2019

Pine, Aphids, and Parasitoid Wasps: Patterns of Cospeciation and Host Switches in a Tri-trophic System

Amber Bass
University of Central Florida

PINE, APHIDS, AND PARASITOID WASPS: PATTERNS OF COSPECIATION AND HOST SWITCHES IN A TRI-TROPHIC SYSTEM

by

AMBER IRENE HILDA BASS
B.S. University of Manitoba, 2015

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2019

ABSTRACT

Ecological interactions may drive speciation events, and the processes that drive these speciation events can leave behind patterns in the phylogenies of interacting taxa. These patterns have been studied extensively in herbivores and host plants, as well as parasites and their hosts, but rarely in tri-trophic systems. Here, we examine three closely related groups of interacting taxa, including parasitoid wasps (*Pauesia*), aphid herbivores (*Cinara*), and pine trees (*Pinus*) to determine if the patterns between each interacting taxa indicate that cospeciation or host switches are more dominant. We create phylogenies of *Cinara* and *Pauesia* in the southeastern United States using ddRADseq data and analyze publicly available data for *Pinus*. Most *Cinara* and *Pauesia* were specialized, with no species utilizing more than three hosts, indicating that this system is well suited to cophylogenetic study, and host interactions likely play a role in the speciation of these taxa. *Pauesia* was slightly more specialized on *Pinus*, suggesting phytochemistry may constrain the host breadth of these wasps and lead to coevolutionary patterns between *Pauesia* and *Pinus*. Distance-based cophylogenetic analyses suggest that aphids and pine, and wasps and aphids have dependent phylogenies, but these analyses differ in regards to wasps and pine. However, event-based methods show that cospeciation events and host switches both present, often in nearly equal proportions, and duplications and sorting events occurred at a lower frequency if at all. Both *Cinara* and *Pauesia* require revisions and the development of updated taxonomic resources for identification. This system presents an ideal model group to study coevolutionary patterns and multi-trophic community dynamics across macro- and microevolutionary time scales.

This thesis is dedicated to those who sparked my interest in Entomology and pushed me to pursue it. First, my mother, Heather Bass, for spending so much time with me, walking around the forests and prairies of my childhood home, watching wildlife, discussing the diversity of plant and animal life, and collecting beetles and butterflies. Second, my best friend, Heather Mitchell, who convinced me to pursue a minor in Entomology at the University of Manitoba. Without her, I would have missed out on so many incredible experiences and opportunities. Finally, for my husband, Ryan Jones, who continuously pushed me to keep going when coffee was not enough to make it through the frustration and tedium of graduate school.

ACKNOWLEDGMENTS

Many people have helped me throughout my time as a graduate student at UCF. First, I would like to thank every graduate student, undergraduate, volunteer, technician, and post doc from the Sharanowski lab. Each of you have contributed in some way to my successful completion of my thesis. Most notably are Miles Zhang, Ryan Ridenbaugh, and Joshua Hogan. Miles was involved in developing the ddRADseq protocol that was used for this research and has helped with many other aspects of my work. Miles was one of my first science mentors while I was still an undergraduate and has always made himself available to give advice or a helpful hand whenever I need it throughout the process. He has been an instrumental part of my success as a graduate student. Ryan was a member of the three-person team who collected wasps and aphids for this study. He stayed nights on end in tents in areas with no cell service, drove for hours, scaled mountains, and put up with me stressing on particularly unsuccessful days to ensure that I had enough samples for my project. Joshua Hogan was my undergraduate mentee and assisted with many small and tedious tasks that saved me a lot of time. He was a pleasure to mentor, a great help, and a wonderful conversationalist.

I was so lucky that Lauren Caspers and Havalend Steinmuller accepted me as their friends. They have provided an outlet for my sorrows on the bad days, and a place to laugh and tell stories on the good ones. Their friendship was so important to me during my graduate degree.

Ryan Jones, my husband, took the sacrifice of putting a halt on his career to move to Florida with me as I studied at UCF. He is my biggest supporter and put in overtime to help with many things behind the scenes during my degree. Without him, I certainly would have been eating ramen noodles every day.

Ken Fedorka introduced me to the University of Central Florida and was a member of my committee. I am very thankful for him welcoming me to the department, providing valuable advice about my research, and being an advocate for graduate students.

Eric Hoffman was a member of my committee, and besides my own I spent the most time in his lab. He was an avid supplier of food and advice and an excellent teacher. I am thankful for the time he dedicated to help me prepare for presentations and provide career advice. I learned a lot from him.

Finally, my advisor, Barb Sharanowski, has been the most important influence in my scientific career. She saw potential in me as an undergraduate and offered me a position in her lab. She has never wavered from being extremely supportive of me, even of my least productive days. When imposter syndrome had me trapped, she always knew what to say to bring me out of it and give me the motivation and confidence to keep going. With her guidance, I have transitioned from a shy, insecure undergraduate into an outgoing and confident professional. In addition to being an advisor, and I am very thankful for the time I have spent learning from her.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Understanding host use patterns and community dynamics on macroevolutionary timescales are important components to assessing how biodiversity is generated (Ehrlich and Raven, 1964; Janz et al., 2006; Mitter et al., 1991; Thompson, 1999). There is clear evidence that ecological processes can drive speciation (Matsubayashi et al., 2010; Schluter, 2000), which has been demonstrated for numerous herbivorous insects (Farrell, 1998; Funk et al., 2002; Janz et al., 2006; Nyman et al., 2006) and interactions with other symbiotic taxa, such as mutualists (Clark et al., 2000), natural enemies and prey (Abrahamson and Blair, 2008; Forbes et al., 2009; Hamerlinck et al., 2016; Stireman et al., 2006), or parasites and hosts (Hoberg and Brooks, 2008; Ricklefs and Fallon, 2002). These ecological processes leave behind macroevolutionary patterns in paired phylogenies, the most apparent of which is congruent phylogenies where links between interacting taxa do not cross over one another (Page, 2003). This pattern is present in cospeciating taxa, where divergence and speciation events are shared between symbiotic taxa (Clark et al., 2000; Hosokawa et al., 2006). Caution must be used when inferring processes from these patterns because different processes can result in similar macroevolutionary patterns. For example, cospeciation can be the result of multiple processes including coevolution, co-vicariance, and phylogenetic tracking (Althoff et al., 2014). Other patterns are less evident, involving incongruent trees with linkages between terminal taxa that cross over one another, and include the following evolutionary events: a switch to a novel and unrelated host (host-switching), a symbiont not speciating when a

host does, or going extinct on a host lineage (sorting), a symbiont speciating without a host undergoing a speciation event (duplication) (Page, 2003).

The relative importance of cospeciation, host switching, sorting, or duplication is likely to differ depending on the interacting taxa that are examined. Taxa with high host fidelity that are intimately associated with their hosts are more likely to demonstrate the macroevolutionary pattern of cospeciation (Bernays and Graham, 1988; Page, 2003). For example, phylogenetic patterns of lice often show high levels of cospeciation with avian (Hughes et al., 2007) or mammalian hosts (Demastes and Hafner, 1993). Herbivores and host plants have also been commonly studied. A classic herbivore-host systems that consistently displays high levels of cospeciation is yucca moths and Joshua tree plants (Smith et al., 2008a).

However, just as often as cospeciation is the dominant pattern, so too host switches are common patterns in similar systems. Host switches occur when a symbiont (parasite or mutualist) accepts a new host, thus becoming reproductively isolated from the population on the ancestral host, and over time divergence continues to the point of speciation (Drès and Mallet, 2002). New hosts can be accepted if: (1) the host shares recognition and defensive traits with the ancestral hosts; (2) the host has been used in the past, the symbiont having a genetic memory of that host; or (3) the symbiont fortuitously possesses capabilities to use the host as a novel resource (Agosta et al., 2010). Therefore, in specialists these switches are often constrained to closely related species that share traits that will be recognized (Agosta et al., 2010; Janz et al., 2001). Systems that have shown host switches as the dominant pattern include nematode parasites and their beetle

hosts (Mayer et al., 2009), and viruses (Charleston and Robertson, 2002; Gottschling et al., 2011). Evidence for host-switching as a dominant factor in phytophagous insect diversification is less compelling (Futuyma and Agrawal, 2009; Winkler and Mitter, 2008), but it may be that taxa go through episodic periods of host-switching followed by an adaptive radiation on specific host taxa (Hoberg and Brooks, 2008; Janz and Nylin, 2008)

Price et al. (1980), when referring to plant host-herbivore systems, stressed the importance of including predators and parasitoids in tri-trophic studies, as parasitoids interact indirectly with host plants, and thus are integral to understanding host-herbivore interactions. If cospeciation or host-switching is the dominant pattern in herbivorous insects, does the same pattern extend to their parasitoids? Recently, it has been demonstrated that the parasitoids of herbivores may diversify in a similar manner to their host insects, causing cascading diversification (Feder and Forbes, 2010; Forbes et al., 2009; Stireman et al., 2006). However, most studies have been limited to one or a few species (Althoff, 2008; Forbes et al., 2009; Stireman et al., 2006) or examine a suite of mostly unrelated parasitoids (Leppänen et al., 2013; Nyman et al., 2007), making evolutionary inferences on cascading co-speciation and host switching in the third trophic level less valid. Macroevolutionary studies that do investigate related and specialized parasitoids have found mixed patterns, where no one pattern dominates (Deng et al., 2013; Hall et al., 2017; Hamerlinck et al., 2016) or where there is a dissimilar pattern between herbivores and plants relative to herbivores and their parasitoids (Peralta et al., 2015; Wilson et al., 2012). In this study, we investigate macroevolutionary patterns of

cladogenesis between closely related taxa with high-host fidelity across three trophic levels: pine trees (*Pinus* spp.), aphid herbivores (*Cinara* spp.), and their parasitoids (*Pauesia* spp.).

Pinus is the only genus in the subfamily Pinoideae, within family Pinaceae. Species relationships were recently described in an eight gene phylogeny representing 115 species (Saladin et al., 2017). *Pinus* is split into two subgenera, *Pinus* and *Strobus,* that are well supported morphologically and phylogenetically (Gernandt et al., 2005; Leslie et al., 2012; Syring et al., 2005). In North America, there are 36 native *Pinus* species (Kershner et al., 2008).

Cinara (Hemiptera: Aphididae: Lachininae) is the second largest genus of aphids with over 200 described species (Favret, 2013), approximately 150 of which are North American (Blackman and Eastop, 2000). The genus is likely much more diverse than current descriptions; a barcoding survey in southeast China found 94 candidate species across all conifer hosts (Chen et al., 2016b). Of these, only 13.8% were previously described (Chen et al., 2016b). *Schizolachnus*, a genus of needle feeding conifer aphids consistently falls within *Cinara* in all phylogenetic reconstructions (Chen et al., 2016a; Meseguer et al., 2015b). *Cinara*, excluding *Schizolachnus*, are found on roots, shoots, trunks, and branches of eight plant genera from two families: Pinaceae and Cupressaceae (Blackman and Eastop, 2000).

The role of host use in shaping the patterns of speciation has previously been studied in *Cinara*. Studies focused on this topic have examined *Cinara* on all conifers and looked at switches between conifer genera (Jousselin et al., 2013; Meseguer et al.,

2015a). These studies report that allopatric speciation and host switches to novel conifer genera were the most dominant drivers of speciation in the evolutionary history of *Cinara* (Jousselin et al., 2013; Meseguer et al., 2015a). However, focusing on conifer genera rather than species may not accurately represent evolutionary dynamics of diversification of herbivores. Another study focused in on species of *Cinara* that specialize on *Pinus monophylla* and *Pinus edulis*; the authors determined that host switches have been an important driver in speciation for these aphids (Favret and Voegtlin, 2004b).

Pauesia spp. (Hymenoptera: Braconidae: Aphidiinae) are koinobiont endoparasitoids of aphids. Host records of the 21 North American species range from one to five host species primarily within *Cinara*, but other records exist in the genera *Schizolachnus* and *Lachnus* and one moth that is certainly a misidentification of another braconid wasp (Yu et al., 2012). *Schizolachnus* has been supported as a subgenus of *Cinara* in previous phylogenies (Chen et al., 2016a; Meseguer et al., 2015a), and *Lachnus* was previously applied to *Cinara* aphids (Blackman and Eastop, 2000), thus *Pauesia* are specialists of *Cinara*. Taxonomic changes in this group have primarily been based on morphology (Pike and Starý, 1996; Pike et al., 2002; Pike et al., 1996), with one phylogenetic study revealing many of these described species are invalid under the phylogenetic species concept (Sanchis et al., 2001). There are 99 *Pauesia* species described worldwide, 21 of which have been recorded in North America (Yu et al., 2012). Given the low number of *Cinara* species in North America, there are likely many more undescribed species in this genus.

Two types of symbiotic relationships exist between the three taxa included in this study. First, parasitism exists between both wasps and aphids and aphids and pine. In these interactions, the parasite (wasp or aphid) benefits, and the host (aphid or pine) is harmed. Second, a potential mutualistic relationship exists between *Pauesia* and *Pinus*, where both parties benefit from the interaction. Many parasitoids locate their hosts through semiochemical cues released by plants in response to herbivore feeding (Tumlinson et al., 1993; Tumlinson et al.). In this relationship, the parasitoid benefits by receiving access to its host while *Pinus* benefits from a reduced parasite load. Thus, specialized parasitoids may be constrained in host use by related host plants as much as their herbivore hosts.

The main objective of this study is to determine the relative dominance of cospeciation versus host switching patterns between 1) aphids and wasps, 2) aphids and pines, and 3) wasps and pines. Using cophylogenetic analyses and NGS data, we assess whether these parasitoid wasps cospeciate more with their host aphids or host plants and whether these patterns cascade, such that parasitoids will show similar diversification patterns on their aphid hosts as do the aphids on their hosts plants. We also investigate the host breadth of both aphids and wasps and indicate the need for revisionary work in these taxa.

METHODS

Specimen Collection

Southwestern United States supports a high diversity of all taxa in our system, so sampling was conducted in June of 2017 in the following states: California, Arizona, New Mexico, Colorado, and Utah. For example, fourteen of the 21 American species of *Pauesia* have been recorded in western USA (Yu et al., 2012), and 22 of the 36 naturally pine species are found in western USA (Kershner et al., 2008). The branches, shoots, needles, and trunks of each pine tree encountered were searched for aphid colonies. Ten aphids were collected from each colony and immediately stored in 95% ethanol and the site of location recorded. Parasitized aphids were stored individually in 1.5 ml microcentrifuge tubes and checked daily for emergence. Upon emergence, adult wasps were placed in 95% ethanol. If *Pauesia* did not emerge from their hosts as adults, wasp larvae or pupae were manually removed from *Cinara* mummies. *Pauesia* used in subsequent analyses were collected in California, Colorado, and New Mexico, while *Cinara* were also collected in Utah and Arizona (Fig. 1, Fig. 2). All specimens were deposited in the University of Central Florida Collection of Arthropods (UCFC).

All aphid colonies that had successfully emerged wasps were included in DNA extraction. In addition, aphids from unique localities and on different pine species were included to have a more complete phylogeny of this group. Both taxa were non-destructively extracted, besides an incision made along the ventral length of each aphid as previously described (Favret, 2005) to ensure the body was preserved for identification.

The DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) was used to extract DNA with a modified protocol to increase the DNA yield (Dal Molin and Menard, 2010). DNA concentrations were tested using a Qubit™ dsDNA high sensitivity assay kit.

Morphological and DNA Barcoding Specimen Identification

Both morphological and genetic methods were utilized to assign specimens to previously described species or putative new species. Each group in this tri-trophic system were handled differently. Pine trees were identified to species using morphology (Kershner et al., 2008). Morphological characters used to distinguish between species included the size of the tree, the number of needles per bundle, the length of the needles, the size and shape of the cones, the color and texture of the bark, along with the geographical distributions. The only two species that were difficult to distinguish in some instances were *Pinus ponderosa* and *Pinus jeffreyi*, as they share many characteristics and can hybridize (Haller, 1962). Instances in which the tree could be either *P. ponderosa* or *P. jeffreyi* were recorded as both species (*ponderosa/jeffreyi*).

Identification keys for *Cinara* are only available for adult females (Bradley, 1951; Favret and Voegtlin, 2014). Thus, nymphs were identified using the barcoding region of *COI*, which were amplified using universal primers LCO1490 and HCO2198 (Folmer et al., 1994) following the protocol in Namin et al. (2014). Amplicons were purified and bi-directionally sequenced at UK Healthcare Genomics Core Laboratory using the BigDye Terminator Cycle Sequencing Kit. Sequences were assembled in Geneious v11.1.4 (Kearse et al., 2012).

Two methods were utilized to identify putative species. First, *COI* sequences were compared to barcode records on BOLD, the Barcode of Life Database (Ratnasingham and Hebert, 2007) using species level barcode records. Some taxa were confidently assigned to one barcode index number (BIN) and fell in a monophyletic clade in the neighbor joining tree output in BOLD and thus could be confidently assigned a species epithet (e.g. *C. brevispinosa*, Table 1). However, several taxa could not be confidently placed within a BIN (Table 1) and were labeled as near (nr.) the best matching species (Table 1). Some species names also fell into two distinct BINS (e.g. *C. ponderosae*, ABY4171 and AAI3985, Table 1), suggesting that these species need revisionary work. In this case, species were labeled with the species epithet but given numbers to differentiate taxa that fell into distinct BINS and monophyletic clades (e.g. *C. ponderosae* 1). In the second method, we utilized the Automatic Barcode Gap Discovery (ABGD) algorithm, which recursively partitions taxa into putative species based on prespecified interspecific genetic distance gap and a range of intraspecific divergences (Puillandre et al., 2012). Default settings were used except distances were calculated under a Kimura 2 parameter model (Kimura, 1980) and 20 steps. Both methods were compared with the phylogenetic results from the ddRADSeq data (discussed below) and ecological data from each specimen to determine the most likely number of putative species.

Morphological identification of species of *Pauesia* was not possible because: (1) existing keys do not contain all currently described species (Pike et al., 2002; Smith, 1944); (2) morphological characters are subjective, such as broad versus narrow, and require slide mounting (Pike et al., 2002); and (3) as with most parasitoid Hymenoptera,

likely many species remain undescribed (Dolphin and Quicke, 2001; Forbes et al., 2018). Bold identifications were also not possible because of the 97 published records on BOLD, only three were identified to species. Thus, we used the ABGD algorithm (using settings as above) to delimit the number of species based on *COI* data. Only 18 of the 33 *Pauesia* samples were amplified and sequenced (as described above) due to limited DNA amounts after preparing ddRADSeq libraries. These results were also compared with the following data to determine the most likely number of species: (1) the phylogenies obtained from the ddRADSeq data using a phylogenetic species concept (Nixon and Wheeler, 1990); (2) the ecological information for each specimen; and (3) calculated inter- and intraspecific distances under a K2P model.

