
	Simpson Township Barrens	37.479640	-88.739077	Simpson	Johnson
	South of Buckley Right-of-Way Prairie	40.581729	-88.044895	Buckley	Iroquois
	South of Ludlow Right-of-Way Prairie	40.357046	-88.139563	Ludlow Township	Champaign
	South of Paxton Right-of-Way Prairie	40.429704	-88.108996	Paxton	Ford
	Stephen A. Forbes State Recreational Area	38.722212	-88.772616	Kinmundy	Marion
	Vermilion River Observatory	40.059779	-87.564859	Danville, IL	Vermilion
	War Buff Valley Sanctuary	37.445882	-88.492191	Lusk	Pope
YES	Weston Cemetery Prairie Nature Preserve	40.747006	-88.614788	Chenoa	McLean
	Wildcat Hollow State Habitat Area	38.995783	-88.618410	Mason	Effingham
	Woodford State Fish and Wildlife Area	40.878619	-89.446638	Low Point	Woodford

GenBank Submission

```
LOCUS       OK626637               1456 bp    DNA     linear   INV 26-OCT-2021
DEFINITION  Neotibicen auriferus voucher INHS 837721 cytochrome c oxidase
            subunit I (COX1) gene, partial cds; mitochondrial.
ACCESSION   OK626637
VERSION     OK626637
KEYWORDS    .
SOURCE      mitochondrion Neotibicen auriferus
  ORGANISM  Neotibicen auriferus
            Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta;
            Pterygota; Neoptera; Paraneoptera; Hemiptera; Auchenorrhyncha;
            Cicadoidea; Cicadidae; Cicadinae; Cryptotympanini; Neotibicen.
REFERENCE   1  (bases 1 to 1456)
  AUTHORS   Dana,C.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2021) Illinois Natural History Survey, University
            of Illinois at Urbana-Champaign, 1816 S Oak St, Champaign, IL
            61820, USA
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..1456
                     /organism="Neotibicen auriferus"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /specimen_voucher="INHS 837721"
                     /db_xref="taxon:1699720"
     gene            <1..>1456
                     /gene="COX1"
     CDS             <1..>1456
                     /gene="COX1"
                     /codon_start=2
                     /transl_table=5
                     /product="cytochrome c oxidase subunit I"
```

/protein_id="UDL18945"
/translation="YFIFGIWSGMIGTALSTLIRIELGIPGSFIGDDQIYNVIVTAHA
FIMIFFMVMPIMMGGFGNWLVPLMIGAPDMAFPRMNNMSFWLLPPSLTLLLVGSLVDN
GAGTGWTVYPPLSSYMFHSGSCVDLTIFSLHLAGVSSILGAVNFISTIFNMRSTGMGL
DKTPLFVWSVLITAFLLLLSLPVLAGAITMLLTDRNLNTCFFDPSGGGDPILYQHLFW
FFGHPEVYILILPGFGLISHIITQESGKIESFGSLGMIYAMMSIGILGFVVWAHHMFT
VGMDVDTRAYFTSATMIIAVPTGIKVFSWLATLNGSKMKMSSSILWSLGFVFLFTMGG
LTGVILANSSIDIVLHDTYYVVAHFHYVLSMGAVFAILGSFVHWYSLFTGISLNPKWL
KIQFSIMFIGVNLTFFPQHFLGLSGMPRRYSDYPDSYMTWNIISSLGSVISLVGIMML
MFIVWESFISMRIVTFSKNMSSSVEWLQKFPPSEH"

ORIGIN

```
      1 atattttatt tttggcattt gatcgggaat aattggtaca gctcttagaa ccttaattcg
     61 aattgaatta ggtattcctg gctcatttat tggtgacgat caaatttata atgttattgt
    121 tacagctcat gcatttatta taatttttt tatagttatg cctattataa taggtggctt
    181 tggtaattgg ttagtacctt taataattgg agctcctgat atagcttttc ctcgaataaa
    241 taatatgaga ttttgacttc ttcctccttc tttaactttg ttattagtag gtagattagt
    301 tgataatggt gctggaactg gttgaacagt ttatccacca ttatcaagat acatgtttca
    361 ttctggttca tgtgttgatt taacaatttt ttctttacat ttggcaggtg tatcatcaat
    421 tctaggagct gtaaatttta ttagaacaat ttttaatata cgttcaactg gcataggtct
    481 tgataagact cctttatttg tttgatcagt tttaattact gcatttttat tattattatc
    541 attaccagtt ttagctggtg caattactat attattaact gatcgtaatc taaatacatg
    601 ttttttgat ccatctggag ggggtgatcc tattctttat caacatctgt tttgattttt
    661 tggacatccn gaagtttata ttttaatttt acctggattt ggattaattt cacatattat
    721 tactcaagaa agaggtaaaa ttgaatcatt tggttcatta ggaataattt atgcaataat
    781 gtcaattggt attcttggat ttgtagtttg agctcatcat atatttacag ttggaataga
    841 tgttgatact cgagcctatt ttacatctgc tactataatt attgctgttc caacaggaat
    901 taaagttttt agatgattag caacattaaa tggcagaaaa atgaaaatga gttcatctat
    961 tttatgatct ttaggatttg tatttttatt tacaatagga ggtttaactg gtgttatttt
   1021 ggctaattca tcaattgata ttgttttaca tgatacttat tatgttgttg ctcattttca
   1081 ttatgtttta tcaataggag cggtatttgc aattttaggt agatttgttc attgatattc
   1141 attgtttaca ggaatttcct taaatccaaa atgattaaaa attcaatttt caattatatt
   1201 tattggggtt aatttaacat tttttcctca acatttttg ggattaagag gaatacctcg
   1261 acgatattct gactatccag atagatatat aacatgaaat attatttctt cattaggaag
   1321 agtaatttca ttagttggaa ttatgatgtt aatatttatc gtatgagaaa gatttatttc
   1381 aatacgtatc gtaacttttt ccaaaaatat gagttcatca gtagaatgat tacaaaaatt
   1441 cccaccatct gaacat
```

```
//
LOCUS       OK626638               779 bp    DNA     linear   INV 26-OCT-2021
DEFINITION  Neotibicen auriferus voucher INHS 837722 cytochrome c oxidase
            subunit I (COX1) gene, partial cds; mitochondrial.
ACCESSION   OK626638
VERSION     OK626638
KEYWORDS    .
SOURCE      mitochondrion Neotibicen auriferus
  ORGANISM  Neotibicen auriferus
            Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta;
            Pterygota; Neoptera; Paraneoptera; Hemiptera; Auchenorrhyncha;
            Cicadoidea; Cicadidae; Cicadinae; Cryptotympanini; Neotibicen.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Dana,C.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2021) Illinois Natural History Survey, University
            of Illinois at Urbana-Champaign, 1816 S Oak St, Champaign, IL
            61820, USA
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..779
                     /organism="Neotibicen auriferus"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /specimen_voucher="INHS 837722"
                     /db_xref="taxon:1699720"
     gene            <1..>779
                     /gene="COX1"
     CDS             <1..>779
                     /gene="COX1"
                     /codon_start=2
                     /transl_table=5
                     /product="cytochrome c oxidase subunit I"
                     /protein_id="UDL18946"
```

/translation="YILILPGFGLISHIITQESGKIESFGSLGMIYAMMSIGILGFVV
WAHHMFTVGMDVDTRAYFTSATMIIAVPTGIKVFSWLATLNGSKMKMSSSILWSLGFV
FLFTMGGLTGVILANSSIDIVLHDTYYVVAHFHYVLSMGAVFAILGSFVHWYSLFTGI
SLNPKWLKIQFSIMFIGVNLTFFPQHFLGLSGMPRRYSDYPDSYMTWNIISSLGSVIS
LVGIMMLMFIVWESFISMRIVTFSKNMSSSVEWLQKFPPSE"
ORIGIN
1 ttatatttta attttacctg gatttggatt aatttcacat attattactc aagaaagagg
61 taaaattgaa tcatttggtt cattaggaat aatttatgca ataatgtcaa ttggtattct
121 tggatttgta gtttgagctc atcatatatt tacagttgga atagatgttg atactcgagc
181 ctattttaca tctgctacta taattattgc tgttccaaca ggaattaaag tttttagatg
241 attagcaaca ttaaatggca gaaaaatgaa aatgagttca tctattttat gatctttagg
301 atttgtattt ttatttacaa taggaggttt aactggtgtt attttggcta attcatcaat
361 tgatattgtt ttacatgata cttattatgt tgttgctcat tttcattatg ttttatcaat
421 aggagcggta tttgcaattt taggtagatt tgttcattga tattcattgt ttacaggaat
481 ttccttaaat ccaaaatgat taaaaattca attttcaatt atatttattg gggttaattt
541 aacatttttt cctcaacatt ttttgggatt aagaggaata cctcgacgat attctgacta
601 tccagataga tatataacat gaaatattat ttcttcatta ggaagagtaa tttcattagt
661 tggaattatg atgttaatat ttatcgtatg agaaagattt atttcaatac gtatcgtaac
721 tttttccaaa aatatgagtt catcagtaga atgattacaa aaattcccac catctgaac
//

```
LOCUS       OK626639               1440 bp    DNA     linear   INV 26-OCT-2021
DEFINITION  Neotibicen auriferus voucher INHS 837724 cytochrome c oxidase
            subunit I (COX1) gene, partial cds; mitochondrial.
ACCESSION   OK626639
VERSION     OK626639
KEYWORDS    .
SOURCE      mitochondrion Neotibicen auriferus
  ORGANISM  Neotibicen auriferus
            Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta;
            Pterygota; Neoptera; Paraneoptera; Hemiptera; Auchenorrhyncha;
            Cicadoidea; Cicadidae; Cicadinae; Cryptotympanini; Neotibicen.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Dana,C.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2021) Illinois Natural History Survey, University
            of Illinois at Urbana-Champaign, 1816 S Oak St, Champaign, IL
            61820, USA
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..1440
                     /organism="Neotibicen auriferus"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /specimen_voucher="INHS 837724"
                     /db_xref="taxon:1699720"
     gene            <1..>1440
                     /gene="COX1"
     CDS             <1..>1440
                     /gene="COX1"
                     /codon_start=1
                     /transl_table=5
                     /product="cytochrome c oxidase subunit I"
                     /protein_id="UDL18947"
                     /translation="IWSGMIGTALSTLIRIELGIPGSFIGDDQIYNVIVTAHAFIMIF
```

FMVMPIMMGGFGNWLVPLMIGAPDMAFPRMNNMSFWLLPPSLTLLLVGSLVDNGAGTG
WTVYPPLSSYMFHSGSCVDLTIFSLHLAGVSSILGAVNFISTIFNMRSTGMGLDKTPL
FVWSVLITAFLLLLSLPVLAGAITMLLTDRNLNTCFFDPSGGGDPILYQHLFWFFGHP
EVYILILPGFGLISHIITQESGKIESFGSLGMIYAMMSIGILGFVVWAHHMFTVGMDV
DTRAYFTSATMIIAVPTGIKVFSWLATLNGSKMKMSSSILWSLGFVFLFTMGGLTGVI
LANSSIDIVLHDTYYVVAHFHYVLSMGAVFAILGSFVHWYSLFTGISLNPKWLKIQFS
IMFIGVNLTFFPQHFLGLSGMPRRYSDYPDSYMTWNIISSLGSVISLVGIMMLMFIVW
ESFISMRIVTFSKNMSSSVEWLQKFPPSEH"

ORIGIN

```
      1 atttgatcgg gaataattgg tacagctctt agaaccttaa ttcgaattga attaggtatt
     61 cctggctcat ttattggtga cgatcaaatt tataatgtta ttgttacagc tcatgcattt
    121 attataattt tttttatagt tatgcctatt ataataggtg gctttggtaa ttggttagta
    181 cctttaataa ttggagctcc tgatatagct tttcctcgaa taaataatat gagattttga
    241 cttcttcctc cttctttaac tttgttatta gtaggtagat tagttgataa tggtgctgga
    301 actggttgaa cagtttatcc accattatca agatacatgt ttcattctgg ttcatgtgtt
    361 gatttaacaa ttttttcttt acatttggca ggtgtatcat caattctagg agctgtaaat
    421 tttattagaa caatttttaa tatacgttca actggcatag gtcttgataa gactccttta
    481 tttgtttgat cagttttaat tactgcattt ttattattat tatcattacc agttttagct
    541 ggtgcaatta ctatattatt aactgatcgt aatctaaata catgtttttt tgatccatct
    601 ggaggggggtg atcctattct ttatcaacat ctgttttgat tttttggaca tcctgaagtt
    661 tatattttaa ttttacctgg atttggatta atttcacata ttattactca agaaagaggt
    721 aaaattgaat catttggttc attaggaata atttatgcaa taatgtcaat tggtattctt
    781 ggatttgtag tttgagctca tcatatattt acagttggaa tagatgttga tactcgagcc
    841 tattttacat ctgctactat aattattgct gttccaacag gaattaaagt ttttagatga
    901 ttagcaacat taaatggcag aaaaatgaaa atgagttcat ctattttatg atctttagga
    961 tttgtatttt tatttacaat aggaggttta actggtgtta ttttggctaa ttcatcaatt
   1021 gatattgttt tacatgatac ttattatgtt gttgctcatt ttcattatgt tttatcaata
   1081 ggagcggtat ttgcaatttt aggtagattt gttcattgat attcattgtt tacaggaatt
   1141 tccttaaatc caaaatgatt aaaaattcaa ttttcaatta tatttattgg ggttaattta
   1201 acatttttc ctcaacattt tttgggatta agaggaatac ctcgacgata ttctgactat
   1261 ccagatagat atataacatg aaatattatt tcttcattag gaagagtaat ttcattagtt
   1321 ggaattatga tgttaatatt tatcgtatga gaaagattta tttcaatacg tattgtaact
   1381 ttttccaaaa atatgagttc atcagtagaa tgattacaaa aattcccacc atctgaacat
//
```

LOCUS OK626640 1470 bp DNA linear INV 26-OCT-2021
DEFINITION Neotibicen auriferus voucher INHS 837725 cytochrome c oxidase subunit I (COX1) gene, partial cds; mitochondrial.
ACCESSION OK626640
VERSION OK626640
KEYWORDS .
SOURCE mitochondrion Neotibicen auriferus
ORGANISM Neotibicen auriferus
Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Auchenorrhyncha; Cicadoidea; Cicadidae; Cicadinae; Cryptotympanini; Neotibicen.
REFERENCE 1 (bases 1 to 1470)
AUTHORS Dana,C.E.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-2021) Illinois Natural History Survey, University of Illinois at Urbana-Champaign, 1816 S Oak St, Champaign, IL 61820, USA
COMMENT ##Assembly-Data-START##
Sequencing Technology :: Sanger dideoxy sequencing
##Assembly-Data-END##
FEATURES Location/Qualifiers
source 1..1470
/organism="Neotibicen auriferus"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/specimen_voucher="INHS 837725"
/db_xref="taxon:1699720"
gene <1..>1470
/gene="COX1"
CDS <1..>1470
/gene="COX1"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="UDL18948"
/translation="DIGTLYFIFGIWSGMIGTALSTLIRIELGIPGSFIGDDQIYNVI

VTAHAFIMIFFMVMPIMMGGFGNWLVPLMIGAPDMAFPRMNNMSFWLLPPSLTLLLVG
SLVDNGAGTGWTVYPPLSSYMFHSGSCVDLTIFSLHLAGVSSILGAVNFISTIFNMRS
TGMGLDKTPLFVWSVLITAFLLLLSLPVLAGAITMLLTDRNLNTCFFDPSGGGDPILY
QHLFWFFGHPEVYILILPGFGLISHIITQESGKIESFGSLGMIYAMMSIGILGFVVWA
HHMFTVGMDVDTRAYFTSATMIIAVPTGIKVFSWLATLNGSKMKMSSSILWSLGFVFL
FTMGGLTGVILANSSIDIVLHDTYYVVAHFHYVLSMGAVFAILGSFVHWYSLFTGISL
NPKWLKIQFSIMFIGVNLTFFPQHFLGLSGMPRRYSDYPDSYMTWNIISSLGSVISLV
GIMMLMFIVWESFISMRIVTFSKNMSSSVEWLQKFPPSEH"

ORIGIN

```
        1 gatattggaa ctttatattt tatttttggc atttgatcgg gaataattgg tacagctctt
       61 agaaccttaa ttcgaattga attaggtatt cctggctcat ttattggtga cgatcaaatt
      121 tataatgtta ttgttacagc tcatgcattt attataattt tttttatagt tatgcctatt
      181 ataataggtg gctttggtaa ttggttagta cctttaataa ttggagctcc tgatatagct
      241 tttcctcgaa taaataatat gagattttga cttcttcctc cttctttaac tttgttatta
      301 gtaggtagat tagttgataa tggtgctgga actggttgaa cagtttatcc accattatca
      361 agatacatgt ttcattctgg ttcatgtgtt gatttaacaa ttttttcttt acatttggca
      421 ggtgtatcat caattctagg agctgtaaat tttattagaa caatttttaa tatacgttca
      481 actggcatag gtcttgataa gactccttta tttgtttgat cagttttaat tactgcattt
      541 ttattattat tatcattacc agttttagct ggtgcaatta ctatattatt aactgatcgt
      601 aatctaaata catgtttttt tgatccatct ggagggggtg atcctattct ttatcaacat
      661 ctgttttgat tttttggaca tccwgaagtt tatattttaa ttttacctgg atttggatta
      721 atttcacata ttattactca agaaagaggt aaaattgaat catttggttc attaggaata
      781 atttatgcaa taatgtcaat tggtattctt ggatttgtag tttgagctca tcatatattt
      841 acagttggaa tagatgttga tactcgagcc tattttacat ctgctactat aattattgct
      901 gttccaacag gaattaaagt ttttagatga ttagcaacat taaatggcag aaaaatgaaa
      961 atgagttcat ctattttatg atctttagga tttgtatttt tatttacaat aggaggttta
     1021 actggtgtta ttttggctaa ttcatcaatt gatattgttt tacatgatac ttattatgtt
     1081 gttgctcatt ttcattatgt tttatcaata ggagcggtat ttgcaatttt aggtagattt
     1141 gttcattgat attcattgtt tacaggaatt tccttaaatc caaaatgatt aaaaattcaa
     1201 ttttcaatta tatttattgg ggttaattta acatttttc ctcaacattt tttgggatta
     1261 agaggaatac ctcgacgata ttctgactat ccagatagat atataacatg aaatattatt
     1321 tcttcattag gaagagtaat ttcattagtt ggaattatga tgttaatatt tatcgtatga
     1381 gaaagattta tttcaatacg tatcgtaact ttttccaaaa atatgagttc atcagtagaa
     1441 tgattacaaa aattcccacc atctgaacat
//
```

APPENDIX B: CHAPTER 2 SUPPLEMENTARY MATERIALS

Table B.1. Samples used in the initial creation of library, including their barcode and index information. Retained reads is after processing with stacks. The year was the year collected and a general site code is given for the latitude, longitude, and year information. If an individual was kept for final analysis and not discarded for low reads or poor coverage, it is indicated in the Final Analysis column as "Yes".

