2019

# Haplotype Diversity and Population Genetic Structure of Antistrophus Gall Wasps Associated with Two Silphium Species and Implication for Host Mediated Speciation

Jiaxin Deng
*Eastern Illinois University*

The Graduate School
EASTERN ILLINOIS UNIVERSITY

## Thesis Maintenance and Reproduction Certificate

FOR: Graduate Candidates Completing Theses in Partial Fulfillment of the Degree
Graduate Faculty Advisors Directing the Theses

RE: Preservation, Reproduction, and Distribution of Thesis Research

Preserving, reproducing, and distributing thesis research is an important part of Booth Library's responsibility to provide access to scholarship. In order to further this goal, Booth Library makes all graduate theses completed as part of a degree program at Eastern Illinois University available for personal study, research, and other not-for-profit educational purposes. Under 17 U.S.C. § 108, the library may reproduce and distribute a copy without infringing on copyright; however, professional courtesy dictates that permission be requested from the author before doing so.

Your signatures affirm the following:

- The graduate candidate is the author of this thesis.

- The graduate candidate retains the copyright and intellectual property rights associated with the original research, creative activity, and intellectual or artistic content of the thesis.

- The graduate candidate certifies her/his compliance with federal copyright law (Title 17 of the U. S. Code) and her/his right to authorize reproduction and distribution of all copyrighted materials included in this thesis.

- The graduate candidate in consultation with the faculty advisor grants Booth Library the nonexclusive, perpetual right to make copies of the thesis freely and publicly available without restriction, by means of any current or successive technology, including but not limited to photocopying, microfilm, digitization, or internet.

- The graduate candidate acknowledges that by depositing her/his thesis with Booth Library, her/his work is available for viewing by the public and may be borrowed through the library's circulation and interlibrary loan departments, or accessed electronically. The graduate candidate acknowledges this policy by indicating in the following manner:

__X__ Yes, I wish to make accessible this thesis for viewing by the public

_____ No, I wish to quarantine the thesis temporarily and have included the ***Thesis Withholding Request Form***

- The graduate candidate waives the confidentiality provisions of the Family Educational Rights and Privacy Act (FERPA) (20 U. S. C. § 1232g; 34 CFR Part 99) with respect to the contents of the thesis and with respect to information concerning authorship of the thesis, including name and status as a student at Eastern Illinois University. I have conferred with my graduate faculty advisor. My signature below indicates that I have read and agree with the above statements, and hereby give my permission to allow Booth Library to reproduce and distribute my thesis. My adviser's signature indicates concurrence to reproduce and distribute the thesis.

Graduate Candidate Signature

Jiaxin Deng
Printed Name

Biology Sciences
Graduate Degree Program

Faculty Adviser Signature

Zhiwei Liu
Printed Name

May 28, 2019
Date

***Please submit in duplicate.***

Haplotype diversity and population genetic structure of Antistrophus gall wasps associated with two Silphium species and implication for host mediated speciation

(TITLE)

BY

Jiaxin Deng

**THESIS**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

Master of Science in Biological Sciences

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

2019

YEAR

I HEREBY RECOMMEND THAT THIS THESIS BE ACCEPTED AS FULFILLING
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

| | | | |
|---|---|---|---|
| THESIS COMMITTEE CHAIR | 5-16-19 DATE | DEPARTMENT/SCHOOL CHAIR OR CHAIR'S DESIGNEE | 5/29/19 DATE |
| THESIS COMMITTEE MEMBER | 5-16-19 DATE | THESIS COMMITTEE MEMBER | DATE |
| THESIS COMMITTEE MEMBER | 5-16-19 DATE | THESIS COMMITTEE MEMBER | DATE |

# Haplotype diversity and population genetic structure of *Antistrophus* gall wasps associated with two *Silphium* species and the implication for host mediated speciation

Jiaxin Deng

Candidate for M. Sci. in Biology

Department of Biological Sciences, Eastern Illinois University

Charleston, IL, USA

**Thesis Advisor:**

Dr. Zhiwei Liu

**Committee Members:**

Dr. Ann Fritz

Dr. Gordon Tucker

## Abstract

Gall making insects form a special feeding guild of phytophagous animals, and by manipulating host plant tissue differentiation, are able to avoid plant chemical defenses and thus have no need for counter defense mechanisms. Host plant selection is crucial to the evolution of these insects because successful gall formation is largely dependent on host plant ability to respond to stimuli. In Illinois and neighboring states, *Antistrophus* gall wasps associated with the rosin weed (*S. integrifolium*) and the cup plant (*S. perfoliatum*) are morphologically indistinguishable and thus have been treated as belonging to single species. However, the wasps from the host plant species display strong host preference to the host plant species from which they are reared as well as other life history differences, suggesting that they may represent two distinct, although cryptic, species. In order to test the competing hypotheses regarding the identity of these wasps, I investigated whether there exists genetic discontinuity between wasp populations defined by host plant species based on molecular data. Wasps associated with either host plant species, mostly from sympatric populations, were collected from Illinois, Wisconsin, Indiana, Missouri and Iowa and extracted total DNA were PCR amplified and sequenced for three DNA loci, including two mitochondrial genes, (Cytb and COI) and one nuclear gene (28S D2). Phylogenetic reconstruction for all samples based on the Cytb and COI genes did not recover any monophyletic lineage consisting exclusively of wasps sampled from either of the two host-plant species but revealed a clear pattern of lineage bias toward host plant species in the tree resolved from Cytb gene. Examination of haplotype diversity revealed five haplotypes among our Cytb (n= 60) and COI (n= 59) sequences, respectively and three haplotypes for the 28S gene (N= 65). Further analyses

of the genealogical relationship among the haplotypes using network method uncovered a distinct host affiliation pattern for Cytb, but not for the COI gene, similar to the results of the phylogenetic analyses. For the more conservative 28S gene, the wasps associated with the two host plant species can be readily separated based on a single indel event at position 181, except for one single sample (R28), which was associated with rosin weed, but without the extra T at 181 position found in other wasps associated with rosinweed. Except for this indel and an additional transitional mutation for R28, there was no variation throughout the entire 524 bp length of the 28S gene. These results do not provide unambiguous support for the cryptic species hypothesis, but revealed significant discontinuity between the gall wasp gene pools associated with the two host plant species, which may indicate incipient host plant mediated speciation.

## Acknowledgement

This research was supported by the Eastern Illinois University and Central South University of Forestry and Technology in China. Many people have made invaluable contributions, both directly and indirectly to my research. I would like to express my gratitude to all those who helped me during the writing of this thesis. My deepest gratitude goes first and foremost to Dr. Zhiwei Liu, my supervisor, for his constant encouragement and guidance. He has walked me through all the stages of the writing of this thesis. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

I am also grateful to my committee members Dr. Ann Fritz and Dr. Gordon Tucker for providing me with valuable advice and for their teaching, their continuous encouragement and massive help during these two years. I wish to thank Mr. Doug Ginder for sharing with me experimental materials and relevant information. I would like to thank Dr. Eden L Effert-Fanta especially for her patient instruction and expert guidance when I started to work as her TA for Anatomy and Physiology II lab, and for her unfaltering belief in me by requesting me as her TA throughout the two years of my study at EIU. I wish to say "Thanks!" to all the faculty and staffs in the Department of Biological Sciences, from whose devoted teaching and enlightening lectures I have benefited a lot and academically prepared for the thesis.

Finally, I want to express my deepest gratitude to my parents for their love and great confidence in me all through these years of studies in China and in the USA. I would also like to thank my friends for their unfailing support, continuous encouragement, and warming friendship throughout my two years of study and through

the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Thank you all!

# Table of Contents

# List of Figures

# List of Tables

# Introduction

Plant galls are abnormal plant tissue growth formed at specific reactive sites of vascular plants, thought to entirely to benefit the gall inducing organisms, typically insects and enable the gall makers to complete their life cycles. As such, a gall is a complex, subtle biological system that involves two important participants: gall making insects and their host plants (Raman et al 2005). During the process of gall development, the gall maker inhibits the normal plant tissue growth by triggering new morphological and physiological responses of their host plants (Meyer and Maresquelle 1983). The parasitic action of the gall making organisms consequently induces all types of original growth, including some abnormal differentiation, in the tissues of the host plant (Meyer 1987). In this sense, plant galls are manifestation of an intricate relationship between the gall making insects and their host plants, where the gall making insects are intimately dependent on their host plants. Ultimately, the biology, ecology and evolution of gall making insects can be influenced by the characteristics of their host plants, resulting in different patterns in the interactions between gall making insects and their host plants (Schoonhoven et al 1998).

The evolution of gall maker and host plant interactions is the driving force behind the creation of gall phenotype which is influenced by both insect genotype and plant genotype, where the insect phenotype represents a stimulus and plant phenotype represents a response (Weis 1986, 1988). From the host plants' aspect, the formation of insect galls originates from the environmental stimuli, and from the insects' aspect, the insect galls are extended phenotypes. As an extended phenotype, the gall making insect can to some extent control the morphology of the galls, and the physiological structure

changes in host plant tissues under the influence of the gall maker contribute to the insect gall formation (Stern 1995). Multiple studies have demonstrated that the insect genotype, rather than the plant genotype, is the determining factor for insect gall formation and gall shape, representing extended insect phenotypes of various insect groups, including wasps, flies, aphids and thrips (Cook 1902, Dawkins 1982, Kinsey 1920, Stern 1995, Stone and Cook 1998).

As a special guild of plant feeders, the gall making insects are tremendously diverse; there may exist up to 211,000 species of gall making insects in 7 orders and 20 families in the world (Espirito-Santo and Fernandes 2007). Cynipid wasps (Cynipidae, Hymenoptera), better known as gall wasps, represents one of the largest radiations of gall making insects, consisting of over 1,500 described species alone (Ronquist et al. 2015), especially noticeable on the oak host plant group (Cornell 1983). Members of the family are primarily gall making wasps, and also include inquiline wasps that do not make their own galls, but live in the galls made predominantly by other cynipid wasps (Ronquist et al 2015) and rarely by other insect so (van Noort et al. 2007, Wachi et al. 2011). It is generally accepted that the species diversity of Cynipidae is probably much higher than what is known at the present and may have as many as at least 3,000 species worldwide if properly studied (Ronquist 1999, Ronquist and Liljeblad 2001, Ronquist et al 2015).

Because of the intricate nature of the interactions between a gall maker and its host plant, it is not surprising that most gall making insect species are highly host and even plant organ specific (Dreger-Jauffret & Shorthouse 1992). In Cynipidae, this specificity of host plant associations is seen at two different levels - they are extremely phylogenetically conservative in host-plant use at higher levels as compared to most other

phytophagous insects (Ronquist et al 2015). Cynipidae are highly diversified at the species level, not only host plant species specific, but also showing specificity in tissues / organs (Ronquist and Liljeblad 2001), resulting in the high species diversity associated with limited host plant lineages (Cornell 1983, Askew 1984). In fact, the 1,500 known species of cynipid wasps utilize plants belonging to a limited number of plant families. The tribes of Cynipini and Diplolepidini, for instance, are each exclusively associated with Fagaceae and Rosaceae host plants, respectively (Ronquist and Liljeblad 2001, Ronquist et al, 2015).

In spite of the high species diversity and high degree of host specificity, the general structure of cynipid galls are nonetheless very similar, and the typical cynipid gall consist of three major tissue layers, the outer most peripheral tissue layer that forms the diverse external morphology of the galls and usually contains tannins in considerably higher concentrations than the normal plant tissue from which they are derived (Cornell 1983), the middle sclerenchyma consisting of tightly aligned small and thick-walled cells, and the inner most layer of nutritive tissue of large, thin-walled cells that are rich in sugar, lipids, and protein contents and free of tannins and other secondary metabolites (Cornell 1983, Bronner 1992, Ronquist and Liljeblad 2001). The outer tissue layers for physical protection from the developing gall wasp larva receives from desiccation for the developing gall wasp larva (Fernandes and Price 1992, Miller et al 2009), extreme temperatures (Williams et al 2002), and natural enemies (Price and Pschorn-Walcher 1988) while the developing larva derives all its nutrient needs exclusively from the inner nutritive tissue (Cornell 1983, Bronner 1992).

The interactions between plants and their specialist herbivores have long been considered important for the evolution of both, and the coevolutionary arms race model has been constructed based on the interactions between butterflies and their host plants (Ehrlich and Raven, 1964). According to the arms race coevolution model, herbivores feeding on plants select for individuals that have novel defensive compounds, which would provide full or partial protection against the herbivores, while those not producing antiherbivore compounds would have a full set of insect herbivores, resulting in reduced fitness relative to the mutants. The novel plant genotypes will have increased fitness through reducing the herbivore insect damage and therefore become more common. Through this process, plants can escape some or all of their herbivores, entering a "new adaptive zone", which represent an empty niche available for colonization by the insect herbivores that would subsequently evolve the ability to use them. Further defense trait evolution in the host plant would be necessary to counter the new attacks by the new insect herbivore genotype, engaging a cycle of reciprocal evolutionary responses between the two parties, resulting in host plants with increasingly effective defense chemistry against their herbivores and the herbivores with increasingly effective counter defense mechanisms against host plant in return (Ehrlich and Raven 1964, Marquis et al 2016). Because insects and angiosperm plants are the two most diverse organismal groups on the earth, much of the terrestrial biological diversity could be accounted for by the coevolutionary process that drives both the evolution and speciation of herbivorous insects and their host plants.