RADseq Library Preparation

Double digest restriction-site associated sequencing (ddRADseq) libraries were created for both wasp and aphid samples. This type of sequence data has been shown to be effective for reconstructing robust phylogenies at shallow (Gilman and Tank, 2018; Zhang et al., 2018) and deeper timescales (Cariou et al., 2013; Leaché et al., 2015) and is appropriate for phylogenetic reconstruction of closely related species (Cariou et al., 2013; Gilman and Tank, 2018; Zhang et al., 2018). Wasp samples that had been manually removed from mummies were sequenced for the barcoding region of COI as described above to determine if they were hyperparasitoids and/or whether the sample was contaminated with fungus. Any samples with BLAST results that were not identified as Aphidiinae were removed from the ddRADseq library prep.

Six potential enzyme pairs (EcoRI-MspI, NlaIII-MluCl, PstI-MseI, PstI-MspI, SbfI-MspI, NlaII-EcoRI) were selected for testing. Two of these pairs were selected from Peterson et al. (2012) who indicated these enzymes may be appropriate for Hymenoptera. We performed in silico tests to assess the number of loci each enzyme pair would create based on three braconid (*Diachasma alloeum* NW_015145002.1, *Fopius arisanus* NW_011887740.1, and *Microplitis demolator* NW_014463857.1) and six aphid (*Homalodisca vitripennis* KK961494.1, *Acyrthosiphon pisum* NW_003383499.1 *Diuraphis noxia* NW_015368243.1, *Myzus persicae* LXJY01000001.1, *Oncopeltus fasciatus* KK854002.1, and *Piezodorus guildinii* JTEQ01000322.1) genomes with the package SimRad (Lepais and Weir, 2014) in R v3.4.3 (R Core Team, 2017). NlaIII and MluCl were chosen for both aphids and wasps as the simulations demonstrated that thousands of loci would be produced, which are necessary to resolve clades and produce robust nodes in shallow phylogenies.

Library preparation followed Peterson et al. (2012) with the following modifications. We used 100 and 50 ng total genomic DNA for *Cinara* and *Pauesia,* respectively. Size selection of 216-336 bp fragments was conducted on 6 pools of 8 individuals each using a PippinHT. Final libraries contained 48 samples. Six *Cinara* samples were run in the *Pauesia* lane with 50 ng genomic DNA as they did not have 100 ng of total DNA, so could not be included in the *Cinara* library. Two lanes were sequenced at Sanford-Burnham Prebys Medical Discovery Institute on an Illumina MiSeq with 150 bp paired-end reads.

Data Filtration

The raw sequence files were denovo demultiplexed, filtered, and assembled in PyRAD v3.0.66 (Eaton, 2014). Multiple runs with different parameter values were tested to maximize the amount of data retrieved while minimizing the individuals that were discarded because of missing data. No mismatches in barcodes were allowed in either dataset to avoid counting a sequence as the wrong individual. Settings that varied between *Cinara* and *Pauesia* datasets included minimum depth required for base calls, the clustering threshold, and the minimum number of samples per locus required for PyRAD to keep it in the final dataset. These values for the wasps and aphids respectively were: 6, 0.85, 4 and 8, 0.875, 6. Both raw libraries are available on FigShare: https://figshare.com/account/home#/projects/57272.

The final libraries had a lot of missing data, typical for ddRADseq. The aphid dataset had an average of 1511 loci per individual. The average number of loci for wasps was 4241. Individuals with less than 100 loci in the final assembly were removed from the dataset. After removing these individuals, the average loci per individual increased from 1511 to 1703 for the aphid dataset, and from 4241 to 4932 for the wasp dataset. PyRAD output files were transformed in the package phrynomics (Banbury and Leache, 2014) in R v3.4.3 (R Core Team, 2017) to prepare them for phylogenomic analyses in different programs. Each program has different filtering requirements, thus total SNPs vary slightly depending on the program. After filtering, the final wasp dataset included 27 individuals and 21340 loci. The nexus file for MrBayes included 54425 SNPs, and the phylip file for RaxML included 53973 SNPs. The final aphid dataset included 48

individuals and 9980 loci. The nexus file for MrBayes included 59635 SNPs, and the phylip file for RaxML included 58379 SNPs.

Aphid and Wasp Phylogenetic Reconstruction

Both Maximum Likelihood (ML) and Bayesian inference (BI) analyses were conducted to determine if trees shared the same topology across methods. Both methods were run with the Lewis ascertainment bias (Lewis, 2001) to control for SNP data. For ML reconstruction, RaxML v8.2.10 (Stamatakis, 2014) was used on the CIPRES gateway (Miller et al., 2010). For BI, we ran MrBayes v3.2.6 (Ronquist et al., 2012) for 2,000,000 generations, sampling every 1,000th generation. Tracer plots (Rambaut et al., 2018) were examined to determine if runs had proper mixing and had reached stationarity and convergence.

Data filtration removed the aphid outgroup taxa, so *Cinara* phylogenies were midpoint rooted. The outgroup for the *Pauesia* tree was *Xenostigmus*. Final trees were viewed in Figtree v1.4.3 (Rambaut, 2014), and edited with the ggtree (Yu et al., 2016) and ggplot2 (Wickham, 2016) in R v3.4.3 (R Core Team, 2017) and Adobe Photoshop CC.

Pine Phylogeny

Sequences of two chloroplast genes (*rbcL* and *matk*) and one nuclear rRNA gene (*ITS2*) were downloaded from GenBank (Clark et al., 2016) for the 11 *Pinus* species that were hosts of *Cinara* in this study (Table 3). The outgroup was *Picea meyeri* as there is high support for *Pinus* as sister to a group containing *Picea* and *Cathaya* (Leslie et al.,

2012; Wang et al., 2000), and all three genes were publicly available for *P. meyeri*. Sequences were aligned by MUSCLE using default settings. The final *rbcL*, *matk*, and *ITS2* datasets were 1454, 1718, and 242 base pairs in length, respectively. These genes were run on PartitionFinder2 (Lanfear et al., 2016) on the CIPRES gateway (Miller et al., 2010) using a greedy algorithm and with partitions of each codon position for each protein gene designated *a priori*. The best model for all codon positions of *ITS2* was HKY. The best model for all codon positions of *matk*, and the third codon of *rbcL* was GTR+G. The best model for the first and second codon positions of *rbcL* was F81+I and JC+I, respectively. ML and BI analyses, assessment of convergence diagnostics, and final tree preparation were completed as described above for the aphid and wasp phylogenies.

Coevolutionary Analyses

Two main methods of cophylogenetic analysis are utilized: event-based methods and distance-based methods (Balbuena et al., 2013; de Vienne et al., 2013). Event-based methods reconstruct the evolutionary events (cospeciation, host switches, sorting, and duplications) in the history of the parasite taxa onto the host phylogeny. Each event is assigned a cost, and events are reconstructed based on the parsimony criterion, where the least cost scenario is considered the best reconstruction. This reconstruction can then be compared to permutations to determine if the events in the best reconstruction differ from random (Balbuena et al., 2013; de Vienne et al., 2013; Merkle et al., 2010). Distance-based methods test the congruence of host and parasite phylogenies, with the null hypothesis being that taxa will have incongruent phylogenies. If phylogenies between

two taxa show a pattern of congruence, this indicates those taxa have a shared evolutionary history and suggests the taxa are cospeciating (Balbuena et al., 2013; de Vienne et al., 2013). These methods will be used to determine the patterns of speciation in all three of the interacting taxa in our tri-trophic system.

For distance-based analyses, two different programs were chosen to ensure that each program was reaching the same conclusion. PACo (Balbuena et al., 2013) was used with 100,000 permutations using the Bayesian phylogenies to produce distance matrices. Then, AxParafit and AxPcoords (Legendre et al., 2002; Stamatakis et al., 2007) was run in Copycat (Meier-Kolthoff et al., 2007) with 99,999 permutations, also using the Bayesian phylogenies to produce distance matrices. Significance of congruence is tested by inputting distance matrices, here created from the Bayesian phylogenies of each taxa, and the associations between the taxa in each phylogeny (Table 2). These distance matrices are transformed into principal coordinates, combined with the host association matrix to create extended principal coordinates and develop a Procrustes plot. A global goodness of fit is provided and tested using permutations where the association matrix is randomly changed. The null hypothesis of incongruence was rejected if the p value was below 0.05. Both programs use this method to determine if the phylogenies are significantly congruent, but AxParafit and AxPcoords also test if each link between terminal taxa significantly contributes to the global goodness of fit of the phylogenies.

For event-based methods, CoRe-Pa was chosen because it allows more than one association per taxa, taxa without associations, and permutations with random interactions (Merkle et al., 2010). Many of these features are not available for other

event-based methods (de Vienne et al., 2013; Martínez-Aquino, 2016). The species clades of aphids and wasps were collapsed for these analyses. This controls for overestimation of cospeciation that may occur when multiple individuals of a given species are included in the analyses. Each pair of interacting taxa were analyzed with the standard event costs (cospeciation: 0, sorting: 1, duplication: 2, host switch: 3) for 10,000 random cycles to determine the most parsimonious reconstructions of the evolutionary events of the parasite. Then 100,000 permutations were conducted with the same event cost settings and topologies, but the interactions between terminal taxa randomized. Finally, subsection Ponderosae in the *Pinus* tree (*P. ponderosa*, *P. jeffreyi*, and *Pinus coulteri*) was collapsed so that the interactions where the tree species identification was not possible (*P. ponderosa*/*P. jeffreyi*) could be included in the analyses.

RESULTS AND DISCUSSION

Species Delimitation, Phylogenies, and Taxonomic Revisions

When conducting coevolutionary analyses, having the correct species identifications is important, as changes in the number of species can change the results of the analyses and thus the interpretation of the results. The insect groups studied here have been described based on one or few lines of evidence, so careful consideration was taken using an integrative framework and multiple independent forms of evidence to determine the number of *Pauesia* and *Cinara* species.

The integrative taxonomic framework resulted in 7 putative *Pauesia* species. The ABGD results for *COI* sequences of *Pauesia* recovered different numbers of putative species depending on the prior maximal intraspecific distance (PMID). Six, five, or four clusters were recovered at PMIDs of 0.1 - 0.128%, 0.162% - .0395%, and 0.886 – 2.98%, respectively. These species numbers were examined with the results from the ddRADseq phylogeny, ecological data, and intra- and interspecific distances to make a final determination on the number of species. The BI (Fig. 3) and ML (Fig. 4) trees reconstructed from the ddRADseq dataset for *Pauesia* were largely identical save some minor intra-clade relationships within species. Thus, all future references to the *Pauesia* phylogeny refer to the Bayesian tree (Fig. 3). Six or seven monophyletic clades of *Pauesia* can be defined from the phylogeny (Fig. 3), depending on how finely the clades are divided. Under a phylogenetic species concept, defining the smallest diagnosable units with shared evolutionary history, eight species can be delimited (Fig. 3). The six

groups indicated by ABGD are monophyletic clades in the Bayesian phylogeny and thus supported as putative species. An additional putative species that lacked a *COI* sequence and thus was not included in the ABGD analysis is represented by one taxon (AIB071), but branch lengths between this taxon and its nearest neighbor indicate high support for it being an additional species. The geographic ranges and hosts of these species further supports delimitation into eight putative species. For example, *Pauesia* sp. 1, 2 are clustered together, but are found in different states or on different host plants. Thus, we accepted seven putative species for subsequent co-evolutionary analyses (Table 2, Fig. 3, Fig. 4). Of the seven putative species for which we had COI data, the interspecific distances ranged from 0.78 - 12.95% and all intraspecific distances were ≤ 0.29% (Table 4).

Without a taxonomic revision of *Pauesia*, we cannot currently ascertain if our collected *Pauesia* are previously described or new species. Within the states we sampled (CA, CO, NM, UT), *Pauesia* have only been recorded from California and Colorado, and of the *Pauesia* listed, none have host records for these states. Future efforts should collect *Pauesia* across the western range and gather geographic data from entomological collections to develop complete species lists and distributions. As with most parasitic Hymenoptera, there are likely several of these species that are new to science (Dolphin and Quicke, 2001; Forbes et al., 2018; Rodriguez et al., 2013), particularly as very few taxonomic studies have focused on North American species of *Pauesia*. There is also a high prevalence of cryptic species complexes in Braconidae (Boring et al., 2011; Derocles et al., 2016; Peixoto et al., 2018; Ridenbaugh et al., 2018; Rodriguez et al.,

2013; Smith et al., 2008b; Zhang et al., 2017). Our data indicates that *Pauesia* are largely monophagous (see host breadth below) and there are many more described *Cinara* species available as hosts. Thus, either many species of *Cinara* have escaped parasitism by *Pauesia*, or there are numerous more *Pauesia* to be discovered with more intensive sampling. Subsequent papers will focus on expanding our knowledge of *Pauesia* distributions, describing new species, revising existing species and developing enhanced identification tools.

Using the BOLD database, we were able to obtain positive species level identifications for five aphid species (*C. brevispinosa*, *C. contortae*, *C. glabra*, *C. ponderosae*, and *C. terminalis*). However, specimens matching to two of these species epithets (*C. ponderosae*, and *C. terminalis*) fell into two different BINS per species name. Further, some specimens had *C. contortae* as the best match but failed to fall within the BIN labelled with that epithet (Table 1). Several specimens fell into clades in the output neighbour joining trees with other labeled species, including *C. contortae*, *C. medispinosa*, or *Cinara murrayanae*. Specimens labelled with a best match to *C. schwarzii* matched 100% to a non-identified taxon, and therefore did not fall within a BIN, but all fell within a well-defined monophyletic group for this species with all other *C. schwarzii* taxa on BOLD. For clades labeled as *C.* nr. *apini* (1 and 2) and *C.* nr *anelia* (1 and 2), there were several other species labels within clusters on the output neighbour joining tree on BOLD, including the following *Cinara* species: *apini*, *anelia*, *moketa*, *kuchea*, *wahtolca*. This likely indicates a need for further revisionary work on these species and subsequent updates to the identified species names on BOLD.

The BI (Fig. 5) and ML (Fig. 6) *Cinara* phylogenies were almost completely congruent, except the placement of two clades were reversed (*C*. nr. *contortae* 1 and *C*. nr *contortae* 2 were reversed) and the location of *Schizolachnus pineti* varies. The BI tree was used as the final tree and in further analyses because it had higher support and fewer polytomies than the ML phylogeny. Fifteen well supported (posterior probabilities >.95) monophyletic clades of *Cinara* can be defined from the phylogeny (Fig. 5), although all *C*. *terminalis* clades had short branch lengths between them. The ABGD results for *COI* sequences of *Cinara* also recovered 15 putative species (including *Schizolachnus* nr. *pineti*) for all prior maximal intraspecific distances (PMIDs) greater than 0.43%. For PMIDs under 0.43%, an additional species was recovered (AIB185). Despite AIB185 being collected from a different state and host plant than other members within the *C*. nr. *anelia* 1 clade (Fig. 5), we chose to agree with the 15 putative *Cinara* species because it matched the well supported monophyletic clades in the BI phylogeny (Fig. 5), the AIB185 *COI* sequence was a messy, and the lowest PMID value for 15 species was quite small (0.43%). Further, with AIB185 included within the *C*. nr. *anelia* 1 clade, the maximum intraspecific distance was 0.29% (Table 5).

Our results paired with recent *Cinara* phylogenetic studies show that revisions are necessary within this genus. Many currently described *Cinara* species are based on characters that vary with host tree utilized or are based simply on the tree they are found on (Favret and Voegtlin, 2004a). These host variable characters are often a part of the dichotomous keys used to identify these species (Bradley, 1951; Favret and Voegtlin, 2014) potentially causing some *Cinara* species to appear more specialized than they are.

Cinara ponderosae, *C. terminalis*, and *C. brevispinosa* were monophyletic here and are consistently monophyletic in other phylogenies (Jousselin et al., 2013; Meseguer et al., 2015a). However, here we show evidence that both *C. terminalis* and *C. ponderosae* may be cryptic species complexes containing multiple species with different geographic distributions or host plants. Other species that may require taxonomic revisions include *C. anelia* and *C. apini*, which were also near *C. kuchea*, *C. moketa*, and *C. wahtolca.* These species form a paraphyletic clade in a previous *Cinara* phylogeny (Jousselin et al., 2013), however the terminal names are based on host plants and not a phylogenetic species concept, causing issues with valid species delimitation. A similar example is *C. contortae*, which BOLD identified as near three species (*C. contortae*, *C. medispinosa*, or *C. murrayanae*), all which feed on *P. contorta*. These species were recovered as paraphyletic in Jousselin et al. (2013) and should be revised. Finally, *Schizolachnus* should be incorporated into *Cinara*, as it consistently falls within this genus in phylogenies (Chen et al., 2016a; Meseguer et al., 2015a).

The Maximum Likelihood phylogeny of *Pinus* (Fig. 7) was well supported across all nodes except between *P. edulis* and *Pinus cembroides* and within subsection Ponderosae, which was returned as a polytomy. The Bayesian phylogeny of *Pinus* (Fig. 8) had similar areas of low support, but subsection Ponderosae was better resolved, with *P. ponderosa* highly supported as the sister to *P. jeffreyi* and *P. coulteri*. Subsection Ponderosae has varied in previous phylogenies of this genus, especially the placement of *P. ponderosa* (Hernández-León et al., 2013; Leslie et al., 2012; Saladin et al., 2017), which may account for the polytomy in the Maximum likelihood phylogeny (Fig. 7).

However, the relationships in our Bayesian tree, even those with low support, are the same as the most recent phylogenies with the largest genetic datasets (Gernandt et al., 2009; Hernández-León et al., 2013; Saladin et al., 2017). Thus, the Bayesian phylogeny (Fig. 8) was utilized in all cophylogenetic analyses. For brevity, for the remainder of the paper we refer to putative species as species.

Host Breadth

Pauesia species from this study were reared from 10 of the 15 collected aphid species, and an additional two species identified in BOLD as *Cinara edulis* and near *Schizolachnus piniradiatae* that were excluded from the final dataset due to missing data. Most *Pauesia* species were specialized on a single aphid species, although three of the eight *Pauesia* were reared from two aphid species (Table 2). Koinobiont parasitoids are expected to be specialists (Quicke, 2015), and this appears to be the case in most of the wasp species we collected here. Interestingly, when a given wasp species was reared from two aphid species, those aphids were not closely related but were often found on the same tree (e.g. *Pauesia* sp. 2 attacks *C*. *terminalis* 3 and *C*. nr. *anelia* 1 both on *P*. *edulis* (Table 2)). Most *Pauesia* (6 of 8) were also specific to a single host tree. The two exceptions were found on related pines. *Pauesia* sp. 4 was found on *P*. *lambertiana* and *P*. *monophylla*, both in subgenus *Strobus*. *Pauesia* sp. 2 was found on *P. monophylla* and *P. edulis*, both in subsection *Cembroides*. This suggests that *Pauesia* are utilizing phytochemicals to detect host tree species, but once found they may be able to utilize

multiple aphid species on that host plant. However, this was not tested in this study and will require future study.

