THE INFLUENCE OF VICARIANCE AND DISPERSAL ON THE DIVERSIFICATION AND EVOLUTION OF SPRINGTAILS (COLLEMBOLA)

BY

ARON D. KATZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Entomology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Doctor Mark A. Davis, Chair and Co-Director of Research
Doctor Steven J. Taylor, Co-Director of Research
Professor Andrew V. Suarez
Professor James B. Whitfield

ABSTRACT

Vicariance and dispersal are major drivers of the evolution of biodiversity, yet the relative impact of these processes, and the factors that influence them, often remain elusive. Identifying and understanding evolutionary processes responsible for generating diversity is essential for understanding ecological patterns and for the development and implementation of management strategies intended to conserve biodiversity. My dissertation research is focused on the historical, ecological and evolutionary underpinnings driving the origin, diversification, and maintenance of biodiversity in springtails (Collembola). This class of tiny, wingless insect-like hexapods includes some of the most abundant (and perhaps, most diverse) arthropods on earth. They have colonized, diversified, and adapted to nearly every terrestrial habitat from marine intertidal zones and tropical rainforests to polar deserts and caves, yet the evolutionary mechanisms behind their ecological success are poorly understood.

Ecological specialization is a central theme in my first two chapters. Species with obligate ecological associations offer simple systems to evaluate biogeographical hypotheses and also provide an ecological context to test the effects and consequences of specialization on patterns of diversity. In Chapter 1, I identify and compare spatial and temporal patterns of molecular diversity for two ecologically distinct and codistributed genera of cavernicolous springtails (cave-obligate vs. cave-facultative species) from a regional cave-bearing karst system spanning the Mississippi River in Illinois and Missouri. Phylogeographic analysis revealed that evolutionary processes of vicariance and dispersal were both major influences on patterns of cave springtail diversity, but the effects of these processes are also strongly influenced by intrinsic ecological factors, in this case, the degree of cave-dependence. Estimates of genetic structure and divergence times also implicated climatic and geological processes involved in the

formation of the modern Mississippi River valley as major factors driving the isolation of cave-obligate species, but cave-facultative species have been able to maintain genetic connectivity across this barrier.

In Chapter 2, I developed a molecular dataset for marine littoral-obligate springtails collected along the Pacific and Atlantic coasts of Panama to identify transisthmian sister taxa, determine processes driving and maintaining their isolation, and to evaluate their timing of origin. I was able to identify multiple geminate species pairs spanning the Isthmus and molecular analyses revealed examples of pre-Pliocene vicariance, post-closure dispersal, and cryptic diversification across the Isthmus of Panama. This study not only demonstrates that ecological specialization can (but not always) reduce genetic connectivity across geographic barriers, but also corroborates recent (and controversial) geological and biogeographical estimates of an early Miocene closure of the Panama Isthmus.

These works demonstrate the utility of incorporating ecologically specialized springtails in evolutionary investigations. However, independent timing information is essential for assessing historical factors influencing contemporary patterns of diversity. Unfortunately, springtails (and most other small, soft-bodied organisms) lack a useful fossil record for this purpose. As a result, employing "universal" rates of molecular evolution to estimate divergence times is common, even though evolutionary rates can vary considerable across taxa. In Chapter 3, I assess the validity of the generalized arthropod rate assumption by conducting a Bayesian phylogenetic analysis to evaluate the relative rate of molecular evolution across all major hexapod groups. I found that substitution rates in Collembola are not significantly different from most other hexapod groups and suggest that the use of "universal" insect molecular clocks are appropriate for estimating collembolan evolutionary timescales.

An additional yet fundamental challenge impacting all fields of biology is the fact that most biodiversity remains to be discovered or is poorly understood—including many species that are of potential conservation concern. This is exacerbated in collembolan taxonomy, due to the lack of variation in discrete morphological characters, and the general shortage of taxonomic expertise in North America, reflected in the limited number of taxonomic tools available to researchers (and the public) for identifying species in this region. To help address this challenge, I produced detailed morphological taxonomic descriptions for all New World species of the springtail genus *Willowsia*, including a new species that is endemic to Florida in Chapter 4. Most members in this genus are from Asia, but comparative morphological analysis revealed two unique character states shared only by endemic New World *Willowsia* and *Americabrya*, providing prima facie evidence of their independent evolution from a common New World ancestor.

To my amazing wife, Monique

who has put up with a lot for me to be here.

Without her loving support, this would not have been possible.

ACKNOWLEDGEMENTS

Although I may have set a departmental record for the most advisors, I am extremely fortunate to have been given the chance to work with all of them, as they each have provided me with support and mentorship, but also different perspectives and experiences. I want to thank my first advisor, Felipe Soto-Adames, who introduced me to the wonderful world of Collembola, and, without whom I would not have become a member of the highly exclusive springtail taxonomist society of North America. I am grateful to Steve Taylor, advisor #2, for teaching me how to be a cave biologist, develop collaborative research projects, and become a more productive scientist. I also want to thank my current advisor, Mark Davis, who has provided generous scientific resources, and always showed interest in my progress, offered encouragement, and devoted the time to quickly return all my last-minute manuscript edits. I am thankful to all the current and former members of the Davis, Taylor, and Heads Labs at the Annex, and particularly, Tara Hohoff, Matt Safford, Dan Swanson, Nathalie Baena, Matt Niemiller, Sam Heads, Jared Thomas, Katie Dana, Tyler Hedlund, and Mary Best for their friendship, camaraderie, support, and generally making the Annex a great place to work. I also thank my dissertation committee members, Andy Suarez and Jim Whitfield, for their valuable feedback and advice regarding the direction and focus of this project. Special thanks to the faculty, staff, and graduate students of the Department of Entomology at the University of Illinois at Urbana-Champaign for their incredible support over the years.

TABLE OF CONTENTS

CHAPTER 1

At the confluence of vicariance and dispersal: phylogeography of cavernicolous springtails (Collembola: Arrhopalitidae, Tomoceridae) codistributed across a geologically complex karst landscape in Illinois and Missouri[1]

Abstract

The processes of vicariance and dispersal are central to our understanding of diversification, yet determining the factors that influence these processes remains a significant challenge in evolutionary biology. Intrinsic ecological differences among cavernicolous organisms, such as the degree of cave-dependence, are thought to be major factors influencing patterns of genetic isolation in caves. Following a comparative phylogeographic approach, I employed mitochondrial and nuclear markers to assess the impacts of ecological and geological factors on the evolutionary history of two ecologically distinct groups of terrestrial cave-dwelling springtails (Collembola) in the genera *Pygmarrhopalites* (Arrhopalitidae) and *Pogonognathellus* (Tomoceridae) that are codistributed in caves throughout the Salem Plateau—a once continuous karst region, now bisected by the Mississippi River valley in Illinois and Missouri. Contrasting phylogeographic patterns recovered for troglobiotic *Pygmarrhopalites* sp. and eutroglophilic *Pogonognathellus* sp. suggests that obligate associations with cave habitats can restrict dispersal across major geographic barriers such as rivers and valleys but may also facilitate subterranean dispersal between neighboring cave systems. *Pygmarrhopalites* sp. populations spanning the Mississippi River Valley were estimated to have diverged 2.90–4.76 Ma, which I attribute to isolation due to vicariance initiated, and subsequently maintained, by climatic and geological processes involved in Mississippi River valley formation beginning during the late Pliocene/early Pleistocene. Lastly, I conclude that the detection of many deeply divergent,

[1] *Ecology and Evolution* (in press, 2018)

morphologically cryptic, and microendemic lineages highlights our poor understanding of microarthropod diversity in caves and exposes potential conservation concerns.

Introduction

Subterranean ecosystems are attractive systems for biologists seeking to understand the evolutionary processes that shape patterns of biological diversity (Culver & Pipan, 2009). These isolated, dark, low-energy habitats promote adaptation and diversification (Ortiz et al., 2014), and provide the ecological context for examining mechanisms underlying divergence and speciation (e.g., Niemiller et al., 2008; Juan et al., 2010; Schönhofer et al., 2015; Gómez et al., 2016). In contrast to their surface relatives, many cave-dwelling species have morphological, physiological, and behavioral adaptations that limit or prevent surface dispersal (White & Culver, 2012). Often restricted to small, discontinuous ranges, these species can exhibit high levels of population structure, short-range endemism, and morphologically cryptic species (e.g., Christman et al., 2005; Zagmajster et al., 2008; Niemiller et al., 2012; Faille et al., 2015).

Genetic isolation is a primary driver of molecular divergence and, ultimately, speciation, but determining the factors that promote or constrain genetic diversity remains a significant challenge in evolutionary biology. Patterns of diversity in caves are often attributed to vicariance or dispersal, but the relative influence these processes have on the evolution and contemporary distributions of cave fauna has been widely debated (see Porter, 2007; Culver et al., 2009). However, it is generally accepted that patterns of diversity in caves are likely shaped by a complex interaction of intrinsic factors (e.g., species-specific differences in ecology, life history, or biology) that can influence dispersal capacity and extrinsic factors (e.g., geographic barriers or climate change) that can enhance or limit dispersal opportunity (Porter, 2007; Juan et al., 2010).

Phylogeography, the study of processes that influence the contemporary geographic distributions of species' populations by utilizing genetic data, can provide insights into the relative influences of evolutionary factors driving patterns of genetic isolation and divergence in biological communities (Avise et al., 1987; Avise, 2000). For instance, phylogeographic congruence among codistributed species can implicate vicariance caused by 'hard' geographic barriers or environmental changes affecting entire communities (Lapointe & Rissler, 2005), whereas conflicting phylogeographic patterns may be attributable to intrinsic differences that can affect species dispersal capacity across 'soft' potential genetic barriers (e.g., Hodges et al., 2007; Goldberg & Trewick, 2011; Hurtado et al., 2013). With cave organisms, the majority of research studies have been limited to single species (e.g., Dörge et al., 2014; Faille et al., 2015) or cryptic species complexes with allopatric distributions (e.g., Rastorgueff et al., 2014; Gómez et al., 2016). Few studies have incorporated phylogeographic approaches that consider intrinsic differences among codistributed cave-dwelling species (see Weckstein et al., 2016; Pérez-Moreno et al., 2017).

The arthropod class Collembola (springtails) offers a nearly unparalleled opportunity for elucidating the interplay of factors that affect speciation and molecular diversification in subterranean ecosystems. These small, wingless, insect-like arthropods are among the most abundant, diverse, and well-adapted organisms in caves (Christiansen, 1965; Thibaud & Deharveng, 1994), and are considered important subterranean examples of adaptive radiations (Christiansen & Culver, 1969) and parallel speciation (Christiansen, 1961, 1965; Christiansen & Culver, 1968). Their small size (body length often less than 1 mm), low vagility, and close associations with cave habitats facilitates their isolation, resulting in a high degree of endemism (Niemiller & Zigler, 2013) and cryptic species (Juan & Emerson, 2010). For example, the genus

Pseudosinella alone contains more than 100 species found in caves worldwide, many of which are known only from a single cave system (Hopkin, 1997). Most importantly, cave-dwelling springtails have varying levels of ecological specificity to, and dependence upon, cave habitats. Although surface species are commonly found in caves as accidentals (i.e., they may fall or get washed into caves, but cannot maintain populations in caves), the majority of collembolans occurring in caves can maintain permanent subterranean populations and are either classified as troglobionts (i.e., obligate cave-dwellers that are never encountered on the surface and often have conspicuous troglomorphic adaptations associated with cave habitats) or eutroglophiles (i.e., facultative cave-dwellers that also occur in surface habitat and usually lack apparent troglomorphy). Because troglobiotic and eutroglophilic springtails can be codistributed (Soto-Adames & Taylor, 2013; Katz et al., 2016), extrinsic evolutionary processes are likely exerting similar selective pressures upon them. Therefore, opposing patterns of genetic structure among these species distributed across the same geographic area can reflect intrinsic factors, such as differences in the degree of ecological association with cave habitats (cave-dependence), that can affect a species' capacity to disperse across geographic barriers (Weckstein et al., 2016; Pérez-Moreno et al., 2017). Disparate geographic distributions among closely related surface springtails provides some indirect evidence that varying dispersal capacity may be associated with differences in species-specific traits (Costa et al., 2013; Katz et al., 2015) and Christiansen and Culver's (1987) biogeographic study of cave springtails revealed that more pronounced troglomorphy can be correlated with smaller geographic ranges.

Long-term local persistence and small geographic ranges are typical for troglobionts, and by definition, these species cannot maintain surface populations to facilitate dispersal between discontinuous subterranean habitats. Therefore, patterns of genetic differentiation in troglobionts

are likely driven primarily by isolation due to physical barriers and reflect vicariance. On the contrary, I expect isolation-by-distance (IBD) to be the primary driver of genetic variation in eutroglophiles owing to their propensity to disperse across surface habitats.

To test these predictions, I incorporate a suite of molecular-based approaches to 1) delimit cryptic species in the focal complexes, 2) detect molecular signatures of isolation to identify potential genetic barriers, and 3) estimate evolutionary relationships and divergence times to elucidate the roles of vicariance and dispersal in shaping patterns of cave-dwelling springtail diversity throughout the Salem Plateau—a major cave-bearing karst region spanning the Mississippi River valley in Illinois and Missouri. Phylogeographic patterns that emerge from this study are used to investigate the impact of intrinsic and extrinsic factors (e.g., the degree of cave-dependence and geographic barriers) on the evolution of cave organisms, broaden our limited understanding of subterranean microarthropod diversity, and assess biogeographic interpretations that may help clarify the complex, yet poorly understood, geological history of the Salem Plateau karst region that spans the Mississippi River valley.

Materials and Methods

Study system, focal taxa, and field collections

The complex geological landscape of the Salem Plateau (Fig. 1.1) provides the ecological context for testing biogeographic hypotheses of vicariance and dispersal. This once continuous karst region, now bisected by the Mississippi River Valley, is located south of St. Louis and covers just eight counties but contains thousands of sinkholes and includes the largest cave systems in Illinois and Missouri (Panno et al., 1997). Here I examine patterns of molecular diversity of two codistributed and ecologically distinct genera of springtails—*Pygmarrhopalites*

Vargovitch, 2009 (Arrhopalitidae) and *Pogonognathellus* Paclt, 1944 (Tomoceridae) (Fig. 1.2)—both highly abundant groups that likely comprise the majority of cave-dwelling hexapods found in this region. Most *Pygmarrhopalites* species in caves are classified as troglobionts because they are usually troglomorphic and have never been reported in surface habitats, instead, occurring primarily on drip pool surfaces or organic debris in cave dark zones. In contrast, *Pogonognathellus* species are not troglobionts and include only a few eutroglophilic species that can maintain permanent populations in caves. Eutroglophilic *Pogonognathellus* occur in both surface and cave habitats, often in abundance on organic debris and rock surfaces in cave entrances and twilight zones, and less frequently and in smaller numbers in cave dark zones.

To date, twelve species of *Pygmarrhopalites* and four species of *Pogonognathellus* have been reported in Salem Plateau caves (Peck & Lewis, 1978; Lewis et al., 2003; Zeppelini & Christiansen, 2003; Soto-Adames & Taylor, 2010, 2013), including nine troglobiotic species of *Pygmarrhopalites* and a single, but widespread, eutroglophilic species complex—the *Pogonognathellus* pale complex Felderhoff et al., 2010, formerly recognized as Nearctic populations of *Pogonognathellus flavescens* (Tullberg), 1871.

Invertebrate surveys were conducted in 25 caves located throughout the Salem Plateau karst in Illinois and Missouri during the summer of 2016 (Table 1.1; Fig. 1.1). At each cave site, the dominant habitat types where springtails are known to occur were opportunistically sampled from cave entrance, twilight, and dark zones. Habitats such as rock surfaces, scat, and drip pools were sampled manually using an aspirator and organic debris was collected in bags for Berlese extraction. Ecological data, such as sample orientation (wall, floor), substrate, habitat, and cave zone, were recorded for each sample. All Collembola recovered from raw samples were subsequently sorted and identified to morphospecies using a Leica MZ12.5 research

stereomicroscope. The most abundant morphospecies (i.e., morphologically indistinguishable under stereomicroscopy) of *Pygmarrhopalites* and *Pogonognathellus* were retained for DNA extraction. All sorted samples, including non-target specimens, were stored in 95% EtOH at 4°C.

Because caves contain sensitive resources, including federally endangered species, specific locations are not included in supporting material—these data must be requested from the Illinois Speleological Survey or the Missouri Speleological Survey.

DNA extraction, PCR amplification, DNA sequencing, and alignment

Chaetotaxy (i.e., the arrangement and morphology of setae) and other small cuticular morphological characters are critical for springtail species identification, and it is often impossible to identify springtails to the species-level without first making slides—a process that destroys DNA. Therefore, I used non-destructive DNA extraction methods to maximize DNA concentrations while preserving morphology to associate genetic sequences to voucher specimens for species identifications. I extracted DNA from specimens representing the most abundant morphospecies for *Pygmarrhopalites* (n=43) and *Pogonognathellus* (n=26) using the following modifications to the DNeasy Blood & Tissue kit protocol (Qiagen Inc.,Valencia, CA): 1) whole bodied specimens (with head removed) were incubated overnight at 56 °C after the addition of ATL buffer with proteinase-K; 2) after the addition of EtOH, the samples were stored at 4°C overnight in order to maximize DNA precipitation; and 3) prior to centrifugation, buffer containing DNA was carefully removed and added to a column using a pipette, taking care to not lose or damage the specimens which were left at the bottom of the tube and preserved in 95% EtOH. Digestion of tissues and pigments by the lysis buffer resulted in very delicate and clear specimens, ready for slide mounting without additional preparation, but the fragile cuticles were

easily damaged when handled and small individuals were nearly invisible making them difficult to recover. Therefore, the heads of specimens, which include important diagnostic morphology, were dissected and stored separately prior to DNA extraction as back up vouchers for those cases where the now-translucent bodies were not recovered.

This study incorporates two mitochondrial (COI and 16S) and three nuclear loci (28S D1-3, 28S D7-10 and Histone-3). COI and 16S are particularly useful for evaluating population level variation as they exhibit high levels of genetic variation and have been used extensively for species and population-level phylogenetic research in springtails (Hogg & Herbert, 2004). Collembola are generally characterized by extremely high levels of molecular diversity (Katz et al., 2015), therefore, more slowly evolving loci, 28S and Histone-3, were included to provide stronger phylogenetic signal among more distantly related taxa. Histone-3 and 28S D1-3 were excluded for *Pogonognathellus* due to inconsistent amplification. See Table 1.2 for list of all taxa included in this study, including sample information and all sequences with corresponding GenBank (Benson et al., 2013) accession numbers.

Gene fragments were amplified by PCR with 12.5 μL of Go Taq Master Mix (Promega Corporation, Madison, WI), 8.5 μL of water, 1 μL of 10 μM forward and reverse primers (Table 1.3), and 2 μL of genomic DNA. Thermocycler settings for all primer combinations were as follows: 95°C 5 min; 40 cycles of 95°C 45 sec, 50°C 1.5 min, 72°C 1.5 min; and a final extension step for 72°C 10 min. Annealing temperature (50°C) was adjusted as needed for problematic amplifications. Successful amplification was verified via gel electrophoresis (90 V, 400 mA for 45 min) using a 1% agarose gel stained with GelRed (Biotium Inc., Hayward, CA). If multiple bands were observed, PCR was repeated with increasing annealing temperatures until single bands were obtained. Single-band PCR products were cleaned with Exo Sap-It Express

(Affymetrix Inc., Santa Clara, CA) following manufactures protocol. Cleaned PCR products were pre-mixed with primers (1 μM) and sent to Eurofins Genomics LLC (Louisville, KY) for sequencing.

Forward and reverse sequences were assembled with Sequencher v. 5.4 (Gene Codes Corporation, Ann Arbor, MI). COI and H3 were aligned using the G-INSI-i alignment method in MAFFT v. 7.273 (Katoh & Standley, 2013) and translated to amino acids to check for stop codons. The GUIDANCE2 online server (Sela et al., 2015) was used to align 16S and 28S using MAFFT (max iterate=1000, 6mer pairwise alignment method, 100 bootstraps) to identify unreliably aligned positions for removal below default guidance score (cutoff=0.93). The outgroup taxa listed in Table 1.4 were chosen based on their affinities with the target taxa and availability of sequences in GenBank.

Detecting cryptic diversity and OTU delimitation

The presence of cryptic diversity was detected by incorporating a number of different tests. First, I calculated pairwise COI distance frequencies with PAUP* 4.0a build 159 (Swofford, 2002) and plotted distance frequency histograms to detect the presence of interspecific variation within each targeted morphospecies. A gap between the greatest putative intraspecific and smallest putative interspecific pairwise distances can be interpreted as the boundary between species and population-level variation (Meier et al., 2008).

To determine how interspecific COI variation was geographically distributed, I performed a hierarchical analysis of molecular variance (AMOVA) for all taxa sampled for each target morphospecies using Arlequin v. 3.5.2.2 (Excoffier & Lischer, 2010). Haplotypes were grouped within samples, among samples in caves, and among caves with 50,000 permutations

performed to assess significance. The presence of strong genetic structuring within samples or among samples in caves can be an indicator of cryptic diversity because sexual isolation is typically required to maintain high levels of genetic variation occurring in sympatry.

I also delimited putative species boundaries using a General Mixed Yule Coalescent (GMYC) analysis (Pons et al., 2006). This method uses ultrametric gene trees to identify the interface between population and species-level branching patterns and demarcates genetically cohesive clades as independent evolutionary units known as operational taxonomic units (OTUs). Bayesian inference of COI gene trees for both genera was performed with BEAST2 v. 2.4.8 (Bouckaert et al., 2014) with the following parameters: site model averaging implemented in bModelTest (Bouckaert & Drummond, 2017) was used to accommodate uncertainty in the model of sequence evolution (default parameters); a strict clock rate set to 1 for relative branch length estimation; Yule tree model; monophyletic constraint prior on the ingroup taxa; Markov chain Monte Carlo (MCMC) for 100 million generations; sampling statistics and trees every 1000 generations (10% burn-in). Effective sample size (ESS) for all parameters were determined to be greater than 200 with Tracer v1.6.0 (Rambaut & Drummond 2007). Maximum clade credibility trees were inferred with TreeAnnotator v2.4.8 (Bouckaert et al., 2014). bModelTest site model distributions and statistics are reported in Tables 1.5–1.6.

The GMYC analysis was performed in R (R Core Team, 2017) using the splits package (Ezard et al., 2009) with the single threshold delimitation method. Inter- and intra-OTU uncorrected genetic distances for COI and 16S were computed in PAUP* and plotted in R.

Representative specimen vouchers recovered during DNA extraction for each OTU were directly slide-mounted with Hoyer's Medium (Mari Mutt, 1979) for morphological examination using a Nikon Eclipse Ni-U upright microscope with phase contrast to check for morphological

differentiation among OTUs, the presence of troglomorphy, and to provide preliminary species identifications for focal OTUs.

Tests for genetic structure

The relative role of cave-dependence and its influence on springtail dispersal capacity remains unclear, in part, because the identities of genetic barriers are not known for cave-dwelling springtails. To identify barriers to *Pygmarrhopalites* and *Pogonognathellus* dispersal, I evaluated and compared levels of genetic structure across cave boundaries and the Mississippi River valley. In addition, I also included sinkhole area boundaries in the genetic structure analyses. Because cave density in karst regions can be correlated with sinkhole density (Shofner et al., 2001), areas without sinkholes may lack sufficient cave habitat for subterranean species dispersal. Therefore, I assigned discontinuous sinkhole karst areas in Illinois (Panno et al., 1997; Panno et al., 1999; Venarsky et al., 2009) and Missouri (Panno et al., 1999; Burr et al., 2001) (neighboring karst subregions were combined) to each cave for genetic structure analyses (Table 1.1; Fig. 1.1).

The most sampled OTUs for each target morphospecies, identified by the GMYC analysis, were chosen as focal OTUs for population analyses to avoid attributing deeply divergent and structured lineages to population-level variation, rather than to species-level variation (Fouquet et al., 2007). Hierarchical AMOVAs were performed independently with Arlequin for COI and 16S for both focal OTUs by grouping haplotypes within samples, among samples within barriers, and among samples across barriers. Significance was assessed with 50,000 permutations.

Patterns of population structure resulting from dispersal and genetic drift, rather than of vicariance across geographic barriers, are common in animals with low mobility and can usually be attributed to a model of isolation-by-distance (IBD) (Timmermans et al., 2005; Cotsa et al., 2013). To determine if geographic distance is significantly correlated with genetic distance, I performed a Mantel test (Mantel, 1967) for each locus. I also evaluated the significance of genetic structure across barriers while controlling for IBD using a partial Mantel test (Mantel, 1967), which allows for the comparison of two variables (i.e., pairwise genetic distances and position relative to geographic barrier) while controlling a third (i.e., geographic distances). The partial Mantel tests required matrices of pairwise uncorrected genetic distances for both focal OTUs, geographic distances (great-circle distance) between each cave location, and matrices with variables coded to indicate whether each pair of specimens occurred together or on different sides of each geographic barrier. All simple and partial Mantel tests were calculated with zt v1.1 (Bonnet and Van de Peer, 2002) with 100,000 permutations.

Templeton-Crandall-Sing (TCS) haplotype networks (Clement et al., 2002) for COI and 16S were estimated with PopART (Leigh & Bryant, 2015) to visualize and compare phylogeographic structure across genetic barriers for *Pygmarrhopalites* and *Pogonognathellus* focal OTUs.

Phylogenetic inference, divergence time estimation, and topology tests

To further investigate the interplay of vicariance and dispersal capacity on cave springtail diversity, I conducted a Bayesian phylogenetic analysis using BEAST 2 to infer evolutionary relationships and to estimate divergence times for all sampled lineages of *Pygmarrhopalites* and *Pogonognathellus*. Two independent datasets were analyzed and compared: the

Pygmarrhopalites dataset (COI, 16S, 28S D1-3, 28S D7-10, Histone-3; 3358 total bp) and the *Pogonognathellus* dataset (COI, 16S, 28S D7-10; 2059 total bp). External rates were used for molecular clock calibrations rather than fossil information because springtails lack an adequate fossil record and phylogenetic framework for calibrating molecular clocks. Brower's (1994) estimate of 2.3% divergence per million years for COI is the most widely used external rate calibration for inferring arthropod divergence times. However, this rate has been recently criticized for poor statistical rigor (Papadopoulou et al., 2010) and in light of recent evidence of potential accelerated mitochondrial evolution in springtails compared to other arthropods (Cicconardi et al., 2010), I chose the faster, more statistically robust 'insect' rates estimated by Papadopoulou et al. (2010) for molecular clock calibration (COI=3.54%/Ma; 16S=1.08%/Ma; 28S=0.12%/Ma). Clock.rate parameters in BEAST 2 were set to 0.0168, 0.0054, and 0.0006 respectively (substitutions/site/Ma). For the *Pygmarrhopalites* dataset, 28S regions D1-3 and D7-10 were combined into a single 28S clock model partition, but site models were estimated independently with bModelTest for each region. Rates for Histone-3 (*Pygmarrhopalites* only) were estimated. To determine whether the use of a strict clock is appropriate for each gene partition, I conducted a preliminary analysis using a relaxed log normal clock model for all partitions to accommodate for potential rate variation and to estimate the coefficient of variation (COV) for each clock partition. Coefficient of variation values approaching zero indicate clock-like evolution among lineages, whereas higher values indicate substantial rate variation among lineages (Drummond & Bouckaert, 2015). The resulting COV values (95% HPD) for *Pygmarrhopalites* (COI=0.34–0.79; 16S=0–0.45; 28S=1.30–2.47; Histone-3=1.39–3.61) and *Pogonognathellus* (COI=0.43–1.99; 16S=0.2–1.86; 28S=0.53–1.856) show that *Pygmarrhopalites* COI and 16S are relatively clock-like (COV < 1), while all other loci have

relatively high levels of rate variation among lineages. Therefore, I used strict clock models for the *Pygmarrhopalites* COI and 16S and relaxed log normal clock models for all other gene partitions. The analyses were run for 200 million generations, sampling statistics and trees every 5000 generations. All additional parameters were the same as those used to estimate COI gene trees for the GMYC analysis (see above). Convergence (ESS>200) and burn-in (10%) were assessed with Tracer and maximum clade credibility trees were inferred with TreeAnnotator. bModelTest site model distributions and statistics are reported in Tables 1.5–1.6.

I also compared different topological models using Bayes factors (BF) to further test for reciprocal monophyly across the Mississippi River for *Pygmarrhopalites* and *Pogonognathellus* focal OTUs. Informed topology was strictly constrained in the prior for all hypotheses because irrelevant background signal in an unconstrained analysis can bias BF tests for monophyly (Bergsten et al., 2013). For the *Pygmarrhopalites* dataset, I compared three different hypotheses regarding the uncertain placement of focal OTU Illinois lineages from Indian Cave (INC) (see results): H_0, all focal OTU Illinois lineages (including INC) were constrained to be monophyletic and sister to all focal OTU Missouri lineages [(IL+INC)+MO]; H_1, all focal OTU Missouri lineages and focal OTU lineages from INC were constrained as monophyletic and sister to all other focal OTU Illinois lineages [(MO+INC)+IL]; and H_2, focal OTU INC lineages were constrained to be sister to all other focal OTU lineages [INC+(IL+MO)]. Two models were compared for the *Pogonognathellus* dataset: H_0, focal OTU Illinois and Missouri lineages were each constrained as monophyletic and sister to each other [IL+MO]; H_1, relationships that group focal OTU Missouri and Illinois lineages resulting from the unconstrained analysis (best tree) were constrained [TMC+(ILC+HSC) & (AC1+AC2)+(PAC+FPC+MJP)] (See Table 1.1 for cave abbreviations). Marginal log likelihoods were estimated with stepping-stone MCMC sampling

using the MODEL_SELECTION v. 1.3.4 package in BEAST2 (alpha=0.3; steps=100; chain length=2 million; all other parameters as default). Number of steps and chain length were increased until there was no significant difference in marginal likelihood estimates. Excluding topological constraint priors, all other parameters were identical for each model; same as those used for divergence time estimations (see above). Following guidelines proposed by Kass and Raftery (1995), a twice logarithm BF difference ($2\times\log_e BF$) of higher than 6 was considered strong evidence against the null hypothesis.

Results

Evidence for cryptic diversity

Uncorrected pairwise COI distance frequency histograms revealed extraordinarily high genetic distances for both taxa: up to 35% for *Pygmarrhopalites* and 18% for *Pogonognathellus*. COI distances above 10–15% in springtails are typically recognized as interspecific when used in combination with independent evidence (Katz et al., 2015). There are also two distinct clusters of COI distances, each separated by a 10% gap (Fig. 1.3). This can be interpreted as a boundary between intra- and interspecific genetic variation, providing preliminary support for the presence of cryptic diversity within both target morphospecies.

The results of the initial AMOVA that incorporated all sampled taxa identified high levels of genetic structure within caves and within samples, supporting the presence of sympatric species (Table 1.7a): between 39.39% and 56.08% of genetic variation in COI and 16S was structured among samples within the same cave for both genera. 23.71% and 21.44% of genetic variation in COI and 16S respectively was also structured within samples for *Pygmarrhopalites*, but this pattern was not recovered for *Pogonognathellus* (COI, 2.43%; 16S, 0%).

Bayesian phylogenetic inference of COI for both genera resulted in gene trees with a high number of deeply divergent and well-supported clades (Figs 1.4–1.5). The GMYC analyses based on the COI gene trees revealed 14 putative species (Figs 1.4–1.5): 10 *Pygmarrhopalites* OTUs (A1-10) and 4 *Pogonognathellus* OTUs (T1-4). Intra-OTU % distances for COI ranged from 17.92–34.84% (mean=26.59%) for *Pygmarrhopalites* and 14.61–17.61% (mean=16.24%) for *Pogonognathellus*, while intra-OTU % distances for COI ranged from 0–8.89% (mean=3.97%) for *Pygmarrhopalites* and 0–2.97% (mean=1.57%) for *Pogonognathellus*. Although *Pygmarrhopalites* had more variable and notably higher genetic distances for COI and 16S compared to *Pogonognathellus*, there is no overlap between intra- and inter-OTU distances for both genera (Fig. 1.6). *Pygmarrhopalites* A10 and *Pogonognathellus* T4 were chosen for comparative phylogeographic analysis because they included the largest number of sampled lineages. It is worth noting that because *Pygmarrhopalites* A10 could not be differentiated from other OTUs under stereomicroscopy, targeted sequencing to increase their sample sizes was not possible.

Morphological examination of slide-mounted DNA voucher specimens of *Pygmarrhopalites* revealed similar, but distinct and unique morphologies for all OTUs (e.g., differentiation of the female subanal appendage and claw morphology) that do not fit any known species description. *Pygmarrhopalites* A10 (Fig. 1.2a) had moderate troglomorphy (e.g., elongated antennae and thread-like unguiculus) and was determined to be an undescribed new species, most similar to *Pygmarrhopalites pavo* Christiansen & Bellinger, 1996, a troglobiont reported from caves in Virginia (Christiansen & Bellinger, 1996), West Virginia (Fong et al., 2007), Tennessee (Lewis, 2005), and Missouri (Zeppelini et al., 2009). I believe the unique differences in morphology, in combination with molecular evidence, supports the recognition of

all *Pygmarrhopalites* OTUs as distinct and potentially new species. Because some cryptic lineages may be of higher conservation concern, it is imperative to identify and describe these lineages for potential management initiatives (Niemiller et al. 2013a; Delić et al., 2017). However, I chose to refrain from giving OTUs formal species names at this time because a comprehensive taxonomic review is required to describe new species and to clarify the status of existing species—a task beyond the scope of this study.

All sampled lineages of *Pogonognathellus* were identified as members of the *Pogonognathellus* pale species complex Felderhoff et al., 2010, (Fig. 1.2b), a common eutroglophile inferred to be comprised of multiple cryptic species that cannot be differentiated without molecular data (Felderhoff et al., 2010), a finding that is also supported here by the recovery of four deeply divergent molecular lineages (T1-T4) with indistinguishable morphology under compound light microscopy.

Genetic structure

Results of the hierarchical AMOVAs identified that the majority of genetic variation in COI and 16S was structured among caves for both *Pygmarrhopalites* A10 (COI, 88%; 16S, 91%) and *Pogonognathellus* T4 (COI, 92%; 16S, 98%) (Table 1.7b). Genetic variation in COI and 16S was also strongly structured among sinkhole areas (COI, 94%; 16S, 96%) and regions east and west of the Mississippi River (COI, 58%; 16S, 73%) for *Pygmarrhopalites* A10, contrasting sharply with patterns of genetic variation observed for *Pogonognathellus* T4 populations spanning sinkhole area boundaries (COI, 43%; 16S, 51%) and regions across the Mississippi River (COI, 7%; 16S, 7%) (Table 1.7c–d).

The Mantel test recovered significant IBD patterns for both focal OTUs, although the relationship between genetic distance and geographic distance was more strongly correlated for *Pygmarrhopalites* A10 (COI, r=0.79; 16S, r=0.79) compared to *Pogonognathellus* T4 (COI, r=0.37; 16S, r=0.45) (Table 1.8). For *Pygmarrhopalites* A10, genetic distance remained strongly correlated to sinkhole area (COI, r=0.63; 16S, r=0.5) and position relative to the Mississippi River (COI, r=0.54; 16S, r=0.85) after controlling for geographic distance using partial Mantel tests, while *Pogonognathellus* T4 had weakly positive to slightly negative correlations between genetic distance and sinkhole area (COI, r=-0.19; 16S, r=0.28) and position relative to the Mississippi River (COI, r=-0.21; 16S, r=-0.23). However, the strong patterns of genetic structure among caves for *Pygmarrhopalites* A10 recovered by the AMOVA were weakly or not supported after controlling for geographic distance with partial Mantel tests (COI, r=0.27; 16S, r=0.18). In contrast, strong positive correlations between genetic distance and caves were detected for *Pogonognathellus* T4 (COI, r=0.62; 16S, r=0.63) (Table 1.8).

TCS haplotype networks for COI and 16S revealed concordant relationships, but also showed markedly distinct phylogeographic patterns between *Pygmarrhopalites* A10 and *Pogonognathellus* T4 (Fig. 1.7). *Pygmarrhopalites* A10 haplotypes were geographically structured with all haplotypes segregating by position relative to the Mississippi River and sinkhole area (although 16S haplotypes were shared among neighboring sinkhole areas 2 and 3), whereas *Pogonognathellus* T4 haplotypes did not strongly segregate by geographic barrier, with divergent haplotypes occurring together within sinkhole areas and on both sides of the Mississippi River valley. Additionally, the *Pygmarrhopalites* A10 COI haplotype network clearly illustrates significant levels of sequence divergence between populations spanning the

Mississippi Valley (mean=7.57%) and between populations from Indian Cave (INC) and other Illinois caves (mean=7.44%) (Fig. 1.7).

Phylogeny, divergence times, and topology tests

The rate-calibrated phylogenetic analysis based on the multilocus dataset produced trees with high support for all OTUs identified by the GMYC analysis, and molecular divergence time estimates revealed that all OTU diversification predated the Pliocene (Fig. 1.8). The median age of the most recent common ancestor (MRCA) for all sampled lineages of *Pygmarrhopalites* was estimated to be 47.34 Ma [95% highest posterior density (95% HPD)=39.62–55.98 Ma], and the MRCA of all sampled lineages of *Pogonognathellus* is 15.36 Ma [95% HPD=8.59–23.77 Ma]. The median age of *Pygmarrhopalites* A10 MRCA is 3.77 Ma [95% HPD=2.90–4.76 Ma] and *Pogonognathellus* T4 MRCA is 2.67 [95% HPD=1.2–5.6 Ma]. Reciprocal monophyly across the Mississippi River was only recovered for *Pygmarrhopalites* A10, albeit with low clade support grouping lineages from Indian Cave (INC) with other Illinois lineages from Stemmler Cave (STC), Hoppy Speck Cave (HSC) and Pautler Cave (PAC). *Pygmarrhopalites* A10 populations from Indian Cave diverged from other Illinois lineages ~3.34 Ma [95% HPD=2.43–4.24 Ma], similar to divergence times estimated for Illinois and Missouri populations spanning the Mississippi River ~3.77 Ma [95% HPD=2.90–4.76 Ma]. There was no evidence of vicariance for *Pogonognathellus* T4 lineages, as they were not grouped by position relative to the Mississippi River.

The multilocus phylogeny also shows that two additional OTUs (*Pygmarrhopalites* A3 and A4) contain both Illinois and Missouri lineages, but did not form monophyletic groups by region relative to the Mississippi River. All other OTUs were short-range endemics, from a

single cave (A1, A2, A6, A8, A9, T1–3) or from neighboring cave systems within the same sinkhole area (A5, A7) (Fig. 1.8).

The topology tests (Table 1.9) strongly supported a hypothesis of reciprocal monophyly across the Mississippi River (H_0) for *Pygmarrhopalites* A10 over alternative topological hypotheses H_1 ($2\times\log_e BF=19.94$) and H_2 ($2\times\log_e BF=11.46$). However, reciprocal monophyly (H_0) was strongly rejected in favor of topology H_1 for *Pogonognathellus* T4 ($2\times\log_e BF=-104.82$), corroborating with previous results supporting more recent gene flow across the Mississippi River Valley.

Discussion

Comparative phylogeography

I found that the troglobiont, *Pygmarrhopalites* A10, and the eutroglophile, *Pogonognathellus* T4, have different phylogeographic patterns despite being codistributed across the same geological landscape. Population structure analyses (Tables 1.7–1.8; Fig. 1.7), time-calibrated phylogenetic reconstructions (Fig. 1.8), and topology tests (Table 1.9), indicated that intrinsic differences between these species, such as their degree of cave-dependence, have had major impacts on processes involved in promoting and maintaining genetic isolation in this system (i.e., vicariance and dispersal). Specifically, two important patterns emerged from the comparative phylogeographic analysis. First, sinkhole area boundaries and the Mississippi River Valley were identified as significant dispersal barriers for *Pygmarrhopalites* A10 only. Hierarchical AMOVAs initially revealed that more than half of all genetic variation was distributed among caves, sinkhole areas, and across the Mississippi River Valley for *Pygmarrhopalites* A10, but for *Pogonognathellus* T4, comparable levels of genetic structure were recovered only among caves

(Table 1.7b–d). Mantel tests confirmed geographic distance to be a significant driver of genetic isolation for both taxa (Table 1.8), suggesting springtails are weak dispersers regardless of ecological classification. After controlling for geographic distance using partial Mantel tests, I still recovered significant positive correlations between genetic distance and sinkhole areas and between genetic distance and position relative to the Mississippi River for *Pygmarrhopalites*, but not for *Pogonognathellus* (Table 1.8). The haplotype networks, phylogenetic trees, and topology tests also corroborate these findings providing similar patterns of genetic structure across sinkhole area boundaries (Fig. 1.7) and the Mississippi River valley (Table 1.9; Figs 1.7–1.8). These results indicate that isolation maintained by the Mississippi River valley floodplain and sinkhole area boundaries are driving genetic differentiation in troglobiotic *Pygmarrhopalites*, while IBD is the primary driver of genetic differentiation in eutroglophilic *Pogonognathellus*.

Relative to the Mississippi River valley and sinkhole area boundaries, I observed a very different pattern when genetic variation was partitioned among caves: cave boundaries were identified as significant genetic barriers for *Pogonognathellus* only, whereas patterns of genetic structure among caves identified by the AMOVA for *Pygmarrhopalites* A10 (Table 1.7b) were not supported after accounting for geographic distance (Table 1.8). In this case, patterns of genetic structure among caves are driven by IBD for *Pygmarrhopalites* A10 (not *Pogonognathellus* T4) suggesting that troglobiotic *Pygmarrhopalites* are capable of dispersing between caves. Although this finding appears to contradict previous results, it can still be explained by differences in cave habitat preferences.

Aquatic interstitial subterranean connections joining neighboring cave systems may enable subterranean dispersal during flooding events for *Pygmarrhopalites* A10. This is supported by shared 16S haplotypes between neighboring cave systems (PAC, HSC, and STC)

(Fig. 1.7b). Alluvial aquifers have been implicated as ‘interstitial highways’ for a wide range of subterranean arthropods (Ward & Palmer, 1994), but Collembola are not normally considered members of the interstitial groundwater community as they cannot complete life cycles while submerged (Deharveng et al., 2008). However, growing evidence suggests that they are not only present in these habitats, but can occur in abundance and comprise diverse communities (Bretschko & Christian, 1989; Deharveng et al., 2008; Shaw et al., 2011; Palacios-Vargas et al., 2018). Shaw et al. (2011) documented a unique springtail community from a karstic hyporheic zone that includes species previously associated only with caves. Under this scenario, obligate associations with certain cave habitats may actually facilitate dispersal, colonization, and gene flow among caves connected by subterranean passages.

Cave-to-cave subterranean dispersal is unlikely or infrequent for *Pogonognathellus* because species in this genus do not occur in interstitial habitats and prefer floor or wall surfaces near cave entrances rather than dark zone habitats. This is supported by strong genetic structuring among caves for *Pogonognathellus* T4 indicating that cave-to-cave dispersal is extremely rare for this species despite having naturally occurring surface populations that could presumably facilitate gene flow between caves. Cave-to-cave surface dispersal may be also difficult for this species simply because cave entrances are extremely small features within very large landscapes. Long-term local persistence of cave populations coupled with long-distance dispersal and gene flow contributed by surface populations may explain the presence of both isolation-by-distance patterns across sinkhole area boundaries and river barriers and strong isolation-by-cave patterns in *Pogonognathellus*. However, additional sampling of surface populations for genetic analysis is necessary for a better understanding of the mechanisms driving patterns of genetic structure in *Pogonognathellus* cave populations.

To assess the effect of cave-dependence on patterns of molecular variation, I was required to make informed assumptions about species ecology, including the classification of *Pygmarrhopalites* A10 as a troglobiont. For many small cave-dwelling animals, such as springtails, it is often impossible to ascertain with certainty that a species only occurs in caves (Christiansen, 1962); a species reported only from caves could also be a common soil species, having yet to be reported from surface habitats; the distinction between cavernicolous habitats and other sub-surface microhabitats may be weak or nonexistent for small animals; and troglobionts often lack obvious troglomorphy. Despite these concerns, I am confident that the combination of troglomorphy, close morphological affinities to known troglobiotic species, and their exclusive occurrence in dark or deep twilight cave zones (Table 1.2) provides sufficient evidence that *Pygmarrhopalites* A10 is a troglobiont.

The degree of cave-dependence is certainly a major factor influencing dispersal capacity in cave-dwelling organisms, but additional intrinsic differences between *Pygmarrhopalites* and *Pogonognathellus* may have also contributed to the disparate phylogeographic patterns observed in this study. The genera being compared belong to separate orders of Collembola, differing substantially in size, mobility, and life history. *Pygmarrhopalites* are typically much smaller, less mobile, and have markedly shorter generation times compared to species of *Pogonognathellus*. Moore et al.'s (2005) study on cave Arrhopalitidae documented parthenogenesis and sexual maturation occurring as early as first instar in *Arrhopalites caecus* (Tullberg), 1871 from Wind Cave in South Dakota. Due to the absence of male *Pygmarrhopalites* A10 observed during my study, I cannot rule out asexuality. It is possible that parthenogenesis may have contributed to their lack of genetic structure between caves (after controlling for IBD). The ability for a single female to colonize new habitats without the need for males can facilitate dispersal to neighboring

cave systems, possibly via small subterranean passages in the epikarst or fissures in bedrock during flooding events. However, male *Pygmarrhopalites* are usually present, but rarely encountered (Christiansen & Bellinger, 1996), and males have also been reported for *P. pavo*, a species that morphologically similar to *Pygmarrhopalites* A10 (Christiansen & Bellinger, 1996).

Biogeography: evidence for vicariance across the Mississippi River valley

The climatic and geological changes during the Pleistocene and their impacts on the distribution and diversity of North American cave fauna have been well documented (Porter, 2007). For example, the modern course of the Ohio River, formed by changing climate during the Pleistocene, bisects a major cave-bearing karst region along the Indiana-Kentucky border. Niemiller et al. (2013b) demonstrated that this river is a major biogeographic barrier, facilitating the divergence and subsequent isolation and speciation of troglobiotic cavefish populations. Like the Ohio River, the Mississippi River has also been implicated as a "hard" geographic barrier to dispersal for many surface species (e.g., Soltis et al., 2006), but its influence on the evolutionary history of cave-dwelling organisms has yet to be evaluated, in part, because the geological history of the Mississippi River and its influence on regional cave-bearing karst remains poorly understood.

Molecular divergence times of *Pygmarrhopalites* A10 populations spanning the Mississippi (Figs 1.8–1.9), patterns of genetic structure (Tables 1.7d, 1.8; Fig. 1.7), and topology tests (Table 1.9) are consistent with the hypothesis that vicariance is the primary driver of genetic isolation in this group—providing prima facie evidence of vicariance across the Mississippi River for terrestrial cave arthropods, and accordingly, the first biogeographic evidence for the initial timing of Mississippi River entrenchment and bisection through Salem Plateau karst in

Illinois and Missouri. An emerging chronological snapshot of ancestral Mississippi River geology coincides with my divergence time estimates. According to upland gravel distributions, the Mississippi River entrenched along its entire length during late Pliocene or early Pleistocene possibly due to a glacioeustatic lowering of sea level (Cupples & Van Arsdale, 2014). The combination of an increased late Pliocene Mississippi River discharge (Cox et al., 2014) and subsequent Pleistocene glacial melt cycles beginning ~2.5 Ma (Balco et al., 2005; Balco & Rovey, 2008) could have facilitated valley growth and the erosion of karst habitat during the Pleistocene in this region. This scenario is further supported by the ages of burial sediments deposited in Mammoth Cave (Granger et al., 2001), which provide a compelling parallel history of karst entrenchment by the Green River, a tributary within the Mississippi River watershed. Granger et al. (2001) attributed the oldest sediment deposit (~3.3 Ma) to a glacioeustatic lowering in sea level, immediately followed by a period of Green River excavation and bedrock incision lasting ~0.9 Ma. It is not unreasonable to consider that these climatic processes had the same effects on the geology of other karst regions occurring within the Mississippi River watershed (e.g., the Salem Plateau karst). Vicariance can also explain the strong patterns of isolation observed for the Indian Cave *Pygmarrhopalites* A10 populations (Fig. 1.6), which are separated from other Illinois populations by the Kaskaskia River. Divergence time estimates between Indian Cave (INC) and other Illinois populations (~3.34 Ma) are similar to those estimated between Illinois and Missouri populations (~3.77 Ma), which makes sense given the Kaskaskia River is a tributary of the Mississippi River: the karst entrenchment processes must have occurred during, or shortly after, it began for the Mississippi River.

The corroboration of timing information derived from both biological and geological data (Fig. 1.9) supports the hypothesis that climatic and geological events beginning in the late

Pliocene initiated and maintained genetic isolation between troglobiotic springtail populations in Illinois and Missouri, but the exact mode of gene flow across the preglacial Mississippi River and tributaries, prior to their genetic isolation, is not known. It is plausible that sections of karst were periodically isolated and rejoined by shifting meanders and periods of low flow, later removed by Plio-Pleistocene entrenchment and excavation, providing intermittent subterranean passage for cave organisms until the late Pliocene or early Pleistocene.

The lack of genetic structure across the Mississippi River (Tables 1.7d, 1.8; Fig. 1.7) and non-monophyly (Table 1.9; Fig. 1.8) for Illinois and Missouri populations of *Pogonognathellus* T4 cannot be explained by this scenario and is instead more consistent with a hypothesis of dispersal, rather than vicariance. Species in the genus *Pogonognathellus* are not ecologically restricted to caves and can maintain large surface populations, thus they are unlikely to be affected by surface barriers to the same degree as obligate cave-dwellers. Reports of springtails travelling vast distances via water surfaces (Coulson et al., 2002), rafting (Hawes et al., 2008), and air currents (e.g., Freeman, 1952; Blackith & Disney, 1988; Coulson et al., 2003; Hawes et al., 2007) highlight a number of potential means for *Pogonognathellus* to passively disperse across the Mississippi River valley that are typically unavailable to obligate subterranean springtails.

Cryptic diversity, short-range endemism, and implications for conservation

Recent discoveries of cryptic species have challenged our current understanding of biological diversity (Fišer et al., 2018), and this paradigm shift is particularly evident in subterranean habitats where ideal conditions have fostered widespread cryptic speciation, including examples of recent divergence in cavefish (Niemiller et al., 2013b), morphological stasis in amphipods

(Trontelj et al., 2009), and morphological convergence in springtails (Christiansen, 1961). Therefore, it was important in this study to detect the presence of cryptic diversity and delimit OTUs prior to phylogeographic comparisons, to avoid interpreting interspecific variation as population-level genetic structure. Large gaps in genetic distance frequencies (Fig. 1.3) and the presence of strong interspecific genetic structure within caves (Table 1.7a) provided evidence of highly divergent and sympatric lineages in cave samples. GMYC analysis identified 14 putative species within two morphospecies (10 *Pygmarrhopalites* and 4 *Pogonognathellus* OTUs), corroborated by the presence of large gaps between inter- and intra-OTU distances for COI and 16S (Fig. 1.6). Lastly, minute differences in morphology among *Pygmarrhopalites* OTUs were also observed under compound light microscopy, providing additional support for the recognition of 10 distinct (and possibly new) species for this genus.