The arms race coevolutionary model between herbivores and host plants requires the existence of plant defense chemicals to force host plant specialization of the

herbivorous insects. These host defense chemicals belong to one of the two major groups of internal plant compounds – secondary metabolites that may have various functions, but nonetheless are not essential for the basic plant growth, when compared to the nutrient compounds that are essential to plants' basic growth (thus the term primary metabolites) (Bernays & Chapman 1995). These secondary plant compounds function as toxins and deterrents to the herbivores (Bernays & Chapman 1995, Strauss and Zangerl 2001). Although the plant families whose members are potentially hosts to gall making insects contain various secondary defense chemicals (Bernays & Chapman 1995), these chemicals would be reallocated in differentiated cell layers of the galls. For example, concentrations of tannins, which are digestion inhibiting to herbivores, are found to be higher in the outer layers of cynipid galls than in the normal oak organs from which the galls are derived (Cornell 1983) while almost absent from the inner nutritive cells (Bronner 1992, Ronquist and Liljeblad 2001). This tissue differentiation of the galls would create an effective defense barrier against gall predators, fungi and microorganisms (Cornell 1983) while the nutritive cell layer free of secondary chemicals would relax the selection pressure for the gall making insects to deal with plants' chemical defense, effectively disrupting reciprocal coevolutionary cycle involving the regular herbivore-host plant relationship. As such, host plant specificity of gall making insects could not be explained as a response to the host plants' chemical defense.

For the successful formation of galls, the female insects have to choose an appropriate location for oviposition, a process that involves two criteria – the correct host plant species and a reactive site on the plant (Weis et al. 1988). Because of the intricate nature between the gall making insect and its host plants as discussed above, the

successful gall formation depends on the appropriate response of the host plant to the stimuli of the gall making insects (Weis et al 1988, Raman et al 2005). In cynipid gall wasps, the oviposition of the female into the right plant tissue and at the right time is essential for successful gall induction (Ronquist and Liljeblad 2001), mostly because gall formation is organ specific (Cornell 1983, Askew 1984) and requires undifferentiated plant tissue (Weis et al 1988). In this sense, plant defenses might be more evolutionarily labile than being driven by herbivore traits related to the association and not constrained by the gall makers while the gall makers are under strong pressure to choose the right host plant. A mutation that causes the gall maker female to lay eggs on the wrong host plant species or genotype, if successful, may lead to new species event, especially when the original host is in short supply while the "wrong" host are abundant (Larsson and Ekbom 1995). This would result in an asymmetry in the evolutionary interactions between the gall makers and their host plants, leading to a pattern more consistent with ecological fitting or resource tracking hypothesis (Agosta et al, 2008, 2009) than with the arms race model of coevolution. This is considered especially likely for the relationship between insect herbivores and their host plants, leading to what has been known as ecological speciation for the insect herbivores (Matsubayashi et al 2010) as have been observed in multiple scenarios (Funk 1998; Berlocher & Feder 2002; Dres & Mallet 2002; Jiggins 2008). A recent study by Endara et al (2017) on the tropical *Inga* species and their lepidopteran herbivores provided strong evidence for the resource tracking "coevolutionary" model. In the relationship between cynipid gall makers and their host plant, an asymmetrical evolutionary process is highly likely with the gall wasps track

their host plants, and host specificity and host shifting play a major role in the formation of new species and overall evolution of the group.

The acceptance or rejection of a plant by insect herbivores depends on their behavioral responses to the plant's physical or chemical traits. All plants release a characteristic mixture of volatile molecules, including various short chain alcohols, aldehydes, ketones, esters, aromatic phenols, and lactones, as well as mono- and sesquiterpenes, which gives a plant its distinctive odor (Bernays and Chapman 1995). This characteristic mixture of volatiles varies with, and therefore, reveals significant information about plant identity, such as plant species and varieties within a species (Vivaldo et al 2017, Ahmed et al 2019) as well as plant conditions such as herbivore damages (McDaniel et al., 2016). Many insect herbivores use these plant volatiles as cues for locating their preferred host plants (e.g., Ahmed et al 2019, Gray et al 2015, Tooker and Hanks 2004, Tooker et al 2005). Because the mostly sessile life history of gall making insects, the host selection for most gall makers are completely reliant on the choices of the egg-laying female and thus the responses of the female to plant volatile cues can potentially lead to host isolation and hence speciation of the gall-maker.

Gall making insects, due to their concealed habits, frequently show some specialized characteristics in their life history that may contribute to their high reproductive success. The typical life history of gall making insects can be categorized into four major types: multivoltine, bivoltine, univoltine and overlapping generations, of which the univoltine life history is the most common type (Weis et al. 1988). The univoltine life history is apparently an adaptation to the annual phenology of the host plants, allowing the gall makers to synchronize its development with the seasonal growth

of their host plants, especially when the host plants are annual or perennial herbs. As a specific example of univoltine gall making insects, adult females of *Eurosta solidaginis* gall fly laid an egg into the bud of the host plant in late spring. After hatching, the larva bores into the stem and resides in the central chamber, which will be surrounded by 8-9 mm thick tissue layer within two months. The larva reaches full size in October and adult emergence from the gall takes place the next May. Cynipid gall wasps associated with herbaceous host plants are typically univoltine (Askew 1984).

Highly similar life history can also lead to morphological convergence. This phenomenon is particularly obvious in the gall wasp family of Cynipidae: all cynipid gall wasps are round in general body shape, with a relatively large and round abdomen, with weak flight capability, and lacking characteristic external morphology other than being smooth, which renders species identification extremely difficult in Cynipidae (Ronquist 1994, 1999, Ronquist et al., 2015). Taxonomy and species identification in Cynipidae often requires molecular data (Nicholls et al., 2018, Zhu et al., 2015, Yang et al, 2019).

The genus *Antistrophus* (Cynpoidea: Cynipidae) comprises 10 valid species, of which eight species are primarily found in Illinois and neighboring states in the Midwestern USA, while only two species, i.e., *A. chrysothamni* and *A. microseris*, are from Arizona from California, respectively (Table 1) (Liljeblad 2018, Ginder & Liu, unpublished data). When life history information is available, all species of the genus are univoltine (Burks 1979, Fay & Hartnett, 1991, 1993; Fay et al 1996, Fay & Throop 2005, Tooker and Hanks 2004, Tooker et al 2005, Liu, unpublished data). Most of these galls are sealed galls formed inside the stems of the host plant, except for one species, *A. silphii*, which induces enlarged galls terminally on the stems of the host plant species

*Silphium perfoliatum* (Gillette 1891, Burks 1979). Fay and colleagues reported the effects of gall formation by *A. silphii* on a different host plant species, In a series studies on the effects of gall formation on host plant vegetative growth, Fay and colleagues for the first reported *A. silphii* on a different host plant species, the rosinweed *S. integrifolium* (Fay & Hartnett, 1991, 1993; Fay et al 1996, Fay & Throop 2005). The species identification, however, were apparently not subjected to solid taxonomic studies.

Recent studies showed that wasps of *Antistrophus* spp. from different *Silphium* species in the Midwestern United States were reproductively isolated because of volatiles of the host plants (*Silphium* spp. ) (Tooker and Hanks 2004, Tooker et al 2005). Adult females rely on olfactory reception to distinguish between host plant species by receiving volatiles; females then choose an appropriate host plant location for oviposition (Tooker and Hanks 2004). The plant volatiles appear to serve as important behavioral cues for mate location of adult gall wasps and females of *A. rufus* were able to discriminate the monoterpene blend emitted by the stems of *S. laciniatum* (Tooker et al 2005). Adult males of the wasps species are protandrous, *i.e.,* males emerge before the females, and as such are challenged for finding the appropriate mate location. Field data suggested that the males apparently used the plant volatiles as surrogates for sex hormone to locate the females by relying on plant volatiles to locate where females may choose to oviposit on stems of the host plant species (Tooker and Hanks 2004). In fact, the reproductive isolation between wasp populations associated with different host plant species were found to be complete such that what were considered to represent host races were convincingly demonstrated to be separate species (Tooker et al 2004).

In the case involving *Antistrophus silphii* complex associated with rosin weed (*S. integrifolium*) and cup plant (*S. perfoliatum*), the galls from the two *Silphium* species have different morphological characteristics, for example, galls on cup plants, compared to those on rosinweed plants, are larger on the average, softer, and lack rough pubescence and resin. The wasps from the cup plant galls appear to have a slight, but distinctly earlier emergence peak. Lab experiments have also demonstrated that the female wasps from both host plant species had strong preference for their own original host plant species (Ginder, unpublished data). Nonetheless, the wasps reared from the two host plant species are morphologically indistinguishable. The discrepancies between the life history traits and the adult morphology of the gall wasps may indicate two competing scenarios – either the gall wasps from both *Silphium* species belong to the same species, as previously suggested (Fay & Hartnett, 1991, 1993; Fay et al 1996, Fay & Throop 2005), and the differences in gall morphology and life history are phenotypic variations as a result of the interactions with the host plants. Alternatively, the differences in gall morphology and life history do represent genotypic differentiation associated with host plant species, but relatively recent divergence time and convergent evolution, due to similar hosts and life history, resulted in two morphologically very similar species (cryptic species) or host races.

DNA sequence data have been shown to especially useful in solving taxonomic problems that are not possible based on morphology and are used extensively in the discovery of DNA-based species identification of specimens, and characterization of genetic diversity of species as well as for distinguishing intraspecific from interspecific genetic variations (Sheth and Thaker 2017, Joy and Crespi 2007). The technique of DNA

barcoding relies on analyzing DNA sequences and can be used as a supplementary system for traditional taxonomy to distinguish millions of species, particularly those species without properly studied or morphological characteristic details or simply lacking such characteristics (Vogler and Monaghan 2007, Sheth and Thaker 2017, Hebert et al 2003a, Hebert et al 2003b, Hebert and Gregory 2005, Nicholls et al 2010). The DNA barcoding system not only has the potential to database the DNA sequences, but also contains a wealth of analytical and bioinformatics tools to derive large numbers of meaningful conclusions about biological diversity as well as facilitates specimen identification by using simple molecular protocols (Teletchea 2010, Kress and Erickson 2008, Ji et al 2013, Bhargava and Sharma 2013). The synergistic use of DNA barcoding and traditional taxonomy can be beneficial in constructing a proper mechanism for biodiversity conservation and overcome drawbacks of each method alone (Sheth and Thaker 2017).

The mitochondrial DNA barcodes are commonly used in systematic and ecological studies for species-level identification by sequencing high variable gene regions, such as cytochrome oxidase I gene (COI) and cytochrome b gene (Cytb). Their applications were found in studies of phylogeny reconstruction, population genetics, evolutionary mechanisms of phytophagy and other life history traits, and coevolutionary relationship between host plant and gall making insects as well as evolution and geographic expansion of invasive pests including some gall making insects (Ndong et al 2015, Ballard and Whitlock 2004, Detwiler and Janovy 2008, Avise 2009, Muths et al 2008). COI barcodes in particular have been shown to reliable for defining phylogeographic groups within species and identifying taxa from species complexes

(Hebert et al 2003a, Porter and Hajibabael 2018) while Cytb sequences can provide information on within-species variation as well as between-species relationships (Rogers and Gonzalez 2010). Because both COI and Cytb genes have strong interspecific variability and intraspecific polymorphism, they are commonly more discriminating than nuclear genes in the study of DNA barcoding (Borsa et al 2009, Acs et al 2010) and have become the standard genetic markers for species identification and lower level phylogeny reconstruction (e.g., Kjer et al. 2014, Zhu et al., 2015, Zhou et al. 2016, Nicholls et al., 2018, Huang et. al, 2019, Yang et al., 2019).

In studies on phylogenetic relationship at lower taxonomic levels above species, some moderately conservative nuclear gene markers, such as the large nuclear ribosomal RNA gene 28Sgene, are considered more useful (Acs et al, 2010, León et al 2016). Sequence data of these genes are often combined with mitochondrial sequence data to determine the species-level identification and description (Friedrich and Tautz 1997, Acs et al 2010, Li et al 2010, Ronquist et al., 2015, Schwéger et al 2015a, 2015b, Yang et al. 2019). Because the ribosomal 28S gene is relatively conserved limited intraspecific and interspecific variation (Low et al 2014), it is not suitable for DNA barcoding. However, non-coding regions of nuclear DNA genes may also be highly variable and thus useful for genealogical as well as taxonomic studies at the species level or below, such as the internal transcribed spacer (ITS) gene makers ITS1 and ITS2 (De Rojas et al. 2006, Zhang et al. 2019). For that reason, ITS1 and ITS2 DNA were also sequenced as part of the effort to decode the identity of the *Antistrophus* gall wasps associated with *S. integrifolium* and *S. perfoliatum* plants (Ginder and Liu, unpublished data). For my thesis project, I used DNA sequence data of three gene markers, including mitochondrial (COI

and Cytb) and one nuclear (28S) to examine whether the populations of *Antistrophus* gall wasps associated the two *Silphium* species represent distinct gene pools and discuss the implication of the results regarding host plant species mediated speciation of the gall making insects as a specialized herbivores.

