A Molecular and Morphological Investigation of the Springtail Genus *Orchesella* (Collembola: Entomobryomorpha: Entomobryidae)

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Catherine Louise Smith
May 2015

DEDICATION

I would like to thank Dr. Moulton, who as my adviser gave me guidance, and shared his wealth of knowledge with me.

I would also like to thank my other graduate committee members, Dr. Bernard and Dr. Jurat-Fuentes both of whom gave encouragement to me during an unexpectedly large project.

I would like to also thank Dr. Felipe Soto-Ademas who shared his expertise and knowledge with me.

Fourthly, to Gary Phillips, Rob Pivar, Jeremy Blanschke, and Joshua Granger who helped me in the lab and field. Best of luck in your futures, gentlemen!

Finally thanks go to my friends and family who gave me support and encouragement during this endeavor. Thank you for letting me roam your backyards, sending me leaf litter samples, or joining me on my collecting trips.

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge all of the help I have received from the Department of Entomology and Plant Pathology.

ABSTRACT

Chapter 1 is devoted to a molecular overview of the North American members of the springtail genus *Orchesella* using mitochondrial and nuclear DNA. Both genes strongly support the presence of four major clades within *Orchesella.* Chapter 2 is a morphological revision of the species near *Orchesella celsa* Christiansen & Bellinger. Five new species are described based largely on chaetotaxonomical differences in. Chapter 3 is an investigation into the origins of United States populations of two introduced European members of the genus, *Orchesella cincta* Linnaeus and *Orchesella villosa* Linnaeus. Mitochondrial data between populations from the two continents indicates multiple introductions for *O. cincta*, and the possibility of cryptic species in the European populations of *O. villosa*.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Genus Orchesella

Classified in the family Entomobryidae, subfamily Orchesellinae, and tribe Orchesellini, members of *Orchesella* are recognized by the following traits: abdominal bothriotricha present on abdominal segments II, III, and IV with 2, 3, 2 bothriotricha respectively on each side, more than 15 setae on the trochanteral organ, absence of the Post Antennal Organ, presence of a clavate tenent hair on each tibiotarsus, fourth abdominal segment less than three times longer than the third abdominal segment, lack of scales, bidentate mucro with basal spine, six apparent antennal segments sue to the subdivision of the first and second segments, one tenacular setae, lack of complex microsetae around the bothriotricha, and setae on the venter of the dens (Soto-Adames et. al 2008).

Orchesella is a Holarctic genus of springtails: the native North America species are not found west of the Great Plains and the Colorado Plateau, nor above the Arctic Circle. In Eurasia, members of this genus are typically not found east of the Ural Mountains, nor above the Arctic Circle (Collembola of the World, www.collembola.org). They are typically found in damp moss and lichens, or under the bark of decaying deciduous trees in relatively undisturbed areas. There are currently 18 described species found in North America: two introduced invasive species and 16 native species (Christiansen and Bellinger 1998; Table 1). Members of this genus have adapted to a wide range of altitudes, from sea level to the top of Clingmans Dome (2025m), and in landscapes where trees are not present they have been found feeding on fungi growing on dead, damp grass.

In Europe, populations of *Orchesella cincta* Linnaeus have been studied due to their adaptation to cadmium-contaminated soils (Posthuma et. al 1992). Other work with *Orchesella* spp. in Europe includes determinion of population groups within *O.cincta* (Timmermanns et. al 2005), and the molecular and morphological determination of new species (Frati et. al 2000).

Table 1. *Orchesella* of North America.

***Orchesella* species in North America**	**Type Locality**
Orchesella ainsliei Folsom	USA: Illinois, Champaign Co.
Orchesella albosa Guthrie	USA: Minnesota, Hennepin Co.
Orchesella alpa Christiansen & Tucker	USA: New Mexico, Taos Co.
Orchesella annulicornis Mills	USA: Iowa, Wapello Co.
Orchesella bulba Christiansen & Tucker	USA: Florida, Dade Co.
Orchesella carneiceps Packard	USA: Tennessee, Knox Co.
Orchesella celsa Christiansen & Tucker	USA: North Carolina, New Hanover Co.
Orchesella cincta Linnaeus	France
Orchesella fishmani Christiansen &Tucker	USA: Texas, Calhoun Co.
Orchesella flora Christiansen & Tucker	USA: Florida, Liberty Co.
Orchesella folsomi Maynard	USA: New York,
Orchesella gloriosa Snider	USA: North Carolina, Swain Co.
Orchesella hexfasciata Harvey	USA: Maine, Penobscot Co.
Orchesella imitari Snider	Canada: Quebec, Parc de la Gaspesie
Orchesella manitobae Mari Mutt	Canada: Manitoba, Fort Whyte
Orchesella texensis Snider	USA: Texas
Orchesella villosa Linnaeus	France: Paris
Orchesella zebra Guthrie	USA: Minnesota, Grey Cloud Island

Morphological and Molecular Identifications

Identification of springtails is hampered by a lack of experts in the field (Rougerie et. al, 2009), and the need for dexterity and access to restricted chemicals to clear specimens and make the slide mounts. Additionally, time and access to adequate microscopes are also factors in recognizing species of *Orchesella*, relying as it does on the meticulous mapping of macrosetae, pseudopores and bothriotrichia. Species descriptions and macrosetal diagrams do not exist; the description of *Orchesella albosa* Guthrie, for instance, consists of only an ink painting of the type material exists (Guthrie, 1903).

Until recently, it was thought that many springtail species had large ranges, but recent molecular studies have shown that some species with large ranges may actually consist of several ranges of similarly marked, or cryptic

species, each with a smaller range (Frati et. al, 2000; Park, 2009; Porco et. al, 2010; Porco et. al, 2012; Porco et. al, 2012; Cicconardi et. al, 2013; Shaw et. al, 2013) Members of *Orchesella* have also been depicted as having large ranges, such as *Orchesella flora* Christiansen & Tucker. The type locality for this species is Liberty County, Florida, with other specimens collected from Indiana, North Carolina, and Virginia (Christiansen and Bellinger, 1998). Based on the molecular work carried out as part of this thesis, the range of *O.flora* consists of multiple species having varying but often overlapping markings.

The objectives of this research were to molecularly characterize every Nearctic *Orchesella* species, correlate molecular results with morphology, and for select undescribed species provide detailed morphological descriptions.

Lastly, phylogentic relationships among the Nearctic species were inferred from nuclear and mitochondrial markers. All DNA sequences generated in this study will be uploaded to the National Center for Biotechnology Information (NCBI).

Phylogenetic Analysis

A mitochondrial gene and a nuclear gene were analyzed phylogenetically. Phylogenetic trees were constructed using the molecular data to elucidate relationships between the native species and the European invasive species, as well as the relationships among the native species.