Only 38 species of *Pygmarrhopalites* and 11 species of *Pogonognathellus* are currently reported for all of North America (Christiansen & Bellinger, 1998; Zeppelini & Christiansen, 2003; Zeppelini et al., 2009; Felderhoff et al., 2010; Park et al., 2011; Soto-Adames & Taylor, 2013). Hence, the discovery of 14 putative species was surprising given the relatively small geographic scale of this study (~150km stretch along the banks of the Mississippi River valley). Molecular data also revealed that 7 of the 10 *Pygmarrhopalites* OTUs may be single-site endemics or have restricted ranges (Fig. 1.8), raising potential conservation concerns.

The detection of short-range endemics, genetic isolation, and apparent cryptic diversity has major conservation implications. Reduced dispersal capacity observed for *Pygmarrhopalites* can increase their susceptibility to human disturbances such as land use practices, climate change, pollution, and invasive species—all of which pose major threats to fragile cave ecosystems (Culver & Pipan, 2009; Taylor & Niemiller, 2016). In fact, growing concerns of

karst groundwater contamination (Panno et al., 1996) prompted *Pygmarrhopalites madonnensis* (Zeppelini & Christiansen), 2003, a troglobiotic springtail known from a single cave in Monroe Co., Illinois, to be listed as state endangered (Mankowsi, 2010). This is concerning considering that my data indicates that single-site endemics are not only extremely common but may also comprise the majority of troglobiotic springtail diversity throughout this region. Lastly, unrecognized cryptic species complexes with allopatric ranges, presumed to be a single widely distributed species, may lead to misguided biodiversity conservation and management decisions.

Conclusions

Salem Plateau caves and their springtail inhabitants provide a model system for comparative phylogeographic studies to address important questions in evolution and subterranean biogeography. I characterized and compared patterns of molecular diversity between species in the genera *Pygmarrhopalites* and *Pogonognathellus*, which led to three important findings. First, conflicting phylogeographic patterns between troglobiotic and eutroglophilic species distributed across the same geographic barriers suggests that different degrees of cave-dependence can have major impacts on the dispersal capacity and genetic connectivity of cave organisms. Second, estimates of genetic structure and molecular divergence indicate that climatic and geological processes during the late Pliocene/early Pleistocene were major factors driving isolation between populations of troglobiotic cave organisms in Salem Plateau karst spanning the Mississippi River in Illinois and Missouri. Lastly, the large number of deeply divergent lineages and high rates of short-range endemism detected in this study exposes a major knowledge gap in our understanding of cave microarthropod diversity and highlights potential conservation concerns under growing threats to cave biodiversity. Additional phylogeographic research on cave

arthropods will further contribute to our understanding of how and why organisms occupy, persist in, and adapt to cave environments—information critical for the development and implementation of conservation strategies needed to manage and protect cave biodiversity (Porter, 2007).

Acknowledgments

This work would not have been possible without the help of the Southeast Missouri Grotto (SEMO), the Middle Mississippi Valley Grotto (MMV), the Meramec Valley Grotto (MVG), the Little Egypt Grotto (LEG), the Missouri Speleological Survey (MSS), the Illinois Speleological Survey (ISS), the Missouri Department of Conservation (MDC), the Illinois Department of Natural Resources (IDNR), the Saint Louis County Department of Parks, the Illinois Nature Preserves Commission (INPC), Clifftop, and many others, who provided us with access to cave locations, permission to collect, assistance in field, and/or collection and research permits (IDNR Research Permit # SS16-050, MDC Wildlife Collector's Permit #17003, INPC Special-use Permit). For this, I would like to especially thank Chad McCain, Ray Shaw, Dan Lamping, Tony Schmidt, Mark Brewer, Bob Weck, Aaron Addison, Jim Sherrell, Erin Ellison, Matt Safford, Dan Swanson, Laura Bartol Belarbi, Jon Beard, Shelly Colatskie, Jason Crites, Dane Driskill, Tony Elliott, Phillip Ellison, Rick Haley, Tom Hellauer, Derik Holtmann, Scott House, Mihai Lefticariu, Joe Light, Andrew Lloyd, Lisa Meisel, Brett Meisel, Matt Niemiller, Bob Osburn, Paul Stevens, Mick Sutton, Shawn Williams, Russ Clendenin, and Richard Young. I am also grateful to Brandon Curry, Sam Panno, and Joe Devera for their geological expertise. This project was supported by the National Speleological Society (Ralph W. Stone Graduate Fellowship), the Cave Conservancy Foundation (Graduate Fellowship in Karst Studies), the

University of Illinois at Urbana-Champaign (Doctoral Dissertation Completion Fellowship), the Illinois Natural History Survey (Ross Memorial Fund), and the Subterranean Ecology Institute. I would also like to thank the faculty, staff, and graduate students of the Department of Entomology at the University of Illinois at Urbana-Champaign for their support.

Tables

Table 1.1 Caves sampled for this study including abbreviated cave names, locality information, and number of *Pygmarrhopalites* (A) and *Pogonognathellus* (T) specimens sequenced from each cave.

Cave	Cave abbrev.	State	County	Sinkhole area[†]	A(n)	T(n)
Ava Cave	AVA	IL	Jackson	5	1	5
Polystichum acrostichoides Sink Cave	PSC	IL	Jackson	5	1	1
Auctioneer Cave	AUC	IL	Monroe	3	0	0
Fogelpole Cave	FPC	IL	Monroe	3	6	8
Hoppy Speck Cave	HSC	IL	Monroe	3	4	3
Illinois Caverns	ILC	IL	Monroe	3	1	2
Mead Jars Pit	MJP	IL	Monroe	3	0	1
Pautler Cave	PAC	IL	Monroe	3	4	1
White Rock Mine	WRM	IL	Monroe	3	1	0
Indian Cave	INC	IL	Randolph	5	2	1
Sphalloplana Stream Cave	SPC	IL	St. Clair	2	0	0
Stemmler Cave	STC	IL	St. Clair	2	1	0
Apple Creek 1	AC1	MO	Cape Girardeau	6	0	1
Apple Creek 3	AC3	MO	Cape Girardeau	6	0	1
Apple Creek 4	AC4	MO	Cape Girardeau	6	0	0
Apple Creek 5	AC5	MO	Cape Girardeau	6	0	0
Flat Rock Creek Cave	FRC	MO	Cape Girardeau	6	0	0
Apple Creek 2	AC2	MO	Perry	6	0	0
Berome Moore Cave	BMC	MO	Perry	6	5	0
Mystery Cave	MYC	MO	Perry	6	4	0
Seventy-Six Cave	SSC	MO	Perry	6	8	0
Streiler City Cave	SCC	MO	Perry	6	1	0
Tom Moore Cave	TMC	MO	Perry	6	1	1
Cliff Cave	CFC	MO	St. Louis	1	1	0
Esoteric Cave	ESC	MO	Ste. Genevieve	4	1	0

[†]Refers to high density sinkhole areas in the study area defined for Illinois (Panno et al., 1997; Panno et al., 1999; Venarsky et al., 2009) and Missouri (Panno et al., 1999; Burr et al., 2001) (See Fig. 1.1).

Table 1.2. List of all taxa used in this study, including locality, sample information, and GenBank accession numbers for all sequences.

Specimen ID[†]	State	County	Cave[‡]	Zone[§]	Habitat[¶]	COI	16S	28S D1-3	28S D7-	Histone-3
Pygmarrhopalites										
AVC_179IL1	IL	Jackson	AVA	T	LL	MH269587	MH269419	MH269482	MH269524	MH269655
BMC_091MO1	MO	Perry	BMC	D	DP	MH269588	MH269420	MH269483	MH269525	MH269656
BMC_091MO3	MO	Perry	BMC	D	DP	MH269589	MH269421	MH269484	MH269526	MH269657
BMC_091MO4	MO	Perry	BMC	D	DP	MH269590	—	MH269485	MH269527	MH269658
BMC_093MO2	MO	Perry	BMC	D	LL	MH269591	MH269422	MH269486	MH269528	MH269659
BMC_097MO2	MO	Perry	BMC	D	RS	MH269592	MH269423	MH269487	MH269529	MH269660
CFC_108MO1	MO	St. Louis	CFC	D	DP	MH269593	MH269424	MH269488	MH269530	MH269661
ESC_084MO1	MO	Ste. Genevieve	ESC	D	DP	MH269594	MH269425	MH269489	MH269531	MH269662
FPC_027IL1	IL	Monroe	FPC	D	LL	MH269595	MH269426	MH269490	MH269532	MH269663
FPC_027IL2	IL	Monroe	FPC	D	LL	MH269596	MH269427	MH269491	MH269533	MH269664
FPC_027IL3	IL	Monroe	FPC	D	LL	MH269597	MH269428	MH269492	MH269534	MH269665
FPC_028IL1	IL	Monroe	FPC	T	LL	MH269598	MH269429	MH269493	MH269535	MH269666
FPC_029IL2	IL	Monroe	FPC	E	LL	MH269599	MH269430	MH269494	MH269536	MH269667
FPC_M10IL1	IL	Monroe	FPC	D	LL	MH269600	—	MH269495	—	MH269668
HSC_006IL1	IL	Monroe	HSC	DT	LL	MH269601	MH269431	MH269496	MH269537	MH269669
HSC_006IL2	IL	Monroe	HSC	DT	LL	MH269602	MH269432	MH269497	MH269538	MH269670
HSC_006IL3	IL	Monroe	HSC	DT	LL	MH269603	MH269433	MH269498	—	MH269671
HSC_007IL2	IL	Monroe	HSC	T	LL	MH269604	MH269434	MH269499	MH269539	MH269672
ILC_193IL1	IL	Monroe	ILC	T	LL	MH269605	MH269435	MH269500	MH269540	MH269673
INC_158IL1	IL	Randolph	INC	D	LL	MH269606	MH269436	MH269501	MH269541	MH269674
INC_162IL1	IL	Randolph	INC	D	ST	MH269607	MH269437	MH269502	MH269542	MH269675
MYC_121MO1	MO	Perry	MYC	D	LL	MH269608	MH269438	MH269503	MH269543	MH269676
MYC_122MO2	MO	Perry	MYC	D	ST	MH269609	MH269439	MH269504	MH269544	MH269677
MYC_124MO2	MO	Perry	MYC	D	DP	MH269610	MH269440	MH269505	MH269545	MH269678
MYC_126MO2	MO	Perry	MYC	D	LL	MH269611	MH269441	MH269506	MH269546	MH269679
PAC_217IL1	IL	Monroe	PAC	D	ST	MH269612	MH269442	MH269507	MH269547	MH269680
PAC_217IL4	IL	Monroe	PAC	D	ST	MH269613	MH269443	MH269508	MH269548	MH269681
PAC_217IL5	IL	Monroe	PAC	D	ST	MH269614	MH269444	MH269509	MH269549	MH269682
PAC_217IL6	IL	Monroe	PAC	D	ST	MH269615	MH269445	MH269510	MH269550	MH269683
PSC_155IL1	IL	Jackson	PSC	E	LL	MH269616	MH269446	MH269511	MH269551	MH269684
SCC_143MO1	MO	Perry	SCC	D	DP	MH269617	MH269447	MH269512	MH269552	MH269685
SSC_141MO1	MO	Perry	SSC	D	LL	MH269618	MH269448	MH269513	MH269553	MH269686
SSC_141MO2	MO	Perry	SSC	D	LL	MH269619	MH269449	MH269514	MH269554	MH269687
SSC_141MO3	MO	Perry	SSC	D	LL	MH269620	MH269450	MH269515	MH269555	MH269688
SSC_142MO1	MO	Perry	SSC	DT	LL	MH269621	MH269451	MH269516	MH269556	MH269689
SSC_165MO1	MO	Perry	SSC	D	GN	MH269622	MH269452	MH269517	MH269557	MH269690
SSC_165MO2	MO	Perry	SSC	D	GN	MH269623	MH269453	MH269518	MH269558	MH269691
SSC_165MO3	MO	Perry	SSC	D	GN	MH269624	MH269454	MH269519	MH269559	MH269692
SSC_167MO1	MO	Perry	SSC	D	LL	MH269625	MH269455	MH269520	MH269560	MH269693
STC_199IL1	IL	St. Clair	STC	D	DP	MH269626	MH269456	MH269521	MH269561	MH269694
TMC_101MO2	MO	Perry	TMC	D	RS	MH269627	—	MH269522	MH269562	MH269695
WRM_033IL1	IL	Monroe	WRM	E	LL	MH269628	MH269457	MH269523	MH269563	MH269696
Pogonognathellus										
AC1_133MO1	MO	Cape Girardeau	AC1	T	RS	MH269629	MH269458	—	MH269564	—
AC3_134MO1	MO	Cape Girardeau	AC3	T	RS	MH269630	MH269459	—	MH269565	—
AVC_171IL1	IL	Jackson	AVA	T	LL	MH269631	MH269460	—	MH269566	—
AVC_172IL1	IL	Jackson	AVA	T	LL	MH269632	—	—	—	—
AVC_173IL1	IL	Jackson	AVA	D	LL	MH269633	MH269461	—	MH269567	—
AVC_175IL2	IL	Jackson	AVA	D	LL	MH269634	MH269462	—	MH269568	—
AVC_179IL1	IL	Jackson	AVA	T	LL	MH269635	MH269463	—	MH269569	—
FPC_027IL1	IL	Monroe	FPC	D	LL	MH269636	MH269464	—	MH269570	—
FPC_027IL2	IL	Monroe	FPC	D	LL	MH269637	MH269465	—	MH269571	—
FPC_028IL1	IL	Monroe	FPC	T	LL	MH269638	MH269466	—	MH269572	—
FPC_029IL1	IL	Monroe	FPC	E	LL	MH269639	MH269467	—	MH269573	—
FPC_064IL1	IL	Monroe	FPC	D	LL	MH269640	MH269468	—	MH269574	—
FPC_068IL1	IL	Monroe	FPC	D	LL	MH269641	MH269469	—	MH269575	—
FPC_073IL1	IL	Monroe	FPC	D	LL	MH269642	MH269470	—	MH269576	—
FPC_073IL2	IL	Monroe	FPC	D	LL	MH269643	MH269471	—	MH269577	—
FPC_M20IL1	IL	Monroe	FPC	D	LL	MH269644	MH269472	—	MH269578	—
HSC_002IL1	IL	Monroe	HSC	T	LL	MH269645	MH269473	—	—	—
HSC_005IL1	IL	Monroe	HSC	T	LL	MH269646	MH269474	—	MH269579	—
HSC_006IL1	IL	Monroe	HSC	DT	LL	MH269647	MH269475	—	MH269580	—
ILC_192IL1	IL	Monroe	ILC	D	LL	MH269648	MH269476	—	MH269581	—
ILC_193IL1	IL	Monroe	ILC	T	LL	MH269649	MH269477	—	MH269582	—
INC_163IL1	IL	Randolph	INC	E	LL	MH269650	MH269478	—	MH269583	—

Table 1.2 (continued)

MJP_013IL1	IL	Monroe	MJP	T	LL	MH269651	—	—	—	—
PAC_213IL1	IL	Monroe	PAC	D	LL	MH269652	MH269479	—	MH269584	—
PSC_155IL1	IL	Jackson	PSC	E	LL	MH269653	MH269480	—	MH269585	—
TMC_106MO1	MO	Perry	TMC	D	LL	MH269654	MH269481	—	MH269586	—

[†]Sampled specimens for each genus are labeled according to Cave_SampleStateSpecimen.
[‡]Cave abbreviations are listed in Table 1.1.
[§]Zone abbreviations: E, entrance; T, twilight; DT, deep twilight; D, dark.
[¶]Habitat abbreviations: E, entrance; T, twilight; DT, deep twilight; D, dark. Habitat abbreviations: LL, leaf litter and other decaying plant material; DP, drip pools; RS, rock surfaces; ST, scat; GN, guano.

Table 1.3. PCR and sequencing primers used in this study, including original references.

Gene	Primer	Sequence (5'–3')	Length[†] (bp)	Reference
COI	jgLCO1490	TITCIACIAAYCAYAARGAYATTGG	597, 674	Geller et al., 2013
	jgHCO2198	TAIACYTCIGGRTGICCRAARAAYCA		Geller et al., 2013
16S	LR-J-12887M	CCGGTCTGAACTCAAATCATGT	315, 528	Zhang et al., 2014
	LR-N-12887M	CGACTGTTTAACAAAAACAT		Zhang et al., 2014
28S D1–3[‡]	28SrD1.2a	CCCSSGTAATTTAAGCATATTA	1272	Whiting, 2002
	28Sbout	CCCACAGCGCCAGTTCTGCTTACC		Giribet et al., 2001
28S D7–10	AS1	CCGCAGCAGGTCTCCAAGGTGAA	845, 857	Xiong et al., 2008
	OP4	CCGCCCCAGTCAAACTCCC		Xiong et al., 2008
Histone-3[2]	H3F1	ATGGCTCGTACCAAGCAGACVGC	329	Colgan et al., 1998
	H3R2	GTAACAATCATGCCCAAGGAYAT		Colgan et al., 1998

[†]Length of final alignment (*Pygmarrhopalites*, *Pogonognathellus*)
[‡]Due to inconsistent amplification, these loci were excluded for *Pogonognathellus*.

Table 1.4. Outgroup taxa and GenBank sequence accession numbers for each locus included the (a) *Pygmarrhopalites* and (b) *Pogonognathellus* datasets.

Outgroups	COI	16S	28S D1-3	28S D7-10	Histone-3
(a)					
Allacma fusca	KT808323	—	EU376054	EU376054	—
Sminthurinus bimaculatus	AY555545	AY555555	AF483398	—	AY555566
Sminthurus viridis	NC_010536	NC_010536	EF199973	EF199973	—
(b)					
Cryptopygus antarcticus	NC_010533	NC_010533	—	EF199971	—
Folsomia candida	KU198392	KU198392	—	JN981046	—
Lepidophorella sp.	KJ716832	—	—	—	—

Table 1.5. The 95% HPD of site models selected by bModelTest for (**a**) the COI gene tree analyses used for GMYC species delimitation and (**b**) the rate-calibrated multilocus phylogenic analyses used for divergence time estimations. PS, posterior support for a model; CS, cumulative posterior support for a model; M, model number implemented in bModelTest. See Appendix in Bouckaert & Drummund (2017) for more information regarding specific model numbers.

(a)	COI		**(b)**	16S			COI			28S D7-10			28S D1-3			Histone-3		
PS	CS	M	PS	CS	M	PS	CS	M	PS	CS	M	PS	CS	M	PS	CS	M
Pygmarrhopalites																		
26.67%	26.67%	121123	13.33%	13.33%	123421	27.65%	27.65%	121123	11.48%	11.48%	123121	28.19%	28.19%	123123	22.08%	22.08%	121121
11.90%	38.56%	121323	12.58%	25.91%	123121	11.40%	39.05%	121323	10.34%	21.82%	123123	21.33%	49.52%	123423	8.07%	30.15%	121131
11.90%	48.55%	121134	7.73%	33.65%	123124	9.66%	48.71%	121134	9.46%	31.27%	121121	19.05%	68.57%	123124	6.55%	36.70%	123121
11.90%	55.53%	121324	6.07%	39.71%	123123	7.00%	55.71%	121324	7.06%	38.33%	123323	13.18%	81.74%	123425	6.45%	43.15%	123321
11.90%	60.10%	121131	5.57%	45.28%	121321	5.08%	60.79%	121121	5.46%	43.79%	123424	4.65%	86.39%	123143	6.19%	49.33%	123323
11.90%	64.65%	121343	5.41%	50.69%	123425	4.74%	65.53%	121131	5.34%	49.12%	123124	3.55%	89.95%	123453	5.81%	55.15%	121123
11.90%	69.15%	123124	4.84%	55.53%	123451	4.55%	70.08%	123124	5.06%	54.19%	123321	3.31%	93.26%	123421	5.37%	60.52%	121321
11.90%	73.27%	121121	4.61%	60.14%	123143	4.18%	74.26%	121343	4.41%	58.60%	121321	3.30%	96.56%	123145	4.86%	65.37%	123123
11.90%	77.17%	123324	4.51%	64.65%	123423	3.81%	78.07%	123324	4.29%	62.88%	123423				4.61%	69.98%	121323
11.90%	79.53%	121345	4.44%	69.09%	123141	2.17%	80.24%	121345	4.20%	67.09%	123421				2.34%	72.33%	123341
11.90%	81.44%	123145	4.25%	73.34%	123424	1.72%	81.96%	123141	3.11%	70.20%	121123				2.32%	74.64%	123141
11.90%	83.31%	123141	3.87%	77.22%	123145	1.66%	83.62%	123424	2.90%	73.10%	123141				2.25%	76.90%	123343
11.90%	85.06%	123424	2.66%	79.88%	121324	1.66%	85.28%	123121	2.86%	75.96%	121323				2.05%	78.94%	121341
11.90%	86.74%	123345	2.60%	82.48%	123453	1.56%	86.84%	123145	2.86%	78.82%	123324				2.05%	80.99%	121134
11.90%	88.41%	123341	2.31%	84.79%	121323	1.55%	88.39%	123345	2.60%	81.41%	123143				1.97%	82.96%	123324
11.90%	90.03%	123121	2.24%	87.03%	123456	1.51%	89.91%	123341	2.41%	83.82%	121131				1.93%	84.89%	123421
11.90%	91.53%	123321	1.94%	88.97%	121121	1.32%	91.23%	123321	2.34%	86.17%	123425				1.92%	86.82%	123124
11.90%	92.65%	123425	1.75%	90.72%	123454	1.11%	92.34%	121341	1.87%	88.03%	123343				1.75%	88.57%	123424
11.90%	93.76%	121341	1.71%	92.43%	121341	1.06%	93.40%	123425	1.47%	89.51%	123454				1.67%	90.24%	121343
11.90%	94.86%	121321	1.46%	93.88%	123321	1.05%	94.45%	121321	1.43%	90.93%	123145				1.67%	91.91%	123143
11.90%	95.69%	123454	1.15%	95.04%	123323	0.96%	95.41%	123323	1.32%	92.26%	123341				1.57%	93.47%	123423
									1.27%	93.53%	121324				1.52%	94.99%	121324
									1.15%	94.68%	123451				0.75%	95.74%	123451
									1.11%	95.79%	121341						
Pogonognathellus																	
31.25%	31.25%	123141	24.67%	24.67%	123124	34.26%	34.26%	123141	12.50%	12.50%	123421						
21.69%	52.93%	123121	21.37%	46.03%	123145	15.82%	50.08%	123121	12.44%	24.94%	121323						
10.42%	63.35%	123451	9.53%	55.57%	123425	13.28%	63.36%	123451	10.18%	35.12%	121324						
7.55%	70.89%	123145	8.62%	64.18%	123424	8.45%	71.81%	123145	8.85%	43.97%	123424						
7.33%	78.22%	123454	7.98%	72.16%	123456	8.21%	80.01%	123454	7.81%	51.78%	123121						
7.12%	85.34%	123421	7.11%	79.27%	123454	6.34%	86.36%	123421	7.63%	59.41%	121321						
5.06%	90.40%	123424	7.00%	86.28%	123121	4.09%	90.44%	123424	6.94%	66.35%	123425						
5.06%	95.46%	123124	6.01%	92.28%	123141	3.58%	94.03%	123124	4.44%	70.78%	123451						
			4.20%	96.49%	123421	3.17%	97.19%	123456	4.21%	74.99%	121343						
									3.47%	78.46%	121345						
									3.04%	81.50%	123454						
									2.77%	84.27%	121341						
									2.69%	86.96%	123141						
									2.61%	89.57%	123423						
									2.52%	92.08%	123124						
									2.32%	94.41%	123456						
									1.00%	95.41%	123453						

Table 1.6. Site model statistics for (**a**) the COI gene tree analyses used for GMYC species delimitation and (**b**) the rate-calibrated multilocus phylogenic analyses, selected by bModelTest.

Site Partition	Summary Statistic	BMT Model Indicator	BMT gamma Shape	BMT Proportion Invariable	has Gamma Rates	has Invariable Sites	Active Prop Invariable	Active Gamma Shape	has Equal Freqs
(a)									
Pygmarrhopalites COI	mean	7.68	0.4791	0.2880	1	0.9981	0.2879	0.4791	0
	stderr	0.03	0.0015	0.0008		0.0004	0.0008	0.0015	
	stdev	7.09	0.1170	0.0635		0.0438	0.0639	0.1170	
	variance	50.26	0.0137	0.0040		0.0019	0.0041	0.0137	
	median	6	0.4677	0.2938		1	0.2938	0.4677	
	95% HPD	1–23	0.2505–0.7189	0.1597–0.4091		1–1	0.1597–0.4091	0.2505–0.7189	
Pogonognathellus COI	mean	17.74	0.9482	0.4055	1	0.9573	0.4029	0.9482	0
	stderr	0.03	0.0098	0.0034		0.0052	0.0037	0.0098	
	stdev	6.60	0.5460	0.1302		0.2022	0.1370	0.5460	
	variance	43.60	0.2981	0.0170		0.0409	0.0188	0.2981	
	median	14	0.8378	0.4431		1	0.4431	0.8378	
	95% HPD	11–29	0.1793–1.9852	0.0804–0.5794		1–1	0–0.5457	0.1793–1.9852	
(b)									
Pygmarrhopalites COI	mean	7.76	0.4142	0.2666	1	0.9927	0.2661	0.4142	0
	stderr	0.04	0.0012	0.0009		0.0012	0.0010	0.0012	
	stdev	7.03	0.0867	0.0623		0.0853	0.0640	0.0867	
	variance	49.36	0.0075	0.0039		0.0073	0.0041	0.0075	
	median	6	0.4123	0.2742		1	0.2742	0.4123	
	95% HPD	1–23	0.2137– 0.5682	0.1305–0.3831		1–1	0.1297–0.3842	0.2137–0.5682	
Pygmarrhopalites 16S	mean	16.67	0.6809	0.2428	1	0.8279	0.2315	0.6809	0
	stderr	0.06	0.0070	2.83E-03		0.0068	0.0032	0.0070	
	stdev	7.77	0.3710	0.1337		0.3775	0.1484	0.3710	
	variance	60.42	0.1377	0.0179		0.1425	0.0220	0.1377	
	median	15	0.5944	0.2663		1	0.2661	0.5944	
	95% HPD	5–30	0.2232–1.4050	0.0002–0.4260		0–1	0–0.4259	0.2232–1.4050	
Pygmarrhopalites 28S D1-3	mean	18.23	0.2302	0.4300	0.2397	0.9912	0.7378	0.2182	0.4066
	stderr	0.04	0.0011	0.0134	0.0047	0.0022	0.0066	0.0051	0.0096
	stdev	6.36	0.0985	0.4585	0.1786	0.0936	0.4398	0.1968	0.3978
	variance	40.51	0.0097	0.2102	0.0319	0.0088	0.1935	0.0387	0.1582
	median	15	0.2067	0.2779	0.2071	1	1	0.1950	0.2733
	95% HPD	12–28	0.0823–0.4460	0.0859–1.2548	0–0.5659	1–1	0–1	0–0.5659	0–1
Pygmarrhopalites 28S D7-10	mean	13.81	0.3008	0.3569	0.9914	0.8878	0.3431	0.2820	0.9527
	stderr	0.04	0.0112	0.0063	0.0017	0.0043	0.0067	0.0091	0.0131
	stdev	7.93	0.4254	0.2163	0.0924	0.3156	0.2322	0.3697	0.2124
	variance	62.96	0.1810	0.0468	0.0085	0.0996	0.0539	0.1367	0.0451
	median	13	0.1738	0.3444	1	1	0.3437	0.1695	1
	95% HPD	1–26	0.0010–1.0265	0.0012–0.7078	1–1	0–1	0–0.7064	0–0.9446	1–1
Pygmarrhopalites Hitone-3	mean	9.71	0.5719	0.3467	0.9988	0.8299	0.3339	0.5698	0.9979
	stderr	0.04	0.0110	0.0053	0.0004	0.0070	0.0058	0.0109	0.0007
	stdev	8.04	0.5224	0.2062	0.0345	0.3757	0.2227	0.5203	0.0462
	variance	64.71	0.2729	0.0425	0.0012	0.1411	0.0496	0.2707	0.0021
	median	8	0.3938	0.3864	1	1	0.3864	0.3926	1
	95% HPD	1–24	0.0990–1.6257	0–0.6238	1–1	0–1	0–0.6238	0.0968–1.6300	1–1
Pogonognathellus COI	mean	18.38	0.6127	0.3685	1	0.9781	0.3671	0.6127	0
	stderr	0.04	0.0054	0.0023		0.0025	0.0024	0.0054	
	stdev	6.65	0.3087	0.1139		0.1462	0.1177	0.3087	
	variance	44.20	0.0953	0.0130		0.0214	0.0138	0.0953	
	median	14	0.5428	0.3886		1	0.3886	0.5428	
	95% HPD	11–29	0.1793–1.2172	0.1325–0.5566		1–1	0.1273–0.5573	0.1793–1.2172	
Pogonognathellus 16S	mean	19.23	0.5593	0.1219	0.9995	0.5457	0.0906	0.5578	0
	stderr	0.04	0.0068	0.0022	0.0003	0.0069	0.0025	0.0066	
	stdev	6.64	0.3012	0.1010	0.0230	0.4979	0.1151	0.2965	
	variance	44.14	0.0907	0.0102	0.0005	0.2479	0.0132	0.0879	
	median	16	0.4640	0.0939	1	1	0.0256	0.4639	
	95% HPD	11–30	0.2569–1.1272	0–0.3258	1–1	0–1	0–0.3256	0.2569–1.1272	
Pogonognathellus 28S D7-10	mean	15.42	0.4011	0.2757	0.9835	0.7210	0.2320	0.3755	0.9936
	stderr	0.05	0.0074	0.0030	0.0015	0.0040	0.0034	0.0055	0.0023
	stdev	8.88	0.4919	0.1941	0.1275	0.4485	0.2208	0.4488	0.0798
	variance	78.81	0.2420	0.0377	0.0163	0.2011	0.0488	0.2014	0.0064
	median	11	0.2507	0.2437	1	1	0.1898	0.2409	1
	95% HPD	5–29	0.0010–1.2880	0–0.6346	1–1	0–1	0–0.6344	0–1.1828	1–1

Table 1.7. Hierarchical analysis of molecular variance (AMOVA) of COI and 16S for *Pygmarrhopalites* and *Pogonognathellus* grouped by (**a**) cave for all sampled taxa, (**b**) cave for focal OTUs only, (**c**) sinkhole area for focal OTUs only, and (**d**) by region relative to the Mississippi River for focal OTUs only. Significance is based on 50,000 permutations: *P < 0.05, **P < 0.01, ***P < 0.001.

Source of variation	COI					16S				
	df †	SS‡	VC§	V%¶	φ-Statistics	df †	SS‡	VC§	V%¶	φ-Statistics
(a) All taxa, by cave										
Pygmarrhopalites spp.										
Among caves	15	1785.7	20.53	28.62	φct = 0.29**	14	658.5	6.6	22.48	φct = 0.23*
Among samples within caves	13	829.91	34.2	47.68	φsc = 0.67**	12	343.73	16.46	56.08	φsc = 0.72**
Within samples	13	221.08	17.01	23.71	φst = 0.76***	12	75.5	6.29	21.44	φst = 0.79***
Pogonognathellus spp.										
Among caves	10	337.9	11.61	56.37	φct = 0.56*	9	193.08	7.66	60.61	φct = 0.61*
Among samples within caves	13	130.02	8.49	41.21	φsc = 0.94	12	67.5	4.98	39.39	φsc = 1
Within samples	2	1	0.5	2.43	φst = 0.98*	2	0	0	0	φst = 1*
(b) Focal OTUs only, by cave										
Pygmarrhopalites A10										
Among caves	4	170.58	15.68	87.98	φct = 0.88	4	31.25	2.93	90.72	φct = 0.91
Among pops within caves	2	5.25	1.93	10.85	φsc = 0.9	2	0.75	0.3	9.28	φsc = 1
Within populations	6	1.25	0.21	1.17	φst = 0.99***	6	0	0	0	φst = 1***
Pogonognathellus T4										
Among caves	8	104.05	5.87	92.47	φct = 0.93***	7	40	2.49	97.91	φct = 0.98***
Among pops within caves	11	5.22	-0.02	-0.35	φsc = -0.05	11	0.67	0.53	2.09	φsc = 1
Within populations	2	1	0.5	7.88	φst = 0.92*	2	0	0	0	φst = 1*
(c) Focal OTUs only, by sinkhole area										
Pygmarrhopalites A10										
Among sinkhole areas	3	170.16	19.08	94.16	φct = 0.95*	3	31.25	3.54	96.06	φct = 0.96**
Among pops within areas	3	5.67	0.98	4.82	φsc = 0.82	3	0.75	0.15	3.94	φsc = 1
Within populations	6	1.25	0.21	1.03	φst = 0.99***	6	0	0	0	φst = 1***
Pogonognathellus T4										
Among sinkhole areas	2	36.67	3	43.33	φct = 0.43***	2	16.27	1.42	50.83	φct = 0.51***
Among pops within areas	17	72.6	3.42	49.44	φsc = 0.87	16	24.4	1.38	49.17	φsc = 1
Within populations	2	1	0.5	7.23	φst = 0.93*	2	0	0	0	φst = 1
(d) Focal OTUs only, by region										
Pygmarrhopalites A10										
Among regions	1	99.38	13.59	58.36	φct = 0.58*	1	21.92	3.38	72.7	φct =0.73*
Among pops within regions	5	76.44	9.49	40.75	φsc = 0.98**	5	10.83	1.27	27.3	φsc = 1*
Within populations	6	1.25	0.2	0.89	φst = 0.99***	6	0	0	0	φst = 1***
Pogonognathellus T4										
Among regions	1	7.29	0.38	6.77	φct = 0.07	1	2.83	0.15	6.79	φct =0.07
Among pops within regions	18	101.98	4.7	84.26	φsc = 0.9	17	37.83	2.01	93.21	φsc = 1
Within populations	2	1	0.5	8.97	φst = 0.91*	2	0	0	0	φst = 1

†df, degrees of freedom
‡SS, sum of squares
§VC, variance components
¶V%, percent of variation

Table 1.8. Mantel test results (**a,** COI; **b,** 16S) to identify isolation-by-distance (IBD) patterns and correlations between genetic distance and geographic barriers after controlling for geographic distance in *Pygmarrhopalites* A10 and *Pogonognathellus* T4. Significance is based on 100,000 permutations: *P < 0.05, **P < 0.01, ***P < 0.001. Bold values indicate strongly positive correlations.

	r value	
Barrier	*Pygmarrhopalites* A10	*Pogonognathellus* T4
(a)		
Geographic distance	**0.79*** **	0.37**
Cave boundaries	0.27*	**0.62*** **
Sinkhole area boundaries	**0.63*** **	-0.19*
Mississippi River valley	**0.54*** **	-0.21*
(b)		
Geographic distance	**0.79*** **	0.45**
Cave boundaries	0.18	**0.63*** **
Sinkhole area boundaries	**0.5*** **	0.28**
Mississippi River valley	**0.85*** **	-0.23**

Table 1.9. Bayes factor comparisons of marginal likelihood estimates from stepping-stone sampling analysis for each topological hypothesis (H_0–H_2) to determine support for reciprocal monophyly across the Mississippi Valley (H_0) for *Pygmarrhopalites* A10 and *Pogonognathellus* T4. BF values in bold indicate strong support for H_0.

	H_0[†]	H_1[‡]	H_2[§]	$2 \times \log_e$ BF H_0 vs. H_1	$2 \times \log_e$ BF H_0 vs. H_2
Pygmarrhopalites A10	-16060.28	-16070.25	-16066.01	**19.94**	**11.46**
Pogonognathellus T4	-7349.78	-7297.37	—	-104.82	—

[†]H_0 = IL+MO.
[‡]H_1 (*Pygmarrhopalites*) = (INC+SSC)+(STC+HSC+PAC); H_1 (*Pogonognathellus*) = TMC+(ILC+HSC) & (AC1+AC2)+(PAC+FPC+MJP).
[§]H_2 = INC+((SCC)+(STC+HSC+PAC)).

Figures

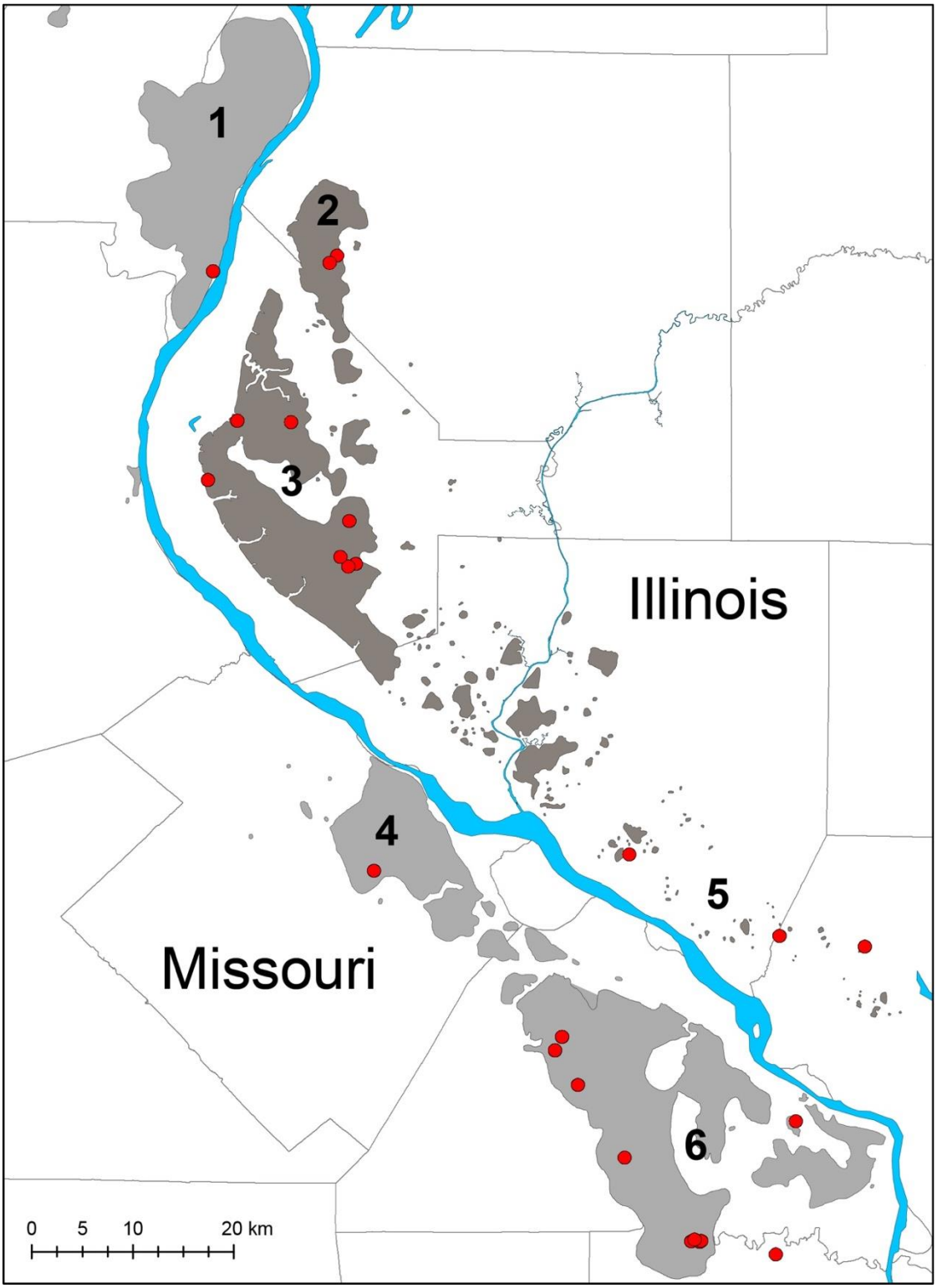


Figure 1.1 Salem Plateau cave-bearing karst spanning the Mississippi River border of Illinois (dark gray) and Missouri (light gray) (adapted from Panno et al., 1997; 1999). Sinkhole karst areas labeled 1–6; caves sampled as red dots.

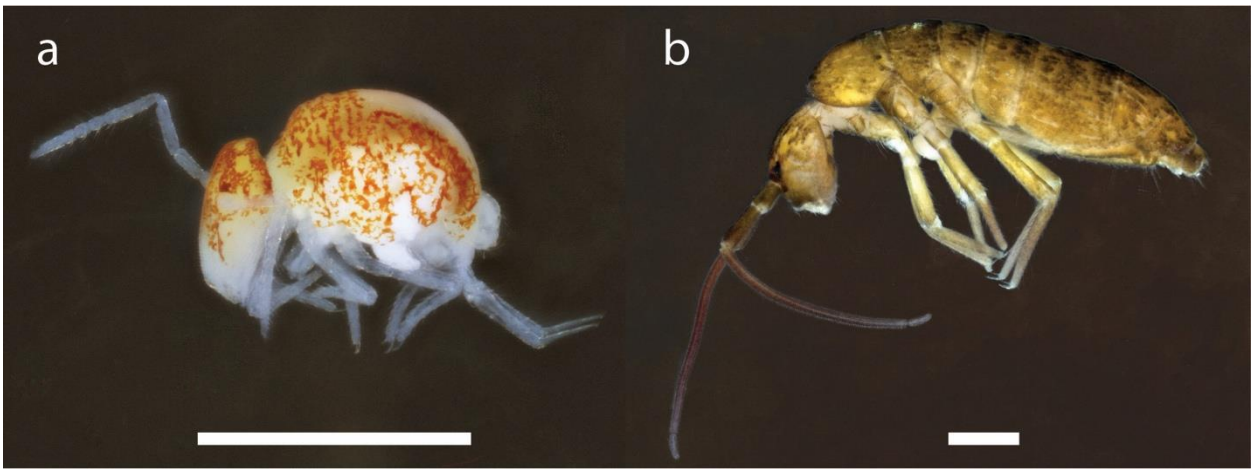


Figure 1.2. Photographs of (**a**) *Pygmarrhopalites* sp. collected from Pautler Cave in Illinois and (**b**) *Pogonognathellus* pale species complex collected from Fogelpole Cave in Illinois. Scale bars = 500 µm.

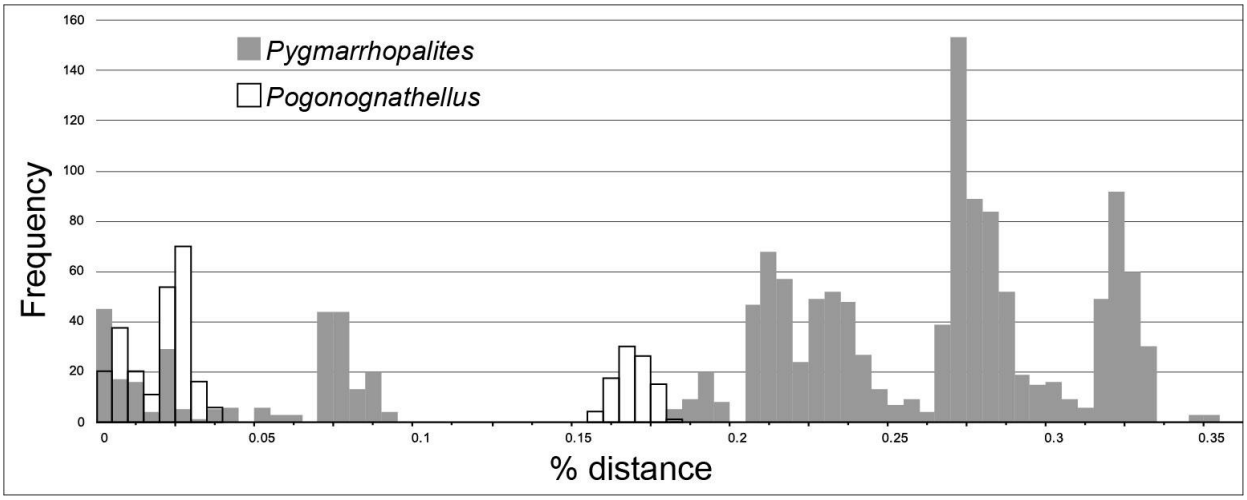


Figure 1.3. COI pairwise distance (uncorrected) frequency histogram for *Pygmarrhopalites* (gray) and *Pogonognathellus* (white).

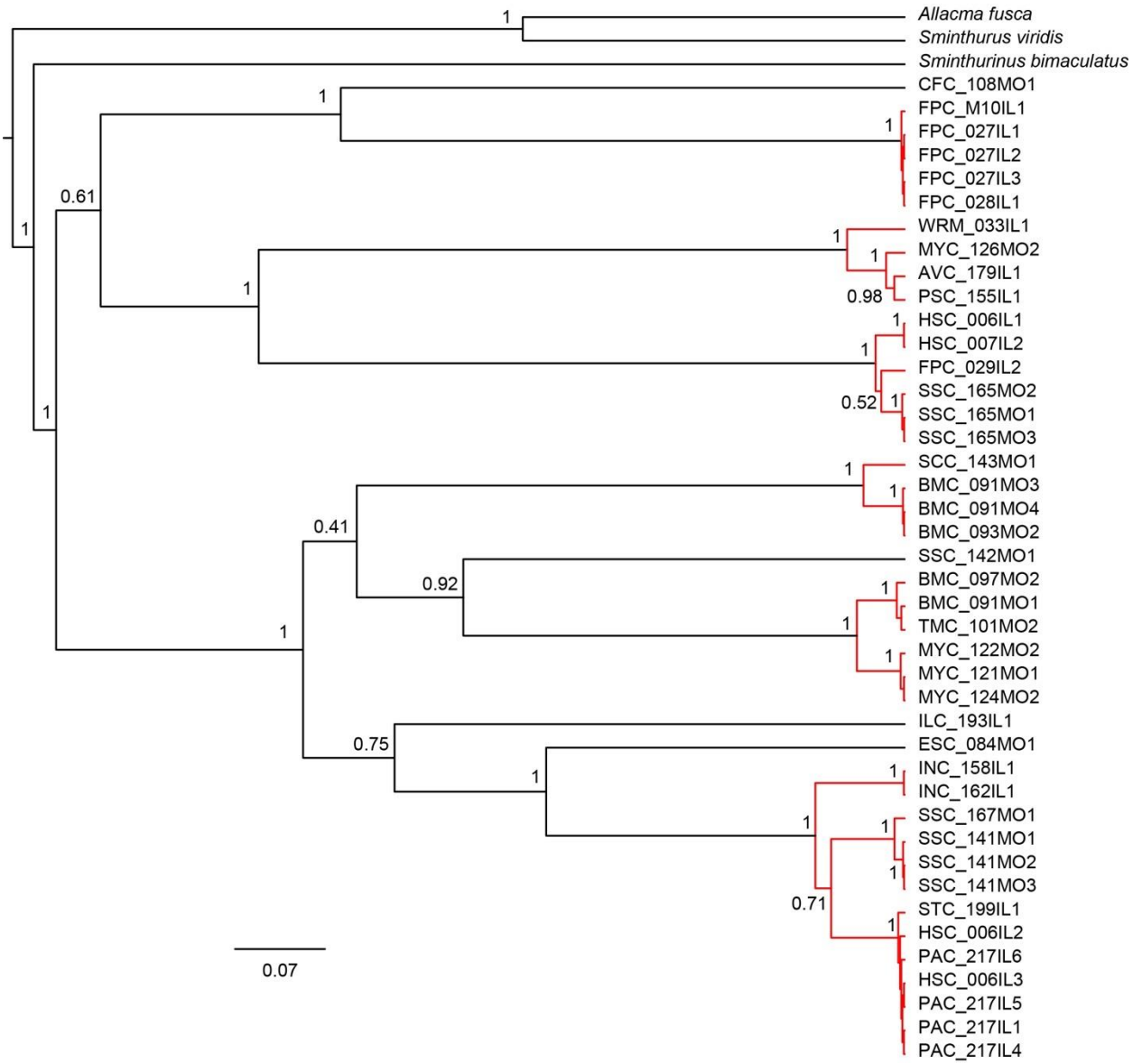


Figure 1.4. Bayesian COI gene tree and GMYC species delimitation results for *Pygmarrhopalites*. Clade posterior probabilities are indicated at each node. Red branches indicate lineages grouped as OTUs by GMYC analysis. Taxon labels correspond to cave name abbreviation, sample #, state, specimen # (See Table 1.1 for cave abbreviations and Table 1.2 for sample information). Scale represents substitutions/site/time.

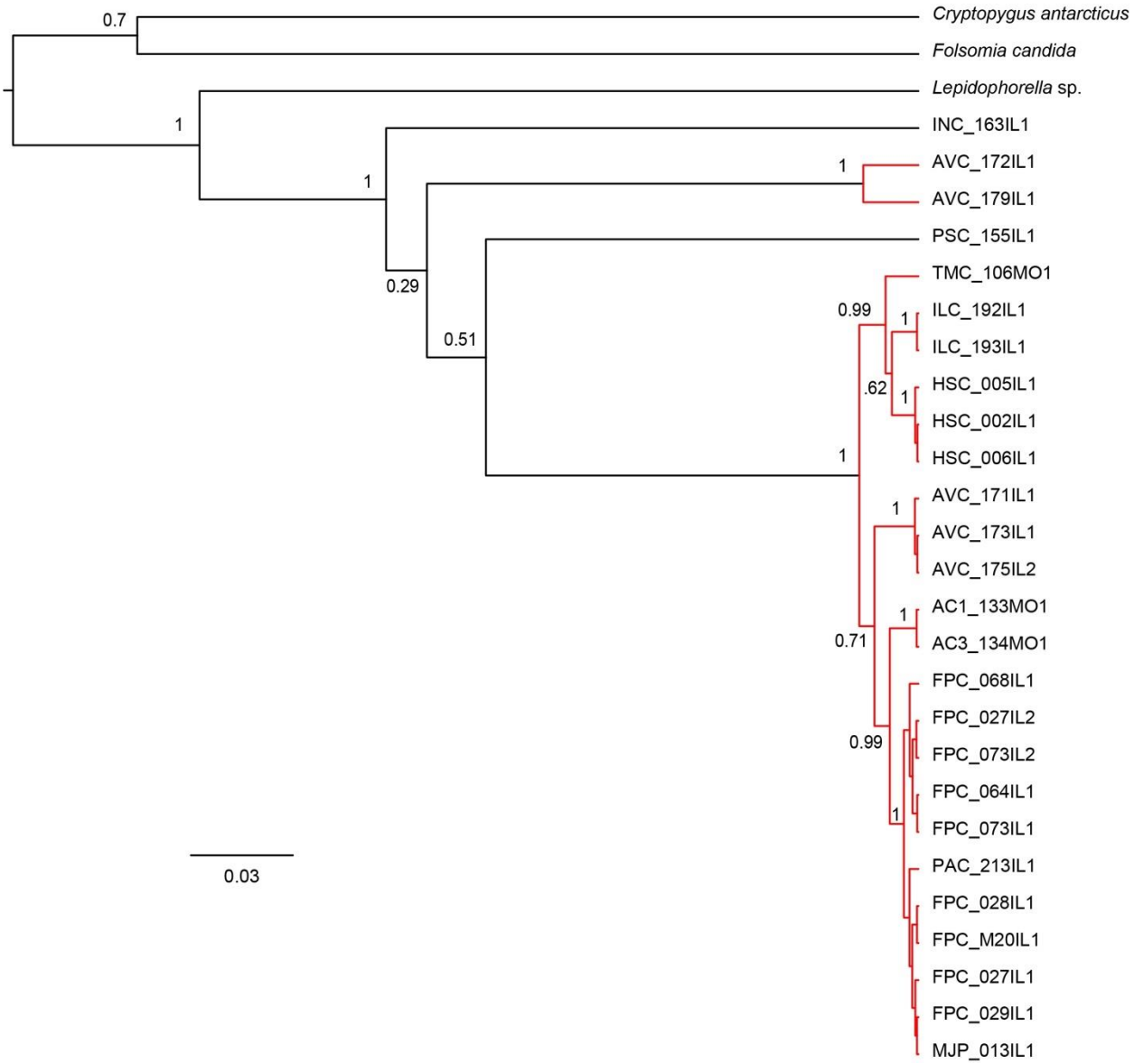


Figure 1.5. Bayesian COI gene tree and GMYC species delimitation results for *Pogonognathellus*. Clade posterior probabilities are indicated at each node. Red branches indicate lineages grouped as OTUs by GMYC analysis. Taxon labels correspond to cave name abbreviation, sample #, state, specimen # (See Table 1.1 for cave abbreviations and Table 1.2 for sample information). Scale represents substitutions/site/time.