## Materials and methods

### *Sampling*

Galls were collected in the winter months of the previous year or in early spring before May and kept in screened containers. Adult wasps were killed and fresh preserved in 100% ethanol after emergence and specimens were stored in freezer at -20 °C or lower temperature. The entire set of samples consisted of 68 *Antistrohpus* species (35 from rosin weed and 33 from cup plant) collected as galls from Batavia, Illinois and adults from Charleston, Illinois; Clinton, Indiana; Marshall, Iowa; Trempealeau, Wisconsin; Missouri (Ginder and Liu, unpublished data). These samples were code named as R1 to R37 and C1 to C34, representing their respective host association with the rosin weed (*S. integrifolium*) and the cup plant (*S. perfoliatum*). The entire collection includes 4 samples from cup plant and 8 from rosin weed from Fermilab, Batavia, Illinois, 10 samples from rosin weed and 2 from cup plant from Marshall, Iowa, 12 samples from rosin weed and 8 from cup plant from Trempealeau, Wisconsin, 5 samples from rosin weed from Missouri, and 17 and 2 samples from cup plant from Clinton, Indiana and Lincoln Log Cabin, Charleston, Illinois, respectively (Fig.1, Table 2).

### *DNA extraction and sequencing*

Whole genomic DNA was extracted from the entire insect with wings removed using proteinase K protocol (Schwenk et al 1998). Extracted DNA was eluted in RNAse free water and stored in -20 °C freezer until being used for PCR amplification.

In this study, two mitochondrial genes, cytochrome b (Cytb) and cytochrome c oxidase subunit I (COI), and one nuclear gene, ribosome RNA gene D2 region 28S D2 (28S). A Cytb fragment about 492 bp long was amplified using the forward primer CB1

(5' TAT GTA CTA CCA TGA GGA CAA ATA TC 3') and reverse primer CB2 (5' ATT ACA CCT CCT AAT TTA TTA GGA AT 3') (Jermiin and Crozier 1994). A COI fragment about 703 bp long was amplified using the forward primer LCO-1490 (5' GGT CAA CAA ATC ATA AAG ATA TTG 3') and reverse primer HCO-2198 (5' TAA ACT TCA GGG TGA CCA AAA AAT CA 3') (Folmer et al 1994). A 524 bp long fragment of the D2 expansion of 28S gene was amplified using forward primer D2f (5' CGT GTT GCT TGA TAG TGA AGC 3') and reverse primer D2r (5' TCA AGA CGG GTC CTG AAA GT 3') (Hancock et al 1988). For all genes, each 25µl polymerase chain reaction (PCR) mix cocktail was carried out in a PTC-100 Programmable Thermal Controller (MJ Research) and consisted of 12.5µl 2X GoTaq Green MasterMix, 2.0µl of each primer, 7.5µl RNase Free water and 1.0µl DNA template. MasterMix are ready-to-use 2X solutions that contain optimal concentrations of GoTaq DNA polymerase, dNTPs, $MgCl_2$ and reaction buffer for efficient amplification of DNA templates. The Cytb mix cocktails were denatured initially with 5 min at 95°C, followed by 35 cycles of 1 min denaturation at 95°C, 1 min annealing at 48°C and 1 min elongation at 72°C, with final step of 5 min at 72°C for final extension. The cycling conditions of COI gene were almost the same as for Cytb, except that only 30 thermal cycles were used. The PCR thermal regime of 28S D2 was very similar to that for Cytb amplification, except that a lower annealing temperature (45°C) was used. The amplified PCR products were then loaded onto precast 1% agarose gel (0.8g molecular grade agarose and 40ml TBE buffer) with 1µl 6 x GelRed loading dye (Biotium, USA) added and run on Owl EasyCast Minigel System B2 (Thermal Fisher, USA) at 120V DC for detection of successful amplification. After running for 30 min, the gel was visualized using Gel Doc XR+ gel imager (Bio-Rad,

USA). Successfully amplified PCR products were sequenced by Invitrogen (Shanghai, China).

*Sequence analysis and Phylogenetic reconstruction*

All samples sequence by Invitrogen were reviewed for sequence quality and sequences with ambiguous peaks were discarded. Out of the 204 sequenced samples (68 for each genes), a total of 185 sequences were kept for further analysis, including 60 for Cytb, 59 for COI, and 65 for 28S D2. For outgroup comparison, Cytb sequence of *Periclistus brandtii* (GenBank accession number: AF395141.1) and COI gene sequence of *Antistrophus rufus* (GenBank accession number: DQ012626.1) were retrieved from GenBank. All new sequences acquired in this project were also submitted to GenBank for permanent depository (Table 2).

Multiple sequence alignment was conducted separately for gene using Clustal X implemented in MEGA X using default parameters (Kumar et al 2018). Each sequence set was then trimmed to remove excessive leading and trailing gaps and the final aligned sequence lengths for Cytb, COI, and 28S were 390bp, 638bp, and 524bp, respectively. The best-fit model of base substitutions was tested using MrModeltest 3.7 (Nylander 2004) using the Akaike information criterion (AIC). The chosen best-fit models of base substitution, GTR for Cytb, HKY for COI, and F81 for 28S D2 were incorporated into subsequent phylogenetic analysis. All the phylogenetic analysis was carried out in Mrbayes 3.2 (Ronquist et al 2011) and ran under the parameters of the above best models. GTR model (nst = '6', rates = 'equal') for Cytb, HKY model (nst = '2', rates = 'equal') for COI and F81 model (nst = '1', rates = 'equal') for 28S D2. For all analyses, 1,000,000 generations were run and were sampled every 1000$^{th}$ generations (10,000 trees

were sampled). A majority consensus tree was calculated after discarding the initial 30 percent sampled trees from all trees saved in Bayesian analysis using Monte Carlo Markov Chain sampling. The final phylogenetic trees were manipulated for better presentation in Figtree 1.4.3 (Rambaut 2012). Host plant association of wasps were optimized onto Cytb, COI and 28S D2 gene trees using PAUP* 4.0b10 (Swofford 2003) to examine whether phylogeny matches with host association, and hence provide evidence for cryptic species or genotypic isolation associated with host plant species.

*Haplotypes diversity and genealogical relationship among gene haplotypes*

Aligned sequence data for each gene were subjected to PAUP* 4.0b10 (Swofford 2003) for detecting haplotypes. Since genetic exchange between population of same species and even between closely related species, although to a much less degree, the genealogical relationship among samples from different populations is rarely bifurcating. Tather, population level genealogical relationship is very likely to be reticulate due to genetic exchange between organism. Therefore, network methods have been extensively used to infer population level genealogies based on nucleotide sequences, especially when divergence level are low. In this study, genealogical relationship network among the haplotypes of the three gene markers was each estimated using TCS 1.21, a software package useful for estimating relationships among organisms with traditional methods or for exploring phylogeographic history of organisms with a nested analysis procedure (Clement et al 2005). The final networks of haplotype relationships were incorporated with information on host plant association as well as frequency of association types.

## Results

*Phylogenetic analyses*

The phylogenetic tree among Cytb sequences found using Bayesian method resolved one major monophyletic clade consisting of all wasps associated with rosin weed, together with 6 samples associated with cup plant (Fig 2). Shown in the unrooted tree, the Cytb sequences is divided into two clusters, one consisting of two haplotypes exclusively associated with cup plant and other consisting of three haplotypes predominantly associated with rosin weed host plant on (Fig 3). The phylogenetic tree among COI sequences did not show much resolution nor reveal any consistent host association pattern (Fig 4). The radiate phylogenetic tree based on COI also did not show much resolution pattern (Fig 5). No relationship was resolved at all for the 28S D2 sequences (Fig 6).

*Haplotype diversity and network analyses*

Five haplotypes were detected for Cytb and COI, respectively (Table 3). The five Cytb haplotypes can be further grouped into two sets based on genealogical relationship: CT1 and satellites (CT3 and CT5), which are predominantly associated with rosin weed, and CT2 and satellite CT4, which are exclusively associated with cup plant (Fig. 7). CT1 as the largest clade comprised 26 samples from rosin weed and only 2 samples from cup plant, whereas CT2 comprised 20 samples from rosin weed and 0 from cup plant. Clade CT2 only showed one mutation from clade CT1, which illustrated a C-to-T substitution at nucleotide position 44 in the 390bp Cytb segment. Clade CT3 consisted of 2 samples from cup plant with single mutation from clade CT1, and the substitution was transferred from C to T at 233 diagnostic nucleotide position. Clade CT5 comprised 5 samples from

rosin weed and 2 samples from cup plant. It also showed one single mutation from clade CT1, which is a T-to-C substitution at 230 diagnostic nucleotide position. Clade CT4 comprised 2 samples from cup plant that revealed two mutations from clade CT1 and one mutation from clade CT2. One mutation was the same as the C-to-T substitution of CT2 from CT1, and the other substitution was a T-to-A at 390 diagnostic nucleotide position.

For COI, most sequences (87%) belong to the single dominant haplotype CO1 and no haplotype with a membership size larger than 3 display exclusive or dominant host plant association or affinity (Fig. 8). Clade CO2 with only one sample from rosin weed revealed 4 mutations from clade CO1. These substitutions were transferred from A to T at 624 diagnostic nucleotide position in the 638bp COI segment, and deleted three codons A, T and A from 636-638 diagnostic nucleotide positions, respectively. Clade CO3 with only one sample from cup plant revealed one G-to-A substitution at 404 diagnostic nucleotide position from clade CO1. Clade CO4 comprised 3 samples from rosin weed and also revealed one mutation from clade CO1. It identified as a G-to-A substitution at 554 diagnostic nucleotide position. Clade CO5 comprised 3 samples from rosin weed and 2 samples from cup plant. It contained the same mutation as CO4 and also had another G-to-A substitution from clade CO1

Only three haplotypes were detected with 28S D2 sequences and they are separated from each other by only a single mutation: 28S1 separated from 28S2 and 28S3 by one single INDEL, with an extra T at nucleotide 181 position while the monotypic 28S3 is separated from 28S2 and 28S3 haplotypes by one single transition at 148 diagnostic nucleotide position – the 28S3 has "T" instead of "G". Based on comparison with outgroup *Hedickiana_levantina*, "T" at 148 position for sample R28 is

autoapomorphic while an insertion of a "T" at 181 position is synapomorphic for all samples associated with rosin weed except R28, which lacks the assumedly inserted "T" (Fig. 10). With regard to host plant association, 28S2 is exclusively associated with cup plant, while 28S1 and 28S3 are both exclusively associated with rosin weed (Fig. 9).

## Discussion

Phylogenetic trees can represent the phylogenetic relationships among the studied populations and phylogenetic reconstruction of barcode sequences can reveal susceptibility at the species level (Ndong et al 2015). Populations that share more novel molecular variations are more likely to belong the same species and therefore conclusions about species identity is dependent on the level of variation detected in used gene regions. Gene markers differ in their evolution rate and therefore are suitable for resolving phylogenetic relationships at different levels – the more conservative genes are more suitable to resolve relationships rooted in evolutionary time while the less conservative genes are more useful for resolving relationships that are relatively recent in evolutionary time (Simon et al 1994, Brower & DeSalle 1994). Because of the uncertainty as to whether the *Antistrophus* wasps associated with rosin weed and cup plant belong to the same or two different species, our team has chosen 4 relatively more variable gene makers (its1, its2, Cytb, and COI) that are considered and often used in resolving genealogical relationship at within species or at species level or slightly beyond (Simon et al 1994, Acs et 1010, Yang et al 2019). In addition, we also included a relatively conservative gene marker, the 28S rNRA nuclear gene D2 region (Brower & DeSalle 1994, Acs et al 2010), which may display limited variation among closely related species (Acs et al 2010, Schwéger et al 2015a, 2015b, Yang et al 2019). The variation of its1 and its2 genes is currently studied in a parallel sub-project and not yet clear at this stage. For the gene makers used in my thesis project, 28S D2 gene was only variable at two of the 628 molecular loci, which was not sufficient to provide any resolution about

phylogenetic relationship among the included sample sequences. This is expected given the that this is a conservative gene as discussed above. COI has been frequently used for DNA barcoding (Borsa et al 2009, Acs et al 2010, Kjer et al. 2014, Zhu et al., 2015, Zhou et al. 2016, Nicholls et al., 2018, Huang et. al, 2019, Yang et al., 2019). Phylogenetic reconstruction using the COI sequences of the samples included in this study resulted in only limited resolution; the only resolved relationship is a clade consisting of 8 samples, of which 6 were from rosin weed and 2 from cup plant (Fig. 4). The phylogenetic tree based on Cytb gene marker were able to reveal more detail about relationships among the sequences - most samples (63%) were shown to form a monophyletic clade, within which the small clades were supported. It is interesting to notice that all samples of one of the two smaller supported clades of the Cytb tree, are the same samples supported as a monophyletic clade on COI tree, except one extra sample in the COI tree (R4) (Fig. 5). Cytb, and especially COI, are commonly used for DNA barcoding, especially for Cynipoidea and other insect groups (Acs et al 2010, Kjer et al. 2014, Zhu et al., 2015, Zhou et al. 2016, Nicholls et al., 2018, Huang et. al, 2019, Yang et al., 2019). The extremely low variation of the two gene markers among our samples may indicate all samples belong to a single species. Nonetheless, the Cytb tree did indicate some genetic differentiation, all rosin weed associated samples, together with 6 samples, or 21.4% of all that were associated with cup plants, form a monophyletic clade (Fig. 5).