The mitochondrial gene cytochrome oxidase subunit I (COI, or COX I) has often been used for genetic species delineation, (Rougerie et al., 2009), aiding in the discovery of new species, often cryptic species. It has been used repeatedly within springtails (Cicconardi et al., 2013; Porco et al., 2010; Porco et al., 2012; Shaw et al., 2013). In this study, however, cytochrome oxidase subunit II, (COXII) was used instead, to make use of primers and sequences from Timmermanns et al., 2005.

Nuclear genes can often more accurately reflect species differences, but nuclear data is difficult to obtain from springtails due to a large number of introns (Song et al., 2010). In addition, mitochondrial genes mutate more rapidly than nuclear genes (Moritz et al., 1987), potentially leading to an over-estimation of species richness. By using a nuclear and a mitochondrial gene, a more robust test of species status can be obtained.

For this study a fragment of the nuclear gene 3’-5’ exoribonuclease II was used. This gene was used successfully in a previous systematic study of the genus *Pogonathellus* (Family Tomoceridae), (Felderhoff, 2007).

Morphological Descriptions

Morphological descriptions of Collembola are a combination of structural characteristics, color, pigment patterns and chaetotaxy (arrangement of setae). Within many entomobryid genera, including *Orchesella*, the structure of characters such as mouthparts, antennae, foot complex, and furcula do not vary significantly among species. In these genera color, pattern and chaetotaxy are the major differentiating characters. The chaetotaxy of Entomobryidae was thoroughly studied by Szeptycki (1979) and elaborated by Christiansen and Bellinger (1980) and Jordana and Baquero (2005). Different authors have used differing nomenclatures for different setal elements and to denote the position of the element on the organism, leading to multiple chaetotaxic descriptions for some species. For the purposes of this paper, the chaetotaxic nomenclature and scheme used is that of Potapov and Kremenitsa (2008), in their description of a new species of *Orchesella* from the Caucasus.

References

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

Checklist of the Collembola, www.collembola.org.

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, **22**:5382-5396.

Felderhoff, Kelly Lynn, "A Revision of Tomoceridae (Insecta: Collembola) in Great Smoky Mountains National Park and Southern Appalachians Using Morphological and Molecular Approaches". Master's Thesis, University of Tennessee, 2007.

Frati F, Dell'Ampio E. Casasanta S, Carapelli A, Fanciulli PP (2000) Large amounts of genetic divergence among Italian species of the Genus *Orchesella* (Insecta, Collembola) and the relationship of the new species. *Molecular Phylogenetics and Evolution,* **17**(3):456-461.

Guthrie, JE (1903) The collembola of Minnesota. Geological and Natural History Survey of Minnesota.

Hopkin, SP (1997) Biology of the springtails (Insecta: Collembola). Oxford University Press.

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 Beichte des Naturkundemuseums Görlitz,* **76**(2):117-134.

Moritz C, Dowling TE, Brown WM (1987) Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. *Annual Review of Ecology and Systematics*, **18**:269-292.

Park, K-H (2009) Molecular biological study on speciation and phylogeny of the order Entomobryomorpha (Collembola: Hexapoda). *Entomological Research*, **39**(2009):334-340.

Porco D, Bedos A, Deharveng L (2010) Description and DNA barconding assessment of the new species *Deutonura gibbosa* (Collembola: Neanuridae: Neanurinae), a common springtail of the Alps and Jura. *Zootaxa*, **2639**:59-68.

Porco D, Bedos A, Greensdale P, Janion C, Skarżyński D, Stevens MI, Jansen van Vuuren B, Deharveng L (2012) Challenging species delimitation in Colembola: cryptic diversity among common springtails unveiled by DNA barcoding. *Invertebrate Systematics*, **26**:470-477.

Porco D, Potapov M, Bedos A, Busmachiu G, Weiner WM, Hamra-Kroua S, Deharveng L (2012) Cryptic Diversity in the Ubiquist Species *Parisotoma notabilis* (Colembola, Isotomidae): a Long-Used Chimeric Species? *PLOS One*, **7**(9):1-7.

Posthuma L, Hogervorst RF, Van Straalen NM (1992) Adaptation to soil pollution by cadmium excretion in natural populations of *Orchesella cincta* (L.) (Collembola). *Environmental Contamination and Toxicology*, **22**(1):146-156.

Potapov M, Kremenitsa A (2008) Comments on the chaetotaxy of the genus *Orchesella* (Collembola, Entomobryomorpha) with a redefinition of the '*spectabilis*' group and description of a new species of *Orchsella* from the Caucasus. *Soil Organisms*, **80**(1):99-115.

Rougerie R, Decaëns T, Deharveng L, Porco D, James SW. Chang C, Richard B, Potapov M, Suhardjono Y, Hebert PDN (2009) DNA barcodes for soil animal taxonomy. *Pesquisa Agropecuária Brasileira*, **44**(8):789-801.

Salmon JT (1964) An Index to the Collembola. *Royal Society of New Zealand*, **7**(2).

Shaw P, Faria C, Emerson B (2013) Updating taxonomic biogeography in the light of new methods-examples from Collembola. *Soil Organisms*, **85**(3): 161-170.

Song JM, Nam K, Sun YU, Kang MH, Kim CG, Kwon ST, Lee J, Lee YH (2010) Molecular and biochemical characterizations of a novel arthropod endo β-1,3-glucanase from the Antarctic springtail, *Cryptopygus antarcticus*, horizonatally acquired from bacteria. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **155**(4): 403-412.

Soto-Adames FN, Barra J, Christiansen K, Jordana R (2008) Suprageneric classification of Collembola Entomobryomorpha. *Annals of the Entomological Society of America*, **101**(3): 501-513.

Szeptycki A (1979) Chaetotaxy of the Entomobryidae and its phylogenetical significance. Morphosystematic studies of Collembola. IV. *Polska Akademia Nauk, Zaklad Zoologii Systematycznej Doswiadczalnej*.

Timmermanns MJTN, Ellers J, Mariën J, Verhoef SC, Ferwerda EB, van Straalen NM (2005) Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular Ecology*, **14**:2017-2024.