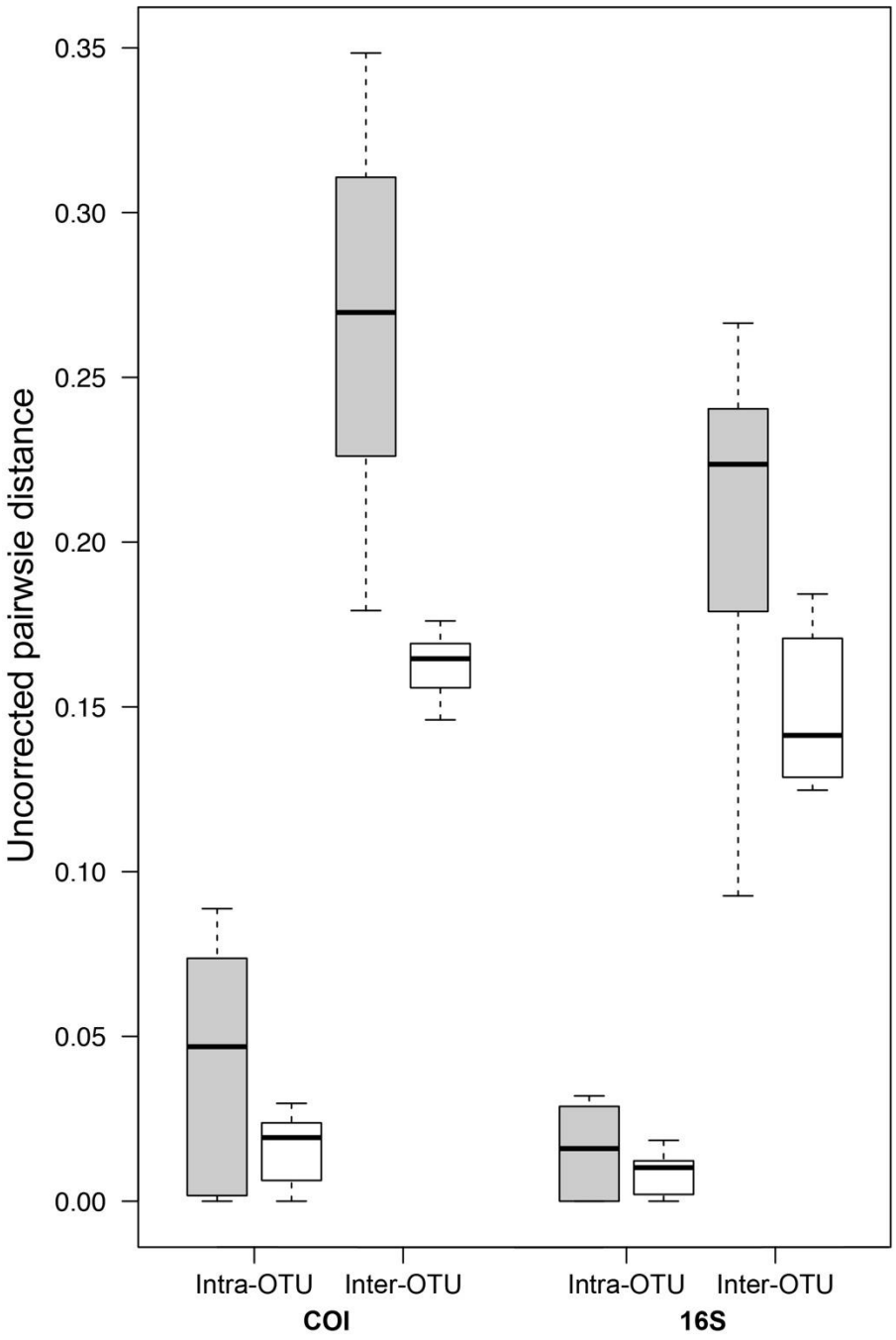


Figure 1.6. Boxplots comparing inter- and intra-OTU genetic distances (uncorrected) for *Pygmarrhopalites* (gray) and *Pogonognathellus* (white).

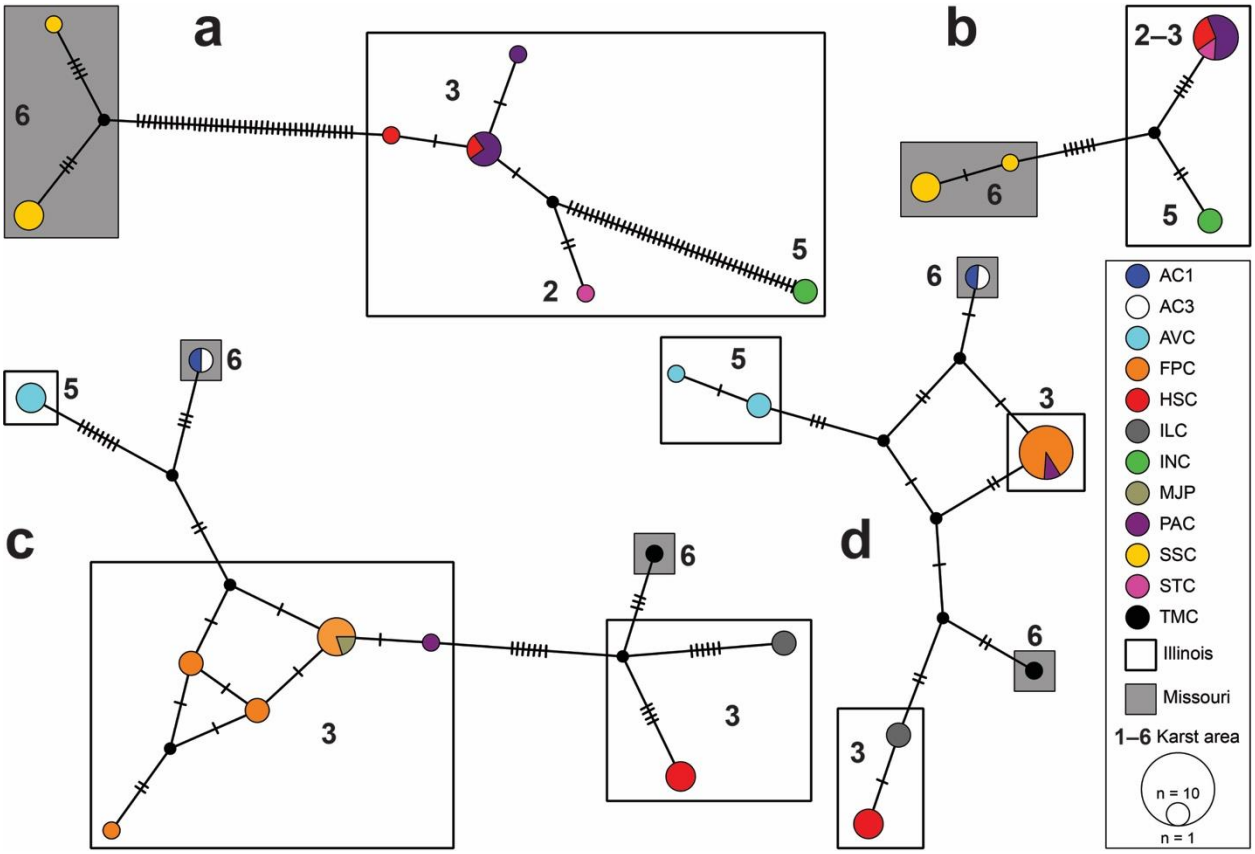


Figure 1.7. TCS haplotype networks of *Pygmarrhopalites* A10 (**a**, COI; **b**, 16S) and *Pogonognathellus* T4 (**c**, COI; **d**, 16S). Hatch marks represent mutational steps between haplotypes; circle color indicates cave locality as illustrated in the legend; numbers 1–6 represent sinkhole karst area (see Table 1.1; Fig. 1.1); rectangle color indicates position relative to the Mississippi River (white, Illinois; gray, Missouri); node diameter represents sample size.

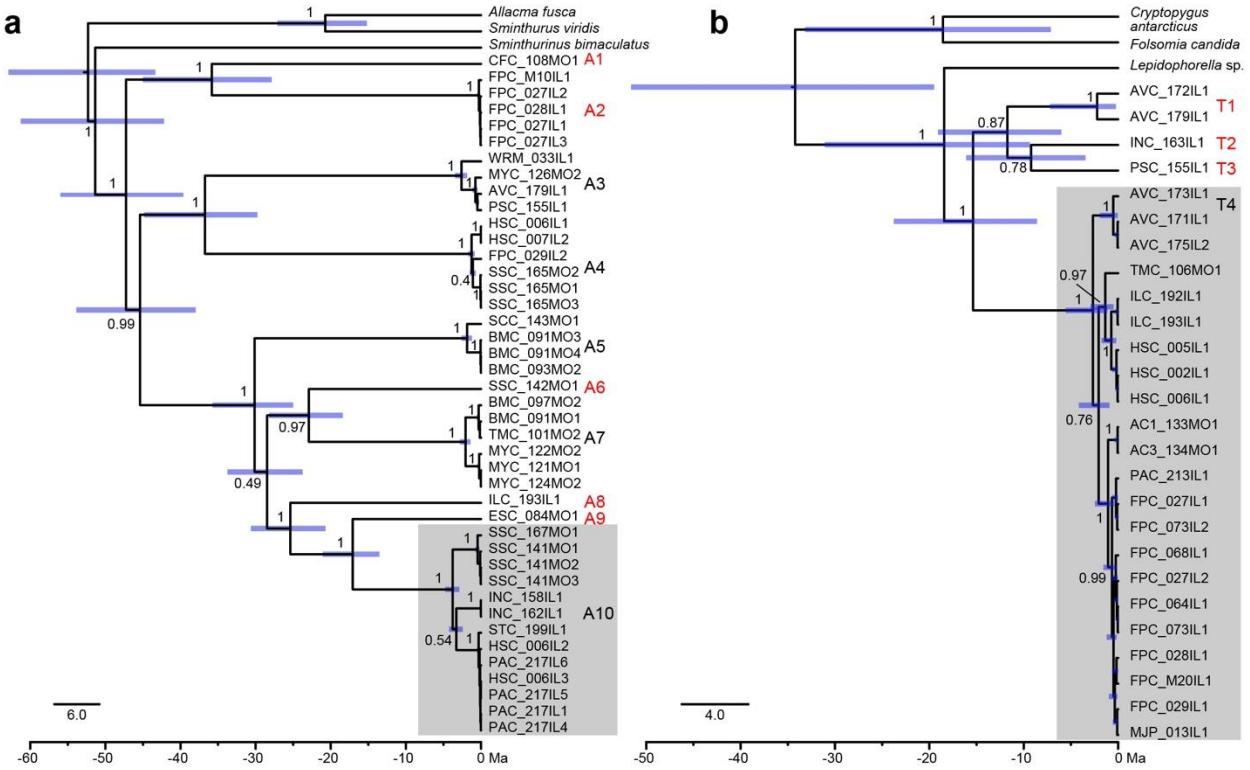


Figure 1.8. Time-calibrated trees for (**a**) *Pygmarrhopalites* and (**b**) *Pogonognathellus* inferred by Bayesian phylogenetic analysis. Clade posterior probabilities are indicated at each node. Divergence times are represented by blue bars at each node with their length corresponding to the 95% HPD of node ages. OTUs identified by the GMYC analysis are indicated to the right of each clade (A1–A10 and T1–T4). Focal OTUs chosen for population structure analyses (A10 and T4) are highlighted in gray boxes (See Figure 1.9 for close-up of A10). Single-site endemic OTUs are labeled in red. Taxon labels correspond to cave name abbreviation, sample #, state, specimen # (See Table 1.1 for cave abbreviations and Table 1.2 for sample information). Scale bars represent substitutions/site/Ma.

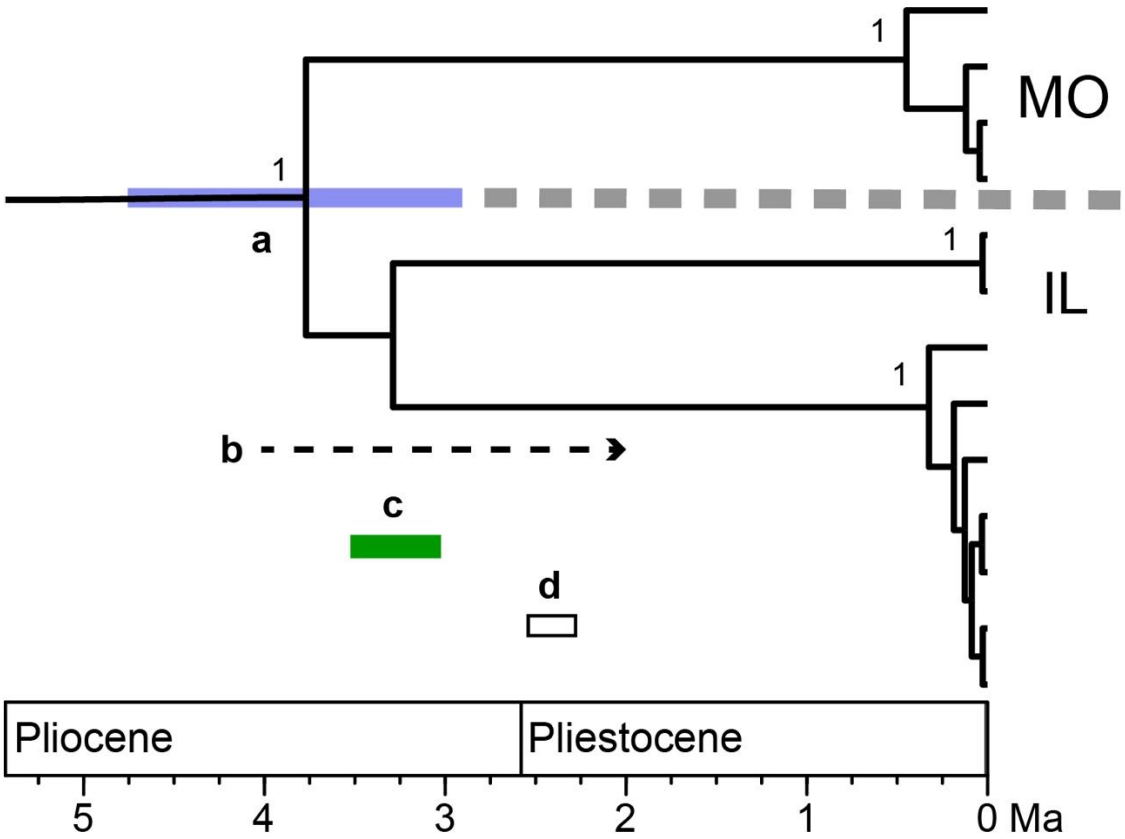


Figure 1.9. Close-up of clade *Pygmarrhopalites* A10 from Figure 1.8a illustrating timing information from estimates of molecular divergence and geological evidence supporting vicariance across the Mississippi River valley: (**a**) 2.90–4.76 Ma (95% HDP) (blue bar) divergence time between Missouri and Illinois lineages (separated by gray dashed line), posterior probabilities at each node lower than 1 are not displayed; (**b**) Late Pliocene/early Pleistocene timing (dashed arrow) of initial Mississippi River entrenchment (Cupples & Van Arsdale, 2014) and increased river discharge (Cox et al., 2014); (**c**) 3.25 ±0.26 Ma (green bar) timing of initial Green River karst incision and excavation (Granger et al., 2001); (**d**) 2.41±0.14 Ma (white bar) timing of first glacial melt (Balco et al., 2005). Holocene (0.01 Ma to present) is not labeled.

References

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press: Cambridge.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, *18*(1), 489–522. doi: 10.1146/annurev.es.18.110187.002421

Balco, G., & Rovey, C. W. (2008). An isochron method for cosmogenic-nuclide dating of buried soils and sediments. *American Journal of Science*, *308*(10), 1083–1114. doi: 10.2475/10.2008.02

Balco, G., Rovey, C. W., & Stone, J. O. H. (2005). The First Glacial Maximum in North America. *Science*, *307*(5707), 222. doi: 10.1126/science.1103406

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. *Nucleic Acids Research*, *41*(D1), D36–D42. doi: 10.1093/nar/gks1195

Bergsten, J., Nilsson, A. N., & Ronquist, F. (2013). Bayesian tests of topology hypotheses with an example from diving beetles. *Systematic Biology*, *62*(5), 660–673. doi: 10.1093/sysbio/syt029

Blackith, R. E., & Disney, R. H. L. (1988). Passive aerial dispersal during moulting in tropical Collembola. *Malayan Nature Journal*, *41*(4), 529–531.

Bonnet, E., & Van de Peer, Y. (2002). zt: a software tool for simple and partial Mantel tests. *Journal of Statistical Software*, *7*(10), 1–12. doi: 10.18637/jss.v007.i10

Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, *17*, 42. doi: 10.1186/s12862-017-0890-6

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, R., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. doi: 10.1371/journal.pcbi.1003537

Bretschko, G., & Christian, E. (1989). Collembola in the Bed Sediments of an Alpine Gravel Stream (RITRODAT-Lunz Study Area, Austria). *Internationale Revue Der Gesamten Hydrobiologie Und Hydrographie*, *74*(5), 491–498.

Brower, A. V. Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences*, *91*(14), 6491–6495. doi: 10.1073/pnas.91.14.6491

Burr, B. M., Adams, G. L., Krejca, J. K., Paul, R. J., & Warren, M. L. (2001). Troglomorphic sculpins of the *Cottus carolinae* species group in Perry County, Missouri: distribution, external morphology, and conservation status. *Environmental Biology of Fishes*, *62*(1–3), 279–296. doi: 10.1023/A:1011819922403

Christiansen, K. (1961). Convergence and Parallelism in Cave Entomobryinae. *Evolution*, *15*(3), 288–301. doi: 10.1111/j.1558-5646.1961.tb03156.x

Christiansen, K. (1962). Proposition pour la classification des animaux cavernicoles. *Spelunca Mémoires*, *2*, 75–78.

Christiansen, K. (1965). Behavior and form in the Evolution of Cave Collembola. *Evolution*, *19*(4), 529–537. doi: 10.1111/j.1558-5646.1965.tb03328.x

Christiansen, K., & Bellinger, P. (1996). Cave *Arrhopalites*: New to science. *Journal of Cave and Karst Studies*, *58*(3), 168–180.

Christiansen, K. & Bellinger, P. (1998). *The Collembola of North America north of the Rio Grande; A taxonomic analysis*. 2nd Edition. Grinnell College: Grinnell.

Christiansen, K., & Culver, D. (1968). Geographical Variation and Evolution in *Pseudosinella hirsuta*. *Evolution*, *22*(2), 237–255. doi: 10.1111/j.1558-5646.1968.tb05891.x

Christiansen, K., & Culver, D. (1969). Geographical Variation and Evolution in *Pseudosinella violenta* (Folsom). *Evolution*, *23*(4), 602–621. doi: 10.1111/j.1558-5646.1969.tb03544.x

Christiansen, K., & Culver, D. (1987). Biogeography and the Distribution of Cave Collembola. *Journal of Biogeography*, *14*(5), 459–477. doi: 10.2307/2844976

Christman, M. C., Culver, D. C., Madden, M. K., & White, D. (2005). Patterns of endemism of the eastern North American cave fauna. *Journal of Biogeography*, *32*(8), 1441–1452. doi: 10.1111/j.1365-2699.2005.01263.x

Cicconardi, F., Nardi, F., Emerson, B. C., Frati, F., & Fanciulli, P. P. (2010). Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology*, *19*(2), 386–400. doi: 10.1111/j.1365-294X.2009.04457.x

Clement, M., Snell, Q., & Walker, P. (2002). TCS: estimating gene genealogies. *Parallel and Distributed Processing Symposium, International Proceedings*, *2*, 184.

Colgan, D. J., McLauchlan, A., Wilson, G. D. F., Livingston, S. P., Edgecombe, G. D., Macaranas, J., Cassis, G., & Gray, M.R. (1998). Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology*, *46*(5), 419–437. doi: 10.1071/ZO98048

Costa, D., Timmermans, M. J. T. N., Sousa, J. P., Ribeiro, R., Roelofs, D., & Van Straalen, N. M. (2013). Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. *Applied Soil Ecology*, *68*, 61–66. doi: 10.1016/j.apsoil.2013.03.003

Coulson, S. J., Hodkinson, I. D., & Webb, N. R. (2003). Aerial dispersal of invertebrates over a high-Arctic glacier foreland: Midtre Lovénbreen, Svalbard. *Polar Biology*, *26*(8), 530–537. doi: 10.1007/s00300-003-0516-x

Coulson, S. J., Hodkinson, I. D., Webb, N. R., & Harrison, J. A. (2002). Survival of terrestrial soil-dwelling arthropods on and in seawater: implications for trans-oceanic dispersal. *Functional Ecology*, *16*(3), 353–356. doi: 10.1046/j.1365-2435.2002.00636.x

Cox, R. T., Lumsden, D. N., & Van Arsdale, R. B. (2014). Possible Relict Meanders of the Pliocene Mississippi River and Their Implications. *The Journal of Geology*, *122*(5), 609–622. doi: 10.1086/676974

Culver, D. C., & Pipan, T. (2009). *The Biology of Caves and Other Subterranean Habitats*. Oxford University Press: New York.

Culver, D. C., Pipan, T., & Schneider, K. (2009). Vicariance, dispersal and scale in the aquatic subterranean fauna of karst regions. *Freshwater Biology*, *54*(4), 918–929. doi: 10.1111/j.1365-2427.2007.01856.x

Cupples, W., & Van Arsdale, R. (2014). The Preglacial "Pliocene" Mississippi River. *The Journal of Geology*, *122*(1), 1–15. doi: 10.1086/674073

Deharveng, L., D'Haese, C. A., & Bedos, A. (2008). Global diversity of springtails (Collembola; Hexapoda) in freshwater. *Hydrobiologia*, *595*(1), 329–338. doi: 10.1007/s10750-007-9116-z

Delić, T., Trontelj, P., Rendoš, M., & Fišer, C. (2017). The importance of naming cryptic species and the conservation of endemic subterranean amphipods. *Scientific Reports*, *7*(1), 3391. doi: 10.1038/s41598-017-02938-z

Dörge, D. D., Zaenker, S., Klussmann-Kolb, A., & Weigand, A. M. (2014). Traversing worlds - dispersal potential and ecological classification of *speolepta leptogaster* (Winnertz, 1863) (Diptera, Mycetophilidae). *Subterranean Biology*, *13*, 1–16. doi: 10.3897/subtbiol.13.6460

Drummond, A. J., & Bouckaert, R. R. (2015). *Bayesian evolutionary analysis with BEAST*. Cambridge University Press: Cambridge.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Ezard, T., Fujisawa, T., & Barraclough, T. G. (2009). Package 'splits: SPecies' LImits by Threshold Statistics' version 1.0-19. Available from http://R-Forge.R-project.org/projects/splits/

Faille, A., Tänzler, R., & Toussaint, E. F. A. (2015). On the Way to Speciation: Shedding Light on the Karstic Phylogeography of the Microendemic Cave Beetle *Aphaenops cerberus* in the Pyrenees. *Journal of Heredity*, *106*(6), 692–699. doi: 10.1093/jhered/esv078

Felderhoff, K. L., Bernard, E. C., & Moulton, J. K. (2010). Survey of *Pogonognathellus* Börner (Collembola: Tomoceridae) in the Southern Appalachians Based on Morphological and Molecular Data. *Annals of the Entomological Society of America*, *103*(4), 472–491. doi: 10.1603/AN09105

Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, *27*(3), 613–635. doi: 10.1111/mec.14486

Fong, D. W., Culver, D. C., Hobbs III, H. H., & Pipan, T. (2007). *The Invertebrate Cave Fauna of West Virginia*. *West Virginia Speleological Survey Bulletin 16* (2nd ed.). West Virginia Speleological Survey (WVASS): Barrackville, West Virginia.

Fouquet, A., Vences, M., Salducci, M. D., Meyer, A., Marty, C., Blanc, M., & Gilles, A. (2007). Revealing cryptic diversity using molecular phylogenetics and phylogeography in frogs of the *Scinax ruber* and *Rhinella margaritifera* species groups. *Molecular Phylogenetics and Evolution*, *43*(2), 567–582. doi: 10.1016/j.ympev.2006.12.006

Freeman, J. A. (1952). Occurrence of Collembola in the air. *Proceedings of the Royal Entomological Society of London. Series A, General Entomology*, *27*(1–3), 28. doi: 10.1111/j.1365-3032.1952.tb00142.x

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, *13*(5), 851–861. doi: 10.1111/1755-0998.12138

Giribet, G., Edgecombe, G. D., & Wheeler, W. C. (2001). Arthropod phylogeny based on eight molecular loci and morphology. *Nature*, *413*(6852), 157–161. doi: 10.1038/35093097

Goldberg, J., & Trewick, S. A. (2011). Exploring phylogeographic congruence in a continental Island system. *Insects*, *2*(3), 369–399. doi: 10.3390/insects2030369

Gómez, R. A., Reddell, J., Will, K., & Moore, W. (2016). Up high and down low: Molecular systematics and insight into the diversification of the ground beetle genus *Rhadine* LeConte. *Molecular Phylogenetics and Evolution*, *98*, 161–175. doi: 10.1016/j.ympev.2016.01.018

Granger, D. E., Fabel, D., & Palmer, A. N. (2001). Pliocene−Pleistocene incision of the Green River, Kentucky, determined from radioactive decay of cosmogenic 26Al and 10Be in Mammoth Cave sediments. *Geological Society of America Bulletin*, *113*(7), 825–836. doi: 10.1130/0016-7606(2001)113<0825:PPIOTG>2.0.CO;2

Hawes, T. C., Worland, M. R., Bale, J. S., & Convey, P. (2008). Rafting in Antarctic Collembola. *Journal of Zoology*, *274*(1), 44–50. doi: 10.1111/j.1469-7998.2007.00355.x

Hawes, T. C., Worland, M. R., Convey, P., & Bale, J. S. (2007). Aerial dispersal of springtails on the Antarctic Peninsula: implications for local distribution and demography. *Antarctic Science*, *19*(1), 3–10. doi: 10.1017/S0954102007000028

Hodges, K. M., Rowell, D. M., & Keogh, J. S. (2007). Remarkably different phylogeographic structure in two closely related lizard species in a zone of sympatry in south-eastern Australia. *Journal of Zoology*, *272*(1), 64–72. doi: 10.1111/j.1469-7998.2006.00244.x

Hogg, I. D., & Hebert, P. D. N. (2004). Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Canadian Journal of Zoology*, *82*(5), 749–754. doi: 10.1139/z04-041

Hopkin, S. P. (1997). *Biology of springtails (Insecta: Collembola)*. Oxford University Press: New York.

Hurtado, L. A., Lee, E. J., & Mateos, M. (2013). Contrasting phylogeography of sandy vs. rocky supralittoral isopods in the megadiverse and geologically dynamic Gulf of California and adjacent areas. *PLoS ONE*, *8*(7), e67827. doi: 10.1371/journal.pone.0067827

Juan, C., & Emerson, B. C. (2010). Evolution underground: Shedding light on the diversification of subterranean insects. *Journal of Biology*, *9*(3), 1–5. doi: 10.1186/jbiol227

Juan, C., Guzik, M. T., Jaume, D., & Cooper, S. J. B. (2010). Evolution in caves: Darwin's "wrecks of ancient life" in the molecular era. *Molecular Ecology*, *19*(18), 3865–3880. doi: 10.1111/j.1365-294X.2010.04759.x

Kass, R. E., & Raftery, A. E. (1994). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

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

Katz, A. D., Giordano, R., & Soto-Adames, F. N. (2015). Operational criteria for cryptic species delimitation when evidence is limited, as exemplified by North American *Entomobrya* (Collembola: Entomobryidae). *Zoological Journal of the Linnean Society*, *173*(4), 810–840. doi: 10.1111/zoj.12220

Katz, A. D., Taylor, S. J., Soto-Adames, F. N., Addison, A., Hoese, G. B., Sutton, M. R., & Toulkeridis, T. (2016). New records and new species of springtails (Collembola: Entomobryidae, Paronellidae) from lava tubes of the Galápagos Islands (Ecuador). *Subterranean Biology*, *17*, 77–120. doi: 10.3897/subtbiol.17.7660

Lapointe, F., & Rissler, L. J. (2005). Congruence, Consensus, and the Comparative Phylogeography of Codistributed Species in California. *The American Naturalist*, *166*(2), 290–299. doi: 10.1086/431283

Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. doi: 10.1111/2041-210X.12410

Lewis, J. J. (2005). *Bioinventory of Caves of the Cumberland Escarpment Area of Tennessee*. Final Report to Tennessee Wildlife Resources Agency & The Nature Conservancy of Tennessee: Lewis & Associates, 158 p.

Lewis, J. J., Moss, P., Tecic, D., & Nelson, M. E. (2003). A conservation focused inventory of subterranean invertebrates of the Southwestern Illinois Karst. *Journal of Cave and Karst Studies*, *65*(1), 9–21.

Mankowski, A. (2010). *Endangered and Threatened Species of Illinois: Status and Distribution, Volume 4 - 2009 and 2010 Changes to the Illinois List of Endangered and Threatened Species*. Illinois Endangered Species Protection Board: Springfield, Illinois.

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, *27*(2), 209–220. doi: 10.1002/stem.684.

Mari-Mutt, J. A. (1979). A Revision of the Genus *Dicranocentrus* Schött (Insecta: Collembola: Entomobryidae). *Bulletin of the University of Puerto Rico*, *259*, 1–79.

Meier, R., Zhang, G., & Ali, F. (2008). The Use of Mean Instead of Smallest Interspecific Distances Exaggerates the Size of the "Barcoding Gap" and Leads to Misidentification. *Systematic Biology*, *57*(5), 809–813. doi: 10.1080/10635150802406343

Moore, J. C., Saunders, P., Selby, G., Horton, H., Chelius, M. K., Chapman, A., & Horrocks, R. D. (2005). The distribution and life history of *Arrhopalites caecus* (Tullberg): Order: Collembola, in Wind Cave, South Dakota, USA. *Journal of Cave and Karst Studies*, *67*(2), 110–119.

Niemiller, M. L., Fitzpatrick, B. M., & Miller, B. T. (2008). Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Molecular Ecology, 17*(9), 2258–2275. doi: 10.1111/j.1365-294X.2008.03750.x

Niemiller, M. L., Graening, G. O., Fenolio, D. B., Godwin, J. C., Cooley, J. R., Pearson, W. D., Fitzpatrick, B. M., & Near, T. J. (2013a). Doomed before they are described? The need for conservation assessments of cryptic species complexes using an amblyopsid cavefish (Amblyopsidae: *Typhlichthys*) as a case study. *Biodiversity and Conservation*, *22*(8), 1799–1820. doi: 10.1007/s10531-013-0514-4

Niemiller, M. L., McCandless, J. R., Reynolds, R. G., Caddle, J., Near, T. J., Tillquist, C. R., Pearson, W. D., & Fitzpatrick, B. M. (2013b). Effects of climatic and geological processes during the pleistocene on the evolutionary history of the northern cavefish, *Amblyopsis spelaea* (Teleostei: Amblyopsidae). *Evolution*, *67*(4), 1011–1025. doi: 10.1111/evo.12017

Niemiller, M. L., Near, T. J., & Fitzpatrick, B. M. (2012). Delimiting species using multilocus data: Diagnosing cryptic diversity in the southern cavefish, *typhlichthys subterraneus* (teleostei: Amblyopsidae). *Evolution*, *66*(3), 846–866. doi: 10.1111/j.1558-5646.2011.01480.x

Niemiller, M. L., & Zigler, K. S. (2013). Patterns of Cave Biodiversity and Endemism in the Appalachians and Interior Plateau of Tennessee, USA. *PLoS ONE*, *8*(5), e64177. doi: 10.1371/journal.pone.0064177

Ortiz, M., Legatzki, A., Neilson, J. W., Fryslie, B., Nelson, W. M., Wing, R. A., Soderlund, C. A., Pryor, B. M., Maier, R. M. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *The ISME Journal*, *8*(2), 478–491. doi: 10.1038/ismej.2013.159

Paclt, J. 1944. Nomina nova in Collembola. *Entomologické Listy*, **7**, 92.

Palacios-Vargas, J. G., Cortés-Guzmán, D., & Alcocer, J. (2018). Springtails (Collembola, Hexapoda) from Montebello Lakes, Chiapas, Mexico. *Inland Waters*. doi: 10.1080/20442041.2018.1439863

Panno, S. V., Krapac, I. G., Weibel, C. P., & Bade, J. D. (1996). Groundwater contamination in karst terrain of Southwestern Illinois. *Illinois State Geological Survey, Environmental Geology, 151*, 1-43.

Panno, S. V., Weibel, C. P., & Li, W. (1997). *Karst regions of Illinois*. Open File Series *1997-2*. Illinois State Geological Survey: Champaign.

Panno, S. V., Weibel, C. P., Wicks, C. M., & Vandike, J. E. (1999). Geology, Hydrogeology, and Water Quality of the Karst Regions of Southwestern Illinois and Southeastern Missouri. *Geological Field Trip Guidebook No. 2 for the 33rd Meeting of the North–Central Geological Society of America, Champaign-Urbana, IL, April 22–23, 1999. Illinois State Geological Survey, ISGS Guidebook 27*, 1–38.

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: The mid-aegean trench calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. doi: 10.1093/molbev/msq051

Park, K.-H., Bernard, E. C., & Moulton, J. K. (2011). Three new species of *Pogonognathellus* (Collembola: Tomoceridae) from North America. *Zootaxa*, *3070*, 1–14.

Peck, S. B., & Lewis, J. J. (1978). Zoogeography and evolution of the subterranean invertebrate faunas of Illinois and southeastern Missouri. *The NSS Bulletin*, *40*(2), 39–63.

Pérez-Moreno, J. L., Balázs, G., Wilkins, B., Herczeg, G., & Bracken-Grissom, H. D. (2017). The role of isolation on contrasting phylogeographic patterns in two cave crustaceans. *BMC Evolutionary Biology*, *17*, 247. doi: 10.1186/s12862-017-1094-9

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609. doi: 10.1080/10635150600852011

Porter, M. L. (2007). Subterranean biogeography: What have we learned from molecular techniques? *Journal of Cave and Karst Studies*, *69*(1), 179–186.

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Tracer v1.7, Available from http://tree.bio.ed.ac.uk/software/tracer/

Rastorgueff, P.-A., Chevaldonné, P., Arslan, D., Verna, C., & Lejeusne, C. (2014). Cryptic habitats and cryptic diversity: unexpected patterns of connectivity and phylogeographical breaks in a Mediterranean endemic marine cave mysid. *Molecular Ecology*, *23*(11), 2825–2843. doi: 10.1111/mec.12776

Schonhöfer, A. L., Vernesi, C., Martens, J., & Marshal, H. (2015). Molecular phylogeny, biogeographic history, and evolution of cave-dwelling taxa in the European harvestman genus *Ischyropsalis* (Opiliones: Dyspnoi). *The Journal of Arachnology*, *43*(1), 40–53. doi: 10.1636/H14-39.1

Sela, I., Ashkenazy, H., Katoh, K., & Pupko, T. (2015). GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Research, 43*(W1), W7–W14. doi: 10.1093/nar/gkv318

Shaw, P., Dunscombe, M., & Robertson, A. (2011). Collembola in the hyporheos of a karstic river: an overlooked habitat for Collembola containing a new genus for the UK. *Soil Organisms, 83*(3), 507–514.

Shofner, G. A., Mills, H. H., & Duke, J. E. (2001). A simple map index of karstification and its relationship to sinkhole and cave distribution in Tennessee. *Journal of Cave and Karst Studies, 63*(2), 67–75.

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology, 15*(14), 4261–4293. doi: 10.1111/j.1365-294X.2006.03061.x

Soto-Adames, F. N., & Taylor, S. J. (2010). Status assessment survey for springtails (Collembola) in Illinois caves: The Salem Plateau. *INHS Technical Report, 13*, 1–76.

Soto-Adames, F. N., & Taylor, S. J. (2013). New Species and New Records of Springtails (Hexapoda: Collembola) from Caves in the Salem Plateau of Illinois, USA. *Journal of Cave and Karst Studies, 75*(2), 146–175. doi: 10.4311/2011LSC0257

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

Taylor, S. J., & Niemiller, M. L. (2016). Biogeography and conservation assessment of *Bactrurus* groundwater amphipods (Crangonyctidae) in the central and eastern United States. *Subterranean Biology, 17*, 1–29. doi: 10.3897/subtbiol.17.7298

Thibaud, J., & Deharveng, L. (1994). Collembola. *Encyclopaedia Biospeologica, tome I* (ed. by C. Juberthie and V. Decu), pp. 267–276. Société Internationale del Biospéologie, Moulis: France.

Timmermans, M. J. T. N., Ellers, J., Mariën, J., Verhoef, S. C., Ferwerda, E. B., & van Straalen, N. M. (2005). Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular Ecology, 14*(7), 2017–2024. doi: 10.1111/j.1365-294X.2005.02548.x

Trontelj, P., Douady, C. J., Fišer, C., Gibert, J., Gorički, Š., Lefébure, T., Sket, B., & Zakšek, V. (2009). A molecular test for cryptic diversity in ground water: How large are the ranges of macro-stygobionts? *Freshwater Biology, 54*(4), 727–744. doi: 10.1111/j.1365-2427.2007.01877.x

Tullberg, T. (1871). Förteckning öfver Svenska Podurider. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar, 28*(1), 143–155.

Vargovitsh, R. S. (2009). New Cave Arrhopalitidae (Collembola: Symphypleona) from the Crimea (Ukraine). *Zootaxa*, *2047*, 1–47.

Venarsky, M. P., Anderson, F. E., & Wilhelm, F. M. (2009). Population genetic study of the U.S. federally listed Illinois cave amphipod, *Gammarus acherondytes*. *Conservation Genetics*, *10*(4), 915–921. doi:10.1007/s10592-008-9579-0

Ward, J. V., & Palmer, M. A. (1994). Distribution patterns of interstitial freshwater meiofauna over a range of spatial scales, with emphasis on alluvial river-aquifer systems. *Hydrobiologia*, *287*(1), 147–156. doi: 10.1007/BF00006903

Weckstein, J. D., Johnson, K. P., Murdoch, J. D., Krejca, J. K., Takiya, D. M., Veni, G., Reddell, J. R., & Taylor, S. J. (2016). Comparative phylogeography of two codistributed subgenera of cave crickets (Orthoptera: Rhaphidophoridae: *Ceuthophilus* spp.). *Journal of Biogeography*, *43*(7), 1450–1463. doi: 10.1111/jbi.12734

White, W. B., & Culver, D. C. (2012). *Encyclopedia of Caves*. Academic Press: Chennai.

Whiting, M. F. (2002). Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta*, *31*(1), 93–104. doi: 10.1046/j.0300-3256.2001.00095.x

Xiong, Y., Gao, Y., Yin, W., & Luan, Y. (2008). Molecular phylogeny of Collembola inferred from ribosomal RNA genes. *Molecular Phylogenetics and Evolution*, *49*(3), 728–735. doi: 10.1016/j.ympev.2008.09.007

Zagmajster, M., Culver, D. C., & Sket, B. (2008). Species richness patterns of obligate subterranean beetles (Insecta: Coleoptera) in a global biodiversity hotspot - Effect of scale and sampling intensity. *Diversity and Distributions*, *14*(1), 95–105. doi: 10.1111/j.1472-4642.2007.00423.x

Zeppelini, D., & Christiansen, K. (2003). *Arrhopalites* (Collembola: Arrhopalitidae) in U.S. caves with the description of seven new species. *Journal of Cave and Karst Studies*, *65*(1), 36–42.

Zeppelini, D., Taylor, S. J., & Slay, M. E. (2009). Cave *Pygmarrhopalites* Vargovitsh, 2009 (Collembola, Symphypleona, Arrhopalitidae) in United States. *Zootaxa*, *2204*, 1–18.

Zhang, F., Chen, Z., Dong, R.-R., Deharveng, L., Stevens, M. I., Huang, Y.-H., & Zhu, C.-D. (2014). Molecular phylogeny reveals independent origins of body scales in Entomobryidae (Hexapoda: Collembola). *Molecular Phylogenetics and Evolution*, *70*, 231–239. doi: 10.1016/j.ympev.2013.09.024

CHAPTER 2

Phylogeography of marine littoral springtails (Collembola): Cryptic diversification, pre-Pliocene vicariance, and post-closure dispersal across the Isthmus of Panama

Abstract

The intertidal zone and its geminate species provide a model system for exploring how and when paleogeographic processes leading to Isthmus closure have shaped the evolution of Panama's biological communities. Within a Panamanian context, geminate species are reciprocally monophyletic transisthmian sister taxa, initially isolated by paleogeographic events associated with the formation of the Isthmus of Panama. I tested the geminate species concept for marine littoral-obligate springtails—a highly abundant, diverse, and ecologically restricted group of tiny, wingless, insect-like arthropods. To investigate the influence of Panama's formation on marine littoral springtail communities, I developed a multi-locus dataset for species collected along the Pacific and Atlantic coasts of Panama to identify transisthmian sister taxa, determine processes driving and maintaining their isolation, and to evaluate the timing of their origin. Phylogeographic analysis recovered multiple deeply divergent, morphologically cryptic, and reciprocally monophyletic 'Atlantic' and 'Pacific' sister lineages for three codistributed morphospecies (*Seira nicoya*, *Archisotoma* sp. A, and *Axelsonia tubifera*). The putative geminate species identified within each morphospecies were estimated to have originated prior to the Pliocene, corroborating recent studies that propose the existence of early Miocene land connection between Central and South America. However, the lack of genetic structure and low molecular divergence observed for transisthmian lineages of *Psammisotoma dispar* appear to result from gene flow or range expansion across the Isthmus within the last 2 Ma, providing evidence of post-closure dispersal during the Pleistocene. This study suggests that events

associated with Isthmus formation have had a significant impact on the evolution of Panama's marine littoral springtail fauna, and the Isthmus has been a major genetic barrier for most, but not all, marine littoral springtail species.

Introduction

The rise of the Isthmus of Panama, and its profound impact on marine and terrestrial biodiversity, is widely considered the most well-studied paleogeographic event in scientific history (Lessios, 2008). The Panamanian Isthmus is thought to have gradually formed over 30 million years (Ma), closing completely by the late Pleistocene, approximately 2.8 Ma (O'Dea et al., 2016). The resulting land bridge joining North and South America not only formed a terrestrial connection between continents leading to the Great American Biotic Interchange (Stehli & Webb, 1985), but also closed the only interoceanic migratory channel between the Caribbean Sea and eastern Pacific Ocean, creating a formidable barrier to genetic exchange between marine organisms remaining on either side of the Isthmus (Lessios, 2008). Ecological changes that followed the closure are linked to mass extinctions and the diversification, and subsequent speciation, of many, now separated marine lineages (Leigh et al., 2014).

Numerous transisthmian sister species have been identified as descendants of ancestral populations that were initially isolated by Panama's closure. These 'geminate species', originally recognized and named by Jordan (1908), have emerged as powerful tools for testing evolutionary hypotheses of vicariance (Collins, 1996), and routinely used in molecular clock calibrations, based on the timing of Panama's closure, for estimating molecular rates and divergence times (Lessios, 2008). However, recent biological evidence of temporary Miocene land connections appearing as early as 24 Ma (e.g., Bacon et al., 2013; Elmer et al., 2013; Bacon et al., 2015a;

Thacker, 2017; Winston et al., 2017; Stange et al., 2018), and geological evidence of tectonic collision beginning 25–23 Ma with final closure around 13–15 Ma (e.g., Montes et al., 2012a; 2012b; 2015), have challenged the conventional and largely accepted view of a more recent collision 14–12 Ma (Coates et al., 2004) and final Pliocene closure 2.8–4 Ma (Coates & Obando, 1996; Jackson & O'Dea, 2013; O'Dea et al., 2016). This has ignited a fierce debate concerning the geological chronology of events leading to Panama's formation, including the precise timing of its final closure (Bacon et al., 2015b; Lessios, 2015; Marko et al., 2015; O'Dea et al., 2016; Jaramillo et al., 2017; Molnar, 2017). The emergence of two models of Isthmus closure, each supported (and apparently contradicted) by substantial evidence (Hoorn & Flantua, 2015), has reinforced earlier cautions against the use of its unresolved timing for molecular clock calibrations (e.g., Marko, 2002; Kodandaramaiah, 2011), but has also provided new opportunities to assess biogeographical data in light of these competing hypotheses.

Organisms inhabiting the marine intertidal zone have long attracted evolutionary biologists seeking to understand the effects of Panama's formation on biodiversity (Lessios, 2008). This tiny strip of shoreline is regularly inundated by ocean tides and is host to diverse biological communities that include species unique to this extreme and transitional environment (Johnson & Baarli, 2012). Collembola (springtails) are among the most abundant and diverse members of the intertidal meso/meiofauna (Haynert et al., 2017) and include many halophilic and psammophilic species that thrive in the sandy, high-salt conditions typical of marine littoral habitats (Joosse, 1976). These minute, wingless, insect-like arthropods, are particularly well-suited candidate taxa for biogeographic investigations of vicariance across the Isthmus of Panama because 1) many species are strictly associated with intertidal habitats, presumably preventing transisthmian migration, 2) these marine littoral-obligate species can be extremely

abundant (Leinaas & Ambrose, 1992), 3) they often occur only in distinct, well-defined habitats that permit exhaustive sampling efforts (e.g., organic debris that accumulates at the high water mark), and 4) these groups are widely distributed throughout the Caribbean basin and Pacific shores of Mexico and Central America (Christiansen & Bellinger, 1988; Soto-Adames, 1988; Soto-Adames & Guillén, 2011). Given the above, multiple marine littoral-obligate species are likely to be co-distributed along both of Panama's shorelines. Moreover, because the intertidal fauna is thought to be among the last marine communities to be isolated by the closure of the Isthmus, patterns of divergence between intertidal geminate species will more closely reflect the timing of final closure (Knowlton et al., 1993; Knowlton & Weigt, 1998, Miura et al., 2010).

Their great abundance, low vagility, and remarkable ecological specificity renders springtails excellent sources of information for inferring biogeographic processes (Garrick, 2008). Nearly all recent phylogeographic studies on springtails have detected deeply divergent and geographically structured genetic lineages, sometimes at very fine spatial scales (e.g., Garrick 2008; Katz et al., 2018)—an indication that it is relatively common for springtail lineages to persist locally for millions of years (Cicconardi et al., 2010; 2013). Their capacity to survive dramatic environmental transformations, such as the Messinian Salinity Crisis (Cicconardi et al., 2010) and the climatic oscillations of the Pleistocene (Cicconardi et al., 2013), suggests that springtails are robust to ecological change—a desirable trait for evaluating historical evolutionary processes. For example, unrecognized extinct lineages can obfuscate true sister relationships, presenting a significant source of error when determining the status of geminate species (Lessios, 2008). To date, only one study has utilized Collembola to explore evolutionary processes shaping patterns of molecular diversity in Panama (i.e., Cicconardi et al. 2013). Among their findings, Cicconardi et al. (2013) attribute an increase in lineage

diversification for a group of forest-dwelling springtails, occurring 15–20 Ma, to the emergence of the first colonizable lands during the formation of Isthmus. However, these results support both proposed models of Isthmian closure (Miocene vs. Pliocene), as land masses were present for both models at the timing of divergence (i.e., as a near complete peninsula or closed Isthmus, Montes et al., 2012b; 2015; or as an island chain, Coates & Obando, 1996). The genetic history of marine littoral springtails may provide greater insights into the timing of final closure, because at that time, gene flow would have been permanently cut off between populations on either side, leaving behind a genetic signature of isolation, detectable via evaluation of patterns in the contemporary distribution of molecular diversity, that can be dated using molecular clocks. However, the molecular diversity of Panama's marine littoral springtail communities has not been evaluated, and therefore, the evolutionary mechanisms responsible for shaping their diversity has yet to be explicitly tested.

The aim of this study is to determine if spatial and temporal patterns of molecular diversity of Panama's marine littoral springtail fauna are consistent with the geminate species concept. To accomplish this, I performed targeted surveys for marine littoral springtails at twelve intertidal localities along the Atlantic and Pacific coasts of Panama to develop molecular datasets for phylogeographic analysis: one for marine littoral-obligate taxa in the genus *Seira* Lubbock, 1870 (Superfamily Entomobryoidea), that includes DNA sequences from the mitochondrial cytochrome c oxidase I (COI) gene, nuclear 28S ribosomal gene, and nuclear elongation factor-1α (ef1α) gene; and an additional dataset of mtDNA COI sequences for taxa in marine littoral-obligate genera of the Superfamily Isotomoidea, including *Archisotoma* Linnaniemi, 1912, *Axelsonia* Börner, 1906, *Psammisotoma* Greenslade & Deharveng, 1986, and *Spinacteletes* Soto-Adames, 1988. These data were used to test the hypothesis that taxa on opposite sides of the

Isthmus are geminate species pairs by 1) evaluating patterns of genetic structure to identify processes driving and maintaining isolation; 2) reconstructing phylogenetic relationships to identify transisthmian sister taxa; and 3) estimating divergence times to determine the age of lineages spanning the Isthmus.

Methods

Focal taxa and field collections

I targeted marine littoral-obligate species of the genus *Seira* (Entomobryoidea), and species in the marine littoral-obligate genera of the Subfamily Isotomoidea, *Archisotoma*, *Axelsonia*, *Psammisotoma*, and *Spinactaletes* (Fig. 2.1). The genus *Seira* includes two neotropical marine littoral-obligate species, *Seira fulva* (Schött), 1896 and *Seira nicoya* Christiansen & Bellinger, 1988, reported only from organic debris on sandy beaches and mudflats in Mexico (Schött, 1896; Christiansen & Bellinger, 1988; Vázquez & Palacios-Vargas, 1990) and Colombia (Riascos, 2002). Species in the genera *Archisotoma* and *Psammisotoma* have been reported throughout the Neotropics (Mari Mutt & Bellinger, 1990) in sand or organic debris deposited on beaches at the high tide mark (e.g., Thibaud, 1993; Abrantes & Mendonça, 2007, 2009). *Axelsonia tubifera* Strenzke, 1958, occurs exclusively on mangrove roots or mangrove leaf litter exposed at low tide and is reported from Brazil (Strenzke, 1958), Jamaica (Christiansen & Bellinger, 1998), and Puerto Rico (Samalot Roque, 2006). Many species of *Spinactaletes* have been reported from Puerto Rico (Soto-Adames, 1988), Mexico (Soto-Adames, 1987), Jamaica (Bellinger, 1962), and Venezuela (Najt & Rapoport, 1972), and are usually restricted to rocky seashores with exposed limestone (Soto-Adames & Guillén, 2011).