Aphids were collected from either 10 or 11 *Pinus* species in this study, depending on the identity of the *ponderosa/jeffreyi* trees. For the positively identified species of aphids from the BOLD database, there was only one new host record. *Cinara terminalis* 3 was found attacking *Pinus longaeva*, which is in a different subsection of *Pinus* than all other host records for this species (Blackman and Eastop, 2000). The aphid species ranged from attacking one to three *Pinus* species. Many of the aphid species that could be found on more than one host tree were on *Pinus* species from different parts of the *Pinus* phylogeny (Table 2). One of the aphid species, *C.* nr. *contortae* 1, was found on *Pinus* from both the *Pinus* and *Strobus* subgenera. This suggests that some species are very specialized and thus cospeciation would be expected, but several other species demonstrate some host flexibility, suggesting that host switches may be equally as likely, or that the evolution of one taxa does not depend on the other. At least some species of *Cinara* can recognise multiple species as suitable hosts, stressing that describing new species based on host associations should be avoided.

Coevolutionary Analyses

Our distance-based co-evolutionary analyses demonstrated that 2/3 paired phylogenies were significantly congruent (Table 6) using both PACo and AxParafit/AxPcoords. There is phylogenetic congruence between two of the pairs of taxa in this system, suggesting there is some cospeciation between those taxa. However, while

PACo analyses found that wasps and pines were significantly congruent, AxParafit/AxPcoords did not. The AxParafit analyses further demonstrate the links between interacting taxa that support the hypothesis that the phylogenies are congruent. In all three cophylogenetic analyses, the percent of significant links ranged from 36.4 – 70.4% (Table 6), suggesting that other patterns of speciation are occurring between these taxa.

To test which patterns of speciation may be occurring we utilized event-based co-phylogenetic reconstruction with CoRe-Pa and collapsed individual taxa to species clades to prevent over-estimation of events. These analyses confirm that co-speciation is not the dominant pattern of speciation across the interacting taxa. Rather, cospeciation and host switching occur in similar frequencies between all interacting taxa, and in some cases host switching is more frequent than cospeciation (Table 7). The most parsimonious reconstruction between the aphids and wasps had 2 cospeciation, 2 sorting, 1 duplication, and 3 host switching events (Fig. 9, Table 7). None of these events, nor the total cost of the reconstruction, varied from the mean and standard deviation of the 100,000 permutations, and thus no events occurred at a frequency larger or smaller than expected if the associations between taxa were random. Though most wasps in this study were monophagous, those that were not were on distantly related aphids, and the patterns from event-based methods suggests that host switches were equally as important as cospeciation in forming wasp species (Table 7). Sorting events (where the host speciates and the parasite does not) must be taken with great caution as sampling can specifically affect these patterns. Our sampling occurred across four states and over one month. Thus,

for aphids where no wasp was reared, it is possible that we sampled too early in the season before colonies were parasitized, which would appear as a sorting event as opposed to a limitation of sampling. The one duplication event reconstructed between wasps and aphids is likely a historical vicariance event. *Pauesia* sp. 5 and 6 both attack *Cinara schwarzii*, however, the different *Pauesia* species occur in different geographic regions.

CoRe-Pa results between the aphids and pines inferred that the total cost of the reconstructions and all events besides duplication differed from random (Fig. 10, Table 7). We collapsed the subsection Ponderosae of the *Pinus* phylogeny and re-ran the analyses to account for the links between aphids attacking pines that could not be accurately identified (*P. ponderosae*/*P. jeffreyi*). When subsection Ponderosae was collapsed, two reconstructions (Fig. 11, Fig. 12, Table 7) had the same cost, and again most events fell outside the range of random permutations. The sorting events were likely somewhat overinflated in the un-collapsed analyses because of the missing links between some aphid taxa and *P. ponderosae*/*P. jeffreyi*. Cospeciation in our aphid-pine reconstruction was slightly higher than the range of random permutations, while host switches were slightly lower (Table 7). This suggests that cospeciation has been a prevalent pattern between aphids and pines, which was also suggested by the relatively high number of links supporting congruence in the paired phylogenies in the distance-based analysis (Table 6). This is contrary to previous research on Cinara that found that cospeciation was not a dominant factor in aphid speciation when examining host relationships with conifer genera (Jousselin et al., 2013; Meseguer et al., 2015a). The

high number of sorting events may be due to limitations of sampling or indicative of extinction events in aphids. The finding that there were fewer host switched in aphids compared to random permutations is contrary to previous *Cinara* cophylogenetic studies (Jousselin et al., 2013; Meseguer et al., 2015a), and suggests that aphids may not be as restricted to plant phytochemistry as is seen with their parasitoids. However, our results also demonstrate that no one pattern is dominant and that many different ecological and non-ecological factors may be driving speciation in aphids.

The CoRe-Pa reconstruction inferred one most parsimonious reconstruction, and two when subsection Ponderosae was collapsed (Fig. 13, Fig. 14, Fig. 15, Table 7). As there were several links between wasps and pines identified as *P. ponderosae*/*P. jeffreyi*, the collapsed Ponderosae clade served to deflate the number of inferred sorting events (Table 7). There were no duplication events when the ponderosae clade was not collapsed, but 1-2 when they were, thus these events can be attributed to multiple wasps associated with a collapsed clade containing 3 species of pines. This suggests that vicariant events are not important for wasps relative to their hosts plants, and that ecological factors likely play a larger role in speciation. The subsection Ponderosae collapsed and un-collapsed analyses did not change the number of cospeciation events, while one of the subsection Ponderosae collapsed reconstructions resulted in one fewer host-switching event (Table 7). The higher number of host switches relative to cospeciation events was not expected given that all of our *Pauesia* species were specialized on one or two closely related host plants. Future plant phytochemistry studies

may reveal that even in the case of host plant switches, the trees have very similar chemical profiles.

Though we now have evidence of the macroevolutionary patterns of speciation in this tri-trophic system, we cannot confirm the processes that led to those patterns. Distance-based methods only test for congruence, which indicates that those taxa have a shared evolutionary history and have had some level of cospeciation. Event-based methods can infer which nodes in our paired phylogenies show a pattern of cospeciation. However, multiple processes can result in the pattern of cospeciation (Althoff et al., 2014), including coevolution (Smith et al., 2008a), phylogenetic tracking (Bentz et al., 2006), or co-vicariance (Koop et al., 2014). We see some evidence for allopatric speciation in both the wasps and the aphids. For example, the two *C. ponderosae* clades appear to be split based on geography, with one clade in California, and the other distributed across New Mexico, Colorado, and Utah. In *Pauesia*, there appears to be a recent divergence within *Pauesia* sp. 5 and 6, where one clade was found in California and another in New Mexico. The sample size of each species were quite low, so a wider sampling would have to be undertaken to confirm that these species are geographically separated, but the Rocky Mountains may be a significant barrier separating the two species.

Finally, available methods to study patterns of cospeciation and host switches between interacting taxa are flawed. The current methods, event- and distance-based methods, are limited in their ability to detect cospeciation. Most event-based methods use maximum parsimony, assuming the simplest reconstruction is correct. Often multiple

reconstructions have an equal total cost and suggest different patterns, making interpretation difficult. Furthermore, the standard cost settings vary between different programs and tend to be low for cospeciation and higher for host switches but have no biological reason for those costs (de Vienne et al., 2013). To our knowledge, there is no study that supports that cospeciation is easier and thus should be a lower cost than host switches. Event-based methods that calculate the best event costs based on the dataset often create more biased cost settings, where host switches are hundreds of times less likely to occur than the other three events (Vanhove et al., 2015). The null hypothesis for distance-based methods is incongruence in paired phylogenies, which is extremely unlikely in closely interacting species such as parasitoids and their hosts and host plants. These taxa obviously share evolutionary history, so there is likely to be some level of congruence in phylogenies that results in rejecting the null hypothesis. Thus, the results of each of these methods must be taken with caution. Of course, linking processes with patterns seen in phylogenies is inherently difficult, as there is always missing information when looking at such distant timescales: entire lineages may have gone extinct in this time, changing the patterns we can detect with these types of analyses. However, these analyses can be improved by incorporating more taxa in the phylogenies, dating the phylogenies, and looking for divergence at the population level.

Future Studies

The results of this study indicate that this system is worthy of further exploration. Some species in our phylogenies show evidence that they are currently or have recently

undergone cospeciation or host switches. *Cinara* nr. *anelia* 1 and 2 should be examined further to test recent for host associated differentiation. Future studies will require more sampling to have a more complete phylogeny for all three taxa. Further, dated trees would help assess if cospeciation is truly occurring. For example Sorenson et al. (2004) found that although coevolutionary analyses suggested parasitic birds coevolved with their host birds, the dates of these divergences were not aligned across trophic levels as they should be if the species were coevolving. However, for the aphids and wasps in these analyses, there is no reliable way to date the phylogenies without inclusion into a larger dataset with numerous fossil evidence. Full species distributions are also necessary to robustly test for geographic speciation, highlighting the need for more intensive sampling across multiple months and years. If taxa have non-overlapping distributions, this may be evidence of co-vicariance rather than coevolution. However, if recently diverged species show evidence of reproductive isolation by having little to no gene flow between them, there may be a pre- or post- zygotic barriers preventing reproduction due to traits that are coevolving with their hosts. Plant phytochemistry should be incorporated into future studies to determine if volatile chemical profiles are related to host use, such that species are restricted to hosts with similar profiles, and host switches occur between the most chemically similar host plants. Finally, these taxa have great potential as a model multi-trophic system for several reasons, including: (1) each lineage are a group of fairly specialized, closely related organisms with high host fidelity, which make them suitable for cophylogenetic analyses; (2) there are other associations in this system, including aphid-tending ants, hyperparasitoids, and a diverse array of aphid

endosymbionts (Burke et al., 2009); and (3) unlike most parasitoid systems, collection and rearing are relatively easy making this system excellent for a multitude of additional experiments on individual species and populations.

CONCLUSIONS

In this study we investigated a plant-herbivore-parasitoid system, involving pine trees, *Cinara* aphids, and *Pauesia* wasps to test for interactions among taxa and cophylogenetic patterns across all trophic levels using massively parallel sequencing data. This study revealed that species in this system are largely specialized with most of the aphids and wasps confined to one or two hosts. As expected with koinobiont endoparasitoids, but rarely tested on a macroevolutionary scale, *Pauesia* wasps were almost exclusively monophagous. When the parasitoids deviated from this pattern, they were found on unrelated aphids but on the same or closely related pine trees, suggesting that phytochemical cues in host plants may constrain the niche breadth of these parasitoids. Most paired phylogenies, between wasps and aphids, aphids and pines, and wasps and pines, demonstrated significant congruence using distance-based methods of cophylogenetic analyses. Cospeciation appears to be a prevalent pattern in all three interacting taxa, but host switches are more common than assumed in a specialized system with high host fidelity. We suggest that cophylogenetic analyses are flawed and can be difficult to interpret but can pinpoint taxa that may be undergoing ecological speciation. At least one group of interacting taxa demonstrate cascading speciation which should be further tested using population level data. Contrary to previous research, the two taxa that have undergone the most cospeciation are *Cinara* and *Pinus*. Additional research with more sampling, phytochemical data, population genetics and biogeography will assist in better understanding the patterns of speciation in this group and may begin to unravel the processes behind these patterns. This system has great potential as a model

multi-trophic system for exploring community dynamics and mechanisms on the genesis of biodiversity because of the specialized interactions and relatively easy sampling and rearing protocol. Exploring this system further will undoubtedly lead to new species descriptions, and revisions are called for both species of *Pauesia* and *Cinara*.

APPENDIX A: FIGURES

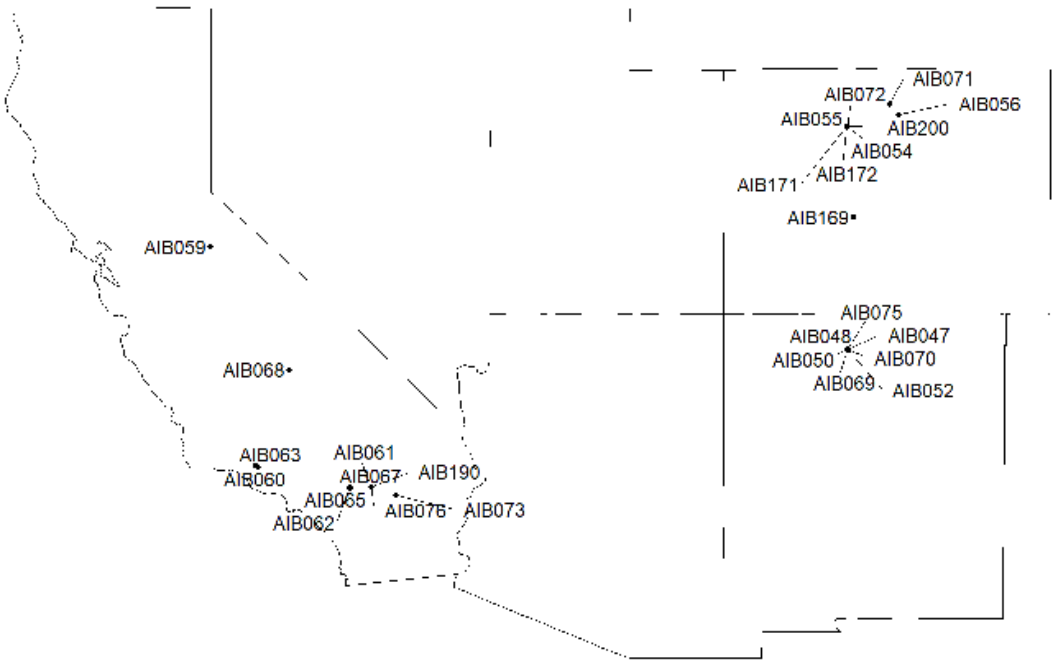

Figure 1: Map of collection localities of wasps used in our analyses. The map was generated in R (R Core Team, 2017) using the following packages: ggplot2 (Wickham, 2016) and ggmap (Kahle and Wickham, 2013).

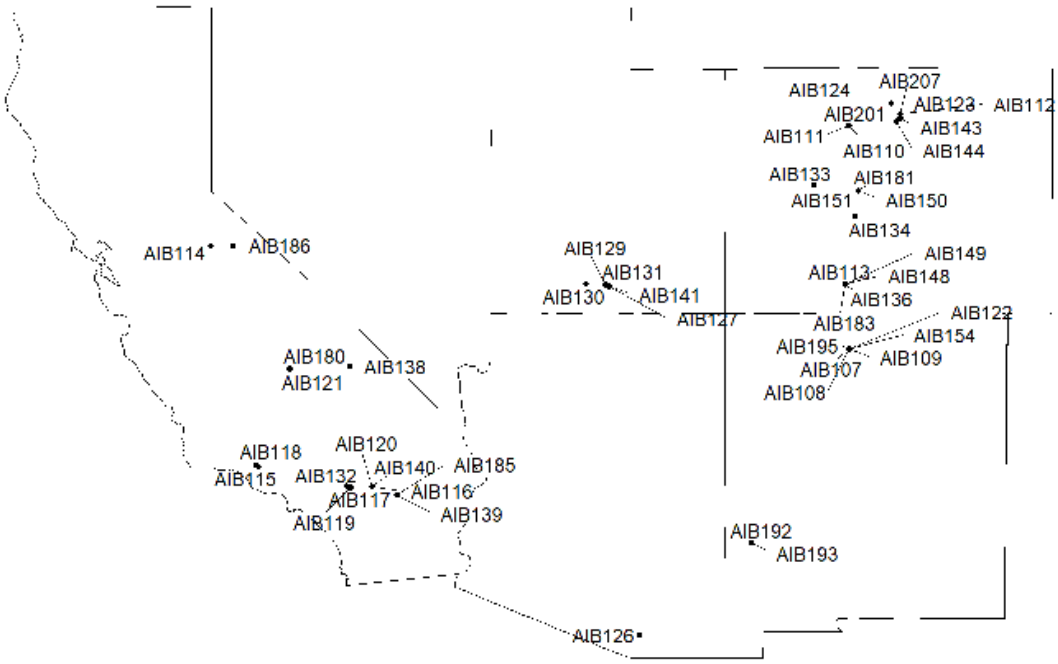

Figure 2: Map of collection localities of aphids used in our analyses. The map was generated in R (R Core Team, 2017) using the following packages: ggplot2 (Wickham, 2016) and ggmap (Kahle and Wickham, 2013).