CHAPTER I

A MOLECULAR OVERVIEW OF THE GENUS *ORCHESELLA* (CLASS COLEMBOLA, FAMILY ENTOMOBRYIDAE)

Abstract

Native North American specimens from the springtail genus *Orchesella* were collected from throughout their ranges along with two introduced European members of the genus and analyzed phylogenetically using a mitochondrial (cytochrome oxidase subunit II) and a nuclear (3'-5' exoribonuclease II) gene. Phylogenetic trees were constructed for each gene separately and from a combined data set. Four clades were consistently recovered: the European invasives clade, the *hexfasciata/flora* clade, the Texas clade, and the *celsa* clade. The *hexfasciata/flora* clade shows little support for recognized morphospecies. The Texas clade supports recognized morphospecies, but demonstrates the existence of multiple undescribed species, and illustrates the need for more extensive collecting west of the Mississippi. The *celsa* clade contains *O.gloriosa* plus several undescribed cryptic species near *O.celsa sensu stricto*.

Introduction

The springtail genus *Orchesella* (Entomobryomorpha: Entomobryidae), has a long history of being studied in Europe, with the descriptions of *Orchesella cincta* and *Orchesella villosa* by Linnaeus in the 1760s (Stach, 1960). New species continue to be described in this group (Frati et al., 2000; Potapov & Kremenitsa, 2008). *Orchesella cincta* has been used extensively in Europe as a model organism to study the effects of heavy metal soil contamination on soil arthropods (Posthuma et al., 1993; Roelofs et al., 2007). In North America, this genus is not well studied. Despite multiple species descriptions published by Christiansen & Tucker (1977), and Snider (1989), it is expected that many more species are awaiting description.

Molecular studies of springtails in North America have focused on a geographic region (Soto-Adames, 2002; Porco et al., 2013), or upon a single species (van der Wurff, 2005; Cicconardi et al., 2013; Porco et al., 2012). These studies have shown an unexpectedly high level of diversity in springtails, particularly of cryptic species. Studies of *Pogonognathellus* in eastern North America revealed several undescribed species (Felderhoff et. al, 2010).

As of 2014, *Orchesella* within the Nearctic region was comprised of 16 described native species plus two presumed introduced species. Many morphospecies posses large ranges covering numerous states or provinces (Christiansen & Bellinger, 1998). Some of these species have been considered to have extremely broad variation in colors and pigment patterns. Several collections of Orchesella specimens that fit no known pigment pattern led me to

the hypothesis that a number of unrecognized species existed in North America and possibly accounted for the apparent large ranges of individual variable species. The objectives of this study were to investigate species boundaries within Nearctic *Orchesella* using primarily a molecular phylogenetic approach, but also chaetotaxy in the instance of the *O.celsa* clade, to elucidate species status. Two markers from unrelated genomes, a 555 basepair fragment of the mitochondrial gene cytochrome oxidase subunit II (COXII) (Timmermans et al., 2005) and a 950 basepair fragment from the nuclear gene; 3'-5' exoribonuclease II (Felderhoff et al., 2010) were selected to increase the robustness of the results. COXII was used for this genetic survey due the availability of published primers and data (Timmermans et al., 2005).

Modern phylogenetic methods were used to construct trees illustrating the relationships between the introduced European specimens and the native specimens, as well as the relationships within the native species.

Methods and Materials

Taxon sampling

Specimens for this study came from two sources: collecting done during 2013-2014, or from the University of Tennessee (ECB Research Collection) if they were collected within the last decade and in suitable ethanol.

The majority of collected specimens were collected from moist lichens, rotted logs, and damp rotting grass mats by sieving, aspiration, and then stored in 95% non-denatured ethanol. The rest of the specimens were extracted from leaf litter samples by means of Tullgren funnels and preserved in 95% ethanol. Where possible, type localities were visited for collection of topotypes.

Molecular techniques

Total genomic DNA for all specimens was extracted using the ThermoScientific GeneJET Genomic DNA Extraction (ThermoScientific, Waltham, MA) kit following the manufacturer's suggested protocol, except for the final step where the gDNA was eluted in 70µL of elution buffer warmed to 52°C. Purified DNA samples were stored at -20°C. The initial forward primer for the mitochondrial aspect of this project was one used by Timmermans et al. (2005): TL2-J-3037 (5'-AATATGGCAGATTAGTGC-3') (Simon et. al, 1994), and later two custom forward primers designed by J.K. Moulton: Orch_CoxII_NEW F (5'-CCWTCAGAMCACTCWTACTC-3'), and Orch_CoxII_F3 (5'-CARAAYAAYCCYCCYTCRGA-3'). One reverse primer was taken from Timmermans et al. (2005): (COII-R: 5'-CCACAGATTTCTGAGCATTGACC-3'), and two custom reverse primers which were designed by J.K. Moulton: Orch_CoxII_NewR (5'-TCRTTNGCNCCRCARATYTC-3'), and Orch_CoxII_R2 (5'-ATYGGYATRAARCTRTGRTT-3'). These primers amplify a 563 base pair

fragment from the mitochondrial genome comprising the extreme 3' end of tRNA-leucine (i.e., the first 15 bases of polished final sequences) and the ca. 5' half of cytochrome oxidase II (COXII). This gene region has been used previously and successfully for phylogenetic purposes in Collembola (Frati and Carapelli, 1999; Frati et al. 2001; McGaughran et al. 2010; Luque et al. 2011).

The primers for the 1200 base pair fragment of 5'-3' exoribonuclease II were designed by J.K. Moulton. Four forward primers were used: 5-3Exo2_203F (5'-CCNGGNGARGGNGARCAYAA-3'), New_Orch_FWD (5'-GGNGARGGNGARCAYAARAT-3'), 5-3E2_Orch_P1F1 (5'-GARCAYAARATYATGGAYTA-3'), and 5-3E2_CollP1.5F (5'-GAYGAYTGGGTNTTYATGTG-3'). Two reverse primers, also designed by J.K. Moulton, were used: 5-3E2_Coll_P1R (5'-ACRTCRAAYTTYGAYTCRTARTA-3'), and 5-3E2_OG_P1R (5'-GCRAAYTGNCCTTCYGGRATRAA-3'). The nuclear sequence typically contained two introns of highly variable sequence separated by less variable exon, but some specimens possessed up to five introns.

Amplifications were performed in GenePro (Bioer Technology Co., Hangzhou, China) thermal cyclers, using TaKaRa Ex Taq Hotstart DNA polymerase (Takara Bio, Shiga, Japan) per the manufacturer's suggested protocol with 1.3µL template DNA, and 3µL of (7mM) of forward and reverse primer. Thermal cycling parameters were as follows: 1.5 min denaturation soak at 94°C; 5 cycles of 94°C for 30s, 52°C for 30 s, and 72°C for 45s; 15 cycles of 94°C for 30s, 47°C for 25s, and 72°C for 45s; 35 cycles of 94°C for 30s, 42°C for 20s, 72°C for 45s, 72°C soak for 3 min, and 13°C hold. PCR products were electrophoresed in 1% agarose, excised from the gel, and purified using a QiaQuick Gel Extraction Kit (Qiagen, Valencia, CA). The reverse strand of each product was cycled sequenced in 20µL reactions using 16-fold diluted Big Dye 3.1 (Applied Biosystems, Foster City, CA). Sequencing reactions were cleaned using Centrisep columns (Prinecton Separations, Adelphia, NJ) and dried in a Centrivap Concentrator (LABCONCO, Kansas City, MO). Sequencing of samples was performed by the University of Tennessee-Knoxville Molecular Biology Resource Facility. Sequences were verified for accuracy using Sequencher 4.7 (GeneCodes, Ann Arbor, MI).