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	AGAATCG-GACCTAAC	7810096	YES	2015	NDor1501F	F	40.5274	-88.0761	2015_LODA
PLATE_1	AGAACTACA-GACCTAAC	10941606	YES	2015	NDor1503F	F	40.5274	-88.0761	2015_LODA
PLATE_1	CCACATGCA-GACCTAAC	8592434	YES	2015	NDor1504M	M	40.5274	-88.0761	2015_LODA
PLATE_1	CCTCACG-GACCTAAC	11738183	YES	2015	NDor1507F	F	40.5274	-88.0761	2015_LODA
PLATE_1	AGTGGTCGGT-GACCTAAC	6499534	YES	2015	NDor1510M	M	40.5274	-88.0761	2015_LODA
PLATE_1	CCTCCAGA-GACCTAAC	6159652	YES	2015	NDor1512F	F	40.5274	-88.0761	2015_LODA
PLATE_1	TTCTGACCA-GACCTAAC	11161938	YES	2015	NDor1516M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	CCAATGA-GACCTAAC	12459964	YES	2015	NDor1517F	F	40.3801	-89.9300	2015_GLEASON
PLATE_1	CTTGTTGTAA-GACCTAAC	10306802	YES	2015	NDor1520M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	CTTATG-GACCTAAC	10736421	YES	2015	NDor1521M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	ACCACTG-GACCTAAC	11289009	YES	2015	NDor1522M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	GGTGCCA-GACCTAAC	9503588	YES	2015	NDor1523M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	GCTGGA-GACCTAAC	9061007	YES	2015	NDor1524M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	CACCTAGCA-GACCTAAC	5080409	YES	2015	NDor1525M	M	40.3801	-89.9300	2015_GLEASON

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	CCTCAGCA-GACCTAAC	10373732	YES	2015	NDor1526M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	AAGATCAA-GACCTAAC	9107056	YES	2015	NDor1527F	F	40.3801	-89.9300	2015_GLEASON
PLATE_1	CGGTGGTGGA-GACCTAAC	6954379	YES	2015	NDor1528M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	TTCCTGCCA-GACCTAAC	10541073	YES	2015	NDor1529M	M	40.3801	-89.9300	2015_GLEASON
PLATE_1	AGTGGCG-GACCTAAC	10368369	YES	2015	NDor1531M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	GAGTACG-GACCTAAC	12936073	YES	2015	NDor1532M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	TTCTTGAA-GACCTAAC	11885438	YES	2015	NDor1533M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	TTCTGACA-GACCTAAC	11859129	YES	2015	NDor1534M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	GGTGGCCA-GACCTAAC	8071036	YES	2015	NDor1535M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	GCGGTCCA-GACCTAAC	9555102	YES	2015	NDor1536F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	GGTGGACCA-GACCTAAC	9417466	YES	2015	NDor1537F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	GGTGACACA-GACCTAAC	10799836	YES	2015	NDor1538F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	GAACAT-GACCTAAC	4774178	YES	2015	NDor1539F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	CCACCACTCG-GACCTAAC	5605896	YES	2015	NDor1540F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	ACCTCG-GACCTAAC	5592801	YES	2015	NDor1541F	F	40.4296	-88.1091	2015_SPAX
PLATE_1	TGAACA-GACCTAAC	2547769	YES	2015	NDor1542M	M	40.4296	-88.1091	2015_SPAX
PLATE_1	TGCGGCACA-GACCTAAC	11933412	YES	2015	NDor1543M	M	39.0609	-88.7048	2015_BNC
PLATE_1	AATCAG-GACCTAAC	12048097	YES	2015	NDor1544M	M	39.0609	-88.7048	2015_BNC
PLATE_1	GAACAACAAT-GACCTAAC	9757233	YES	2015	NDor1549F	F	40.5274	-88.0761	2015_LODA

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	TTCAAGCA-GACCTAAC	9286794	YES	2015	NDor1550M	M	40.5274	-88.0761	2015_LODA
PLATE_1	CTTCTGA-GACCTAAC	7983223	YES	2015	NDor1554F	F	40.5274	-88.0761	2015_LODA
PLATE_1	AAGAACGAAT-GACCTAAC	8543627	YES	2015	NDor1564M	M	40.5274	-88.0761	2015_LODA
PLATE_1	TTGTACCA-GACCTAAC	6437309	YES	2015	NDor1567F	F	40.5274	-88.0761	2015_LODA
PLATE_1	AAGATACCA-GACCTAAC	501239	NO	2015	NDor1569M	M	40.5274	-88.0761	2015_LODA
PLATE_1	TGTTAACA-GACCTAAC	7753867	YES	2015	NDor1570M	M	38.7684	-88.8011	2015_SOL
PLATE_1	GTGGCCGCA-GACCTAAC	7043131	YES	2015	NDor1571F	F	38.7684	-88.8011	2015_SOL
PLATE_1	CCAACTGA-GACCTAAC	8960795	YES	2015	NDor1575M	M	39.0260	-88.1649	2015_JASPER
PLATE_1	GGTGTCGGTA-GACCTAAC	9728352	YES	2015	NDor1576M	M	39.0260	-88.1649	2015_JASPER
PLATE_1	ACTGGTGGTT-GACCTAAC	11118732	YES	2015	NDor1577M	M	38.9305	-88.1923	2015_DONNELLEY
PLATE_1	TCTTAG-GACCTAAC	12539483	YES	2015	NDor1578M	M	38.9305	-88.1923	2015_DONNELLEY
PLATE_1	TGCGGA-GACCTAAC	12530165	YES	2015	NDor1579M	M	38.7438	-88.7812	2015_FORBES
PLATE_1	GAAGATCCA-GACCTAAC	7371835	YES	2016	NDor1603M	M	40.5274	-88.0761	2016_LODA
PLATE_1	CGTGGAA-GACCTAAC	9657556	YES	2016	NDor1624F	F	40.5318	-88.0666	2016_NLODA
PLATE_1	GGCGGTAGGT-GACCTAAC	9601355	YES	2016	NDor1625F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	GAGACT-GACCTAAC	11173875	YES	2016	NDor1626M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	TTCTGCA-GACCTAAC	8565273	YES	2016	NDor1627M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	CCACCATCAG-GACCTAAC	4124145	YES	2016	NDor1628F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	CCTACCACAG-GACCTAAC	11098850	YES	2016	NDor1634M	M	39.0609	-88.7048	2016_BNC

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	ACCACCATCG-GACCTAAC	5614601	YES	2016	NDor1636M	M	40.3801	-89.9300	2016_GLEASON
PLATE_1	GAGGTCA-GACCTAAC	11120550	YES	2016	NDor1637F	F	40.3801	-89.9300	2016_GLEASON
PLATE_1	GAGTGCCA-GACCTAAC	5987251	YES	2016	NDor1638F	F	39.0609	-88.7048	2016_BNC
PLATE_1	TTCAGT-GACCTAAC	11329275	YES	2016	NDor1642F	F	40.5274	-88.0761	2016_LODA
PLATE_1	AACATG-GACCTAAC	10620585	YES	2016	NDor1643F	F	40.5274	-88.0761	2016_LODA
PLATE_1	AGAACTCA-GACCTAAC	9713300	YES	2016	NDor1644F	F	40.5274	-88.0761	2016_LODA
PLATE_1	AACCGT-GACCTAAC	10370194	YES	2016	NDor1645F	F	40.5274	-88.0761	2016_LODA
PLATE_1	ACACCACCTG-GACCTAAC	5781463	YES	2016	NDor1651M	M	40.5274	-88.0761	2016_LODA
PLATE_1	CCATCCGCA-GACCTAAC	9459655	YES	2016	NDor1652M	M	40.5274	-88.0761	2016_LODA
PLATE_1	AAGCTCG-GACCTAAC	7957755	YES	2016	NDor1653M	M	40.5274	-88.0761	2016_LODA
PLATE_1	CGTGAA-GACCTAAC	3548465	YES	2016	NDor1654M	M	40.5274	-88.0761	2016_LODA
PLATE_1	CCACGT-GACCTAAC	3018039	YES	2016	NDor1665F	F	39.0609	-88.7048	2016_BNC
PLATE_1	AAGAATCA-GACCTAAC	6624617	YES	2016	NDor1666F	F	40.3801	-89.9300	2016_GLEASON
PLATE_1	GAGCATA-GACCTAAC	10522790	YES	2016	NDor1667M	M	40.3801	-89.9300	2016_GLEASON
PLATE_1	TGAGGCA-GACCTAAC	34735028	NO	2016	NDor1668M	M	40.5502	-88.0601	2016_MAINLINE
PLATE_1	CCACTGACA-GACCTAAC	2398217	NO	2016	NDor1670M	M	39.0609	-88.7048	2016_BNC
PLATE_1	CTGTAT-GACCTAAC	9875707	YES	2016	NDor1672M	M	39.0609	-88.7048	2016_BNC
PLATE_1	GAGGCTG-GACCTAAC	8688727	YES	2016	NDor1673F	F	39.0609	-88.7048	2016_BNC
PLATE_1	ACACTAG-GACCTAAC	7338324	YES	2016	NDor1674F	F	39.0609	-88.7048	2016_BNC

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	TTCTCAG-GACCTAAC	10474873	NO	2016	NDor1675F	F	39.0609	-88.7048	2016_BNC
PLATE_1	TTCTTGTTCA-GACCTAAC	16083998	YES	2016	NDor1676F	F	39.0609	-88.7048	2016_BNC
PLATE_1	GAACAAGATG-GACCTAAC	6771767	YES	2016	NDor1678M	M	39.0609	-88.7048	2016_BNC
PLATE_1	GAGAACAACT-GACCTAAC	2767233	YES	2016	NDor1679F	F	39.0609	-88.7048	2016_BNC
PLATE_1	TTGTGTTCGA-GACCTAAC	12415839	YES	2016	NDor1680M	M	39.0609	-88.7048	2016_BNC
PLATE_1	AACCTGCA-GACCTAAC	4430618	YES	2016	NDor1692M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	GGATCT-GACCTAAC	9981107	YES	2016	NDor1693M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	CCTACAGCA-GACCTAAC	9511664	YES	2016	NDor1694M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	TTGTTGTCTA-GACCTAAC	8368246	YES	2016	NDor1695F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	TTCACGA-GACCTAAC	9349106	YES	2016	NDor16100M	M	40.4296	-88.1091	2016_SPAX
PLATE_1	AGGTGGACA-GACCTAAC	6290097	YES	2016	NDor16101F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	TTCCATGCA-GACCTAAC	8531752	YES	2016	NDor16102F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	AACAAGAACT-GACCTAAC	8944932	YES	2016	NDor16103F	F	40.4296	-88.1091	2016_SPAX
PLATE_1	CACAGTCA-GACCTAAC	7139505	YES	2016	NDor16110F	F	40.3801	-89.9300	2016_GLEASON
PLATE_1	CCATAAG-GACCTAAC	7022180	NO	2016	NDor16111F	F	40.5274	-88.0761	2016_LODA
PLATE_1	CTGGTA-GACCTAAC	6537913	YES	2016	NDor16112F	F	40.5274	-88.0761	2016_LODA
PLATE_1	AAGAATGCA-GACCTAAC	6250764	YES	2016	NDor16113M	M	40.5274	-88.0761	2016_LODA
PLATE_1	ACACTG-GACCTAAC	8183803	YES	2016	NDor16122M	F	40.3801	-89.9300	2016_GLEASON
PLATE_1	TTCTCGACA-GACCTAAC	8610720	YES	2017	NDor1701M	M	40.4296	-88.1091	2017_SPAX

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_1	GTAGGCAA-GACCTAAC	10700343	YES	2017	NDor1702F	F	40.4296	-88.1091	2017_SPAX
PLATE_1	TTGTTGCTGA-GACCTAAC	104	NA		CONTROL_PLATE1				
PLATE_2	AGAATCG-ACCAACTG	10759439	YES	2017	Ndor1703M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	AGAACTACA-ACCAACTG	10172991	YES	2017	Ndor1704M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	CCACATGCA-ACCAACTG	8483380	YES	2017	Ndor1705M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	CCTCACG-ACCAACTG	9636610	YES	2017	Ndor1706M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	AGTGGTCGGT-ACCAACTG	9349299	YES	2017	Ndor1707M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	CCTCCAGA-ACCAACTG	8618231	YES	2017	Ndor1708M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	TTCTGACCA-ACCAACTG	7567525	YES	2017	Ndor1709F	F	40.4296	-88.1091	2017_SPAX
PLATE_2	CCAATGA-ACCAACTG	12678381	YES	2017	Ndor1710M	M	40.4296	-88.1091	2017_SPAX
PLATE_2	CTTGTTGTAA-ACCAACTG	9894708	YES	2017	NDor1729F	F	40.4663	-87.8330	2017_RANKIN
PLATE_2	CTTATG-ACCAACTG	7659925	YES	2017	NDor1730M	M	40.4663	-87.8330	2017_RANKIN
PLATE_2	ACCACTG-ACCAACTG	9712240	YES	2017	NDor1732F	F	40.4727	-87.8292	2017_RANKIN
PLATE_2	GGTGCCA-ACCAACTG	9271360	YES	2017	NDor1733M	M	40.4727	-87.8292	2017_RANKIN
PLATE_2	GCTGGA-ACCAACTG	11955275	YES	2017	NDor1734M	M	40.4727	-87.8292	2017_RANKIN
PLATE_2	CACCTAGCA-ACCAACTG	5446110	YES	2017	NDor1737M	M	40.4727	-87.8292	2017_RANKIN
PLATE_2	CCTCAGCA-ACCAACTG	6484271	YES	2017	NDor1738F	F	40.4727	-87.8292	2017_RANKIN
PLATE_2	AAGATCAA-ACCAACTG	8155301	YES	2017	NDor1741F	F	40.3801	-89.9300	2017_GLEASON
PLATE_2	CGGTGGTGGA-ACCAACTG	9844239	YES	2017	NDor1742M	M	40.3801	-89.9300	2017_GLEASON

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_2	TTCCTGCCA-ACCAACTG	9728502	YES	2017	NDor1743M	M	40.3801	-89.9300	2017_GLEASON
PLATE_2	AGTGGCG-ACCAACTG	8621964	YES	2017	Ndor1744M	M	40.3801	-89.9300	2017_GLEASON
PLATE_2	GAGTACG-ACCAACTG	9629952	YES	2018	NDor1806F	F	40.4296	-88.1091	2018_SPAX
PLATE_2	TTCTTGAA-ACCAACTG	4690797	YES	2018	NDor1811M	M	40.4296	-88.1091	2018_SPAX
PLATE_2	TTCTGACA-ACCAACTG	8451658	YES	2018	NDor1812F	F	40.4296	-88.1091	2018_SPAX
PLATE_2	GGTGGCCA-ACCAACTG	5118691	YES	2018	NDor1813M	M	40.4296	-88.1091	2018_SPAX
PLATE_2	GCGGTCCA-ACCAACTG	9391041	YES	2018	NDor1819M	M	40.5274	-88.0761	2018_LODA
PLATE_2	GGTGGACCA-ACCAACTG	7756104	YES	2018	NDor1825M	M	40.5274	-88.0761	2018_LODA
PLATE_2	GGTGACACA-ACCAACTG	8579429	YES	2018	NDor1826M	M	40.5274	-88.0761	2018_LODA
PLATE_2	GAACAT-ACCAACTG	4343720	YES	2018	NDor1833M	M	40.5274	-88.0761	2018_LODA
PLATE_2	CCACCACTCG-ACCAACTG	10870645	YES	2018	NDor1834M	M	40.5274	-88.0761	2018_LODA
PLATE_2	ACCTCG-ACCAACTG	8918661	YES	2018	NDor1838F	F	40.5274	-88.0761	2018_LODA
PLATE_2	TGAACA-ACCAACTG	9603955	YES	2018	NDor1845F	F	40.5274	-88.0761	2018_LODA
PLATE_2	TGCGGCACA-ACCAACTG	7310660	YES	2018	NDor1846F	F	40.5274	-88.0761	2018_LODA
PLATE_2	AATCAG-ACCAACTG	7145139	YES	2018	NDor1848F	F	40.5274	-88.0761	2018_LODA
PLATE_2	GAACAACAAT-ACCAACTG	5336748	YES	2018	NDor1849M	M	40.5274	-88.0761	2018_LODA
PLATE_2	TTCAAGCA-ACCAACTG	5973664	YES	2018	NDor1851M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	CTTCTGA-ACCAACTG	6499047	YES	2018	NDor1860F	F	40.3801	-89.9300	2018_GLEASON
PLATE_2	AAGAACGAAT-ACCAACTG	7953815	YES	2018	NDor1861F	F	40.3801	-89.9300	2018_GLEASON