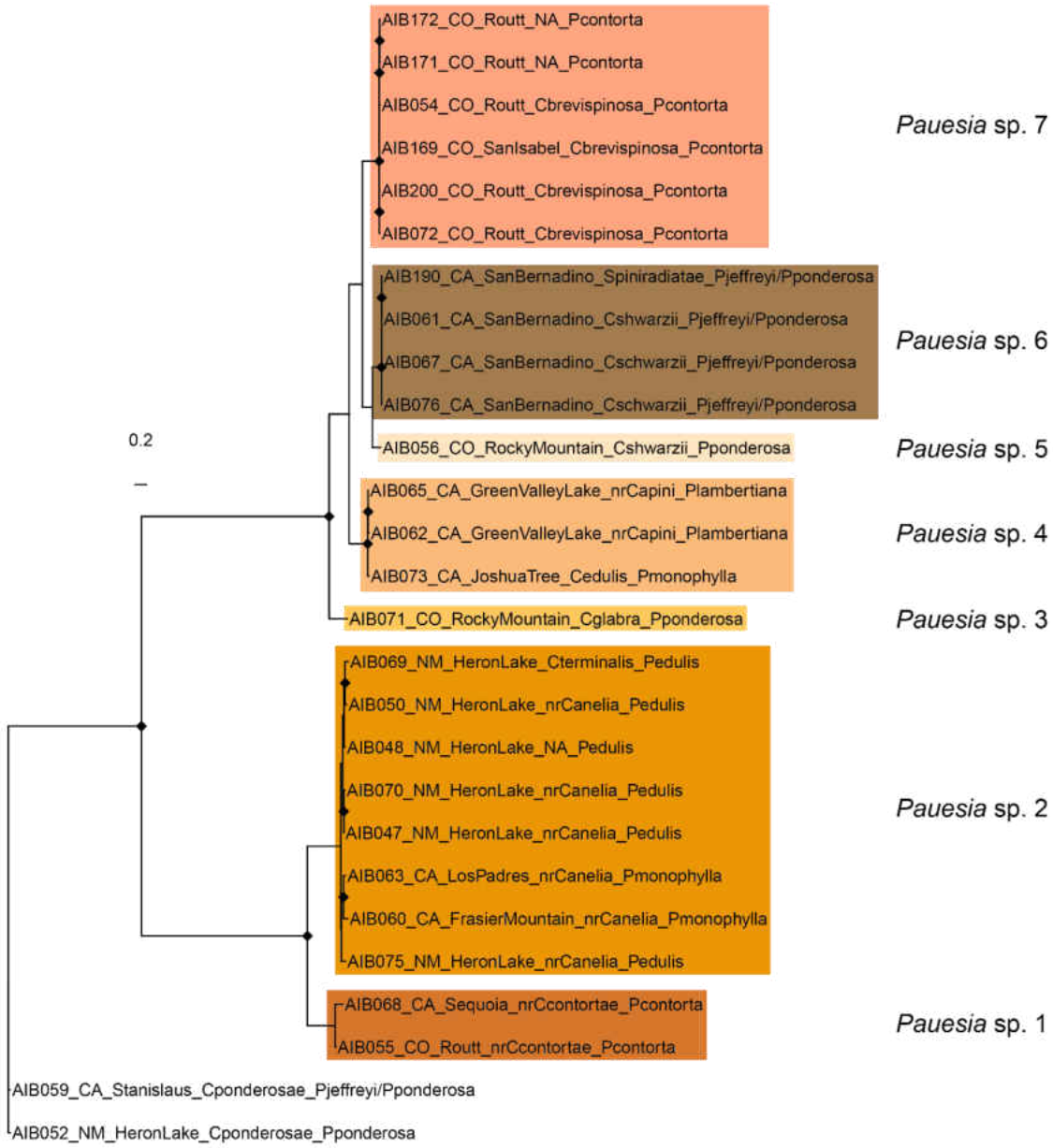

Figure 3: Bayesian phylogeny of *Pauesia* created in MrBayes v3.2.6 from 54425 SNPS. The outgroup is *Xenostigmus* sp. Nodes with a black diamond have a posterior probability of ≥ 95. Monophyletic clades are colored, and the putative species name is presented on the right. Terminal taxa are labelled with the following information, in this order, with an underscore between each: DNA voucher, State, locality, *Cinara* host species identified from the BOLD species level barcodes database, and *Pinus* host species. Terminal taxa with NA rather than a host species identified did not have a COI sequence, so could not be identified.

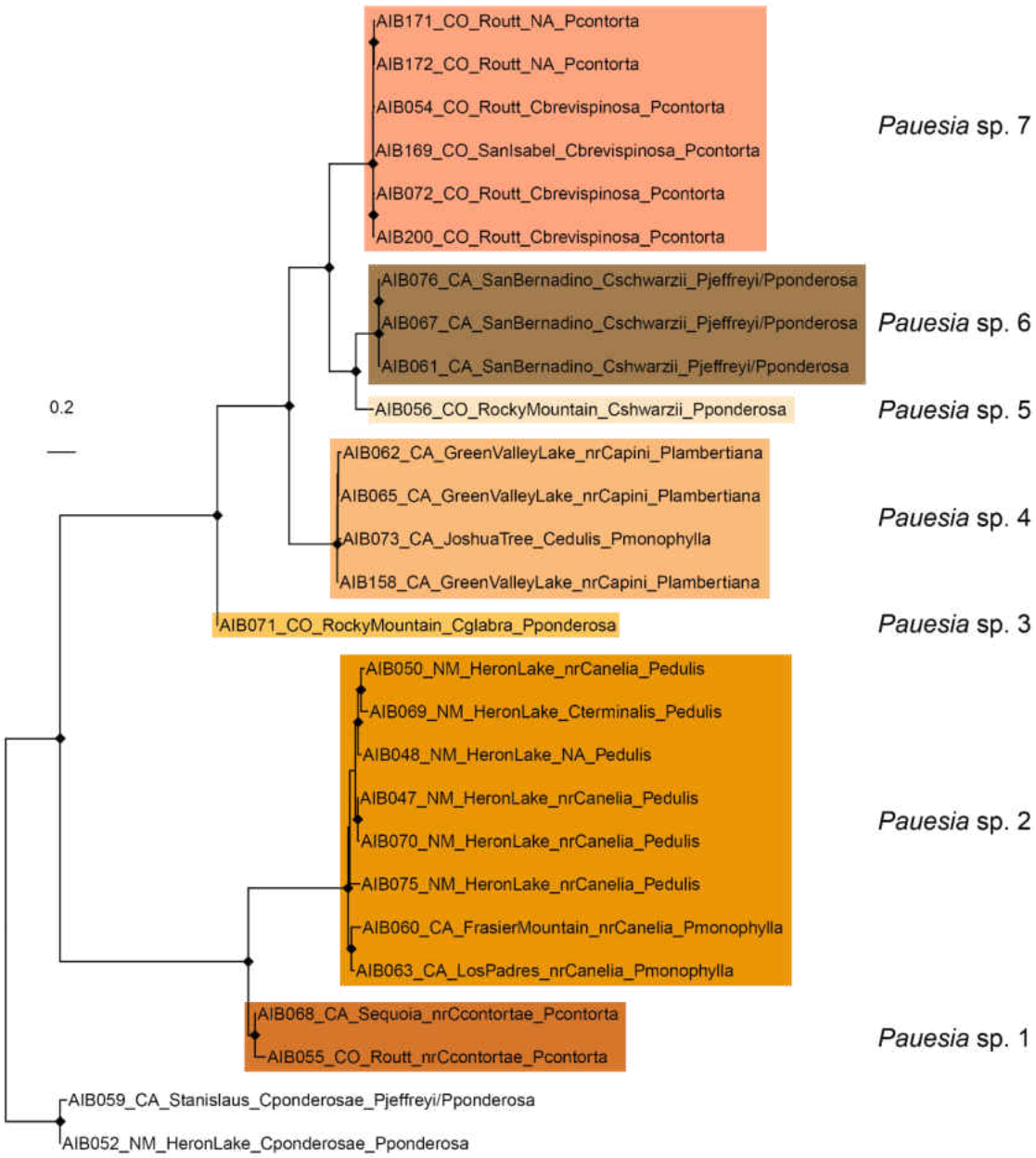

Figure 4: Maximum likelihood phylogeny of *Pauesia* created in RaxML v8.2.10 with 53973 SNPs. The outgroup is *Xenostigmus* sp. Nodes with a black diamond have a bootstrap value of ≥ 90. Monophyletic clades are colored, and the putative species name is presented on the right. Terminal taxa are labelled with the following information, in this order, with an underscore between each: DNA voucher, State, locality, *Cinara* host species identified from the BOLD species level barcodes database, and *Pinus* host species. Terminal taxa with NA rather than a host species identified did not have a COI sequence, so could not be identified.

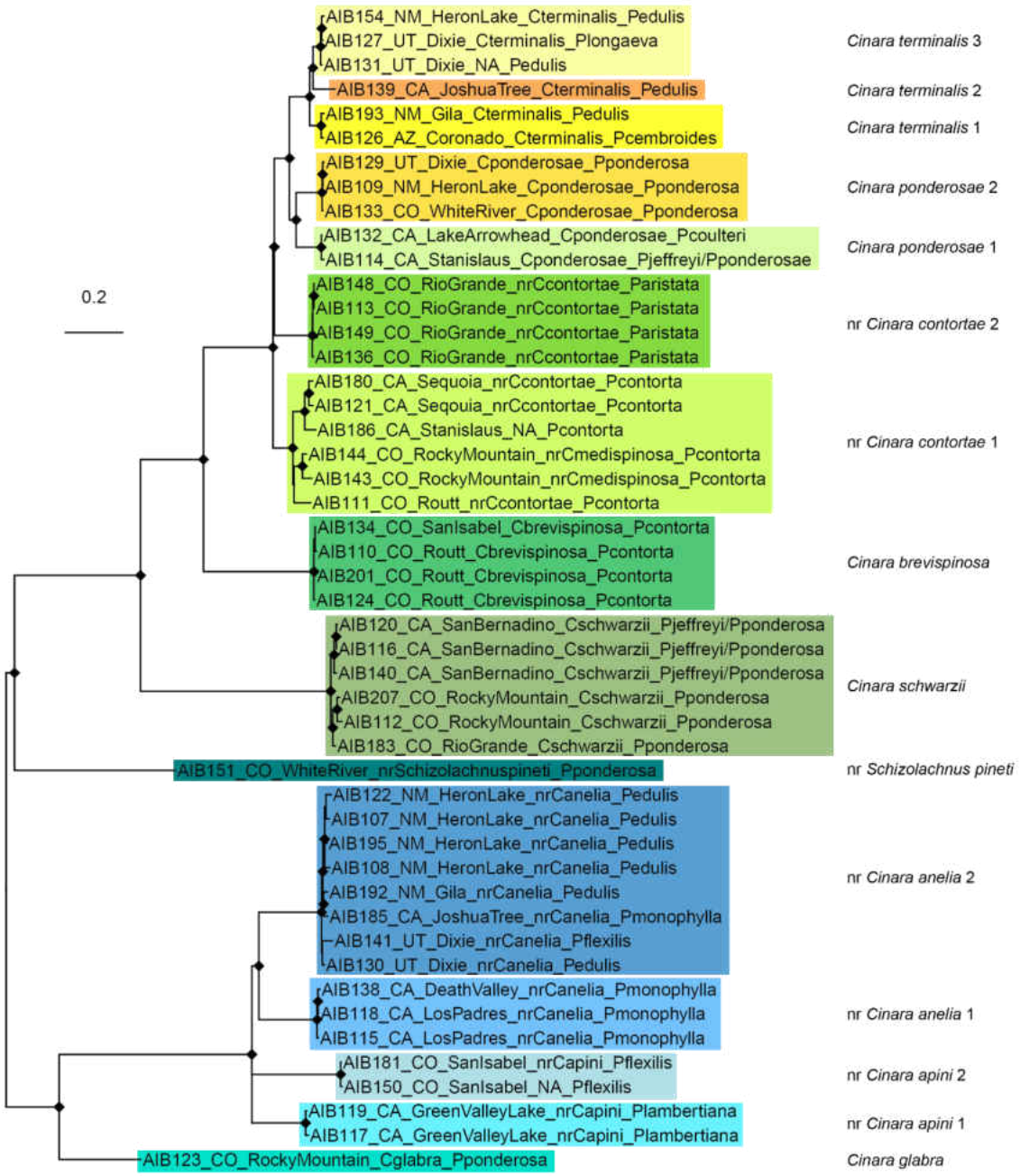

Figure 5: Bayesian phylogeny of *Cinara* created in MrBayes v3.2.6 from 59635 SNPS. Nodes with a black diamond have a posterior probability of ≥ 95. Clades of putative species are colored and labelled. Terminal taxa are labelled with the following information, in this order, with an underscore between each: DNA voucher, State, locality, *Cinara* species identified from BOLD species level barcodes database, and *Pinus* host species. Terminal taxa with NA rather than a *Cinara* species identified did not have a COI sequence, so could not be identified through the BOLD database.

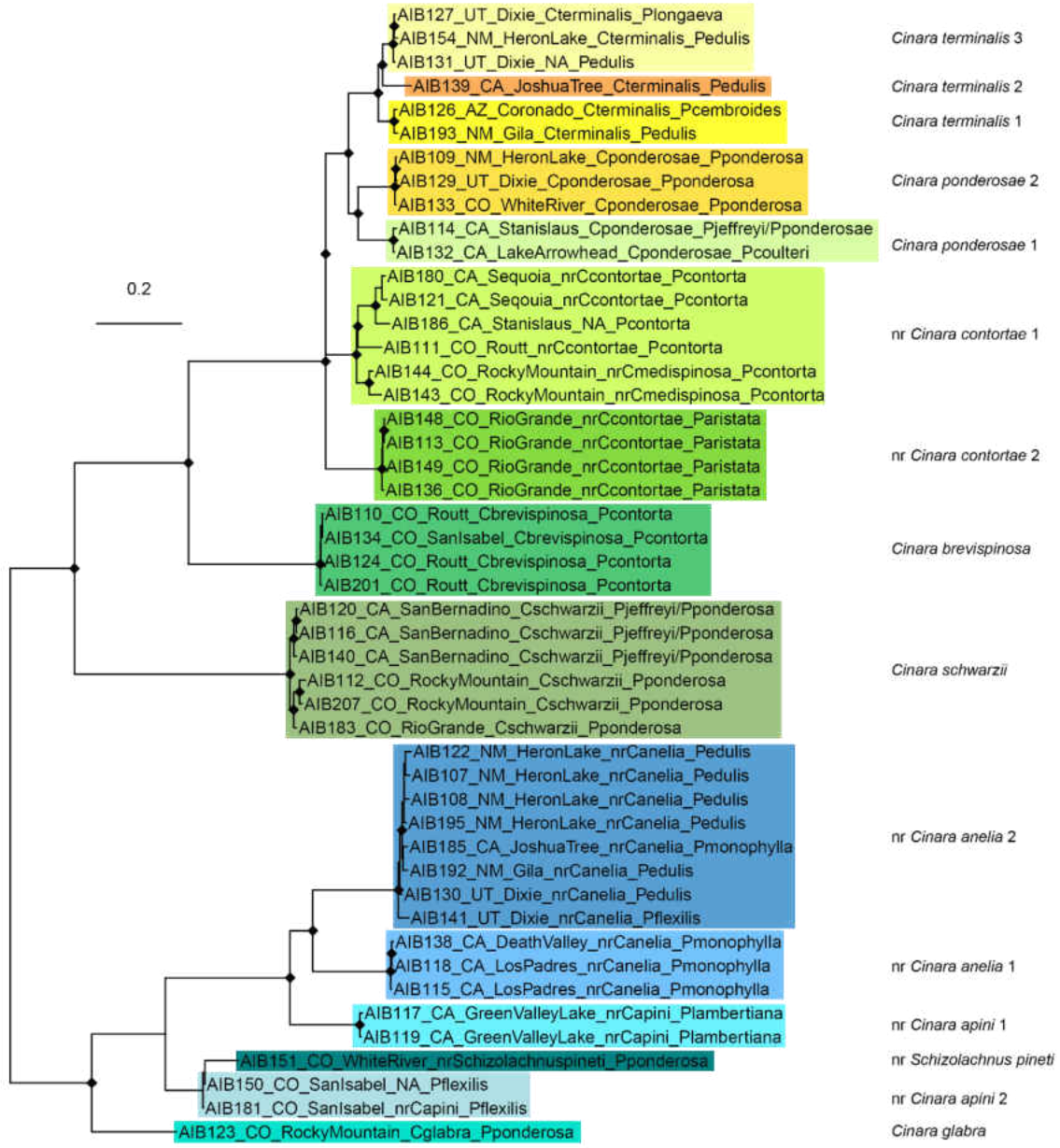

Figure 6: Maximum likelihood phylogeny of *Cinara* created in RaxML v8.2.10 with 58379 SNPs. Nodes with a black diamond have a bootstrap value of ≥ 90. Clades of putative species are colored and labelled. Terminal taxa are labelled with the following information, in this order, with an underscore between each: DNA voucher, State, locality, *Cinara* species identified from BOLD species level barcodes database, and *Pinus* host species. Terminal taxa with NA rather than a *Cinara* species identified did not have a COI sequence, so could not be identified through the BOLD database.

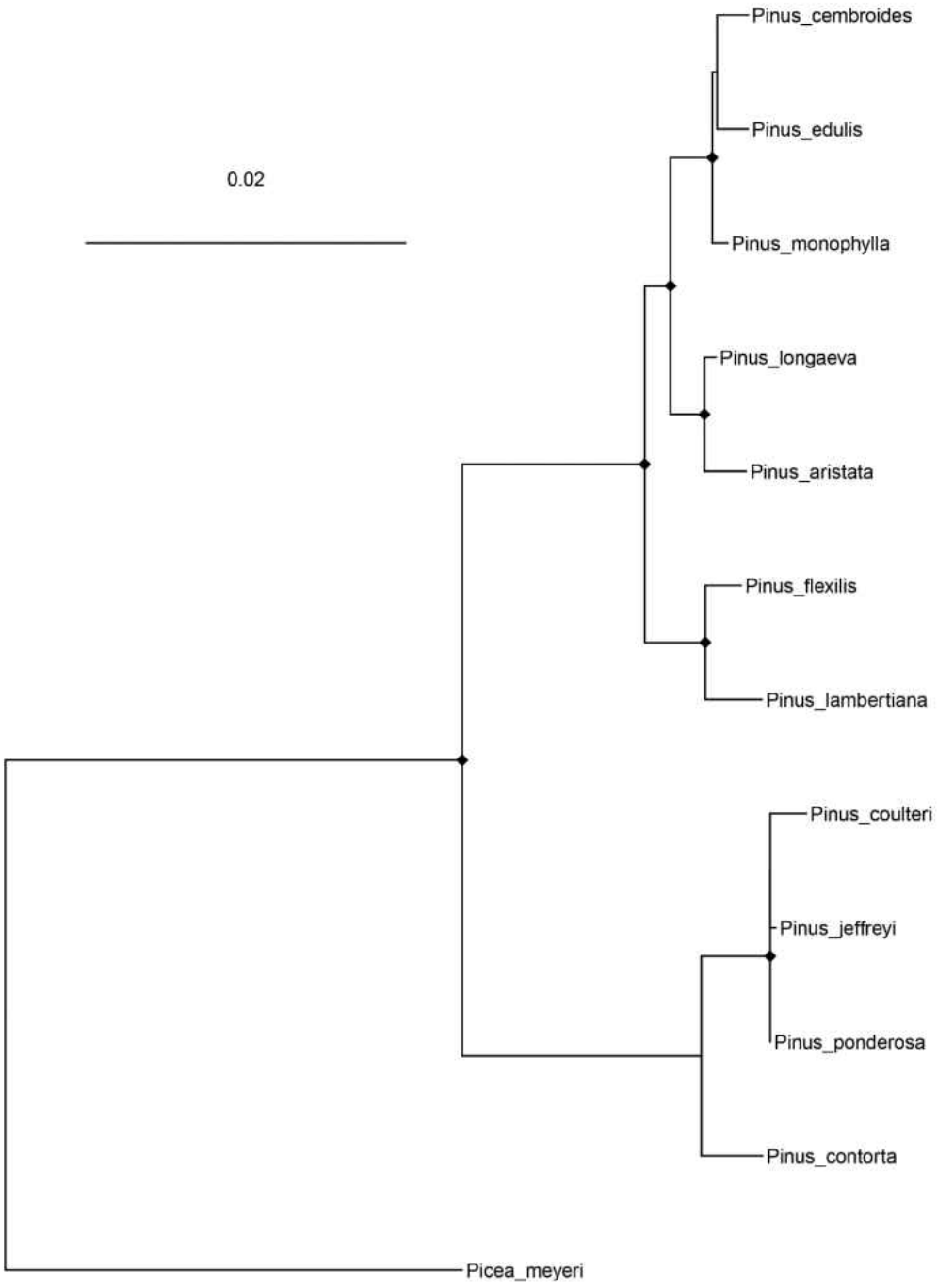

Figure 7: Maximum likelihood phylogeny of *Pinus* species created in RaxML v8.2.10 from three genes: ITS2, rbcL, and matK. Nodes with black diamonds have bootstrap values of 90 or greater. Terminal taxa are labelled by species.