DNA alignment and phylogenetic analysis

Mitochondrial sequences and nuclear sequences were aligned and analyzed separately, then concatenated and analyzed together.

Sequences from all North American *Orchsella* specimens collected were assembled into a matrix along with those from specimens of the European springtails of *O. cincta* and *O.vilosa* and selected outgroups (Table 10). Alignment of the sequences was straightforward, requiring no indels. Mesquite (Maddison and Maddison, 2011) was used to partition the data set into codon positions. PAUP* (Swofford, 2001) was used to calculate pairwise sequence

divergence. HKY85 corrected distances were calculated because this model was selected as the best fit by jModelTest.

Bayesian and maximum likelihood analyses were performed on the nucleotides. The optimal evolutionary model was determined using jModelTest 2.1.4 (Guindon and Gascuel, 2003; Darriba et al., 2012). For the mitochondrial and nuclear total evidence analysis and the nuclear only analysis, the TVM+I+G was selected as the best fit model based on the Bayesian information criterion. Mitochondrial and nuclear total evidence analysis, [-lnL=25360.7105;K=213; AIC=51147.4210; f(A)=0.3280; f(C)=0.2020; f(G)=0.1206; f(T)=0.3494; I=0.3570; G=0.5050]. Nuclear only analysis, [-lnL=25360.7105;K=213; AIC=51147.4210; f(A)=0.3280; f(C)=0.2020; f(G)=0.1206; f(T)=0.3494; I=0.3570; G=0.5050]. For the mitochondrial analysis, the GTR+I+G was selected as the best fit model based on the Bayesian information criterion [-lnL=12893.5140;K=184; AIC=26155.0281; f(A)=0.3249; f(C)=0.2120; f(G)=0.0681; f(T)=0.3951; I=0.3900; G=0.4950].

Best-fit models were implemented in Bayesian analyses using MrBayes 3.2.2 (Ronquist and Huelsenbeck, 2003). Every Markov chain in the Bayesian search was started from a random tree and set to compute 1×10^7 generations, sampling every 1000^{th} one from the chain, resulting in a total of 1000 trees. Three hot chains and 1 cold chain were run simultaneously, with pre-stationarity trees discarded as burn-in. Each simulation was run twice. Default settings for the priors were used, and the base frequencies were estimated from the data. Tracer 1.6 (Rambaut et al., 2013) was used to parse and combine the log files, determine at which point the Markov Chain Monte Carlo (MCMC) began to sample from the stationarity distribution, and to check that effective sample sizes (ESSs) were sufficient for all parameters. To reduce the probability of convergence on local optima, multiple starting points for each chain were used. Maximum likelihood analysis was performed using RAxML-HPC2 (Stamatakais, 2006; Stamatakis et al., 2008), as implemented in CIPRES-XSEDE (Miller et al., 2010). Analyses were conducted using the evolutionary model GTRGAMMAI (GTR+I+G) as well as the default model (GTRCAT), each with 1,000 bootstrap replicates, with the data partitioned by codon position and gene. There were no discernible differences between approaches. Bayesian posterior probabilities and nonparametric bootstrap proportions were used to assess node support (Felsenstein, 1985).

Results

Orchesella specimens used in this analysis grouped naturally into four major clades: The Europen invasives in purple, the *hexfasciata/flora* in green, the Texas clade in red, and the *celsa* clade in blue. Each of these clades has modest to strong node support: the European invasives clade (BS=63, PP=0.96), the *hexfasciata/flora* clade (BS=85, PP=1.0), the Texas clade (BS=95, PP=1.0), and the *celsa* clade (BS=100, PP=1.0) (see Figure 1).