With these habitats in mind, prospective sampling sites were identified via satellite imagery and selected based on accessibility and habitat type (i.e. sandy beaches, rocky shores, mangroves, or exposed reef close to shore). Twelve primary localities along the Atlantic and Pacific coasts of Panama were surveyed between June 8th and July 4th, 2014 (Fig. 2.2). Marine littoral springtail habitats were exhaustively sampled at each site. If seaweed, leaf litter, and/or other organic debris was present on the beach, it was collected into sealable plastic bags, processed overnight via Berlese funnel, and stored in 95% ethanol. Specimens were also manually collected directly into ethanol from rock surfaces, mangrove roots, and beach litter using an aspirator and preserved directly in 95% ethanol. All specimens were sorted to putative morphospecies using a Leica MZ12.5 research dissecting microscope and representative individuals of each morphospecies from each locality were subsequently slide mounted in Hoyer's Medium (Mari Mutt, 1979) for more detailed morphological observation using a Nikon Eclipse E400 upright microscope with phase contrast. Additional representatives of each morphospecies (determined by dissecting microscopy) from each sampling locality were retained for DNA extraction.

DNA extraction, PCR, DNA sequencing, and alignment

DNA was extracted from whole specimens in the genera *Seira* (n=43), *Archisotoma* (n=20), *Axelsonia* (n=4), *Psammisotoma* (n=7), and *Spinactaletes* (n=12) using the following modifications to the DNeasy Blood & Tissue kit protocol (Qiagen Inc.,Valencia, CA) to increase genomic DNA yields: samples were left to incubate at 56°C overnight in ATL buffer with proteinase-K; stored at 4°C overnight in ethanol to maximize DNA precipitation; Promega

Wizard SV Minicolumns (Promega Corporation, Madison, WI, USA) were used in place of those supplied with the Qiagen kit; and 25–28 μL of AE buffer was used for the final elution step.

Mitochondrial (COI) and nuclear (28S and ef1a) gene fragments were amplified by PCR using illustra PuReTaq Ready-To-Go PCR beads (GE Healthcare, Buckinghamshire, UK), 21 μL of water, 1 μL of 10 μM forward and reverse primers (Table 2.1), and 2 μL of genomic DNA. The PCR reactions were run with the following thermocycler protocol: an initial denaturing temperature of 95°C for 2 min followed by 40 cycles of (95°C for 30s, 50–53°C (depending on primer TM) for 30s, and an extension temperature of 72°C for 2 min). Successful amplification was verified via gel electrophoresis (90 V, 400 mA for 35 min) using a 1% agarose gel stained with GelGreen (Biotium Inc., Hayward, CA, USA). PCR products were cleaned with the following modifications to the QIAquick PCR Purification Kit protocol (Qiagen Inc., Valencia, CA): Promega Wizard SV Minicolumns were used instead of supplied Qiagen kit columns and 22–28 μL of buffer was used for the final elution amplified DNA. If multiple bands were observed, PCR was repeated, bands were separated in a 1% agarose gel, excised and cleaned using the Zymoclean Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, CA, USA) and 10 μL of H_2O for final elution. Cleaned PCR products were sequenced in both directions using 3 μL of a mixture of BigDye Terminator v3.1, dGTP BigDye Terminator v3.0, and buffer in a ratio of 2:1:1 respectively, 1.6 μL of 2 μM primer primers, and 2 μL DNA. Sequencing reactions were run using the following thermocycler protocol: 96°C 2 min (95°C 20s; 50°C 5s; 60°C 240s) for 25 cycles. Sequence reaction products were cleaned using Performa ultra 96-well column plates (Edge Biosystems Inc., Gaithersburg, MD, USA) and centrifuged for 5 min at 964 g. Sanger sequencing was completed at the W. M. Keck Center (University of Illinois at Urban-Champaign, Urbana, IL, USA). Outgroup and additional ingroup taxa were chosen based on their

affinities with the focal taxa and the availability of sequences in GenBank (Benson et al., 2013). See Table 2.2 for list of all specimens used in this study and their corresponding sequence GenBank accession numbers.

Forward and reverse sequences were assembled with Sequencher v. 5.4 (Gene Codes Corporation, Ann Arbor, MI). All sequences were aligned using the MAFFT v7.273 (Katoh & Standley, 2013) (G-INSI-i method) algorithm implemented in Geneious v11.1.4 (Kearse et al., 2012) and translated to amino acids to check for stop codons.

Geminate species criteria

The recognition of geminate species in a Panamanian context is contingent upon two primary conditions: 1) they are transisthmian sister taxa, and 2) the rise of the Panamanian Isthmus initiated, and subsequently maintained, their isolation (Collins, 1996). In practice, these conditions are typically recognized by using a combination of independent criteria that together can provide evidence of recent divergence, geographic isolation, and vicariance across the Isthmus (Collins, 1996). For the purpose of this study, putative geminate species are analyzed with respect to morphological, molecular, and biogeographic criteria. To be considered geminate species, marine littoral-obligate springtail taxa distributed on both sides of the Isthmus must be morphologically similar, be reciprocally monophyletic sister taxa, exhibit patterns of genetic structure across the Isthmus, have divergence times that coincide with the timing of proposed models of Isthmian closure, and have similar divergence times to those estimated for other, codistributed geminate springtail species.

Species identification

Focal genera utilized for this study (i.e., *Seira*, *Archisotoma*, *Axelsonia*, *Psammisotoma*, and *Spinactaletes*) were recognized using a range of diagnostic characters with respect to body shape and size, chaetotaxy, color pattern, and the morphology of mouth parts, claw, and mucro (Christiansen & Bellinger, 1998). However, the diversity of neotropical species within these genera is poorly understood, and species-level diagnostic keys for this geographical area are taxonomically limited (e.g., Christiansen & Bellinger, 1988). Therefore, for species identification I referred to available taxonomic revisions for *Archisotoma* (Poinsot, 1965; Thibaud & Palacios-Vargas, 2001), *Psammisotoma* (Greenslade & Deharveng, 1986), and *Spinactaletes* (Soto-Adames, 1988), while paying particular attention to descriptions that were available for neotropical species in the focal genera (e.g., Strenzke, 1955; Delamare Deboutteville, 1956; Strenzke, 1958; Bellinger, 1962; Najt & Rapoport, 1972; Christiansen & Bellinger, 1988; Christiansen & Bellinger, 2000; Soto-Adames, 1987; Thibaud, 1993; Abrantes & Mendonça, 2007; Abrantes & Mendonça, 2009; Soto-Adames & Guillén, 2011).

Cryptic species complexes are believed to comprise the vast majority of springtail diversity (Cicconardi et al., 2013), and because their detection is often difficult without genetic data, traditional morphological characters should be augmented with molecular information to produce a more realistic estimate of species diversity (Katz et al., 2015). Therefore, to define operating taxonomic units (OTUs) for species tree analyses (see below), I used the General Mixed Yule Coalescent (GMYC) (Pons et al., 2006). The GMYC analyses were performed in R (R Core Team, 2017) using the splits package (Ezard et al., 2009) with the single threshold delimitation method (default parameters). Ultrametric COI gene trees used for the GMYC analyses were estimated with BEAST2 v. 2.4.8 (Bouckaert et al., 2014) with the following

parameters: bModelTest (Bouckaert & Drummond, 2017) for site model averaging (default parameters); the strict clock rate was set to 1; Yule tree prior; monophyletic constraint prior on all ingroup taxa; MCMC for 200 million generations; and sampling trees and statistics every 5000 generations. Focal taxa that could not be assigned to a known species, but which were morphologically distinct, were given a letter to signify putative specific status (e.g., sp. A). All OTUs that could not be differentiated by morphology are considered to be part of a putative cryptic complex but are referred to as a 'morphospecies' throughout this paper.

Tests for genetic structure

Two molecular datasets were developed for this study and utilized for phylogeographic analyses (Table 2.2): a *Seira* dataset including the mitochondrial locus COI (1440 bp) and nuclear loci, 28S (671 bp) and ef1a (789 bp); and 2) a COI (1500 bp) dataset for focal taxa in the Superfamily Isotomoidea, including the genera *Archisotoma*, *Axelsonia*, *Psammisotoma*, and *Spinactaletes*. Only COI was amplified for the Isotomoidea dataset, due to limited material.

To meet the geminate species criteria that divergence is driven by isolation due to vicariance, rather than dispersal, I evaluated and compared patterns of genetic structure for each morphospecies to identify potential barriers (the Isthmus vs. geographic distance) driving genetic isolation between transisthmian lineages. First, I performed an analysis of molecular variance (AMOVA) independently for each gene using Arlequin v3.5.2.2 (Excoffier & Lischer, 2010) by grouping haplotypes within and among geographic region relative to the Isthmus (north vs. south). In addition, I also performed hierarchical AMOVAs by grouping haplotypes among regions (north or south of the Isthmus), among localities within regions, and within localities to

determine if populations are structured by locality, rather than, or in addition to, region. Significance was assessed with 50,000 permutations.

Isolation-by-distance (IBD) resulting from dispersal and genetic drift, rather than vicariance, often produces strong patterns of genetic structure in non-vagile organisms (Costa et al., 2013). Therefore, I also evaluated the significance of genetic structure across the Isthmus while controlling for geographic distance by conducting partial Mantel tests (Mantel, 1967) for each gene using zt v1.1 (Bonnet & Van de Peer, 2002). Three matrices containing uncorrected pairwise genetic distances, pairwise geographic distances (great-circle distance), and binary variables coded to indicate whether each pair of specimens occurred together or on opposite sides of the Isthmus, were inputted into zt to calculate the Mantel test statistic (r) for each focal morphospecies. 100,000 permutations were performed to assess significance. All pairwise genetic distances (uncorrected p-distances) were calculated in Paup* v4.0a build 161 (Swofford, 2002). COI p-distances were also estimated for taxa occurring on opposite sides of the Isthmus and for taxa occurring on the same side of the Isthmus and plotted in R to visualize and compare distances among morphospecies.

Phylogenetic inference and divergence time estimation

A Bayesian phylogenetic approach was employed to estimate evolutionary relationships and divergence times for all sampled lineages of *Seira*, *Archisotoma*, *Axelsonia*, *Psammisotoma,* and *Spinactaletes* to further evaluate the hypothesis that events associated with the closure of the Isthmus of Panama shaped the evolutionary history of marine littoral springtail assemblages, and, specifically, to determine whether or not: 1) transisthmian lineages are reciprocally monophyletic sister taxa, 2) divergence times for transisthmian lineages correspond to a Pleistocene or

Miocene closure model, and 3) transisthmian lineage divergence times are similar among codistributed and related genera. I conducted Bayesian phylogenetic analyses for both the *Seira* and Isotomoidea datasets using BEAST2 with the following parameters: unlinked site and clock models for each gene partition; linked tree models for all gene partitions; site model averaging implemented in bModelTest to accommodate uncertainty in the model of sequence evolution (default parameters); monophyletic constraint prior on all ingroup taxa; MCMC for 200 million generations; and sampling trees and statistics every 5000 generations.

To assign absolute dates to divergence times, a molecular clock needs to be calibrated against independent evidence (Wilke et al., 2009). Dated fossils are the most common source of timing information for molecular clock calibrations, but for springtails and other small and soft-bodied organisms, adequate fossil records for molecular clock calibration are typically not available. Rather than using the widely implemented insect rate calibration (2.3%/Ma) for COI (Brower, 1994), I decided to use the more statistically robust and faster insect rates estimated by Papadopoulou et al. (2010) for molecular clock calibration: COI=3.54%/Ma, 28S=0.12%/Ma. This conservative (faster) approach for clock calibration would underestimate, rather than overestimate, divergence times, which could then be interpreted as minimum ages (Cicconardi et al., 2010). Clock.rate parameters were set to 0.0168 and 0.0006 (substitutions/site/Ma), respectively, and ef1a rate was set to be estimated using a gamma distribution (alpha=0.001, beta=1000) prior. To determine whether the use of a strict clock model is appropriate for each gene partition, I conducted preliminary phylogenetic analyses using the same parameters as above with the exception that I applied an uncorrelated log normal clock model to all gene partitions to estimate the coefficient of variation (COV) for each clock partition. The resulting 95% highest posterior density (HPD) COV values for COI indicated high levels of rate variation

among lineages for both *Seira* and Isotomoidea datasets (95% HPD COV=1.39–3.6 and 1.59–3.07, respectively), whereas COV values for 28S and ef1a in the *Seira* dataset were relatively low, indicating a more clock-like evolution among lineages (95% HPD COV=0.33–0.99, 0.00–0.67). Accordingly, I used uncorrelated log normal clock models for both *Seira* and Isotomoidea COI partitions and strict clock models for the *Seira* 28S and ef1a partitions.

When sequence alignments are largely composed of intraspecific diversity, a multispecies coalescent tree prior, which combines coalescent (population) and birth-death (species) tree models, may be more appropriate for estimating phylogeny and divergence times because it can accommodate for both species and population-level branching patterns and differences between gene trees due to incomplete lineage sorting (Drummond and Bouckaert, 2015). Therefore, in addition to the aforementioned concatenated gene tree analyses, I also took a multispecies coalescent approach to estimate phylogeny and divergence times for both datasets using the starBEAST2 v0.14.0 (Ogilvie et al., 2017) package implemented in BEAST2 with the following parameters: unlinked site, clock, and tree models for each gene partition; gene ploidy COI=0.5, 28S=2.0, ef1a=2.0; analytical population size integration; bModelTest for site model averaging (default parameters); uncorrelated lognormal clock model for COI (0.0168 substitutions/site/Ma); strict clock models for 28S (0.0006 substitutions/site/Ma) and ef1a (estimated with a gamma distribution prior, alpha=0.001, beta=1000); Yule tree prior; monophyletic constraint prior on all ingroup taxa; MCMC for 200 million generations; and sampling trees and statistics every 5000 generations. The starBEAST2 XML files were manually edited to provide starting species and gene trees with monophyletic ingroups. OTUs identified by GMYC analysis were used for the assignment of taxon sets required for estimating species trees in starBEAST2.

For all phylogenetic analyses, the effective sample size (ESS) for all parameters were determined to be greater than 200 (10% burn-in) using Tracer v1.6.0 (Rambaut & Drummond, 2007) and maximum clade credibility trees with median node heights were inferred with TreeAnnotator v2.4.8 (Bouckaert et al., 2014) using a 10% burn-in.

Results

Species identification

A total of 7 morphospecies in the focal genera *Seira*, *Archisotoma*, *Axelsonia*, *Psammisotoma*, and *Spinactaletes* (Fig. 2.1) were identified from 8 of the 12 sampling localities along the Panama's Pacific and Caribbean shorelines (i.e., Boca Brava, Fort Sherman, Nombre de Dios, Playa Boca del Drago, Playa Diablo, Playa Torio, Playa Tortuguilla, and Punta Chame) (Table 2.3). *Seira*, *Archisotoma*, and *Psammisotoma* were found in seaweed and other organic debris deposited along the high-water mark on sandy beaches (Fig. 2.3a) from both the Pacific and Caribbean coasts, and often occurred together in the same sample. *Axelsonia* was collected on mangrove roots exposed at low tide (Fig. 2.3b) from both coasts, but only from two localities (i.e., Boca Brava and Playa Diablo) and in very small numbers. *Spinactaletes* were collected from three Caribbean localities (Nombre de Dios, Playa Boca del Drago, and Play Diablo) exclusively from the surface of coralliferous limestone reef outcrops exposed during low tide (Fig. 2.3c) or along old docks lined with dredged reef (Fig. 2.3d) and were not present on other substrates. *Spinactaletes* was not collected from any Pacific localities because this specific habitat (coralliferous limestone or dredged corals) was not encountered along the southern coast of Panama.

Morphological observation of slide mounted specimens yielded no discernable differences among specimens representing the genera *Seira* or *Psammisotoma,* suggesting all specimens within each genus represent a single morphospecies: *Seira nicoya* and *Psammisotoma dispar* Christian & Bellinger, 1988. In contrast, collections from Torio, yielded two *Archisotoma* morphospecies, *Archisotoma* sp. A and *Archisotoma* sp. B, which could not be identified to a known species, but are clearly differentiated from each other by chaetotaxy and additional diagnostic morphological characters (Thibaud & Palacios-Vargas, 2001). *Archisotoma* sp. A was identified from 3 additional localities (Nombre de Dios, Playa Tortuguilla, and Fort Sherman), but *Archisotoma* sp. B was only found in a single sample of organic debris collected from Torio and was not recovered from other localities. All slide-mounted specimens of the genus *Axelsonia* were identified as *Axelsonia tubifera*. This species was collected in abundance from Boca Brava, but only a single specimen was recovered from Playa Diablo, which was subsequently destructively sampled for DNA extraction prior to slide mounting for morphological examination. Other than slight differences in antennal pigmentation, all specimens of *A. tubifera* were morphologically identical under stereomicroscopy. *Spinactaletes* collected from Playa Boca del Drago (*Spinactaletes* sp. A) were found to be morphologically distinct from those collected from Playa Diablo and Nombre de Dios (*Spinactaletes* sp. B) and could be clearly differentiated by metatibiotarsus spur morphology—a well-established species-level diagnostic character (Soto-Adames, 1988; Soto-Adames & Guillén, 2011). In summary, my sampling efforts yielded four marine littoral-obligate morphospecies that have transisthmian distributions: *S. nicoya*, *Archisotoma* sp. A, *A. tubifera*, and *P. dispar*.

Bayesian analysis of the COI gene for both *Seira* and Isotomoidea datasets produced highly supported gene trees with strong posterior clade support for each of the seven

morphospecies (posterior probability (pb)=1) (Figs 2.4–2.5). The GMYC analysis based on the COI gene trees identified 11 OTUs for the 7 focal morphospecies (Figs 2.4–2.5). *Seira nicoya* was determined to have 3 distinct genetic lineages, each forming a clade (pb=1) by sample locality—Torio, Punta Chame, and all Caribbean localities—with Torio and Punta Chame OTUs forming a 'Pacific' clade that is reciprocally monophyletic and sister to the 'Caribbean' OTU clade. *Archisotoma* sp. A formed two clades (pb=1) grouping Pacific and Caribbean lineages as two reciprocally monophyletic sister lineages, each identified as an OTU. The same pattern was observed in *A. tubifera*, with Pacific lineages from Boca Brava forming a clade (pb=1) that is sister to a single lineage from Playa Diablo on the Caribbean side, each identified as a distinct OTU. *Archisotoma* sp. B, *Spinactaletes* sp. A, *Spinactaletes* sp. B, and *P. dispar*, all form their own clades (pb=1) that were each identified as an OTU. However, of all the OTUs identified, only *P. dispar* includes both Pacific and Caribbean lineages, albeit they do not group into distinct, reciprocally monophyletic 'Pacific' and 'Caribbean' clades.

Genetic structure across the Isthmus

Uncorrected pairwise COI distances (p-distance) boxplots (Fig. 2.6) revealed remarkably high genetic distances between morphologically identical lineages occurring on opposite sides of the Isthmus: *S. nicoya*, mean=0.196 (n=216), range=0.186–0.206; *Archisotoma* sp. A, mean=0.155 (n=18), range=0.153–0.159; and *A. tubifera*, mean=0.151 (n=3), range=0.151–0.151. Mean distances among lineages occurring on the same side of the Isthmus were relatively low: 0.027 (n=312; range=0–0.174), 0.008 (n=37; range=0.001–0.016), and 0.0004 (n=3; range=0–0.0007), for *S. nicoya*, *Archisotoma* sp. A, and *A. tubifera*, respectively. Although, distances between *S. nicoya* lineages from the Pacific localities, Torio and Punta Chame, were closer to estimated

transisthmian distances (mean=0.172; n=20; range=17.15–17.43), they suggest additional, species level differentiation. In contrast, genetic distances between transisthmian lineages of *P. dispar* (mean=0.011; n=10; range=0.0006–0.028) were much lower compared to *S. nicoya*, *Archisotoma* sp. A, and *A. tubifera*, and were not significantly different from those estimated between lineages occurring together on the same side of the Isthmus (mean=0.014; n=11; range=0.0007–0.029).

AMOVAs (Table 2.4) identified the majority of genetic variation in *S. nicoya* and *Archisotoma* sp. A to be strongly structured among regions north and south of the Isthmus when haplotypes were grouped by region only (% variation for *S. nicoya* COI=79%, 28S=87%, ef1a=95%; *Archisotoma* sp. A COI=95%) and when haplotypes were grouped by locality and region (% variation for *S. nicoya* COI=71%, 28S=83%, ef1a=93%; *Archisotoma* sp. A COI=95%). This pattern was not recovered for *P. dispar*, which lacked genetic structure across the Isthmus resulting in negative variance among regions (Table 2.4). Partial mantel tests (Table 2.5) also show strong, positive correlations between genetic distances and region relative to the Isthmus after controlling for geographic distance for *S. nicoya* (COI, r=0.93; 28S, r=0.88; ef1a, r=0.99) and *Archisotoma* sp. A (COI, r=0.97). However, a correlation between genetic distance and position relative to the Isthmus was also not supported for *P. dispar* (COI, r=-0.08). Genetic structure was not evaluated for *Spinactaletes*, which was not collected from the Pacific coast, and AMOVAs and Mantel tests were not performed for *Axelsonia*, which was only collected from two localities.

Phylogeny and Divergence times

The gene tree and species tree analyses produced congruent estimates of phylogeny and divergence times. For transisthmian lineages of *S. nicoya* (Fig. 2.7a) and for *Archisotoma* sp. A and *A. tubifera*, both approaches recovered topologies consistent with COI gene trees estimated for the GMYC analyses (Figs 2.4–2.5) with reciprocally monophyletic sister clades, each with strong clade support (pb=1) (Fig. 2.7b). Divergence time estimates for these transisthmian sister lineages were very similar, suggesting that they were initially isolated during the early Miocene or late Oligocene. Gene tree divergence time estimates places the median age of the most recent common ancestors (MRCA) for transisthmian populations of *S. nicoya*, *Archisotoma* sp. A, and *A. tubifera*, at 26.55 Ma [95% highest posterior density (95% HPD) =19.54–34.45 Ma], 25.40 Ma [95% HPD=14.72–37.95 Ma], and 16.07 Ma [95% HPD=14.72–37.95 Ma], respectively. The species tree produced equivalent divergence times, though were slightly younger and with smaller confidence intervals: *S. nicoya*, [median=24.76 Ma; 95% HPD=18.75–30.84 Ma], *Archisotoma* sp. A [median=18.87 Ma; 95% HPD=11.21–28.8 Ma], and *A. tubifera* [median=12.90 Ma; 95% HPD=7.44–19.44 Ma). Pacific populations of *S. nicoya* were grouped by locality forming two distinct subclades, with mean divergence times estimated from the gene tree and species tree analyses at 9.57 Ma [95% HPD=5.91–13.42] and 11.14 Ma [95% HPD=7.39–15.12], respectively. It is also worth noting that the divergence time intervals (95% HPD) for these transisthmian pairs show significant overlap (Table 2.6; Fig. 2.7), suggesting simultaneous divergence, though the confidence intervals are relatively wide. However, among these taxa, the minimum bound of the most recent divergence time interval predates the Pliocene (i.e., *Axelsonia tuberifa*, 95% HPD=7.44–19.44 Ma). In contrast, reciprocal monophyly was not recovered for transisthmian lineages of *P. dispar* (Fig. 2.7b) and median time of lineage

diversification from both gene tree and species tree analyses, 1.04 Ma [95% HPD=0.66–1.57 Ma] and 1.26 Ma [95% HPD=0.71–1.96 Ma, respectively, are very recent compared to those estimated for the other transisthmian focal taxa, indicating that lineage diversification within *P. dispar* occurred during the Pleistocene.

Discussion

Cryptic geminate species pairs support an early Isthmus

Molecular analyses of marine littoral-obligate springtails collected from 12 localities along Panama's coastlines uncovered reciprocal monophyletic sister relationships between morphologically cryptic Pacific and Atlantic lineages (Fig. 2.7), with high levels of genetic structure partitioned by regions North and South of the Isthmus of Panama (Tables 2.4–2.5, Fig. 2.6), and divergence times between transisthmian lineages dated to the Oligocene/Miocene (Fig. 2.7). Collectively, these data support predictions of the geminate species concept for three morphospecies: *S. nicoya*, *Archisotoma* sp. A, and *A. tubifera* (Table 2.6). The majority of genetic variation was distributed among regions North and South of the Isthmus (Table 2.4), suggesting that the Isthmus serves as a major barrier to gene flow between these geminate species pairs. Mantel tests confirmed these results by excluding geographic distance as a significant driver of genetic isolation (Table 2.5).

Phylogeographic congruence among three codistributed putative geminate species pairs lends further support to a hypothesis of vicariant speciation across the Isthmus because extrinsic evolutionary factors (e.g., geographic barriers to gene flow) are likely to exert equal selective pressures on codistributed and closely related groups (Lapointe & Rissler, 2005). However, the interruption of gene flow may have occurred at different times during Panama's closure (Lessios,

1998). Several authors have attributed differences in sequence divergence among geminate species pairs across the Isthmus of Panama to differences in habitat depth, or proximity to the ocean (Knowlton et al., 1993; Knowlton and Weigt, 1998, Miura et al., 2010). In this study, median divergence times estimated for transisthmian lineages of mangrove-dwelling *A. tubifera* (12.9–16.1 Ma) were more recent compared to those observed for transisthmian lineages of sand/beach debris-dwelling *S. nicoya* (24.8–26.6 Ma) and *Archisotoma* sp. A (18.9–25.4 Ma), which would have validity if mangrove communities were among the last to be isolated via Isthmian closure (Knowlton et al., 1993; Knowlton and Weigt, 1998, Miura et al., 2010). Lastly, all divergence time intervals for geminate species pairs predate a Pliocene closure (i.e., 2.8–4 Ma) by at least 3.5 Ma. Together, these results indicate that the geographic distribution of genetic diversity in marine littoral-obligate springtails cannot be explained by the traditional model of Pliocene closure, and instead, is more likely to have been shaped by paleogeographic events associated with a Miocene closure—a model supported by geological evidence that proposes the existence of a nearly complete Isthmus and/or temporary land bridges during the late Oligocene/early Miocene (20–25 Ma) with final closure 13–15 Ma (Montes et al., 2012a; 2012b; 2015)—with marine littoral springtails now added to the growing list of biological evidence supporting a Miocene land connection (e.g., Bacon et al., 2013; Elmer et al., 2013; Bacon et al., 2015a; Thacker, 2017; Winston et al., 2017; Stange et al., 2018).

In a recent meta-analysis of 169 dated molecular phylogenies, Bacon et al. (2015a) suggests that the emergence of Panama was more geologically complex than previously thought, providing multiple opportunities for vicariance between Atlantic and Pacific lineages of marine organisms. Bacon et al. (2015a) identified three potential large scale vicariant events for marine organisms at 23.73 (19.9–27.41), 7.96 (7.776–8.96), and 2.06 (1.03–4.35) Ma, two of which are

consistent with divergence times identified in this study (Table 2.6): *S. nicoya* and *Archisotoma* sp. A split during the initial event (11–27 Ma) and *A. tubifera* overlapped with the middle event (7–24 Ma). These dates are also similar to those reported for some intertidal invertebrates along Panama's coastlines such as 9.4–17.9 Ma in marine flatworms (Scarpa et al., 2015) and 14–18.8 Ma in bivalves (Marko & Moran, 2009). Additionally, out of 115 putative geminate species pairs (of crustaceans, sea urchins, fishes, and mollusks), Lessios (2008) suggests that more than half had diverged prior to the Pliocene.

Axelsonia tubifera had relatively lower median divergence times compared to *S. nicoya* and *Archisotoma* sp. A, but due to significant overlap in 95% HPD intervals (Table 2.6), a hypothesis of simultaneous divergence cannot be rejected. However, lineage divergence may not be correlated with barrier formation, even when multiple codistributed transisthmian taxon pairs exhibit similar patterns of divergence across a given barrier (Heads, 2005). Therefore, I cannot ascertain with certainty whether my sampling was sufficiently extensive to reliably detect true sister taxa isolated by Isthmus formation (Marko et al., 2015), though I did incorporate multiple criteria to assess geminate species status to minimize this possibility (Table 2.6). I also employed an external insect molecular clock proposed by Papadopoulou et al. (2010) to estimate divergence times, but I used an uncorrelated log normal clock model to account for possible rate variation among lineages. Moreover, the incorporation of a relatively conservative (faster) COI rate prior (3.54%/Ma) will likely underestimate, rather than overestimate divergence times to infer the minimum age of lineages—an approach used in other springtail biogeography studies (e.g., Cicconardi et al., 2010; 2013).

Evidence of post-closure dispersal across the Isthmus

Patterns of lineage diversification within *P. dispar* were in stark contrast to those observed for the other transisthmian morphospecies. Transisthmian lineages of *P. dispar* were not recovered as reciprocally monophyletic 'Pacific' and 'Atlantic' clades (Fig. 2.7b), had relatively low levels of genetic divergence (<3% COI p-distance) (Fig. 2.6), completely lacked geographic genetic structure across the Isthmus (Tables 2.4–2.5), and lineage diversification was estimated to have occurred within the last 2 Ma (Table 2.6; Fig. 2.6), well after either proposed models of Isthmus closure. These results clearly reject the hypothesis that geographic isolation was initiated and maintained by the emergence of the Isthmus, and therefore, they cannot be considered geminate species. In fact, patterns of molecular diversity in *P. dispar* cannot be explained by vicariance, regardless of final closure timing. Instead, their lack of reciprocal monophyly, low genetic divergence, and absence of genetic structure north and south of the Isthmus, are more consistent with a scenario of post-closure dispersal and evidence of recent gene flow and/or range expansion across the Isthmus.

The phylogeographic discordance between *P. dispar* and the 3 other codistributed transisthmian species pairs was relatively surprising, given that *Psammisotoma* is presumably a marine littoral-obligate genus (Greenslade & Deharveng, 1986; Christiansen & Bellinger, 1988). Differences in habitat specificity among species has been shown to have large impacts on patterns of genetic isolation and connectivity in springtails (Katz et al., 2018) and other invertebrates (Ayre et al., 2009). Therefore, it is possible that the contrasting phylogeographic patterns observed for *P. dispar* reflect unrecognized relaxed ecological restrictions that enable limited active dispersal across the Isthmus. Contradicting this interpretation, is the fact that, to my knowledge, this species has only been reported from marine intertidal habitats in Florida

(Christiansen & Bellinger, 1988) and Puerto Rico (Samalot Roque, 2006), and active dispersal across terrestrial habitats is unlikely. An alternative explanation is that this species has an increased propensity for passive dispersal. Long distance dispersal via oceanic and air currents have been well-documented in springtails (Freeman, 1952; Blackith & Disney, 1988; Coulson et al., 2002; Coulson et al., 2003; Hawes et al., 2007; Hawes et al., 2008), and may serve as a potential mode of transisthmian or circumglobal dispersal for this species. Muira et al. (2012) provide compelling molecular evidence of multiple post-closure dispersal events in a group of snails believed to be mediated by birds, lending credibility to the proposition that springtails could be hitching rides across the Isthmus on birds (or on their food). Human-mediated dispersal via the Panama Canal in ship ballast seawater is an additional scenario that may have provided this species a direct path across the Isthmus with minimal exposure to terrestrial or aquatic habitats (Roy & Sponer, 2002).

Pacific vs. Atlantic lineages of Seira nicoya

Bayesian phylogenetic analyses revealed substantial differences in the patterns of genetic diversity between Pacific and Atlantic lineages of *S. nicoya*. The Atlantic clade, comprised of taxa collected from Boca del Drago through Nombre de Dios (300+ km apart), can be characterized as genetically homogeneous, with little genetic variation (<4% COI p-distance) (Fig. 2.6) among lineages, which are not grouped by locality. However, the Pacific clade, comprised of taxa from Torio and Punta Chame (~180 km apart), each form a genetically distinct sub-clade by locality. Despite occurring together on the same side of the Isthmus, these Pacific sub-clades are estimated to have been isolated for ~10 Ma (Fig. 2.7), accumulating ~17% p-distance in COI (represented as outliers in Fig. 2.6). The closure of the Central American Seaway

is thought to be associated with massive changes to intertidal ecosystems in Panama, such as the loss of reef in the Eastern Pacific (Leigh et al., 2014). It is possible that reduced habitat connectivity has impacted the genetic cohesion of Pacific populations, leading to long-term isolation and subsequent speciation. The loss of reef habitat in the Eastern Pacific also may explain why *Spinactaletes* were not encountered at any of the Pacific localities in Panama, as they appear to be strictly limited to coastal outcrops of coralliferous limestone. Nonetheless, additional sampling is required to more fully adjudicate potential factors responsible for the disparate patterns of genetic structure observed between Atlantic and Pacific marine littoral springtail lineages.

Conclusions

Patterns of molecular diversity of marine littoral springtails inhabiting Panama's intertidal zones have revealed morphologically cryptic, deeply divergent, and reciprocally monophyletic 'Atlantic' and 'Pacific' sister lineages for three codistributed morphospecies, *S. nicoya*, *Archisotoma* sp. A, and *A. tubifera*. The timing of lineage diversification for these putative geminate species pairs is in agreement with recent and controversial biological and geological studies that propose the existence of pre-Pliocene land connections between South and Central America, but I express caution in my biogeographic interpretations due to the assumptions made regarding molecular rates, and the possibility of divergence prior to the formation of the Isthmus. However, I also discovered evidence of recent gene-flow or range expansion across the Isthmus within the last 2 Ma for *P. dispar* and suggest dispersal of this species was passively mediated by ocean or air currents, or by animals such as birds or humans. This is significant, as it suggests some species seem to have greater dispersal potential and can maintain gene flow across large

distances and seemingly inhospitable habitats, despite having seemingly strict ecological associations. Finally, disparate patterns of genetic diversity between Atlantic and Pacific lineages of *S. nicoya* suggests that differential habitat connectivity may have influenced genetic cohesion in these marine littoral springtail lineages, but further investigation is required to identify factors promoting isolation for populations along the Pacific coast.

This is the first attempt to characterize patterns of molecular diversity in marine littoral springtails to address biogeographic questions. Additional sampling of springtail lineages along the coasts of North and South America, the development of springtail genomic data, and if possible, a robust springtail molecular clock, is needed to improve the precision and accuracy of divergence time analysis for future biogeographic inferences of Panama's geological history.

Acknowledgments

I thank Bruno Zachrisson and Onesio Martinez for their hospitality and assistance with field collections in Panama. Special thanks to Katherine Noble for providing help with field work in Panama. This project was supported by the Center for Latin American and Caribbean Studies at the University of Illinois at Urbana Champaign (Tinker Fellowship to A.D.K), the Department of Entomology at the University of Illinois at Urbana-Champaign, National Science Foundation (grant 0956255 to F.N.S) and the Illinois Natural History Survey (Ross Memorial Fund to A.D.K).

Tables

Table 2.1. PCR and sequencing primers used in this study, including references.

Primer	Direction	Sequence 5'–3'	Reference
COI (all taxa)			
LCO1490	F	GGT CAA CAA ATC ATA AAG ATA TTG G	Folmer et al., 1994
TL2-N-3014	R	TCC AAT GCA CTA ATC TGC CAT ATT A	Simon et al., 1994
C1-J-1718	internal F	GGA GGA TTT GGA AAT TGA TTA GTT CC	Simon et al., 1994
C1-J-2195	internal F	TTG ATT TTT TGG TCA TCC AGA AGT	Simon et al., 1994
COI (Seira only)			
670-1F	internal F	TGT TCA TGA GAC CAA ACA CC	Present study
860-1R	internal R	CGG TCA CCC TGA GGT GTA T	Present study
COI (Archisotoma only)			
ArcF1	internal F	TGA CCC AGC AGG TGG GGG CGA CCC CAT TC	Present study
ArcR1	internal R	TAC CAG GCT TTG GGA TAA TTT CAC ATA TTG	Present study
COI (Axelsonia only)			
AxeF1	internal F	TGA CCC AGC GGG TGG AGG AGA CCC TAT TC	Present Study
AxeR1	internal R	TCT ACC GGG ATT CGG AAT AAT CTC CCA C	Present Study
COI (Psammisotoma only)			
PsaF1	internal F	CGA CCC TTC TGG AGG AGG TGA CCC AAT CT	Present Study
PsaR1	internal R	TCC TTC CTG GAT TTG GAA TAA TTT CTC	Present Study
COI (Spinactaletes only)			
SpiF1	internal F	TGA CCC TGC AGG AGG GGG TGA CCC TAT TT	Present Study
SpiR1	internal R	CTT ATT CTC CCC GGG TTC GGT ATA GTT TCT	Present Study
28S (Seira only)			
28Sc1pF	F	CCC GCT GAA TTT AAG CAT	Present study
28Sd2R	R	CCC GTC TTG AAA CAC GGA	Present study
ef1a (Seira only)			
Ef1a44F	F	CCA CCA CCG GAC ATT TGA TC	Present study
Ef1a854dR	R	GTG GAR ATG CAY CAC GAA GC	Present study

Table 2.2. List of all taxa in the *Seira* (a) and Isotomoidea (b) datasets used for this study, including their corresponding GenBank sequence accession numbers. Specimen sequences developed for this study (and collected from Panama) are marked with an asterisk. All other sequences were obtained from GenBank.

(a) Taxa (*Seira* dataset)	COI	28S	ef1a
Outgroups			
Lepidocyrtus (*Ascocyrtus*) sp.	KM978345	KC236269	—
Lepidocyrtus cyaneus	KF642091	KY235260	—
Lepidocyrtus lanuginosus	MF611198	—	GU128387
Lepidocyrtus lignorum	KF642048	LK024354	GU128395
Lepidocyrtus paradoxus	KT706447 (5') AF398089 (3')	JN981068	—
Orchesella cincta	KT985987	LK024376	AJ009857
Orchesella villosa	EU016195	EF199972	GU128428
Pseudosinella alba	KM978368	KC236295	—
Pseudosinella tumula	KM978367	KC236294	—
Focal taxa			
**Seira nicoya* Fort Sherman 29_1	MH579480	MH579523	MH579437
**Seira nicoya* Fort Sherman 29_2	MH579481	MH579524	MH579438
**Seira nicoya* Fort Sherman 29_3	MH579482	MH579525	MH579439
**Seira nicoya* Fort Sherman 29_4	MH579483	MH579526	MH579440
**Seira nicoya* Fort Sherman 29_5	MH579484	MH579527	MH579441
**Seira nicoya* Fort Sherman 29_6	MH579485	MH579528	MH579442
**Seira nicoya* Fort Sherman 29_7	MH579486	MH579529	MH579443
**Seira nicoya* Nombre de Dios 1_2	MH579468	MH579511	MH579425
**Seira nicoya* Nombre de Dios 1_3	MH579469	MH579512	MH579426
**Seira nicoya* Nombre de Dios 1_4	MH579470	MH579513	MH579427
**Seira nicoya* Nombre de Dios 1_5	MH579471	MH579514	MH579428
**Seira nicoya* Nombre de Dios 1_7	MH579472	MH579515	MH579429
**Seira nicoya* Nombre de Dios 1_8	MH579473	MH579516	MH579430
**Seira nicoya* Nombre de Dios 1_9	MH579474	MH579517	MH579431
**Seira nicoya* Nombre de Dios 1_10	MH579467	MH579510	MH579424
**Seira nicoya* Playa Boca del Drago 39_1	MH579488	MH579531	MH579445
**Seira nicoya* Playa Boca del Drago 39_2	MH579489	MH579532	MH579446
**Seira nicoya* Playa Boca del Drago 39_3	MH579490	MH579533	MH579447
**Seira nicoya* Playa Boca del Drago 39_4	MH579491	MH579534	MH579448
**Seira nicoya* Playa Boca del Drago 39_5	MH579492	MH579535	MH579449
**Seira nicoya* Playa Boca del Drago 39_6	MH579493	MH579536	MH579450
**Seira nicoya* Playa Diablo 46_1	MH579498	MH579541	MH579455
**Seira nicoya* Playa Diablo 46_3	MH579499	MH579542	MH579456
**Seira nicoya* Playa Tortuguila 30_1	MH579487	MH579530	MH579444
**Seira nicoya* Punta Chame 25_1	MH579475	MH579518	MH579432
**Seira nicoya* Punta Chame 25_2	MH579476	MH579519	MH579433
**Seira nicoya* Punta Chame 25_3	MH579477	MH579520	MH579434
**Seira nicoya* Punta Chame 25_4	MH579478	MH579521	MH579435
**Seira nicoya* Punta Chame 25_5	MH579479	MH579522	MH579436
**Seira nicoya* Torio 40_1	MH579494	MH579537	MH579451
**Seira nicoya* Torio 40_2	MH579495	MH579538	MH579452
**Seira nicoya* Torio 40_3	MH579496	MH579539	MH579453
**Seira nicoya* Torio 40_4	MH579497	MH579540	MH579454
Additional ingroups			
Seira barnardi (South Africa)	KU508096	KC236296	—
Seira delamarei (China)	KM978370	KC236292	—
**Seira delamarei* 7_1	MH579463	MH579506	MH579420
**Seira delamarei* 7_2	MH579464	MH579507	MH579421
**Seira delamarei* 7_3	MH579465	MH579508	MH579422
**Seira delamarei* 7_4	MH579466	MH579509	MH579423
Seira dowlingi (USA)	KM610133	—	—
**Seira* sp. 1 32_1	MH579500	MH579543	MH579457
**Seira* sp. 1 32_2	MH579501	MH579544	MH579458
**Seira* sp. 2 11_1	MH579502	MH579545	MH579459
**Seira* sp. 2 11_2	MH579503	MH579546	MH579460
**Seira* sp. 2 17_2	MH579504	MH579547	MH579461
**Seira* sp. 2 17_3	MH579505	MH579548	MH579462

(b) Taxa (Isotomoidea dataset)	COI
Outgroups	
**Spinactaletes* sp. A 38_1	MH579576
**Spinactaletes* sp. A 38_2	MH579577
**Spinactaletes* sp. A 38_3	MH579578
**Spinactaletes* sp. A 38_4	MH579579
**Spinactaletes* sp. A 38_5	MH579580
**Spinactaletes* sp. B 2_1	MH579581
**Spinactaletes* sp. B 2_2	MH579582
**Spinactaletes* sp. B 43_1	MH579583
**Spinactaletes* sp. B 43_2	MH579584
**Spinactaletes* sp. B 43_3	MH579585
**Spinactaletes* sp. B 43_4	MH579586
**Spinactaletes* sp. B 43_5	MH579587
Focal taxa	
**Archisotoma* sp. A Fort Sherman 29_1	MH579552
**Archisotoma* sp. A Fort Sherman 29_2	MH579553
**Archisotoma* sp. A Fort Sherman 29_3	MH579554
**Archisotoma* sp. A Nombre de Dios 1_2	MH579549
**Archisotoma* sp. A Nombre de Dios 1_3	MH579550
**Archisotoma* sp. A Nombre de Dios 1_4	MH579551
**Archisotoma* sp. A Playa Tortuguilla	MH579555
**Archisotoma* sp. A Playa Tortuguilla	MH579556
**Archisotoma* sp. A Playa Tortuguilla	MH579557
**Archisotoma* sp. A Torio 40_1	MH579558
**Archisotoma* sp. A Torio 40_2	MH579559
**Axelsonia tubifera* Boca Brava 18_1	MH579565
**Axelsonia tubifera* Boca Brava 18_3	MH579566
**Axelsonia tubifera* Boca Brava 18_4	MH579567
**Axelsonia tubifera* Playa Diablo 44_1	MH579568
**Psammisotoma dispar* Fort Sherman	MH579573
**Psammisotoma dispar* Nombre de Dios	MH579569
**Psammisotoma dispar* Nombre de Dios	MH579570
**Psammisotoma dispar* Playa Tortuguilla	MH579574
**Psammisotoma dispar* Playa Tortuguilla	MH579575
**Psammisotoma dispar* Punta Chame	MH579571
**Psammisotoma dispar* Punta Chame	MH579572
Additional ingroups	
**Archisotoma* sp. B 35_1	MH579560
**Archisotoma* sp. B 35_2	MH579561
**Archisotoma* sp. B 35_3	MH579562
**Archisotoma* sp. B 35_4	MH579563
**Archisotoma* sp. B 35_5	MH579564
Archisotoma besselsi (Russia)	JN981071
Archisotoma Polaris (Canada)	AY665342
Archisotoma pulchella (UK)	KT808327
Cryptopygus antarcticus	NC_010533
Cryptopygus terranovus	KX863671
Folsomia candida	KU198392
Folsomotoma octooculata	NC_024155

Table 2.3. List of all samples collected for marine littoral specimens sequenced for this study, including sample ID#, coordinates, habitat, date collected, and number of marine littoral specimens sequenced for each genus in each sample.

Sample site	Sample ID#	Coordinates N	Coordinates W	Habitat	Date collected	*Seira*	*Archisotoma*	*Axelsonia*	*Psammisotoma*	*Spinactaletes*
Boca Brava	18	8.21249	82.21021	mangrove roots	17June2014	0	0	3	0	0
Fort Sherman	29	9.35915	79.94874	organic beach debris	23June2014	7	3	0	1	0
Nombre de Dios	01	9.58639	79.46926	organic beach debris	08June2014	8	3	0	2	0
Nombre de Dios	02	9.59325	79.46914	coralliferous limestone	08June2014	0	0	0	0	2
Playa Boca del Drago	38	9.41432	82.33150	coralliferous limestone	28June2014	0	0	0	0	5
Playa Boca del Drago	39	9.41642	82.32809	organic beach debris	28June2014	6	0	0	0	0
Playa Diablo	43	9.36519	79.96356	coralliferous limestone	01July2014	0	0	0	0	5
Playa Diablo	44	9.36646	79.96286	mangrove roots	01July2014	0	0	1	0	0
Playa Diablo	46	9.36434	79.96419	organic beach debris	01July2014	2	0	0	0	0
Playa Tortuguilla	30	9.32644	80.00322	organic beach debris	23June2014	1	3	0	2	0
Punta Chame	25	8.61415	79.74516	organic beach debris	22June2014	5	0	0	2	0
Torio	35	7.54462	80.95005	organic beach debris	26June2014	0	5	0	0	0
Torio	40	7.54462	80.95005	organic beach debris	30June2014	4	2	0	0	0

Table 2.4. AMOVA results for morphospecies *Seira nicoya* (COI, 28S, ef1a), *Archisotoma* sp. A (COI), and *Psammisotoma dispar* (COI), with haplotypes grouped by (**a**) region only and **(b**) by region and locality. Significance is based on 50,000 permutations: *P < 0.05, **P < 0.01, ***P < 0.001.

(a) Source of variation	df †	SS‡	VC§	V%¶	φst
Seira nicoya COI					
Among regions	1	1358.51	101.66	78.56	0.79***
Within regions	31	859.92	27.74	21.44	
Seira nicoya 28S					
Among regions	1	42.0	3.17	87.32	0.87***
Within regions	31	14.28	0.46	12.68	
Seira nicoya ef1a					
Among regions	1	291.37	22.16	94.73	0.95***
Within regions	31	38.236	1.23	5.27	
Archisotoma sp. A COI					
Among regions	1	366.35	110.09	94.78	0.95***
Within regions	9	54.56	6.06	5.22	
Psammisotoma dispar COI					
Among regions	1	9.27	-0.07	-0.7	-0.01
Within regions	5	47.3	9.46	100.7	

(b) Source of variation	df †	SS‡	VC§	V%¶	φ-statistics
Seira nicoya COI					
Among regions	1	1358.51	90.62	70.82	φct=0.71*
Among localities within regions	5	751.68	33.18	25.93	φsc=0.88***
Within localities	26	108.24	4.16	3.25	φst=0.96***
Seira nicoya 28S					
Among regions	1	42.00	3.00	83.01	φct=0.83*
Among localities within regions	5	12.06	0.53	14.63	φsc=0.86**
Within localities	26	2.21	0.09	2.36	φst=0.98***
Seira nicoya ef1a					
Among regions	1	291.37	21.65	92.80	φct=0.92*
Among localities within regions	5	34.61	1.54	6.60	φsc=0.91**
Within localities	26	3.62	0.14	0.60	φst=0.99***
Archisotoma sp. A COI					
Among regions	1	366.35	109.91	94.67	φct=0.95
Among localities within regions	2	13.89	0.38	0.33	φsc=0.06
Within localities	7	40.67	5.81	5.00	φst=0.95**
Psammisotoma dispar COI					
Among regions	1	9.27	-6.33	-78.41	φct=-0.78
Among localities within regions	2	45.30	13.74	170.15	φsc=0.95
Within localities	3	2.00	0.67	8.26	φst=0.92*

†df, degrees of freedom
‡SS, sum of squares
§VC, variance components
¶V%, percent of variation

Table 2.5. Mantel test results for the identification of correlation between genetic distance and region relative to the Isthmus after controlling for geographic distance in *Seira nicoya*, *Archisotoma* sp. A, and *Psammisotoma dispar*. Significance is based on 100,000 permutations: *P < 0.05, **P < 0.01, ***P < 0.001.

Morphospecies	r value		
	COI	28S	ef1a
Seira nicoya	0.93***	0.88***	0.99***
Archisotoma sp. A	0.97***	—	—
Psammisotoma dispar	-0.08	—	—

Table 2.6. Summary of geminate species criteria satisfied, as indicated by check marks (✓), for all morphospecies collected for this study. Criteria marked (×) indicates a lack of supporting evidence and (—) indicates that the criterion is not applicable.

Criteria	*Seira nicoya*	*Archisotoma* sp. A	*Archisotoma* sp. B	*Axelsonia tubifera*	*Psammisotoma dispar*	*Spinactaletes* sp. A & B
1 Marine littoral specialists	✓	✓	✓	✓	✓	✓
Habitat	beach debris	beach debris	beach debris	mangrove roots	beach debris	coral limestone
2 Transisthmian distribution	✓	✓	×	✓	✓	×
3 Morphologically identical	✓	✓	—	✓[†]	✓	—
4 Reciprocally monophyletic sister taxa	✓	✓	—	✓	×	—
5 Genetic structure across the Isthmus	✓	✓	—	✓[‡]	×	—
COI p-distance (range)	0.186–0.206	0.153–0.159	—	0.151–0.151	0.001–0.028	—
AMOVA (V%)	0.79–0.95	0.95	—	—	-0.01	—
Mantel test (r)	0.88–0.99***	0.97***	—	—	-0.08	—
6 MRCA and Isthmus have similar age	✓	✓	—	✓	×	—
Gene tree (median)	26.55 Ma	25.40 Ma	—	16.07 Ma	1.04 Ma	—
95% HPD	19.54–34.45 Ma	14.72–37.95 Ma	—	9.32–23.90 Ma	0.66–1.57 Ma	—
Species tree (median)	24.76 Ma	18.87 Ma	—	12.90 Ma	1.26 Ma	—
95% HPD	18.75–30.84 Ma	11.21–28.28 Ma	—	7.44–19.44 Ma	0.71–1.96 Ma	—
Closure model	Miocene	Miocene	—	Miocene	×	—
7 Divergence times are similar to codistributed geminate species pairs	✓	✓	—	✓	×	—

[†]Only a single specimen for *Axelsonia tubifera* was collected from Play Diablo, which was destroyed during DNA extraction process. Therefore, I was unable to examine and compare morphology between Pacific and Caribbean localities. Despite small differences in antennal pigmentation, they were otherwise identical under stereomicroscopy.