The fact that five haplotypes were identified among the Cytb and COI sequences while only three haplotypes were identified for 28S D2 sequences (Table 3) is consistent with the current literature in that 28S is more conservative than Cytb and COI (Bower & DeSalle 1994, Simon et al 1994, Acs et al 2010, Yang et al 2019). For the two relatively

variable genes COI and Cytb, the network analyses of haplotype genealogical relationship did not reveal any haplotype divergence associated with the two host plant species among COI haplotypes (Fig. 8) whereas a distinct pattern of haplotype divergence associated with host plant species was observed for the Cytb haplotypes - the Cytb haplotypes can be divided into two groups (Fig. 7). The first Cytb haplotype group consists of CTI, the most common haplotype and its two satellite haplotypes CT3 and CT5 and is predominantly associated with *S. integrifolium* (87%). The second Cytb haplotype group consists of CT2, the second most common haplotype, and CT4, and is completely associated *S. perfoliatum.* Although both COI and Cytb genes are frequently used for DNA barcoding, COI is generally considered to be more reliable in defining species, especially for insects and other invertebrate groups (Acs et al 2010, Kjer et al. 2014, Zhu et al., 2015, Zhou et al. 2016, Nicholls et al., 2018, Huang et. al, 2019, Yang et al., 2019). The lack of COI sequence divergence in correspondence with host plant species association might suggest that the populations of the *Antistrophus* gall wasps associated with the both host plant species belong to one single species. Nonetheless, the clear pattern of Cytb sequence divergence in correspondence with host plant species association strongly indicate the lack of gene flow between the populations of the *Antistrophus* gall wasps associated with the both host plant species belong to one single species, suggesting high level of reproduction isolation of the wasp populations associated with the two host plant species, and thus the existence of two distinct species, or highly isolated host races.

With regard to 28S D2 gene marker, it is not surprising to see the low variability among the sequences, and in fact, the gene marker was not much expected to be useful at

all. The limited, yet consistent variation of this gene among the *Antistrophus* wasp samples in the current study provides strong support, like the Cybt gene marker, for the hypothesis of the existence of two distinct species, especially considering the fact that 28S is highly conservative gene (Acs et al 2010, Yang et al 2019).

As 28S D2 gene is concerned, there are two equally plausible evolutionary scenarios concerning the genealogical relationship among the wasp samples, whether they are considered to belong to one or two distinct species. It is possible that the ancestral host association of these wasps was with rosin weed and a rosin weed in host to association with cup plant, accompanied by the 181 position indel, separated 28S1 from the ancestor of 28S2 and 28S3, which subsequently reversed back to rosin weed , accompanied with a G-to-T transition at 148 diagnostic nucleotide position. Alternatively, the ancestral host association of the complex was with cup plant, the transition occurred in 28S3, in accordance with host association change while the181 position indel event took place in accordance with host shift to rosin weed.

Plant volatiles as the most important parameter to identify the chemical defense of host association in the outer cell layers of galls especially the cue of secondary defense chemicals (Cornell 1983). A phylogenetic study based molecular data has shown that rosin weed and cup plant belong to a crown group with another *Silphium* species (Clevinger and Panero 2000). Therefore, the two species may share similar metabolic mechanisms to respond to stimuli by egg laying female as well as developing larva of any gall wasp species that may be able to form galls on any of the species. Therefore, errors by ovipositing females are likely to succeed, especially at high gall density (Larson & Ekbom 1985) and yield fertile offspring. When the emerging females prefer the original

host due to simply being used to the volatile blend characteristic of the original host and males rely on their olfaction to locate females (Tooker and Hanks 2004, Tooker et al 2005), assortative mating could occur, which would eventually lead to the formation of new species given enough time. The results of my study demonstrated that this is highly likely, although the exact mechanisms of genotype divergence with host association need to be examined using well control host selection experiments.

## Reference

Ács, Z., Challis, R., Bihari, P., Blaxter, M., Hayward, A., Melika, G., Csóka, Gy., Pénzes, Z., Pujade-Villar, J., Nieves-Aldrey, J.L., Schönrogge, K. & Stone, G.N. (2010). Phylogeny and DNA barcoding of inquiline oak gallwasps (Hymenoptera: Cynipidae) of the Western Palaearctic. *Molecular Phylogenetics and Evolution* **55**, 210–225. doi:10.1016/j.ympev.2009.12.004

Agosta, Salvatore J.; Jeffrey A. Klemens. (2008). Ecological fitting by phenotypically flexible genotypes: implications for species associations, community assembly and evolution. *Ecology Letters*. 11 (11): 1123–1134.

Agosta, Salvatore J.; J. A. Klemens. (2009). Resource specialization in a phytophagous insect: no evidence for genetically based performance trade-offs across hosts in the field or laboratory. *Journal of Evolutionary Biology*, 22: 907–912.

Ahmed, N., Darshanee, H. L. C., Khan, I. A., Zhang, Z-F., Liu, T.-X. (2019). Host Selection Behavior of the Green Peach Aphid, Myzus persicae, in Response to Volatile Organic Compounds and Nitrogen Contents of Cabbage Cultivars. *Frontiers in Plant Science*, 10:79.doi: 10.3389/fpls.2019.00079.

Askew, R. R. (1984). The biology of gall wasps. In: Ananthakrishnan, T. N. (Ed.). *Biology of Gall Insects*, pp. 223-271, Edward Arnold, London.

Avise, J.C. (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography,* vol.36, pp.3–15.

Ballard, W.J.O., and M.C. Whitlock. (2004). The incomplete natural history of mitochondria. *Molecular Ecology,* vol.13, pp.729–744.

Bassett, H.F. (1900). New species of North American Cynipidae. *Transactions of the American Entomological Society.* 26: 310-36.

Berlocher S. H. & Feder J.L. (2002). Sympatric speciation in phytophagous insects: moving beyond controversy? Annual Review of Entomology 47: 773–815.

Bernays, E.A. and R.F. Chapman. (1995). *Host-plant Selection by Phytophagous Insects.* Chapman and Hall, New York.

Beutenmueller. (1908). A new Cynipid from Arizona. *Journal of the New York Entomological Society.* 16: 45.

Bhargava, M. and A. Sharma. (2013). DNA barcoding in plants: Evolution and applications of *in silico* approaches and resources. *Molecular Phylogenetics and Evolution.* 67(3): 631-641.

Bronner R. (1992). The role of nutritive cells in the nutrition of cynipids and cecidomyiids. In: Shorthouse JD, Rohfritsch O, eds. Biology of insect-induced galls. Oxford: Oxford University Press, 118–140.

Brower, A. V. Z., and DeSalle, R. (1994). Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. Ann. Entomol. Soc. Am. 87: 702–716.

Burks, B. D. (1979). Superfamily Cynipoidea. In. K. V. Krombein et al, P. D. Hurd, Jr., D. R. Smith, and B. D. Burks (eds.), *Catalog of Hymenoptera in America north of Mexico*, vol.1, Symphyta and Apocrita, 1045-1107. Smithsonian Institution Press, Washington, DC.

Borsa, P., S. Lemer, W.J. Chen, A. Collet, A.Carassou, and D. Ponton. (2009). Identification par barcode (séquences nucléotidiques d'un fragment du gène ducytochrome *b*) des larves de Lethrinidae capturées en baies de Dumbéa et Ouano (Nouvelle-Calédonie). Programme d'évaluation des ressources marines de la zone économique de Nouvelle-Calédonie, IRD, UR 128, *Nouméa, Nouvelle-Calédonie*.

Clement, M, Posada, D, and Crandall KA. (2000). TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657–1660. Doi: 10.1046/j.1365-294x.2000.01020

Clevinger, J., & Panero, J. (2000). Phylogenetic Analysis of *Silphium* and Subtribe Engelmanniinae (Asteraceae: Heliantheae) Based on ITS and ETS Sequence Data. American Journal of Botany, 87(4), 565-572. Retrieved from http://www.jstor.org/stable/2656600

Cornell, J.L., Paicos, P. (1983). Downy woodpecker predation of goldenrod galls. *Journal of Field Ornithology.* 56: 56-64.

Cook, M.T. (1902). Galls and the insects producing them. *Ohio Nat.* II: 263-278.

Dawkins, R. (1982). *The extended phenotype*. Oxford University. Press, Oxford, U.K.

De Rojas, Manuel, Úbeda, José Manuel, Cutillas, Cristina, Mora, Ma. Dolores, Ariza, Concepción, and Guevara, Diego. (2007). Utility of ITS1–5.8S–ITS2 and 16S mitochondrial DNA sequences for species identification and phylogenetic inference within the Rhinonyssus coniventris species complex (Acari: Rhinonyssidae). Parasitology Research, 100: 1041–1046.

Detwiler, D., and J. Janovy. (2008). The role of phylogeny and ecology in experimental host specificity: insights from a eugregarine-host system. *Journal of Parasitology.*, vol.94(1), pp.7–12.

Dreger-Jauffret, F. and J. D. Shorhouse. (1992). *Diversity of gall-inducing insects and their galls.* In: Shorthouse JD, Rohfritsch O, eds. Biology of insect-induced galls. Oxford: Oxford University Press, 8–33.

Dre´s M. and Mallet J. (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transaction of the Royal Society of London B: Biological Sciences*, 357: 471–492.

Endara, M., Coley, P.D., Ghabash, G., Nicholls, J.A., Dexter, k.G., Donoso, D.A., Stone, G.N., Pennington, R.T. and Kursar, T.A. (2017). Coevolutionary arms race versus host defense chase in a tropical herbivore-plant system. *Proceedings of the National Academy of Sciences.* 114(36): 7499-7505. DOI: 10.1073/pnas.1707727114

Ehrlich, P.R. and P.H. Raven. (1964). Butterflies and plants: a study in coevolution. *Evolution* 18: 586–608.

Espírito-Santo, M. M. and G. W. Fernandes. (2007). How Many Species of Gall-Inducing Insects Are There on Earth, and Where Are They? *Annals of the Entomological Society of America*, 100: 95–99.

Fay, P. A. and D. C. Hartnett. (1991). Constraints on growth and allocation patterns of *Silphium integrifolium* (Asteraceae) caused by a cynipid gall wasp. *Oecologia.* 88: 243-250.

Fay, P. A., D. C. Hartnett, and A. K. Knapp. (1993). Increased photosynthesis and water potentials in *Silphium integrifolium* galled by cynipid wasps. *Oecologia.* 93: 114-120.

Fay, P.A., Hartnett, D.C., and Knapp, A.K. (1996). Plant tolerance of gall-insect attack and gall-insect performance. *Ecology*. 77(2): 521-534.

Fay, P.A. and Throop, H.L. (2005). Branching responses in *Silphium integrifolium* (Asteraceae) following mechanical or gall damage to apical meristems and neighbor removal. *American Journal of Botany.* 92(6): 954-959.

Fernandes GW and Price PW. (1992). The adaptive significance of insect gall distribution: survivorship of species in xeric and mesic habitats. *Oecologia.* 90: 14–20.

Folmer, O., Black, M., Wr, H., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial Cytochrome C oxidase subunit I from diverse metazoan invertebrates. Molecular marine biology and biotechnology. 3. 294-9.

Friedrich, M., and Tautz, D. (1997). Evolution and phylogeny of the Diptera: a molecular phylogenetic analysis using 28S rDNA sequences. *Systematic Biology.* 46(4): 674-698.

Funk JD, Filchak KE, and Feder JL. (1998). Herbivorous insects: model systems for the comparative study of speciation ecology. *Genetica* 116: 251-267.

Gillette C. P. (1891). Descriptions of new Cynipidae in the collection of the Illinois State Laboratory of Natural History. *Bulletin of the Illinois State Laboratory of Natural History.* 3: 191–197.

Gray, C.A, Runyon, J.B, Jenkins, M.J, Giunta A.D. (2015). Mountain Pine Beetles Use Volatile Cues to Locate Host Limber Pine and Avoid Non-Host Great Basin Bristlecone Pine. *PLoS ONE.* 10(9): e0135752.

Hancock, J. M., and G. A. Dover. (1988). Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. Mol. Biol. Evol. 5:377-392.

Hebert, P.D. & T.R. Gregory. (2005). The promise of DNA barcoding for taxonomy. *Systematic Biology*. 54(5): 852-859.

Hebert, P. D., Cywinska, A., and Ball, S.L. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological. Sciences*. 270(1512): 313-321.

Hebert, P.D., S. Ratnasingham, and J. R. deWaard. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*. 270: S96-S99.

Huang, T-F, Liu, Z., Gong, X-Y, Wu, T., Liu, H., Deng, J-X., Zhang, Y-X, Peng, Q-Z., Zhang, L-B., & Liu, Z-X. (2019). Vampire in the darkness: a new genus & species of land leech exclusively attacking cave-dwelling bats from China (Hirudinia: Arhyinchobdellida: Haemadipsidae). Zootaxa 4560 (2): 257–272. DOI: 10.11646/zootaxa.4560.2.2

Jermiin, L.S. and Crozier, R.H. (1994). The cytochrome *B* region in the mitochondrial DNA of the ant *Tetraponera rufoniger*—sequence divergence in Hymenoptera may be associated with nucleotide content. J Mol Evol 38:282–294

Ji, Y., L. Ashton, S.M. Pedley, D.P. Edwards, Y. Tang, A. Nakamura, et al. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters.* 16(10): 1245-1257.