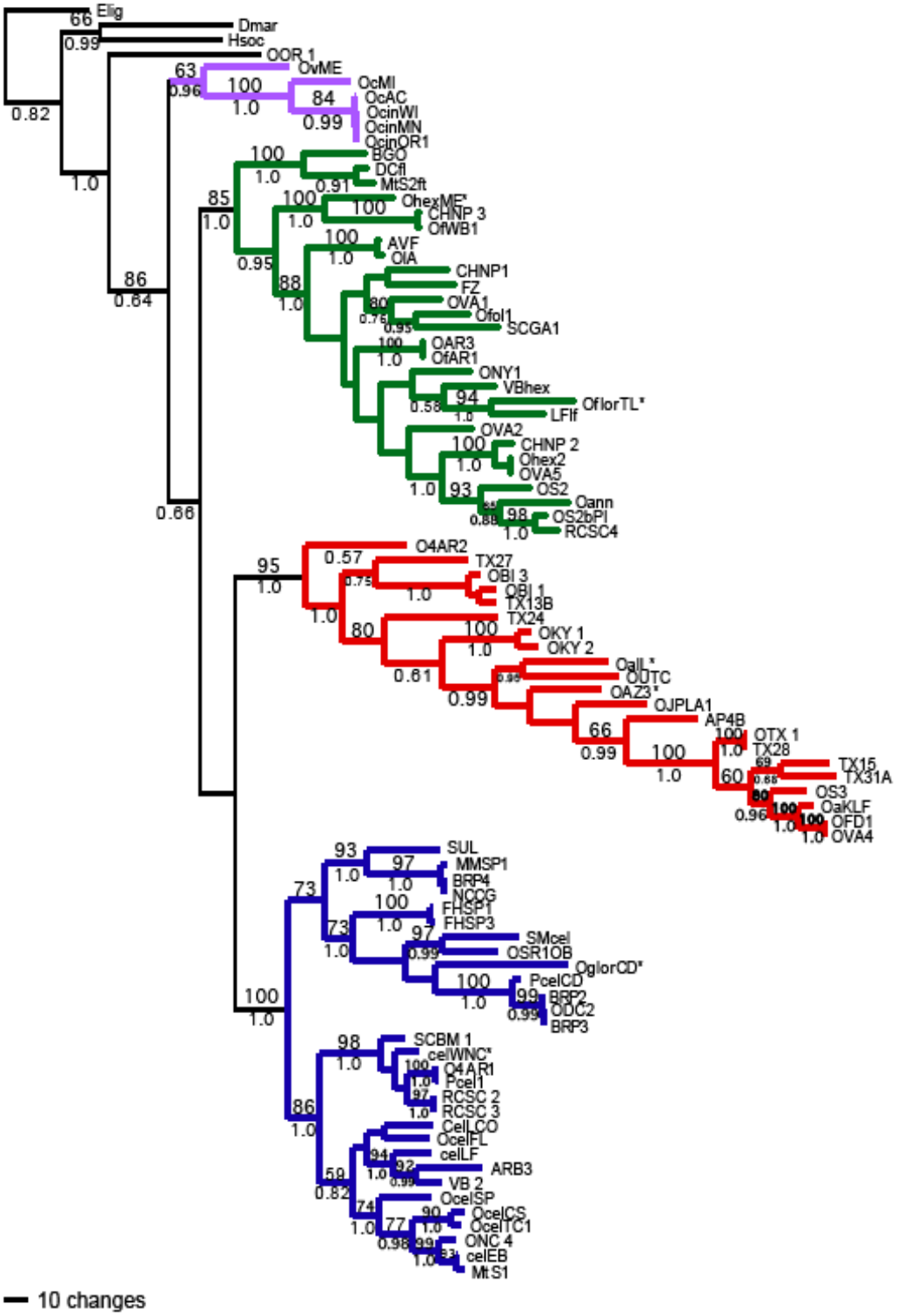

Figure 1. COXII tree with bootstrap and posterior probability values.
*indicates specimens collected from the type locality

All three trees showed strong support for monophyly of each European species with the two specimens of *Orchesella villosa* having bootstrap scores (BS) of 100 and posterior probabilities (PP) of 1.0 and the eight specimens of *Orchesella cincta* having bootstrap values of 100 and PP of 1.0 from the combined nuclear and mitochondrial tree (Figure 2).

All inferred trees (MtDNA, nuclear, and total evidence) exhibited strong support for monophyly of each European species. This is evidenced by the pairing of the two specimens of *Orchesella villosa* (Total Evidence: BS=52, PP=1.0, Mitochondrial Only: BS=63, PP= 0.96) and the eight specimens of *O. cincta* sampled in this study (Figures 2 and 3).

Within the *hexfasciata/flora* clade, the presumed morphospecies based on dorsal color pattern were not recovered as monophyletic in any of the three inferred phylogenetic trees. *Orchesella annulicornis* (specimen Oann), which was sampled from a single locality, was also recovered within this group.

The Texas clade shows strong support for a group of three species, *Orchesella bulba* (specimens OBI, OBI3, and TX13) and two currently undescribed species (specimens TX27 and O4AR2). Together, these three species form the sister group to all the remaining specimens in this clade. The remainder of this clade is composed of a series of very strongly supported single OTUs that represent species or even species groups. Further taxon sampling in this area of the Nearctic region is needed to clarify species limits in this area of the tree.

The *celsa* clade consists of two major groups with moderate to strong support for each. The first group contains *Orchesella gloriosa* (specimen OglorCD) and four currently undescribed species each having strong node support. The second group contains *Orchesella celsa* (specimen celWNC) and one currently undescribed species, both of which possess strong node support (Figures 2 and 3).

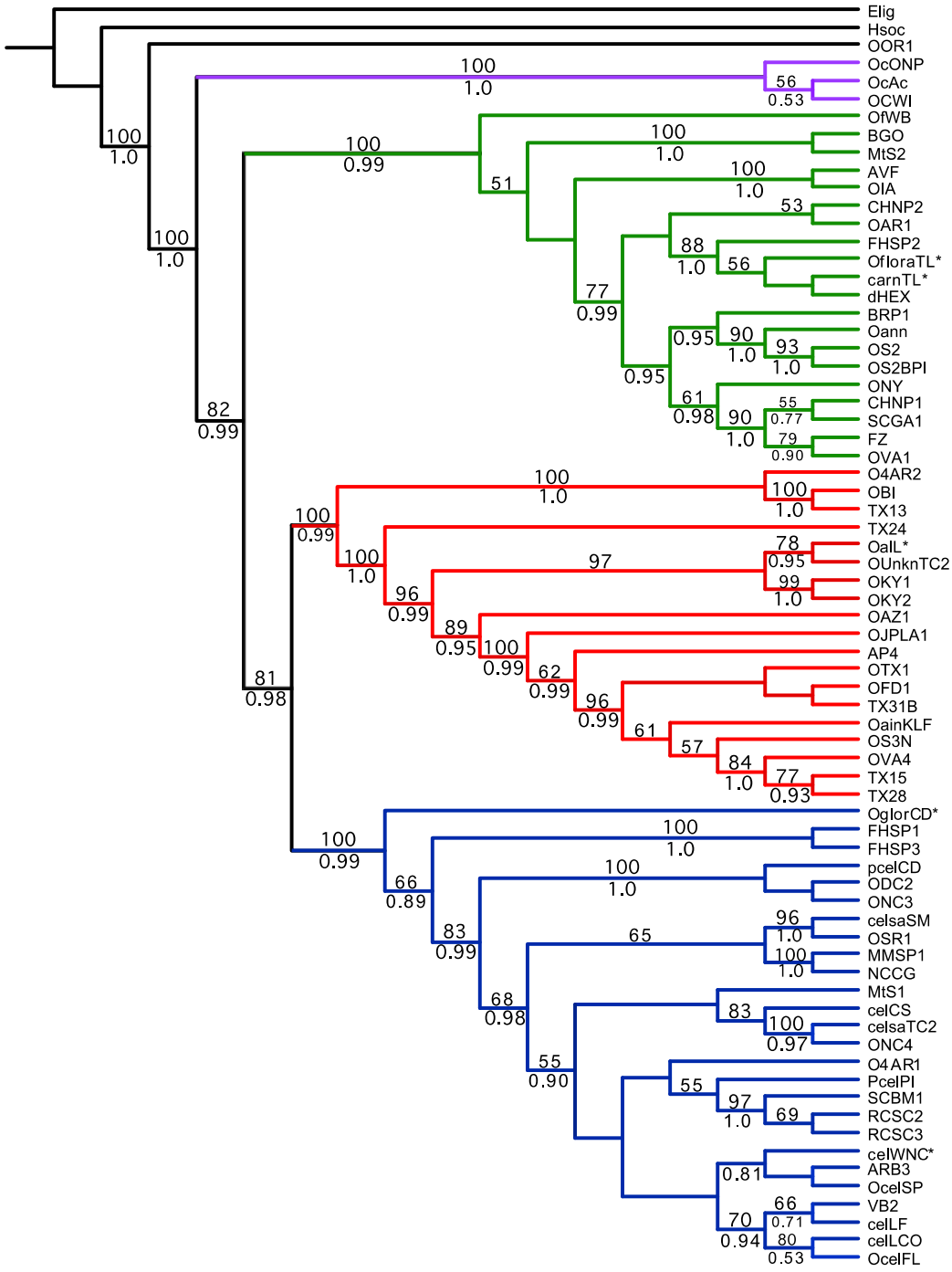

Figure 2. Tree from combined mitochondrial and nuclear sequences with bootstrap and posterior probability scores.