[‡]Genetic structure was not analyzed for *Axelsonia tubifera* because few specimens were collected from only 2 localities. However, high genetic distances between transisthmian populations (and low intra-isthmian distances) suggests genetic structure across the Isthmus of Panama is likely present.

Figures

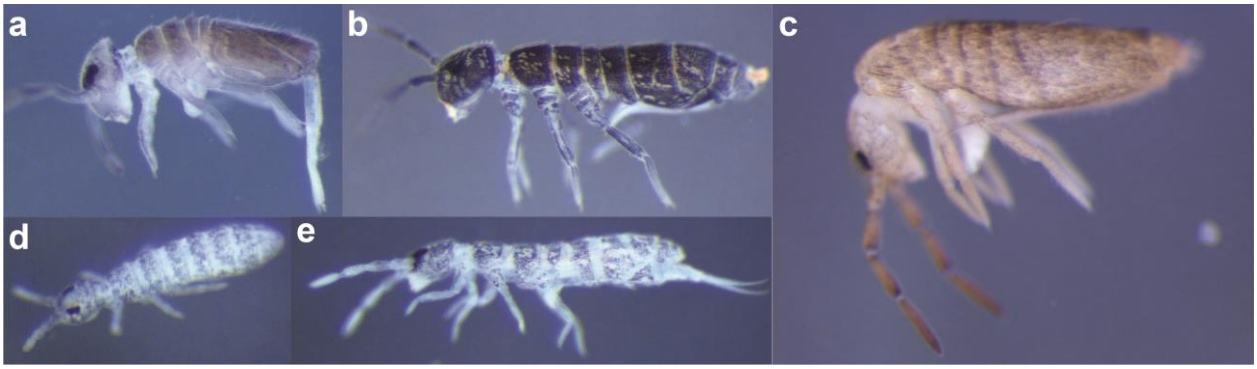


Figure 2.1. (**a**) *Spinactaletes*; (**b**) *Axelsonia*; (**c**) *Seira*; (**d**) *Archisotoma*; and (**e**) *Psammisotoma*; collected from marine littoral habitats in Panama for this study.

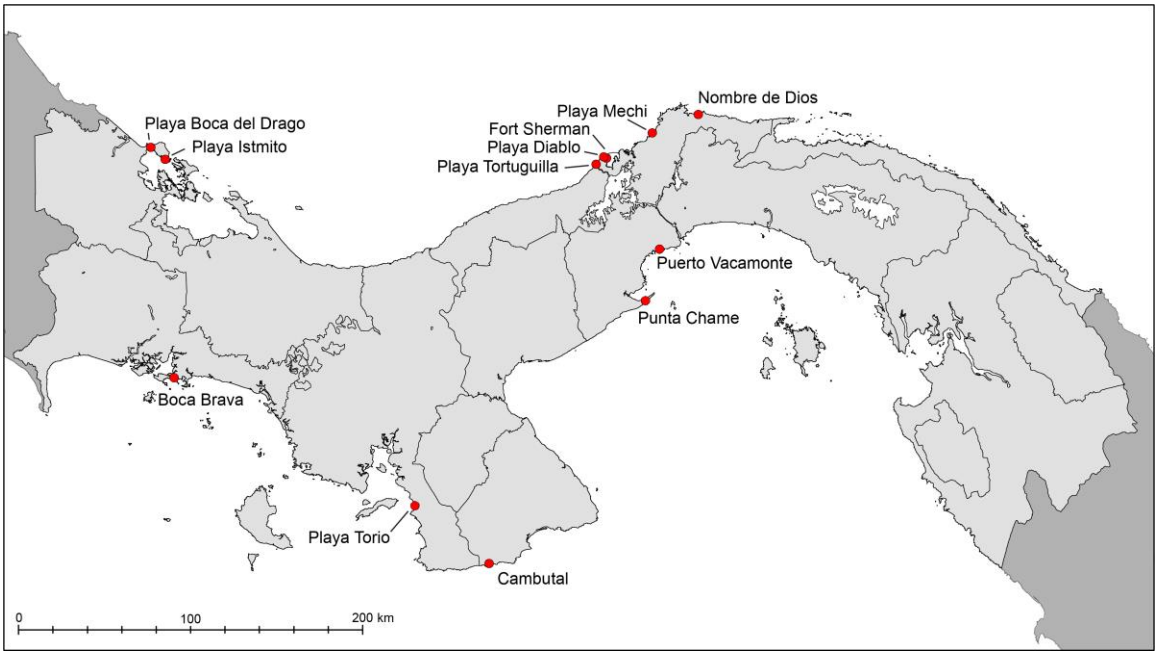


Figure 2.2. Map of Panama (light gray) and intertidal localities sampled for this study (red dots).

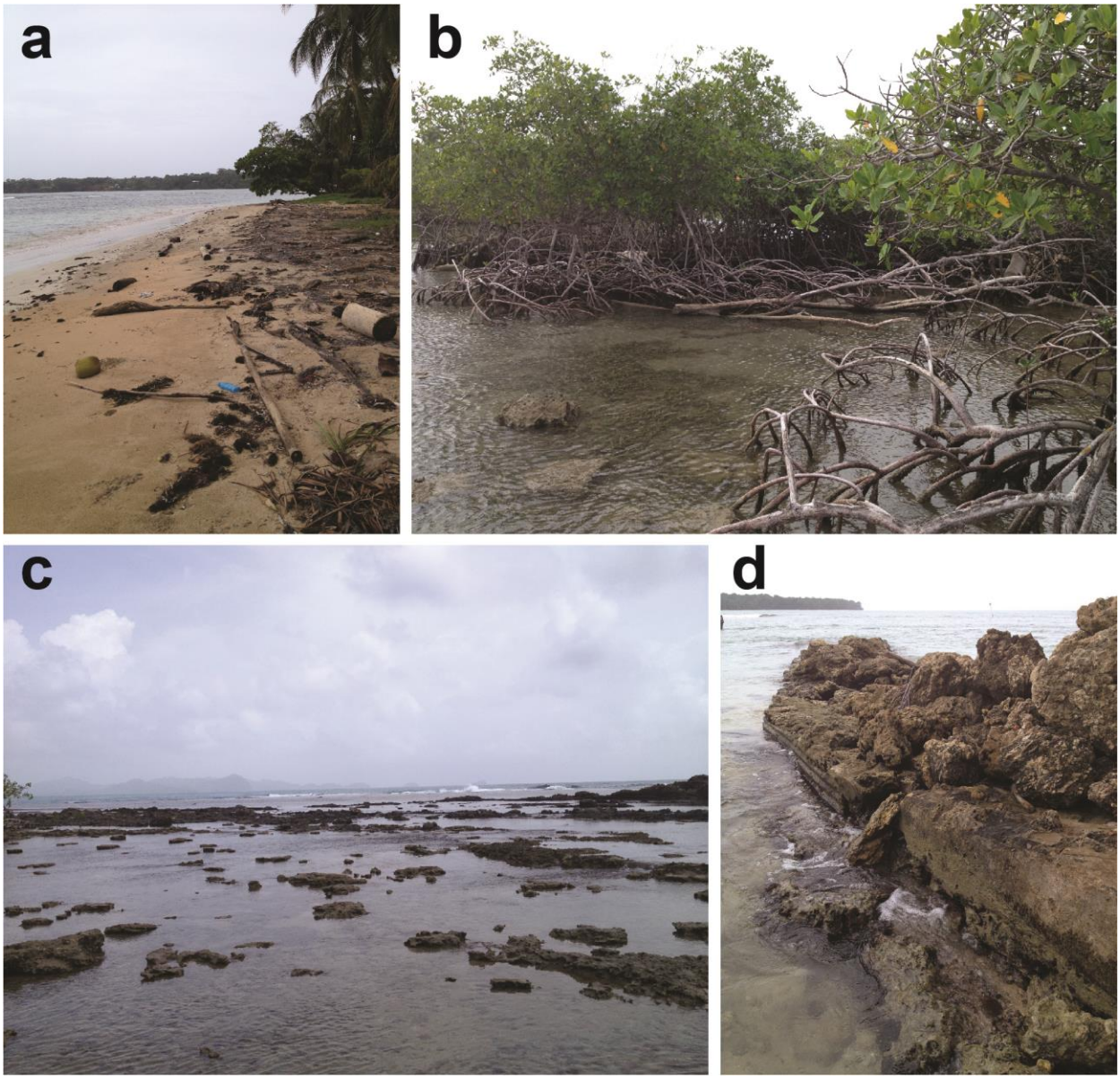


Figure 2.3. Marine littoral habitats at Panama sampling sites: (**a**) seaweed and other organic debris accumulated on sand at the high tide water mark at Playa Boca del Drago, typical habitat for species in the genera *Seira*, *Archisotoma*, and *Psammisotoma*; (**b**) mangrove roots exposed during low tide near Playa Diablo, typical habitat for species in the genus *Axelsonia*; (**c**) coralliferous limestone reef exposed during low tide near Nombre de Dios and (**d**) old docks lined with dredged reef at Playa Boca del Drago, typical habitats for species in the genus *Spinactaletes*.

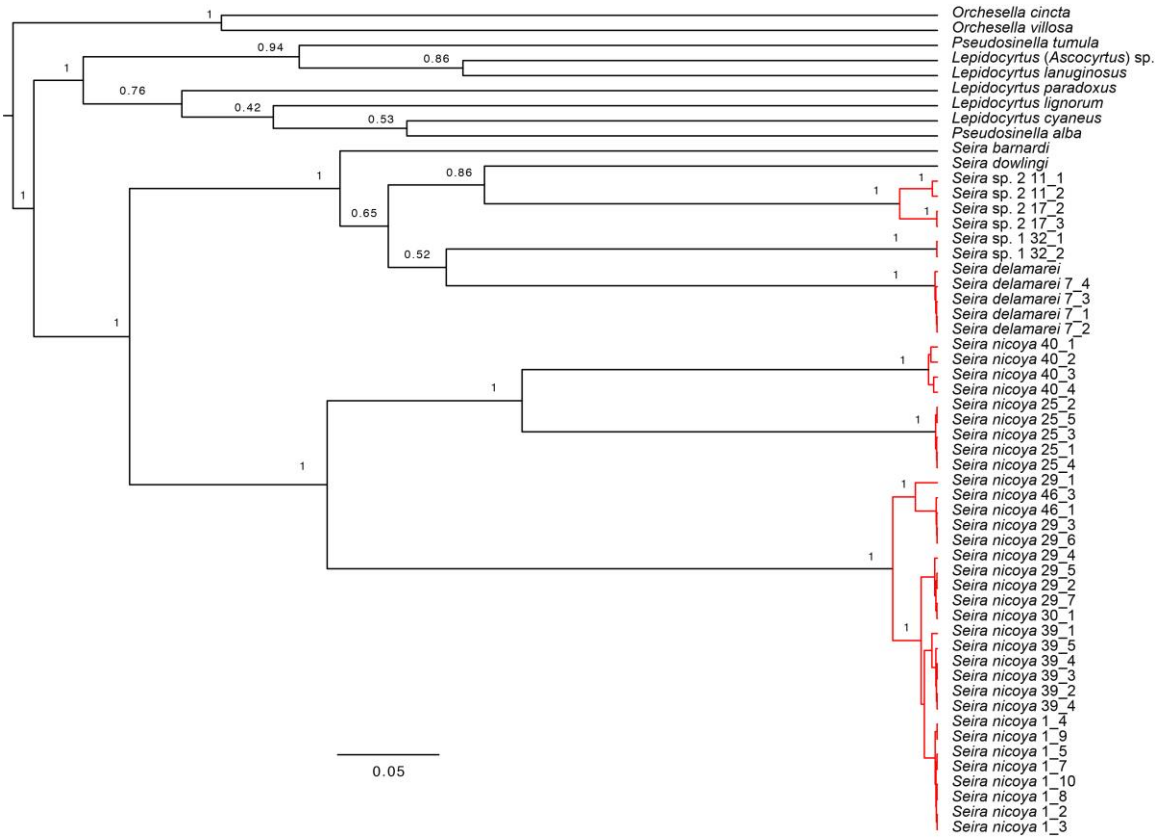


Figure 2.4. Bayesian COI gene tree and GMYC species delimitation results for focal taxa in the genus *Seira* (Superfamily Entomobryoidea). Clade posterior probabilities are indicated at each node. Red branches indicate lineages grouped as OTUs by GMYC analysis. Scale represents substitutions/site/time.

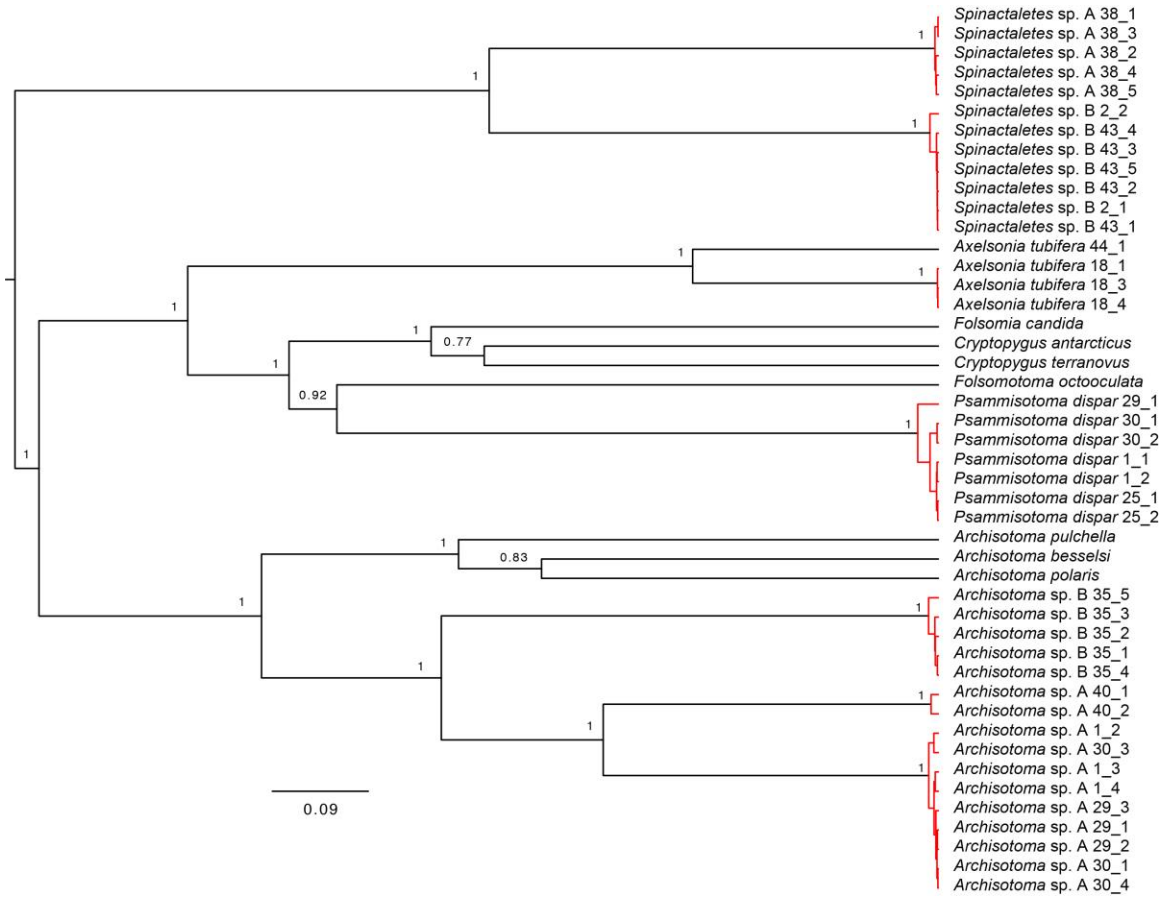


Figure 2.5. Bayesian COI gene tree and GMYC species delimitation results for focal taxa in the Superfamily Isotomoidea. Clade posterior probabilities are indicated at each node. Red branches indicate lineages grouped as OTUs by GMYC analysis. Scale represents substitutions/site/time.

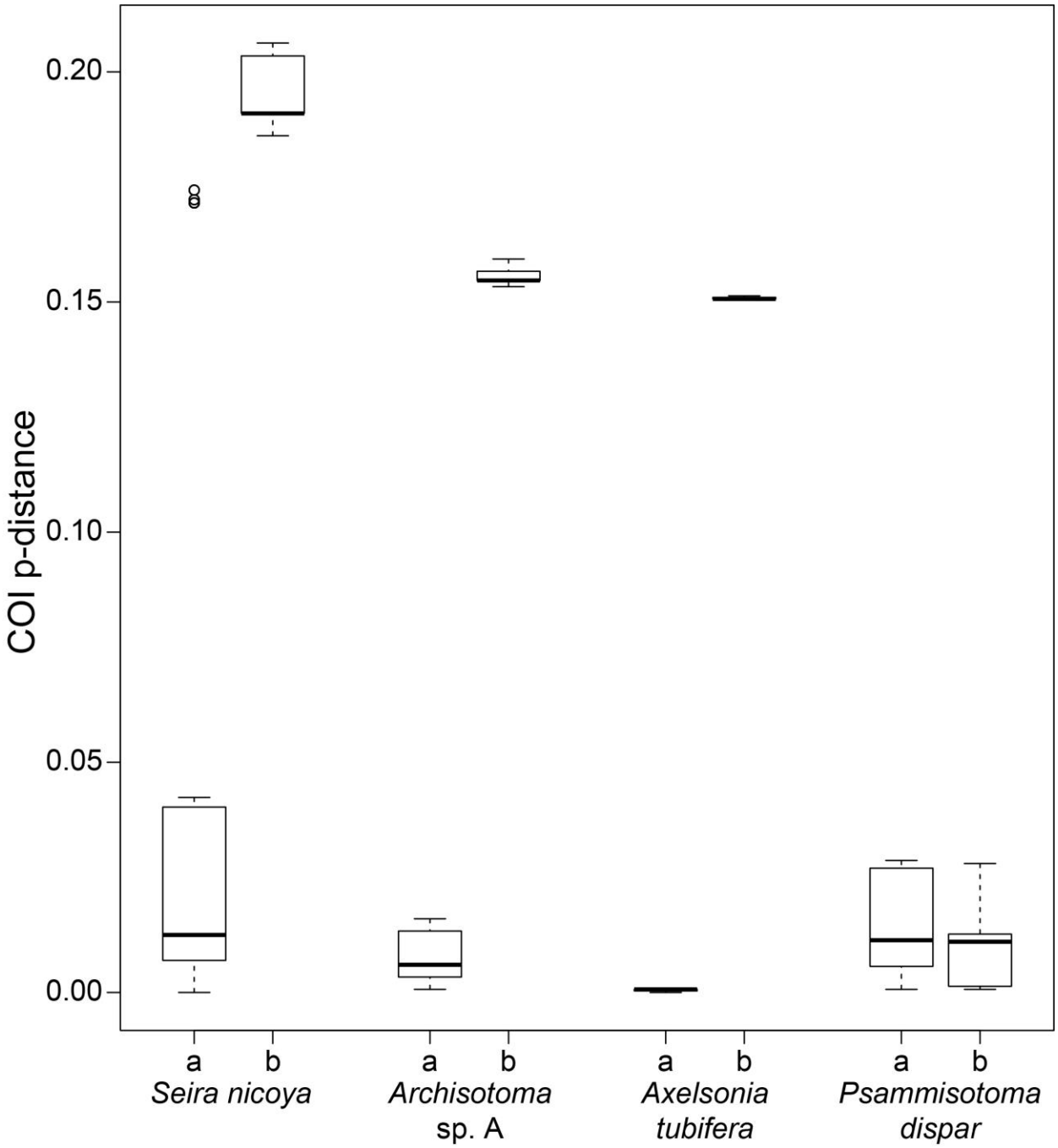


Figure 2.6. Boxplots comparing COI p-distances (uncorrected) estimated between (**a**) taxa occurring on the same side of the Isthmus (intra-isthmian distance) and (**b**) taxa occurring on opposite sides of the Isthmus (inter-isthmian distance).

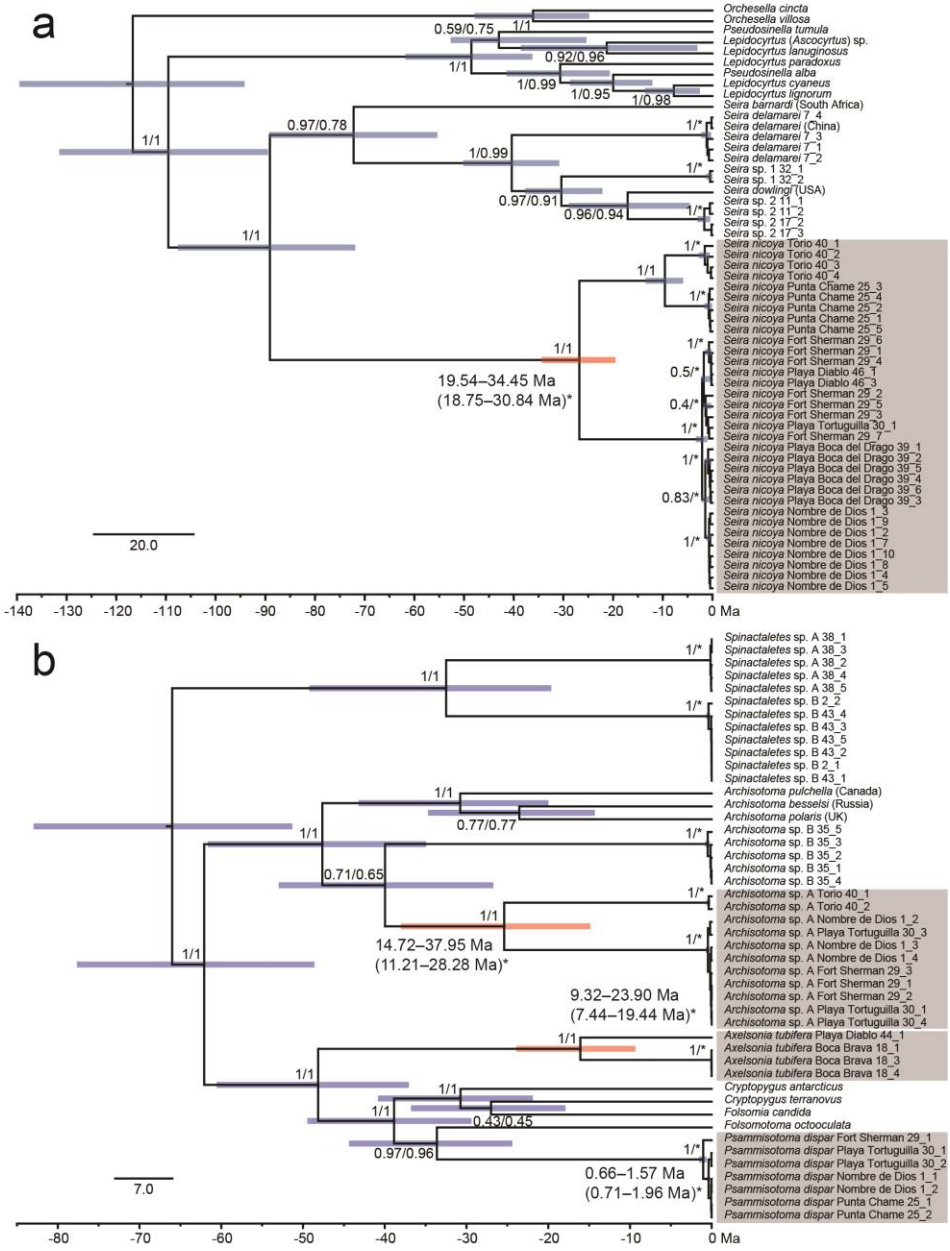


Figure 2.7. Bayesian time-calibrated gene trees for the (**a**) *Seira* dataset and (**b**) Isotomoidea dataset estimated with BEAST (concatenated gene tree approach). Clade posterior probabilities (gene tree/species tree) are indicated at each node. Asterisks in place of species tree support values denote clades within species taxon sets (defined by OTUs from GMYC analysis) used for species tree analyses, thus, support values are not available at these nodes. Divergence times are represented by blue bars at each node, with their lengths corresponding to the 95% HPD of node ages. Red bars represent divergence times for putative geminate species. Divergence time intervals (95% HPD) resulting from BEAST labeled next to divergence time bars for transisthmian morphospecies (species tree estimates in parenthesis with asterisk). Morphospecies with transisthmian distributions are highlighted in gray boxes. Taxon labels for samples collected in Panama and sequenced for this study include a specimen ID (i.e., sample#_specimen#). See Tables 2.2–2.3 for additional sample information. Scale bars in substitutions/site/Ma.

References

Abrantes, E. A. & Mendonça, M. C. de (2007), New species and a new record of Isotomidae (Collembola) from the coast of Brazil. *Zootaxa*, *1500*, 55–60.

Abrantes, E. A. & Mendonça, M. C. de (2009), A new species of *Psammisotoma* Greenslade and Deharveng (Collembola: Isotomidae) from Brazil. *Zootaxa*, *2295*, 25–30.

Ayre, D. J., Minchinton, E., & Perrin, C. (2009). Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Molecular Ecology*, *18*(9), 1887–1903.

Bacon, C. D., Mora, A., Wagner, W. L., & Jaramillo, C. A. (2013). Testing geological models of evolution of the Isthmus of Panama in a phylogenetic framework. *Botanical Journal of the Linnean Society*, *171*(1), 287–300. doi: 10.1111/j.1095-8339.2012.01281.x

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., & Antonelli, A. (2015a). Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences*, *112*(19), 6110–6115. doi: 10.1073/pnas.1423853112

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., & Antonelli, A. (2015b). Reply to Lessios and Marko et al.: Early and progressive migration across the Isthmus of Panama is robust to missing data and biases. *Proceedings of the National Academy of Sciences*, *112*(43), E5767–E5768. doi: 10.1073/pnas.1515451112

Bellinger, P. F. (1962). A dimorphic species of *Actaletes* (Collembola). *Journal of the New York Entomological Society*, *70*, 88–91.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, *41*, D36–D42. doi: 10.1093/nar/gks1195

Blackith, R. E. & Disney, R. H. L. (1988). Passive aerial dispersal during moulting in tropical Collembola. *Malayan Nature Journal*, *41*(4), 529–531.

Bonnet, E. & Van de Peer, Y. (2002). zt: a software tool for simple and partial Mantel tests. *Journal of Statistical Software*, *7*(10), 1–12. doi: 10.18637/jss.v007.i10

Börner, C. (1906). Das System der Collembolen nebst Beschreibung neuer Collembolen des Hamburger Naturhistorischen Museums. *Mitteilungen aus dem Naturhistorischen Museum in Hamburg*, *23*, 147–188.

Bouckaert, R. R. & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, *17*(42), 1–11. doi: 10.1186/s12862-017-0890-6

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. doi: 10.1371/journal.pcbi.1003537

Brower, A. V. Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences*, *91*(14), 6491–6495. doi: 10.1073/pnas.91.14.6491

Christiansen, K. & Bellinger, P. (1988). Marine Littoral Collembola of North and Central America. *Bulletin of Marine Science*, *42*(2), 215–245.

Christiansen, K. & Bellinger, P. (1998). *The Collembola of North America north of the Rio Grande; A taxonomic analysis*. 2nd Edition. Grinnell College: Grinnell.

Christiansen, K. & Bellinger, P. (2000). A Survey of the Genus Seira (Hexapoda: Collembola: Entomobryidae) in the Americas. *Caribbean Journal of Science*, *36*(1–2), 39–75.

Cicconardi, F., Nardi, F., Emerson, B. C., Frati, F., & Fanciulli, P. P. (2010). Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology*, *19*(2), 386–400. doi: 10.1111/j.1365-294X.2009.04457.x

Cicconardi, F., Fanciulli, P. P., & Emerson, B.C. (2013). Collembola, the biological species concept and the underestimation of global species richness. *Molecular ecology*, *22*(21), 5382–5396. doi: 10.1111/mec.12472

Coates, A. G., Collins, L. S., Aubry, M.-P., & Berggren, W. A. (2004). The Geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. *Geological Society of America Bulletin*, *116*(11–12), 1327–1344. doi: 10.1130/B25275.1

Coates, A. G. & Obando, J. A. (1996). The geologic evolution of the Central American Isthmus. *Evolution and environment in tropical America* (ed. by J.B.C. Jackson, A.F. Budd, and A.G. Coates), pp. 21–56. The University of Chicago: Chicago.

Collins, T. (1996). Molecular comparisons of transisthmian species pairs: rates and patterns of evolution. *Evolution and environment in tropical America* (ed. by J.B.C. Jackson, A.F. Budd, and A.G. Coates), pp. 303–334. The University of Chicago: Chicago.

Costa, D., Timmermans, M. J. T. N., Sousa, J. P., Ribeiro, R., Roelofs, D., & Van Straalen, N. M. (2013). Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. *Applied Soil Ecology*, *68*, 61–66. doi: 10.1016/j.apsoil.2013.03.003

Coulson, S. J., Hodkinson, I. D., & Webb, N. R. (2003). Aerial dispersal of invertebrates over a high-Arctic glacier foreland: Midtre Lovénbreen, Svalbard. *Polar Biology*, *26*(8), 530–537. doi: 10.1007/s00300-003-0516-x

Coulson, S. J., Hodkinson, I. D., Webb, N. R., & Harrison, J. A. (2002). Survival of terrestrial soil-dwelling arthropods on and in seawater: implications for trans-oceanic dispersal. *Functional Ecology*, *16*(3), 353–356. doi: 10.1046/j.1365-2435.2002.00636.x

Delamare Deboutteville, C. (1956). Études sur la faune interstitielle des iles Bahamas récoltée par madame Renaud-Debyser. II. Un Nouveau Collembole Marin. *Vie et Milieu*, *7*(3), 397–399.

Drummond, A. J. & Bouckaert, R. R. (2015). *Bayesian evolutionary analysis with BEAST*. Cambridge University Press: Cambridge.

Elmer, K. R., Bonett, R. M., Wake, D. B., & Lougheed, S. C. (2013). Early Miocene origin and cryptic diversification of South American salamanders. *BMC Evolutionary Biology*, *13*(1), 59. doi: 10.1186/1471-2148-13-59

Excoffier, L. & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Ezard, T., Fujisawa, T., & Barraclough, T. G. (2009). Package 'splits: SPecies' LImits by Threshold Statistics' version 1.0-19. Available from http://R-Forge.R-project.org/projects/splits/

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. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299.

Freeman, J. A. (1952). Occurrence of Collembola in the air. *Proceedings of the Royal Entomological Society of London. Series A, General Entomology*, *27*(1–3), 28. doi: 10.1111/j.1365-3032.1952.tb00142.x

Garrick, R. C., Rowell, D. M., Simmons, C. S., Hillis, D. M., & Sunnucks, P. (2008). Fine-scale phylogeographic congruence despite demographic incongruence in two low-mobility saproxylic springtails. *Evolution*, *62*(5), 1103–1118. doi: 10.1111/j.1558-5646.2008.00349.x

Greenslade, P. & Deharveng, L. (1986). *Psammisotoma*, a new genus of Isotomidae (Collembola) from marine littoral habitats. *Proceedings of the Royal Society of Queensland*, *97*, 89–95.

Hawes, T. C., Worland, M. R., Bale, J. S., & Convey, P. (2008). Rafting in Antarctic Collembola. *Journal of Zoology*, *274*(1), 44–50. doi: 10.1111/j.1469-7998.2007.00355.x

Hawes, T. C., Worland, M. R., Convey, P., & Bale, J. S. (2007). Aerial dispersal of springtails on the Antarctic Peninsula: implications for local distribution and demography. *Antarctic Science*, *19*(1), 3–10. doi: 10.1017/S0954102007000028

Haynert, K., Kiggen, M., Klarner, B., Maraun, M., & Scheu, S. (2017). The structure of salt marsh soil mesofauna food webs – The prevalence of disturbance. *PLOS ONE*, *12*(12), e0189645. doi: 10.1371/journal.pone.0189645

Hoorn, C. & Flantua, S. (2015). An early start for the Panama land bridge. *Science*, *348*(6231), 186–187. doi: 10.1126/science.aab0099

Jackson, J. B. C. & O'Dea, A. (2013). Timing of the oceanographic and biological isolation of the Caribbean Sea from the tropical eastern Pacific Ocean. *Bulletin of Marine Science*, *89*(4), 779–800. doi: 10.5343/bms.2012.1096

Jaramillo, C., Montes, C., Cardona, A., Silvestro, D., Antonelli, A., & Bacon, C. D. (2017). Comment (1) on "Formation of the Isthmus of Panama" by O'Dea et al . *Science Advances*, *3*(6), e1602321. doi: 10.1126/sciadv.1602321

Johnson, M. E. & Baarli, B. G. (2012). Development of intertidal biotas through Phanerozoic time. In *Earth and Life* (pp. 63–128). Springer: Netherlands.

Joosse, E. N. G. (1976). Littoral apterygotes (Collembola and Thysanura). *Marine Insects* (ed. by L. Cheng), pp. 151–186. North-Holland Publishing Company: Amsterdam-Oxford.

Jordan, D. S. (1908). The law of geminate species. *The American Naturalist*, *42*(494), 73–80.

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

Katz, A. D., Giordano, R., & Soto-Adames, F. N. (2015). Operational criteria for cryptic species delimitation when evidence is limited, as exemplified by North American *Entomobrya* (Collembola: Entomobryidae). *Zoological Journal of the Linnean Society*, *173*(4), 810–840. doi: 10.1111/zoj.12220

Katz, A. D., Taylor, S. J., & Davis, M. A. (2018). At the confluence of vicariance and dispersal: phylogeography of cavernicolous springtails (Collembola: Arrhopalitidae, Tomoceridae) codistributed across a geologically complex karst landscape in Illinois and Missouri. *Ecology and Evolution*, In Press.

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*(12), 1647–1649. doi: 10.1093/bioinformatics/bts199

Knowlton, N. & Weigt, L. A. (1998). New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London. Series B, Biological sciences*, *265*(1412), 2257–2263. doi: 10.1098/rspb.1998.0568

Knowlton, N., Weigt, L. A., Solórzano, L. A., Mills, D. K., & Bermingham, E. (1993). Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science*, *260*(5114), 1629–1632. doi: 10.1126/science.8503007

Kodandaramaiah, U. (2011). Tectonic calibrations in molecular dating. *Current Zoology*, *57*(1), 116–124. doi: 10.1093/czoolo/57.1.116

Lapointe, F. & Rissler, L. J. (2005). Congruence, Consensus, and the Comparative Phylogeography of Codistributed Species in California. *The American Naturalist*, *166*(2), 290–299. doi: 10.1086/431283

Leigh, E. G., O'Dea, A., & Vermeij, G. J. (2014). Historical biogeography of the Isthmus of Panama. *Biological Reviews*, *89*(1), 148–172. doi: 10.1111/brv.12048

Leinaas, H. P. & Ambrose Jr, W. G. (1992). Utilization of different foraging habitats by the purple sandpiper *Calidris maritima* on a Spitsbergen beach. *Fauna Norvegica, Series C*, 15(2), 85–91.

Lessios, H. A. (2008). The Great American Schism: Divergence of Marine Organisms After the Rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 63–91. doi: 10.1146/annurev.ecolsys.38.091206.095815

Lessios, H. A. (2015). Appearance of an early closure of the Isthmus of Panama is the product of biased inclusion of data in the metaanalysis. *Proceedings of the National Academy of Sciences*, *112*(43), E5765–E5765. doi: 10.1073/pnas.1514719112

Linnaniemi, W. M. (1912). Die Apterygotenfauna Finlands, II. Spezieller Teil. *Acta Societatis Scientiarum Fennicæ*, *40*(5), 1–359.

Lubbock, J. (1870). Notes on the Thysanura. Part IV. *Transactions of the Linnean Society of London (Zoology)*, *27*(2), 277–297.

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, *27*(2), 209–220. doi: 10.1002/stem.684.

Mari-Mutt, J. A. (1979). A Revision of the Genus Dicranocentrus Schött (Insecta: Collembola: Entomobryidae). *Bulletin of the University of Puerto Rico*, *259*, 1–79.

Mari Mutt, J. A. & Bellinger, P. F. (1990). *A catalog of the Neotropical Collembola, including Nearctic areas of Mexico.* Sandhill Crane Press: Gainesville.

Marko, P. B. (2002). Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, *19*(11), 2005–2021. doi: 10.1093/oxfordjournals.molbev.a004024

Marko, P. B., Eytan, R. I., & Knowlton, N. (2015). Do large molecular sequence divergences imply an early closure of the Isthmus of Panama? *Proceedings of the National Academy of Sciences*, *112*(43), E5766–E5766. doi: 10.1073/pnas.1515048112

Marko, P. B. & Moran, A. L. (2009). Out of sight, out of mind: high cryptic diversity obscures the identities and histories of geminate species in the marine bivalve subgenus Acar. *Journal of Biogeography*, *36*(10), 1861–1880.

Miura, O., Torchin, M. E., & Bermingham, E. (2010). Molecular phylogenetics reveals differential divergence of coastal snails separated by the Isthmus of Panama. *Molecular Phylogenetics and Evolution*, *56*(1), 40–48. doi: 10.1016/j.ympev.2010.04.012

Miura, O., Torchin, M. E., Bermingham, E., Jacobs, D. K., & Hechinger, R. F. (2012). Flying shells: historical dispersal of marine snails across Central America. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1731), 1061–1067. doi: 10.1098/rspb.2011.1599

Molnar, P. (2017). Comment (2) on "Formation of the Isthmus of Panama" by O'Dea et al. *Science Advances*, *3*(6), e1602320. doi: 10.1126/sciadv.1602320

Montes, C., Bayona, G., Cardona, A., Buchs, D. M., Silva, C. A., Morón, S., Hoyos, N., Ramírez, D. A., Jaramillo, C. A., & Valencia, V. (2012a). Arc-continent collision and orocline formation: Closing of the Central American seaway. *Journal of Geophysical Research: Solid Earth*, *117*(4), 1–25. doi: 10.1029/2011JB008959

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, V., Ayala, C., Pérez-Angel, L. C., Rodriguez-Parra, L. A., Ramirez, V., & Niño, H. (2015). Middle Miocene closure of the Central American Seaway. *Science*, *348*(6231), 226–229. doi: 10.1126/science.aaa2815

Montes, C., Cardona, A., McFadden, R., Morón, S. E., Silva, C. A., Restrepo-Moreno, S., Ramírez, D. A., Hoyos, N., Wilson, J., Farris, D., Bayona, G. A., Jaramillo, C. A., Valencia, V., Bryan, J., & Flores, J. A. (2012b). Evidence for middle Eocene and younger land emergence in central Panama: Implications for Isthmus closure. *Geological Society of America Bulletin*, *124*(5–6), 780–799. doi: 10.1130/B30528.1

Najt, J. & Rapoport, E. H. (1972). Una nueva especie de Actaletidae de Venezuela (Insecta, Collembola). *Physis*, *31*(82), 219–221.

O'Dea, A., Lessios, H. A., Coates, A. G., et al. (2016). Formation of the Isthmus of Panama. *Science Advances*, *2*(8), e1600883. doi: 10.1126/sciadv.1600883

Ogilvie, H. A., Bouckaert, R. R., & Drummond, A. J. (2017). StarBEAST2 Brings Faster Species Tree Inference and Accurate Estimates of Substitution Rates. *Molecular Biology and Evolution*, *34*(8), 2101–2114. doi: 10.1093/molbev/msx126

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: The mid-aegean trench calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. doi: 10.1093/molbev/msq051

Poinsot, N. (1965). Révision du genre *Archisotoma* Linnaniemi 1912. *Revue d'Écologie et de Biologie du Sol*, *2*(3), 453–459.

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Tracer v1.7, Available from http://tree.bio.ed.ac.uk/software/tracer/

Riascos, J. M. (2002). Changes in the macrobenthos of a sandy beach during "El Niño" 1997–98 in the Malaga Bay, Colombian Pacific. *Ciencias Marinas*, *28*(1), 13–25.

Roy, M. S. & Sponer, R. (2002). Evidence of a human-mediated invasion of the tropical western Atlantic by the "world"s most common brittlestar'. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1495), 1017–1023. doi: 10.1098/rspb.2002.1977

Samalot Roque, B. (2006). *Diversidad de Collembola (Hexapoda) asociados a Rhizophora mangle en manglares de Puerto Rico.* M.S. Thesis, Department of Biology, Unisersidad de Puerto Rico.

Schött, H. (1896). North American Apterygogenea. *Proceedings of the California Academy of Sciences, Series 2*, *6*, 169–196.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, *87*(6), 651–700.

Soto-Adames, F. N. (1987). A new species of *Actaletes* from Mexico (Collembola: Actaletidae). *The Pan-Pacific entomologist*, *63*(1), 52–55.

Soto-Adames, F. N. (1988). Revisión de la familia Actaletidae Borner, 1902 (Insects: Collembola). *Caribbean Journal of Science*, *24*(3–4), 161–196.

Soto-Adames, F. N. & Guillén, C. (2011). Two new species of the marine littoral springtail genus *Spinactaletes* (Collembola: Actaletidae) from Costa Rica. *Bulletin of Marine Science*, *87*(3), 463–483. doi: 0.5343/bms.2010.1078

Stange, M., Sánchez-Villagra, M. R., Salzburger, W., & Matschiner, M. (2018). Bayesian divergence-time estimation with genome-wide SNP data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. *Systematic Biology*, 1–19. doi: 10.1101/102129

Stehli, F. G., & Webb, S. D. (1985). *The Great American Biotic Interchange. Topics in Geobiology*. Plenum Press: New York.

Strenzke, K. (1955). Thalassobionte und thalassophile Collembola. *Die Tierwelt der Nord- und Ostsee*, *36*(11), 1–52.

Strenzke, K. (1958). *Axelsonia tubifera* n. sp., ein neuer arthropleoner Collembole mit Geschlechtsdimorphismus aus der brasilianischen Mangrove. *Acta Zoologica Cracoviensia*, *2*(26), 608–619.

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

Thacker, C. E. (2017). Patterns of divergence in fish species separated by the Isthmus of Panama. *BMC Evolutionary Biology*, *17*(1), 111. doi: 10.1186/s12862-017-0957-4

Thibaud, J.-M. (1993). Les Collemboles des Petites Antilles. VI. Interstitiels Terrestres et Marins. *Revue française d'Entomologie*, *15*(2), 69–80.

Thibaud, J.-M. & Palacios-Vargas, J. G. (2001). Révision du genre *Archisotoma* Linnaniemi, 1912 (Collembola: Isotomidae). *Annales de la Société Entomologique de France*, *37*(3), 347–356.

Vázquez, M. M. & Palacios-Vargas, J. G. (1990). Nuevos registros y aspectos biogeográficos de los colémbolos de la Sierra de la Laguna, B.C.S., México. *Folia Entomológica Mexicana*, *78*, 5–22.

Wilke, T., Schultheiß, R., & Albrecht, C. (2009). As Time Goes by: A Simple Fool's Guide to Molecular Clock Approaches in Invertebrates. *American Malacological Bulletin*, *27*(1–2), 25–45. doi: 10.4003/006.027.0203

Winston, M. E., Kronauer, D. J. C., & Moreau, C. S. (2017). Early and dynamic colonization of Central America drives speciation in Neotropical army ants. *Molecular Ecology*, *26*(3), 859–870. doi: 10.1111/mec.13846

CHAPTER 3

Relative rates of molecular evolution in Class Hexapoda (Arthropoda): Do springtails (Collembola) have accelerated rates of mitochondrial DNA substitution?

Abstract

External rates of molecular evolution derived from independent taxa are often used to estimate divergence times for taxa that lack useful independent information for molecular clock calibrations. A major limitation of this approach is that molecular rates can vary considerably among taxonomic groups and few studies evaluate whether an external clock calibration is appropriate for the study group prior to its application. Pervasive cryptic diversification uncovered in Collembola—a group of highly abundant hexapods that lack informative information for molecular clock calibration—may be linked to accelerated rates of molecular evolution. Therefore, applications of external clocks estimated for insects may produce misleading estimates of evolutionary timescales. Here I attempt to establish whether an external "insect" clock for COI is appropriate for estimating evolutionary timescales for Collembola. I conducted a Bayesian phylogenetic analysis using a fixed tree topology and a relaxed molecular clock with a mean rate set to 1 to estimate relative rates of the COI gene for nearly all orders of Hexapoda that were then compared to identify groups with elevated rates of COI substitution and determine if their rates differ significantly from those of Collembola. I found that substitution rates in Collembola do not significantly differ from most hexapod groups, suggesting that external insect clocks for COI are appropriate for estimating evolutionary timescales for Collembola. However, a few hexapod groups exhibited accelerated rates in COI, up to 3 times faster than other arthropods, underscoring the need for caution when applying

external molecular clocks to estimate divergence times. Lastly, this study exemplifies a relatively simple method for evaluating rate constancy within a taxonomic group to determine whether the use of an external "universal" clock is appropriate.

Introduction

The concept of the molecular clock has revolutionized the field of evolutionary biology by providing a foundation for evaluating the tempo of biological processes and mechanisms shaping patterns of biodiversity (Ho & Duchêne, 2014). Molecular clocks have been widely implemented to estimate divergence times, determine evolutionary rates, and to assess biogeographic hypotheses. But in practice, this powerful statistical tool requires independent information to calibrate rates and timescales into units of absolute time (Wilke et al., 2009). The most common approach for calibrating molecular clocks is to constrain the minimum age of phylogenetic relationships to dates derived from the fossil record (Hipsley & Müller, 2014) or to the timing of biogeographic events associated with lineage divergence (e.g., Papadopoulou et al., 2010). However, for most small, soft-bodied organisms, informative fossils for calibration are generally unavailable. Furthermore, the use of biogeographic calibrations requires strong assumptions regarding the connection between phylogenetic and biogeographic events, as well as the age of the biogeographic event itself (Simon et al., 2015), information that is often unreliable at best, and unavailable at worst. In cases where informative fossil and biogeographic information are unavailable, external clocks, or "universal" substitution rates derived from previous studies, can be used to fix evolutionary rates or inform a prior rate distribution for divergence time estimation (e.g., Marshall et al., 2016). External

clocks have been widely implemented in the absence of independent timing information (Hipsley & Müller, 2014). However, because rates of genetic change can vary considerably among organisms and genes (Ho & Duchêne, 2014), rate constancy, a fundamental assumption for external clocks, is easily violated (Wilke et al., 2009). External clocks estimated for the taxa of interest (or for closely related taxa) are ideal because rates are more likely to be similar (Weir & Schluter, 2008), but taxon-specific molecular clocks are unavailable for most organisms, and "universal" clocks for larger groups, like invertebrates, do not meet the assumption of rate constancy (Thomas et al., 2006).

Collembola present an interesting, yet challenging, case of exploring evolutionary hypotheses for taxa without available node calibrations or external taxon-specific molecular clocks. These minute, insect-like microarthropods, commonly called springtails, are among the most abundant organisms in the soil mesofauna, occupy nearly all terrestrial ecosystems and habitats, and serve vital roles in detrital decomposition, food webs, and soil structure (Hopkin, 1997). Despite their ecological importance and ubiquity in nature, springtail diversity is poorly understood, and to date, less than 2% of species are thought to have been described (Cicconardi et al., 2013). The diversity deficit in springtails (and other members of the soil mesofauna) has been attributed to a lack of scientific priority, inadequate taxonomic resolution used in soil habitat research, and geographic under-sampling (André, 2002). Cryptic species have further confounded our current understanding of biological diversity (Fišer et al., 2018), as is particularly evident in Collembola, within which deeply divergent, yet morphologically identical lineages are routinely discovered and often have molecular distances equivalent to those between

genera or even families of other arthropods (e.g., Frati et al., 1994; 1995; 2000; Carapelli et al., 1995; 2005; Simonsen et al., 1999; Soto-Adames, 2002; Timmermans et al., 2005; Garrick et al., 2007; 2008; Cicconardi et al., 2010; 2013; Felderhoff et al., 2010; Emerson et al., 2011; Porco et al., 2012a; 2012b; Katz et al., 2015; 2018). The existence of apparent widespread cryptic diversity in springtails has recently prompted global species richness estimates of more than 500,000 species (Cicconardi et al., 2013), but the evolutionary processes driving these patterns of cryptic diversity have not yet been explored for this group.

Patterns of morphological stasis, or 'cryptic speciation', can be driven by directional selection on non-morphological traits such as sexual recognition, physiology, or behavior (Bickford et al., 2007), and/or stabilizing selection resulting from extreme, and possibly invariant, selection pressures (Rothschild & Mancinelli, 2001), but it remains unclear as to why this discordance between genetic and morphological variation is so prevalent in springtails. Populations may have diversified at the molecular level without morphological change over long periods of time, possibly due to invariant selection pressures on morphology (Sánchez-García & Engel, 2016). The fossil record supports this 'ancient relic' hypothesis, with many examples of fossils being placed into extant families or genera (Mari Mutt, 1983; Christiansen & Pike, 2002; Christiansen & Nascimbene, 2006; Sánchez-García & Engel, 2016). Alternatively, species with highly genetically structured populations, as is common among springtails, may have an increased probability of mutations reaching fixation, which can lead to accelerated rates of molecular evolution (Frean et al., 2013). For example, deep phylogeographic structure, large genetic distances, rapid speciation, and accelerated rates of molecular evolution

have been attributed to founder effects in parasitoid wasps (Kaltenpoth et al., 2012) and *Drosophila* flies (DeSalle & Templeton, 1988). Wessel et al. (2013) also attributes the rapid radiation and cryptic speciation of Hawaiian cave planthoppers to this phenomenon. These studies suggest that low mobility and/or high degrees of ecological specialization, coupled with passive long-distance migrations, may predispose some species to founder events, spatially structured populations, and ultimately, accelerated rates of molecular evolution.

Collembola present a perfect example of these peculiar population dynamics: they are non-vagile, with high degrees of ecological specificity, and are seemingly robust to climatic oscillations and ecological change (Cicconardi et al., 2010). These characteristics enable them to persist locally, often at very fine spatial scales, for long periods of time (Garrick et al., 2008; Katz et al., 2018), but springtails can also passively disperse vast distances by air (e.g., Freeman, 1952; Blackith & Disney, 1988; Coulson, 2003; Hawes et al., 2007) and water (e.g., Hawes et al. 2008; Coulson et al. 2002) potentially facilitating recurrent founder events resulting in strongly structured populations. Biological factors such as shorter generation time (Thomas et al., 2010), smaller effective population size (Lanfear et al., 2014), inefficient DNA repair mechanisms (Britten, 1986), smaller body size (Martin & Palumbi, 1993; Gillooly et al., 2005; but see Thomas et al., 2006), asexual reproduction (Neiman et al., 2010), and increased metabolic rate (Martin & Palumbi, 1993; Gillooly et al., 2005; but see Lanfear et al., 2007) have also been hypothesized to stimulate accelerated rates of molecular evolution. Some of these characteristics are present in springtails: some species are asexual with extremely short generation times (Moore et al., 2005), they are miniscule

(body length often less than 1 mm), and they may have surprisingly small effective population sizes due to their highly structured, isolated populations.