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_2	TTGTACCA-ACCAACTG	7998743	YES	2018	NDor1862F	F	40.3801	-89.9300	2018_GLEASON
PLATE_2	AAGATACCA-ACCAACTG	7074965	YES	2018	NDor1863F	F	40.3801	-89.9300	2018_GLEASON
PLATE_2	TGTTAACA-ACCAACTG	4909196	YES	2018	NDor1864M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	GTGGCCGCA-ACCAACTG	5973680	YES	2018	NDor1865F	F	40.3801	-89.9300	2018_GLEASON
PLATE_2	CCAACTGA-ACCAACTG	5004004	YES	2018	NDor1866M	M	40.4320	-88.1082	2018_SPAX
PLATE_2	GGTGTCGGTA-ACCAACTG	5403082	YES	2018	NDor1867F	F	40.4320	-88.1082	2018_SPAX
PLATE_2	ACTGGTGGTT-ACCAACTG	4686547	YES	2018	NDor1868M	M	40.4320	-88.1082	2018_SPAX
PLATE_2	TCTTAG-ACCAACTG	8266867	YES	2018	NDor1872M	M	40.4727	-87.8292	2018_RANKIN
PLATE_2	TGCGGA-ACCAACTG	11296304	YES	2018	NDor1874M	M	40.4727	-87.8292	2018_RANKIN
PLATE_2	GAAGATCCA-ACCAACTG	5428375	YES	2018	NDor1875M	M	40.4727	-87.8292	2018_RANKIN
PLATE_2	CGTGGAA-ACCAACTG	7381738	YES	2018	NDor1876F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	GGCGGTAGGT-ACCAACTG	6105493	YES	2018	NDor1877F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	GAGACT-ACCAACTG	10090829	YES	2018	NDor1878F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	TTCTGCA-ACCAACTG	8503769	YES	2018	NDor1879F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	CCACCATCAG-ACCAACTG	4777458	YES	2018	NDor1882M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	CCTACCACAG-ACCAACTG	9066850	YES	2018	NDor1885M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	ACCACCATCG-ACCAACTG	7044950	YES	2018	NDor1888M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	GAGGTCA-ACCAACTG	8377112	YES	2018	NDor1889M	M	40.3801	-89.9300	2018_GLEASON
PLATE_2	GAGTGCCA-ACCAACTG	4594728	YES	2018	NDor1894M	M	40.4320	-88.1082	2018_SPAX

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_2	TTCAGT-ACCAACTG	8100180	YES	2018	NDor1895F	F	40.4320	-88.1082	2018_SPAX
PLATE_2	AACATG-ACCAACTG	7732034	YES	2018	NDor1896F	F	40.4320	-88.1082	2018_SPAX
PLATE_2	AGAACTCA-ACCAACTG	4609062	YES	2018	MDor1897F	F	40.4320	-88.1082	2018_SPAX
PLATE_2	AACCGT-ACCAACTG	9421140	YES	2018	MDor1898M	M	38.7882	-88.8274	2018_NKIN
PLATE_2	ACACCACCTG-ACCAACTG	7606565	YES	2018	MDor1899M	M	38.7882	-88.8274	2018_NKIN
PLATE_2	CCATCCGCA-ACCAACTG	8397240	YES	2018	MDor18100M	M	38.7882	-88.8274	2018_NKIN
PLATE_2	AAGCTCG-ACCAACTG	781637	NO	2018	MDor18102F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	CGTGAA-ACCAACTG	2206794	NO	2018	MDor18103M	M	38.7882	-88.8274	2018_NKIN
PLATE_2	CCACGT-ACCAACTG	8863291	YES	2018	MDor18104M	M	38.7882	-88.8274	2018_NKIN
PLATE_2	AAGAATCA-ACCAACTG	8754873	NO	2018	MDor18105F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	GAGCATA-ACCAACTG	8482095	YES	2018	MDor18106F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	TGAGGCA-ACCAACTG	7747454	YES	2018	MDor18107F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	CCACTGACA-ACCAACTG	8018286	YES	2018	MDor18108F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	CTGTAT-ACCAACTG	7443941	YES	2018	MDor18109F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	GAGGCTG-ACCAACTG	7274625	YES	2018	MDor18110F	F	38.7882	-88.8274	2018_NKIN
PLATE_2	ACACTAG-ACCAACTG	7013943	YES	2018	MDor18111F	F	40.3801	-89.9300	2018_GLEASON
PLATE_2	TTCTCAG-ACCAACTG	9949638	YES	2018	MDor18113M	M	40.4287	-88.1092	2018_SPAX
PLATE_2	TTCTTGTTCA-ACCAACTG	9504646	YES	2018	MDor18166M	M	40.4727	-87.8292	2018_RANKIN
PLATE_2	GAACAAGATG-ACCAACTG	7166307	YES	2018	MDor18167M	M	40.4727	-87.8292	2018_RANKIN

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_2	GAGAACAACT-ACCAACTG	4945219	YES	2018	MDor18168M	M	40.4727	-87.8292	2018_RANKIN
PLATE_2	TTGTGTTCGA-ACCAACTG	7940595	YES	2018	MDor18169F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	AACCTGCA-ACCAACTG	12234174	NO	2018	MDor18170F	F	40.4727	-87.8292	2018_RANKIN
PLATE_2	GGATCT-ACCAACTG	10165080	YES	2018	MDor18171M	M	40.5274	-88.0761	2018_LODA
PLATE_2	CCTACAGCA-ACCAACTG	5465486	YES	2018	MDor18172F	F	40.5274	-88.0761	2018_LODA
PLATE_2	TTGTTGTCTA-ACCAACTG	9872916	YES	2018	MDor18173M	M	40.2259	-90.0516	2018_LONGBRANCH
PLATE_2	TTCACGA-ACCAACTG	11241823	YES	2018	MDor18174M	M	40.2259	-90.0516	2018_LONGBRANCH
PLATE_2	AGGTGGACA-ACCAACTG	8321018	YES	2019	MDor1903M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	TTCCATGCA-ACCAACTG	1981034	YES	2019	MDor1904M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	AACAAGAACT-ACCAACTG	8250223	YES	2019	MDor1907M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	CACAGTCA-ACCAACTG	5171339	YES	2019	MDor1908M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	CCATAAG-ACCAACTG	3472409	YES	2019	MDor1909M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	CTGGTA-ACCAACTG	6014501	YES	2019	MDor1910M	M	40.4450	-88.0978	2019_PROSPECT
PLATE_2	AAGAATGCA-ACCAACTG	8993641	YES	2019	MDor1912F	F	40.4450	-88.0978	2019_PROSPECT
PLATE_2	ACACTG-ACCAACTG	11610861	YES	2019	MDor1913F	F	40.4450	-88.0978	2019_PROSPECT
PLATE_2	TTCTCGACA-ACCAACTG	10385727	YES	2019	MDor1914F	F	40.4450	-88.0978	2019_PROSPECT
PLATE_2	GTAGGCAA-ACCAACTG	6513198	YES	2019	MDor1915F	F	40.4450	-88.0978	2019_PROSPECT
PLATE_2	TTGTTGCTGA-ACCAACTG	56	NA	NA	CONTROL_PLATE1	NA	NA	NA	NA
PLATE_3	AGAATCG-TTCGCTGA	4869465	YES	2019	MDor1916F	F	40.4450	-88.0978	2019_PROSPECT

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_3	AGAACTACA-TTCGCTGA	7205555	YES	2019	MDor1919M	M	40.5274	-88.0761	2019_LODA
PLATE_3	CCACATGCA-TTCGCTGA	11130825	YES	2019	MDor1928M	M	40.5274	-88.0761	2019_LODA
PLATE_3	CCTCACG-TTCGCTGA	906555	NO	2019	MDor1932M	M	40.5274	-88.0761	2019_LODA
PLATE_3	AGTGGTCGGT-TTCGCTGA	12767181	YES	2019	MDor1936M	M	40.5274	-88.0761	2019_LODA
PLATE_3	CCTCCAGA-TTCGCTGA	6290551	YES	2019	MDor1945F	F	40.5274	-88.0761	2019_LODA
PLATE_3	TTCTGACCA-TTCGCTGA	13729858	YES	2019	MDor1953F	F	40.5274	-88.0761	2019_LODA
PLATE_3	CCAATGA-TTCGCTGA	14222360	YES	2019	MDor1954F	F	40.5274	-88.0761	2019_LODA
PLATE_3	CTTGTTGTAA-TTCGCTGA	10526645	YES	2019	MDor1957F	F	40.5274	-88.0761	2019_LODA
PLATE_3	CTTATG-TTCGCTGA	9485834	YES	2019	MDor1959F	F	38.6685	-89.3242	2019_ELDON
PLATE_3	ACCACTG-TTCGCTGA	12531101	YES	2019	MDor1960F	F	38.6685	-89.3242	2019_ELDON
PLATE_3	GGTGCCA-TTCGCTGA	9677651	YES	2019	MDor1961M	M	38.6685	-89.3242	2019_ELDON
PLATE_3	GCTGGA-TTCGCTGA	15111513	YES	2019	MDor1962M	M	38.6685	-89.3242	2019_ELDON
PLATE_3	CACCTAGCA-TTCGCTGA	10640062	YES	2019	MDor1963M	M	38.6685	-89.3242	2019_ELDON
PLATE_3	CCTCAGCA-TTCGCTGA	13634486	YES	2019	MDor1964F	F	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	AAGATCAA-TTCGCTGA	10486634	YES	2019	MDor1965F	F	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	CGGTGGTGGA-TTCGCTGA	11091108	YES	2019	MDor1966F	F	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	TTCCTGCCA-TTCGCTGA	9406477	YES	2019	MDor1967M	M	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	AGTGGCG-TTCGCTGA	16248979	YES	2019	MDor1968M	M	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	GAGTACG-TTCGCTGA	10081895	YES	2019	MDor1969M	M	40.5813	-88.0456	2019_BUCKLEY

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_3	TTCTTGAA-TTCGCTGA	7627933	YES	2019	MDor1970M	M	40.5813	-88.0456	2019_BUCKLEY
PLATE_3	TTCTGACA-TTCGCTGA	8575978	YES	2019	MDor1971F	F	40.5507	-88.0596	2019_MAINLINE
PLATE_3	GGTGGCCA-TTCGCTGA	7587596	YES	2019	MDor1972F	F	40.5507	-88.0596	2019_MAINLINE
PLATE_3	GCGGTCCA-TTCGCTGA	146780	NO	2019	MDor1983M	M	40.5815	-88.0449	2019_BUCKLEY
PLATE_3	GGTGGACCA-TTCGCTGA	459995	NO	2019	MDor1984M	M	40.5815	-88.0449	2019_BUCKLEY
PLATE_3	GGTGACACA-TTCGCTGA	8925689	YES	2019	MDor1985M	M	40.5815	-88.0449	2019_BUCKLEY
PLATE_3	GAACAT-TTCGCTGA	4570949	YES	2019	MDor1987F	F	40.5815	-88.0449	2019_BUCKLEY
PLATE_3	CCACCACTCG-TTCGCTGA	9014694	YES	2019	MDor1990F	F	40.5815	-88.0449	2019_BUCKLEY
PLATE_3	ACCTCG-TTCGCTGA	11845289	YES	2019	MDor1991F	F	40.5274	-88.0761	2019_LODA
PLATE_3	TGAACA-TTCGCTGA	4852111	YES	2019	MDor1992M	M	40.5274	-88.0761	2019_LODA
PLATE_3	TGCGGCACA-TTCGCTGA	11168707	YES	2019	MDor1994M	M	40.5274	-88.0761	2019_LODA
PLATE_3	AATCAG-TTCGCTGA	4893740	YES	2019	MDor1995F	F	40.5274	-88.0761	2019_LODA
PLATE_3	GAACAACAAT-TTCGCTGA	6142365	YES	2019	MDor1996M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	TTCAAGCA-TTCGCTGA	10779204	YES	2019	MDor1997M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	CTTCTGA-TTCGCTGA	9683911	YES	2019	MDor1998M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	AAGAACGAAT-TTCGCTGA	14157194	YES	2019	MDor1999M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	TTGTACCA-TTCGCTGA	10878468	YES	2019	MDor19100M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	AAGATACCA-TTCGCTGA	13266713	YES	2019	MDor19101M	M	38.8198	-88.7884	2019_TRACT5
PLATE_3	TGTTAACA-TTCGCTGA	6739399	YES	2019	MDor19102M	M	38.8198	-88.7884	2019_TRACT5

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_3	GTGGCCGCA-TTCGCTGA	6099714	NO	2019	MDor19104F	F	38.9149	-88.6716	2019_ORIOLE
PLATE_3	CCAACTGA-TTCGCTGA	19947218	YES	2019	MDor19105M	M	38.9149	-88.6716	2019_ORIOLE
PLATE_3	GGTGTCGGTA-TTCGCTGA	8531386	YES	2019	MDor19106M	M	40.4726	-87.8291	2019_RANKIN
PLATE_3	ACTGGTGGTT-TTCGCTGA	9028794	YES	2019	MDor19107M	M	40.4726	-87.8291	2019_RANKIN
PLATE_3	TCTTAG-TTCGCTGA	6347606	YES	2019	MDor19108M	M	40.4726	-87.8291	2019_RANKIN
PLATE_3	TGCGGA-TTCGCTGA	15012548	YES	2019	MDor19109M	M	40.4726	-87.8291	2019_RANKIN
PLATE_3	GAAGATCCA-TTCGCTGA	2363705	YES	2019	MDor19110M	M	40.4726	-87.8291	2019_RANKIN
PLATE_3	CGTGGAA-TTCGCTGA	10729120	YES	2019	MDor19111F	F	40.4726	-87.8291	2019_RANKIN
PLATE_3	GGCGGTAGGT-TTCGCTGA	11683478	YES	2019	MDor19112F	F	40.4726	-87.8291	2019_RANKIN
PLATE_3	GAGACT-TTCGCTGA	8769143	YES	2019	MDor19113F	F	40.4726	-87.8291	2019_RANKIN
PLATE_3	TTCTGCA-TTCGCTGA	7489196	YES	2019	MDor19114F	F	40.4726	-87.8291	2019_RANKIN
PLATE_3	CCACCATCAG-TTCGCTGA	5611928	YES	2019	MDor19115F	F	40.4726	-87.8291	2019_RANKIN
PLATE_3	CCTACCACAG-TTCGCTGA	4022397	NO	2019	MDor19116F	F	38.6691	-89.3258	2019_ELDON
PLATE_3	ACCACCATCG-TTCGCTGA	11165389	YES	2019	MDor19120M	M	40.4320	-88.1082	2019_SPAX
PLATE_3	GAGGTCA-TTCGCTGA	4144899	YES	2019	MDor19121M	M	40.4320	-88.1082	2019_SPAX
PLATE_3	GAGTGCCA-TTCGCTGA	4093907	YES	2019	MDor19122M	M	40.4320	-88.1082	2019_SPAX
PLATE_3	TTCAGT-TTCGCTGA	11593848	YES	2019	MDor19123M	M	40.4287	-88.1092	2019_SPAX
PLATE_3	AACATG-TTCGCTGA	10765369	YES	2019	MDor19124F	F	40.4320	-88.1082	2019_SPAX
PLATE_3	AGAACTCA-TTCGCTGA	10110250	YES	2019	MDor19125F	F	40.4320	-88.1082	2019_SPAX

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_3	AACCGT-TTCGCTGA	4830338	YES	2019	MDor19126F	F	40.4320	-88.1082	2019_SPAX
PLATE_3	ACACCACCTG-TTCGCTGA	7969952	YES	2019	MDor19127F	F	40.4320	-88.1082	2019_SPAX
PLATE_3	CCATCCGCA-TTCGCTGA	10924574	YES	2019	MDor19128F	F	40.4320	-88.1082	2019_SPAX
PLATE_3	AAGCTCG-TTCGCTGA	9413820	YES	2019	MDor19130F	F	40.4287	-88.1092	2019_SPAX
PLATE_3	CGTGAA-TTCGCTGA	2225371	YES	2019	MDor19132M	M	38.6691	-89.3258	2019_ELDON
PLATE_3	CCACGT-TTCGCTGA	3267120	YES	2019	MDor19133F	F	38.6685	-89.3242	2019_ELDON
PLATE_3	AAGAATCA-TTCGCTGA	8639014	YES	2019	MDor19136M	M	40.3801	-89.9300	2019_GLEASON
PLATE_3	GAGCATA-TTCGCTGA	5163649	YES	2019	MDor19137M	M	40.3801	-89.9300	2019_GLEASON
PLATE_3	TGAGGCA-TTCGCTGA	5306914	YES	2019	MDor19138M	M	40.3801	-89.9300	2019_GLEASON
PLATE_3	CCACTGACA-TTCGCTGA	6871877	YES	2019	MDor19139F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	CTGTAT-TTCGCTGA	7650884	YES	2019	MDor19140F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	GAGGCTG-TTCGCTGA	10535426	YES	2019	MDor19141F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	ACACTAG-TTCGCTGA	8151547	YES	2019	MDor19142F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	TTCTCAG-TTCGCTGA	16191233	YES	2019	MDor19143F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	TTCTTGTTCA-TTCGCTGA	14772876	YES	2019	MDor19144F	F	40.3801	-89.9300	2019_GLEASON
PLATE_3	GAACAAGATG-TTCGCTGA	15289379	YES	2019	MDor19147M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	GAGAACAACT-TTCGCTGA	1344843	NO	2019	MDor19148M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	TTGTGTTCGA-TTCGCTGA	10274362	YES	2019	MDor19149M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	AACCTGCA-TTCGCTGA	12478516	YES	2019	MDor19150M	M	38.9384	-88.6425	2019_TRACT2