*indicates described species from its type locality

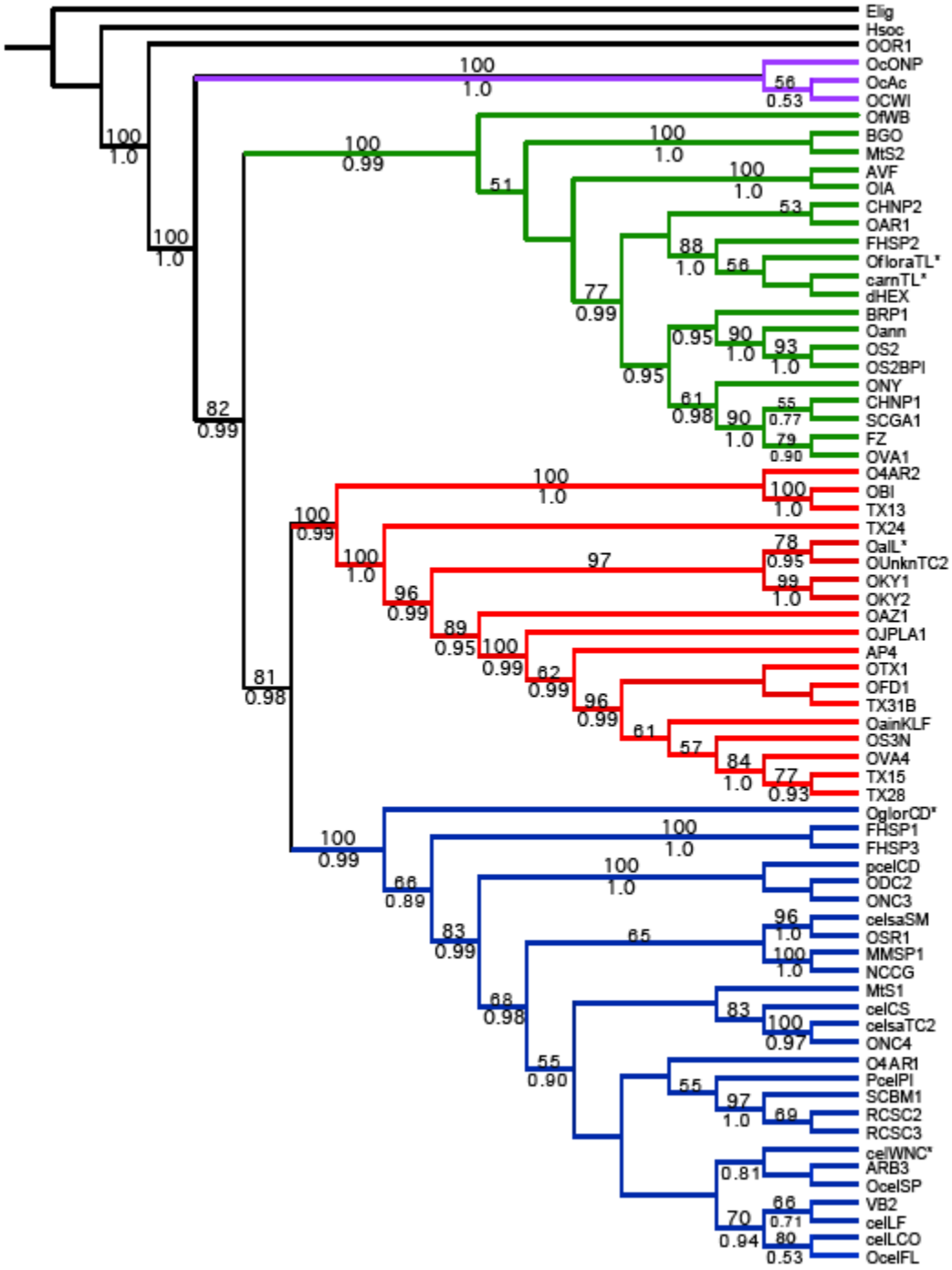

Figure 3. 5'-3' Exoribonuclease II tree with bootstrap and posterior probability scores.

*indicates described species from its type locality

Discussion

As shown by all three phylogenetic trees, the European invasives are more closely related to each other than they are to any of the native North American specimens. Given that the final separation of the eastern half of the North American continent and European continent was about 50 million years ago during the Eocene (McKenna, 1972), these two groups show remarkably little morphological differentiation.

The three trees indicate a need to bring North American species concepts in this genus into a more modern species concept; one not based entirely upon dorsal color patterns and geographic location, but upon fine-scale, detailed chaetotaxonomic examination and molecular analysis. By bringing a more balanced approach to species determination in this genus, it will be possible to describe monophyletic, scientifically robust species.

Within the *hexfasciata/flora* clade (Figure 4), dorsal markings appear meaningless in terms of species identifications and relationships (Figure 4). For instance, *Orchesella hexfasciata* (specimen OhexME), has six transverse stripes along the length of its body and traditionally all *Orchesella* specimens with six stripes have been considered to be *O. hexfasciata*. The *O. hexfasciata* from the type locality in Maine, USA, consistently pairs with specimen OfWB from Illinois, USA(a specimen with white head, white mesothorax, black metathorax, and black abdomen with white fifth and sixth abdominal segments). Other *Orchesella* specimens with six transverse stripes are Ohex2 and CHNP2, which pair together but are more closely related to the *Orchesella flora* (specimen OfloraTL) group than to true *O. hexfasciata*. Lack of monophyly among similarly patterned specimens occurs repeatedly throughout this clade for all described species, except *O. annulicornis* which could not be evaluated for the presence of cryptic species as it was represented in this study by only a single individual.