Despite the circumstantial plausibility for accelerated molecular rates in Collembola, "universal" insect mitochondrial (COI) clocks, such as Brower's (1994) estimate of 2.3% divergence per million years (Ma) or Papadopoulou et al.'s (2010) more recent and statistically robust estimate of 3.54% divergence/Ma (in tenebrionid beetles), are routinely used for evolutionary investigations of springtail biodiversity (e.g., Stevens & Hogg, 2003; Nolan et al., 2006; Stevens et al., 2006; McGaughran et al., 2008; Bennett et al., 2016; von Saltzwedel et al., 2016; 2017; Ding et al., 2018; Katz et al., 2018). Estimates of evolutionary rates in units of absolute time are essential for evaluating the tempo of molecular evolution in springtails, but without independent timing information for molecular clock calibration, rates in absolute time cannot be determined, thus estimates of evolutionary timescales for springtails using external clocks remain dubious at best.

Nevertheless, estimating molecular rates in relative, rather than absolute, time can offer an alternative way to test evolutionary hypotheses concerning the rate of genetic change or molecular divergence for groups that lack calibrations. This comparative approach has been used to estimate relative divergence times for hydrobiid snails (Wilke et al., 2009) and members of a solute carrier protein family (Geyer et al., 2006) to determine which groups had originated most recently. Other studies have incorporated modern relaxed molecular clock models to accommodate observed rate heterogeneity among lineages for estimating relative divergence times *and* relative rates of molecular evolution in sucrose sugar transporter proteins (Cabezas-Cruz et al., 2015).

Although not within the primary scope of their study, Ciccondardi et al. (2010) estimated relative rates among groups of hexapods to identify an appropriate external clock rate for estimating divergence times in Mediterranean springtails. However, their analysis had limitations: they used COII, a gene less commonly used for springtail phylogenetics; their taxonomic coverage of Hexapoda was incomplete (98 taxa for 3/4 hexapod classes, including 15/28 insect orders); they did not test for saturation, which may have impacted rate estimation; phylogenetic relationships among classes and orders were not constrained (likely because many of these relationships remained unresolved at this time; see Meusemann et al., 2010; Trautwein et al., 2012), which may have produced incorrect tree topologies (which they do not report) that could have affected their rate estimates; they did not report the model of sequence evolution used, but it was likely inferred using maximum likelihood methods (due to the unavailability of Bayesian substitution model averaging methods at this time); and lastly, they did not provide adequate summaries of their findings (only graphical representations of mean rates and standard error), which lack 95% highest posterior density (HPD) values—an important credibility interval for assessing significance within a Bayesian framework (Box & Tiao, 1992).

I expand the work of Cicconardi et al. (2010) by using the cytochrome c oxidase I (COI) gene, rather than COII, to estimate relative rates of substitution in Hexapoda. COI is the most widely implemented gene used for external clock calibrations in springtails (and most other invertebrates) due to the popularity of its 'barcoding' region, making it an ideal candidate gene for evaluating relative rates in hexapods. Here I performed a Bayesian phylogenetic analysis using COI sequences for 188 taxa representing every

hexapod class and insect order, a fixed hexapod tree topology based on robust relationships that were independently inferred with genomic data (Misof et al., 2014), and an uncorrelated log normal "relaxed" clock model set to a mean rate of 1 (subs/site/time) to allow branch rates to vary (with respect to each other and relative to 1). This approach allowed us to avoid potentially violating assumptions inherent with the use of "universal" clock rates (e.g., rate constancy) or node calibrations (e.g., timing of biogeographic event), while still addressing the basic question of whether or not Collembola have accelerated rates of molecular substitution (compared to other hexapods) and ultimately, whether or not the use of a "universal" arthropod/insect COI clock is appropriate for estimating absolute evolutionary timescales in this group.

Methods

Sequence acquisition and alignment

Complete cytochrome c oxidase I (COI) sequences for a total of 204 taxa were acquired from GenBank (Benson et al., 2013) (Table 3.1): 188 taxa representing all classes within Hexapoda (Collembola, Protura, Diplura and all 28 insect orders) and 16 arthropod outgroup taxa. Taxa were chosen to correspond to species used by Misof et al. (2014) because I used their phylogeny to topologically constrain the Bayesian analysis (see below). If COI was not available for a given taxon, COI sequences for a closely related taxon were used instead (i.e., from the same genus, or same subfamily, etc.). Sequences were haphazardly chosen if multiple complete COI sequences were available in GenBank for a given taxon. For some groups (e.g., Collembola) additional taxa were included if available. All nucleotide sequences were then aligned by amino acids using MAFFT

(Katoh et al., 2005) implemented in TranslatorX (Abascal et al., 2010). I used the GBlocks (Castresana, 2000) option in TranslatorX to analyze and remove columns with ambiguous homology from the nucleotide alignment based on the amino acid alignment.

Tests for COI substitution saturation

Substitution saturation is a major concern when estimating phylogenetic parameters across deep evolutionary time (Arbogast et al., 2002), especially when using faster evolving loci typically used to estimate species level relationships (e.g., COI). Therefore, to make sure COI is an appropriate marker to evaluate relative substitution rates among groups of arthropods, some of which have been independently evolving for more than 400 million years (Misof et al., 2014), I used two different approaches to test for substitution saturation in the COI sequence dataset (1491 bp for 204 arthropod taxa). First, I performed simple linear regression analyses to identify correlations between uncorrected genetic distances (uncorrected p-distances) and genetic distances corrected with a general time reversible (GTR) model of sequence evolution (model-corrected p-distances), independently for each codon position. If the relationship between uncorrected and model-corrected p-distances is approximately linear, then there is no saturation. If saturation is present, the relationship will deviate and begin to plateau because uncorrected p-distances will be underestimating the number of substitutions between taxa due to a loss of information resulting from multiple substitutions at single sites (Arbogast et al., 2002). At these evolutionary time scales, the third codon position for COI is likely to exhibit severe substitution saturation due to increased numbers of synonymous mutations. However, synonymous mutations should conform better to neutral theory of

molecular evolution (Xia & Lemey, 2009), and therefore, their inclusion may benefit phylogenetic estimation (Yang, 1996). To determine if the inclusion of codon 3 would impair phylogenetic estimation, I compared linear relationships for codons 1 & 2 combined with those for all codon positions combined. R^2 values were calculated in Microsoft Excel to measure the fit of linear models for all distance plots. All uncorrected and corrected distances were calculated in Paup* v4.0a build 161 (Swofford, 2002). Next, I assessed substitution saturation for COI (all codons) using the substitution saturation test developed by Xia et al. (2003) with DAMBE v6.0.0 (Xia, 2017) using 100 jackknife replicates. This test evaluates whether the observed saturation is significantly lower than critical saturation values derived from simulation studies (Xia et al., 2003) assuming different tree topologies (Xia & Lemy, 2009).

Bayesian phylogenetic inference and relative rate estimation

To estimate relative rates of COI substitution across Hexapoda I performed a Bayesian phylogenetic analysis using a fixed tree topology (Fig. 3.1) and an uncorrelated log normal "relaxed" clock model with a mean rate equal to 1 unit time. A fixed tree topology allows us to mitigate the effect of COI's weak resolving power for deeper nodes. Fortunately, Misof et al.'s (2014) phylogenomic study produced a highly resolved and well-supported hexapod phylogeny that can be used to constrain the tree topology for my analysis (Fig. 3.2). Well-supported relationships from Misof et al. (2014) were constrained to be monophyletic, while those with weak support were collapsed to allow them to be estimated in our analysis (Fig. 3.1). This approach effectively limits the 'tree space' the MCMC algorithm needs to search, which reduces the time needed for

convergence and minimizes the error in rate estimation due to incorrect estimates of phylogenetic relationships. I used an uncorrelated log normal clock model and set the mean rate equal to 1 to allow rates to vary among branches relative to a mean of 1 unit time for all branch rates. This allows identification of groups with faster or slower rates of substitution relative to other hexapod groups.

The Bayesian phylogenetic analysis was performed using BEAST2 v. 2.4.8 (Bouckaert et al., 2014) with the following parameters: bModelTest (Bouckaert & Drummond, 2017) for site model averaging to accommodate uncertainty in the model of sequence evolution (default parameters); an uncorrelated log normal clock to allow rate variation among branches; clock.rate set to 1 to estimate branch lengths relative to 1 unit time; Yule tree prior; multimonophyletic constraint prior to fix tree topology for well-supported relationships estimated by Misof et al. (2014) (Figs 3.1–3.2); MCMC for 300 million generations; sampling trees and statistics every 5000 generations; all other parameters were left as default. After applying a 10% burn-in, the effective sample size (ESS) for all parameters were determined to be greater than 200 with Tracer v1.6.0 (Rambaut & Drummond, 2007), a total of 54001 trees were sampled for analysis, and a maximum clade credibility tree with median node heights was inferred with TreeAnnotator v2.4.8 (Bouckaert et al., 2014). Trees were cosmetically modified in FigTree v.1.4.3 (Available at http://tree.bio.ed.ac.uk/software/figtree/) and Adobe Illustrator CC. Rate summary statistics for Arthropod groups were extracted from the BEAST2 output log file using TreeStat v1.8.4 from the BEAST v1.8.4 software package (Drummond et al., 2012). Mean rate bar graphs were created using Microsoft Excel.

Significant differences between mean relative rates among hexapod groups were assessed by comparing 95% highest posterior density (HPD) rate intervals.

Results

COI substitution saturation tests

Simple linear regression plots of genetic distances (uncorrected vs. model corrected) (Fig. 3.3) identified significant substitution saturation for COI codon position 3 only (R^2=16.98). Plots for codon 1, codon 2, and codons 1 and 2 combined, indicate strong correlations between uncorrected and model-corrected genetic distances (R^2=0.9564, 0.989, and 0.9774, respectively). When codon 3 was combined with codons 1 and 2, the correlation coefficient remained relatively high (R^2=0.9153) indicating that the inclusion of codon 3 does not significantly impact accurate estimations of substitution, thus should be included for the Bayesian analysis for estimating evolutionary rates. Results from Xia et al.'s (2003) test for substitution saturation provided an additional measure of support to justify the inclusion of codon 3 for the relative rate analysis. Iss (simple index of substitution saturation) was significantly lower than Iss.c (critical Iss value) under the assumption of both symmetrical and asymmetrical tree topologies ($P<0.01$) (Table 3.2), thus the hypothesis of significant substitution saturation in the combined (codons 1–3) sequence alignment can be rejected. Thus, I determined that phylogenetic inferences of the hexapod COI dataset are not strongly impacted by substitution saturation, and I included all three codons in subsequent analyses.

Relative rates of COI substitution among hexapods

The Bayesian phylogenetic analysis of 204 arthropod COI sequences revealed significant variation in relative rates of COI substitution among hexapod groups (Table 3.3; Figs 3.4–3.6). Visual inspection of the maximum clade credibility tree (with branches colored to indicate rate) revealed a six groups (i.e., Protura, Embioptera, Thysanoptera, Psocodea, Hymenoptera, and Strepsiptera) that display highly accelerated rates of COI substitution compared to other Hexapods (Figs 3.4, 3.6). Mean relative rates ranged from 0.508 to 3.395 substitutions/site/time (Table 3.2). However, only a scant few groups had 95% HPD rate intervals that did not include 1 substitutions/site/time (mean tree rate) (i.e., Strepsiptera, Embioptera, Psocodea, Protura, Hymenoptera, and Thysanoptera were higher, while Grylloblattodea, Diplura, Megaloptera and Zygentoma were lower), and 95% HPD rate intervals for all other taxonomic groups, including Collembola, included 1 substitutions/site/time (Fig. 3.5). When compared to Collembola, only five groups had significantly different rates of COI substitution (i.e., Strepsiptera Embioptera, Psocodea, Protura, and Hymenoptera) (Fig. 3.5). Therefore, the hypothesis that springtails have accelerated rates of substitution compared to other arthropods is unsupported.

Discussion

Saturation tests indicated that the phylogenetic analysis would not be significantly impacted by substitution saturation (Table 3.2; Fig. 3.3). This was relatively surprising given the taxonomic depth of the samples (i.e., Hexapoda + outgroups). Simple linear regression analyses (Fig. 3.3) indicated a slight reduction in the strength of linear correlation (R^2) from 0.9774 for codon position 1 and 2 combined, to 0.9153 after adding

codon position 3. However, Philippe et al. (2011) suggests that the phylogenetic inference can be maximized when sequences are only slightly saturated. It is possible that my comprehensive sampling approach, which incorporated most major lineages of Hexapoda, provided sufficient coverage of nucleotide diversity needed to facilitate accurate approximation multiple COI substitutions (Philippe et al., 2011).

The comparative rate analysis of the COI gene (Table 3.3; Figs 3.4–3.5) revealed that Collembola exhibit rates within the range of rate variation for most other hexapod groups. This suggests that the application of external insect clocks may be appropriate for springtails. Relative rates estimated in this study were also remarkably consistent with those estimated by Cicconardi et al.'s (2010), despite differences in methodology and gene choice. Although most hexapod groups were determined to have relatively similar rates of COI substitution, I did identify five groups with significantly higher rates compared to springtails (i.e., Strepsiptera, Embioptera, Psocodea, Protura, and Hymenoptera), supporting critics of "universal" clocks (e.g., Thomas et al., 2006) and those that stress the importance of taxon-specific clocks (e.g., Weir & Schluter, 2008). It is also noteworthy that accelerated rates of mitochondrial substitution have been previously documented for most of these groups (e.g., Hafner et al., 1994; Johnson et al., 2003; Kaltenpoth et al., 2012; Chen et al., 2017), signifying that, despite the low precision associated with my estimates, this method has still produced accurate results that are neither unique nor unusual. In some of these groups there is a relatively high prevalance of parasitism (i.e., Strepsiptera, Psocodea, and Hymenoptera)—a life history trait thought to be linked to founder effects and strongly structured populations that can lead to accelerated evolutionary rates (Kaltenpoth et al., 2012). Other studies suggest that

high rates of substitution in Thysanoptera, Psocodea, and Embioptera are related to frequent gene rearrangements (Shao et al., 2003; Kômoto et al., 2012).

The high rates for Protura were more perplexing. Protura and Collembola are phylogenetic sister taxa that share many similarities in their biology: they are both members of the soil mesofauna, are very small, and have very low dispersal capabilities. In addition, large genetic distances (~20%) between presumably conspecific taxa are commonly detected in both groups (e.g., Resch et al., 2014; Katz et al., 2015), suggesting they have similar patterns of pervasive cryptic diversity. Therefore, I was surprised to see that Protura substitution rates were 1.5–2 times higher than Collembola rates. Many aspects of proturan biology, ecology, and population dynamics remain enigmatic, thus additional research is needed to identify potential factors responsible for these observed differences.

Because springtail substitution rates were found to be generally similar to most other hexapods, their apparent morphological-molecular disparity cannot be explained by accelerated rates of molecular evolution, and this disparity is instead more consistent with the hypothesis of long-term morphological stasis. However, given the limited taxonomic sampling (average rates of clades may be biased) and low precision in the rate estimates, my findings should be considered with caution, especially for hypotheses concerning mechanisms driving cryptic diversity within the Collembola. This study serves as a useful first attempt to understanding the rate of molecular evolution in springtails with respect to patterns of cryptic diversity and provides some limited validation of previous studies that have used “universal” insect clocks for estimating divergence times in springtails. More importantly, this analysis exemplifies a relatively simple method for evaluating rate

constancy within a taxonomic group to determine whether the use of an external "universal" clock is appropriate—a practice that is becoming increasingly scrutinized (e.g., Coppard & Lessios, 2017), despite being essential for addressing evolutionary hypotheses in groups lacking fossils or other independent timing information for molecular clock calibration (Wilke et al., 2009).

Acknowledgements

This project was supported by the Graduate School at the University of Illinois at Urbana, the Department of Entomology at the University of Illinois at Urbana-Champaign, and the Illinois Natural History Survey at the Prairie Research Institute.

Tables

Table 3.1. All taxa used in this study with corresponding GenBank accession numbers.

Subphylum	Class/Subclass	Order	Family	Species	GenBank #
Chelicerata	Arachnida	Ixodida	Ixodidae	*Ixodes pavlovskyi*	NC_023831
Crustacea	Branchiopoda	Cladocera	Daphniidae	*Daphnia pulex*	NC_000844
Crustacea	Copepoda	Siphonostomatoida	Caligidae	*Lepeophtheirus salmonis*	NC_007215
Crustacea	Malacostraca	Decapoda	Ocypodidae	*Ocypode cordimanus*	NC_029725
Crustacea	Malacostraca	Decapoda	Ocypodidae	*Leptuca pugilator*	AF466700
Crustacea	Malacostraca	Decapoda	Penaeidae	*Litopenaeus vannamei*	NC_009626
Crustacea	Ostracoda	Myodocopida	Cypridinidae	*Vargula hilgendorfii*	NC_005306
Crustacea	Remipedia	Nectiopoda	Speleonectidae	*Speleonectes tulumensis*	NC_005938
Hexapoda	Collembola	Entomobryomorpha	Entomobryidae	*Entomobrya assuta*	KM610047
Hexapoda	Collembola	Entomobryomorpha	Entomobryidae	*Entomobryoides dissimilis*	KM610126
Hexapoda	Collembola	Entomobryomorpha	Entomobryidae	*Homidia socia*	KM610128
Hexapoda	Collembola	Entomobryomorpha	Entomobryidae	*Willowsia nigromaculata*	KM610130
Hexapoda	Collembola	Entomobryomorpha	Isotomidae	*Cryptopygus antarcticus*	NC_010533
Hexapoda	Collembola	Entomobryomorpha	Isotomidae	*Folsomotoma octooculata*	NC_024155
Hexapoda	Collembola	Entomobryomorpha	Lepidocyrtidae	*Pseudosinella violenta*	KM610132
Hexapoda	Collembola	Entomobryomorpha	Orchesellidae	*Orchesella villosa*	NC_010534
Hexapoda	Collembola	Entomobryomorpha	Seiridae	*Seira dowlingi*	KM610133
Hexapoda	Collembola	Poduromorpha	Hypogastruridae	*Gomphiocephalus hodgsoni*	NC_005438
Hexapoda	Collembola	Poduromorpha	Neanuridae	*Bilobella aurantiaca*	NC_011195
Hexapoda	Collembola	Poduromorpha	Neanuridae	*Friesea grisea*	NC_010535
Hexapoda	Collembola	Poduromorpha	Onychiuridae	*Onychiurus orientalis*	NC_006074
Hexapoda	Collembola	Poduromorpha	Onychiuridae	*Tetrodontophora bielanensis*	NC_002735
Hexapoda	Collembola	Poduromorpha	Poduridae	*Podura aquatica*	NC_006075
Hexapoda	Collembola	Symphypleona	Sminthuridae	*Sminthurus viridis*	NC_010536
Hexapoda	Diplura	Dicellurata	Japygidae	*Japyx solifugus*	NC_007214
Hexapoda	Diplura	Dicellurata	Japygidae	*Occasjapyx japonicus*	NC_022674
Hexapoda	Diplura	Dicellurata	Parajapygidae	*Parajapyx emeryanus*	NC_022673
Hexapoda	Diplura	Rhabdura	Campodeidae	*Campodea fragilis*	NC_008233
Hexapoda	Diplura	Rhabdura	Campodeidae	*Lepidocampa weberi*	NC_022675
Hexapoda	Diplura	Rhabdura	Octostigmatidae	*Octostigma sinensis*	NC_022672
Hexapoda	Insecta	Archaeognatha	Machilidae	*Pedetontus silvestrii*	NC_011717
Hexapoda	Insecta	Archaeognatha	Machilidae	*Petrobius brevistylis*	NC_007688
Hexapoda	Insecta	Archaeognatha	Machilidae	*Songmachilis xinxiangensis*	NC_021384
Hexapoda	Insecta	Archaeognatha	Machilidae	*Trigoniophthalmus alternatus*	NC_010532
Hexapoda	Insecta	Archaeognatha	Meinertellidae	*Nesomachilis australica*	NC_006895
Hexapoda	Insecta	Blattodea	Archotermopsidae	*Zootermopsis nevadensis*	NC_024658
Hexapoda	Insecta	Blattodea	Blaberidae	*Blaptica dubia*	NC_29224
Hexapoda	Insecta	Blattodea	Blattidae	*Periplaneta americana*	NC_016956
Hexapoda	Insecta	Blattodea	Cryptocercidae	*Cryptocercus kyebangensis*	NC_030191
Hexapoda	Insecta	Blattodea	Mastotermitidae	*Mastotermes darwiniensis*	NC_018120
Hexapoda	Insecta	Blattodea	Rhinotermitidae	*Prorhinotermes canalifrons*	KP026256
Hexapoda	Insecta	Coleoptera	Carabidae	*Carabus mirabilissimus*	NC_016469
Hexapoda	Insecta	Coleoptera	Curculionidae	*Dendroctonus terebrans*	JQ005146
Hexapoda	Insecta	Coleoptera	Gyrinidae	*Macrogyrus oblongus*	NC_13249
Hexapoda	Insecta	Coleoptera	Hydroscaphidae	*Hydroscapha granulum*	NC_012144
Hexapoda	Insecta	Coleoptera	Meloidae	*Epicauta chinensis*	NC_29192
Hexapoda	Insecta	Coleoptera	Meloidae	*Hycleus chodschenticus*	KT808466
Hexapoda	Insecta	Coleoptera	Meloidae	*Lytta caraganae*	NC_033339
Hexapoda	Insecta	Coleoptera	Ommatidae	*Tetraphalerus bruchi*	NC_011328
Hexapoda	Insecta	Coleoptera	Staphylinidae	*Aleochara* sp.	KT780622
Hexapoda	Insecta	Coleoptera	Tenebrionidae	*Tribolium castaneum*	NC_003081
Hexapoda	Insecta	Dermaptera	Anisolabididae	*Euborellia arcanum*	NC_032075
Hexapoda	Insecta	Dermaptera	Pygidicranidae	*Challia fletcheri*	NC_018538
Hexapoda	Insecta	Diptera	Bibionidae	*Bibio xanthopus*	KC192959
Hexapoda	Insecta	Diptera	Bombyliidae	*Bombylius major*	KC192961
Hexapoda	Insecta	Diptera	Chloropidae	*Thaumatomyia notata*	KC192976
Hexapoda	Insecta	Diptera	Culicidae	*Aedes aegypti*	NC_035159
Hexapoda	Insecta	Diptera	Culicidae	*Anopheles gambiae*	NC_002084
Hexapoda	Insecta	Diptera	Drosophilidae	*Drosophila melanogaster*	NC_024511
Hexapoda	Insecta	Diptera	Glossinidae	*Glossina morsitans*	KC192971
Hexapoda	Insecta	Diptera	Psychodidae	*Phlebotomus papatasi*	NC_028042
Hexapoda	Insecta	Diptera	Sarcophagidae	*Sarcophaga crassipalpis*	NC_026667
Hexapoda	Insecta	Diptera	Tephritidae	*Rhagoletis pomonella*	DQ006862
Hexapoda	Insecta	Diptera	Tipulidae	*Tipula cockerelliana*	NC_030520

Table 3.1 (Continued)

Hexapoda	Insecta	Diptera	Trichoceridae	*Trichocera bimacula*	NC_016169
Hexapoda	Insecta	Embioptera	Oligotomidae	*Aposthonia japonica*	AB639034
Hexapoda	Insecta	Ephemeroptera	Baetidae	*Baetis* sp.	GU936204
Hexapoda	Insecta	Ephemeroptera	Ephemeridae	*Ephemera orientalis*	NC_012645
Hexapoda	Insecta	Ephemeroptera	Heptageniidae	*Parafronurus youi*	NC_011359
Hexapoda	Insecta	Ephemeroptera	Isonychiidae	*Isonychia ignota*	HM143892
Hexapoda	Insecta	Ephemeroptera	Siphlonuridae	*Siphlonurus immanis*	NC_013822
Hexapoda	Insecta	Ephemeroptera	Vietnamellidae	*Vietnamella dabieshanensis*	HM067837
Hexapoda	Insecta	Grylloblattodea	Grylloblattidae	*Grylloblatta sculleni*	DQ241796
Hexapoda	Insecta	Hemiptera	Aleyrodidae	*Bemisia tabaci*	NC_006279
Hexapoda	Insecta	Hemiptera	Aleyrodidae	*Trialeurodes vaporariorum*	NC_006280
Hexapoda	Insecta	Hemiptera	Aphididae	*Acyrthosiphon pisum*	NC_011594
Hexapoda	Insecta	Hemiptera	Aphididae	*Aphis gossypii*	NC_024581
Hexapoda	Insecta	Hemiptera	Cicadidae	*Diceroprocta semicincta*	KM000131
Hexapoda	Insecta	Hemiptera	Cicadidae	*Gaeana maculata*	KM244671
Hexapoda	Insecta	Hemiptera	Cicadidae	*Magicicada tredecim*	KM000130
Hexapoda	Insecta	Hemiptera	Cicadidae	*Tettigades auropilosa*	KM000129
Hexapoda	Insecta	Hemiptera	Cydnidae	*Macroscytus gibbulus*	NC_012457
Hexapoda	Insecta	Hemiptera	Delphacidae	*Nilaparvata lugens*	NC_021748
Hexapoda	Insecta	Hemiptera	Miridae	*Adelphocoris fasciaticollis*	NC_023796
Hexapoda	Insecta	Hemiptera	Miridae	*Apolygus lucorum*	NC_023083
Hexapoda	Insecta	Hemiptera	Miridae	*Creontiades dilutus*	NC_030257
Hexapoda	Insecta	Hemiptera	Miridae	*Lygus hesperus*	NC_024641
Hexapoda	Insecta	Hemiptera	Nepidae	*Laccotrephes robustus*	NC_012817
Hexapoda	Insecta	Hemiptera	Nepidae	*Nepa hoffmanni*	NC_028084
Hexapoda	Insecta	Hemiptera	Pentatomidae	*Dolycoris baccarum*	NC_020373
Hexapoda	Insecta	Hemiptera	Pentatomidae	*Eurydema gebleri*	NC_027489
Hexapoda	Insecta	Hemiptera	Pentatomidae	*Graphosoma rubrolineatum*	NC_033875
Hexapoda	Insecta	Hemiptera	Pentatomidae	*Halyomorpha halys*	NC_013272
Hexapoda	Insecta	Hemiptera	Pentatomidae	*Nezara viridula*	NC_011755
Hexapoda	Insecta	Hemiptera	Peloridiidae	*Hackeriella veitchi*	NC_020309
Hexapoda	Insecta	Hemiptera	Peloridiidae	*Hemiodoecus leai*	NC_025329
Hexapoda	Insecta	Hemiptera	Peloridiidae	*Xenophyes cascus*	NC_024622
Hexapoda	Insecta	Hemiptera	Plataspidae	*Coptosoma bifaria*	NC_012449
Hexapoda	Insecta	Hemiptera	Plataspidae	*Megacopta cribraria*	NC_015342
Hexapoda	Insecta	Hemiptera	Tessaratomidae	*Eusthenes cupreus*	NC_022449
Hexapoda	Insecta	Hemiptera	Triozidae	*Bactericera cockerelli*	NC_030055
Hexapoda	Insecta	Hemiptera	Triozidae	*Paratrioza sinica*	NC_024577
Hexapoda	Insecta	Hemiptera	Urostylididae	*Urochela quadrinotata*	NC_020144
Hexapoda	Insecta	Hemiptera	Veliidae	*Paravelia conata*	KX821865
Hexapoda	Insecta	Hemiptera	Veliidae	*Platyvelia brachialis*	KX821864
Hexapoda	Insecta	Hemiptera	Veliidae	*Stridulivelia strigosa*	KX821866
Hexapoda	Insecta	Hymenoptera	Apidae	*Apis mellifera ligustica*	NC_001566
Hexapoda	Insecta	Hymenoptera	Apidae	*Bombus terrestris*	KT368150
Hexapoda	Insecta	Hymenoptera	Braconidae	*Cotesia vestalis*	NC_014272
Hexapoda	Insecta	Hymenoptera	Chrysididae	*Chrysis fulgida*	KU854924
Hexapoda	Insecta	Hymenoptera	Figitidae	*Leptopilina boulardi*	KU665622
Hexapoda	Insecta	Hymenoptera	Formicidae	*Atta laevigata*	KC346251
Hexapoda	Insecta	Hymenoptera	Orussidae	*Orussus occidentalis*	NC_012689
Hexapoda	Insecta	Hymenoptera	Pteromalidae	*Nasonia vitripennis*	EU746609
Hexapoda	Insecta	Hymenoptera	Tenthredinidae	*Tenthredo tienmushana*	KR703581
Hexapoda	Insecta	Lepidoptera	Bombycidae	*Bombyx mori*	NC_002355
Hexapoda	Insecta	Lepidoptera	Hepialidae	*Ahamus yunnanensis*	NC_018095
Hexapoda	Insecta	Lepidoptera	Hepialidae	*Endoclita signifer*	NC_029873
Hexapoda	Insecta	Lepidoptera	Hepialidae	*Napialus hunanensis*	NC_024424
Hexapoda	Insecta	Lepidoptera	Hepialidae	*Thitarodes pui*	NC_023530
Hexapoda	Insecta	Lepidoptera	Lycaenidae	*Celastrina hersilia*	NC_018049
Hexapoda	Insecta	Lepidoptera	Lycaenidae	*Cupido argiades*	NC_020779
Hexapoda	Insecta	Lepidoptera	Lycaenidae	*Shijimiaeoides divina*	NC_029763
Hexapoda	Insecta	Lepidoptera	Papilionidae	*Parides eurimedes*	AY804371
Hexapoda	Insecta	Lepidoptera	Sphingidae	*Manduca sexta*	NC_010266
Hexapoda	Insecta	Lepidoptera	Yponomeutidae	*Yponomeuta malinellus*	YMU09206
Hexapoda	Insecta	Lepidoptera	Zygaenidae	*Rhodopsona rubiginosa*	NC_025761
Hexapoda	Insecta	Mantodea	Mantidae	*Mantis religiosa*	NC_030265
Hexapoda	Insecta	Mantophasmatodea	Mantophasmatidae	*Sclerophasma paresisense*	NC_007701
Hexapoda	Insecta	Mecoptera	Bittacidae	*Bittacus pilicornis*	NC_015118
Hexapoda	Insecta	Mecoptera	Boreidae	*Boreus elegans*	NC_015119

Table 3.1 (Continued)

Hexapoda	Insecta	Mecoptera	Nannochoristidae	*Nannochorista philpotti*	HQ696580
Hexapoda	Insecta	Mecoptera	Panorpidae	*Neopanorpa pulchra*	NC_013180
Hexapoda	Insecta	Megaloptera	Corydalidae	*Corydalus cornutus*	NC_011276
Hexapoda	Insecta	Megaloptera	Sialidae	*Sialis hamata*	NC_013256
Hexapoda	Insecta	Neuroptera	Chrysopidae	*Apochrysa matsumurae*	NC_015095
Hexapoda	Insecta	Neuroptera	Chrysopidae	*Chrysopa pallens*	NC_019618
Hexapoda	Insecta	Neuroptera	Chrysopidae	*Chrysoperla nipponensis*	NC_015093
Hexapoda	Insecta	Neuroptera	Coniopterygidae	*Semidalis aleyrodiformis*	KT425067
Hexapoda	Insecta	Neuroptera	Myrmeleontidae	*Bullanga florida*	NC_032298
Hexapoda	Insecta	Neuroptera	Myrmeleontidae	*Epacanthaclisis banksi*	NC_025905
Hexapoda	Insecta	Neuroptera	Myrmeleontidae	*Myrmeleon immanis*	NC_024826
Hexapoda	Insecta	Neuroptera	Osmylidae	*Thyridosmylus langii*	NC_021415
Hexapoda	Insecta	Odonata	Calopterygidae	*Atrocalopteryx atrata*	NC_027181
Hexapoda	Insecta	Odonata	Calopterygidae	*Vestalis melania*	NC_023233
Hexapoda	Insecta	Odonata	Coenagrionidae	*Ischnura pumilio*	NC_021617
Hexapoda	Insecta	Odonata	Coenagrionidae	*Megaloprepus caerulatus*	NC_031823
Hexapoda	Insecta	Odonata	Epiophlebiidae	*Epiophlebia superstes*	NC_023232
Hexapoda	Insecta	Odonata	Euphaeidae	*Euphaea decorata*	NC_026058
Hexapoda	Insecta	Odonata	Gomphidae	*Davidius lunatus*	NC_012644
Hexapoda	Insecta	Odonata	Libellulidae	*Hydrobasileus croceus*	NC_025758
Hexapoda	Insecta	Odonata	Libellulidae	*Orthetrum chrysis*	NC_032048
Hexapoda	Insecta	Odonata	Platycnemididae	*Platycnemis foliacea*	NC_027180
Hexapoda	Insecta	Odonata	Pseudolestidae	*Pseudolestes mirabilis*	NC_020636
Hexapoda	Insecta	Orthoptera	Acrididae	*Chorthippus chinensis*	NC_011095
Hexapoda	Insecta	Orthoptera	Acrididae	*Euchorthippus fusigeniculatus*	NC_014449
Hexapoda	Insecta	Orthoptera	Acrididae	*Gomphocerippus rufus*	NC_014349
Hexapoda	Insecta	Orthoptera	Acrididae	*Gomphocerus licenti*	NC_013847
Hexapoda	Insecta	Orthoptera	Gryllotalpidae	*Gryllotalpa unispina*	NC_029148
Hexapoda	Insecta	Orthoptera	Rhaphidophoridae	*Diestrammena asynamora*	NC_029148
Hexapoda	Insecta	Orthoptera	Rhaphidophoridae	*Troglophilus neglectus*	NC_033989
Hexapoda	Insecta	Orthoptera	Tetrigidae	*Tetrix japonica*	NC_018543
Hexapoda	Insecta	Phasmatodea	Heteropterygidae	*Heteropteryx dilatata*	NC_014680
Hexapoda	Insecta	Phasmatodea	Heteropterygidae	*Orestes mouhotii*	AB477462
Hexapoda	Insecta	Phasmatodea	Timematidae	*Timema californicum*	DQ241799
Hexapoda	Insecta	Plecoptera	Capniidae	*Apteroperla tikumana*	NC_027698
Hexapoda	Insecta	Plecoptera	Capniidae	*Capnia zijinshana*	NC_034661
Hexapoda	Insecta	Plecoptera	Gripopterygidae	*Zelandoperla fenestrata*	NC_034997
Hexapoda	Insecta	Plecoptera	Nemouridae	*Nemoura nankinensis*	NC_034939
Hexapoda	Insecta	Plecoptera	Perlidae	*Dinocras cephalotes*	NC_022843
Hexapoda	Insecta	Plecoptera	Perlidae	*Kamimuria chungnanshana*	NC_028076
Hexapoda	Insecta	Plecoptera	Perlidae	*Acroneuria hainana*	NC_026104
Hexapoda	Insecta	Plecoptera	Pteronarcyidae	*Pteronarcella badia*	NC_029248
Hexapoda	Insecta	Plecoptera	Pteronarcyidae	*Pteronarcys princeps*	NC_006133
Hexapoda	Insecta	Plecoptera	Styloperlidae	*Styloperla spinicercia*	NC_034809
Hexapoda	Insecta	Psocodea	Boopiidae	*Heterodoxus macropus*	NC_002651
Hexapoda	Insecta	Psocodea	Liposcelididae	*Liposcelis sculptilimacula*	KX171073
Hexapoda	Insecta	Psocodea	Pediculidae	*Pediculus humanus capitis*	KC685833
Hexapoda	Insecta	Psocodea	Pediculidae	*Pediculus humanus corporis*	KC685832
Hexapoda	Insecta	Psocodea	Psocidae	*Longivalvus hyalospilus*	JQ910986
Hexapoda	Insecta	Psocodea	Trichopsocidae	*Psococerastis albimaculata*	NC_021400
Hexapoda	Insecta	Raphidioptera	Inocelliidae	*Inocellia fujiana*	KT425085
Hexapoda	Insecta	Raphidioptera	Raphidiidae	*Xanthostigma gobicola*	KT425093
Hexapoda	Insecta	Siphonaptera	Ceratophyllidae	*Jellisonia amadoi*	NC_022710
Hexapoda	Insecta	Strepsiptera	Mengenillidae	*Mengenilla moldrzyki*	NC_018545
Hexapoda	Insecta	Strepsiptera	Stylopidae	*Xenos vesparum*	AM286745
Hexapoda	Insecta	Thysanoptera	Thripidae	*Frankliniella intonsa*	NC_021378
Hexapoda	Insecta	Thysanoptera	Thripidae	*Thrips imaginis*	NC_004371
Hexapoda	Insecta	Trichoptera	Hydropsychidae	*Hydropsyche pellucidula*	NC_029246
Hexapoda	Insecta	Trichoptera	Limnephilidae	*Limnephilus decipiens*	NC_026219
Hexapoda	Insecta	Zoraptera	Zorotypidae	*Zorotypus medoensis*	NC_026077
Hexapoda	Insecta	Zygentoma	Ateluridae	*Atelura formicaria*	NC_011197
Hexapoda	Insecta	Zygentoma	Lepismatidae	*Thermobia domestica*	NC_006080
Hexapoda	Insecta	Zygentoma	Tricholepidiidae	*Tricholepidion gertschi*	NC_005437
Hexapoda	Protura	Acerentomata	Acerentomidae	*Acerella muscorum*	NC_026675
Hexapoda	Protura	Acerentomata	Acerentomidae	*Acerentomon microrhinus*	NC_026666
Myriapoda	Diplopoda	Callipodida	Abacionidae	*Abacion magnum*	NC_021932
Myriapoda	Diplopoda	Julida	Julidae	*Anaulaciulus koreanus*	NC_034656

Table 3.1 (Continued)

Myriapoda	Diplopoda	Julida	Nemasomatidae	*Antrokoreana gracilipes*	NC_010221
Myriapoda	Diplopoda	Platydesmida	Andrognathidae	*Brachycybe lecontii*	NC_021934
Myriapoda	Diplopoda	Polydesmida	Xystodesmidae	*Appalachioria falcifera*	NC_021933
Myriapoda	Diplopoda	Spirobolida	Spirobolidae	*Narceus annularus*	NC_003343
Myriapoda	Symphyla	—	Scolopendrellidae	*Symphylella* sp.	NC_011572
Myriapoda	Symphyla	—	Scutigerellidae	*Scutigerella causeyae*	NC_008453

Table 3.2. Results output for the substitution saturation test for COI implemented in DAMBE. This test evaluates whether the observed Iss (simple index of substitution saturation) is significantly lower than Iss.c (critical Iss value derived from simulation results) assuming a symmetrical (Sym) and asymmetrical (Asym) topology. It uses a heuristic approach to randomly sample different subsets of 4, 8, 16, 32 OTUs (NumOTU) multiple times (100 jackknife replicates) to test for the presence of substitution saturation for each subset. COI substitution saturation can be rejected if Iss is significantly lower than Iss.c ($P<0.5$). The mean proportion of invariant sites required for this test was determined from the Bayesian phylogenetic analysis using BEAST2.

NumOTU	Iss	Iss.cSym	T	DF	P	Iss.cAsym	T	DF	P
4	0.432	0.833	24.970	1177	0.0000	0.803	23.083	1177	0.0000
8	0.434	0.809	21.131	1177	0.0000	0.709	16.226	1177	0.0000
16	0.442	0.792	20.347	1177	0.0000	0.608	9.669	1177	0.0000
32	0.443	0.774	19.299	1177	0.0000	0.491	2.7771	1177	0.0057
Proportion of invariant sites (mean) = 0.21									

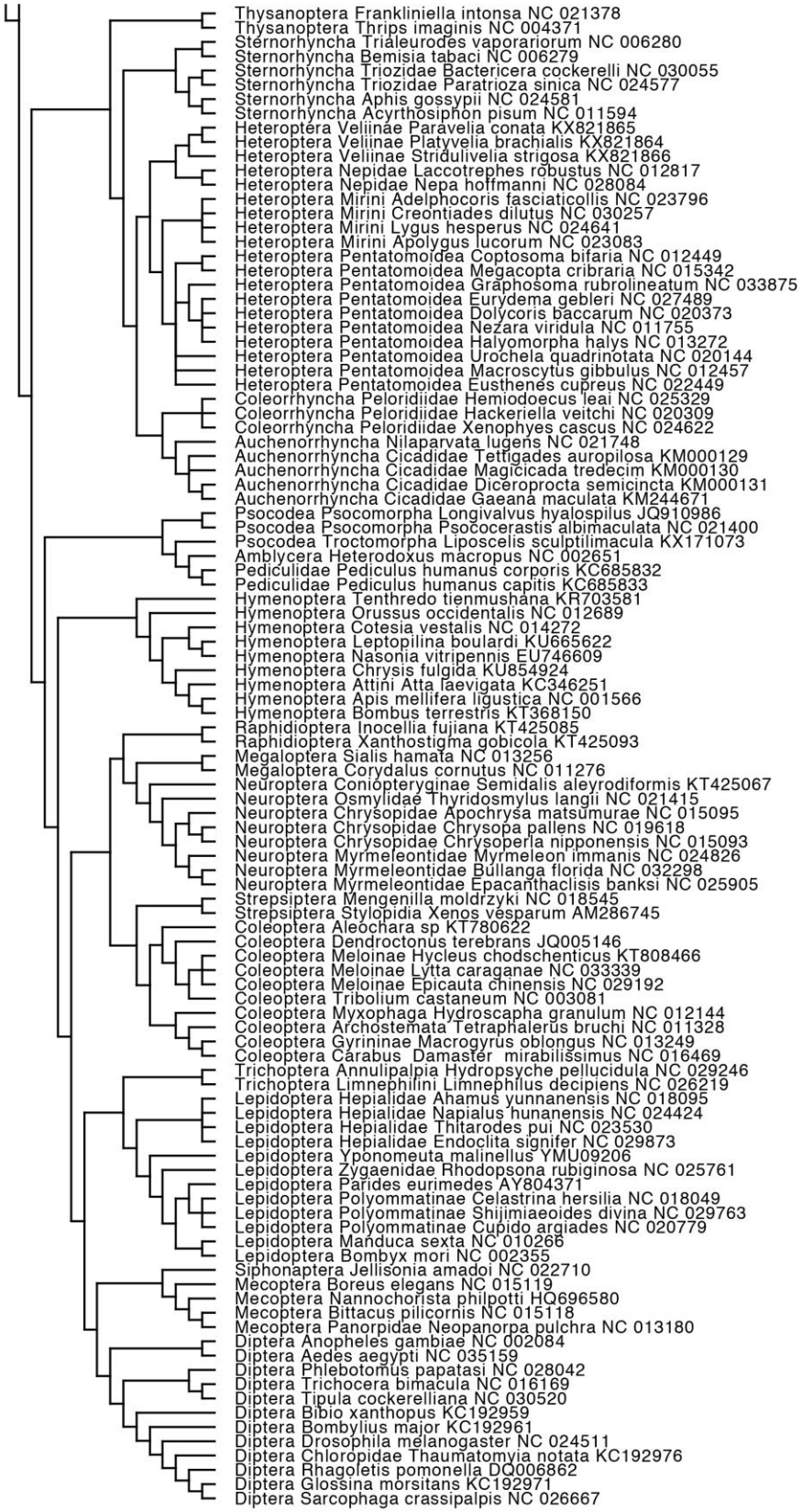
Table 3.3. Relative rate summary statistics for each hexapod group compared in this study based on a total of 54001 sampled trees resulting from the Bayesian phylogenetic analysis. Units in substitutions/site/time.

Taxonomic group	mean	stderr of mean	stdev	variance	median	95% HPD Interval
Archaeognatha	0.774	0.008	0.131	0.017	0.754	0.5543–1.0352
Blattodea	0.829	0.004	0.113	0.013	0.818	0.6215–1.0547
Coleoptera	0.903	0.002	0.077	0.006	0.896	0.7592–1.0571
Collembola	0.940	0.003	0.076	0.006	0.935	0.7952–1.0903
Dermaptera	0.901	0.007	0.184	0.034	0.858	0.6338–1.2793
Diplura	0.721	0.004	0.099	0.010	0.706	0.5552–0.9236
Diptera	0.846	0.003	0.079	0.006	0.841	0.6968–1.0037
Embioptera	1.917	0.012	0.290	0.084	1.896	1.3981–2.5072
Ephemeroptera	0.793	0.005	0.113	0.013	0.778	0.5920–1.0160
Grylloblattodea	0.738	0.004	0.121	0.015	0.725	0.5330–0.9824
Hemiptera	1.089	0.002	0.056	0.003	1.087	0.9798–1.1980
Hymenoptera	1.558	0.003	0.108	0.012	1.552	1.3543–1.7746
Lepidoptera	0.924	0.003	0.088	0.008	0.918	0.7554–1.0955
Mantodea	1.252	0.007	0.202	0.041	1.230	0.8794–1.6597
Mantophasmatodea	0.890	0.005	0.146	0.021	0.873	0.6313–1.1810
Mecoptera	0.890	0.004	0.144	0.021	0.873	0.6262–1.1746
Megaloptera	0.680	0.002	0.113	0.013	0.659	0.5048–0.9117
Neuroptera	0.934	0.003	0.095	0.009	0.925	0.7565–1.1231
Odonata	0.880	0.004	0.091	0.008	0.875	0.7073–1.0614
Orthoptera	0.827	0.002	0.090	0.008	0.820	0.6596–1.0082
Phasmatodea	1.117	0.006	0.162	0.026	1.101	0.8289–1.4416
Plecoptera	0.863	0.005	0.096	0.009	0.855	0.6797–1.0527
Protura	1.646	0.007	0.202	0.041	1.620	1.2945–2.0411
Psocodea	1.898	0.003	0.141	0.020	1.889	1.6361–2.1847
Raphidioptera	0.874	0.004	0.140	0.020	0.844	0.6610–1.1674
Strepsiptera	3.395	0.006	0.256	0.066	3.388	2.9112–3.9146
Thysanoptera	1.347	0.010	0.208	0.043	1.304	1.0312–1.7989
Trichoptera	1.283	0.004	0.172	0.030	1.265	0.9678–1.6226
Zoraptera	1.089	0.003	0.140	0.020	1.074	0.8349–1.3618
Zygentoma	0.642	0.003	0.098	0.010	0.623	0.4899–0.8429

Figures

Figure 3.1. Constraint tree, modeled after Misof et al. (2014), used to constrain tree topology for Bayesian phylogenetic analysis

Figure 3.1 (Continued)



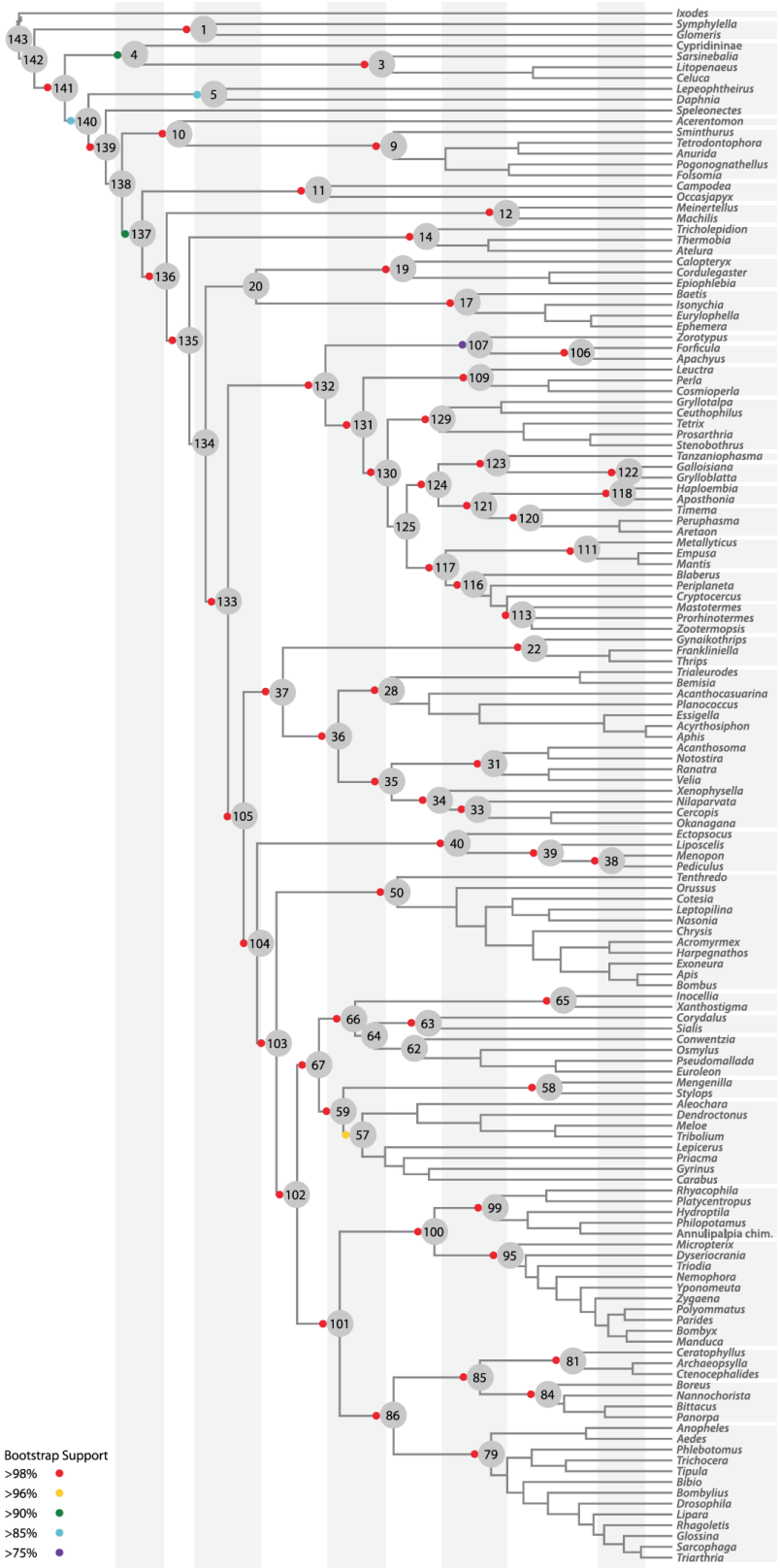


Figure 3.2. Hexapod phylogeny inferred by Misof et al. (2014) using a maximum-likelihood analysis of 413,459 amino acid sites translated from genomic data. Figure taken from Misof et al. (2014).