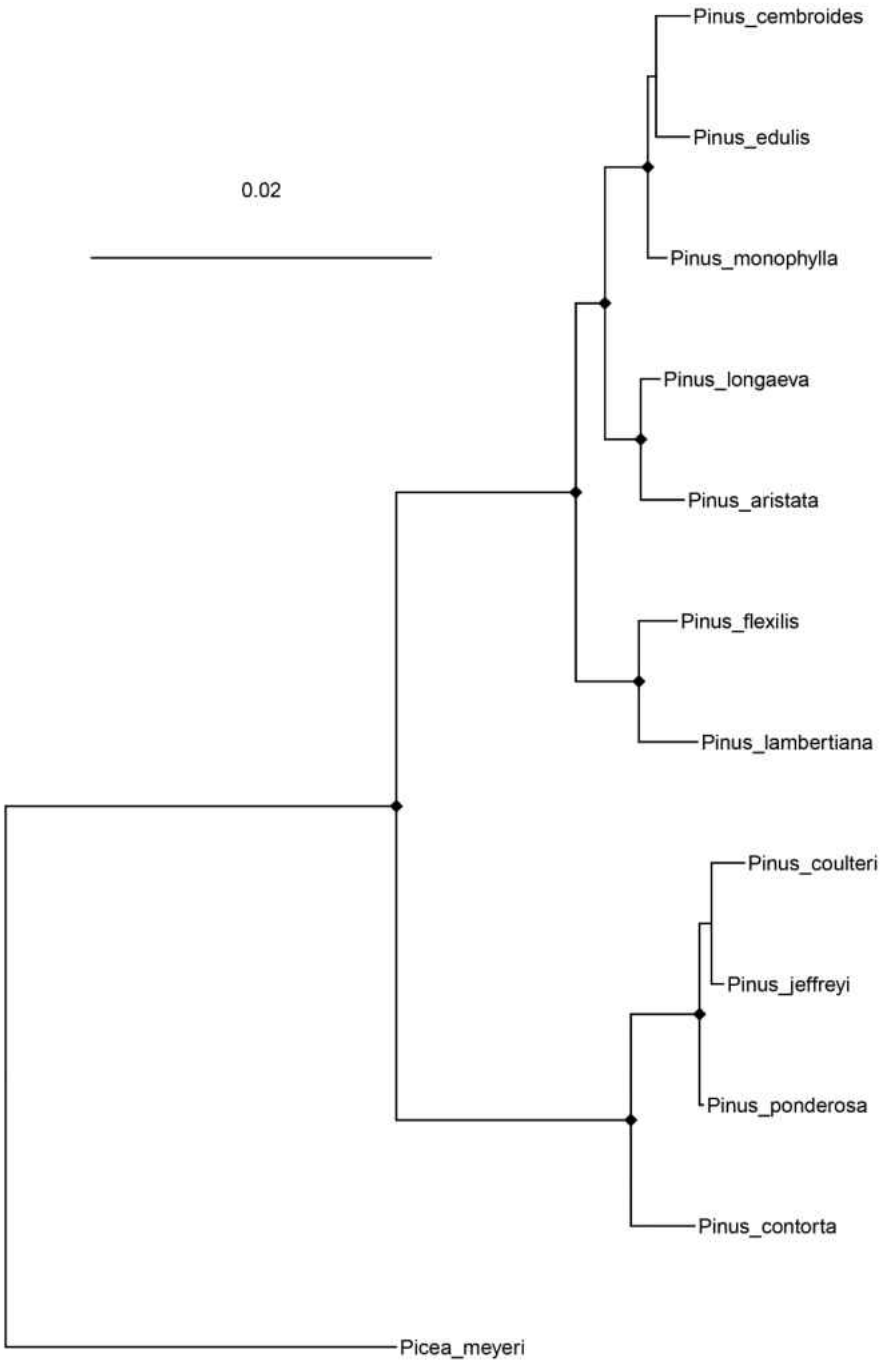

Figure 8: Bayesian phylogeny of *Pinus* species created in MrBayes v3.2.6 from three genes: ITS2, rbcL, and matK. Nodes with black diamonds have posterior probabilities of 95 or greater. Terminal taxa are labelled by species.

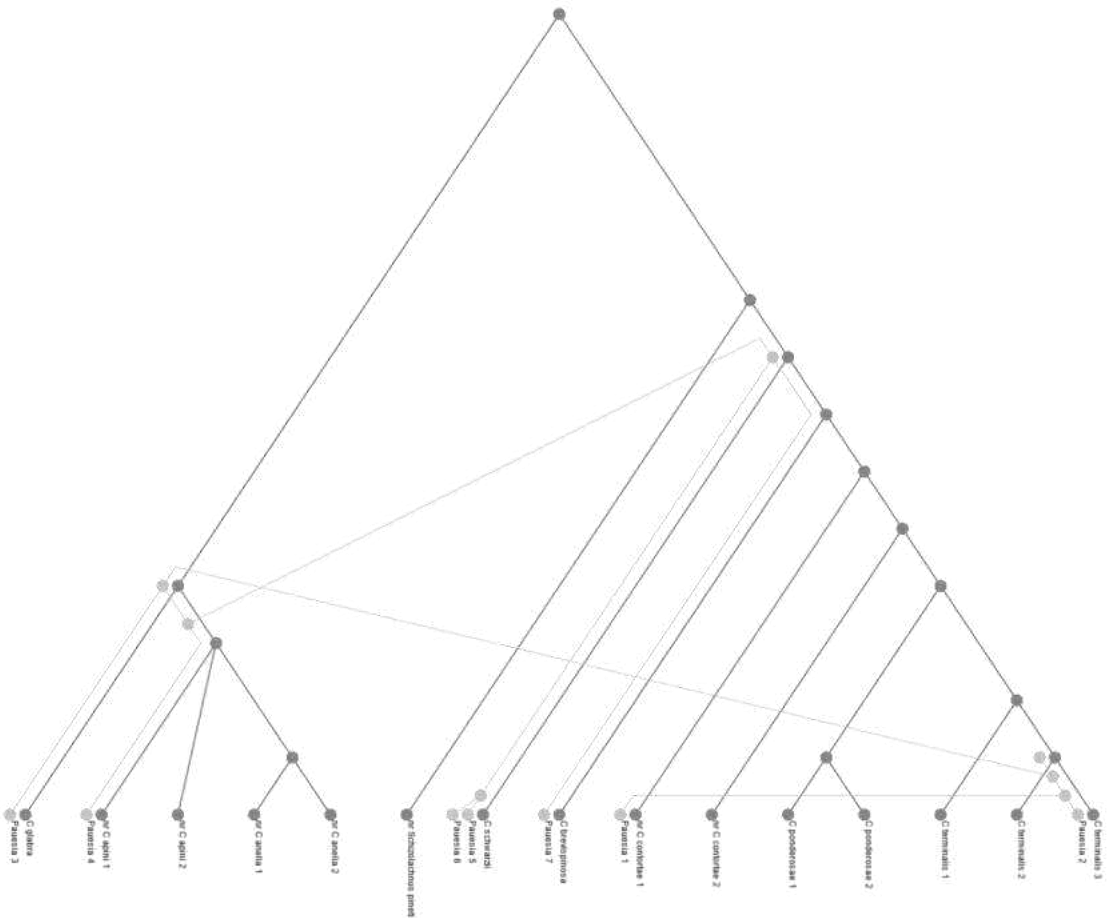

Figure 9: CoRe-Pa lowest cost reconstruction of *Pauesia* and *Cinara* with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 13, and this reconstruction has the following number of events: 2 cospeciation, 2 sorting, 1 duplication, and 3 host switch.

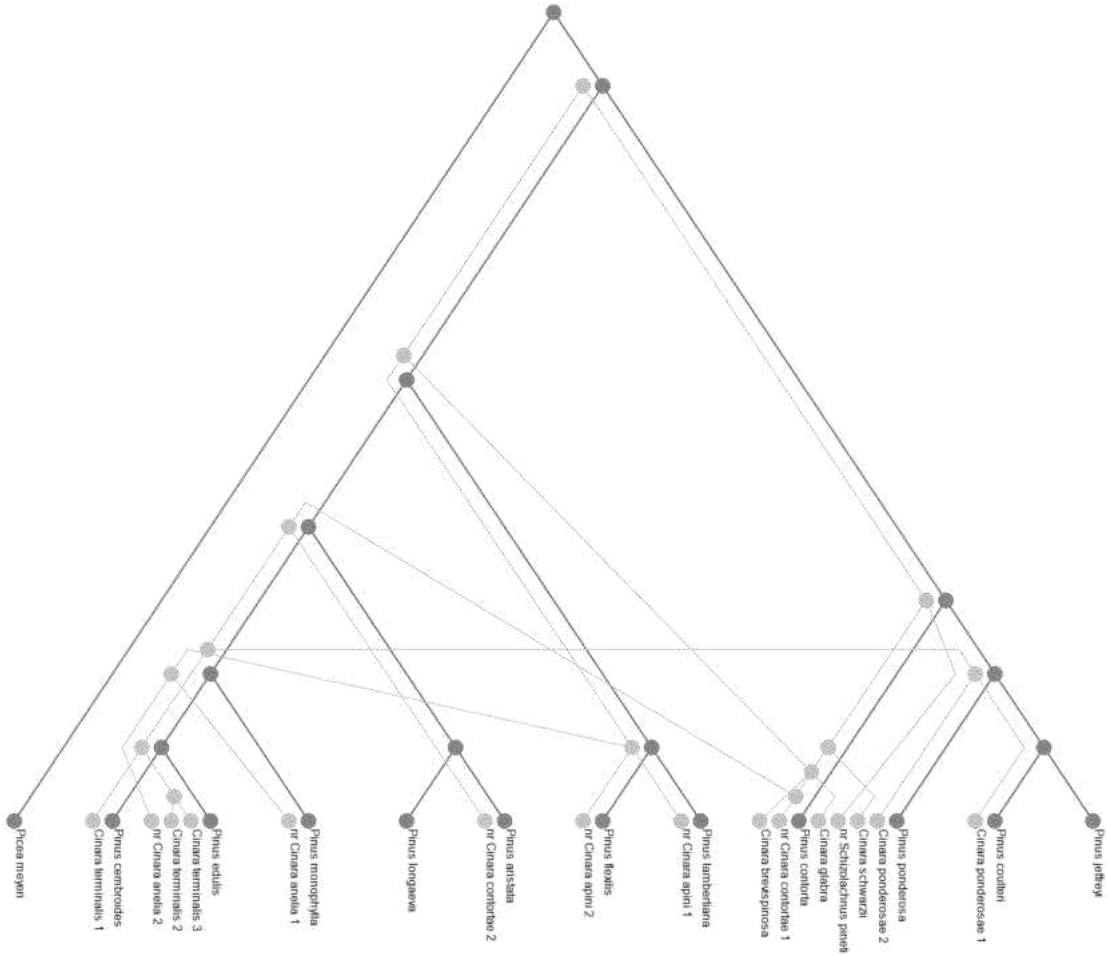

Figure 10: CoRe-Pa lowest cost reconstruction of *Cinara* and *Pinus* with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 25, and this reconstruction has the following number of events: 7 cospeciation, 6 sorting, 2 duplication, and 5 host switch.

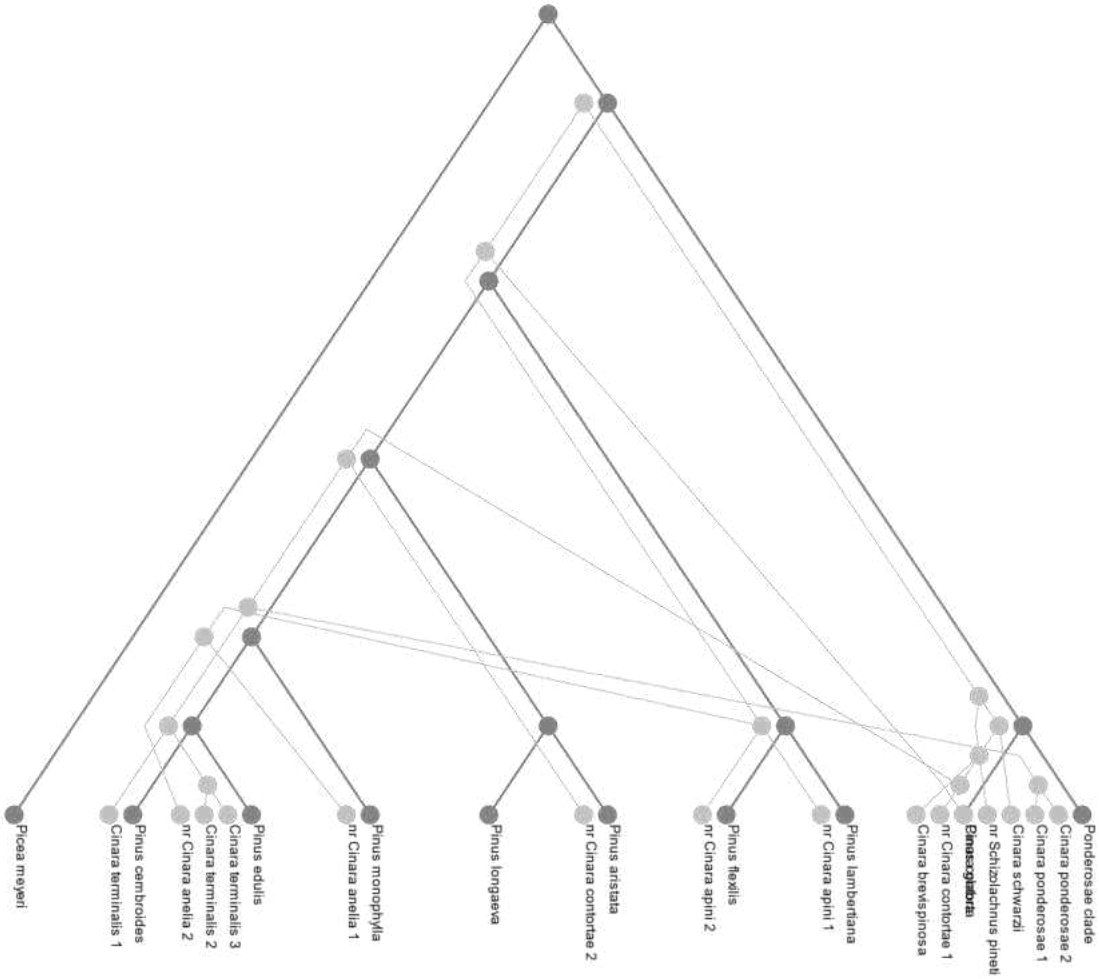

Figure 11: First CoRe-Pa lowest cost reconstruction of *Cinara* and *Pinus* (subsection Ponderosae collapsed) with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 25, and this reconstruction has the following number of events: 6 cospeciation, 5 sorting, 3 duplication, and 5 host switch.

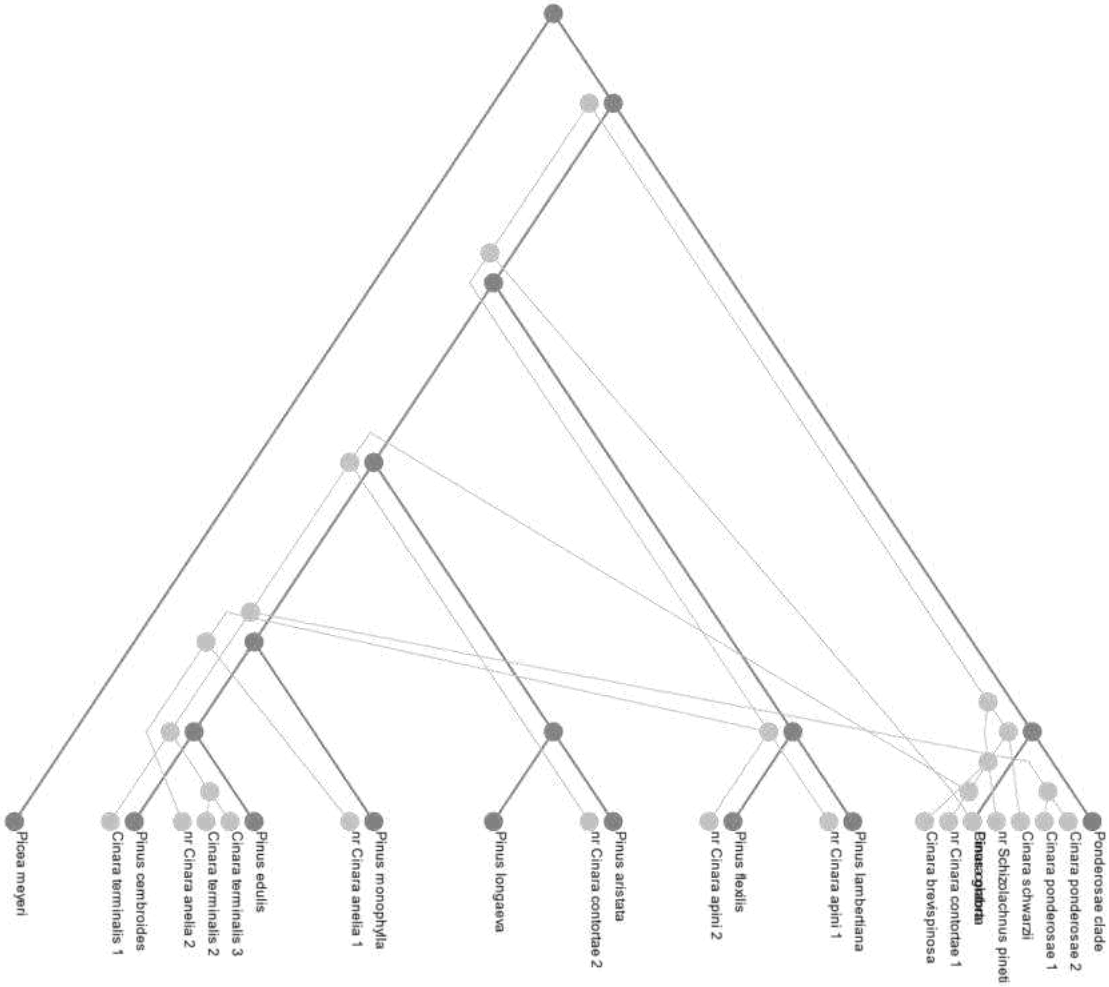

Figure 12: Second CoRe-Pa lowest cost reconstruction of *Cinara* and *Pinus* (subsection Ponderosae collapsed) with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 25, and this reconstruction has the following number of events: 6 cospeciation, 5 sorting, 4 duplication, and 4 host switch.