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_3	GGATCT-TTCGCTGA	1190731	NO	2019	MDor19151M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	CCTACAGCA-TTCGCTGA	5866199	YES	2019	MDor19152M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	TTGTTGTCTA-TTCGCTGA	5457785	YES	2019	MDor19153M	M	38.9384	-88.6425	2019_TRACT2
PLATE_3	TTCACGA-TTCGCTGA	6891891	YES	2019	MDor19155F	F	38.9384	-88.6425	2019_TRACT2
PLATE_3	AGGTGGACA-TTCGCTGA	5449640	YES	2019	MDor19156F	F	38.9384	-88.6425	2019_TRACT2
PLATE_3	TTCCATGCA-TTCGCTGA	7385469	YES	2019	MDor19157F	F	38.9384	-88.6425	2019_TRACT2
PLATE_3	AACAAGAACT-TTCGCTGA	8724070	YES	2019	MDor19158M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	CACAGTCA-TTCGCTGA	13711560	YES	2019	MDor19160M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	CCATAAG-TTCGCTGA	12145562	YES	2019	MDor19161M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	CTGGTA-TTCGCTGA	10289889	YES	2019	MDor19163M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	AAGAATGCA-TTCGCTGA	2635535	YES	2019	MDor19164M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	ACACTG-TTCGCTGA	10071023	YES	2019	MDor19165M	M	38.8701	-88.7267	2019_TRACT3
PLATE_3	TTCTCGACA-TTCGCTGA	8184048	YES	2019	MDor19166F	F	38.8701	-88.7267	2019_TRACT3
PLATE_3	GTAGGCAA-TTCGCTGA	30520	NA	NA	CONTROL_PLATE_3	NA	NA	NA	NA
PLATE_4	AGAATCG-TATCAGCG	7008434	YES	2019	MDor19167F	F	38.8701	-88.7267	2019_TRACT3
PLATE_4	AGAACTACA-TATCAGCG	12644627	YES	2019	MDor19168F	F	38.8701	-88.7267	2019_TRACT3
PLATE_4	CCACATGCA-TATCAGCG	9049555	YES	2019	MDor19169F	F	38.8701	-88.7267	2019_TRACT3
PLATE_4	CCTCACG-TATCAGCG	6683183	YES	2019	MDor19170F	F	38.8701	-88.7267	2019_TRACT3
PLATE_4	AGTGGTCGGT-TATCAGCG	9359524	YES	2019	MDor19171M	M	38.8701	-88.7267	2019_TRACT3

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_4	CCTCCAGA-TATCAGCG	4801141	YES	2019	MDor19172F	F	40.3705	-88.1337	2019_LUDLOW
PLATE_4	TTCTGACCA-TATCAGCG	9545171	YES	2019	MDor19173M	M	40.3705	-88.1337	2019_LUDLOW
PLATE_4	CCAATGA-TATCAGCG	23733804	YES	2019	MDor19174M	M	38.8198	-88.7884	2019_TRACT5
PLATE_4	CTTGTTGTAA-TATCAGCG	4781208	YES	2019	MDor19175F	F	40.5502	-88.0601	2019_MAINLINE
PLATE_4	CTTATG-TATCAGCG	38722475	NO	2019	MDor19176F	F	38.8198	-88.7884	2019_TRACT5
PLATE_4	ACCACTG-TATCAGCG	5104514	NO	2019	MDor19177M	M	40.3801	-89.9300	2019_GLEASON
PLATE_4	GGTGCCA-TATCAGCG	6649760	YES	2019	MDor19178M	M	38.6685	-89.3242	2019_ELDON
PLATE_4	GCTGGA-TATCAGCG	7641578	YES	2019	MDor19180M	M	40.4296	-88.1091	2019_SPAX
PLATE_4	CACCTAGCA-TATCAGCG	5526409	YES	2019	MDor19182M	M	40.3801	-89.9300	2019_GLEASON
PLATE_4	CCTCAGCA-TATCAGCG	4064572	YES	2019	MDor19183M	M	40.3801	-89.9300	2019_GLEASON
PLATE_4	AAGATCAA-TATCAGCG	7715358	YES	2019	MDor19185M	M	38.9384	-88.6425	2019_TRACT2
PLATE_4	CGGTGGTGGA-TATCAGCG	6716629	YES	2019	MDor19186F	F	38.9384	-88.6425	2019_TRACT2
PLATE_4	TTCCTGCCA-TATCAGCG	7491722	YES	2019	MDor19187M	M	38.9735	-88.5997	2019_TRACT1
PLATE_4	AGTGGCG-TATCAGCG	6548490	YES	2019	MDor19189M	M	40.3563	-88.1397	2019_LUDLOW
PLATE_4	GAGTACG-TATCAGCG	2447223	NO	2019	MDor19190M	M	40.4287	-88.1092	2019_SPAX
PLATE_4	TTCTTGAA-TATCAGCG	5528650	YES	2019	MDor19192F	F	40.4450	-88.0978	2019_PROSPECT
PLATE_4	TTCTGACA-TATCAGCG	7900423	YES	2019	MDor19194F	F	38.8422	-88.7611	2019_TRACT4
PLATE_4	GGTGGCCA-TATCAGCG	5877529	YES	2019	MDor19195F	F	38.8422	-88.7611	2019_TRACT4
PLATE_4	GCGGTCCA-TATCAGCG	9106094	YES	2019	MDor19196M	M	38.8422	-88.7611	2019_TRACT4

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_4	GGTGGACCA-TATCAGCG	10306814	YES	2019	MDor19197M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	GGTGACACA-TATCAGCG	8907630	YES	2019	MDor19198M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	GAACAT-TATCAGCG	4677414	YES	2019	MDor19199M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	CCACCACTCG-TATCAGCG	6368917	YES	2019	MDor19200M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	ACCTCG-TATCAGCG	13146079	YES	2019	MDor19201M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	TGAACA-TATCAGCG	8219625	YES	2019	MDor19202F	F	40.3641	-88.1366	2019_LUDLOW
PLATE_4	TGCGGCACA-TATCAGCG	13303050	YES	2019	MDor19203F	F	40.3563	-88.1397	2019_LUDLOW
PLATE_4	AATCAG-TATCAGCG	9928486	YES	2019	MDor19204F	F	40.3705	-88.1337	2019_LUDLOW
PLATE_4	GAACAACAAT-TATCAGCG	10276791	YES	2019	MDor19205M	M	40.3705	-88.1337	2019_LUDLOW
PLATE_4	TTCAAGCA-TATCAGCG	6115514	YES	2019	MDor19206M	M	38.8198	-88.7884	2019_TRACT5
PLATE_4	CTTCTGA-TATCAGCG	9044900	YES	2019	MDor19207M	M	38.8198	-88.7884	2019_TRACT5
PLATE_4	AAGAACGAAT-TATCAGCG	7799100	YES	2019	MDor19208M	M	38.8198	-88.7884	2019_TRACT5
PLATE_4	TTGTACCA-TATCAGCG	550739	NO	2019	MDor19209F	F	38.8422	-88.7611	2019_TRACT4
PLATE_4	AAGATACCA-TATCAGCG	8421083	YES	2019	MDor19210M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	TGTTAACA-TATCAGCG	3129065	YES	2019	MDor19211M	M	38.8422	-88.7611	2019_TRACT4
PLATE_4	GTGGCCGCA-TATCAGCG	4085546	YES	2019	MDor19212F	F	38.8422	-88.7611	2019_TRACT4
PLATE_4	CCAACTGA-TATCAGCG	8231823	YES	2019	MDor19213M	M	40.5502	-88.0601	2019_MAINLINE
PLATE_4	GGTGTCGGTA-TATCAGCG	15420795	NO	2019	MDor19216M	M	38.9735	-88.5997	2019_TRACT1
PLATE_4	ACTGGTGGTT-TATCAGCG	34648592	NO	2019	MDor19217F	F	40.3641	-88.1366	2019_LUDLOW

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_4	TCTTAG-TATCAGCG	8240021	YES	2019	MDor19219F	F	40.3563	-88.1397	2019_LUDLOW
PLATE_4	TGCGGA-TATCAGCG	8957667	YES	2020	MDor2001F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	GAAGATCCA-TATCAGCG	7416553	YES	2020	MDor2002F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	CGTGGAA-TATCAGCG	8350716	YES	2020	MDor2003F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	GGCGGTAGGT-TATCAGCG	8219173	YES	2020	MDor2004F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	GAGACT-TATCAGCG	8177764	YES	2020	MDor2005F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	TTCTGCA-TATCAGCG	10835048	YES	2020	MDor2006F	F	38.9384	-88.6425	2020_TRACT2
PLATE_4	CCACCATCAG-TATCAGCG	3407209	YES	2020	MDor2008M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	CCTACCACAG-TATCAGCG	6486253	YES	2020	MDor2009M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	ACCACCATCG-TATCAGCG	8527463	YES	2020	MDor2010M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	GAGGTCA-TATCAGCG	8753802	YES	2020	MDor2011M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	GAGTGCCA-TATCAGCG	7006719	YES	2020	MDor2012M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	TTCAGT-TATCAGCG	15467175	YES	2020	MDor2014F	F	38.9735	-88.5997	2020_TRACT1
PLATE_4	AACATG-TATCAGCG	4625930	NO	2020	MDor2015M	M	38.9735	-88.5997	2020_TRACT1
PLATE_4	AGAACTCA-TATCAGCG	7745559	YES	2020	MDor2016M	M	38.9735	-88.5997	2020_TRACT1
PLATE_4	AACCGT-TATCAGCG	4147488	NO	2020	MDor2017M	M	38.9735	-88.5997	2020_TRACT1
PLATE_4	ACACCACCTG-TATCAGCG	4027329	NO	2020	MDor2018F	F	40.5274	-88.0761	2020_LODA
PLATE_4	CCATCCGCA-TATCAGCG	3571364	NO	2020	MDor2019F	F	40.5274	-88.0761	2020_LODA
PLATE_4	AAGCTCG-TATCAGCG	2936597	NO	2020	MDor2020F	F	40.5274	-88.0761	2020_LODA

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_4	CGTGAA-TATCAGCG	2352469	YES	2020	MDor2021F	F	40.5274	-88.0761	2020_LODA
PLATE_4	CCACGT-TATCAGCG	4382954	YES	2020	MDor2022F	F	40.5274	-88.0761	2020_LODA
PLATE_4	AAGAATCA-TATCAGCG	7896109	YES	2020	MDor2023M	M	40.5274	-88.0761	2020_LODA
PLATE_4	GAGCATA-TATCAGCG	8864861	YES	2020	MDor2024M	M	40.5274	-88.0761	2020_LODA
PLATE_4	TGAGGCA-TATCAGCG	10001198	YES	2020	MDor2025M	M	40.5274	-88.0761	2020_LODA
PLATE_4	CCACTGACA-TATCAGCG	11434004	YES	2020	MDor2026M	M	40.5274	-88.0761	2020_LODA
PLATE_4	CTGTAT-TATCAGCG	2329451	NO	2020	MDor2027M	M	40.5274	-88.0761	2020_LODA
PLATE_4	GAGGCTG-TATCAGCG	3237165	NO	2020	MDor2028M	M	40.5274	-88.0761	2020_LODA
PLATE_4	ACACTAG-TATCAGCG	4178010	NO	2020	MDor2031M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTCTCAG-TATCAGCG	4315046	NO	2020	MDor2032F	F	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTCTTGTTCA-TATCAGCG	20035985	YES	2020	MDor2033M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	GAACAAGATG-TATCAGCG	5978874	YES	2020	MDor2034M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	GAGAACAACT-TATCAGCG	3365333	NO	2020	MDor2035M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTGTGTTCGA-TATCAGCG	3017139	NO	2020	MDor2037M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	AACCTGCA-TATCAGCG	11655329	YES	2020	MDor2038M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	GGATCT-TATCAGCG	9978835	YES	2020	MDor2039M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	CCTACAGCA-TATCAGCG	6808973	YES	2020	MDor2040M	M	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTGTTGTCTA-TATCAGCG	6470894	YES	2020	MDor2041F	F	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTCACGA-TATCAGCG	8617527	YES	2020	MDor2042F	F	40.4450	-88.0978	2020_PROSPECT

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_4	AGGTGGACA-TATCAGCG	6834625	YES	2020	MDor2043F	F	40.4450	-88.0978	2020_PROSPECT
PLATE_4	TTCCATGCA-TATCAGCG	6024245	YES	2020	MDor2045F	F	38.9735	-88.5997	2020_TRACT1
PLATE_4	AACAAGAACT-TATCAGCG	6342637	YES	2020	MDor2046M	M	38.9384	-88.6425	2020_TRACT2
PLATE_4	CACAGTCA-TATCAGCG	7344323	YES	2020	MDor2048M	M	40.4287	-88.1092	2020_SPAX
PLATE_4	CCATAAG-TATCAGCG	1561920	NO	2020	MDor2049F	F	40.4287	-88.1092	2020_SPAX
PLATE_4	CTGGTA-TATCAGCG	331689	NO	2020	MDor2054M	M	40.4287	-88.1092	2020_SPAX
PLATE_4	AAGAATGCA-TATCAGCG	3636419	YES	2020	MDor2055M	M	40.4287	-88.1092	2020_SPAX
PLATE_4	ACACTG-TATCAGCG	934700	NO	2020	MDor2056M	M	40.4287	-88.1092	2020_SPAX
PLATE_4	TTCTCGACA-TATCAGCG	7838056	YES	2020	MDor2057M	M	40.4287	-88.1092	2020_SPAX
PLATE_4	GTAGGCAA-TATCAGCG	29672	NA	NA	CONTROL_PLATE_4	NA	NA	NA	NA
PLATE_5	AGAATCG-AGGTGTAC	8010842	YES	2020	MDor2058M	M	40.4287	-88.1092	2020_SPAX
PLATE_5	AGAACTACA-AGGTGTAC	9911858	YES	2020	MDor2059M	M	40.4287	-88.1092	2020_SPAX
PLATE_5	CCACATGCA-AGGTGTAC	7253895	YES	2020	MDor2060F	F	40.4287	-88.1092	2020_SPAX
PLATE_5	CCTCACG-AGGTGTAC	7796698	YES	2020	MDor2066M	M	40.4320	-88.1082	2020_SPAX
PLATE_5	AGTGGTCGGT-AGGTGTAC	7513924	YES	2020	MDor2067M	M	40.4320	-88.1082	2020_SPAX
PLATE_5	CCTCCAGA-AGGTGTAC	5204674	YES	2020	MDor2072F	F	40.3801	-89.9300	2020_GLEASON
PLATE_5	TTCTGACCA-AGGTGTAC	11799086	YES	2020	MDor2074M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	CCAATGA-AGGTGTAC	7489165	YES	2020	MDor2075M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	CTTGTTGTAA-AGGTGTAC	9304195	YES	2020	MDor2076F	F	40.3801	-89.9300	2020_GLEASON

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_5	CTTATG-AGGTGTAC	9396374	YES	2020	MDor2077M	M	38.9735	-88.5997	2020_TRACT1
PLATE_5	ACCACTG-AGGTGTAC	10641459	YES	2020	MDor2078F	F	38.9735	-88.5997	2020_TRACT1
PLATE_5	GGTGCCA-AGGTGTAC	10354792	YES	2020	MDor2080F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	GCTGGA-AGGTGTAC	14046464	YES	2020	MDor2081M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	CACCTAGCA-AGGTGTAC	11977413	YES	2020	MDor2083M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	CCTCAGCA-AGGTGTAC	13056728	YES	2020	MDor2084M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	AAGATCAA-AGGTGTAC	13655429	YES	2020	MDor2085M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	CGGTGGTGGA-AGGTGTAC	7007370	YES	2020	MDor2086M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	TTCCTGCCA-AGGTGTAC	11396821	YES	2020	MDor2087M	M	38.8701	-88.7267	2020_TRACT3
PLATE_5	AGTGGCG-AGGTGTAC	10635369	YES	2020	MDor2088F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	GAGTACG-AGGTGTAC	8950590	YES	2020	MDor2089F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	TTCTTGAA-AGGTGTAC	10269180	YES	2020	MDor2090F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	TTCTGACA-AGGTGTAC	10591149	YES	2020	MDor2091F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	GGTGGCCA-AGGTGTAC	6681185	YES	2020	MDor2092F	F	38.8701	-88.7267	2020_TRACT3
PLATE_5	GCGGTCCA-AGGTGTAC	7668867	NO	2020	MDor2097M	M	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	GGTGGACCA-AGGTGTAC	11167629	YES	2020	MDor2098M	M	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	GGTGACACA-AGGTGTAC	10725771	YES	2020	MDor2099M	M	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	GAACAT-AGGTGTAC	3696778	NO	2020	MDor20100M	M	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	CCACCACTCG-AGGTGTAC	4909323	NO	2020	MDor20101M	M	40.5815	-88.0449	2020_BUCKLEY