Jiggins, C.D. (2008). Ecological speciation in mimetic butterflies. *BioScience* 58: 541–548.

Joy, J. B., and B. J. Crespi. (2007). Adaptive radiation of gall-inducing insects within a single host-plant species. *Evolution.* 61: 784-795.

Kjer KM, Zhou X, Frandsen PB, Thomas JA, Blahnik RJ. (2014). Moving toward species-level phylogeny using ribosomal DNA and COI barcodes: an example from the diverse caddisfly genus Chimarra (Trichoptera: Philopotamidae). Arthropod Syst. Phylogeny 72, 345–354.

Kress, W. J. and D. L. Erickson. (2008). DNA barcodes: genes, genomics and bioinformatics. *Proceedings of the National. Academy of Sciences of the United States of America.* 105(8): 2761-2762.

Kinsey, A.C. (1920). Phylogeny of Cynipid genera and biological characteristics. *Bulletin of the American Museum National History.* 42:357-402.

Kumar, Sudhir & Stecher, Glen & Li, Michael & Knyaz, Christina & Tamura, Koichiro. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular biology and evolution. 35. 10.1093/molbev/msy096.

Larsson, S. and B. Ekbom. (1995). Oviposition Mistakes in Herbivorous Insects: Confusion or a Step Towards a New Host Plant? *Oikos*. 72: 155-160.

León, G.P., Pinacho-Pinacho, C.D., Mendoza-Garfias, B., Choudhury, A., & García-Varela, M. (2016). Phylogenetic Analysis Using the 28S rRNA Gene Reveals That the Genus Paracreptotrema (Digenea: Allocreadiidae) Is Not Monophyletic; Description of Two New Genera and One New Species. The Journal of parasitology, 102:131-42.

Li, Y., X. I. N. Zhou, G. Feng, H. Hu, L. Niu, P. D. Hebert, and D. Huang. (2010). COI and ITS2 sequences delimit species, reveal cryptic taxa and host specificity of fig-associated *Sycophila.* (Hymenoptera, Eurytomidae). *Molecular Ecology Resources.* 10: 31-40.

Liljeblad, J. (2018). Hymenoptera Online (HOL). [Online] Available from https://hol.osu.edu/search.html?limit=50&name=Antistrophus. [accessed 27 November 2018].

Low V.L., Tan, T.K., Lim, P.E. et al. (2014). Use of COI, CytB and ND5 genes for intra- and inter-specific differentiation of *Haematobia irritans* and *Haematobia exigua*. *Veterinary Parasitology.* 204: 439-442.

McDaniel, T., Tosh, C. R., Gatehouse, A. M. R., George, D., Robson, M., and Brogan, B. (2016). Novel resistance mechanisms of a wild tomato against the glasshouse whitefly. *Agronomy for Sustainable Development.* 36:14. doi: 10.1007/s13593-016-0351-4.

Matsubayashi, K. W., I. Ohshima, and P. Nosil. (2010). Ecological speciation in phytophagous insects. *Entomologia Experimentalis et Applicata* 134:1–27.

Marquis RJ, Salazar D, Baer C, Reinhardt J, Priest G, Barnett K. (2016). Ode to Ehrlich and Raven or how herbivorous insects might drive plant speciation. *Ecology.* 97: 2939-2951. PMID 27870033 DOI: 10.1002/ecy.1534

McCracken, I., and Egbert, D. (1922). California gall-making Cynipidae with descriptions of new species. *Stanford University Publications, University Series. Biological Sciences*, 3(1): 1-70.

Meyer, J. (1987). *Plant galls and gall inducers.* Gebruder Brontraeger, Stuttgart, Germay.

Meyer, J., and H.J. Maresquelle. (1983). *Anatomie des galles. Handbuch der Pflanzenanatomie,* Volume XIII, 1. Gebruder Brontraeger, Stuttgart, Germay.

Miller, D. G., C. T. Ivey, and J. D. Shedd. (2009). Support for the microenvironment hypothesis for adaptive value of gall induction in the California gall wasp, *Andricus quercuscalifornicus. Entomologia Experimentalis et Applicata.* 132: 126-133.

Morrison, D.A. and Ellis, J.T. (1997). Effects of nucleotide sequence alignment on phylogeny estimation: A case study of 18S rDNAs of apicomplexa. *Molecular Biology and Evolution.* 14, 428-441.

Muths, D., E. Tessier, and J. Bourjea. (2008). Contribution de la génétique des populations à l'étude de la connectivité régionale de certaines espèces de poissons commerciaux présents dans la Réserve Naturelle Marine de La Réunion: étude de faisabilité. *Parc Marin, Réunion.* Pp.23.

Ndong, A., Thiaw, C., Diallo, B., Sarr, M., Diome, T., Kane, M., and Sembene, M. (2015). Barcoding – Comparison of variation degree of COI and Cytochrome b Markers in two species primary maize pests (*Sitophilus zeamais* and *Sitophilus oryzae*). *International Journal of Sciences: Basic and Applied Research.* 2307-4531.

Nicholls J. A., S. Preuss, A. Hayward, Melika, et al. (2010). Concordant phylogeography and cryptic speciation in two Western Palaearctic oak gall parasitoid species complexes. *Molecular Ecology.* 19: 592-609.

Nicholls, J., Melika, G., DeMartini, J. & Stone, G. (2018). New species of *Dryocosmus* Giraud gallwasps from California (Hymenoptera: Cynipidae: Cynipini) galling Chrysolepis Hjelmq. (Fagaceae). *Zootaxa.* 4532. 407. 10.11646/zootaxa.4532.3.6.

Nylander , J. A. A. (2002). MrModeltest v.1.0. Program distributed by the author. Department of Systematic Zoology, Uppsala University, Uppsala, Sweden. Website http://www.ebc.uu.se/systzoo/staff/ nylander.html.

Porter, T. M., & Hajibabaei, M. (2018). Over 2.5 million COI sequences in GenBank and growing. PloS one, 13(9), e0200177. doi:10.1371/journal.pone.0200177

Price PW and Pschorn-Walcher H. (1988). Are galling insects better protected against parasitoids than exposed feeders? A test using tenthredinid sawflies. *Ecological Entomology* 13: 195–205.

Raman, A., Schaefer, C. W., & Withers, T. M. (2005). *Galls and gall-inducing arthropods: An overview of their biology, ecology, and evolution.* In A. Raman, C. W. Schaefer, & T. M. Withers (Eds.), Biology, ecology, and evolution of gall-inducing arthropods (pp. 1–33). New Hampshire: Science Publishers, Inc.

Rambaut, A. (2012). Computer program distributed by the author, website: http://tree.bio.ed.ac.uk/software/figtree/ [accessed January 4, 2011]

Rogers, D. S. and M. W. Gonzalez. (2010). Phylogenetic relationships among spiny pocket mice (*Heteromys*) inferred from mitochondrial and nuclear sequence data. *Journal of Mammalogy.* 91(4): 914-930.

Roman, A., Schaefer, C.W., and Withers, T.M. (2005). *Biology, ecology, and evolution of gall-inducing arthropods.* Enfield, New Hampshire, Science publishers.

Ronquist, F. (1994). Evolution of parasitism among closely related species: phylogenetic relationships and the origin of inquilinism in gall wasps (Hymenoptera, Cynipidae). *Evolution.* 48: 241-266.

Ronquist F. Phylogeny, classification and evolution of the Cynipoidea. *Zoologica Scripta*. (1999). 28: 139–164.

Ronquist F, and Liljeblad J. (2001). Evolution of the gall wasp-host plant association. *Evolution*. 51: 2503–2522.

Ronquist, F., J. Nieves-Aldrey, M.L. Buffington, Z. Liu, J. Liljeblad, and J.A. Nylander. (2015). Phylogeny, evolution and classification of gall wasps: the plot thickens. *Plos one*. 10(5): 1-40.DOI: 10.1371/journal.pone.0123301

Ronquist, F., M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, and J.P. Huelsenbeck. (2012). MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space. Systematic Biology. 61:539-542.

Schoonhoven, L.M., T. Jermy, and J.J.A. van Loon. (1998). *Insect-plant biology*. Chapman &Hall, London, UK.

Schwéger, Sz. Melika, G. Tang, C.T. Bihari, P. Bozsó, M. Stone, G.N. Nicholls, J.A. & Pénzes, Z. (2015a). New species of cynipid inquilines of the genus *Synergus* (Hymenoptera: Cynipidae: Synergini) from the Eastern Palaearctic. *Zootaxa* **3999 (4):** 451–497. doi: 10.11646/zootaxa.3999.4.1

Schwéger, S. Melika, G. Tang, C.T. Yang, M.M. Stone, G.N. Nicholls, J.A. Sinclair, F. Hearn, J. Bozsó, M. & Pénzes, Zs. (2015b). New species of cynipid inquilines of the genus *Saphonecrus* (Hymenoptera: Cynipidae: Synergini) from the Eastern Palaearctic, with a re-appraisal of known species world-wide. *Zootaxa* **4054(1):** 1-84. doi: 10.11646/zootaxa.4054.1.1

Schwenk K, *et al.* (1998). Genetic markers, genealogies and biogeographic patterns in the Cladocera. *Aquat Ecol* **32**:37–51

Sheth, B. P. and V.S. Thaker. (2017). DNA barcoding and traditional taxonomy: an integrated approach for biodiversity conservation. *NRC Research Press*. 60: 618-628.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., and Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87: 651–701.

Stone, G.N., and J.M. Cook. (1998). The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. *Proceedings: Biological Sciences*. 265: 979-988.

Stern, D.L. (1995). Phylogenetic evidence that aphids, rather than plants, determine gall morphology. *Proceedings: Biological Sciences*. 260: 85-89.

Strauss, S.Y. and A.R. Zangerl. (2001). *Invertebrate-plant Interactions*. pp 77-106 In C. Herrera and O. Pellmyr (eds), Plant-Animal Interactions. Blackwell Science.

Swofford, D. L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Teletchea, F. (2010). After 7 years and 1000 citations: comparative assessment of the DNA barcoding and the DNA taxonomy proposals for taxonomists and non-taxonomists. *Mitochondrial DNA*. 21(6): 206-226.

Tooker J. F, Deans AR, and Hanks LM. (2004). Description of the *Antistrophus rufus* (Hymenoptera: Cynipidae) species com- plex, including two new species. *Journal of Hymenoptera Research.* 13: 125–133.

Tooker, J. F. and L. M. Hanks. (2004). Endophytic insect communities of two prairie perennials (Asteraceae: Silphium spp.). *Biodiversity and Conservation.* 13: 2551-2566.

Tooker, J. F., and L. M. Hanks. (2004). Stereochemistry of host plant monoterpenes as mate location cues for the gall wasp. *Journal of Chemical Ecology.*30: 473-477.

Tooker, J. F., A. L. Crumrin, and L. M. Hanks. (2005). Plant volatiles are behavioral cues for adult females of the gall wasp *Antistrophus rufus*. *Chemoecology.* 15: 85-88.

Vogler, A. P. and M. T. Monaghan. (2007). Recent advances in DNA taxonomy. *Journal of Zoological Systematics and Evolutionary Research.* 45(1):1-10.

van Noort, S. Stone, G.N. Whitehead, V.B. & Nieves-Aldrey, J.L. (2007). Biology of *Rhoophilus loewi* (Hymenoptera: Cynipoidea: Cynipidae), with implications for the evolution of inquilinism in gall wasps. *Biological Journal of the Linnaean Society* 90: 153-172.

Vivaldo, G., Masi, E., Taiti, C., Caldarelli, G., Mancuso, S. (2017). *The network of plants volatile organic compounds.* Scientific Reports, 7: 11050. DOI:10.1038/s41598-017-10975-x.

Wachi, N. Ide, T. & Abe, Y. (2011). A new inquiline species of *Saphonecrus* (Hymenoptera: Cynipidae: Synergini) associated with cecidomyiid galls on oak trees in Japan. *Annals of the Entomological Society of America* 104: 369-373.

Walsh, B.D., and Riley, C.V. (1870). Galls and their architects-2d article. *American Entomologist*, 2(3): 70-74.

Weis, A. E. (1996). Variable selection on Eurosta's gall size, III: Can an evolutionary response to selection be detected? *Journal of Evolutionary Biology*. 9: 623-640. doi:10.1046/j.1420-9101.1996.9050623.x.

Weis, A.E., and W.G. Abrahamson. (1986). Evolution of host-plant manipulation by gall makers: ecological and genetic factors in the solidago-eurosta system. *The American Naturalists.* 127: 681-695.

Weis, A.E., R. Walton, C.L. Crego. (1988). Reactive plant tissue sites and the population biology of gall makers. *Annual Review Entomol*ogy. 33: 467-486.

Whitham, T.G. (1979). Territorial behavior of *Pemphigus* gall aphids. *Nature*. 279: 324-325.

Williams, J. B., Joseph D. Shorthouse, Richard E., and Lee, Jr. (2002). Extreme resistance to desiccation and microclimate-related differences in cold-hardiness of gall wasps (Hymenoptera: Cynipidae) overwintering on roses in southern Canada. *Journal of Experimental Biology*. 205: 2115-2124.

Yang, X-H., Liu. Z., Pang, Y. Su, C-Y., Zhu, D. (2019). Three New Species of Synergine Wasps from Same Galls on from Lithocarpus glaber (Thunb.) in Hunan, China (Hymenoptera, Cynipidae). *Insect Systematics and Evolution* (*in press*).