Figure 13: CoRe-Pa lowest cost reconstruction of *Pauesia* and *Pinus* with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 14, and this reconstruction has the following number of events: 2 cospeciation, 0 sorting, 1 duplication, and 4 host switch.

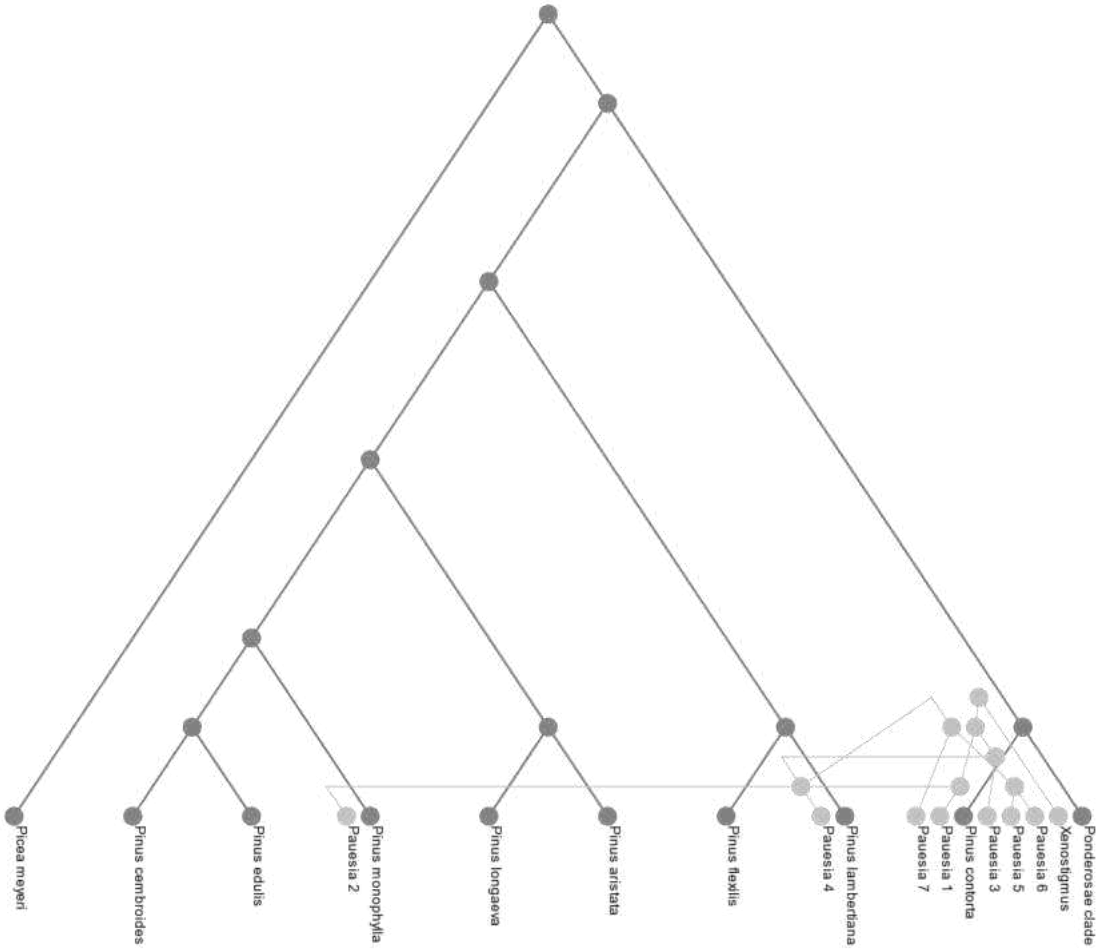

Figure 14: First CoRe-Pa lowest cost reconstruction of *Pauesia* and *Pinus* (subsection Ponderosae collapsed) with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 14, and this reconstruction has the following number of events: 2 cospeciation, 1 sorting, 2 duplication, and 3 host switch.

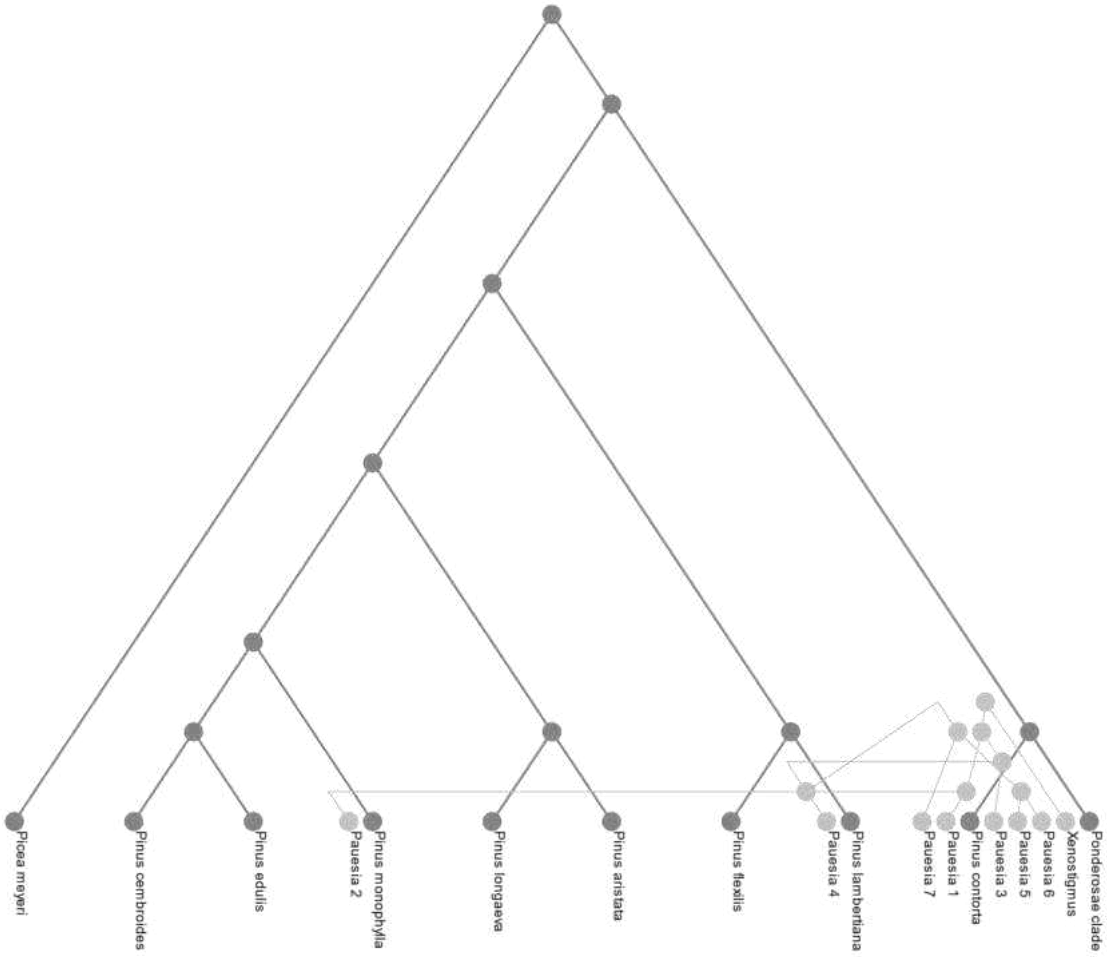

Figure 15: Second CoRe-Pa lowest cost reconstruction of *Pauesia* and *Pinus* (subsection Ponderosae collapsed) with 10,000 simulations, and all other settings standard. The total cost of the reconstruction is 25, and this reconstruction has the following number of events: 6 cospeciation, 5 sorting, 4 duplication, and 4 host switch.

APPENDIX B: TABLES

Table 1: Identification of *Cinara* species used in this study. Identifications were made on the BOLD species level barcodes database. Samples with NA were not placed into a BIN.

Query ID	Best ID	Top %	Low %	BIN	Average distance	Distance to Nearest Neighbor	Nearest Neighbour BIN	Nearest Neighbor ID
AIB107	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB108	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB109	*Cinara ponderosae*	100	96.2	ABY4171	0.07%	2.09%	AAI3985	*Cinara ponderosae*
AIB110	*Cinara brevispinosa*	100	95.85	AAI3975	0.18%	2.89%	AAI3987	*Cinara parvicornis*
AIB111	*Cinara contortae*	100	96.62	NA	NA	NA	NA	NA
AIB112	*Cinara* sp. 3371	100	95.1	NA	NA	NA	NA	NA
AIB113	*Cinara contortae*	97.7	96.77	ABY8476	0.51%	1.93%	AAD8443	*Cinara atlantica*
AIB114	*Cinara ponderosae*	100	96.31	AAI3985	0.15%	2.09%	ABY4171	*Cinara ponderosae*
AIB115	*Cinara anelia*	100	94.27	NA	NA	NA	NA	NA
AIB116	*Cinara schwarzii*	100	94.93	NA	NA	NA	NA	NA
AIB117	*Cinara apini*	99.23	94.35	NA	NA	NA	NA	NA
AIB118	*Cinara anelia*	100	94.27	NA	NA	NA	NA	NA
AIB119	*Cinara apini*	99.23	94.35	NA	NA	NA	NA	NA
AIB120	*Cinara schwarzii*	100	94.93	NA	NA	NA	NA	NA
AIB121	*Cinara contortae*	99.85	96.6	NA	NA	NA	NA	NA
AIB122	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB123	*Cinara glabra*	100	94.75	ABU9394	0.73%	3.05%	ABY8468	*Cinara solitaria*
AIB124	*Cinara brevispinosa*	100	95.85	AAI3975	0.18%	2.89%	AAI3987	*Cinara parvicornis*

AIB126	*Cinara terminalis*	99.85	96.47	ABY8478	0.00%	1.73%	ABY8479	*Cinara terminalis*
AIB127	*Cinara terminalis*	99.85	96.16	ABY8479	0.39%	1.73%	ABY8478	*Cinara terminalis*
AIB129	*Cinara ponderosae*	100	96.2	ABY4171	0.07%	2.09%	AAI3985	*Cinara ponderosae*
AIB130	*Cinara anelia*	100	95.16	NA	NA	NA	NA	NA
AIB132	*Cinara ponderosae*	100	96.16	AAI3985	0.15%	2.09%	ABY4171	*Cinara ponderosae*
AIB133	*Cinara ponderosae*	100	94.01	ABY4171	0.07%	2.09%	AAI3985	*Cinara ponderosae*
AIB134	*Cinara brevispinosa*	100	96.01	AAI3975	0.18%	2.89%	AAI3987	*Cinara parvicornis*
AIB136	*Cinara contortae*	97.38	96.92	ABY8476	0.51%	1.93%	AAD8443	*Cinara atlantica*
AIB138	*Cinara anelia*	100	94.27	NA	NA	NA	NA	NA
AIB139	*Cinara terminalis*	99.08	96.31	ABY8479	0.39%	1.73%	ABY8478	*Cinara terminalis*
AIB140	*Cinara schwarzii*	100	94.93	NA	NA	NA	NA	NA
AIB141	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB143	*Cinara medispinosa*	100	96.76	NA	NA	NA	NA	NA
AIB144	*Cinara medispinosa*	99.85	96.62	NA	NA	NA	NA	NA
AIB148	*Cinara contortae*	97.7	96.77	ABY8476	0.51%	1.93%	AAD8443	*Cinara atlantica*
AIB149	*Cinara contortae*	97.7	96.77	NA	NA	NA	NA	NA
AIB151	*Schizolachnus pineti*	98.99	96.62	ABX5085	0.33%	2.25%	AAF1997	*Schizolachnus curvispinosus*
AIB154	*Cinara terminalis*	99.85	96.15	ABY8479	0.39%	1.73%	ABY8478	*Cinara terminalis*
AIB180	*Cinara contortae*	99.85	96.56	NA	NA	NA	NA	NA
AIB181	*Cinara apini*	100	94.69	NA	NA	NA	NA	NA

AIB183	*Cinara schwarzii*	100	95.18	NA	NA	NA	NA	NA
AIB185	*Cinara anelia*	99.5	95.24	NA	NA	NA	NA	NA
AIB192	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB193	*Cinara terminalis*	99.85	96.77	ABY8478	0%	1.73%	ABY8479	*Cinara terminalis*
AIB195	*Cinara anelia*	100	95.14	NA	NA	NA	NA	NA
AIB201	*Cinara brevispinosa*	100	95.85	ABY8478	0%	1.73%	ABY8479	*Cinara terminalis*
AIB207	*Cinara* sp. 3371	100	95.1	NA	NA	NA	NA	NA

Table 2: Collection record of wasps (*Pauesia* spp.) and aphids (*Cinara* spp.) collected in summer 2017 and included in phylogenomic analyses. The species of pine tree, the state and locality, and the GPS coordinates of the collection are included. An asterisk beside the sample ID for wasps and aphids indicate which individuals were used to test the barcoding gap species concept. Species delimitation results are shown for both the wasps and aphids.

Wasp Sample ID	**Wasp Species Delimitation**	**Aphid Sample ID**	**Aphid Species Delimitation**	**Pine Species**	**State**	**Locality**	**Latitude**	**Longitude**
NA	NA	AIB113*	nr. *Cinara contortae* 2	*Pinus aristata*	CO	Rio Grande National Forest	37.47772	-106.47993
NA	NA	AIB136*	*nr. Cinara contortae* 2	*Pinus aristata*	CO	Rio Grande National Forest	37.46942	-106.49023
NA	NA	AIB148*	nr. *Cinara contortae* 2	*Pinus aristata*	CO	Rio Grande National Forest	37.48079	-106.46378
NA	NA	AIB149*	nr. *Cinara contortae* 2	*Pinus aristata*	CO	Rio Grande National Forest	37.48079	-106.46378
NA	NA	AIB126*	*Cinara terminalis* 1	*Pinus cembroides*	AZ	Coronado National Forest	31.72671	-110.87431
AIB054*	*Pauesia* sp. 7	AIB110*	*Cinara brevispinosa*	*Pinus contorta*	CO	Routt National Forest	40.06723	-106.40648
AIB055*	*Pauesia* sp. 1	AIB111*	nr. *Cinara contortae* 1	*Pinus contorta*	CO	Routt National Forest	40.07423	-106.39637
AIB068*	*Pauesia* sp. 1	AIB121*	nr. *Cinara contortae* 1	*Pinus contorta*	CA	Sequoia National Park	36.08199	-118.32591
AIB072*	*Pauesia* sp. 7	AIB124*	*Cinara brevispinosa*	*Pinus contorta*	CO	Routt National Forest	40.07431	-106.39639
AIB169	*Pauesia* sp. 7	AIB134*	*Cinara brevispinosa*	*Pinus contorta*	CO	San Isabel National Forest	38.58947	-106.26357
AIB171	*Pauesia* sp. 7	NA	NA	*Pinus contorta*	CO	Routt National Forest	40.06715	-106.40703
AIB172	*Pauesia* sp. 7	NA	NA	*Pinus contorta*	CO	Routt National Forest	40.06715	-106.40703
AIB200	*Pauesia* sp. 7	AIB201*	*Cinara brevispinosa*	*Pinus contorta*	CO	Routt National Forest	40.07426	-106.39647
NA	NA	AIB143*	nr. *Cinara contortae* 1	*Pinus contorta*	CO	Rocky Mountain National Park	40.19228	-105.30807
NA	NA	AIB180*	*nr. Cinara contortae* 1	*Pinus contorta*	CO	Sequoia National Park	36.10297	-118.33864

Wasp Sample ID	Wasp Species Delimitation	Aphid Sample ID	Aphid Species Delimitation	Pine Species	State	Locality	Latitude	Longitude
NA	NA	AIB186*	nr. *Cinara contortae* 1	*Pinus contorta*	CA	Stanislaus National Forest	38.10748	-119.54498
NA	NA	AIB132*	*Cinara ponderosae* 1	*Pinus coulteri*	CA	San Bernardino National Forest	34.16700	-117.11131
AIB047*	*Pauesia* sp. 2	AIB107*	nr. *Cinara anelia* 2	*Pinus edulis*	NM	Heron Lake State Park	36.41698	-106.39609
AIB048*	*Pauesia* sp. 2	NA	NA	*Pinus edulis*	NM	Heron Lake State Park	36.41698	-106.39609
AIB050*	*Pauesia* sp. 2	AIB108*	nr. *Cinara anelia* 2	*Pinus edulis*	NM	Heron Lake State Park	36.41698	-106.39609
AIB069*	*Pauesia* sp. 2	AIB154*	*Cinara terminalis* 3	*Pinus edulis*	NM	Heron Lake State Park	36.41505	-106.39803
AIB070	*Pauesia* sp. 2	AIB122*	nr. *Cinara anelia* 2	*Pinus edulis*	NM	Heron Lake State Park	36.41698	-106.39609
AIB075	*Pauesia* sp. 2	AIB195*	nr. *Cinara anelia* 2	*Pinus edulis*	NM	Heron Lake State Park	36.41705	-106.39648
NA	NA	AIB130*	nr. *Cinara anelia* 2	*Pinus edulis*	UT	Dixie National Forest	37.47639	-112.00496
NA	NA	AIB131	*Cinara terminalis* 3	*Pinus edulis*	UT	Dixie National Forest	37.47468	-111.58866
NA	NA	AIB192*	nr. *Cinara anelia* 2	*Pinus edulis*	NM	Gila National Forest	33.23666	-108.49085
NA	NA	AIB193*	*Cinara terminalis* 1	*Pinus edulis*	NM	Gila National Forest	33.23658	-108.49072
NA	NA	AIB141*	nr. *Cinara anelia* 2	*Pinus flexilis*	UT	Dixie National Forest	37.44084	-111.52210
NA	NA	AIB144*	nr. *Cinara contortae* 1	*Pinus flexilis*	CO	Rocky Mountain National Park	40.14527	-105.38940
NA	NA	AIB150	nr. *Cinara apini* 2	*Pinus flexilis*	CO	San Isabel National Forest	39.00137	-106.20421
NA	NA	AIB181*	*nr Cinara apini* 2	*Pinus flexilis*	CO	San Isabel National Forest	39.00137	-106.20421
AIB059	*Xenostigmus* sp.	AIB114*	*Cinara ponderosae* 1	*Pinus jeffreyi/ Pinus ponderosa*	CA	Stanislaus National Park	38.10322	-120.01912
AIB061*	*Pauesia* sp. 6	AIB116*	*Cinara schwarzii*	*Pinus jeffreyi/ Pinus ponderosa*	CA	San Bernardino National Forest	34.15901	-116.56857
AIB067*	*Pauesia* sp. 6	AIB120*	*Cinara schwarzii*	*Pinus jeffreyi/ Pinus ponderosa*	CA	San Bernardino National Forest	34.15901	-116.56857
AIB076*	*Pauesia* sp. 6	AIB140*	*Cinara schwarzii*	*Pinus jeffreyi/ Pinus ponderosa*	CA	San Bernardino National Forest	34.15901	-116.56857