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_5	ACCTCG-AGGTGTAC	11079108	YES	2020	MDor20102M	M	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	TGAACA-AGGTGTAC	9930289	YES	2020	MDor20103F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	TGCGGCACA-AGGTGTAC	10615182	YES	2020	MDor20104F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	AATCAG-AGGTGTAC	8699456	YES	2020	MDor20105F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	GAACAACAAT-AGGTGTAC	13089227	YES	2020	MDor20106F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	TTCAAGCA-AGGTGTAC	9353585	YES	2020	MDor20107F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	CTTCTGA-AGGTGTAC	7304734	YES	2020	MDor20108F	F	40.5815	-88.0449	2020_BUCKLEY
PLATE_5	AAGAACGAAT-AGGTGTAC	8955653	YES	2020	MDor20124F	F	40.5274	-88.0761	2020_LODA
PLATE_5	TTGTACCA-AGGTGTAC	9427913	YES	2020	MDor20130M	M	40.3712	-88.1334	2020_LUDLOW
PLATE_5	AAGATACCA-AGGTGTAC	10638552	YES	2020	MDor20131F	F	40.3712	-88.1334	2020_LUDLOW
PLATE_5	TGTTAACA-AGGTGTAC	8679202	YES	2020	MDor20134M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	GTGGCCGCA-AGGTGTAC	10465812	YES	2020	MDor20135M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	CCAACTGA-AGGTGTAC	9653383	YES	2020	MDor20136M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	GGTGTCGGTA-AGGTGTAC	8174146	YES	2020	MDor20137M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	ACTGGTGGTT-AGGTGTAC	8208996	YES	2020	MDor20138M	M	40.3801	-89.9300	2020_GLEASON
PLATE_5	TCTTAG-AGGTGTAC	10193850	YES	2020	MDor20139F	F	40.3801	-89.9300	2020_GLEASON
PLATE_5	TGCGGA-AGGTGTAC	12380559	YES	2020	MDor20140F	F	40.3801	-89.9300	2020_GLEASON
PLATE_5	GAAGATCCA-AGGTGTAC	8571603	YES	2020	MDor20141F	F	40.3801	-89.9300	2020_GLEASON
PLATE_5	CGTGGAA-AGGTGTAC	10373218	YES	2020	MDor20152M	M	38.8422	-88.7611	2020_TRACT4

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_5	GGCGGTAGGT-AGGTGTAC	8095178	YES	2020	MDor20153M	M	38.8422	-88.7611	2020_TRACT4
PLATE_5	GAGACT-AGGTGTAC	10071624	YES	2020	MDor20154M	M	38.8422	-88.7611	2020_TRACT4
PLATE_5	TTCTGCA-AGGTGTAC	14709062	YES	2020	MDor20155M	M	38.8422	-88.7611	2020_TRACT4
PLATE_5	CCACCATCAG-AGGTGTAC	5213150	YES	2020	MDor20156M	M	38.8422	-88.7611	2020_TRACT4
PLATE_5	CCTACCACAG-AGGTGTAC	13189050	YES	2020	MDor20157M	M	38.8422	-88.7611	2020_TRACT4
PLATE_5	ACCACCATCG-AGGTGTAC	9173024	YES	2020	MDor20158F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	GAGGTCA-AGGTGTAC	11497214	YES	2020	MDor20159F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	GAGTGCCA-AGGTGTAC	7886014	YES	2020	MDor20160F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	TTCAGT-AGGTGTAC	8654015	YES	2020	MDor20161F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	AACATG-AGGTGTAC	8428662	YES	2020	MDor20162F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	AGAACTCA-AGGTGTAC	8033779	YES	2020	MDor20163M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	AACCGT-AGGTGTAC	11652333	YES	2020	MDor20164M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	ACACCACCTG-AGGTGTAC	8660962	YES	2020	MDor20165M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	CCATCCGCA-AGGTGTAC	10947493	YES	2020	MDor20166M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	AAGCTCG-AGGTGTAC	8528343	YES	2020	MDor20167M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	CGTGAA-AGGTGTAC	2918666	YES	2020	MDor20168M	F	38.8198	-88.7884	2020_TRACT5
PLATE_5	CCACGT-AGGTGTAC	15007199	YES	2020	MDor20169F	F	38.8198	-88.7884	2020_TRACT5
PLATE_5	AAGAATCA-AGGTGTAC	8979230	YES	2020	MDor20170F	F	38.8198	-88.7884	2020_TRACT5
PLATE_5	GAGCATA-AGGTGTAC	7934084	YES	2020	MDor20171F	F	38.8198	-88.7884	2020_TRACT5

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_5	TGAGGCA-AGGTGTAC	6365392	YES	2020	MDor20172F	F	38.8198	-88.7884	2020_TRACT5
PLATE_5	CCACTGACA-AGGTGTAC	8981730	YES	2020	MDor20173F	F	38.8198	-88.7884	2020_TRACT5
PLATE_5	CTGTAT-AGGTGTAC	5956141	YES	2020	MDor20180M	M	38.6685	-89.3242	2020_ELDON
PLATE_5	GAGGCTG-AGGTGTAC	8196497	YES	2020	MDor20181M	M	38.6685	-89.3242	2020_ELDON
PLATE_5	ACACTAG-AGGTGTAC	6705665	YES	2020	MDor20182M	M	38.6685	-89.3242	2020_ELDON
PLATE_5	TTCTCAG-AGGTGTAC	4505653	YES	2020	MDor20183M	M	38.6685	-89.3242	2020_ELDON
PLATE_5	TTCTTGTTCA-AGGTGTAC	8631333	YES	2020	MDor20184M	M	38.6685	-89.3242	2020_ELDON
PLATE_5	GAACAAGATG-AGGTGTAC	10989551	YES	2020	MDor20185F	F	38.6685	-89.3242	2020_ELDON
PLATE_5	GAGAACAACT-AGGTGTAC	6808145	YES	2020	MDor20186F	F	38.6685	-89.3242	2020_ELDON
PLATE_5	TTGTGTTCGA-AGGTGTAC	6978081	YES	2020	MDor20187F	F	38.6685	-89.3242	2020_ELDON
PLATE_5	AACCTGCA-AGGTGTAC	5128619	YES	2020	MDor20188F	F	38.6685	-89.3242	2020_ELDON
PLATE_5	GGATCT-AGGTGTAC	6497506	YES	2020	MDor20189F	F	38.6685	-89.3242	2020_ELDON
PLATE_5	CCTACAGCA-AGGTGTAC	6112271	YES	2020	MDor20191M	M	38.8198	-88.7884	2020_TRACT5
PLATE_5	TTGTTGTCTA-AGGTGTAC	4179702	YES	2020	MDor20192F	F	38.8422	-88.7611	2020_TRACT4
PLATE_5	TTCACGA-AGGTGTAC	12642825	YES	2020	MDor20193F	F	40.4320	-88.1082	2020_SPAX
PLATE_5	AGGTGGACA-AGGTGTAC	6308767	YES	2020	MDor20198M	M	38.9735	-88.5997	2020_TRACT1
PLATE_5	TTCCATGCA-AGGTGTAC	9838073	YES	2020	MDor20199F	F	38.9735	-88.5997	2020_TRACT1
PLATE_5	AACAAGAACT-AGGTGTAC	6781253	YES	2020	MDor20202M	M	40.3712	-88.1334	2020_LUDLOW
PLATE_5	CACAGTCA-AGGTGTAC	11868406	YES	2020	MDor20203F	F	40.3712	-88.1334	2020_LUDLOW

Plate_ number	Barcode-Index	Retained_Reads	Final analysis	Year	Specimen_ID	Sex	Latitude	Longitude	Site_Code
PLATE_5	CCATAAG-AGGTGTAC	8805276	YES	2020	MDor20204M	M	40.2259	-90.0516	2020_LONGBRANCH
PLATE_5	CTGGTA-AGGTGTAC	11158090	YES	2020	MDor20206M	M	38.6691	-89.3258	2020_ELDON
PLATE_5	AAGAATGCA-AGGTGTAC	12128919	YES	2020	MDor20207F	F	38.6691	-89.3258	2020_ELDON
PLATE_5	ACACTG-AGGTGTAC	18058034	YES	2020	MDor20208M	M	40.4611	-87.9233	2020_PELLS
PLATE_5	TTCTCGACA-AGGTGTAC	7893571	YES	2020	MDor20209M	M	40.4611	-87.9233	2020_PELLS
PLATE_5	GTAGGCAA-AGGTGTAC	33568	NA	NA	CONTROL_PLATE_5	NA	NA	NA	NA

Table B.2. Calculations used to determine size range of fragments if using PstI and MspI enzymes. Fragmatic estimates ("Predicted Fragments") were based on the most closely related genome available (*Magicicada septendecim*). *Ma. septendecula* was also ran through the fragmatic pipeline but returned fewer fragments for each size category. Genome size factor represents the approximated expected difference between *Megatibicen dorsatus* and *Ma. septendecim*. The paired-end X 2 column is the expected number of reads based on the previous columns. This allows for determining what size flow cell is needed for Illumina sequencing. The final size ranges used are indicated in green.

Min Fragment Size	Max Fragment Size	Predicted Fragments	Genome Size Factor x 3.56	Number of Individuals x 452	Desired Coverage 30x	Paired-end x 2	Size range Pre-PCR	Size range Post-PCR
250	300	16,039	57,099	7,249,628	217,488,840	434,977,680		
250	350	30,325	107,957	13,706,900	411,207,000	822,414,000		
300	400	27,751	98,794	12,543,452	376,303,560	752,607,120		
300	500	50,894	181,183	23,004,088	690,122,640	1,380,245,280		
300	550	61,185	217,819	27,655,620	829,668,600	1,659,337,200		
300	600	70,552	251,165	31,889,504	956,685,120	1,913,370,240	**370-700**	**370-835**

Table B.3. Adapters that attach to PstI and MspI cut sites and identify samples based on barcode found in second column.

Well location	Adapter Name	Adapter Sequence (5' to 3')	Adapter Sequence 2 (5' to 3')
A1	PstI-01-AGAATCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAATCGTGCA	CGATTCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A2	PstI-02-AGAACTACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAACTACATGCA	TGTAGTTCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A3	PstI-03-CCACATGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACATGCATGCA	TGCATGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A4	PstI-04-CCTCACG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTCACGTGCA	CGTGAGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A5	PstI-05-AGTGGTCGGT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTGGTCGGTTGCA	ACCGACCACTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A6	PstI-06-CCTCCAGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTCCAGATGCA	TCTGGAGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A7	PstI-07-TTCTGACCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTGACCATGCA	TGGTCAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A8	PstI-08-CCAATGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAATGATGCA	TCATTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A9	PstI-09-CTTGTTGTAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTTGTAATGCA	TTACAACAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A10	PstI-10-CTTATG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTATGTGCA	CATAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A11	PstI-11-ACCACTG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACCACTGTGCA	CAGTGGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
A12	PstI-12-GGTGCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGCCATGCA	TGGCACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B1	PstI-13-GCTGGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGGATGCA	TCCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B2	PstI-14-CACCTAGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCACCTAGCATGCA	TGCTAGGTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5’ to 3’)	Adapter Sequence 2 (5’ to 3’)
B3	PstI-15-CCTCAGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTCAGCATGCA	TGCTGAGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B4	PstI-16-AAGATCAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGATCAATGCA	TTGATCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B5	PstI-17-CGGTGGTGGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTGGTGGATGCA	TCCACCACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B6	PstI-18-TTCCTGCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCCTGCCATGCA	TGGCAGGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B7	PstI-19-AGTGGCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTGGCGTGCA	CGCCACTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B8	PstI-20-GAGTACG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTACGTGCA	CGTACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B9	PstI-21-TTCTTGAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTTGAATGCA	TTCAAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B10	PstI-22-TTCTGACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTGACATGCA	TGTCAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B11	PstI-23-GGTGGCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGGCCATGCA	TGGCCACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
B12	PstI-24-GCGGTCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCGGTCCATGCA	TGGACCGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C1	PstI-25-GGTGGACCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGGACCATGCA	TGGTCCACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C2	PstI-26-GGTGACACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGACACATGCA	TGTGTCACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C3	PstI-27-GAACAT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAACATTGCA	ATGTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C4	PstI-28-CCACCACTCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACCACTCGTGCA	CGAGTGGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C5	PstI-29-ACCTCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACCTCGTGCA	CGAGGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5' to 3')	Adapter Sequence 2 (5' to 3')
C6	PstI-30-TGAACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGAACATGCA	TGTTCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C7	PstI-31-TGCGGCACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCGGCACATGCA	TGTGCCGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C8	PstI-32-AATCAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAATCAGTGCA	CTGATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C9	PstI-33-GAACAACAAT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAACAACAATTGCA	ATTGTTGTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C10	PstI-34-TTCAAGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCAAGCATGCA	TGCTTGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C11	PstI-35-CTTCTGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTCTGATGCA	TCAGAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
C12	PstI-36-AAGAACGAAT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGAACGAATTGCA	ATTCGTTCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D1	PstI-37-TTGTACCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTACCATGCA	TGGTACAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D2	PstI-38-AAGATACCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGATACCATGCA	TGGTATCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D3	PstI-39-TGTTAACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGTTAACATGCA	TGTTAACAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D4	PstI-40-GTGGCCGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGGCCGCATGCA	TGCGGCCACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D5	PstI-41-CCAACTGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAACTGATGCA	TCAGTTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D6	PstI-42-GGTGTCGGTA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGTCGGTATGCA	TACCGACACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D7	PstI-43-ACTGGTGGTT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGGTGGTTTGCA	AACCACCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5’ to 3’)	Adapter Sequence 2 (5’ to 3’)
D8	PstI-44-TCTTAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTTAGTGCA	CTAAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D9	PstI-45-TGCGGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCGGATGCA	TCCGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D10	PstI-46-GAAGATCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGATCCATGCA	TGGATCTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D11	PstI-47-CGTGGAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGGAATGCA	TTCCACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
D12	PstI-48-GGCGGTAGGT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCGGTAGGTTGCA	ACCTACCGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E1	PstI-49-GAGACT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGACTTGCA	AGTCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E2	PstI-50-TTCTGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTGCATGCA	TGCAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E3	PstI-51-CCACCATCAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACCATCAGTGCA	CTGATGGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E4	PstI-52-CCTACCACAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTACCACAGTGCA	CTGTGGTAGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E5	PstI-53-ACCACCATCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACCACCATCGTGCA	CGATGGTGGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E6	PstI-54-GAGGTCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGGTCATGCA	TGACCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E7	PstI-55-GAGTGCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTGCCATGCA	TGGCACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E8	PstI-56-TTCAGT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCAGTTGCA	ACTGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E9	PstI-57-AACATG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAACATGTGCA	CATGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E10	PstI-58-AGAACTCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAACTCATGCA	TGAGTTCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5’ to 3’)	Adapter Sequence 2 (5’ to 3’)
E11	PstI-59-AACCGT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCGTTGCA	ACGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
E12	PstI-60-ACACCACCTG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACACCACCTGTGCA	CAGGTGGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F1	PstI-61-CCATCCGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCATCCGCATGCA	TGCGGATGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F2	PstI-62-AAGCTCG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGCTCGTGCA	CGAGCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F3	PstI-63-CGTGAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGAATGCA	TTCACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F4	PstI-64-CCACGT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACGTTGCA	ACGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F5	PstI-65-AAGAATCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGAATCATGCA	TGATTCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F6	PstI-66-GAGCATA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGCATATGCA	TATGCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F7	PstI-67-TGAGGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGAGGCATGCA	TGCCTCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F8	PstI-68-CCACTGACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACTGACATGCA	TGTCAGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F9	PstI-69-CTGTAT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTATTGCA	ATACAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F10	PstI-70-GAGGCTG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGGCTGTGCA	CAGCCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F11	PstI-71-ACACTAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACACTAGTGCA	CTAGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
F12	PstI-72-TTCTCAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTCAGTGCA	CTGAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G1	PstI-73-TTCTTGTTCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTTGTTCATGCA	TGAACAAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5’ to 3’)	Adapter Sequence 2 (5’ to 3’)
G2	PstI-74-GAACAAGATG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAACAAGATGTGCA	CATCTTGTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G3	PstI-75-GAGAACAACT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGAACAACTTGCA	AGTTGTTCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G4	PstI-76-TTGTGTTCGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGTTCGATGCA	TCGAACACAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G5	PstI-77-AACCTGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCTGCATGCA	TGCAGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G6	PstI-78-GGATCT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGATCTTGCA	AGATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G7	PstI-79-CCTACAGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTACAGCATGCA	TGCTGTAGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G8	PstI-80-TTGTTGTCTA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTTGTCTATGCA	TAGACAACAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G9	PstI-81-TTCACGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCACGATGCA	TCGTGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G10	PstI-82-AGGTGGACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGGTGGACATGCA	TGTCCACCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G11	PstI-83-TTCCATGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCCATGCATGCA	TGCATGGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
G12	PstI-84-AACAAGAACT	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAACAAGAACTTGCA	AGTTCTTGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H1	PstI-85-CACAGTCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCACAGTCATGCA	TGACTGTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H2	PstI-86-CCATAAG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCATAAGTGCA	CTTATGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H3	PstI-87-CTGGTA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGGTATGCA	TACCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC

Well location	Adapter Name	Adapter Sequence (5’ to 3’)	Adapter Sequence 2 (5’ to 3’)
H4	PstI-88-AAGAATGCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGAATGCATGCA	TGCATTCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H5	PstI-89-ACACTG	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACACTGTGCA	CAGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H6	PstI-90-TTCTCGACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTCGACATGCA	TGTCGAGAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H7	PstI-91-GTAGGCAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGGCAATGCA	TTGCCTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H8	PstI-92-TTGTTGCTGA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTTGCTGATGCA	TCAGCAACAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H9	PstI-93-TTGTCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTCATGCA	TGACAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H10	PstI-94-AGACTACA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGACTACATGCA	TGTAGTCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H11	PstI-95-TCTTGGCCA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTTGGCCATGCA	TGGCCAAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
H12	PstI-96-GGTGCGAA	GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTGCGAATGCA	TTCGCACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATC
NA	MspI-A	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	CGAGATCGGAAGAGCACTTTCTCC
NA	MspI-B	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTA	CGTAGATCGGAAGAGCACTTTCTCC
NA	MspI-C	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTA	CGTAAGATCGGAAGAGCACTTTCTCC
NA	MspI-D	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTA	CGTACAGATCGGAAGAGCACTTTCTCC

Table B.4. Indexing primers (forward and reverse) and corresponding plate number used to amplify the library.