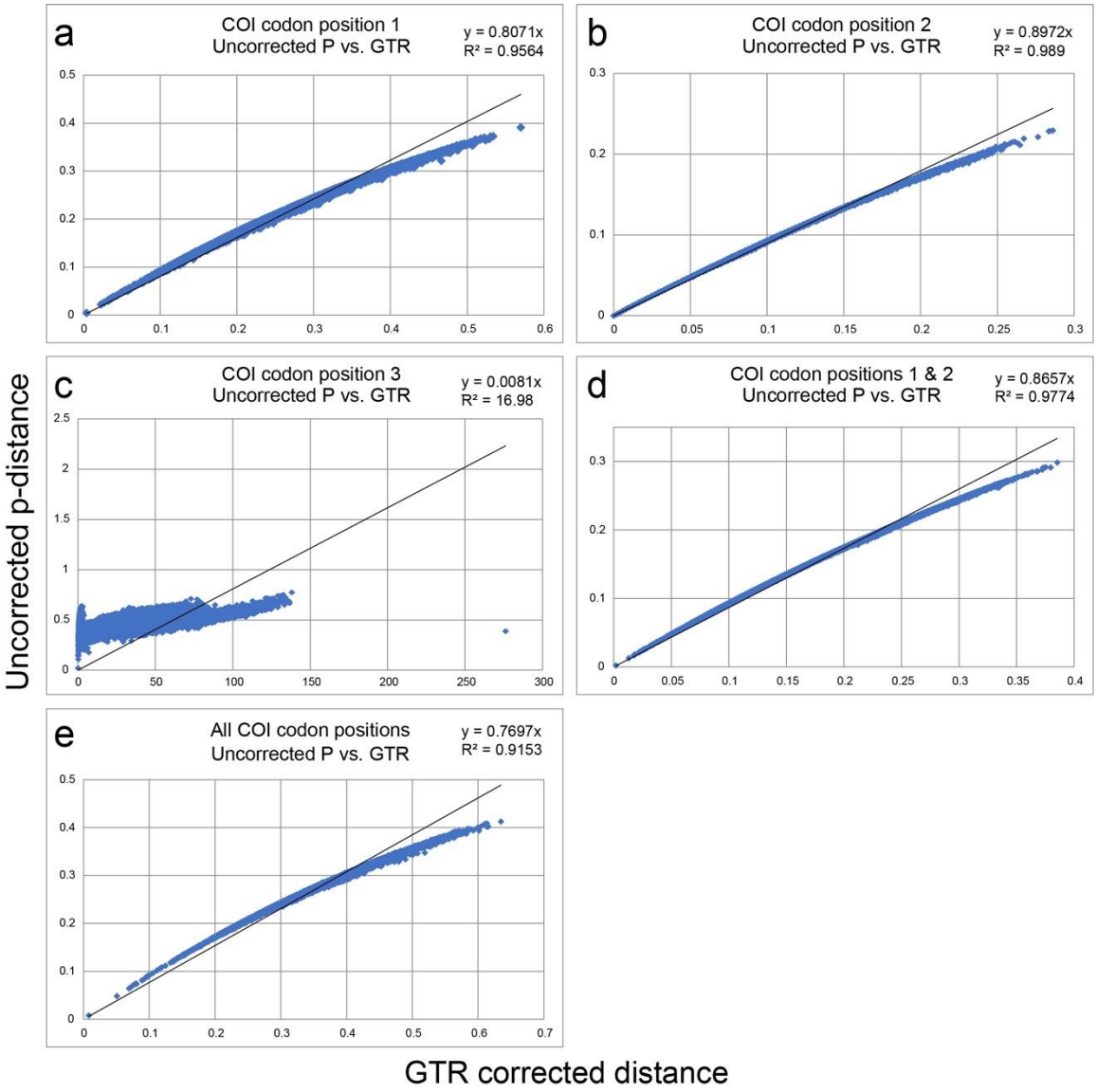


Figure 3.3. COI substitution saturation plots for testing a linear correlation between uncorrected and model-corrected (GTR) genetic distances for (**a**) codon position 1, (**b**) codon position 2, (**c**) codon position 3, (**d**) codon positions 1 and 2 combined, and (**e**) all codon positions. The correlation coefficient (R^2) is indicated for each plot.

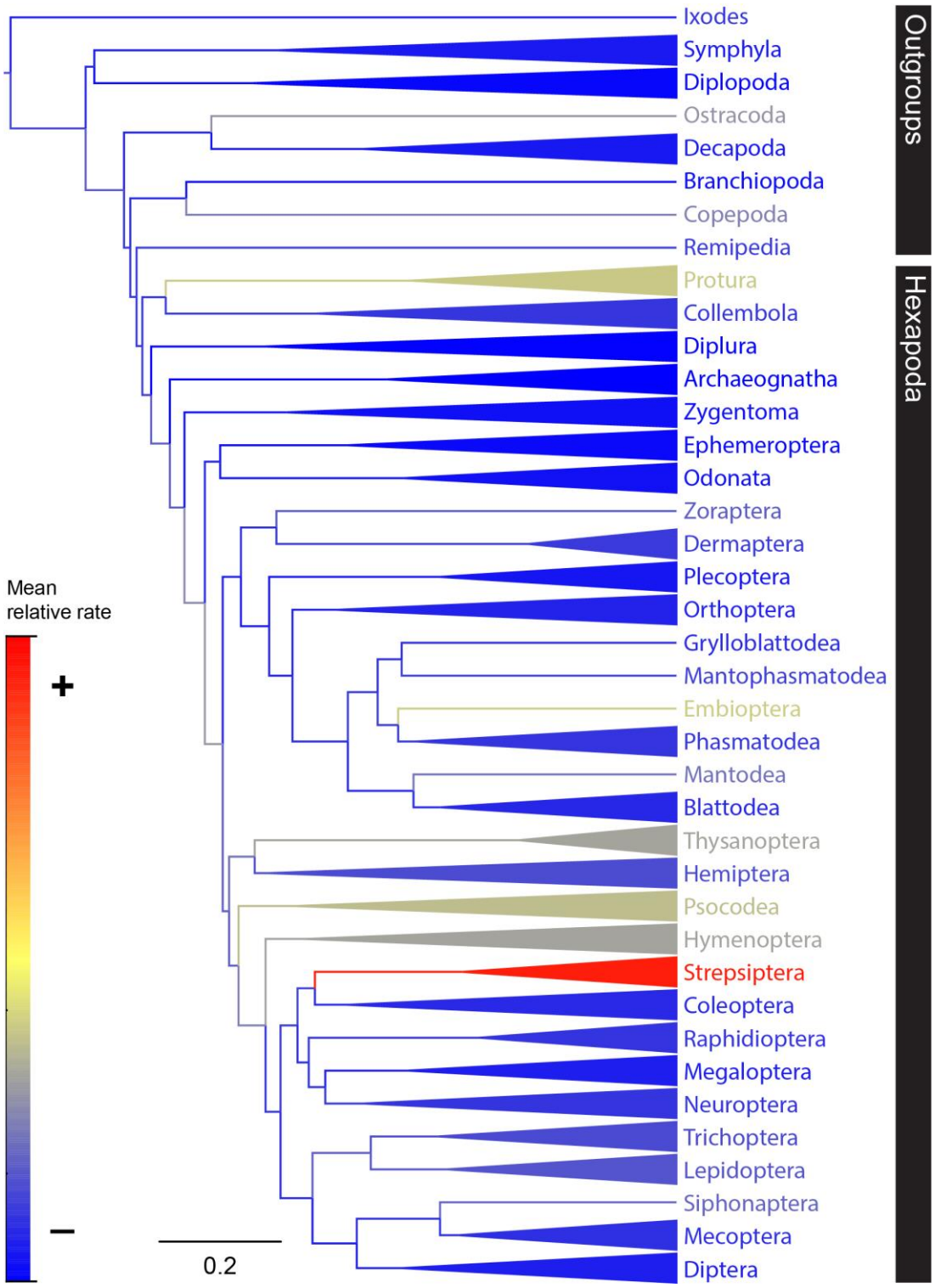


Figure 3.4. Bayesian phylogeny using a topological constraint from Misof et al. (2014) and a molecular clock set to 1 substitutions/site/time to estimate relative rates of COI substitution for all Hexapoda groups. Branch color indicates relative rate (see legend). Rates estimated from a total of 54001 sampled trees from the Bayesian phylogenetic analysis. Group clades with multiple taxa are collapsed. See Figure 3.6 for phylogeny showing all branches for individual taxa with branch support.

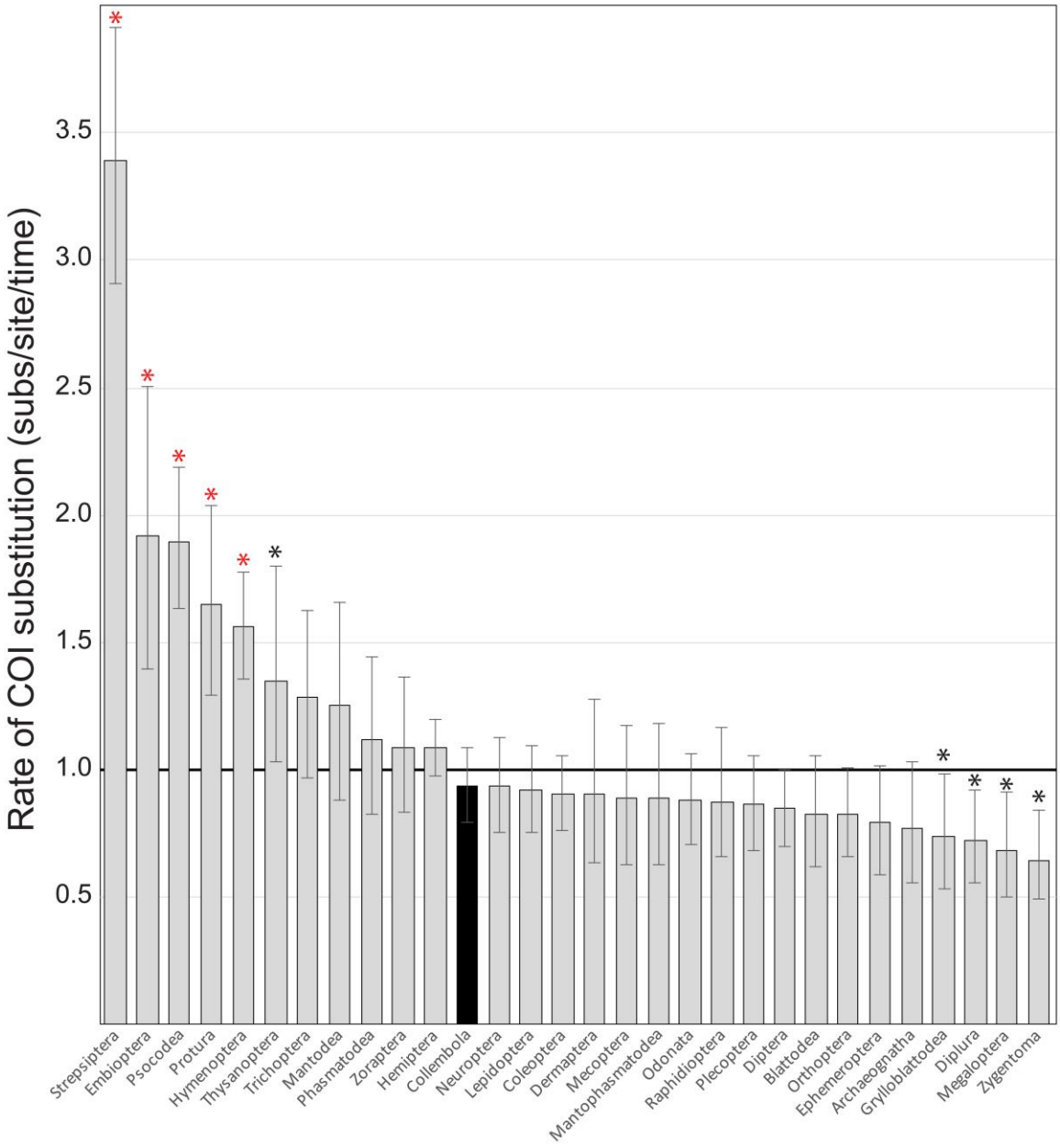


Figure 3.5. Bar chart of relative rates of COI substitution for all major Hexapod groups based on a total of 54001 sampled trees resulting from the Bayesian phylogenetic analysis. Error bars represent the 95% HPD mean rate interval. Asterisks indicate significant difference from the mean tree rate of 1 (horizontal black line); red asterisks indicate significant difference from Collembola rates (black bar).

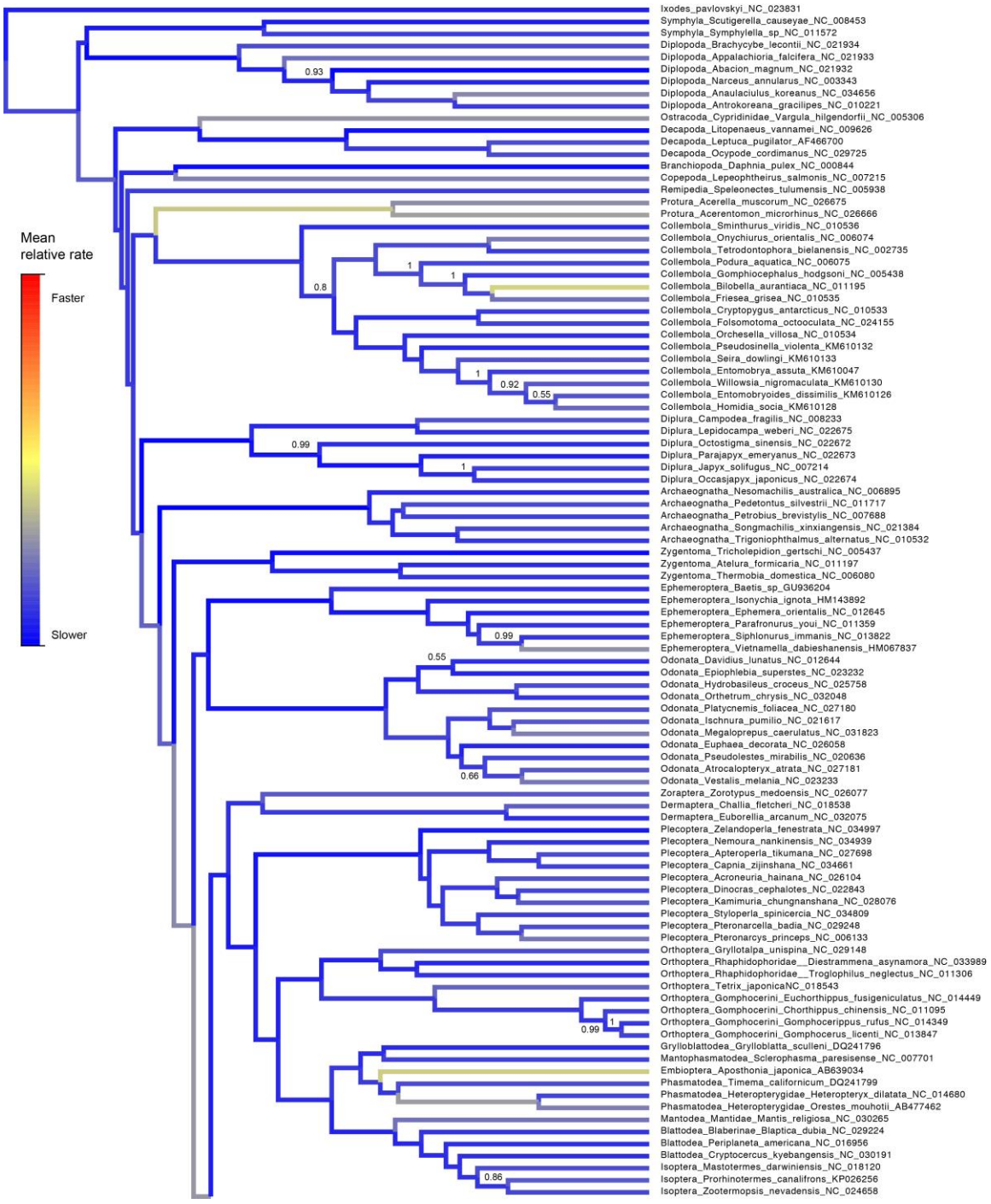
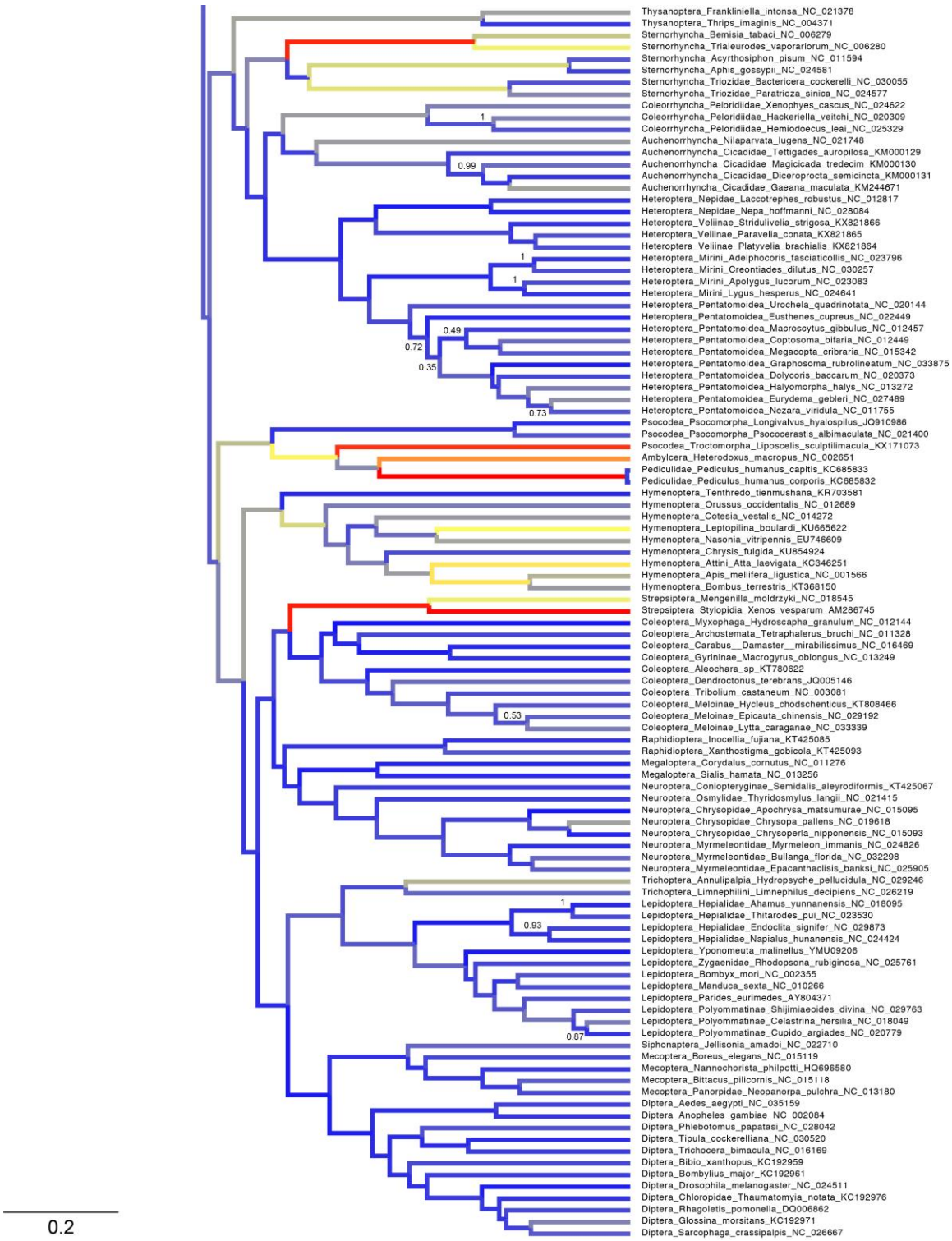


Figure 3.6. Bayesian phylogeny (ungrouped) with fixed topology from Misof et al. (2014) to estimate relative rates of COI substitution for Hexapoda. Branch color indicates relative rate (see legend). Unconstrained nodes have posterior probability support values. Scale bar units in substitutions/site/time.

Figure 3.6 (Continued)



References

Abascal, F., Zardoya, R., & Telford, M. J. (2010). TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*, *38*(suppl 2), W7–W13. doi: 10.1093/nar/gkq291

Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P., & Slowinski, J. B. (2002). Estimating Divergence Times from Molecular Data on Phylogenetic and Population Genetic Timescales. *Annual Review of Ecology and Systematics*, *33*(1), 707–740. doi: 10.1146/annurev.ecolsys.33.010802.150500

André, H. M., Ducarme, X., & Lebrun, P. (2002). Soil biodiversity: myth, reality or conning? *Oikos*, *96*(1), 3–24. doi: 10.1034/j.1600-0706.2002.11216.x

Bennett, K. R., Hogg, I. D., Adams, B. J., & Hebert, P. D. N. (2016). High levels of intraspecific genetic divergences revealed for Antarctic springtails: evidence for small-scale isolation during Pleistocene glaciation. *Biological Journal of the Linnean Society*, *119*, 166-178. doi: 10.1111/bij.12796

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, *41*, D36–D42. doi: 10.1093/nar/gks1195

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, *22*(3), 148–155. doi: 10.1016/j.tree.2006.11.004

Blackith, R. E. & Disney, R. H. L. (1988). Passive aerial dispersal during moulting in tropical Collembola. *Malayan Nature Journal*, *41*(4), 529–531.

Bouckaert, R. R. & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, *17*(42), 1–11. doi: 10.1186/s12862-017-0890-6

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. doi: 10.1371/journal.pcbi.1003537

Box, G. E. P. & Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis.* John Wiley & Sons, Inc.: Hoboken. doi: 10.1002/9781118033197

Britten, R. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science*, *231*(4744), 1393–1398. doi: 10.1126/science.3082006

Brower, A. V. Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences*, *91*(14), 6491–6495. doi: 10.1073/pnas.91.14.6491

Cabezas-Cruz, A., Valdés, J. J., Lancelot, J., & Pierce, R. J. (2015). Fast evolutionary rates associated with functional loss in class I glucose transporters of *Schistosoma mansoni*. *BMC Genomics*, *16*(1), 980. doi: 10.1186/s12864-015-2144-6

Carapelli, A., Frati, F., Fanciulli, P. P., & Dallai, R. (1995). Genetic differentiation of six sympatric species of *Isotomurus* (Collembola, Isotomidae): is there any difference in their microhabitat preference? *European Journal of Soil Biology*, *31*(2), 87–99.

Carapelli, A., Frati, F., Fanciulli, P. P., Nardi, F., & Dallai, R. (2005). Assessing species boundaries and evolutionary relationships in a group of south-western European species of *Isotomurus* (Collembola, Isotomidae) using allozyme data. *Zoologica Scripta*, *34*(1), 71–79. doi: 10.1111/j.1463-6409.2005.00174.x

Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, *17*(4), 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chen, L., Chen, P.-Y., Xue, X.-F., Hua, H.-Q., Li, Y.-X., Zhang, F., & Wei, S.-J. (2018). Extensive gene rearrangements in the mitochondrial genomes of two egg parasitoids, *Trichogramma japonicum* and *Trichogramma ostriniae* (Hymenoptera: Chalcidoidea: Trichogrammatidae). *Scientific Reports*, *8*(1), 7034. doi: 10.1038/s41598-018-25338-3

Christiansen, K. & Pike, E. (2002). Cretaceous Collembola (Arthropoda, Hexapoda) from the Upper Cretaceous of Canada. *Cretaceous Research*, *23*(2), 165–188. doi: 10.1006/cres.2002.0313

Christiansen, K. & Nascimbene, P. (2006). Collembola (Arthropoda, Hexapoda) from the mid Cretaceous of Myanmar (Burma). *Cretaceous Research*, *27*(3), 318–363. doi: 10.1016/j.cretres.2005.07.003

Cicconardi, F., Nardi, F., Emerson, B. C., Frati, F., & Fanciulli, P. P. (2010). Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology*, **1***9*(2), 386–400. doi: 10.1111/j.1365-294X.2009.04457.x

Cicconardi, F., Fanciulli, P. P., & Emerson, B. C. (2013). Collembola, the biological species concept and the underestimation of global species richness. *Molecular ecology*, *22*(21), 5382–5396. doi: 10.1111/mec.12472

Coppard, S. E. & Lessios, H. A. (2017). Phylogeography of the sand dollar genus Encope: implications regarding the Central American Isthmus and rates of molecular evolution. *Scientific Reports*, *7*(1), 11520. doi: 10.1038/s41598-017-11875-w

Coulson, S. J., Hodkinson, I. D., & Webb, N. R. (2003). Aerial dispersal of invertebrates over a high-Arctic glacier foreland: Midtre Lovénbreen, Svalbard. *Polar Biology*, *26*(8), 530–537. doi: 10.1007/s00300-003-0516-x

Coulson, S. J., Hodkinson, I. D., Webb, N. R., & Harrison, J. A. (2002). Survival of terrestrial soil-dwelling arthropods on and in seawater: implications for trans-oceanic dispersal. *Functional Ecology*, *16*(3), 353–356. doi: 10.1046/j.1365-2435.2002.00636.x

DeSalle, R. & Templeton, A. R. (1988). Founder Effects and the Rate of Mitochondrial DNA Evolution in Hawaiian *Drosophila*. *Evolution*, *42*(5), 1076–1084. doi: 10.1111/j.1558-5646.1988.tb02525.x

Ding, Y.-H., Yu, D.-Y., Guo, W.-B., Li, J.-N., & Zhang, F. (2018). Molecular phylogeny of *Entomobrya* (Collembola: Entomobryidae) from China: Color pattern groups and multiple origins. *Insect Science*, 1–11. doi: 10.1111/1744-7917.12559

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. doi: 10.1093/molbev/mss075

Emerson, B. C., Cicconardi, F., Fanciulli, P. P., & Shaw, P. J. A. (2011). Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1576), 2391–2402. doi: 10.1098/rstb.2011.0057

Felderhoff, K. L., Bernard, E. C., & Moulton, J. K. (2010). Survey of *Pogonognathellus* Börner (Collembola: Tomoceridae) in the Southern Appalachians Based on Morphological and Molecular Data. *Annals of the Entomological Society of America*, *103*(4), 472–491. doi: 10.1603/AN09105

Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, *27*(3), 613–635. doi: 10.1111/mec.14486

Frati, F., Carapelli, A., & Fanciulli, P. P. (1995). The genus *Isotomurus*: where molecular markers help to evaluate the importance of morphological characters for the diagnosis of species. *Polskie Pismo Entomologiczne*, *64*(1–4), 41–51.

Frati, F., Dell'Ampio, E., Casasanta, S., Carapelli, A., & Fanciulli, P. P. (2000). Large amounts of genetic divergence among Italian species of the genus *Orchesella* (Insecta, Collembola) and the relationships of two new species. *Molecular Phylogenetics and Evolution*, *17*(3), 456–461. doi: 10.1006/mpev.2000.0854

Frati, F., Fanciulli, P. P., & Dallai, R. (1994). Further acquisitions on systematic relationships within the genus *Orchesella* (Collembola, Entomobryidae) using allozymes. *Acta Zoologica Fennica*, *195*, 35–43.

Frean, M., Rainey, P. B., & Traulsen, A. (2013). The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1762), 20130211–20130211. doi: 10.1098/rspb.2013.0211

Freeman, J. A. (1952). Occurrence of Collembola in the air. *Proceedings of the Royal Entomological Society of London. Series A, General Entomology*, *27*(1–3), 28. doi: 10.1111/j.1365-3032.1952.tb00142.x

Garrick, R. C., Sands, C. J., Rowell, D. M., Hillis, D. M., & Sunnucks, P. (2007). Catchments catch all: long-term population history of a giant springtail from the southeast Australian highlands - a multigene approach. *Molecular Ecology*, *16*(9), 1865–1882. doi: 10.1111/j.1365-294X.2006.03165.x

Garrick, R. C., Rowell, D. M., Simmons, C. S., Hillis, D. M., & Sunnucks, P. (2008). Fine-scale phylogeographic congruence despite demographic incongruence in two low-mobility saproxylic springtails. *Evolution*, *62*(5), 1103–1118. doi: 10.1111/j.1558-5646.2008.00349.x

Geyer, J., Wilke, T., & Petzinger, E. (2006). The solute carrier family SLC10: more than a family of bile acid transporters regarding function and phylogenetic relationships. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *372*(6), 413–431. doi: 10.1007/s00210-006-0043-8

Gillooly, J. F., Allen, A. P., West, G. B., & Brown, J. H. (2005). The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences*, *102*(1), 140–145. doi: 10.1073/pnas.0407735101

Hafner, M. S., Sudman, P. D., Villablanca, F. X., Spradling, T. A., Demastes, J. W., & Nadler, S. A. (1994). Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, *265*(5175), 1087–1090. doi: 10.1126/science.8066445

Hawes, T. C., Worland, M. R., Bale, J. S., & Convey, P. (2008). Rafting in Antarctic Collembola. *Journal of Zoology*, 274(1), 44–50. doi: 10.1111/j.1469-7998.2007.00355.x

Hawes, T. C., Worland, M. R., Convey, P., & Bale, J. S. (2007). Aerial dispersal of springtails on the Antarctic Peninsula: implications for local distribution and demography. *Antarctic Science*, *19*(1), 3–10. doi: 10.1017/S0954102007000028

Hipsley, C. A. & Müller, J. (2014). Beyond fossil calibrations: Realities of molecular clock practices in evolutionary biology. *Frontiers in Genetics*, *5*, 1–11. doi: 10.3389/fgene.2014.00138

Hopkin, S. P. (1997). *Biology of springtails (Insecta: Collembola)*. Oxford University Press: New York.

Ho, S. Y. W. & Duchêne, S. (2014). Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology*, *23*(42), 5947–5965. doi: 10.1111/mec.12953

Ho, S. Y. W., Tong, K. J., Foster, C. S. P., Ritchie, A. M., Lo, N., & Crisp, M. D. (2015). Biogeographic calibrations for the molecular clock. *Biology Letters*, *11*(9), 20150194. doi: 10.1098/rsbl.2015.0194

Johnson, K. P., Cruickshank, R. H., Adams, R. J., Smith, V. S., Page, R. D. M., & Clayton, D. H. (2003). Dramatically elevated rate of mitochondrial substitution in lice (Insecta: Phthiraptera). *Molecular Phylogenetics and Evolution*, *26*(2), 231–242. doi: 10.1016/S1055-7903(02)00342-1

Kaltenpoth, M., Showers Corneli, P., Dunn, D. M., Weiss, R. B., Strohm, E., & Seger, J. (2012). Accelerated evolution of mitochondrial but not nuclear genomes of Hymenoptera: new evidence from crabronid wasps. *PloS one*, *7*(3), e32826. doi: 10.1371/journal.pone.0032826

Katoh, K., Kuma, K., Toh, H., & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, *33*(2), 511–518. doi: 10.1093/nar/gki198

Katz, A. D., Giordano, R., & Soto-Adames, F. N. (2015). Operational criteria for cryptic species delimitation when evidence is limited, as exemplified by North American *Entomobrya* (Collembola: Entomobryidae). *Zoological Journal of the Linnean Society*, *173*(4), 810–840. doi: 10.1111/zoj.12220

Kjer, K. M., Ware, J. L., Rust, J., et al. (2015). Response to Comment on "Phylogenomics resolves the timing and pattern of insect evolution." *Science*, *349*(6247), 487–487. doi: 10.1126/science.aaa7136

Kômoto, N., Yukuhiro, K., & Tomita, S. (2012). Novel gene rearrangements in the mitochondrial genome of a webspinner, *Aposthonia japonica* (Insecta: Embioptera). *Genome*, *55*(3), 222–233. doi: 10.1139/g2012-007

Lanfear, R., Thomas, J. A., Welch, J. J., Brey, T., & Bromham, L. (2007). Metabolic rate does not calibrate the molecular clock. *Proceedings of the National Academy of Sciences*, *104*(39), 15388–15393. doi: 10.1073/pnas.0703359104

Lanfear, R., Kokko, H., & Eyre-Walker, A. (2014). Population size and the rate of evolution. *Trends in Ecology & Evolution*, *29*(1), 33–41. doi: 10.1016/j.tree.2013.09.009

Mari Mutt, J. A. (1983). Collembola in amber from the Dominican Republic. *Proceedings of the Entomological Society of Washington*, *85*(3), 575–587.

Marshall, D. C., Hill, K. B. R., Moulds, M., Vanderpool, D., Cooley, J. R., Mohagan, A. B., & Simon, C. (2016). Inflation of Molecular Clock Rates and Dates: Molecular Phylogenetics, Biogeography, and Diversification of a Global Cicada Radiation from Australasia (Hemiptera: Cicadidae: Cicadettini). *Systematic Biology*, *65*(1), 16–34. doi: 10.1093/sysbio/syv069

Martin, A. P. & Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences*, *90*(9), 4087–4091. doi: 10.1073/pnas.90.9.4087

McGaughran, A., Hogg, I. D., & Stevens, M. I. (2008). Patterns of population genetic structure for springtails and mites in southern Victoria Land, Antarctica. *Molecular Phylogenetics and Evolution*, *46*(2), 606–618. doi: 10.1016/j.ympev.2007.10.003

Meusemann, K., von Reumont, B. M., Simon, S., Roeding, F., Strauss, S., Kuck, P., Ebersberger, I., Walzl, M., Pass, G., Breuers, S., Achter, V., von Haeseler, A., Burmester, T., Hadrys, H., Wagele, J. W., & Misof, B. (2010). A Phylogenomic Approach to Resolve the Arthropod Tree of Life. *Molecular Biology and Evolution*, *27*(11), 2451–2464. doi: 10.1093/molbev/msq130

Misof, B., Liu, S., Meusemann, K., et al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*(6210), 763–767. doi: 10.1126/science.1257570

Moore, J. C., Saunders, P., Selby, G., Horton, H., Chelius, M. K., Chapman, A., & Horrocks, R. D. (2005). The distribution and life history of *Arrhopalites caecus* (Tullberg): Order: Collembola, in Wind Cave, South Dakota, USA. *Journal of Cave and Karst Studies*, *67*(2), 110–119.

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M., & Taylor, D. R. (2010). Accelerated Mutation Accumulation in Asexual Lineages of a Freshwater Snail. *Molecular Biology and Evolution*, *27*(4), 954–963. doi: 10.1093/molbev/msp300

Nolan, L., Hogg, I. D., Stevens, M. I., & Haase, M. (2006). Fine scale distribution of mtDNA haplotypes for the springtail *Gomphiocephalus hodgsoni* (Collembola) corresponds to an ancient shoreline in Taylor Valley, continental Antarctica. *Polar Biology*, *29*(10), 813–819. doi: 10.1007/s00300-006-0119-4

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: The mid-aegean trench calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. doi: 10.1093/molbev/msq051

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough. *PLoS Biology*, *9*(3), e1000602. doi: 10.1371/journal.pbio.1000602

Porco, D., Bedos, A., Greenslade, P., Janion, C., Skarżyński, D., Stevens, M. I., Jansen van Vuuren, B., & Deharveng, L. (2012a). Challenging species delimitation in Collembola: cryptic diversity among common springtails unveiled by DNA barcoding. *Invertebrate Systematics*, *26*(26), 470–477. doi: 10.1071/IS12026

Porco, D., Potapov, M., Bedos, A., Busmachiu, G., Weiner, W. M., Hamra-Kroua, S., & Deharveng, L. (2012b). Cryptic diversity in the ubiquist species *Parisotoma notabilis* (Collembola, Isotomidae): a long-used chimeric species? *PloS one*, *7*(9), e46056. doi: 10.1371/journal.pone.0046056

Resch, M. C., Shrubovych, J., Bartel, D., Szucsich, N. U., Timelthaler, G., Bu, Y., Walzl, M., & Pass, G. (2014). Where Taxonomy Based on Subtle Morphological Differences Is Perfectly Mirrored by Huge Genetic Distances: DNA Barcoding in Protura (Hexapoda). *PLoS ONE*, *9*(3), e90653. doi: 10.1371/journal.pone.0090653

Rothschild, L. J. & Mancinelli, R. L. (2001). Life in extreme environments. *Nature*, *409*(6823), 1092–1101. doi: 10.1038/35059215

Sánchez-García, A. & Engel, M. S. (2016). Long-term stasis in a diverse fauna of Early Cretaceous springtails (Collembola: Symphypleona). *Journal of Systematic Palaeontology*, 1–25. doi: 10.1080/14772019.2016.1194575

Shao, R., Dowton, M., Murrell, A., & Barker, S. C. (2003). Rates of Gene Rearrangement and Nucleotide Substitution Are Correlated in the Mitochondrial Genomes of Insects. *Molecular Biology and Evolution*, *20*(10), 1612–1619. doi: 10.1093/molbev/msg176

Simonsen, V., Krogh, P. H., Filser, J., & Fjellberg, A. (1999). Three species of *Isotoma* (Collembola, Isotomidae) based on morphology, isozymes and ecology. *Zoologica Scripta*, *28*(3–4), 281–287. doi: 10.1046/j.1463-6409.1999.00025.x

Soto-Adames, F. N. (2002). Molecular phylogeny of the Puerto Rican Lepidocyrtus and *Pseudosinella* (Hexapoda: Collembola), a validation of Yoshii's "color pattern species". *Molecular phylogenetics and evolution*, *25*(1), 27–42. doi: 10.1016/S1055-7903(02)00250-6

Stevens, M. I., Greenslade, P., Hogg, I. D., & Sunnucks, P. (2006). Southern hemisphere springtails: Could any have survived glaciation of antarctica? *Molecular Biology and Evolution*, *23*(5), 874–882. doi: 10.1093/molbev/msj073

Stevens, M. I. & Hogg, I. D. (2003). Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. *Molecular Ecology*, *12*(9), 2357–2369. doi: 10.1046/j.1365-294X.2003.01907.x

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

Thomas, J. A., Welch, J. J., Lanfear, R., & Bromham, L. (2010). A Generation Time Effect on the Rate of Molecular Evolution in Invertebrates. *Molecular Biology and Evolution*, *27*(5), 1173–1180. doi: 10.1093/molbev/msq009

Thomas, J. A., Welch, J. J., Woolfit, M., & Bromham, L. (2006). There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. *Proceedings of the National Academy of Sciences*, *103*(19), 7366–7371. doi: 10.1073/pnas.0510251103

Timmermans, M. J. T. N., Ellers, J., Mariën, J., Verhoef, S. C., Ferwerda, E. B., & van Straalen, N. M. (2005). Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular ecology*, *14*(7), 2017–2024. doi: 10.1111/j.1365-294X.2005.02548.x

Tong, K. J., Duchene, S., Ho, S. Y. W., & Lo, N. (2015). Comment on "Phylogenomics resolves the timing and pattern of insect evolution." *Science*, *349*(6247), 487–487. doi: 10.1126/science.aaa5460

Trautwein, M. D., Wiegmann, B. M., Beutel, R., Kjer, K. M., & Yeates, D. K. (2012). Advances in Insect Phylogeny at the Dawn of the Postgenomic Era. *Annual Review of Entomology*, *57*(1), 449–468. doi: 10.1146/annurev-ento-120710-100538

von Saltzwedel, H., Scheu, S., & Schaefer, I. (2017). Genetic structure and distribution of *Parisotoma notabilis* (Collembola) in Europe: Cryptic diversity, split of lineages and colonization patterns. *PLOS ONE*, 12(2), e0170909. doi: 10.1371/journal.pone.0170909

von Saltzwedel, H., Scheu, S., & Schaefer, I. (2016). Founder events and pre-glacial divergences shape the genetic structure of European Collembola species. *BMC Evolutionary Biology*, *16*(1), 148. doi: 10.1186/s12862-016-0719-8

Weir, J. T. & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, *17*(10), 2321–2328. doi: 10.1111/j.1365-294X.2008.03742.x

Wessel, A., Hoch, H., Asche, M., von Rintelen, T., Stelbrink, B., Heck, V., Stone, F. D., & Howarth, F. G. (2013). Founder effects initiated rapid species radiation in Hawaiian cave planthoppers. *Proceedings of the National Academy of Sciences*, *110*(23), 9391–9396. doi: 10.1073/pnas.1301657110

Wilke, T., Schultheiß, R., & Albrecht, C. (2009). As Time Goes by: A Simple Fool's Guide to Molecular Clock Approaches in Invertebrates. *American Malacological Bulletin*, *27*(1–2), 25–45. doi: 10.4003/006.027.0203

Xia, X. (2017). DAMBE6: New Tools for Microbial Genomics, Phylogenetics, and Molecular Evolution. *Journal of Heredity*, *108*(4), 431–437. doi: 10.1093/jhered/esx033

Xia, X. & Lemey, P. (2009). Assessing substitution saturation with DAMBE. *The phylogenetic handbook: a practical approach to DNA and protein phylogeny* (ed. by P. Lemey, M. Salemi, and A.-M. Vandamme), pp. 615–630. Cambridge University Press, Cambridge. doi: 10.1017/CBO9780511819049.022

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003). An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, *26*(1), 1–7. doi: 10.1016/S1055-7903(02)00326-3

Yang, Z. (1996). Maximum-likelihood models for combined analyses of multiple sequence data. *Journal of Molecular Evolution*, *42*(5), 587–596. doi: 10.1007/BF02352289

CHAPTER 4

A new endemic species of *Willowsia* from Florida (USA) and descriptive notes on all New World *Willowsia* (Collembola: Entomobryidae)[2]

Abstract

Four species of *Willowsia* have been reported from the Americas (*W. buski*, *W. jacobsoni*, *W. mexicana*, and *W. nigromaculata*), and to date, *W. mexicana* is the only member of the genus endemic to the New World. Here, *Willowsia pyrrhopygia* **sp. nov.** from Florida is described. Like *W. mexicana*, this new species has a native New World distribution and uninterrupted rib scale type, but can be separated by color pattern and chaetotaxy. Dorsal head chaetotaxy and other descriptive notes are provided to compliment to descriptions for *W. buski*, *W. jacobsoni*, *W. mexicana*, and *W. nigromaculata*. Comparative morphological analysis also reveals two unique character states among Entomobryinae—the outer maxillary lobe with two (not three) sublobal hairs and the absence of labial triangle seta r—shared only by endemic New World *Willowsia* and *Americabrya*, providing prima facie evidence for their independent evolution from a common New World ancestor.

Introduction

The genus *Willowsia* Shoebotham, 1917 includes species in the subfamily Entomobryinae Schäffer, 1896 (*sensu* Zhang & Deharveng, 2015 and Zhang, 2016), that have 8+8 eyes, bidentate mucro with a smooth basal spine and apically acuminate scales on the head and body, but lack antennal sub-segmentation, dental scales, dental spines,

[2] Katz, A. D. (2017) A new endemic species of *Willowsia* from Florida (USA) and descriptive notes on all New World *Willowsia* (Collembola: Entomobryidae). *Zootaxa*, **4350**, 549–562.
doi: 10.11646/zootaxa.4350.3.7

and differentiated “smooth” setae on the hind tibiotarsus. Thirty-five species of *Willowsia* have been described worldwide, including four species with New World distributions: *W. buski* (Lubbock, 1870), *W. jacobsoni* (Börner, 1913), *W. mexicana* Zhang, Palacios-Vargas & Chen, 2007, and *W. nigromaculata* (Lubbock, 1873).

Zhang *et al*. (2011) and Pan & Zhang (2016) noted the diversity of *Willowsia* scale rib sculpturing and speculated on its potential evolutionary significance, suggesting that different types of scales may correspond to independent lineages of *Willowsia*. Subsequent molecular analysis supports the growing consensus that *Willowsia* is polyphyletic, having multiple independent origins within Entomobryinae (Zhang *et al*. 2014b; 2015; 2016; Katz *et al*. 2015a; Zhang & Deharveng 2015). To date, *W. mexicana* is the only species in the genus endemic to the New World, and it displays a rare scale morphology in which ribs are complete, without interruptions or brakes along their length. Only two species from New Caledonia (*W. neocaledonia* Zhang, Bedos & Deharveng, 2014a and *W. nigra* Zhang, Bedos & Deharveng, 2014a) share the uninterrupted rib scale type with *W. mexicana*, suggesting that they may share a common independent origin (Zhang *et al*. 2014b).

In this study, a new species of *Willowsia* with the uninterrupted rib scale type is described from Florida. Dorsal head chaetotaxy and other descriptive notes are also provided as complements to the descriptions of *W. buski*, *W. jacobsoni*, *W. mexicana*, and *W. nigromaculata*.

Methods

Willowsia specimens were cleared in Nesbitt's solution and slide-mounted in Hoyer's medium (Mari Mutt 1979) for light microscopy observation. Illustrations were hand-drawn with a camera lucida, scanned, and digitized with Adobe Illustrator CC. Slide mounted specimens of *Americabrya arida* Christiansen & Bellinger 1980 (prepared by Mari Mutt & Palacios-Vargas 1987) were also examined for morphological comparisons. All specimens observed for this study, including type material, were deposited at the Illinois Natural History Survey (INHS) in Champaign, IL.

Descriptions of dorsal body chaetotaxy follow the nomenclature established by Szeptycki (1979); the dorsal head chaetotaxy follows Jordana and Baquero (2005) and Soto-Adames (2008). All descriptions of labial chaetotaxy follow the nomenclature of Chen & Christiansen (1993). Descriptions of tergal S-chaetotaxy follow Zhang & Deharveng (2015). Scale types follow those described by Zhang *et al*. (2011): 1, spinulate type (broad scales with many short and equal length spinules, regularly spaced); 2, short rib type (elongate, pointed scales with short, unequal length and irregularly spaced ribs); 3, long basal rib type (broad, pointed scales with long basal ribs and short distal ribs converging distally); 4, uninterrupted rib type (broad scales, usually pointed, with long, unbroken parallel ribs that do not converge distally).

Abbreviations: Ant. I–IV—antennal segments 1–4; Th. II—mesothorax; Th. III—metathorax; Abd. I–VI—abdominal segments 1–6; mac—macroseta(e); mic—microseta(e); ms—S-microchaeta(e); sens—ordinary S-chaeta(e).

Taxonomy

Genus *Willowsia* Shoebotham, 1917

This genus is characterized by having 8+8 eyes within black or dark blue patches of pigment, a bidentate mucro with a smooth basal spine, and broad scales mostly pointed apically and with sculpturing ranging from many short spinules to more than two long ribs (Zhang *et al*. 2011), but lack scales and spines on the dens, antennal sub-segmentation, and differentiated "smooth" setae on the inner surface of the hind tibiotarsus.

In addition, all species of *Willowsia* found in the Americas share the following characters: apical antennal bulb usually unilobed, rarely bilobed; sensilla 2–3 (i.e., main sensilla) of the apical sense organ of Ant. III smooth and blunt; eyes G and H smallest; five interocular setae (pqrst); labral setae 5,5,4 and smooth; subapical and apical setae of the outer maxillary lobe smooth and subequal; labial palp with 5 smooth proximal setae; post labial setae ciliate; unguis with one outer, two lateral and four inner teeth; unguiculus lanceolate, outer edge serrate; tenent hairs spatulate; tenaculum with 4+4 teeth and one large mac; mucro bidentate, basal spine reaching apex of subapical tooth; accessory mic of bothriotrichal complexes on Abd. II–IV unmodified; tergal S-chaetotaxy (Th. II – Abd. VI) ms as 1,0|1,0,1,0,0,0, sens as 2,2|1,2,2,+,3,0 (Abd. IV with high and variable number of sens).

***Willowsia buski* (Lubbock, 1870)**

Figs 4.1a, 4.2h, 4.3a

Material examined. USA: Wisconsin, Sauk Co., Parfrey's Glen Natural Area, off Bluff Rd., 43.421102, -89.636841, on bark, 12.vi.2011, A. Katz, AK11-52, 4 on slides (INHS 810,141–810,144); Illinois, Champaign Co., Urbana, Meadowbrook organic gardens, 40.08066, -88.20836, hand collected from mulch and leaf litter, 14.iii.2012, A. Katz, AK12-24, 1 on slide (INHS 810,145), 6 in ethanol (INHS 810,146); Illinois, Champaign Co., Urbana, Brownfield Woods, 40.14326, -88.16441, on bark from dead tree, 10.ix.2009, F. Soto-Adames, 1 on slide (INHS 810,147).

Complement to description based on American specimens. Habitus as in Figure 4.1a. Scales long basal rib type (Fig. 4.2h). Scales absent on antennae, legs, and furcula. Dorsal chaetotaxy of the head as in Figure 4.3a: row An with 5 mac; anterior mac A0, A2, A3, A5 present; 3+3 median mac, M1, M2, M4 present, M3 always absent; sutural mac S0, S2, S3, S4i, S4, S5i, S5 present, S6 as mic; posterior mac Ps5 present. Outer maxillary lobe with 3 sublobal hairs. Prelabral setae ciliate. Labral papillae each with single seta-like projection. Labial triangle as MrEL1L2A1–5, r always present. Lateral process of labial palp straight, blunt, and reaching tip of labial papilla E. Cephalic groove with 6+6 ciliate setae.

Ecology. This species occurs under tree bark, in leaf litter, and on vegetation.

Distribution. Cosmopolitan, including the Americas.

Remarks. *Willowsia buski* is the only member of the genus with the following combination of characters: uniformly blue color pattern (Fig. 4.1a); long basal rib type scales (Fig. 4.2h); scales absent on antennae, legs, and furcula; labral papillae with single,

seta-like projections; prelabral setae ciliate; labial triangle seta r present; outer maxillary lobe with 3+3 sublobal hairs; dorsal head with 3+3 median mac, 6+6 sutural mac, mac S0 and Ps5 present (Fig. 4.3a); Th. II with 2+2 median mac and 8+8 posterior mac; Abd. I with 2+2 mac; Abd. II with 3+3 inner mac; Abd. III with 2+2 inner and 3+3 lateral mac; and Abd. IV with 5+5 inner mac.

In the Americas, this species is most similar to *W. nigromaculata*: both have basal long rib scale type (Zhang *et al*., 2011); scales absent on antennae, legs, and manubrium; labral papillae each with a single seta-like projection; ciliate prelabral setae, labial triangle r seta present, 3+3 sublobal hairs of outer maxilla; 3+3 median mac, 6+6 sutural mac, and S0 mac present on dorsal head; 3+3 inner mac on Abd. II; and 3+3 lateral mac on Abd. IV. They can be differentiated by color pattern and chaetotaxy: *W. buski* is uniformly dark blue or purple without a clear pattern (*W. nigromaculata* has a distinct dorsal pigment pattern); has dorsal head mac Ps5 (absent in *W. nigromaculata*), has 2+2 median mac on Th. II (*W. nigromaculata* has 1+1), 2+2 mac on Abd. I (*W. nigromaculata* has 3+3), 2+2 inner mac on Abd. III (*W. nigromaculata* has 3+3), and 5+5 inner mac on Abd. IV (*W. nigromaculata* has 7+7). See Table 4.1 for additional diagnostic characters.

The specimens observed for this study matched morphological descriptions, including dorsal chaetotaxy, by Christiansen & Bellinger (1998) and Zhang *et al*. (2011).

***Willowsia jacobsoni* (Börner, 1913)**

Figs 4.1c–d, 4.2g, 4.3b

Material examined. USA: Puerto Rico, Aguada, Coloso Sugar Cane Mill, on sugar cane litter, 16.vi.1999, F. Soto-Adames, 2 on slides (INHS 810,148–810,149); Puerto Rico, Aguada, Coloso Sugar Cane Mill, on sugar cane litter, 17.vi.1999, F. Soto-Adames, 4 females, 2 (1 male) on slides (INHS 810,150–810,155), 15 in ethanol (INHS 810,156).

Complement to description based on American specimens. Habitus as in Figs 4.1c–d. Scales long basal rib type (Fig. 4.2g) and absent on antennae, legs, and furcula. Dorsal chaetotaxy of the head as in Figure 4.3b: row An with 6 mac; anterior mac A0, A2, A3, A5 present; 3+3 median mac, M1, M2, M4 present, M3 always absent; 4+4 sutural mac, S2, S3, S4, S5 present, S6 as mic; post-sutural mac Ps5 absent. Outer maxillary lobe with 3 sublobal hairs. Prelabral setae ciliate. Labral papillae each with multiple seta-like projections. Labial triangle as MrEL1L2A1–5, r always present. Lateral process of labial palp straight, blunt, reaching tip of labial papilla E. Cephalic groove with 5+5 ciliate setae. Th. II without median mac and 2+2 posterior mac. Abd. I with 1+1 mac.