Wasp Sample ID	Wasp Species Delimitation	Aphid Sample ID	Aphid Species Delimitation	Pine Species	State	Locality	Latitude	Longitude
NA	NA	AIB191	nr. *Schizolachnus piniradiatae*	*Pinus jeffreyi/ Pinus ponderosa*	CA	San Bernardino National Forest	34.15901	-116.56857
AIB062*	*Pauesia* sp. 4	AIB117*	nr. *Cinara apini* 1	*Pinus lambertiana*	CA	San Bernardino National Forest	34.14697	-117.03836
AIB065*	*Pauesia* sp. 4	AIB119*	nr. *Cinara apini* 1	*Pinus lambertiana*	CA	San Bernardino National Forest	34.14714	-117.03758
NA	NA	AIB127*	*Cinara terminalis* 3	*Pinus longaeva*	UT	Dixie National Forest	37.44099	-111.52252
AIB060*	*Pauesia* sp. 2	AIB115*	nr. *Cinara anelia* 1	*Pinus monophylla*	CA	Los Padres National Forest	34.48124	-119.00496
AIB063	*Pauesia* sp. 2	AIB118*	nr. *Cinara anelia* 2	*Pinus monophylla*	CA	Los Padres National Forest	34.50425	-119.05209
AIB073	*Pauesia* sp. 4	AIB125	*Cinara edulis*	*Pinus monophylla*	CA	Joshua Tree National Park	34.02516	-116.04211
NA	NA	AIB138*	nr. *Cinara anelia* 1	*Pinus monophylla*	CA	Death Valley National Park	36.13676	-117.04110
NA	NA	AIB139*	*Cinara terminalis* 2	*Pinus monophylla*	CA	Joshua Tree National Park	34.02516	-116.04211
NA	NA	AIB185*	nr. *Cinara anelia* 2	*Pinus monophylla*	CA	Joshua Tree National Park	34.02516	-116.04211
AIB052	*Xenostigmus* sp.	AIB109*	*Cinara ponderosae* 2	*Pinus ponderosa*	NM	Heron Lake State Park	36.41125	-106.39753
AIB056*	*Pauesia* sp. 5	AIB112*	*Cinara schwarzii*	*Pinus ponderosa*	CO	Rocky Mountain National Park	40.25882	-105.30229
AIB071	*Pauesia* sp. 3	AIB123*	*Cinara glabra*	*Pinus ponderosa*	NM	Heron Lake State Park	40.43146	-105.50345
NA	NA	AIB129*	*Cinara ponderosae* 2	*Pinus ponderosa*	UT	Dixie National Forest	37.47267	-111.59944
NA	NA	AIB133*	*Cinara ponderosae* 2	*Pinus ponderosa*	CO	White River National Forest	39.09737	-107.14959
NA	NA	AIB151*	*nr. Schizolachnus pineti*	*Pinus ponderosa*	CO	White River National Forest	39.09729	-107.14968

Wasp Sample ID	Wasp Species Delimitation	Aphid Sample ID	Aphid Species Delimitation	Pine Species	State	Locality	Latitude	Longitude
NA	NA	AIB183*	*Cinara schwarzii*	*Pinus ponderosa*	CO	Rio Grande National Forest	37.46938	-106.49043
NA	NA	AIB207*	*Cinara schwarzii*	*Pinus ponderosa*	CO	Rocky Mountain National Park	40.25817	-105.30632
AIB190	*Pauesia* sp. 6	NA	NA	*Pinus jeffreyi*/ *Pinus ponderosa*	CA	San Bernardino National Forest	34.15901	-116.56857

Table 3: GenBank accession numbers for the *Pinus* sequences used in phylogenetic analyses. ITS2 was unavailable for *Pinus jeffreyi* and *Pinus longaeva*.

Species	**ITS2**	**rbcL**	**matk**
Pinus ponderosa	GQ434746.1	AY497234.1	AY497270.1
Pinus jeffreyi	NA	AY497235.1	AY497271.1
Pinus aristata	AF037000.2	AY115758.1	AY115794.1
Pinus edulis	AF343993.1	AY115739.1	AY115765.1
Pinus contorta	U23956.1	AY497230.1	AY497266.1
Pinus monophylla	AF343986.1	AY115741.1	AY115768.1
Pinus lambertiana	AF036990.1	AY497224.1	AY497260.1
Pinus coulteri	AF037013.1	AY724759.1	AY724751.1
Pinus flexilis	AF344001.1	AY497222.1	AY497258.1
Pinus longaeva	NA	AY115759.1	AY115796.1
Pinus cembroides	AF343983.1	AY115751.1	AY115781.1
Picea meyeri	GQ865721.1	KP088721.1	AY729948.1

Table 4: Interspecific and intraspecific (in bold) genetic distances of COI under a K2P model between monophyletic groups of *Pauesia* species. Specimens included in this analysis are indicated in Table 1 with an asterisk. *Pauesia* sp. 3 was represented by one specimen and COI was not successfully sequenced for this individual, therefore no genetic distances were available between *Pauesia* sp. 3 and the other six species.

		1	2	3	4	5	6	7
1	*Pauesia* sp. 1	**0.00295**						
2	*Pauesia* sp. 2	0.01552	**0.00299**					
3	*Pauesia* sp. 3	NA	NA	**NA**				
4	*Pauesia* sp. 4	0.11227	0.10796	NA	**NA**			
5	*Pauesia* sp. 5	0.12955	0.12867	NA	0.05357	**0.00075**		
6	*Pauesia* sp. 6	0.12876	0.12769	NA	0.05256	0.00780	**0.00000**	
7	*Pauesia* sp. 7	0.11739	0.11303	NA	0.03774	0.04570	0.04788	**0.00000**

Table 5: Interspecific and intraspecific (in bold) genetic distances of COI under a K2P model between monophyletic groups of *Cinara* species. Specimens included in this analysis are indicated in Table 1 with an asterisk.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	*Cinara terminalis* 1	**0.295**														
2	*Cinara terminalis* 2	1.037	**NA**													
3	*Cinara terminalis* 3	2.244	2.093	**0.295**												
4	*Cinara ponderosae* 1	3.323	2.859	2.556	**0**											
5	*Cinara ponderosae* 2	3.71	3.244	2.939	2.015	**0.147**										
6	nr. *Cinara contortae* 1	2.938	2.785	2.482	2.789	3.018	**0.221**									
7	nr. *Cinara contortae* 2	2.965	2.651	2.905	2.965	3.195	2.704	**0.326**								
8	*Cinara brevispinosa*	4.842	4.368	3.901	4.049	4.598	4.728	4.635	**0.074**							
9	*Cinara schwarzii*	6.134	5.784	6.055	6.05	5.488	6.099	5.584	5.102	**0.139**						
10	nr. *Cinara anelia* 1	7.325	7.515	6.927	6.945	6.87	7.485	7.098	7.469	5.909	**0.121**					
11	nr. *Cinara anelia* 2	8.482	8.731	8.074	7.737	7.665	8.621	7.699	8.189	6.531	1.568	**0**				
12	nr. *Cinara apini* 1	8.479	8.728	7.741	8.065	7.993	8.617	8.294	7.856	6.527	1.567	1.64	**0**			
13	nr. *Cinara apini* 2	8.704	8.613	8.093	8.262	8.185	8.848	8.193	7.866	6.446	1.129	1.106	1.264	**NA**		
14	*Cinara glabra*	6.766	6.762	6.289	7.09	6.848	7.633	7.083	6.896	6.521	6.84	7.767	7.101	7.762	**NA**	
15	*Schizolachnus pineti*	7.652	7.402	6.923	7.079	6.998	7.45	7.233	6.396	5.579	6.015	6.767	6.602	6.875	5.011	**NA**

Table 6: Results of the two distance-based methods between the three interacting taxa. PACo found all interacting taxa had congruent phylogenies, while AxParafit/AxPcoords found all but wasp and pine had congruent phylogenies. The number of significant links between all taxa indicated from AxParafit/AxPcoords are shown.

	PACo		AxParafit/AxPcoords		
Interaction	P value	m^2	P value	ParaFitGlobal	Significant links
wasp-aphid	0.00784	19.86853	0.01912	1190.81	10/22
aphid-pine	0.00001	0.00970	0.00265	0.15064	31/44
wasp-pine	0.00273	0.00273	0.05174	0.39489	8/22

Table 7: Results of reconstructions and permutations from CoRe-Pa for each of the interacting pairs of taxa. The number of events (cospeciation, sorting, duplication, and host switch) and the total cost are shown. The results of 10,000 random cycles are shown. The average +/- standard deviation for 100,000 permutations with random interactions are shown for each interacting pair. Collapsed datasets are those where the clades of the phylospecies are collapsed for the reconstruction. An asterisk shows where the number of events or total cost falls outside of the mean +/- standard deviation of the permutations, indicating that the results of the phylogenies and interactions in this study differ from random.

	Cospeciation	Sorting	Duplication	Host Switch	Total Cost
Pauesia-Cinara	2	2	1	3	13
Pauesia-Cinara permutations	2.10 +/- 0.76	1.89 +/- 1.60	1.01 +/- 0.12	2.88 +/- 0.77	12.57 +/- 1.50
Cinara-Pinus	7*	6*	2	5*	25*
Cinara-Pinus permutations	3.81 +/- 1.25	2.66 +/- 2.18	3.22 +/- 1.41	6.97 +/- 1.11	30.02 +/- 1.81
Cinara-Pinus Ponderosae collapsed	6*	4	3	5*	25*
Cinara-Pinus Ponderosae collapsed	6*	5*	4	4*	25*
Cinara-Pinus Ponderosae collapsed permutations	3.78 +/- 1.17	2.45 +/- 2.09	3.92 +/- 1.34	6.30 +/- 1.13	29.19 +/- 1.72
Pauesia-Pinus	2	2	0	4	14
Pauesia-Pinus permutations	1.67 +/- 0.89	1.43 +/- 1.51	0.74 +/- 0.78	3.59 +/- 0.72	13.67 +/- 1.57
Pauesia-Pinus Ponderosae collapsed	2	0	1	4	14
Pauesia-Pinus Ponderosae collapsed	2	1	2	3	14
Pauesia-Pinus Ponderosae collapsed permutations	1.88 +/- 0.89	1.46 +/- 1.54	1.60 +/- 0.71	3.52 +/- 0.73	15.22 +/- 1.59

LIST OF REFERENCES

Abrahamson, W.G., Blair, C.P., 2008. Sequential radiation through host-race formation: herbivore diversity leads to diversity in natural enemies. In: Tilmon, K.J. (Ed.), Specialization, Speciation, and Radiation: the Evolutionary Biology of Herbivorous Insects. University of California Press, Berkeley, pp. 188-202.

Agosta, S.J., Janz, N., Brooks, D.R., 2010. How specialists can be generalists: resolving the "parasite paradox" and implications for emerging infectious disease. Zoologia 27, 151-162. https://doi.org/10.1590/s1984-46702010000200001.

Althoff, D.M., 2008. A test of host-associated differentiation across the 'parasite continuum' in the tri-trophic interaction among yuccas, bogus yucca moths, and parasitoids. Mol. Ecol. 17, 3917-3927. https://doi.org/10.1111/j.1365-294X.2008.03874.x.

Althoff, D.M., Segraves, K.A., Johnson, M.T.J., 2014. Testing for coevolutionary diversification: Linking pattern with process. Trends Ecol. Evol. 29, 82-89. https://doi.org/10.1016/j.tree.2013.11.003.

Balbuena, J.A., Míguez-Lozano, R., Blasco-Costa, I., 2013. PACo: A novel procrustes application to cophylogenetic analysis. Plos One 8, e61048. https://doi.org/10.1371/journal.pone.0061048.

Banbury, B.L., Leache, A., 2014. phrynomics: An SNP toolkit.

Bentz, S., Sinnappah-Kang, N.D., Lim, L.-H.S., Lebedev, B., Combes, C., Verneau, O., 2006. Historical biogeography of amphibian parasites, genus *Polystoma*

(Monogenea: Polystomatidae). J. Biogeogr. 33, 742-749. https://doi.org/10.1111/j.1365-2699.2005.01402.x.

Bernays, E.A., Graham, M., 1988. On the evolution of host specificity in phytophagous arthropods. Ecology 69, 886-892.

Blackman, R., Eastop, V.F., 2000. Aphids on the world's trees: An identification and information guide.

Boring, C.A.B., Sharanowski, B.J., Sharkey, M.J., 2011. Maxfischeriinae: a new braconid subfamily (Hymenoptera) with highly specialized egg morphology. . Syst. Entomol. 36, 529-548.

Bradley, G.A., 1951. A field key to the species of *Cinara* (Homoptera: Aphididae) of Canada east of the rockies. Can. Entomol. 83, 333-335. https://doi.org/10.4039/Ent83333-12.

Burke, G.R., Normark, B.B., Favret, C., Moran, N.A., 2009. Evolution and diversity of facultative symbionts from the aphid subfamily Lachninae. Appl. Environ. Microbiol. 75, 5328-5335.

Cariou, M., Duret, L., Charlat, S., 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. Ecol. Evol. 3, 846-852. https://doi.org/10.1002/ece3.512.

Charleston, M.A., Robertson, D.L., 2002. Preferential Host Switching by Primate Lentiviruses Can Account for Phylogenetic Similarity with the Primate Phylogeny. Syst. Biol. 51, 528-535. https://doi.org/10.1080/10635150290069940.

Chen, R., Favret, C., Jiang, L., Wang, Z., Qiao, G., 2016a. An aphid lineage maintains a bark-feeding niche while switching to and diversifying on conifers. Cladistics 32, 555-572. https://doi.org/10.1111/cla.12141.

Chen, R., Jiang, L.-Y., Chen, J., Qiao, G.-X., 2016b. DNA barcoding reveals a mysterious high species diversity of conifer-feeding aphids in the mountains of southwest China. Sci. Rep-UK 6, 20123. https://doi.org/10.1038/srep20123.

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2016. GenBank. Nucleic Acids Res. 44, D67-D72. https://doi.org/10.1093/nar/gkv1276.

Clark, M.A., Moran, N.A., Baumann, P., Wernegreen, J.J., 2000. Cospeciation between bacterial endosymbionts (*Buchnera*) and a recent radiation of aphids (*Uroleucon*) and pitfalls of testing for phylogenetic congruence. Evolution 54, 517-525.

Dal Molin, A., Menard, K., 2010. Non-destructive DNA extraction protocol for minute insects. protocols.io (1535930959).

de Vienne, D.M., Refrégier, G., López-Villavicencio, M., Tellier, A., Hood, M.E., Giraud, T., 2013. Cospeciation vs host-shift speciation: Methods for testing, evidence from natural associations and relation to coevolution. New Phytol. 198, 347-385. https://doi.org/10.1111/nph.12150.

Demastes, J.W., Hafner, M.S., 1993. Cospeciation of pocket gophers (*Geomys*) and their chewing lice (*Geomydoecus*). J. Mammal. 74, 521-530. https://doi.org/10.2307/1382271.

Deng, J., Yu, F., Li, H.-B., Gebiola, M., Desdevises, Y., Wu, S.-A., Zhang, Y.-Z., 2013. Cophylogenetic relationships between *Anicetus* parasitoids (Hymenoptera:

Encyrtidae) and their scale insect hosts (Hemiptera: Coccidae). BMC Evol. Biol. 13, 275.

Derocles, S.A.P., Plantegenest, M., Rasplus, J.-Y., Marie, A., Evans, D.M., Lunt, D.H., Le Ralec, A., 2016. Are generalist Aphidiinae (Hym. Braconidae) mostly cryptic species complexes? Syst. Entomol. 41, 379-391. https://doi.org/10.1111/syen.12160.

Dolphin, K., Quicke, D.L.J., 2001. Estimating the global species richness of an incompletely described taxon: An example using parasitoid wasps (Hymenoptera: Braconidae). Biol. J. Linn. Soc. 73, 279-286. https://doi.org/10.1006/bijl.2001.0537.

Drès, M., Mallet, J., 2002. Host races in plant–feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 357, 471. https://doi.org/10.1098/rstb.2002.1059.

Eaton, D.A.R., 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30, 1844-1849. https://doi.org/10.1093/bioinformatics/btu121.

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

Farrell, B.D., 1998. "Inordinate fondness" explained: why are there so many beetles? Science 281, 555 - 559.

Favret, C., 2005. A new non-destructive DNA extraction and specimen clearing technique for aphids (Hemiptera). Proc. Entomol. Soc. Wash. 107, 469-470.

Favret, C., 2013. Aphid species file. Version 5.0/5.0. [May 4, 2018]. http://Aphid.SpeciesFile.org.

Favret, C., Voegtlin, D.J., 2004a. Host-based morphometric differentiation in three *Cinara* species (Insecta: Hemiptera: Aphididae) feeding on *Pinus edulis* and *P. monophylla*. West. N. Am. Nat. 64, 364-375. https://doi.org/10.1603/0013-8746(2004)097[1165:AROTCS]2.0.CO;2.

Favret, C., Voegtlin, D.J., 2004b. Speciation by host-switching in pinyon *Cinara* (Insecta: Hemiptera: Aphididae). Mol. Phylogenet. Evol. 32. https://doi.org/10.1016/j.ympev.2003.12.005.

Favret, C., Voegtlin, D.J., 2014. A revision of the *Cinara* species (Hemiptera: Aphididae) of the United States pinyon pines. Ann. Entomol. Soc. Am. 97, 1165-1197. https://doi.org/10.1603/0013-8746(2004)097[1165:AROTCS]2.0.CO;2.

Feder, J.L., Forbes, A.A., 2010. Sequential speciation and the diversity of parasitic insects. Ecol. Entomol. 35, 67-76. https://doi.org/10.1111/j.1365-2311.2009.01144.x.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294.