Name	Plate #	Index	Sequence (5' to 3')
RADseq-Primer4R	1	TATCAGCG	CAAGCAGAAGACGGCATACGAGATCGCTGATAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
RADseq-Primer4F	1	TATCAGCG	AATGATACGGCGACCACCGAGATCTACACCGCTGATAACACTCTTTCCCTACACGACGCTCTTCCGATCT
RADseq-Primer5R	2	AGGTGTAC	CAAGCAGAAGACGGCATACGAGATGTACACCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
RADseq-Primer5F	2	AGGTGTAC	AATGATACGGCGACCACCGAGATCTACACGTACACCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
RADseq-Primer6R	3	GACCTAAC	AATGATACGGCGACCACCGAGATCTACACGTTAGGTCACACTCTTTCCCTACACGACGCTCTTCCGATCT
RADseq-Primer6F	3	GACCTAAC	CAAGCAGAAGACGGCATACGAGATGTTAGGTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
RADseq-Primer7R	4	ACCAACTG	AATGATACGGCGACCACCGAGATCTACACCAGTTGGTACACTCTTTCCCTACACGACGCTCTTCCGATCT
RADseq-Primer7F	4	ACCAACTG	CAAGCAGAAGACGGCATACGAGATCAGTTGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
RADseq-Primer8R	5	TTCGCTGA	AATGATACGGCGACCACCGAGATCTACACTCAGCGAAACACTCTTTCCCTACACGACGCTCTTCCGATCT
RADseq-Primer8F	5	TTCGCTGA	CAAGCAGAAGACGGCATACGAGATTCAGCGAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC

Chapter 2 code

```
#### cutadapt ##### https://cutadapt.readthedocs.io/

     # use to trim 5' end of paired-end reads to remove sequence
     # before CGG cutsite
     # works sequentially -- run each line in order for each
     # plate R2 file

     module load cutadapt/2.10-IGB-gcc-8.2.0-Python-3.7.2

     # Need this module – pigz - loaded if you are using gz
     # files and multiple cores

     module load pigz/2.4-IGB-gcc-8.2.0

     # Your files need to be uploaded to the server as .gz files
     # (nice and compressed), can be in the same folder for now

     # Adding cores does NOT make this faster, I think there is
     # a bottleneck involved with .gz files.

     # -g indicates this is on the 5' end (i.e. XADAPTER) vs -a
     # would be on the 3' end (i.e. ADAPTERX)

     # Each of the following lines should be run in succession.
     # Took ~6 hours for each line for the Megatibicen data.
     # Make sure to only run the cutadapt on R2 (MspI side) not
     # R1 (PstI side)
     # If you want to see what your sequences looks like prior
     # to and after running cutadapt you can use:

     less filename.fastq.gz

     # THEN type q to exit preview
     # Please see notes after cutadapt code below
```

```
     # Plate_1

     cutadapt -g XGTA --no-indels -e 0 --cores=8 -o
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXGTA.fastq.gz
MD_PLATE_1_GACCTAAC_L002_R2_001.fastq.gz --overlap 3 >
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXGTA.fastq.gz.log

     cutadapt -g XTA --no-indels -e 0 --cores=8 -o
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXTA.fastq.gz
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXGTA.fastq.gz --overlap 2 >
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXTA.fastq.gz.log

     cutadapt -g XA --no-indels -e 0 --cores=8 -o
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXA.fastq.gz
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXTA.fastq.gz --overlap 1 >
MD_PLATE_1_GACCTAAC_L002_R2_001_CleanXA.fastq.gz.log


     # Plate_2

     cutadapt -g XGTA --no-indels -e 0 --cores=8 -o
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXGTA.fastq.gz
MD_PLATE_2_ACCAACTG_L002_R2_001.fastq.gz --overlap 3 >
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXGTA.fastq.gz.log

     cutadapt -g XTA --no-indels -e 0 --cores=8 -o
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXTA.fastq.gz
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXGTA.fastq.gz --overlap 2 >
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXTA.fastq.gz.log

     cutadapt -g XA --no-indels -e 0 --cores=8 -o
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXA.fastq.gz
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXTA.fastq.gz --overlap 1 >
MD_PLATE_2_ACCAACTG_L002_R2_001_CleanXA.fastq.gz.log


     # Plate_3

     cutadapt -g XGTA --no-indels -e 0 --cores=8 -o
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXGTA.fastq.gz
MD_PLATE_3_TTCGCTGA_L002_R2_001.fastq.gz --overlap 3 >
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXGTA.fastq.gz.log

     cutadapt -g XTA --no-indels -e 0 --cores=8 -o
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXTA.fastq.gz
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXGTA.fastq.gz --overlap 2 >
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXTA.fastq.gz.log
```

```
     cutadapt -g XA --no-indels -e 0 --cores=8 -o
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXA.fastq.gz
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXTA.fastq.gz --overlap 1 >
MD_PLATE_3_TTCGCTGA_L002_R2_001_CleanXA.fastq.gz.log


     # Plate_4
     cutadapt -g XGTA --no-indels -e 0 --cores=8 -o
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXGTA.fastq.gz
MD_PLATE_4_TATCAGCG_L002_R2_001.fastq.gz --overlap 3 >
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXGTA.fastq.gz.log

     cutadapt -g XTA --no-indels -e 0 --cores=8 -o
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXTA.fastq.gz
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXGTA.fastq.gz --overlap 2 >
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXTA.fastq.gz.log

     cutadapt -g XA --no-indels -e 0 --cores=8 -o
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXA.fastq.gz
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXTA.fastq.gz --overlap 1 >
MD_PLATE_4_TATCAGCG_L002_R2_001_CleanXA.fastq.gz.log


     # Plate_5
     cutadapt -g XGTA --no-indels -e 0 --cores=8 -o
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXGTA.fastq.gz
MD_PLATE_5_AGGTGTAC_L002_R2_001.fastq.gz --overlap 3 >
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXGTA.fastq.gz.log

     cutadapt -g XTA --no-indels -e 0 --cores=8 -o
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXTA.fastq.gz
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXGTA.fastq.gz --overlap 2 >
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXTA.fastq.gz.log

     cutadapt -g XA --no-indels -e 0 --cores=8 -o
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXA.fastq.gz
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXTA.fastq.gz --overlap 1 >
MD_PLATE_5_AGGTGTAC_L002_R2_001_CleanXA.fastq.gz.log



     # The final files you want to make sure you keep are the
     # files called *_CleanXA.fastq.gz
     # However, for the next steps you will need to remove
     # _CleanXA from the file name so that Stacks is able to
     # associate files together
```

```
# Stacks code for after cutadapt
     # First step -- process_radtags - clean, rescue, and
demultiplex
     # by adapter

     # Stacks Manual Category 4 -- If your data are paired-end
     with an
     # inline barcode on the single-end (in red) and an index
     barcode
     # (in blue):
     # Then specify the --inline_index flag to process_radtags.

#  Example: process_radtags -p ./raw/ -o ./samples/ -b
./barcodes/barcodes_lane3 -e sbfI -r -c -q

# If your data are paired-end, Illumina HiSeq data, in a
directory called raw:
     # then you simply add the -P flag.
# process_radtags understands the Illumina naming scheme and
will figure out how to properly pair the files together.

# Need to rename the cutadapt files so that stacks can find
them, otherwise it isnt able to pair R2 with R1.
     # As mentioned in the cutadapt notes, make sure to remove
_CleanXA from file name
     # This is only relevant for the R2 reads, R1 can be used as
is.

# Create a raw folder within your folder on the cluster.
# Within the raw folder you will need a folder for each "Plate"
with its own index (two files -- R1 and R2)
     # I named my folders P1 thru P5

# Create a barcode folder with files for each index/plate
     # The order of the columns for the barcode file should be:
ADAPTER (tab) INDEX (tab) SAMPLE_ID
     #    AGAATCG   AGGTGTAC  MDor2058M
          AGAACTACA AGGTGTAC  MDor2059M
          CCACATGCA AGGTGTAC  MDor2060F
# Documentation for process_radtags:
https://catchenlab.life.illinois.edu/stacks/comp/process_radtags
.php

module load Stacks/2.54-IGB-gcc-8.2.0

# -P for paired reads
# -p for path to input (raw sequences)
```

```
# -o for path to output (Make sure to create a folder for files
for each plate so you can monitor it as the process runs)
# -b for path to barcodes
# --inline index   barcode is inline with sequence on single-end
read, occurs in FASTQ header for paired-end read
# -r rescue barcodes and RAD-Tags
# -c clean data, remove any read with an uncalled base
# -q discard reads with low quality scores
# -D to capture discarded reads

process_radtags -P -p raw/P1/ -o P1/ -b barcodes/P1 --
inline_index -r -c -q -D --renz_1 pstI --renz_2 mspI
process_radtags -P -p raw/P2/ -o P2/ -b barcodes/P2 --
inline_index -r -c -q -D --renz_1 pstI --renz_2 mspI
process_radtags -P -p raw/P3/ -o P3/ -b barcodes/P3 --
inline_index -r -c -q -D --renz_1 pstI --renz_2 mspI
process_radtags -P -p raw/P4/ -o P4/ -b barcodes/P4 --
inline_index -r -c -q -D --renz_1 pstI --renz_2 mspI
process_radtags -P -p raw/P5/ -o P5/ -b barcodes/P5 --
inline_index -r -c -q -D --renz_1 pstI --renz_2 mspI
```

Chapter 2 R code

```
# transform populations vcf file into "012" matrix with VCFTools
prior to running in R

# Load libraries
library(ggplot2)
library(stringr)
library(editData) # if edits are needed
library(tidyverse)


# Load 012 matrix files
gt_matrix <- read.table(file="populations.snps.012",
                        row.names = 1, header=FALSE)
sample_names <- read.table(file="populations.snps.012.indv",
                           header=FALSE)

# Load a dataframe with the sample and group/category info
# Can do this in R Studio from an excel file, just make sure indiv
match in order and that the groups of interest are included


# PCA plot
pca <- prcomp(gt_matrix)$x[,1:4]
pca <- cbind(pca, sample_info)
ggplot(pca, aes(x=PC1, y=PC2, color=group)) +
     theme_bw() +
     stat_ellipse() +
     geom_point()
```

```
# FST Stats and Mantel test for IBD
# load libraries
library("adegenet")
     library("vegan")
     library("vcfR")
     library("Rtools")
     library("stringr")
     library("hierfstat")
     library("dartR")
     library("vcfR")
     library("tidyverse")
     library("ggplot2")
     library("dplyr")
     library("poppr")
     library("ggpubr")
     library("ggrepel")
     library("readxl")

#get samples names from genind object
indNames <- data.frame(V1=indNames(gi))

#load vcf
vcf <- read.vcfR("populations.snps.vcf")


#load BestKpopmap_brief csv file for pop map and k = 6 info

#change loci names, code from Roberto Cucalon
#Function to change name by adding the word Loci_ at the beginning
#This will be handy for basic statistics analyses using hierfstat
package
     newLociName <- function(vcf){
       newname <- str_sub(string = vcf@fix[,3],end = -3)
       newLoci <- str_c("Locus_",newname)
       vcf@fix[,3] <- newLoci
       return(vcf)
     }

     vcf_NewLociName <- newLociName(vcf)

#Convert to genind
     gi <- vcfR2genind(vcf_NewLociName)
     gi

#assign population to the genind object
     pop(gi) <- BestKpopmap_brief$K6
     pop(gi)
```

```
#create genlight
     gl <- gi2gl(gi)
     gl

# For mantel test
# Add the spatial coordinates separate from the genotype data.
# Create the genind with the genotype data first, then import the
spatial
# coordinates into a separate data frame. Then direct those
coordinates into
# the "other" slot. Here is an example (assumes a tab-delimited
text file):


colnames(latlong)<-c("x", "y")
gi@other$xy<-latlong


mantelgenepop <- genind2genpop(gi)

Dgen <- dist.genpop(mantelgenepop,method=2)
Dgeo <- dist(other(gi))

plot(combo_FST_latlong$FST$Distance)

ibd <- mantel.randtest(Dgen,Dgeo)
ibd
plot(ibd)

dist_lm <- lm(as.vector(Dgen) ~ as.vector(Dgeo))
plot(Dgeo, Dgen)
abline(dist_lm, col="red", lty=2)
title("Isolation by distance plot")


gi_heirdf <- genind2hierfstat(gi,pop = pop(gi))
gi_heirdf[1:10,1:8]

print("--- Pairwise FST ---")
pairwise_FST <- pairwise.WCfst(gi_heirdf)
pairwise_FST[pairwise_FST < 0] <- 0

write.csv(pairwise_FST, file = "pairwise_FST_k6.csv")
```

```
#get basic stats
bs.nc <- basic.stats(gi_heirdf)
bs.nc
```

APPENDIX C: CHAPTER 3 SUPPLEMENTARY MATERIALS

Table C.1. Metadata table used with QIIME2. CALLIOPE = *Cicadettana calliope calliope*; PRUINOSUS = *Neotibicen pruinosus pruinosus*; DORSATUS = *Megatibicen dorsatus*. Pruinosus in "body part" column refers to pruinosity scraped off cicada thorax/abdomen using sterile razorblade.

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-MDE-06	Body	EARLY	DORSATUS	M	MDE	ACAGTCATAT
84-MDE-07	Forewings	EARLY	DORSATUS	M	MDE	CTACGATCAG
84-MDE-08	Hindwings	EARLY	DORSATUS	M	MDE	GCACTAGACA
84-MDE-09	Legs	EARLY	DORSATUS	M	MDE	CTAGCAGATG
84-MDE-10	Head	EARLY	DORSATUS	M	MDE	ATGTATAGTC
84-MDE-12	Forewings	EARLY	DORSATUS	M	MDE	CATGATACGC
84-MDE-13	Hindwings	EARLY	DORSATUS	M	MDE	GCAGCTGTCA
84-MDE-14	Legs	EARLY	DORSATUS	M	MDE	ACGTATCATC
84-MDE-15	Head	EARLY	DORSATUS	M	MDE	ACATATACGT
84-MDE-17	Forewings	EARLY	DORSATUS	F	MDE	AGTATCGTAC
84-MDE-18	Hindwings	EARLY	DORSATUS	F	MDE	GATACACTGA
84-MDE-19	Legs	EARLY	DORSATUS	F	MDE	GCGAGATGTA
84-MDE-20	Head	EARLY	DORSATUS	F	MDE	AGCATCTATA
84-MDE-21	Body	EARLY	DORSATUS	M	MDE	AGACTATATC
84-MDE-22	Forewings	EARLY	DORSATUS	M	MDE	GACTAGTCAG
84-MDE-23	Hindwings	EARLY	DORSATUS	M	MDE	GATGACTACG
84-MDE-24	Legs	EARLY	DORSATUS	M	MDE	CACATACAGT
84-MDE-25	Head	EARLY	DORSATUS	M	MDE	CAGCATCTAG
84-MDE-27	Forewings	EARLY	DORSATUS	M	MDE	CGAGACGACA
84-MDE-28	Hindwings	EARLY	DORSATUS	M	MDE	ATCACTCATA
84-MDE-29	Legs	EARLY	DORSATUS	M	MDE	AGCTCTGTGA
84-MDE-30	Head	EARLY	DORSATUS	M	MDE	ATGTCATGCT
84-MDE-32	Forewings	EARLY	DORSATUS	F	MDE	GCTGACAGAG
84-MDE-33	Hindwings	EARLY	DORSATUS	F	MDE	ATACAGTCTC
84-MDE-34	Legs	EARLY	DORSATUS	F	MDE	CATAGACGTG