Ecology. This species occurs in leaf litter, and is abundant in agroecosystems such as sugar cane fields and banana leave axils.

Distribution. Pantropical, including Hawaii and Puerto Rico.

Remarks. *Willowsia jacobsoni* is the only member of the genus with the following combination of characters: color pattern as in Figs 4.1c–d; long basal rib type scales (Fig. 4.2g); scales absent on antennae, legs, and furcula; labral papillae with multiple projections; prelabral setae ciliate; labial triangle seta r present; outer maxillary

lobe with 3+3 sublobal hairs; dorsal head with 3+3 median mac, 4+4 sutural mac, mac S0 and Ps5 absent (Fig. 4.3b); Th. II with 0+0 median mac and 2+2 posterior mac; Abd. I with 1+1 mac; Abd. II with 2+2 inner mac; Abd. III with 2+2 inner and 3+3 lateral mac; and Abd. IV with 5+5 inner mac.

In scale morphology and scale distribution, this species is very similar *W. buski* and *W. nigromaculata*: they all have long basal rib type scales that are absent on the antennae, legs, and manubrium. However, *W. jacobsoni* can be differentiated from *W. buski* and *W. nigromaculata* by color pattern and head and body chaetotaxy: head mac Ps5 is absent in *W. jacobsoni* (present in *W. buski* and *W. nigromaculata*) and *W. jacobsoni* has 3+3 inner mac on Abd. III (*W. buski* and *W. nigromaculata* have 2+2). See Table 4.1 for additional diagnostic characters.

The material observed for this study is in general agreement with Mari Mutt's (1981) detailed morphological description, except Th. II mac a5 is incorrectly labeled as m4 in Mari Mutt (1981).

***Willowsia mexicana* Zhang, Palacios-Vargas & Chen, 2007**

Figs 4.2d–f, 4.3d

Material examined. *Paratype*, Mexico: Oaxaca, Oaxaca, Mexico, 06/05/1994, ex. *Pinus oaxacana*, E. López Col., 19127, 1 on slide. *Additional material*, Mexico: Cd. de México, L. NW. iztaccihuatl, alt. 3020 m, a:1 m:1, 19057, 27 agosto 2005, 1 on slide; Cd. de México, L. NW. iztaccihuatl, alt. 3440 m, a:6 m:6, 19097, 27 agosto 2005, 1 on slide; Veracruz, EBLT EBLT, 18°58.56N, 95°4'35.62''O, Trampa Pitfall 2, 248 msnm Selva Alta perennifolia, M. Madora, col, 16–21IV–2015, 2 on slides.

Complements to description. Scales present on dorsal side of Ant. I and Ant. II, and middle and hind legs (coxae, trochanter, femora). Outer maxillary lobe with 2 sublobal hairs. Ribs on scales straight, parallel, not converging distally, most running uninterrupted from base to tip (Figs 4.2d–f); some scales with ribs interrupted near tip (Fig. 4.2d). Dorsal chaetotaxy as described in Zhang *et al*. (2007), except dorsal head mac M1 sometimes present (Fig. 4.3d) and Th. II with 4+4 posterior mac (p3–6).

Ecology. This species occurs under tree bark, mangrove leaf litter, and has been collected at high elevations (750–3500 m above sea level).

Distribution. Mexico.

Remarks. *Willowsia mexicana* is the only member of the genus with the following combination of characters: Abd. IV–VI with evenly dark pigment; uninterrupted rib type scales (Figs 4.2d–f); scales present on antennae, legs, and furcula; labral papillae with multiple projections; prelabral setae smooth; labial triangle seta r absent; outer maxillary lobe with 2+2 sublobal hairs; dorsal head with 3+3 (rarely 4+4) median mac, 5+5 sutural mac, mac S0 absent, and mac Ps5 present (Fig. 4.3d); Th. II with 0+0 median mac and 3+3 posterior mac; Abd. I with 1+1 mac; Abd. II with 2+2 inner mac; Abd. III with 2+2 inner and 3+3 lateral mac; and Abd. IV with 4+4 inner mac.

This species is most similar to *W. pyrrhopygia* **sp. nov.** (differences between the two species are listed in the remarks for *W. pyrrhopygia* **sp. nov.**).

W. mexicana was previously described as having scales with only uninterrupted ribs (Zhang *et al*., 2007, 2014a), but new observations of the type material show some larger scales do have broken ribs near the apex. *Willowsia mexicana, W. nigra, W.*

neocaledonia and *W. pyrrhopygia* **sp. nov.** are the only species in this genus with uninterrupted rib type scales.

Zhang *et al*. (2007) also describe scales as present on the ventral face of the manubrium, but do not specify scale morphology (e.g., broad vs. thin). Although, I was unable to confirm the presence of scales (i.e., typical broad scales like those observed on body) on the ventral face of the manubrium in the type material, I did observe setae-like or narrow scales, similar to those observed on the ventral face of the manubrium in other *Willowsia* species (Zhang 2015).

***Willowsia nigromaculata* (Lubbock, 1873)**

Figs 4.1b, 4.2i, 4.3c

Material examined. USA: Illinois, Pope Co., Dixon Springs Ag. Center, 37.43435, -88.66718, elev. 524ft, hand collected on pavement on dying *Arilus cristatus*, 23Sep2011, A. Katz, adk11-159, 6 on slides (INHS 810,157–810,162), 30 in ethanol (INHS 810,163); Illinois, Urbana, 19 Montclair Rd., 40.094873, -88.206853, 6.vii.2008, by kitchen sink, F. Soto-Adames, 1 on slide (INHS 810,164); Illinois, Urbana, 19 Montclair Rd., 40.094873, -88.206853, 2012, F. Soto-Adames, 1 on slide (INHS 810,165).

Complement to description based on American specimens. Habitus as in Figure 4.1b. Scales long basal rib type (Fig. 4.2i) and absent on antennae, legs, and furcula. Dorsal chaetotaxy of the head as in Figure 4.3c: row An with 5 mac; anterior mac A0, A2, A3, A5 present; 3+3 median mac, M1, M2, M4 present, M3 always absent; sutural mac, S0, S2, S3, S4i, S4, S5i, S5 present, S6 as mic; posterior mac Ps5 absent. Outer maxillary lobe with 3 sublobal hairs. Prelabral setae ciliate. Labral papillae each

with single seta-like projection. Labial triangle as MrEL1L2A1–5, r always present. Lateral process of labial palp thin, curved, blunt, reaching tip of labial papilla E. Cephalic groove with 8+8 ciliate setae.

Ecology. This species occurs in leaf litter and is also considered a nuisance species in North America as it is commonly found inside homes.

Distribution. Cosmopolitan, including North and Central America.

Remarks. *Willowsia nigromaculata* is the only member of the genus with the following combination of characters: color pattern as in Figure 4.1b; long basal rib scale type (Fig. 4.2i); scales absent from antennae, legs, and furcula; labral papillae with a single seta-like projection; prelabral setae ciliate; labial triangle seta r present; outer maxillary lobe with 3+3 sublobal hairs; dorsal head with 3+3 median mac, 6+6 sutural mac, mac S0 present, and mac Ps5 absent (Fig. 4.3c); Th. II with 1+1 median mac and 7+7 posterior mac; Abd. I with 3+3 mac; Abd. II with 3+3 inner mac; Abd. III with 3+3 inner and 3+3 lateral mac; and Abd. IV with 7+7 inner mac.

In the Americas, this species is most similar to *W. buski*: both have basal long rib scale type (Zhang *et al*., 2011); scales absent on antennae, legs, and manubrium; labral papillae each with a single seta-like projection; ciliate prelabral setae, labial triangle r seta present, 3+3 sublobal hairs of outer maxilla; 3+3 median mac, 6+6 sutural mac, and mac S0 present dorsally on head; 3+3 inner mac on Abd. II; and 3+3 lateral mac on Abd. IV. They can be differentiated by color pattern and chaetotaxy: *W. nigromaculata* has a distinct dorsal pigment pattern, while *W. buski* is uniformly dark blue or purple without a clear pattern; dorsal head mac Ps5 is absent in *W. nigromaculata* (present in *W. buski*); 1+1 median mac on Th. II (*W. buski* has 2+2); 3+3 mac on Abd. I (*W. buski* has 2+2);

3+3 inner mac on Abd. III (*W. buski* has 2+2), and 7+7 inner mac on Abd. IV (*W. buski* has 5+5). See Table 4.1 for additional diagnostic characters.

The specimens observed for this study matched morphological descriptions, including dorsal chaetotaxy, by Christiansen & Bellinger (1998) and Zhang *et al*. (2011).

***Willowsia pyrrhopygia* sp. nov. Katz**

Willowsia n. sp. 1 Katz *et al*. 2015a

Willowsia sp. nov. 1 Katz *et al*. 2015b

Figs 4.2a–c, 4.3e, 4.4–4.5

Type material. *Holotype*, USA, Florida, Okeechobee Co., Kissimmee Prairie Preserve State Park, 27.5835, -81.05187, on vegetation, 10.viii.2011, A. Katz & J. Cech, AK11-123, 1 male on slide (INHS 810,166).

Allotype, USA, Florida, Taylor Co., Econfina River State Park, 30.059533, -83.907350, on vegetation and twigs, 9.viii.2011, A. Katz, AK11-117, 1 female on slide (INHS 810,167).

Paratypes, USA: Florida, Okeechobee Co., Kissimmee Prairie Preserve State Park, 27.5835, -81.05187, on vegetation, 10.viii.2011, A. Katz & J. Cech, AK11-123, 7 in ethanol (INHS 810,186); Florida, Okeechobee Co., Kissimmee Prairie Preserve State Park, 27.5835, -81.05187, hand collected from bark, 10.viii.2011, A. Katz & J. Cech, AK11-124, 1 on slide (INHS 810,187), 5 in ethanol (INHS 810,188); Florida, Taylor Co., Econfina River State Park, 30.059533, -83.907350, 10.viii.2011, A. Katz, hand collected from bark, adk11-116, 1 on slide (INHS 810,189); Florida, Taylor Co., Econfina River

State Park, 30.059533, -83.907350, on vegetation and twigs, 9.viii.2011, A. Katz, AK11-117, 5 in ethanol (INHS 810,190).

Additional material, USA: Florida, Citrus Co., Chassahowitzka National Wildlife Refuge, W Burnt Bridge Rd., 28.75996, -82.57583, elev. 9ft, hand collected on epiphytes, 12.viii.2011, A. Katz & J. Cech, AK11-135, 2 (1 female) on slides (INHS 810,168–810,169), 17 in ethanol (INHS 810,170); Florida, Citrus Co., Chassahowitzka National Wildlife Refuge, W Burnt Bridge Rd, 28.75996, -85.57583, on vegetation, 12.viii.2011, A. Katz & J. Cech, AK11-134, 1 female on slide (INHS 810,171), 25 in ethanol (INHS 810,172); Florida, Citrus Co., Chassahowitzka National Wildlife Refuge, W Burnt Bridge Rd, 28.75996, -85.57583, hand collected from leaf litter, 12.viii.2011, A. Katz & J. Cech, AK11-133, 1 in ethanol (INHS 810,173); Florida, Duval Co., Jacksonville, Chicopit Bay, West off side of Highway 1A, 30.376267, -81.436500, elev. 9ft, hand collected beating vegetation, 12.viii.2011, A. Katz & J. Cech, adk11-140, 4 in ethanol (INHS 810,174); Florida, Liberty Co., Apalochicola National Forest, 30.176650, -84.677233, 9.viii.2011, A. Katz, hand collected from bark, adk11-114, 1 on slide (INHS 810,175); Florida, Liberty Co., Apalochicola National Forest, 30.176650, -84.677233, 9.viii.2011, A. Katz, leaf litter, adk11-115, 1 in ethanol (INHS 810,176); Florida, Marion Co., Ocala National Forest, N of CR-314, 29.30076, -81.84686, elev. 105ft, hand collected under bark, 12.viii.2011, A. Katz & J. Cech, adk11-137, 4 in ethanol (INHS 810,177); Florida, Marion Co., Ocala National Forest, N of CR-314, 29.30076, -81.84686, on vegetation, 12.viii.2011, A. Katz & J. Cech, AK11-138, 1 female on slide (INHS 810,178), 17 in ethanol (INHS 810,178); Florida, Marion Co., Ocala National Forest, N of CR-314, 29.30076, -81.84686, elev. 105ft, hand collected in leaf litter, 12.viii.2011, A. Katz & J.

Cech, adk11-139, 1 in ethanol (INHS 810,180); Florida, Miami-Dade Co., Everglades National Park, Pinelands Trail, 25.42298, -80.67965, on vegetation, 11.viii.2011, A. Katz & J. Cech, AK11-126, 1 on slide (INHS 810,181), 14 in ethanol (INHS 810,182); Florida, Miami-Dade Co., Everglades National Park, Pinelands Trail, 25.42298, -80.67965, hand collected from bark, 11.viii.2011, A. Katz & J. Cech, AK11-127, 1 in ethanol (INHS 810,183); Florida, Miami-Dade Co., Big Cypress National Preserve, County Rd 94, 25.760167, -81.0386, elev. 4ft, hand collected beating vegetation, 11.viii.2011, A. Katz & J. Cech, adk11-130, 2 in ethanol (INHS 810,184); Florida, Miami-Dade Co., Big Cypress National Preserve, County Rd 94, 25.760167, -81.0386, elev. 4ft, hand collected from epiphytes, 11.viii.2011, A. Katz & J. Cech, adk11-132, 2 in ethanol (INHS 810,185);

Etymology. The specific epithet is derived from the Greek adjective πυρρός, -ά, -όν, Latinized *pyrrhus*, *-a*, *-um*, 'flame-colored, yellowish-red'; the Greek feminine noun, πῡγή, Latinized *pyge*, 'rump, buttocks'; and the Greek adjective-forming suffix -ιος, Latinized *-ius*, *-a*, *-um*, 'of, pertaining to'. The name references the conspicuous reddish-orange color of Abd. VI.

Description. *Body shape and color pattern*. Length up to 1.6 mm. Body dorso-ventrally -flattened (Figs 4.4a–b). Background color white with dark purple pigment forming two lateral bands from eye spot through Abd. I; on distal end of Ant I–III, most of Ant. IV deep purple; distal end of coxa, femora, and tibiotarsus dark purple; manubrium, and lateral margins of Abd. II–VI with small purple patches. Head and body speckled with light orange pigment. Abd. VI usually with bright orange pigment. Dens

without pigment. Living specimens often look dark because of a dorsal cover of black or dark brown scales.

Scales. Scales as uninterrupted ribbed type: broad, rounded, and usually apically acuminate (Figs 4.2a–c). Larger scales with straight parallel edges (Figs 4.2b–c). Nearly all scales with straight, parallel, uninterrupted ribs from base to tip that do not converge distally. Ribs rarely interrupted near tip in larger scales. Scales present dorsally on Ant. I–II, head, body; present on trochanter and femora of middle and hind legs and ventral face of manubrium; absent from Ant. III–IV, ventral side of head, ventral tube, and fore legs.

Head. Apical bulb of Ant. IV simple, sometimes bilobed. Dorsal chaetotaxy of head as in Figure 4.3e: row An with 6 mac; anterior mac A0, A2, A3, A5 present; median mac M1–4 present; sutural mac S2, S4, S5 present, S5i mac sometimes present, S6 as mic; posterior mac Ps5 present. Prelabral setae smooth. Labral papillae with 2–3 projections (Fig. 4.5b). Outer maxillary lobe with 2 sublobal hairs (Fig. 4.5a). Labial triangle as MEL1L2A1–5, r always absent (Fig. 4.5c). Lateral process of labial palp thick, blunt, not reaching tip of labial papilla E (Fig. 4.5d). Cephalic groove with 4+4 ciliate setae.

Thorax. Dorsal chaetotaxy of Th. II with 3 mac: a5, p3, p5 (Fig. 4.4c). Th. III with 5 mac: p3, p6, m5, m6, a6 (Fig. 4.4d). Trochanteral organ with triangular setal pattern with up to 26 smooth spiny setae (Fig. 4.5h). Hind claw complex as in Figure 4.5e.

Abdomen. Abd. I with 1 mac (Fig. 4.4e). Abd. II with 3 mac: m3, m3e, m5 (Fig. 4.4f). Abd. III with 5 mac: a2, m3, p6, pm6, am6 (Fig. 4.4g). Abd. IV with 4 mac internal to bothriotricha and at least 7 lateral mac (Fig. 4.4h). Ventral tube with 15–20 ciliate

setae on each side of anterior face; 5+5 small smooth and 1+1 larger smooth setae on posterior face; lateral flaps with 7+7 smooth setae. Tenaculum as in Figure 4.5f. Mucro as in Figure 4.5g.

DNA Barcode. GenBank KM610131

Ecology. This species occurs under tree bark, in leaf litter, and on grasses, shrubs, epiphytes and other vegetation.

Distribution. *Willowsia pyrrhopygia* **sp. nov.** is known only from Florida, USA. It is very common and abundant throughout the state.

Remarks. *Willowsia pyrrhopygia* **sp. nov.** is the only member of the genus with the following combination of characters: color pattern as in Figs 4.4a–b; uninterrupted rib scale type (Figs 4.2a–c); scales present on antennae, manubrium, and legs; labral papillae with multiple projections (Fig. 4.5b); prelabral setae smooth; labial triangle seta r absent (Fig. 4.5c); outer maxillary lobe with 2+2 sublobal hairs (Fig. 4.5a); dorsal head with 4+4 median mac, 4+4 sutural mac, mac S0 absent, and mac ps5 present (Fig. 4.3e); Th. II with 0+0 median mac and 2+2 posterior mac (Fig. 4.4c); Abd. I with 1+1 mac (Fig. 4.4e); Abd. II with 2+2 inner mac (Fig. 4.4f); Abd. III with 2+2 inner mac and 3+3 lateral mac (Fig. 4.4g); and Abd. IV with 4+4 inner mac (Fig. 4.4h).

It is most similar to *W. mexicana* in morphology and native distribution: both appear to be the only *Willowsia* species endemic to the New World, lacking labial triangle seta r, with 2 sublobal hairs, and share similarly reduced dorsal chaetotaxy. However, they can be separated by color pattern and dorsal chaetotaxay: *W. pyrrhopygia* **sp. nov.** has only small patches of dark blue pigment along the lateral margins of Abd. IV–VI, and Abd. VI has conspicuous orange pigment (Figs 4.4a–b), whereas, in *W.*

mexicana, Abd. IV–VI are mostly covered with dark blue pigment (See figure 8 in Zhang *et al*., 2007); *W. pyrrhopygia* **sp. nov.** always has 4+4 median mac on head, whereas *W. mexicana* usually has 3+3 (M1 usually absent); *W. pyrrhopygia* **sp. nov.** has 4+4 (sometimes 3+3) sutural mac on head, lacking mac S3 and sometimes lacking S5i, whereas *W. mexicana* has 5+5, with S3 and S5i always present; *W. pyrrhopygia* **sp. nov.** has 2+2 posterior mac on Th. II (p3, p5), whereas *W. mexicana* has 3+3 (p3–p5).

Two species from New Caledonia, *W. neocaledonia* and *W. nigra*, are also similar to *W. pyrrhopygia* **sp. nov.** and *W. mexicana*, and together, these four species comprise the uninterrupted rip scale type group (*sensu* Zhang *et al*., 2011): they all have reduced dorsal chaetotaxy; smooth prelabral setae; and labral papillae with multiple projections. However, *W. pyrrhopygia* **sp. nov.** lacks labial triangle seta r, has scales on the legs, and has only 2+2 hairs on the sublobal plate of the outer maxillary lobe, whereas in *W. neocaledonia* and *W. nigra* labial triangle seta r is present, scales are absent on the legs, and there are 3+3 hairs on the sublobal plate of the outer maxillary lobe. See Table 4.1 for additional diagnostic characters.

The genus *Americabrya* Mari Mutt & Palacios-Vargas, 1987, also endemic to the New World, has scales with similar (yet fewer) uninterrupted ribs and shares similar chaetotaxy with native New World *Willowsia*. Morphological comparisons among native New World *Willowsia* and *Americabrya* may provide additional insight into the complex evolutionary history of scaled Entomobryinae. It is apparent that *W. pyrrhopygia* **sp. nov.** and *W. mexicana* share many morphological similarities with *Americabrya arida* (See Table 4.1) that include character states that are unique, or at least undocumented, for Entomobryinae (i.e., 2 hairs on the sublobal plate of the outer maxilla and the absence of

labial seta r) suggesting that *W. pyrrhopygia* **sp. nov.** and *W. mexicana* may be more closely related to *Americabrya* than to their Asian congeners. This hypothesis is supported by recent molecular phylogenetic studies (Zhang *et al*. 2014b; 2015; 2016; Katz *et al*. 2015a; Zhang & Deharveng 2015) that dispute *Willowsia*'s monophyly, proposing that some species of *Willowsia* may have descended from independent lineages of *Entomobrya*. In this case, *W. pyrrhopygia* **sp. nov.**, *W. mexicana* and *Americabrya* spp., all endemic to North America, may share a New World *Entomobrya* ancestor. Furthermore, Zhang *et al*. (2011) suggests that differences in scale morphology (i.e., rib sculpturing and shape) may provide valuable insight regarding the evolution and systematics of *Willowsia*. For example, all *Willowsia* species with uninterrupted rib type scales may share common ancestry with *Americabrya* which have smaller and thinner, seta-like scales with only 2 uninterrupted ribs (Fig. 4.2j). However, a robust, species-level molecular phylogeny is needed to evaluate these hypotheses in the future.

Acknowledgments

Special thanks to Felipe Soto-Adames, Steve Taylor, and Sam Heads for providing helpful comments and access to laboratory equipment; José-Palacios-Vargas for lending me material from Mexico; Douglas Zeppelini for transporting material from Mexico to Illinois; Dan Swanson for his etymological expertise; and Monique DuBray, Juraj Cech, and Benjamin Hottel for their valuable assistance with field collections. This work was supported by the Herbert H. Ross Memorial Fund from the University of Illinois and the Society of Systematic Biologists, Graduate Student Research Award. I would also like to

thank the Department of Entomology at the University of Illinois, Urbana-Champaign and the Illinois Natural History Survey at the Prairie Research Institute.

Table

Table 4.1. Diagnostic characters to separate species of *Willowsia* that occur in the New World and additional species with close affinities with *W. pyrrhopygia* **sp. nov.** Bold character states indicate those shared with *W. pyrrhopygia* **sp. nov.** Underlined character states indicate those unique to native New World *Willowsia* and *Americabrya*.

Characters	***Willowsia pyrrhopygia* sp. nov.**	***Willowsia mexicana***	***Willowsia neocaledonia***	***Willowsia nigra***	***Willowsia jacobsoni***	***Willowsia nigromaculata***	***Willowsia buski***	***Americabrya arida***
Scale rib sculpturing	**uninterrupted**	**uninterrupted**	**uninterrupted**	**uninterrupted**	interrupted	interrupted	interrupted	**uninterrupted**
Scales on antennae	**present**	**present**	absent	**present**	absent	absent	absent	absent
Scales on legs	**present**	**present**	absent	absent	absent	absent	absent	absent
Scales on manubrium	**present**	**present**[1]	absent	absent	absent	absent	absent	absent
Labral papilla projections	**>1**	**>1**	**>1**	**>1**	**>1**	1	1	**>1**
Prelabral setae	**smooth**	**smooth**	**smooth**	**smooth**	ciliate	ciliate	ciliate	**smooth**
Labial triangle seta r	**absent**	**absent**	present	present	present	present	present	**absent**
Sublobal plate hairs	**2+2**	**2+2**	3+3	3+3	3+3	3+3	3+3	**2+2**
Head median (M) mac	**4+4**	**3+3(4+4)**	2+2	1+1	3+3	3+3	3+3	**4+4**
Head sutural (S) mac	**4+4(3+3)**	5+5	**3+3**	2+2	**4+4**	6+6	6+6	6+6
Head mac S0	**absent**	**absent**	**absent**	**absent**	**absent**	present	present	**absent**
Head mac Ps5	**present**	**present**	absent	absent	absent	absent	**present**	**present**
Th. II median mac	**0+0**	**0+0**	1+1	1+1	**0+0**	1+1	2+2	**0+0**
Th. II posterior mac	**2+2**	3+3	**2+2**	**2+2**	**2+2**	7+7	8+8	**2+2**
Abd. I mac	**1+1**	**1+1**	**1+1**	**1+1**	**1+1**	3+3	2+2	**1+1**
Abd. II inner mac	**2+2**	**2+2**	**2+2**	**2+2**	**2+2**	3+3	3+3	**2+2**
Abd. III inner mac	**2+2**	**2+2**	1+1	1+1	**2+2**	3+3	**2+2**	**2+2**
Abd. III lateral mac	**3+3**	**3+3**	2+2	2+2	**3+3**	**3+3**	**3+3**	**3+3**
Abd. IV inner mac	**4+4**	**4+4**	**4+4**	**4+4**	5+5	7+7	5+5	**4+4**

[1] Described by Zhang *et al*. (2007), but not confirmed in this study.

Figures

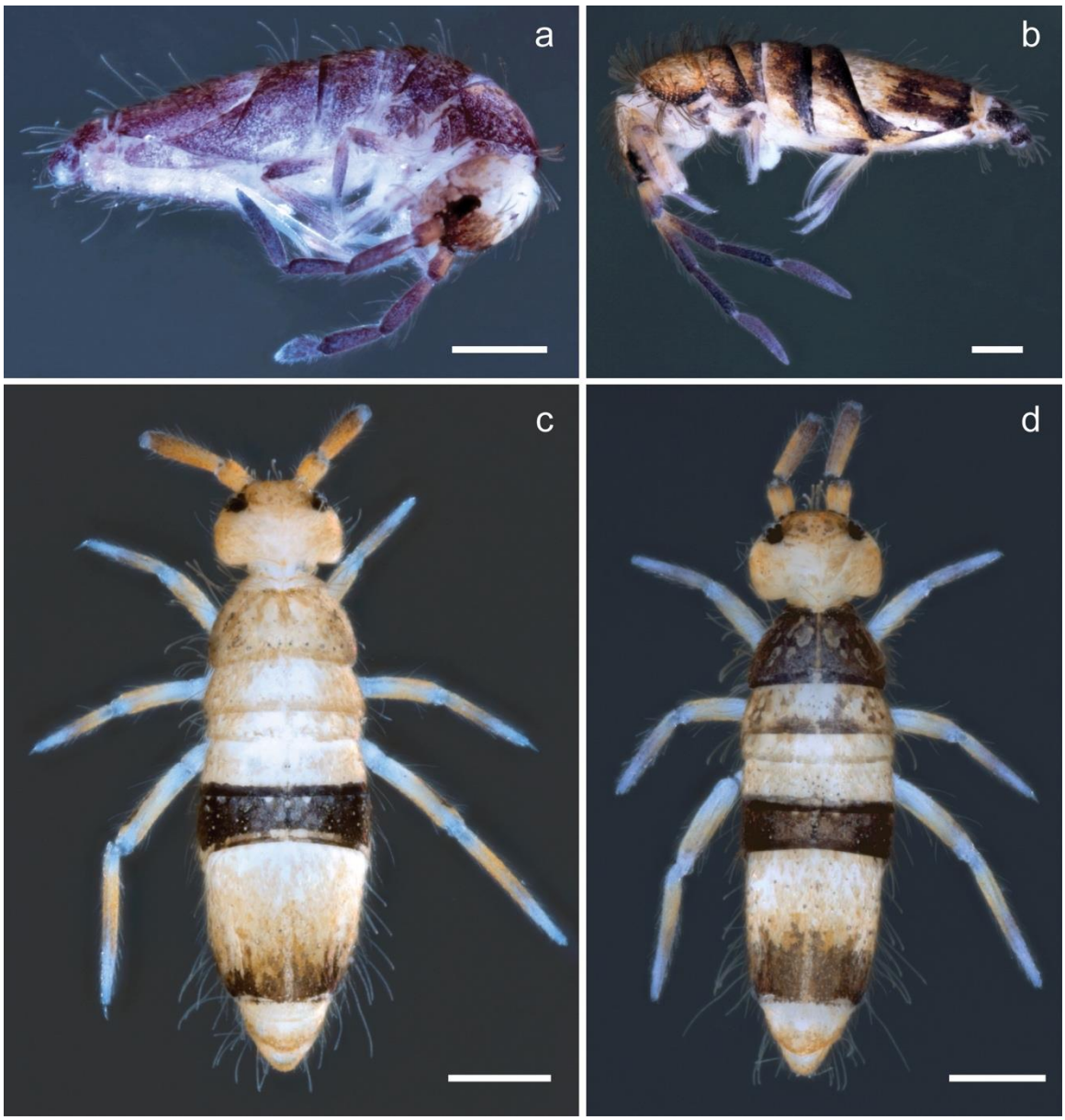


Figure 4.1. Habitus. (**a**) *Willowsia buski*; (**b**) *Willowsia nigromaculata*; (**c**) *Willowsia jacobsoni* (male); (**d**) *Willowsia jacobsoni* (female). Scale bars = 200 μm.

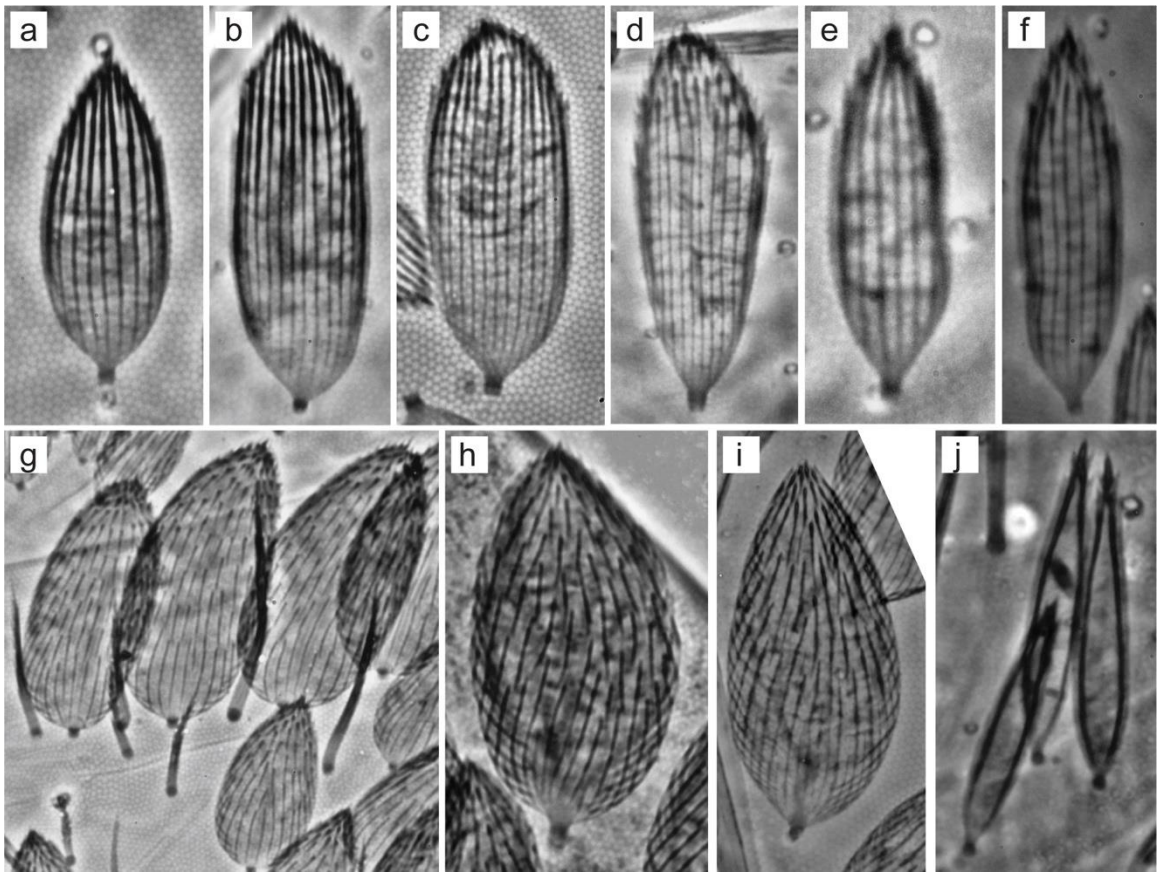


Figure 4.2. Scales. (**a–c**) *Willowsia pyrrhopygia* **sp. nov.**; (**d–f**) *Willowsia mexicana*; (**g**) *Willowsia jacobsoni*; (**h**) *Willowsia buski*; (**i**) *Willowsia nigromaculata*; (**j**) *Americabrya arida*. Scale bars = 1 μm.

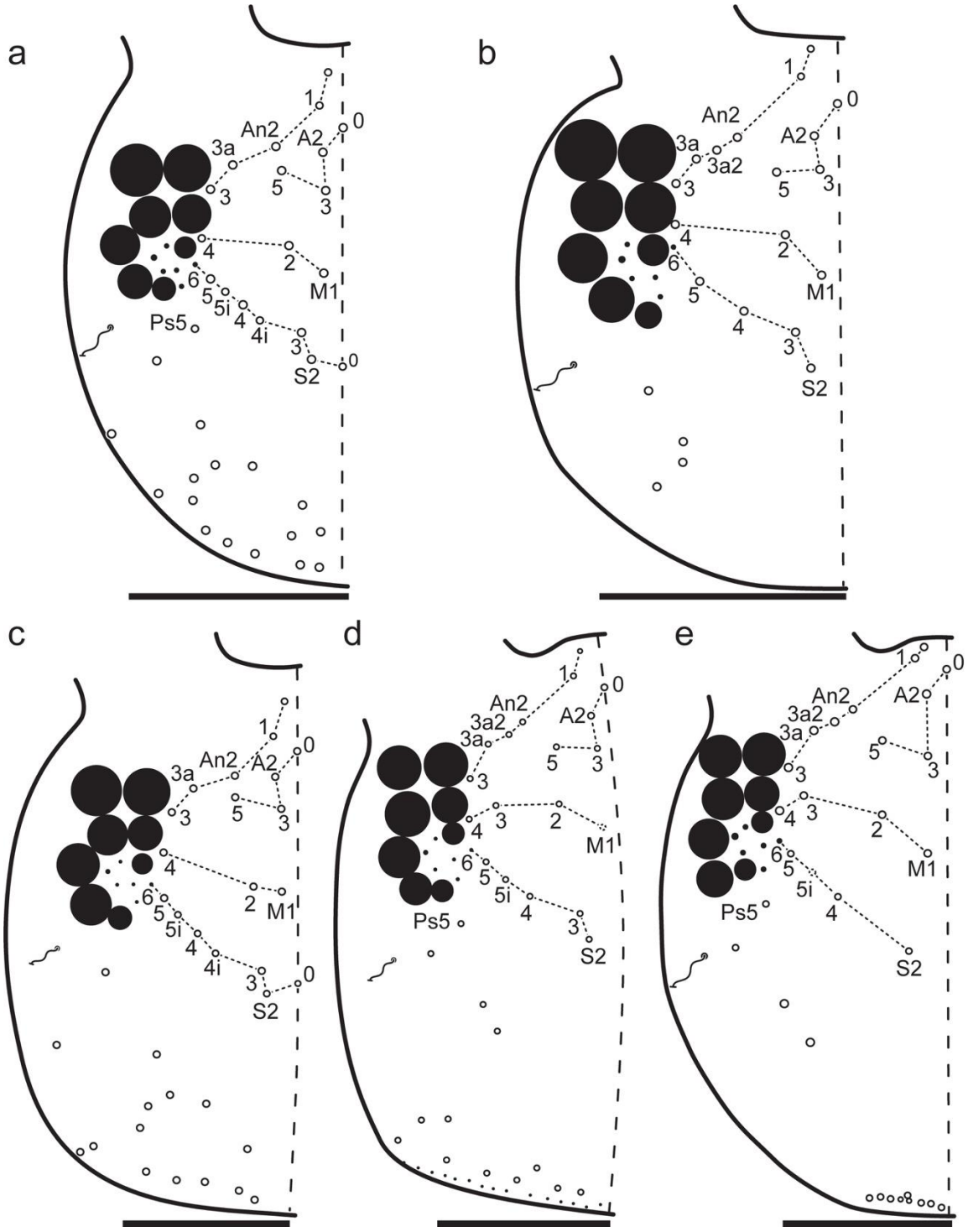


Figure 4.3. Head dorsal chaetotaxy. (**a**) *Willowsia buski*; (**b**) *Willowsia jacobsoni*; (**c**) *Willowsia nigromaculata*; (**d**) *Willowsia mexicana*; (**e**) *Willowsia pyrrhopygia* **sp. nov.** Scale bars = 100 μm.

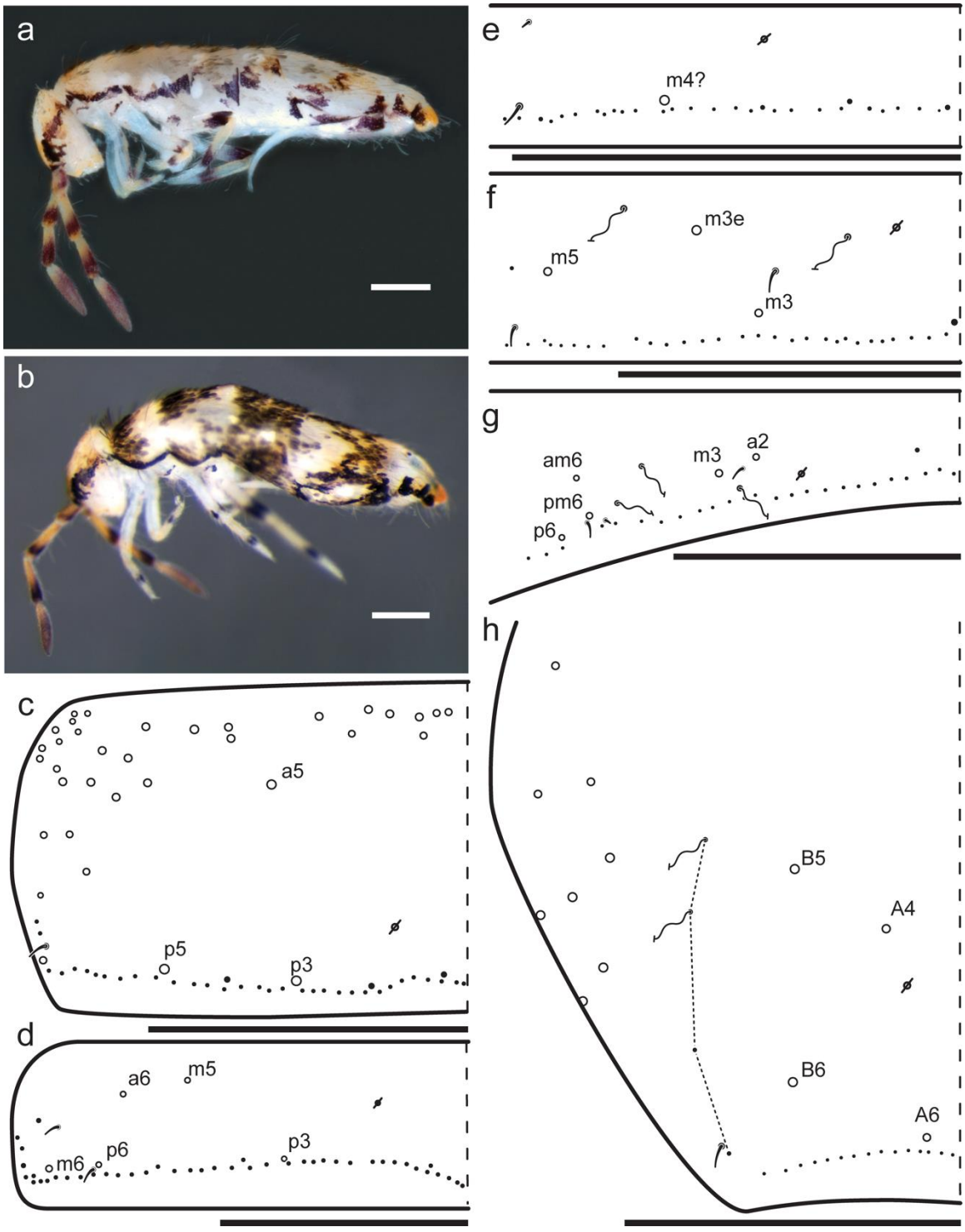


Figure 4.4. *Willowsia pyrrhopygia* **sp. nov. (a–b**) Habitus; (**c**) Th. II dorsal chaetotaxy; (**d**) Th. III dorsal chaetotaxy; (**e**) Abd. I dorsal chaetotaxy; (**f**) Abd. II dorsal chaetotaxy; (**g**) Abd. III dorsal chaetotaxy; (**h**) Abd. IV dorsal chaetotaxy. Scale bars = 100 μm unless indicated.

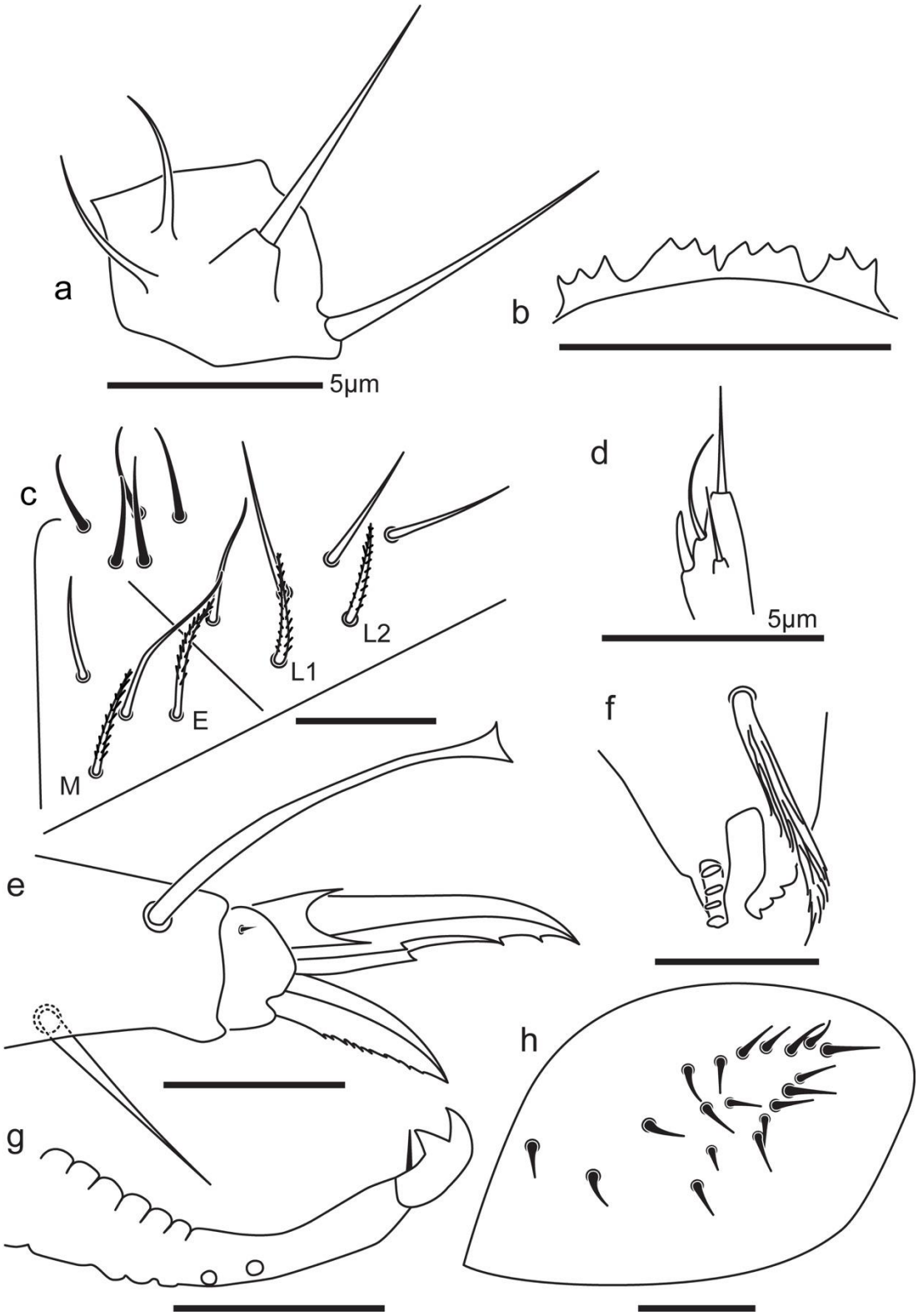


Figure 4.5. *Willowsia pyrrhopygia* **sp. nov. (a**) Sublobal plate of the outer maxillary lobe; (**b**) Labral papillae; (**c**) Labial triangle; (**d**) Lateral process of labial palp; (**e**) Hind claw complex; (**f**) Tenaculum; (**g**) Mucro; (**h**) Trochanteral organ. Scale bars = 25 µm unless indicated.

References

Börner, C. (1913). Zur Collembolenfauna Javas. Das trochanteralorgan der Entomobryiden. *Tijdschrift Voor Entomologie*, *56*, 44–61.

Chen, J.-X. & Christiansen, K. (1993). The genus *Sinella* with special reference to *Sinella s.s.* (Collembola: Entomobryidae) of China. *Oriental Insects*, *27*(1), 1–54. doi: 10.1080/00305316.1993.10432236

Christiansen, K. & Bellinger, P. (1980). *The Collembola of North America, North of the Rio Grande: A taxonomic analysis*. Grinnell College: Grinnell.

Christiansen, K. & Bellinger, P. (1998). *The Collembola of North America, North of the Rio Grande: A taxonomic analysis*. Grinnell College: Grinnell.

Jordana, R. & Baquero, E. (2005). A proposal of characters for taxonomic identification of *Entomobrya* species (Collembola, Entomobryomorpha), with description of a new species. *Abhandlungen und Berichte des Naturkundemuseums Görlitz*, *76*(2), 117–134.

Katz, A. D., Giordano, R. & Soto-Adames, F. (2015a). Taxonomic review and phylogenetic analysis of fifteen North American *Entomobrya* (Collembola, Entomobryidae), including four new species. *ZooKeys*, *525*, 1–75. doi: 10.3897/zookeys.525.6020

Katz, A. D., Giordano, R. & Soto-Adames, F. N. (2015b). Operational criteria for cryptic species delimitation when evidence is limited, as exemplified by North American *Entomobrya* (Collembola: Entomobryidae). *Zoological Journal of the Linnean Society*, *173*(4), 810–840. doi: 10.1111/zoj.12220

Lubbock, J. (1870). Notes on the Thysanura. Part IV. *Transactions of the Linnean Society of London (Zoology)*, *27*(2), 277–297. doi: 10.1111/j.1096-3642.1870.tb00214.x

Lubbock, J. (1873). *Monograph of the Collembola and Thysanura*. Ray Society: London.

Mari-Mutt, J. A. (1979). A revision of the genus *Dicranocentrus* Schött (Insecta: Collembola: Entomobryidae). *Bulletin of the University of Puerto Rico*, *259*, 1–79.

Mari Mutt, J. A. (1981). Redescription of *Willowsia jacobsoni* (Börner), an Entomobryid with conspicuous sexual dimorphism (lnsecta: Collembola). *Journal of Agriculture of the University of Puerto Rico*, *65*, 361–373.

Mari Mutt, J. A. & Palacios-Vargas, J. G. (1987). *Americabrya*, a new genus of Entomobryidae, with a description of *A. arida* based on Mexican specimens and descriptive notes for *A. epiphyta* (Loring). *Journal of the New York Entomological Society*, *95*(1), 99–108.

Pan, Z. & Zhang, F. (2016). Contribution to the *Willowsia* species having body scales of the long basal rib type: four new species and a redescription of *W. qui* (Collembola: Entomobryidae). *European Journal of Taxonomy*, *245*, 1–25. doi: 10.5852/ejt.2016.245

Schäffer, C. (1896). Die Collembola der Umgebung von Hamgurg und benachbarter Gebiete. *Mitteilungen aus dem Naturhistorischen Museum in Hamburg*, *13*, 149–216.

Shoebotham, J. W. (1917). XL.— Notes on Collembola .—Part 4. The classification of the Collembola; with a list of Genera known to occur in the British Isles. *Journal of Natural History Series 8*, *19*(114), 425–436. doi: 10.1080/00222931709486959

Soto-Adames, F. N. (2008). Postembryonic development of the dorsal chaetotaxy in *Seira dowlingi* (Collembola, Entomobryidae); with an analysis of the diagnostic and phylogenetic significance of primary chaetotaxy in *Seira*. *Zootaxa*, *1683*, 1–31.

Szeptycki, A. (1979). *Chaetotaxy of the Entomobryidae and its phylogenetical significance. Morpho-systematic studies on Collembola (Vol. 4)*. Państwowe Wydawnictwo Naukowe: Warsaw.

Zhang, F. (2015). Some *Willowsia* from Nepal and Vietnam (Collembola: Entomobryidae) and description of one new species. *Zootaxa*, *3905*(4), 489–499. doi: 10.11646/zootaxa.3905.4.3

Zhang, F., Bedos, A. & Deharveng, L. (2014a). Two new species of *Willowsia* from New Caledonia (Collembola: Entomobryidae). *Zootaxa*, *3872*(4), 381–386. doi: 10.11646/zootaxa.3872.4.5

Zhang, F., Chen, J.-X. & Deharveng, L. (2011). New insight into the systematics of the *Willowsia* complex (Collembola: Entomobryidae). *Annales de la Société Entomologique de France*, *47*(1–2), 1–20. doi: 10.1080/00379271.2011.10697692

Zhang, F., Chen, Z., Dong, R. R., Deharveng, L., Stevens, M. I., Huang, Y. H. & Zhu, C. D. (2014b). Molecular phylogeny reveals independent origins of body scales in Entomobryidae (Hexapoda: Collembola). *Molecular Phylogenetics and Evolution*, *70*, 231–239. doi: 10.1016/j.ympev.2013.09.024

Zhang, F. & Deharveng, L. (2015). Systematic revision of Entomobryidae (Collembola) by integrating molecular and new morphological evidence. *Zoologica Scripta*, *44*(3), 298–311. doi: 10.1111/zsc.12100

Zhang, F., Palacios-Vargas, J. G. & Chen, J.-X. (2007). The genus *Willowsia* and its Mexican species (Collembola: Entomobryidae). *Annals of the Entomological Society of America*, *100*(1), 36–40. doi: 10.1603/0013-8746(2007)100[36:TGWAIM]2.0.CO;2

Zhang, F., Pan, Z., Wu, J., Ding, Y., Yu, D. & Wang, B. (2016). Dental scales could occur in all scaled subfamilies of Entomobryidae (Collembola): new definition of Entomobryinae with description of a new genus and three new species. *Invertebrate Systematics*, *30*, 598–615. doi: 10.1071/IS16005

Zhang, F., Sun, D.-D., Yu, D.-Y. & Wang, B.-X. (2015). Molecular phylogeny supports S-chaetae as a key character better than jumping organs and body scales in classification of Entomobryoidea (Collembola). *Scientific Reports*, *5*, 12471. doi: 10.1038/srep12471