Zhang X, Sun B, Tang Q, Chen R and Han S. (2019). Molecular Identification and Phylogenetic Analysis of Nuclear rDNA Sequences of Clonorchis sinensis Isolates From Human Fecal Samples in Heilongjiang Province, China. Front. Microbiol. 10:26. doi: 10.3389/fmicb.2019.00026

Zhou X et al. (2016). The Trichoptera barcode initiative: a strategy for generating a species-level Tree of Life. Phil. Trans. R. Soc. B 371: 20160025.doi: 10.1098/rstb.2016.0025

Zhu, D-H., Z. Liu, P-F. Lu. X. Yang, X-H Yang, C-Y Su, P. Liu. (2015). New Gall Wasp Species Attacking Chestnut Trees: Dryocosmus zhuili n. sp. (Hymenoptera: Cynipidae) on *Castanea henryi* from Southeastern China. *Journal of Insect Sciences.* 15: 156; DOI: 10.1093/jisesa/iev118


**Figure 1.** Samples locations from Wisconsin, Illinois, Indiana, Iowa and Missouri. Samples from rosin weed are represented by red pentagrams while samples from cup plant are represented by blue circles. The size of the shapes is proportional to the frequency of the samples.

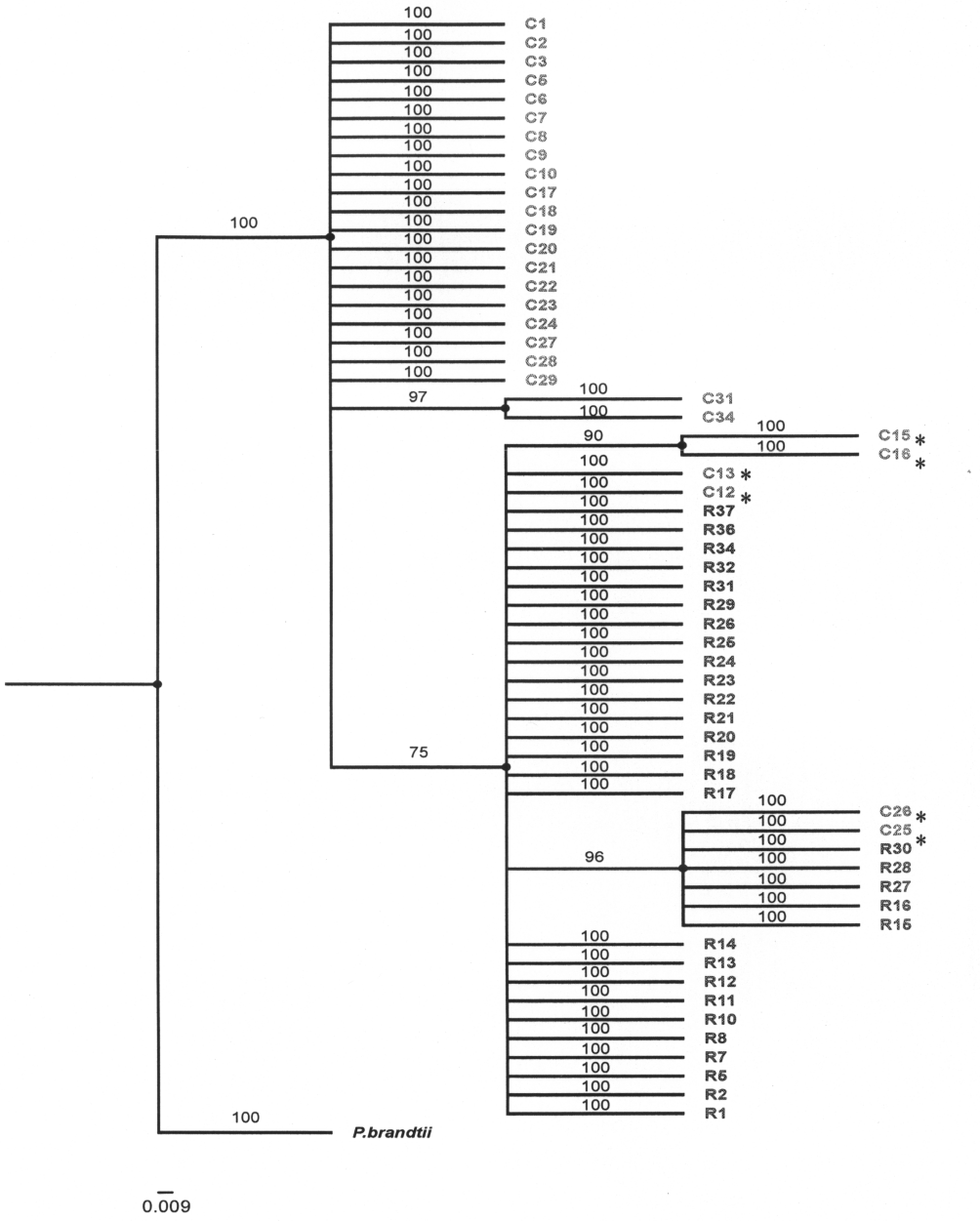


**Figure 2.** Bayesian majority rule consensus phylogenetic tree based on Cytb and rooted with outgroup *P. brandti*. Bayesian analysis assumed a GTR strict clock model of sequence evolution and was run for 1,000,000 generations with a 1/1,000 sampling frequency (with a total of 10,000 trees sampled) and a burnin time of 30%. Numbers on branches indicate estimated posterior clade probabilities (in % units). Samples starting with "C" are from cup plant while samples beginning with "R" are from rosin weed species. The samples with asterisk are those reared from cup plant, but nonetheless belong to the clade consisting mostly of samples from rosin weed.

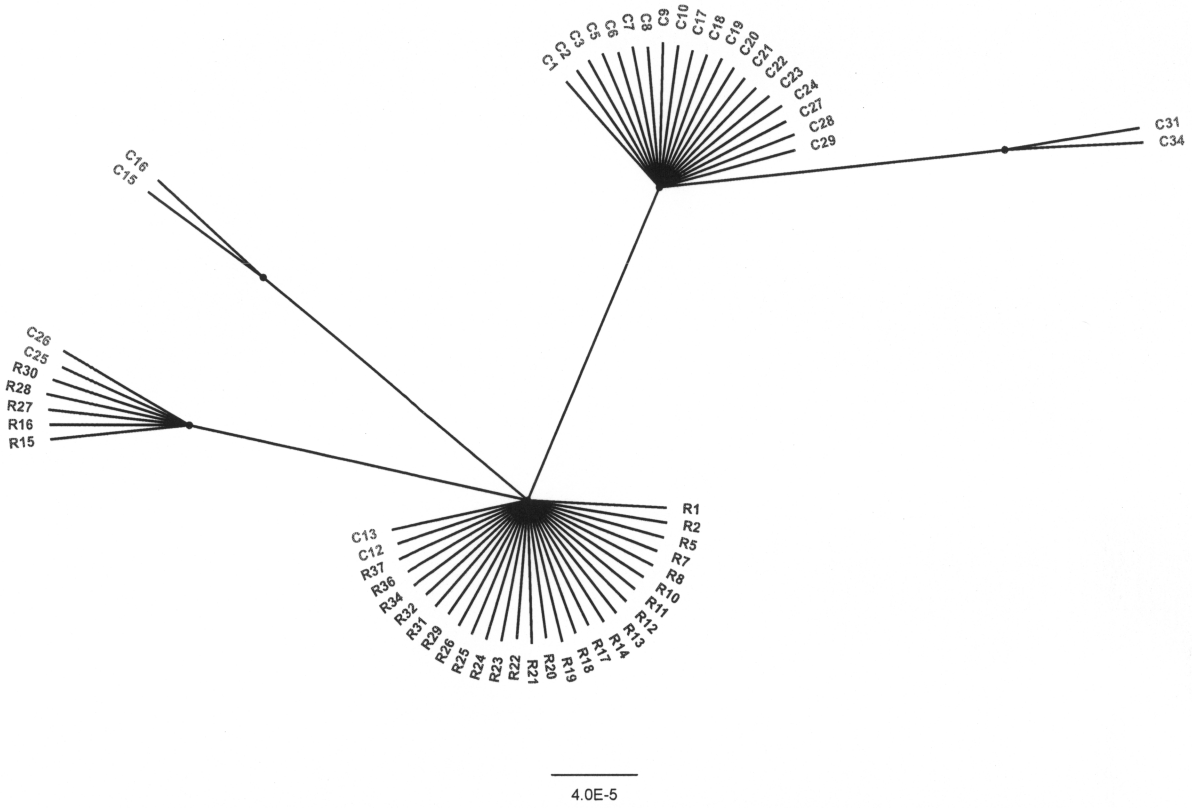


**Figure 3.** Bayesian majority rule consensus radiation tree based on Cytb. This is the same tree as in Figure 2, but presented in this form to show the genealogical relationships among the 5 identified haplotypes. Samples starting with "C" are from cup plant while samples beginning with "R" are from rosin weed species.

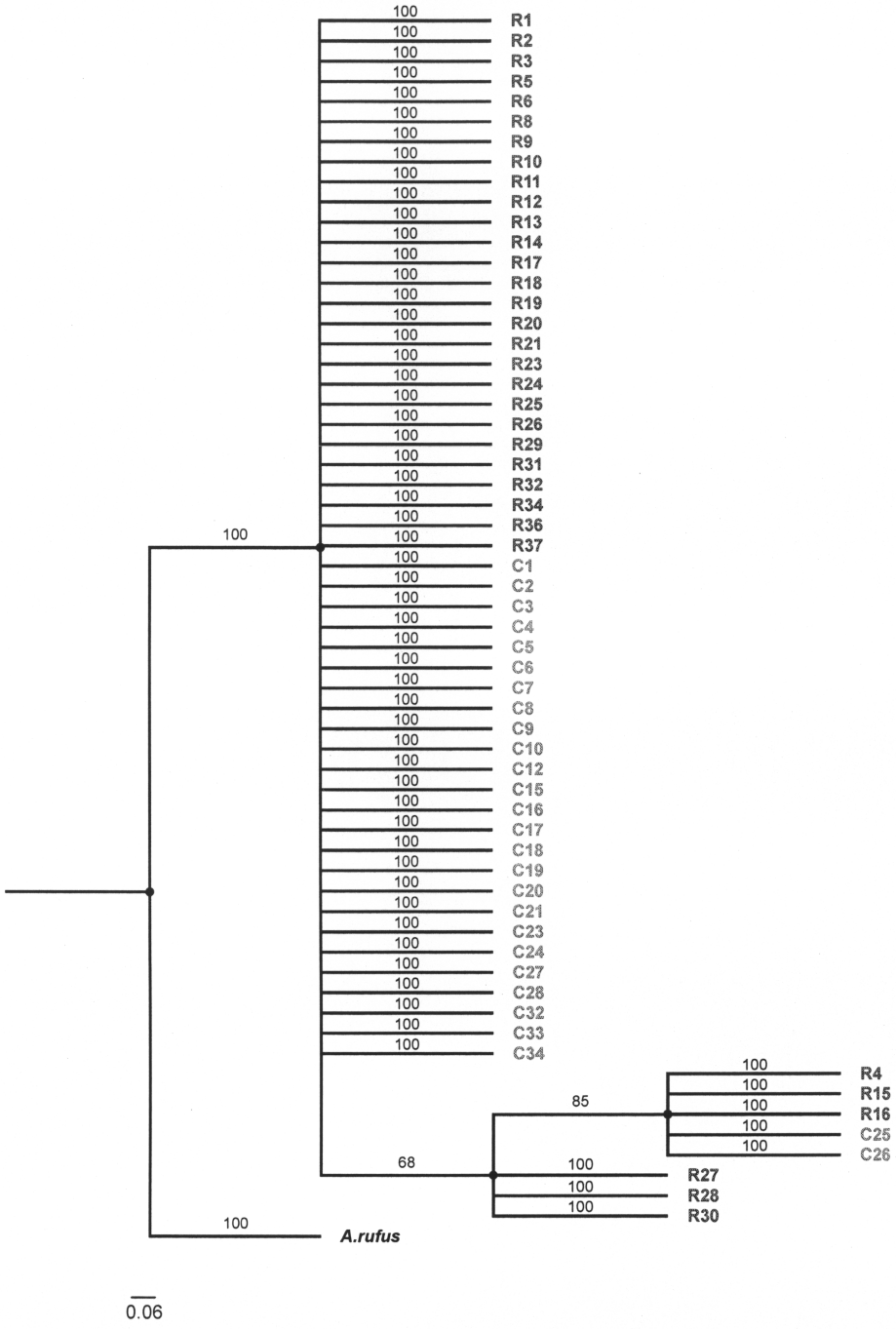


**Figure 4.** Bayesian majority rule consensus phylogenetic tree based on COI and with outgroup *A. rufus*. Bayesian analysis assumed a HKY model of sequence evolution and was run for 1,000,000 generations with a 1/1,000 sampling frequency (with a total of 10,000 trees sampled) and a burnin time of 30%. Numbers on branches indicate estimated posterior clade probabilities (in % units). Samples starting with "C" are from cup plant while samples beginning with "R" are from rosin weed species.