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-MDE-35	Head	EARLY	DORSATUS	F	MDE	AGAGATATCA
84-MDE-36	Body	EARLY	DORSATUS	M	MDE	ATGCTGCGCT
84-MDE-37	Forewings	EARLY	DORSATUS	M	MDE	AGTCAGACGC
84-MDE-38	Hindwings	EARLY	DORSATUS	M	MDE	CTACATACTA
84-MDE-39	Legs	EARLY	DORSATUS	M	MDE	TACACAGTAG
84-MDE-40	Head	EARLY	DORSATUS	M	MDE	GACGATCGCA
84-MDE-42	Forewings	EARLY	DORSATUS	F	MDE	CACAGTGATG
84-MDE-43	Hindwings	EARLY	DORSATUS	F	MDE	CGAGCTAGCA
84-MDE-44	Legs	EARLY	DORSATUS	F	MDE	GAGACTATGC
84-MDE-45	Head	EARLY	DORSATUS	F	MDE	CGCAGAGCAT
84-MDE-46	Body	EARLY	DORSATUS	F	MDE	GTCGTGTACT
84-MDE-47	Forewings	EARLY	DORSATUS	F	MDE	GATGTAGCGT
84-MDE-48	Hindwings	EARLY	DORSATUS	F	MDE	GAGTGATCGT
84-MDE-49	Legs	EARLY	DORSATUS	F	MDE	CGCTATCAGT
84-MDE-50	Head	EARLY	DORSATUS	F	MDE	CGCTGTAGTC
84-MDE-51	Body	EARLY	DORSATUS	M	MDE	GCTAGTGAGT
84-MDE-52	Forewings	EARLY	DORSATUS	M	MDE	GAGCTAGTGA
84-MDE-53	Hindwings	EARLY	DORSATUS	M	MDE	CGTGCTGTCA
84-MDE-54	Legs	EARLY	DORSATUS	M	MDE	GATCGTCTCT
84-MDE-55	Head	EARLY	DORSATUS	M	MDE	GTGCTGTCGT
84-MDE-56	Body	EARLY	DORSATUS	M	MDE	TGAGCGTGCT
84-MDE-57	Forewings	EARLY	DORSATUS	M	MDE	CATGTCGTCA
84-MDE-58	Hindwings	EARLY	DORSATUS	M	MDE	TCAGTGTCTC
84-MDE-59	Legs	EARLY	DORSATUS	M	MDE	GTGCTCATGT
84-MDE-60	Head	EARLY	DORSATUS	M	MDE	CGTATCTCGA
84-MDE-61	Body	EARLY	DORSATUS	M	MDE	GTCATGCGTC
84-MDE-62	Forewings	EARLY	DORSATUS	M	MDE	CTATGCGATC
84-MDE-63	Hindwings	EARLY	DORSATUS	M	MDE	TGCTATGCTG
84-MDE-64	Legs	EARLY	DORSATUS	M	MDE	TGTGTGCATG
84-MDE-65	Head	EARLY	DORSATUS	M	MDE	GAGTGTCACT
84-MDE-66	Molt	EARLY	DORSATUS	M	MDE	CTAGTCTCGT

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-MDE-67	Honeydew	EARLY	DORSATUS	F	MDE	GAGTGCATCT
84-MDL-01	Body	LATE	DORSATUS	F	MDL	TGATACTCTG
84-MDL-02	Forewings	LATE	DORSATUS	F	MDL	CAGCTATAGC
84-MDL-03	Hindwings	LATE	DORSATUS	F	MDL	TCGATGCGCT
84-MDL-04	Legs	LATE	DORSATUS	F	MDL	TTGTTGCTGT
84-MDL-05	Head	LATE	DORSATUS	F	MDL	AGCAGTACTC
84-MDL-100	Pruinosus	LATE	DORSATUS	F	MDL	TCAGCGATAT
84-MDL-101	Pruinosus	LATE	DORSATUS	F	MDL	TGGTGCTGGA
84-MDL-46	Body	LATE	DORSATUS	F	MDL	CACGAGATGA
84-MDL-47	Forewings	LATE	DORSATUS	F	MDL	TGCTACATCA
84-MDL-48	Hindwings	LATE	DORSATUS	F	MDL	AGTGTGTCTA
84-MDL-49	Legs	LATE	DORSATUS	F	MDL	ACGCACATAT
84-MDL-50	Head	LATE	DORSATUS	F	MDL	ACGATCACAT
84-MDL-52	Forewings	LATE	DORSATUS	M	MDL	TCATATCGCG
84-MDL-53	Hindwings	LATE	DORSATUS	M	MDL	CTGCATGATC
84-MDL-54	Legs	LATE	DORSATUS	M	MDL	GTAATGGAGT
84-MDL-55	Head	LATE	DORSATUS	M	MDL	CTCGTTATTC
84-MDL-57	Forewings	LATE	DORSATUS	M	MDL	CGCGTATCAT
84-MDL-58	Hindwings	LATE	DORSATUS	M	MDL	GTATCTCTCG
84-MDL-59	Legs	LATE	DORSATUS	M	MDL	GCTCATATGC
84-MDL-60	Head	LATE	DORSATUS	M	MDL	CTAATCGTGT
84-MDL-62	Forewings	LATE	DORSATUS	M	MDL	CACTATGTCG
84-MDL-63	Hindwings	LATE	DORSATUS	M	MDL	TAGCGCGTAG
84-MDL-64	Legs	LATE	DORSATUS	M	MDL	GGAAGTAAGG
84-MDL-65	Head	LATE	DORSATUS	M	MDL	CGGTGTGTGT
84-MDL-67	Forewings	LATE	DORSATUS	M	MDL	CGTCACAGTA
84-MDL-68	Hindwings	LATE	DORSATUS	M	MDL	TCGCGTGAGA
84-MDL-69	Legs	LATE	DORSATUS	M	MDL	TGTGAATCTC
84-MDL-70	Head	LATE	DORSATUS	M	MDL	CGTCTTCTTA
84-MDL-72	Forewings	LATE	DORSATUS	M	MDL	TACATCGCTG
84-MDL-73	Hindwings	LATE	DORSATUS	M	MDL	GTGAGAGACA

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-MDL-74	Legs	LATE	DORSATUS	M	MDL	CTCTTAGTTC
84-MDL-75	Head	LATE	DORSATUS	M	MDL	GACTGTACGT
84-MDL-77	Forewings	LATE	DORSATUS	M	MDL	GCACGTAGCT
84-MDL-78	Hindwings	LATE	DORSATUS	M	MDL	TCACGCTATG
84-MDL-79	Legs	LATE	DORSATUS	M	MDL	GGATAGGATC
84-MDL-80	Head	LATE	DORSATUS	M	MDL	GTCTCAATGT
84-MDL-82	Forewings	LATE	DORSATUS	M	MDL	CGTACTACGT
84-MDL-83	Hindwings	LATE	DORSATUS	M	MDL	CAGCTGAGTA
84-MDL-84	Legs	LATE	DORSATUS	M	MDL	GGTGTCTTGT
84-MDL-85	Head	LATE	DORSATUS	M	MDL	GATGAGGTAT
84-MDL-86	Body	LATE	DORSATUS	F	MDL	GGTAGAATGA
84-MDL-87	Forewings	LATE	DORSATUS	F	MDL	GAGATCAGTC
84-MDL-88	Hindwings	LATE	DORSATUS	F	MDL	TATCATGTGC
84-MDL-89	Legs	LATE	DORSATUS	F	MDL	GATGGTTGTA
84-MDL-90	Head	LATE	DORSATUS	F	MDL	GGTGTTAGTG
84-MDL-91	Body	LATE	DORSATUS	M	MDL	TTAGTGGTGA
84-MDL-92	Forewings	LATE	DORSATUS	M	MDL	TGCGTAGTCG
84-MDL-93	Hindwings	LATE	DORSATUS	M	MDL	CTGTGTCGTC
84-MDL-94	Legs	LATE	DORSATUS	M	MDL	CATGAGTGTA
84-MDL-95	Head	LATE	DORSATUS	M	MDL	CCTCGTTGTT
84-MDL-96	Pruinosus	LATE	DORSATUS	M	MDL	CTGTAGTGCG
84-MDL-97	Pruinosus	LATE	DORSATUS	M	MDL	GTGCGCTAGT
84-MDL-98	Pruinosus	LATE	DORSATUS	M	MDL	CGTTAGCGTA
84-MDL-99	Pruinosus	LATE	DORSATUS	M	MDL	TACTAGGATC
84-NPE-01	Body	EARLY	PRUINOSUS	F	NPE	CGTCTATGAT
84-NPE-02	Molt	EARLY	PRUINOSUS	F	NPE	AGTCATCGCA
84-NPE-03	Forewings	EARLY	PRUINOSUS	F	NPE	ATGCTAGAGA
84-NPE-04	Hindwings	EARLY	PRUINOSUS	F	NPE	GCACGCGTAT
84-NPE-05	Legs	EARLY	PRUINOSUS	F	NPE	GTGATACTGA
84-NPE-06	Head	EARLY	PRUINOSUS	F	NPE	ATCGCTACAT
84-NPE-07	Body	EARLY	PRUINOSUS	F	NPE	CTAGATCTGA

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-NPE-08	Molt	EARLY	PRUINOSUS	F	NPE	TGCGAGACGT
84-NPE-09	Forewings	EARLY	PRUINOSUS	F	NPE	TGATGTATGT
84-NPE-10	Hindwings	EARLY	PRUINOSUS	F	NPE	GACTCATGCT
84-NPE-11	Legs	EARLY	PRUINOSUS	F	NPE	TCTACGACAT
84-NPE-12	Head	EARLY	PRUINOSUS	F	NPE	TATCAGTCTG
84-NPE-13	Body	EARLY	PRUINOSUS	M	NPE	AGCGAGTATG
84-NPE-14	Molt	EARLY	PRUINOSUS	M	NPE	TAGCATACAG
84-NPE-15	Forewings	EARLY	PRUINOSUS	M	NPE	GATATATGTC
84-NPE-16	Hindwings	EARLY	PRUINOSUS	M	NPE	CTCAGCAGTG
84-NPE-17	Legs	EARLY	PRUINOSUS	M	NPE	ACGATACACT
84-NPE-18	Head	EARLY	PRUINOSUS	M	NPE	TAGTACTAGA
84-NPE-19	Body	EARLY	PRUINOSUS	F	NPE	TATAGAGATC
84-NPE-20	Molt	EARLY	PRUINOSUS	F	NPE	CAGTCTACAT
84-NPE-21	Forewings	EARLY	PRUINOSUS	F	NPE	TCGATATCTA
84-NPE-22	Hindwings	EARLY	PRUINOSUS	F	NPE	TACTGCAGCG
84-NPE-23	Legs	EARLY	PRUINOSUS	F	NPE	TGAGATCATA
84-NPE-24	Head	EARLY	PRUINOSUS	F	NPE	TCAGATGCTA
84-NPE-25	Body	EARLY	PRUINOSUS	F	NPE	TACATGATAG
84-NPE-26	Molt	EARLY	PRUINOSUS	F	NPE	CTGATGCAGA
84-NPE-27	Forewings	EARLY	PRUINOSUS	F	NPE	GTGACGTACG
84-NPE-28	Hindwings	EARLY	PRUINOSUS	F	NPE	CGACGCTGAT
84-NPE-29	Legs	EARLY	PRUINOSUS	F	NPE	AGTAGATCAT
84-NPE-30	Head	EARLY	PRUINOSUS	F	NPE	ACATAGTATC
84-NPE-47	Molt	EARLY	PRUINOSUS	M	NPE	GGTTGGAGTT
84-NPE-48	Forewings	EARLY	PRUINOSUS	M	NPE	TACTGAGCTG
84-NPE-49	Hindwings	EARLY	PRUINOSUS	M	NPE	TAGTAGCGCG
84-NPE-50	Legs	EARLY	PRUINOSUS	M	NPE	GACGTCTGCT
84-NPE-51	Head	EARLY	PRUINOSUS	M	NPE	CATTCTCTGA
84-NPE-53	Molt	EARLY	PRUINOSUS	F	NPE	GTACTCGCGA
84-NPE-54	Forewings	EARLY	PRUINOSUS	F	NPE	TCTGAGCGCA
84-NPE-55	Hindwings	EARLY	PRUINOSUS	F	NPE	TAGACGTGCT

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-NPE-56	Legs	EARLY	PRUINOSUS	F	NPE	GTGACTCGTC
84-NPE-57	Head	EARLY	PRUINOSUS	F	NPE	CATCTGGAGT
84-NPE-59	Molt	EARLY	PRUINOSUS	F	NPE	TGTCGTCATA
84-NPE-60	Forewings	EARLY	PRUINOSUS	F	NPE	TATCTCATGC
84-NPE-61	Hindwings	EARLY	PRUINOSUS	F	NPE	TGTGTCACTA
84-NPE-62	Legs	EARLY	PRUINOSUS	F	NPE	TATCGATGCT
84-NPE-63	Head	EARLY	PRUINOSUS	F	NPE	GAATGGAAGA
84-NPE-65	Molt	EARLY	PRUINOSUS	F	NPE	TAGAGTCTGT
84-NPE-66	Forewings	EARLY	PRUINOSUS	F	NPE	CATGCATCAT
84-NPE-67	Hindwings	EARLY	PRUINOSUS	F	NPE	TGATCAGTCA
84-NPE-68	Legs	EARLY	PRUINOSUS	F	NPE	TCGAGTAGCG
84-NPE-69	Head	EARLY	PRUINOSUS	F	NPE	GGCTGTGATC
84-NPE-70	Body	EARLY	PRUINOSUS	M	NPE	GTATCGTCGT
84-NPE-71	Forewings	EARLY	PRUINOSUS	M	NPE	TGTGCTCGCA
84-NPE-72	Hindwings	EARLY	PRUINOSUS	M	NPE	GATGCGAGCT
84-NPE-73	Legs	EARLY	PRUINOSUS	M	NPE	CTGTACGTGA
84-NPE-74	Head	EARLY	PRUINOSUS	M	NPE	GCGATGATGA
84-NPE-75	Molt	EARLY	PRUINOSUS	M	NPE	TGTCGAGTCA
84-NPL-32	Forewings	LATE	PRUINOSUS	F	NPL	GTGTGGTTGT
84-NPL-33	Hindwings	LATE	PRUINOSUS	F	NPL	TAGGTGGAAT
84-NPL-34	Legs	LATE	PRUINOSUS	F	NPL	TCATCATGCG
84-NPL-35	Head	LATE	PRUINOSUS	F	NPL	TGTAGGTGGA
84-NPL-36	Body	LATE	PRUINOSUS	F	NPL	TATGGTAAGG
84-NPL-37	Forewings	LATE	PRUINOSUS	F	NPL	GTGAAGGTAA
84-NPL-38	Hindwings	LATE	PRUINOSUS	F	NPL	ACTGCGTGTC
84-NPL-39	Legs	LATE	PRUINOSUS	F	NPL	GTTGATGAGT
84-NPL-40	Head	LATE	PRUINOSUS	F	NPL	TGTTGTGGTA
84-NPL-41	Body	LATE	PRUINOSUS	F	NPL	GTTCGATTGT
84-NPL-43	Hindwings	LATE	PRUINOSUS	F	NPL	TACGTATAGC
84-NPL-44	Legs	LATE	PRUINOSUS	F	NPL	GGTCAGTGTA
84-NPL-45	Head	LATE	PRUINOSUS	F	NPL	ATCTAGATCA

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
84-NPL-46	Body	LATE	PRUINOSUS	F	NPL	TGGTGTCCGT
84-NPL-47	Forewings	LATE	PRUINOSUS	F	NPL	GTCTACTGTC
84-NPL-48	Hindwings	LATE	PRUINOSUS	F	NPL	CGTATGATGT
84-NPL-49	Legs	LATE	PRUINOSUS	F	NPL	TAGTCTGTCA
84-NPL-50	Head	LATE	PRUINOSUS	F	NPL	TATGTACGTG
84-NPL-51	Body	LATE	PRUINOSUS	M	NPL	CTATACAGTG
84-NPL-52	Forewings	LATE	PRUINOSUS	M	NPL	CAGTCAGAGT
84-NPL-53	Hindwings	LATE	PRUINOSUS	M	NPL	CGCAGTCTAT
84-NPL-54	Legs	LATE	PRUINOSUS	M	NPL	GTATGAGCAC
84-NPL-55	Head	LATE	PRUINOSUS	M	NPL	TGTCTCTATC
84-NPL-56	Body	LATE	PRUINOSUS	F	NPL	CTAGAGTATC
84-NPL-57	Forewings	LATE	PRUINOSUS	F	NPL	CGAGTGCTGT
84-NPL-58	Hindwings	LATE	PRUINOSUS	F	NPL	TATAGCACGC
84-NPL-59	Legs	LATE	PRUINOSUS	F	NPL	TCATGCGCGA
84-NPL-60	Head	LATE	PRUINOSUS	F	NPL	TATGCGCTGC
84-NPL-61	Body	LATE	PRUINOSUS	M	NPL	CAGAGCTAGT
84-NPL-62	Forewings	LATE	PRUINOSUS	M	NPL	TGTACAGCGA
84-NPL-63	Hindwings	LATE	PRUINOSUS	M	NPL	TCACAGCATA
84-NPL-64	Legs	LATE	PRUINOSUS	M	NPL	CGATCGACTG
84-NPL-65	Head	LATE	PRUINOSUS	M	NPL	ACTAGCTGTC
84-SOIL-19-1-1	SOIL	NA	SOIL	NA	SOIL	TTCTCATCGT
84-SOIL-19-3-1	SOIL	NA	SOIL	NA	SOIL	CTCAATCGTA
84-SOIL-19-5-1	SOIL	NA	SOIL	NA	SOIL	CGCTAATGTA
84-SOIL-20-Control-1	SOIL	NA	SOIL	NA	CONTROL	GCCATGTCAT
84-SOIL-20-MDE-1	SOIL	EARLY	DORSATUS	NA	SOIL	TTCTGTTGCC
84-SOIL-20-MDE-2	SOIL	EARLY	DORSATUS	NA	SOIL	TTGTCCTTGC
84-SOIL-20-MDE-3	SOIL	EARLY	DORSATUS	NA	SOIL	CCTGTGTAGA
84-SOIL-20-MDE-4	SOIL	EARLY	DORSATUS	NA	SOIL	GATAAGAAGG
84-SOIL-20-MDE-5	SOIL	EARLY	DORSATUS	NA	SOIL	CAGGTCACAT
84-SOIL-20-NPE-1	SOIL	EARLY	PRUINOSUS	NA	SOIL	GCGTCTGAAT
AE-CONTROL	CONTROL	CONTROL	CONTROL	NA	CONTROL	CTTAGTTCGC

SAMPLEID	BODYPART	TIME	SPECIES	SEX	CATEGORY	BARCODE
AE-CONTROL-1	CONTROL	CONTROL	CONTROL	NA	CONTROL	ATCGCATAGA
FIELD-CONTROL-2	CONTROL	CONTROL	CONTROL	NA	CONTROL	GTAGTACACA
GENE-JET-BLANK	CONTROL	CONTROL	CONTROL	NA	BLANK	ACTGATGTAG
MD-FIELD-CONTROL-1	CONTROL	CONTROL	CONTROL	NA	FIELDCONTROL	TCTCTGTGCA
MD-FIELD-CONTROL-3	CONTROL	CONTROL	CONTROL	NA	FIELDCONTROL	CAGAGAGTCA
QIAGEN-CONTROL	CONTROL	CONTROL	CONTROL	NA	CONTROL	TACGCTGCTG
ZYMO-CONTROL	CONTROL	CONTROL	CONTROL	NA	CONTROL	TCTGCCTATA

Table C.2. Subset of documented endosymbionts in cicada species as well as their tribe. ND = not determined.