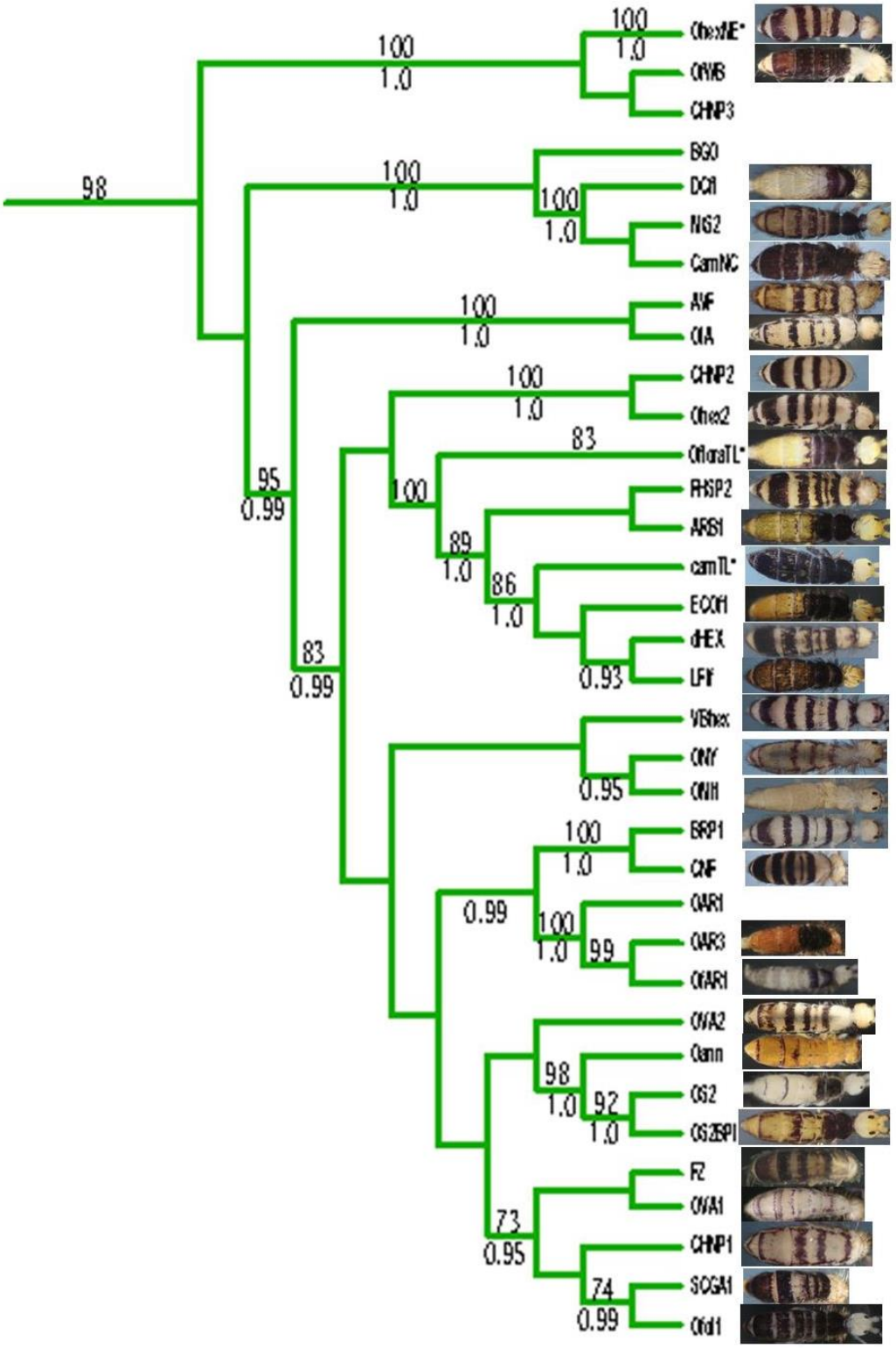

Figure 4. Total evidence tree for the *hexfasciata/flora* clade with photos of dorsal patterns.

Table 2. Collection data for specimens in the *hexfasciata/flora clade*.

Specimen Designation	Collection Data
AVF	WI: Columbia Co., Avalon Farms
BGO	Blue Ridge Parkway, Balsam Gap
BRP1	NC: Jackson Co., Blue Ridge Parkway, Woodfin Cascade
carnTL (*Orchesella carneiceps*)	TN: Knox Co., Cherokee Woodlot
CHNP1	OH: Hocking Co., Crane Hollow Nature Preserve
CHNP2	OH: Hocking Co., Crane Hollow Nature Preserve
CHNP3	OH: Hocking Co., Crane Hollow Nature Preserve
DCfl	NC: Transylvania Co., Devil's Courthouse
dHEX	TN: Cooke Co., Cherokee National Forest
FHSP2	TN: Morgan Co., Frozen Head State Park
FZ	TN: VanBuren Co.,
LFfl	GA: Walker Co., Lula Lake Land Trust
MtS2	NC: Haywood Co., Mt. Sterling
Oann* (*Orchesella annulicornis*)	IL: Will Co., Braidwood Dunes and Savanna Nature Preserve
OAR1	AR: Franklin Co., Ozark National Forest
OAR3	AR: Franklin Co., Ozark National Forest
OfAR1	AR: Franklin Co., Ozark National Forest
Ofol1	NC: Swain Co., Enloe Cemetery
OfloraTL* (*Orchesella flora*)	FL: Liberty Co., Torreya State Park
OfWB	IL: Union Co., Rich's Cave
Ohex2	VA: Fairfax Co., George Washington Memorial Parkway

Table 2. Collection data for specimens in the *hexfasciata/flora* clade (cont.).

Specimen Designation	Collection Data
OhexME* (*Orchesella hexfasciata*)	ME: Hancock Co.
OIA	IA: Dubuque Co.
ONY	NY: Schuyler County, Finger Lakes National Forest
OS2	TN: Sevier Co., GSMP, Twin Creeks Pavilion
OS2bPI	SC: Georgetown Co., Pawley's Island
OVA1	VA: Fairfax Co., George Washington Memorial Parkway
OVA2	VA: Fairfax Co., George Washington Memorial Parkway
OVA5	VA: Fairfax Co., George Washington Memorial Parkway
RCSC4	SC: Chesterfield Co., Near Rocky Creek
SCGA1	GA: Hart Co., near Shoal Creek
VBhex	TN: VanBuren Co.

The Texas clade (Figure 5), named this for the preponderance of Texas OTUs, shows strong support for *Orchesella alpa* (specimen OAZ1), but only weak to moderate support for *Orchesella ainsliei* (specimen OaIL) (See Figure 5). Further testing of different populations of *O. alpa*, and other southwestern species might change the strength of that relationship. There is strong support for the monophyly of *Orchesella bulba* (specimens OBI, OBI3, TX13), a species that truly seems to have a large range in southeastern United States from the Carolinas to Texas along the Atlantic and the Gulf Coast. There is also strong support for two currently undescribed species closely related to *O.bulba* (specimens TX27, and O4AR2). An examination of the O4AR2 specimen shows it has the same post-antennal structure found in the *O.bulba* specimens, which is likely a synapomorphy for this potentially formally recognized species group.