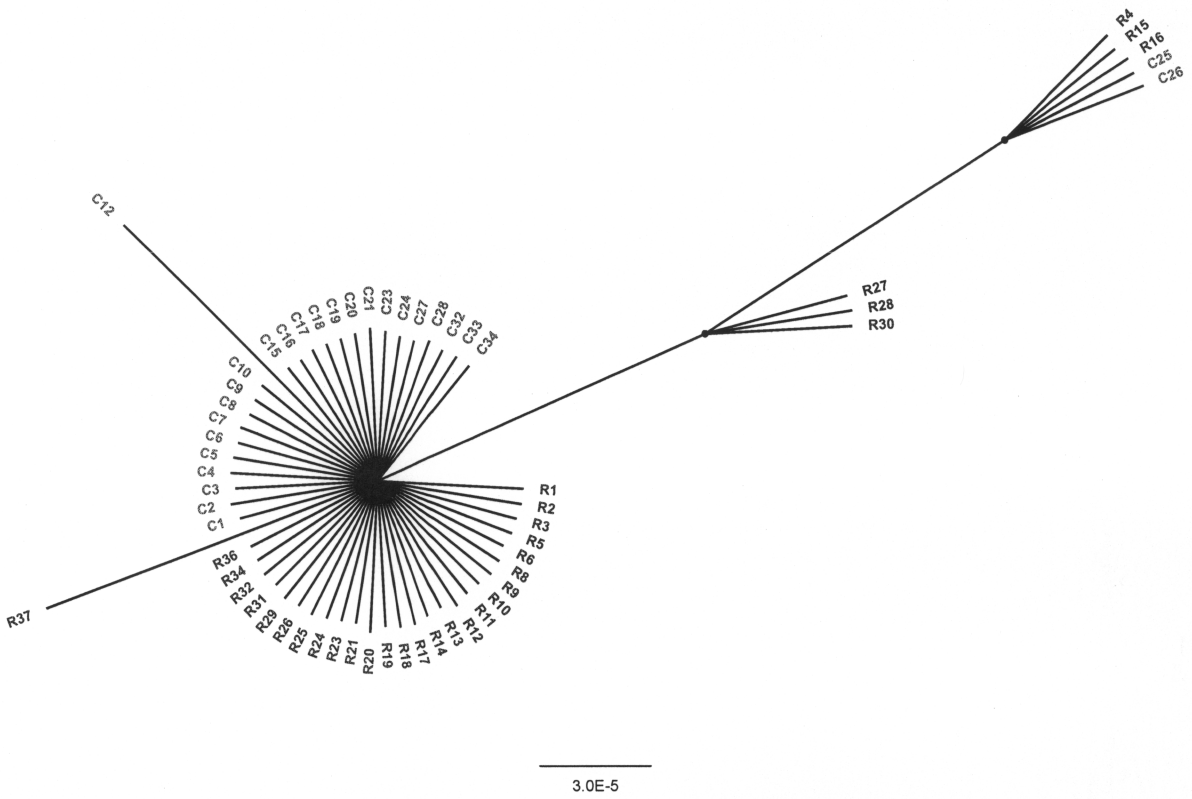


**Figure 5.** Bayesian majority rule consensus radiation tree based on COI. This is the same tree as in Figure 4, but presented in this form to show the genealogical relationships among the 5 identified haplotypes. Samples starting with "C" are from cup plant while samples beginning with "R" are from rosin weed species.

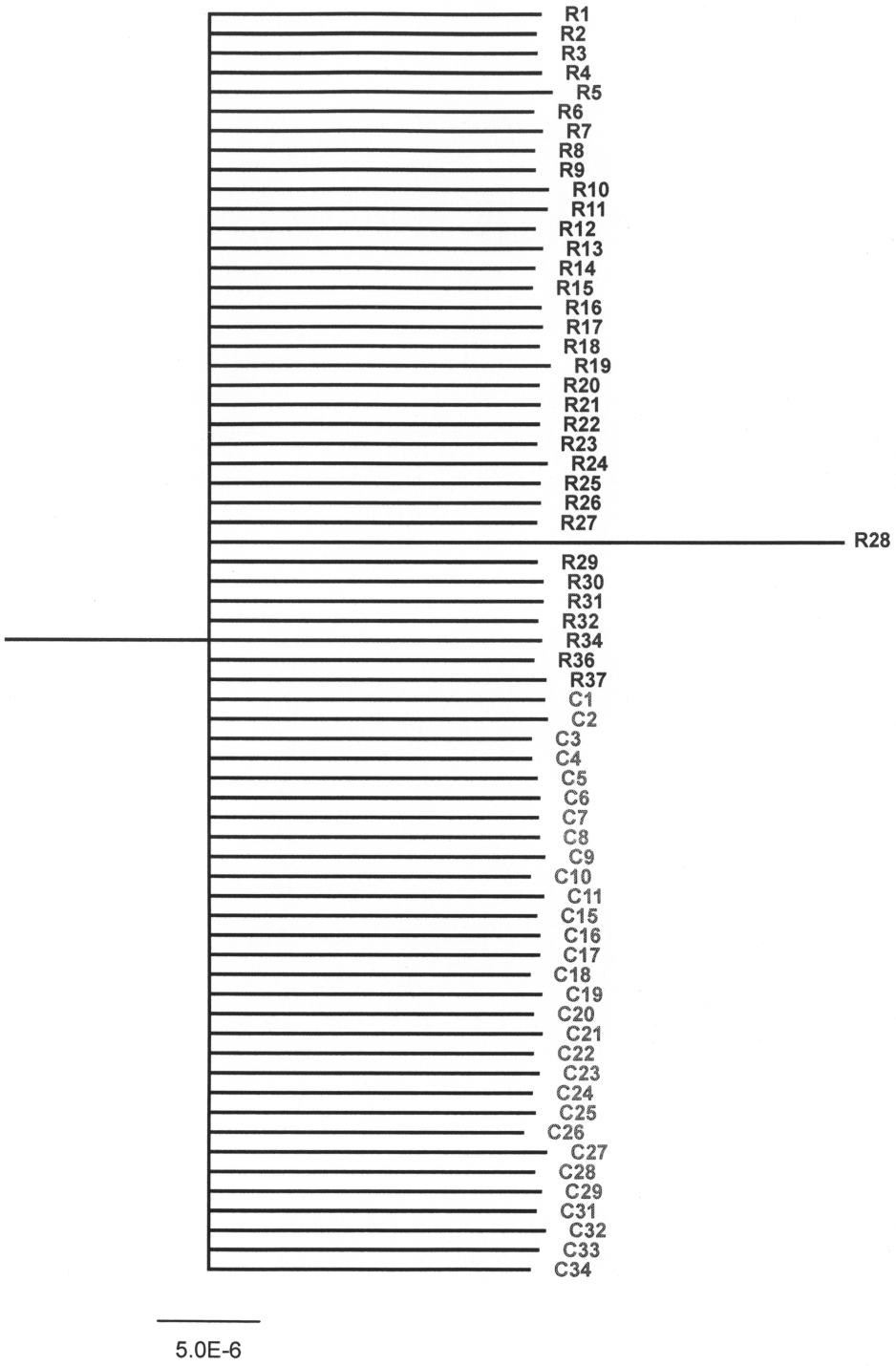


**Figure 6.** Bayesian majority rule consensus phylogenetic tree based on 28S D2. Bayesian analysis assumed an F81 model of sequence evolution and was run for 1,000,000 generations with a 1/1,000 sampling frequency (with a total of 10,000 trees sampled) and a burnin time of 30%. Samples starting with "C" are from cup plant while samples beginning with "R" are from rosin weed species.

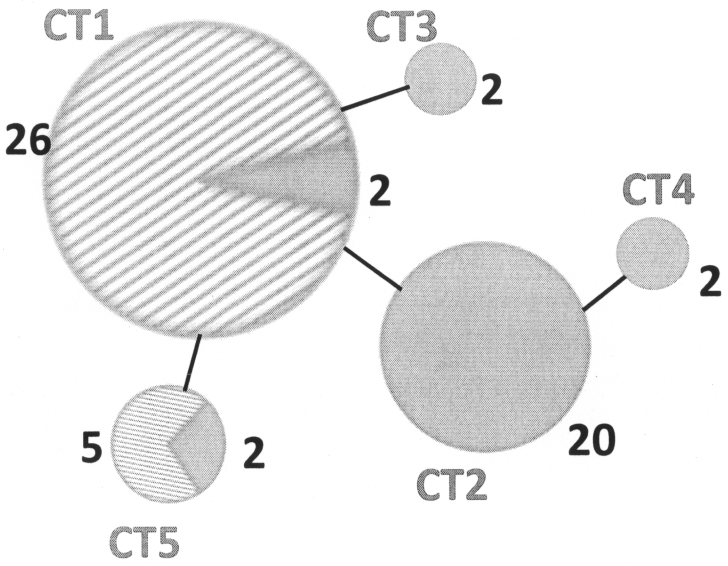


**Figure 7.** TCS network of Cytb haplotypes based on 95% connection limit. Each line segment represents a single mutation. The names of haplotypes begin with letters and the size of the circles is proportional to the frequency of the haplotypes. For each haplotype, the hatched area represents the fraction of samples in association with rosin weed while the solid area represents the fraction of samples in association with cup plant. The number next to a fraction of a circle represents the number of samples in that fraction.

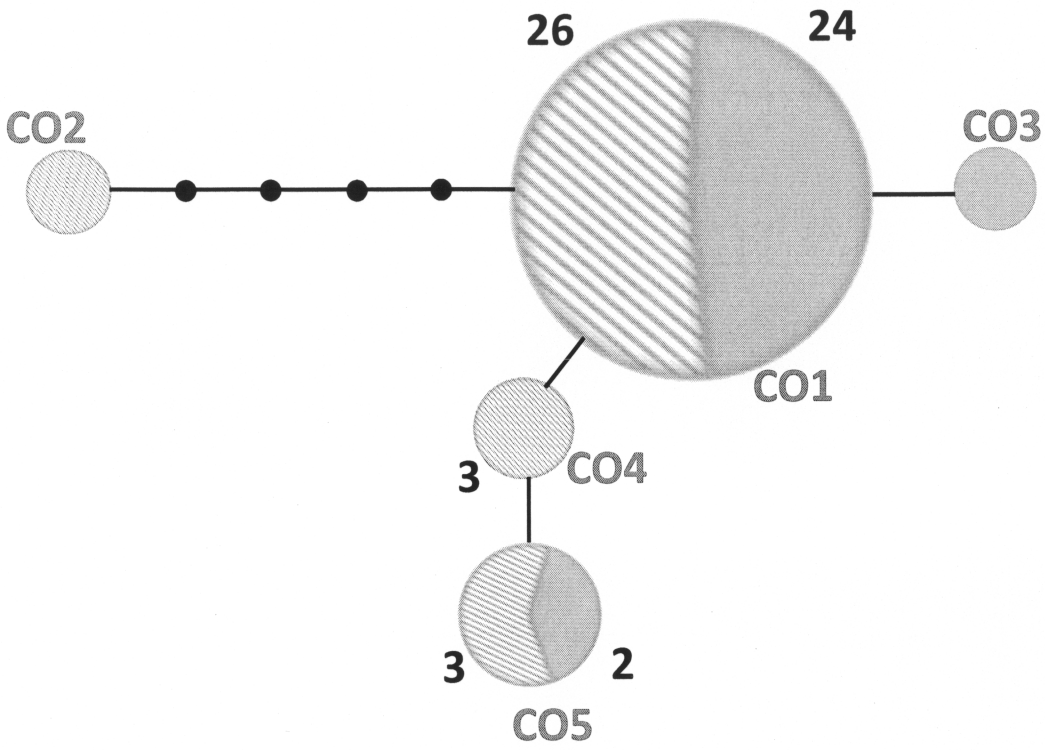


**Figure 8.** TCS network of COI haplotypes based on 95% connection limit. Each line segment represents a single mutation. The names of haplotypes begin with letters and the size of the circles is proportional to the frequency of the haplotypes. For each haplotype, the hatched area represents the fraction of samples in association with rosin weed while the solid area represents the fraction of samples in association with cup plant. The number next to a fraction of a circle represents the number of samples in that fraction.

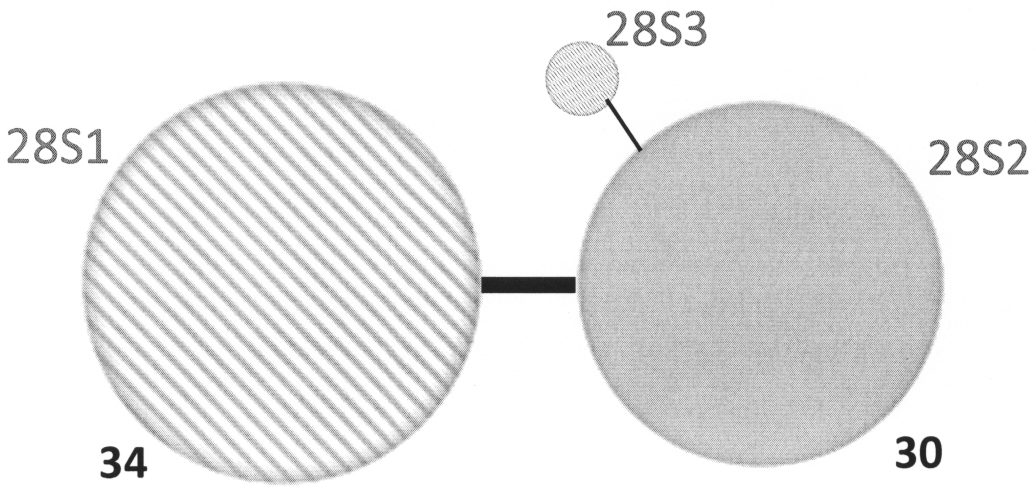


**Figure 9.** TCS network of 28S D2 haplotypes based on 95% connection limit. Each line segment represents a single mutation. The names of haplotypes consist of a letter and numbers while the size of the circles is proportional to the frequency of the haplotypes. For each haplotype, the hatched area represents the fraction of samples in association with rosin weed while the solid area represents the fraction of samples in association with cup plant. The number next to a fraction of a circle represents the number of samples in that fraction.