Subfamily	Tribe	Genus	Species	Hodgkinia	YLS	Sulcia	References
Cicadinae	Cryptotympanini	*Auritibicen*	*japonicus*	YES	ND	YES	Matsuura et al. 2018
Cicadinae	Cryptotympanini	*Auritibicen*	*bihamatus*	YES	ND	YES	Matsuura et al. 2018
Cicadettinae	Cicadettini	*Cicadettana*	*calliope calliope*	YES	ND	YES	This study
Cicadinae	Cryptotympanini	*Cryptotympana*	*facialis*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Cryptotympanini	*Cryptotympana*	*atrata*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Fidicini	*Diceroprocta*	*semicincta*	YES	ND	YES	Matsuura et al. 2018; McCutcheon et al. 2009
Cicadinae	Fidicini	*Diceroprocta*	*swalei*	YES	ND	YES	McCutcheon et al. 2009
Cicadinae	Leptopsaltriini	*Euterpnosia*	*chibensis*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Polyneurini	*Graptopsaltria*	*nigrofuscata*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Polyneurini	*Graptopsaltria*	*bimaculata*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Sonatini	*Hyalessa*	*maculaticollis*	ND	YES	YES	Matsuura et al. 2018
Cicadettinae	Lamotialnini	*Magicicada*	spp.	YES	ND	YES	Brumfield et al. 2022; Campbell et al. 2015; McCutcheon et al. 2009
Cicadinae	Cryptotympanini	*Megatibicen*	*dorsatus*	ND	YES	YES	This study
Cicadinae	Dundubiini	*Meimuna*	*opalifera*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Dundubiini	*Meimuna*	*oshimensis*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Dundubiini	*Meimuna*	*iwasakii*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Dundubiini	*Meimuna*	*kuroiwae*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Platypleurini	*Neotibicen*	*pruinosus pruinosus*	ND	YES	YES	This study
Cicadinae	Platypleurini	*Platypleura*	*kaempferi*	YES	ND	YES	Matsuura et al. 2018
Cicadinae	Platypleurini	*Platypleura*	*keroiwae*	YES	ND	YES	Matsuura et al. 2018
Cicadinae	Platypleurini	*Platypleura*	*yaeyamana*	YES	ND	YES	Matsuura et al. 2018
Cicadinae	Leptopsaltriini	*Tanna*	*japonensis*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Leptopsaltriini	*Terpnosia*	*vacua*	ND	YES	YES	Matsuura et al. 2018
Cicadinae	Leptopsaltriini	*Terpnosia*	*nigricosta*	ND	YES	YES	Matsuura et al. 2018
Tibicininae	Tettigadinii	*Tettigades*	*undata*	YES	ND	YES	Matsuura et al. 2018
Tibicininae	Tettigadinii	*Tettigades*	*ulnaria*	YES	ND	YES	Matsuura et al. 2018
Cicadettinae	Cicadatrini	*Vagitanus*	*terminalis*	YES	ND	YES	Matsuura et al. 2018

Sequences of ASVs found in samples

Ca Hodgkinia V3V4 ASV found in 84 CC 01, 84 CCL 05, 84 CCL 06, 84 CCL 07, 84 CCL 08, and 84 CCL 09
95.96% identity to Accession LC370610.1 *Candidatus* Hodgkinia cicadicola gene for 16S rRNA, partial sequence, clone: KOSYEZ_BA06 (*Kosemia yezoensis*)
Length of read: 419 bp
> ca4e69f2852d642d3728c4a6d74d9d57
TGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGTCGCGTGAGTGATG
AAGGTTGTGTGAATATCGATGAATATTTTAACTGTAAAGCTCTTTTATCGGTTATAAT
GATGATTGAACCGAAGAATAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAG
ACGGGGGGAGCAAGTGTTGTTCGGATTTATTGGGCGTAAAGGGCGCGTAGGCGGTT
TAATAAGTTGATACTTAAATTTTGAAGCTTAACCTCATTGAAAGTGTCAATACTGTT
AAACTAGAACTTAGCTGGGGTGAACGGAATTCCAAGTGTAGAGGTGAAATTCGTAG
ATATTTGGAGGAACACCGGAGGCGAAAGCGGTTCATTATTCTAAGTTGACGCTGATG
CGCGAAAGCGTGGGGAGCAAACA

Ca Hodgkinia V3V4 ASV found in 84 CCL 09
88.39% identity to Accession LC370610.1 *Candidatus* Hodgkinia cicadicola gene for 16S rRNA, partial sequence, clone: KOSYEZ_BA06 (*Kosemia yezoensis*)
Length of read: 401 bp
> 8922fb0b31acbae9d5e451e5e3f835da
TGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGTCGCGTGAGTGATG
AAGGTCGCAATGACCGTAAAGCTCTTTTGTCGGTGATAATGATGATTGAACCGAAGA
ATAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGGGGGAGCAAGTGT
TGTTCGGATTTATTGGGTGTAAAGGGCGCGTAGGCGGTTTGATAAGTTAGCACTCAA
ATCTTGAGGAGCAACTTCATTAAGGTGCTAATACTGTTAAGCTTTGGGCTTAATGGG
GGTGAACGGAATTCCAAGTGTAGAGGTGAAATTCGTAGATATTTGGAGGAACGCCG
GAGGCGAAAGCGGTTCACTATGTTAAGTTGACGCTGATGCGCGAAAGCGTGGGGAG
CAAACA

Yeast-Like Symbiont (YLS) or cf. *Ophiocordyceps* ITS3-4 ASV found in early collected *Neotibicen pruinosus pruinosus* (NPE 01)
Confidence categorical value assigned to ASV: 0.939046063
93.33% identity to *Ophiocordyceps yakusimensis* (Accession AB044643.1)
91.76% identity to *Ophiocordyceps sobolifera* (Accession MT349951.1)
91.29% identity to endosymbiont found in *Meimuna mongolica* (Cicadidae: Tribe Dundubiini) (Accession MT548754.1)
Length of read: 362 bp
> 0e2be2972d14d489379b632d012df883
GAAATGCGACAAGTAATGTGAATTGCAGAATTCAGCGAGTCATCGAATCTTTGAAC
GCACATTGCGCCCGCCAGCATTCTGGCCGGCATGCCTGTCCGAGCGTCATTGTCGGC
CCTCGAGCCCGCCGTGGCGCGCGCTCGGCGTTGGGGGTCCCGGCCCGACCAGGCCG
CCCCCGAAATTCAGTGGCGACACCCGCCGCACGCCTCCCCTGCGCAGTAGCAGACG
GCCCGCATCGGGGGGCGCCGCAACGGAGGTCACGGCCGTAAAAGAAGGGAGCGCC

GGGGGAGGGGGAAACCCCCCCCCGGAGGCGCCGCCATCTCGTGGTTGACCTCGGAT
CAGGTAGGGCTACCCGCTGAACTTAA

Yeast-Like Symbiont (YLS) or cf. *Ophiocordyceps* ITS3-4 ASV found in late collected *Megatibicen dorsatus* (MDL 91)
Confidence categorical value assigned to ASV: 0.995708972
92.77% identity to *Ophiocordyceps yakusimensis* (Accession AB044643.1)
91.21% identity to endosymbiont found in *Meimuna mongolica* (Cicadidae: Tribe Dundubiini) (Accession MT548754.1)
90.68% identity to *Ophiocordyceps longissima* (Accession MG031297.1)
Length of read: 235 bp
>c9e8576242323001dcc6694d43995db3
GGCGCGCGCTCGGCGTTGGGGGTCCCGGCCCGACTAGGCCGCCCCCGAAATTCGGT
GGCGACACCCGCCGCACGCCTCCCCTGCGCAGTGGCAAACGCTCGCATCGGGGGGC
GCCGCAAACGGAGGTCACGGCCGTAGAAGAGGGAGCGCCGAAGGAGGGGGAAAAC
CCTCCCCCCGGAGGCGCCGCCATCTCGTGGTTGGCCTCGGATCAGGTAGGACTTACCC
GCTGAACTTAA

Yeast-Like Symbiont (YLS) or cf. *Ophiocordyceps* ITS3-4 ASV found in early *Neotibicen pruinosus pruinosus* (NPE 19)
Confidence categorical value assigned to ASV: 0.944989654
93.33% identity to *Ophiocordyceps yakusimensis* (Accession AB044643.1)
91.63% identity to *Ophiocordyceps longissima* (Accession AB968406.1)
91.27% identity to endosymbiont found in *Meimuna mongolica* (Cicadidae: Tribe Dundubiini) (Accession MT548754.1)
Length of read: 364
>03dd4d6dab2a619dd50300b64680d8fe
GAAATGCGACAAGTAATGTGAATTGCAGAATTCAGTGAGTCATCGAATCTTTGAAC
GCACATTGCGCCCGCCAGCATTCTGGCCGGCATGCCTGTCCGAGCGTCATTGTCGGC
CCTCGAGCCCGCCGTGGCGCGCGCTCGGCGTTGGGGGTCCCGGCCCGACCAGGCCG
CCCCCGAAATTCAGTGGCGACACCCGCCGCACGCCTCCCCTGCGCAGTAGCAGACG
GCTCGCATCGGGGGGCGCCGCAACGGAGGTCACGGCCGTAAAAGAGGGAGCGCCG
GGGGAGGGGGAAACACACTCCCCCCGGAGGCGCCGCCATCTCGTGGTTGACCTCGG
ATCAGGTAGGACTACCCGCTGAACTTAA

Beauveria bassiana ITS3-4 ASV
Confidence categorical value assigned to ASV: 0.979203403
100% identity to *Beauveria bassiana* (Accession MT533246.1)
Length of read: 308 bp
> 04392b3bab564aa71617113886f58f32
GAAACGCGATAAGTAATGTGAATTGCAGAATCCAGTGAATCATCGAATCTTTGAAC
GCACATTGCGCCCGCCAGCATTCTGGCGGGCATGCCTGTTCGAGCGTCATTTCAACC
CTCGACCTCCCCTTGGGGAGGTCGGCGTTGGGGACCGGCAGCACACCGCCGGCCCT
GAAATGGAGTGGCGGCCCGTCCGCGGCGACCTCTGCGCAGTAATACAGCTCGCACC

GGAACCCCGACGCGGCCACGCCGTAAAACACCCAACTTCTGAACGTTGACCTCGAA
TCAGGTAGGACTACCCGCTGAACTTAA

Chapter 3 code

```
# use trimmomatic to trim primers from forward and reverse reads
# based on specific primer length for region (e.g. ITS)

module load Trimmomatic/0.39-Java-1.8.0_201

trimmomatic SE -threads 4 ITS3_ITS4_R1.fastq
ITS3_ITS4_R1_trimmed.fastq HEADCROP:20
trimmomatic SE -threads 4 ITS3_ITS4_R2.fastq
ITS3_ITS4_R2_trimmed.fastq  HEADCROP:20

trimmomatic SE -threads 4 V3_F357_N_V4_R805_R1.fastq
V3_F357_N_V4_R805_R1_trimmed.fastq  HEADCROP:17
trimmomatic SE -threads 4 V3_F357_N_V4_R805 _R2.fastq
primer_sorted_trimmed/V3_F357_N_V4_R805_R2_trimmed.fastq
HEADCROP:21
```

```
module load QIIME2/2021.4
# files from trimmomatic need to be gzipped and renamed for
# Earth Microbiome Project (EMP) protocol

gzip V3_F357_N_V4_R805_R1_trimmed.fastq

# files should be forward.fastq.gz , reverse.fastq.gz and
# barcodes.fastq.gz and in a folder by themselves, no other
# files or folders can be present

# create .qza file for QIIME

qiime tools import --type EMPPairedEndSequences --input-path ITS
--output-path ITS.qza

qiime tools import --type EMPPairedEndSequences --input-path
V3V4 --output-path V3V4.qza

# demux data using Keemei validated metadata table with barcodes

qiime demux emp-paired --i-seqs  ITS.qza --m-barcodes-file
metadata.tsv --m-barcodes-column Barcode --p-no-golay-error-
correction --o-per-sample-sequences ITS_demux.qza --o-error-
correction-details ITS_err-corr.qza

# run through dada2
#see aron's code, might want to increase truncate based on
quality scores
# added p-n-reads-learn based on aron's notes
qiime dada2 denoise-paired
--i-demultiplexed-seqs ITS_demux_joined_filtered.qza
--p-trunc-len-f 0
--p-trunc-len-r 0
--p-n-threads 8
--p-n-reads-learn 1000000
--o-representative-sequences ITS_seqs_dada2.qza
--o-table ITS_table_dada2.qza
--o-denoising-stats ITS_stats_dada2.qza
--verbose

# Visualize dada2 table
qiime feature-table summarize --i-table ITS_table_dada2.qza --o-
visualization ITS_table_dada2.qzv --m-sample-metadata-file
microbiome-metadata.tsv

# Visualize dada2 sequence table
```

```
qiime feature-table tabulate-seqs --i-data ITS_seqs_dada2.qza --
o-visualization ITS_seqs_dada2.qzv

qiime metadata tabulate --m-input-file ITS_stats_dada2.qza --o-
visualization ITS_stats_dada2.qzv
qiime feature-table summarize --i-table ITS_table_dada2.qza --o-
visualization ITS_table_dada2.qzv
# train a classifier
#first change the upper case issues
awk '/^>/ {print($0)}; /^[^>]/ {print(toupper($0))}'
sh_refs_qiime_ver8_99_10.05.2021_dev.fasta | tr -d ' ' >
sh_refs_qiime_ver8_99_10.05.2021_dev_uppercase.fasta

#then import
qiime tools import --type FeatureData[Sequence] --input-path
sh_refs_qiime_ver8_99_10.05.2021_dev_uppercase.fasta --output-
path unite-ver8-seqs_99_10.05.2021.qza
#import tax
qiime tools import --type FeatureData[Taxonomy] --input-path
sh_taxonomy_qiime_ver8_99_10.05.2021_dev.txt --output-path
unite-ver8-taxonomy_99_10.05.2021.qza --input-format
HeaderlessTSVTaxonomyFormat
#train

qiime feature-classifier fit-classifier-naive-bayes --i-
reference-reads unite-ver8-seqs_99_10.05.2021.qza --i-reference-
taxonomy unite-ver8-taxonomy_99_10.05.2021.qza --o-classifier
unite-ver8-99-classifier-10.05.2021.qza

qiime feature-classifier classify-sklearn --i-classifier unite-
ver8-99-classifier-10.05.2021.qza --i-reads /home/a-
m/cdana2/Exobiome/primer_sorted_trimmed/ITS/results/demux-
duo/ITS_demux.qza --o-classification
ITS_demux_joined_filtered_taxonomy-single-end.qza

qiime tools export --input-file ITS_seqs_dada2_taxonomy.qza --
output-path phyloseq

qiime alignment mafft --i-sequences Arch_seqs_dada2.qza --o-
alignment Arch_seqs_aligned.qza

qiime alignment mask --i-alignment Arch_seqs_aligned.qza --o-
masked-alignment Arch_masked-aligned-rep-seqs.qza

qiime phylogeny fasttree --i-alignment Arch_masked-aligned-rep-
seqs.qza --o-tree Arch_unrooted-tree.qza
```

```
qiime phylogeny midpoint-root --i-tree Arch_unrooted-tree.qza --
o-rooted-tree Arch_rooted-tree.qza

qiime diversity core-metrics-phylogenetic --i-phylogeny
Arch_rooted-tree.qza --i-table Arch_table_dada2.qza --p-
sampling-depth 1500 --m-metadata-file metadata_sub.tsv --output-
dir ITS_metrics_1500_sub


# output to visualize

qiime diversity alpha-group-significance --i-alpha-diversity
metrics/faith_pd_vector.qza --m-metadata-file metadata_sub.tsv -
-o-visualization metrics/faith-pd-group-significance.qzv
```