Forbes, A.A., Bagley, R.K., Beer, M.A., Hippee, A.C., Widmayer, H.A., 2018. Quantifying the unquantifiable: Why Hymenoptera — not Coleoptera — is the most speciose animal order. BMC Ecol. https://doi.org/10.1186/s12898-018-0176-x.

Forbes, A.A., Powell, T.H.Q., Stelinski, L.L., Smith, J.J., Feder, J.L., 2009. Sequential sympatric speciation across trophic levels. Science 323, 776-779. https://doi.org/10.1126/science.1166981.

Funk, D.J., Filchak, K.E., Feder, J.L., 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116, 251-267.

Futuyma, D.J., Agrawal, A.A., 2009. Macroevolution and the biological diversity of plants and herbivores. Proceedings of the National Academy of Sciences 106, 18054-18061. https://doi.org/10.1073/pnas.0904106106.

Gernandt, D.S., Hernández-León, S., Salgado-Hernández, E., Pérez de La Rosa, J.A., 2009. Phylogenetic relationships of *Pinus* subsection Ponderosae inferred from rapidly evolving cpDNA regions. Syst. Bot. 34, 481-491. https://doi.org/10.1600/036364409789271290.

Gernandt, D.S., López, G.G., García, S.O., Liston, A., 2005. Phylogeny and classification of *Pinus*. Taxon 54, 29-42. https://doi.org/10.2307/25065300.

Gilman, I.S., Tank, D.C., 2018. Species tree estimation using ddRADseq data from historical specimens confirms the monophyly of highly disjunct species of *Chloropyron* (Orobanchaceae). Syst. Bot. 43, 701-708. https://doi.org/10.1600/036364418X697418.

Gottschling, M., Göker, M., Stamatakis, A., Bininda-Emonds, O.R.P., Nindl, I., Bravo, I.G., 2011. Quantifying the phylodynamic forces driving papillomavirus evolution. Mol. Biol. Evol. 28, 2101-2113. https://doi.org/10.1093/molbev/msr030.

Hall, A.A., Steinbauer, M.J., Taylor, G.S., Johnson, S.N., Cook, J.M., Riegler, M., 2017. Unravelling mummies: cryptic diversity, host specificity, trophic and coevolutionary interactions in psyllid–parasitoid food webs. BMC Evol. Biol. 17, 127.

Haller, J.R., 1962. Variation and hybridization in ponderosa and jeffrey pines. Univ. Calif. Publ. Bot. 34, 123-165.

Hamerlinck, G., Hulbert, D., Hood, G.R., Smith, J.J., Forbes, A.A., 2016. Histories of host shifts and cospeciation among free-living parasitoids of *Rhagoletis* flies. J. Evol. Biol. 29, 1766–1779.

Hernández-León, S., Gernandt, D.S., Pérez de la Rosa, J.A., Jardón-Barbolla, L., 2013. Phylogenetic relationships and species delimitation in *Pinus* section Trifoliae inferrred from plastid DNA. Plos One 8, e70501. https://doi.org/10.1371/journal.pone.0070501.

Hoberg, E.P., Brooks, D.R., 2008. A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host–parasite systems. J. Biogeogr. 35, 1533-1550. https://doi.org/10.1111/j.1365-2699.2008.01951.x.

Hosokawa, T., Kikuchi, Y., Nikoh, N., Shimada, M., Fukatsu, T., 2006. Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. PLoS Biol. 4, e337.

Hughes, J., Kennedy, M., Johnson, K.P., Palma, R.L., Page, R.D.M., 2007. Multiple cophylogenetic analyses reveal frequent cospeciation between pelecaniform birds

and *Pectinopygus* lice. Syst. Biol. 56, 232-251. https://doi.org/10.1080/10635150701311370.

Janz, N., Nyblom, K., Nylin, S., 2001. Evolutionary dynamics of host-plant specialization: A case study of the tribe Nymphalini. Evolution 55, 783-796. https://doi.org/10.1111/j.0014-3820.2001.tb00814.x.

Janz, N., Nylin, S., 2008. The oscillation hypothesis of host-plant range and speciation. Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects. University of California Press, Berkeley, pp. 203-215.

Janz, N., Nylin, S., Wahlberg, N., 2006. Diversty begets diversity: Host expansions and the diversification of plant-feeding insects. BMC Evol. Biol. 6, 4. https://doi.org/10.1186/1471-2148-6-4.

Jousselin, E., Cruaud, A., Genson, G., Chevenet, F., Foottit, R.G., Cœur d'acier, A., 2013. Is ecological speciation a major trend in aphids? Insights from a molecular phylogeny of the conifer-feeding genus *Cinara*. Front. Zool. 10, 56. https://doi.org/10.1186/1742-9994-10-56.

Kahle, D., Wickham, H., 2013. ggmap: Spatial visualization with ggplot2. The R J. 5, 144-161.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647-1649. https://doi.org/10.1093/bioinformatics/bts199.

Kershner, B., Tufts, C., Matthews, D., Nelson, G., 2008. National Wildlife Federation field guide to trees of North America. Sterling Publishing Company, New York, NY.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120.

Koop, J.A.H., DeMatteo, K.E., Parker, P.G., Whiteman, N.K., 2014. Birds are islands for parasites. Biol. Lett. 10. https://doi.org/10.1098/rsbl.2014.0255.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772-773. https://doi.org/10.1093/molbev/msw260.

Leaché, A.D., Chavez, A.S., Jones, L.N., Grummer, J.A., Gottscho, A.D., Linkem, C.W., 2015. Phylogenomics of phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biol. Evol. 7, 706-719. https://doi.org/10.1093/gbe/evv026.

Legendre, P., Desdevises, Y., Bazin, E., 2002. A statistical test for host–parasite coevolution. Syst. Biol. 51, 217-234. https://doi.org/10.1080/10635150252899734.

Lepais, O., Weir, J.T., 2014. SimRAD: An R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. Mol. Ecol. Resour. 14, 1314-1321. https://doi.org/10.1111/1755-0998.12273.

Leppänen, S.A., Altenhofer, E., Liston, A.D., Nyman, T., 2013. Ecological versus phylogenetic determinants of trophic associations in a plant–leafminer–parasitoid food web. Evolution 67, 1493-1502. https://doi.org/10.1111/evo.12028.

Leslie, A.B., Beaulieu, J.M., Rai, H.S., Crane, P.R., Donoghue, M.J., Mathews, S., 2012. Hemisphere-scale differences in conifer evolutionary dynamics. Proc. Natl. Acad. Sci. USA 109, 16217-16221. https://doi.org/10.1073/pnas.1213621109.

Lewis, P.O., 2001. A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. Syst. Biol. 50, 913-925. https://doi.org/10.1080/106351501753462876.

Martínez-Aquino, A., 2016. Phylogenetic framework for coevolutionary studies: A compass for exploring jungles of tangled trees. Curr. Zool. 62, 393-403. https://doi.org/10.1093/cz/zow018.

Matsubayashi, K.W., Ohshima, I., Nosil, P., 2010. Ecological speciation in phytophagous insects. Entomol. Exp. Appl. 134, 1-27. https://doi.org/10.1111/j.1570-7458.2009.00916.x.

Mayer, W.E., Herrmann, M., Sommer, R.J., 2009. Molecular phylogeny of beetle associated diplogastrid nematodes suggests host switching rather than nematode-beetle coevolution. BMC Evol. Biol. 9, 212. https://doi.org/10.1186/1471-2148-9-212.

Meier-Kolthoff, J.P., Auch, A.F., Huson, D.H., Göker, M., 2007. CopyCat : Cophylogenetic analysis tool. Bioinformatics 23, 898-900. https://doi.org/10.1093/bioinformatics/btm027.

Merkle, D., Middendorf, M., Wieseke, N., 2010. A parameter-adaptive dynamic programming approach for inferring cophylogenies. BMC Bioinformatics 11, S60. https://doi.org/10.1186/1471-2105-11-S1-S60.

Meseguer, A.S., Coeur d'acier, A., Genson, G., Jousselin, E., 2015a. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus. J. Biogeogr. 42, 1482-1492. https://doi.org/10.1111/jbi.12531.

Meseguer, A.S., Coeur d'acier, A., Genson, G., Jousselin, E., 2015b. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus. J. Biogeogr. 42, 1482-1492.

Miller, M.A., Pfeiffer, W., Schwarts, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE). https://doi.org/10.1109/GCE.2010.5676129.

Mitter, C., Farrell, B., Futuyma, D.J., 1991. Phylogenetic studies of insect-plant interactions: insights into the genesis of diversity. Trends Ecol. Evol. 6, 290-293.

Namin, H.H., Iranpour, M., Sharanowski, B.J., 2014. Phylogenetics and molecular identification of the *Ochlerotatus communis* complex (Diptera: Culicidae) using DNA barcoding and polymerase chain reaction-restriction fragment length polymorphism. Can. Entomol. 146, 26-35. https://doi.org/10.4039/tce.2013.60.

Nixon, K.C., Wheeler, Q.D., 1990. An amplification of the phylogenetic species concept. Cladistics 6, 211-223.

Nyman, T., Bokma, F., Kopelke, J.-P., 2007. Reciprocal diversification in a complex plant-herbivore-parasitoid food web. BMC Biology 5, 49. https://doi.org/10.1186/1741-7007-5-49.

Nyman, T., Farrell, B.D., Zinovjev, A.G., Vikberg, V., 2006. Larval habits, host-plant associations, and speciation in nematine sawflies (Hymenoptera: Tenthredinidae). Evolution 60, 1622 - 1637.

Page, R.D.M., 2003. Tangled trees: Phylogeny, cospeciation, and coevolution. University of Chicago Press.

Peixoto, L., Allen, G.R., Ridenbaugh, R.D., Quarrell, S.R., Withers, T.M., Sharanowski, B.J., 2018. When taxonomy and biological control researchers unite: species delimitation of *Eadya* parasitoids (Braconidae) and consequences for biological control of invasive paropsine pests of *Eucalyptus*. PLoS One 13, e0201276. https://doi.org/https://doi.org/10.1371/journal.pone.0201276.

Peralta, G., Frost, C.M., Didham, R.K., Varsani, A., Tylianakis, J.M., 2015. Phylogenetic diversity and co-evolutionary signals among trophic levels change across a habitat edge. J. Anim. Ecol. 84, 364-372.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. Plos One 7, e37135. https://doi.org/10.1371/journal.pone.0037135.

Pike, K.S., Starý, P., 1996. New species of *Pauesia* (Hymenoptera: Braconidae: Aphidiinae) parasitoids on *Cinara* (Homoptera: Aphididae: Lachninae) associated with conifers in the pacific northwest. Proc. Entomol. Soc. Wash. 98, 324-331.

Pike, K.S., Starý, P., Graf, G., Allison, D., 2002. *Pauesia columbiana*, n. sp. (Hymenoptera: Braconidae: Aphidiinae) on juniper aphids, and a key to related species. Proc. Entomol. Soc. Wash. 104, 646-654. https://doi.org/http://biostor.org/reference/55196.

Pike, K.S., Starý, P., Miller, R., Allison, D., Boydston, L., Graf, G., Miller, T., 1996. New species and host records of aphid parasitoids (Hymenoptera: Braconidae: Aphidiinae) from the Pacific Northwest, USA. Proc. Entomol. Soc. Wash. 98, 570-591. https://doi.org/http://biostor.org/reference/56763.

Price, P.W., Bouton, C.E., Gross, P., McPheron, B.A., Thompson, J.N., Weis, A.E., 1980. Interactions among three trophic levels: Influence of plants on interactions between insect herbivores and natural enemies. Annu. Rev. Ecol. Evol. Syst. 11, 41-65. https://doi.org/10.1146/annurev.es.11.110180.000353.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol. Ecol. 21, 1864-1877.

Quicke, D.L.J., 2015. Idiobionts, Koinobionts and Other Life History Traits. In The Braconid and Ichneumonid Parasitoid Wasps, D. L. Quicke (Ed.). In: Quicke, D.L.J. (Ed.), The Braconid and Ichneumonid Parasitoid Wasps. John Wiley & Sons, Chichester, UK.

R Core Team, 2017. R: A language and environment for statistical computing. In: Team, R.C. (Ed.), R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A., 2014. Figtree v1.4.3. Molecular Evolution, Phylogenetics and Epidemiology, University of Edinburgh, Institute of Evolutionaty Biology, Edinburgh.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901-904. https://doi.org/10.1093/sysbio/syy032.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes 7, 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Ricklefs, R.E., Fallon, S.M., 2002. Diversification and host switching in avian malaria parasites. Proceedings of the Royal Society of London B: Biological Sciences 269, 885-892.

Ridenbaugh, R.D., Barbeau, E., Sharanowski, B.J., 2018. Description of four new species of *Eadya* (Hymenoptera: Braconidae), parasitoids of the Eucalyptus Tortoise Beetle (*Paropsis charybdis*) and other Eucalyptus defoliating leaf beetles. Journal of Hymenopteran Research 64, 141-175.

Rodriguez, J.J., Fernández-Triana, J.L., Smith, M.A., Janzen, D.H., Hallwachs, W., Erwin, T.L., Whitfield, J.B., 2013. Extrapolations from field studies and known faunas converge on dramatically increased estimates of global microgastrine

parasitoid wasp species richness (Hymenoptera: Braconidae). Insect Conservation and Diversity 6, 530-536.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539-542. https://doi.org/10.1093/sysbio/sys029.

Saladin, B., Leslie, A.B., Wüest, R.O., Litsios, G., Conti, E., Salamin, N., Zimmermann, N.E., 2017. Fossils matter: Improved estimates of divergence times in *Pinus* reveal older diversification. BMC Evol. Biol. 17, 95. https://doi.org/10.1186/s12862-017-0941-z.

Sanchis, A., Michelena, J.M., Latorre, A., Quicke, D.L.J., Gärdenfors, U., Belshaw, R., 2001. The phylogenetic analysis of variable-length sequence data: Elongation factor–1α introns in European populations of the parasitoid wasp genus *Pauesia* (Hymenoptera: Braconidae: Aphidiinae). Mol. Biol. Evol. 18, 1117-1131. https://doi.org/10.1093/oxfordjournals.molbev.a003882.

Schluter, D., 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Smith, C.F., 1944. The Aphidiinae of North America (Braconidae: Hymenoptera). Ohio State University, Columbus.

Smith, C.I., Godsoe, W.K.W., Tank, S., Yoder, J.B., Pellmyr, O., 2008a. Distinguishing coevolution from covicariance in an obligate pollination mutualism: Asynchronous divergence in Joshua Tree and its pollinators. Evolution 62, 2676-2687. https://doi.org/10.1111/j.1558-5646.2008.00500.x.

Smith, M.A., Rodriguez, J.J., Whitfield, J.B., Deans, A.R., Janzen, D.H., Hallwachs, W., Hebert, P.D.N., 2008b. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proceedings of the National Academy of Sciences 105, 12359-12364. https://doi.org/10.1073/pnas.0805319105.

Sorenson, M.D., Balakrishnan, C.N., Payne, R.B., 2004. Clade-limited colonization in brood parasitic finches (*Vidua* spp.). Syst. Biol. 53, 140-153. https://doi.org/10.1080/10635150490265021.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313. https://doi.org/10.1093/bioinformatics/btu033.

Stamatakis, A., Auch, A.F., Meier-Kolthoff, J., Göker, M., 2007. AxPcoords & parallel AxParafit: Statistical co-phylogenetic analyses on thousands of taxa. BMC Bioinformatics 8, 405. https://doi.org/10.1186/1471-2105-8-405.

Stireman, J.O., Nason, J.D., Heard, S.B., Seehawer, J.M., 2006. Cascading host-associated genetic differentiation in parasitoids of phytophagous insects. Proceedings of the Royal Society of London B: Biological Sciences 273, 523 - 530.

Syring, J., Willyard, A., Cronn, R., Liston, A., 2005. Evolutionary relationships among *Pinus* (Pinaceae) subsections inferred from multiple low-copy nuclear loci. Am. J. Bot., 2086. https://doi.org/10.3732/ajb.92.12.2086.

Thompson, J.N., 1999. Specific hypotheses on the geographic mosaic of coevolution. Am Nat, 1 - 14.

Tumlinson, J.H., Lewis, W.J., Vet, L.E.M., 1993. How parasitic wasps find their hosts. Sci. Am. 3, 100-106. https://doi.org/10.1038/scientificamerican0393-100.

Tumlinson, J.H., Paré, P.W., Lewis, W.J., Plant production of volatile semiochemicals in response to insect-derived elicitors. Wiley Online Library, pp. 95-109.

Vanhove, M.P.M., Pariselle, A., Van Steenberge, M., Raeymaekers, J.A.M., Hablützel, P.I., Gillardin, C., Hellemans, B., Breman, F.C., Koblmüller, S., Sturmbauer, C., Snoeks, J., Volckaert, F.A.M., Huyse, T., 2015. Hidden biodiversity in an ancient lake: phylogenetic congruence between Lake Tanganyika tropheine cichlids and their monogenean flatworm parasites. Sci Rep-UK 5, 13669. https://doi.org/10.1038/srep13669.

Wang, X.Q., Tank, D.C., Sang, T., 2000. Phylogeny and divergence times in Pinaceae: Evidence from three genomes. Mol. Biol. Evol. 17, 773-781. https://doi.org/10.1093/oxfordjournals.molbev.a026356.

Wickham, H., 2016. ggplot2: Elegant graphics for data analysis. Springer, New York, NY.

Wilson, J., Forister, M., Dyer, L., O'connor, J., Burls, K., Feldman, C., Jaramillo, M., Miller, J., Rodríguez-Castañeda, G., Tepe, E., 2012. Host conservatism, host shifts and diversification across three trophic levels in two Neotropical forests. J. Evol. Biol. 25, 532-546.

Winkler, I.S., Mitter, C., 2008. The phylogenetic dimension of insect-plant interactions: a review of recent evidence. In: Tilman, K. (Ed.), Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects. University California Press, Berkley, pp. 240-263.

Yu, D.S.K., Van Achterberg, C., Horstmann, K., 2012. Taxapad 2012, Ichneumonoidea 2011. URL: www. taxapad. com.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2016. ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28-36. https://doi.org/10.1111/2041-210X.12628.

Zhang, M.Y., Ridenbaugh, R.D., Sharanowski, B.J., 2017. Integrative taxonomy improves understanding of native beneficial fauna: revision of the Nearctic *Peristenus pallipes* complex (Hymenoptera: Braconidae) and implications for release of exotic biocontrol agents. Syst. Entomol. 42, 596-608.

Zhang, Y.M., Bass, A.I.H., Fernández, D.C., Sharanowski, B.J., 2018. Habitat or temporal isolation: Unraveling herbivore–parasitoid speciation patterns using double digest RADseq. Ecol. Evol. 8, 9803-9816. https://doi.org/10.1002/ece3.4457.