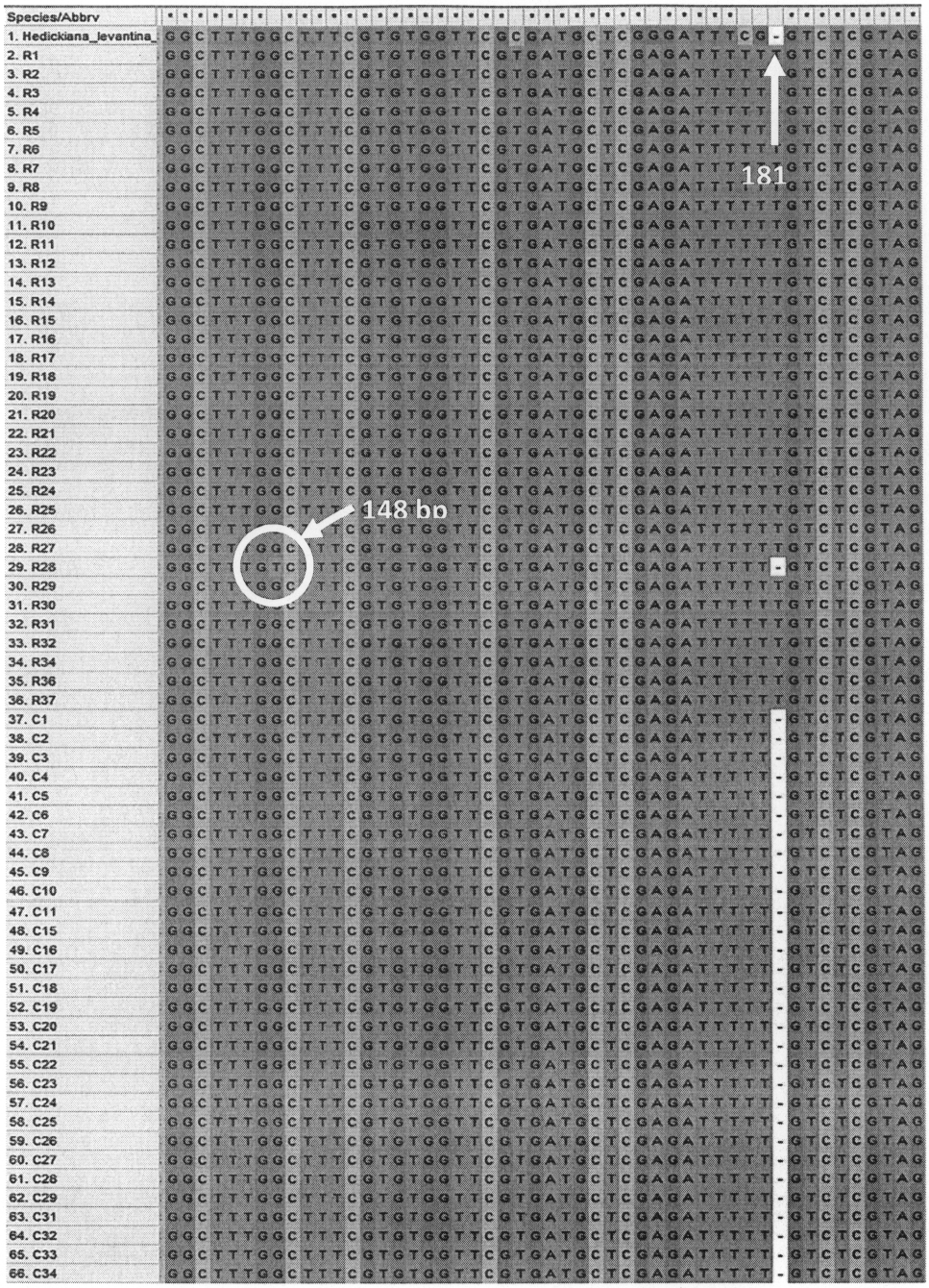


**Figure 10.** Nucleotides 141 to 190bp of the aligned 28S D2 sequences including that of the outgroup *Hedickiana levantina*. A single indel at 181 nucleotide position diagnostic nucleotide position easily separates the samples associated with rosin weed (with T at 181) from those associated with cup plant (with gap, missing T, at 181 position). The only exception was R28 – it is associated with rosin weed, but has a gap, missing T, at 181 position like those associated with cup plant. R28 is also unique in having T, rather than G at 148 nucleotide position.

**Table 1.** List of known *Antistrophus* species with data on known geographic distribution, host plant species, gall forming location on host plant, and source of data. The name of the *Silphium* genus is abbreviated as *S.* when the host plant(s) of a gall wasp species belong(s) to this genus.

| ***Antistrophus* species** | **Distribution area** | **Host plant** | **Gall forming sites** | **Reference** |
|---|---|---|---|---|
| *A. bicolor* | Illinois | *S. integrifolium* | Stems | Gillette 1891 |
| *A. chrysothamni* | Arizona | *Chrysothamnus* spp. | Stems | Beutenmueller 1908 |
| *A. laciniatus* | The Midwestern USA | *S. lacijiialutn* | Stems, flowers | Gillette 1891 |
| *A. lygodesmiarpisum* | The Midwestern USA | *Lygodemnia juncea* | Stems | Walsh 1870 |
| *A. microseris* | California | *Microseris* spp. | Stems | McCracken and Egbert 1922 |
| *A. minor* | The Midwestern USA | *S. laciniatum* | Stems | Gillette 1891 |
| *A. rufus* | The Midwestern USA | *S. laciniatum,*<br>*S. terebinthinaceum,*<br>*S. perfoliatum* | Stems, flowers | Gillette 1891 |
| *A. silphii* | The Midwestern USA | *S. integrifolium*<br>*S. perfoliatum* | Stems | Gillette 1891 |
| *A. jeanae* | The Midwestern USA | *S. terebinthinaceum* | Stems | Tooker and Hanks 2004 |
| *A. meganae* | The Midwestern USA | *S. perfoliatum* | Stems | Tooker and Hanks 2004 |
| *A. pisum (Invalid)* | The Midwestern USA | *Lygodemnia juncea* | Stems | Walsh 1870 |
| *A. leavenworthi (Invalid)* | Virginia | *Lactuca* spp. *& Mulgedium* spp. | Stalks | Bassett 1900 |

**Table 2.** The accession number of three gene markers from GenBank. The collecting place indicates as the state abbreviation format of each sample. The dash in the table represents the sample based on this gene marker does not have the successful DNA sequence. The sequences with asterisk have in-sequence stop codons and thus were not accepted by GenBank. Further work is needed to identify the issues and guarantee that these sequences become available to the public.

| Sample # | Collecting Place | COI | 28S D2 | Cytb |
|---|---|---|---|---|
| **R1** | IL | MK878665 | MK878631 | MK878697 |
| **R2** | IL | MK878666 | MK878632 | MK878698 |
| **R3** | IL | MK878667 | MK878633 | - |
| **R4** | IL | MK878668 | MK878634 | - |
| **R5** | IL | MK878669 | MK878635 | MK878699 |
| **R6** | IL | MK878670 | MK878636 | - |
| **R7** | WI | - | MK878637 | MK878700 |
| **R8** | WI | MK878671 | MK878638 | MK878701 |
| **R9** | WI | MK878672 | MK878639 | - |
| **R10** | WI | MK878673 | MK878640 | MK878702 |
| **R11** | WI | MK878674 | MK878641 | MK878703 |
| **R12** | WI | MK878675 | MK878642 | MK878704 |
| **R13** | WI | MK878676 | MK878643 | MK878705 |
| **R14** | WI | MK878677 | MK878644 | MK878706 |
| **R15** | IL | MK878678 | MK878645 | MK878707 |
| **R16** | IL | MK878679 | MK878646 | MK878708 |
| **R17** | IA | MK878680 | MK878647 | MK878709 |
| **R18** | IA | MK878681 | MK878648 | MK878710 |
| **R19** | IA | MK878682 | MK878649 | MK878711 |
| **R20** | IA | MK878683 | MK878650 | MK878712 |
| **R21** | IA | MK878684 | MK878651 | MK878713 |
| **R22** | IA | - | MK878652 | MK878714 |

| | | | | |
|---|---|---|---|---|
| **R23** | IA | MK878685 | MK878653 | MK878715 |
| **R24** | IA | MK878686 | MK878654 | MK878716 |
| **R25** | IA | MK878687 | MK878655 | MK878717 |
| **R26** | MO | MK878688 | MK878656 | MK878718 |
| **R27** | MO | MK878689 | MK878657 | MK878719 |
| **R28** | MO | MK878690 | * | MK878720 |
| **R29** | MO | MK878691 | MK878658 | MK878721 |
| **R30** | MO | MK878692 | MK878659 | MK878722 |
| **R31** | WI | MK878693 | MK878660 | MK878723 |
| **R32** | WI | MK878694 | MK878661 | MK878724 |
| **R34** | WI | MK878695 | MK878662 | MK878725 |
| **R36** | WI | MK878696 | MK878663 | * |
| **R37** | IA | * | MK878664 | * |
| **C1** | IN | * | * | * |
| **C2** | IN | * | * | MK876298 |
| **C3** | IN | * | * | MK876299 |
| **C4** | IN | * | * | - |
| **C5** | IN | * | * | MK876300 |
| **C6** | IN | * | * | MK876301 |
| **C7** | IN | * | * | MK876302 |
| **C8** | IN | * | * | MK876303 |
| **C9** | IN | * | * | MK876304 |
| **C10** | IN | * | * | MK876305 |
| **C11** | IL | - | * | - |
| **C12** | IL | * | - | MK876306 |
| **C13** | IL | - | - | MK876307 |
| **C14** | IL | - | - | - |
| **C15** | IA | * | * | MK876308 |
| **C16** | IA | * | * | MK876309 |

| | | | | |
|---|---|---|---|---|
| **C17** | WI | * | * | MK876310 |
| **C18** | WI | * | * | MK876311 |
| **C19** | WI | * | * | MK876312 |
| **C20** | WI | * | * | MK876313 |
| **C21** | WI | * | * | MK876314 |
| **C22** | WI | - | * | MK876315 |
| **C23** | WI | * | * | MK876316 |
| **C24** | WI | * | * | MK876317 |
| **C25** | IL | * | * | MK876318 |
| **C26** | IL | * | * | MK876319 |
| **C27** | IN | * | * | MK876320 |
| **C28** | IN | * | * | MK876321 |
| **C29** | IN | - | * | MK876322 |
| **C31** | IN | - | * | MK876323 |
| **C32** | IN | * | * | - |
| **C33** | IN | * | * | - |
| **C34** | IN | * | * | MK876324 |

**Table 3.** List of haplotypes of Cytb, COI, and 28S identified for the successfully sequenced samples and the samples belonging to each haplotype. Samples starting with "C" are from cup plant while samples italicized and beginning with "R" are from rosin weed. RW/CP is the ratio of the number of rosin weed associated samples to the number of cup plant associated samples.

| Gene Markers | Haplotype # | RW/ CP | Samples # |
|---|---|---|---|
| **Cytb** | CT1 | 26/2 | R1, R2, R5, R7, R8, R10, R11, R12, R13, R14, R17, R18, R19, R20, R21, R22, R23, R24, R25, R26, R29, R31, R32, R34, R36, R37, *C12. C13* |
| | CT2 | 0/20 | *C1, C2, C3, C5, C6, C7, C8, C9, C10, C17, C18, C19, C20, C21, C22, C23, C24, C27, C28, C29* |
| | CT3 | 0/2 | *C15, C16* |
| | CT4 | 0/2 | *C31, C34* |
| | CT5 | 5/2 | R15, R16, R27, R28, R30, *C25, C26* |
| **COI** | CO1 | 26/24 | R1, R2, R3, R5, R6, R8, R9, R10, R11, R12, R13, R14, R17, R18, R19, R20, R21, R23, R24, R25, R26, R29, R31, R32, R34, R36, R37, *C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C12, C15, C16, C17, C18, C19, C20, C21, C23, C24, C27, C28, C32, C33, C34* |
| | CO2 | 1/0 | R37 |
| | CO3 | 0/1 | *C12* |
| | CO4 | 3/0 | R27, R28, R30 |
| | CO5 | 3/2 | R4, R15, R16, *C25, C26* |
| **28S D2** | 28S1 | 35/0 | R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11, R12, R13, R14, R15, R16, R17, R18, R19, R20, R21, R22, R23, R24, R25, R26, R27, R29, R30, R31, R32, R34, R36, R37 |
| | 28S2 | 0/30 | *C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C31, C32, C33, C34* |
| | 28S3 | 1/0 | R28 |