The strongly supported laddering of long branches among OTUs in the Texas clade indicates a critical need for more extensive collection in this region and likely the presence of numerous undescribed species.

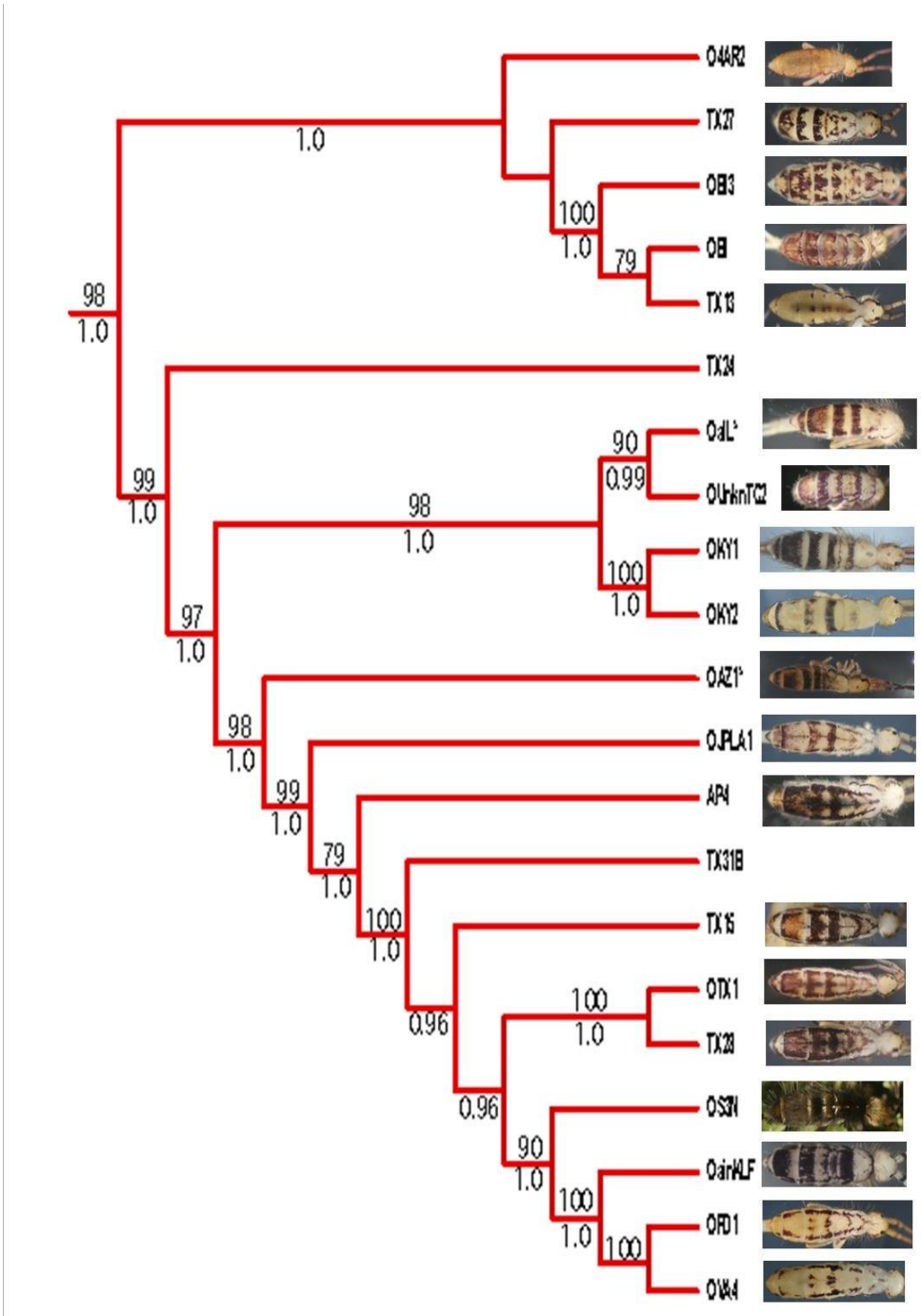

Figure 5. Total evidence tree for the Texas clade with dorsal pattern images.

* indicates described species from Chrsitian and Bellinger 1997

Table 3. Collection data for the specimens in the Texas clade.

Specimen Designation	Collection Data
AP4B	TN: Fayette Co., Ames Plantation
O4AR2	SC: Lancaster Co., 40-Acre Rock WMA
OaIL	IL: Will Co., Braidwood Dunes and Savanna Nature Preserve
OaKLF	AR: Franklin Co., Ozark National Forest
OAZ3	AZ: Apache Co., Government Spring
OBI1	SC: Georgetown Co., Belle Isle Plantation
OBI3	SC: Georgetown Co., Belle Isle Plantation
OFD1	TN: Knox Co., Fort Dickerson Park
OJPLA1	LA: Jackson Parish, Schoolhouse Spring
OKY1	KY: Calloway Co., Terrapin Creek SNP
OKY2	KY: Calloway Co., Terrapin Creek SNP
OS3	TX: Tyler Co., Big Thicket National Preserve, Sundew Trail
OTX1	TX: Travis Co.
OUTC	TN: Sevier Co., GSMP, Twin Creeks Pavilion
OVA4	VA: Fairfax Co., George Washington Memorial Parkway
TX13B	TX: USA: TX: Hardin Co., NW of Kountze
TX15	TX: Tyler Co., Big Thicket National Preserve, Sundew Trail
TX24	TX: Williamson Co., Cedar Elm Preserve
TX27	TX: Hays Co., Road FM-150
TX28	TX: Jackson Co., US59 & Sandy Creek
TX31A	TX: Jackson Co., US59 & Sandy Creek

Members of the Texas clade tend to have two stripes posterior to their eyes and four longitudinal stripes running the length of their thorax and abdomen. The exception to this pattern is specimen O4AR-2, which completely lacks stripes.

The *celsa* clade (Figure 6) is made up of *O. gloriosa* (specimen OglorCD), *Orchesella celsa* (specimen celWNC) both of which were collected from their respective type localities. Other specimens that were identified as "*celsa*-types" based on their dorsal patterns. There is strong support for two groups in the *celsa* clade; one group is consists of specimens from the higher elevations of the southern Appalachian Mountains, and one group consists of specimens from the lower elevations on either side of the Smoky Mountains and the southeastern coastal plain. *Orchesella gloriosa* and four putatively undescribed species are in the high elevation group. *Orchesella celsa* is in the lower elevation group with one putatively undescribed species (see chapter 3).

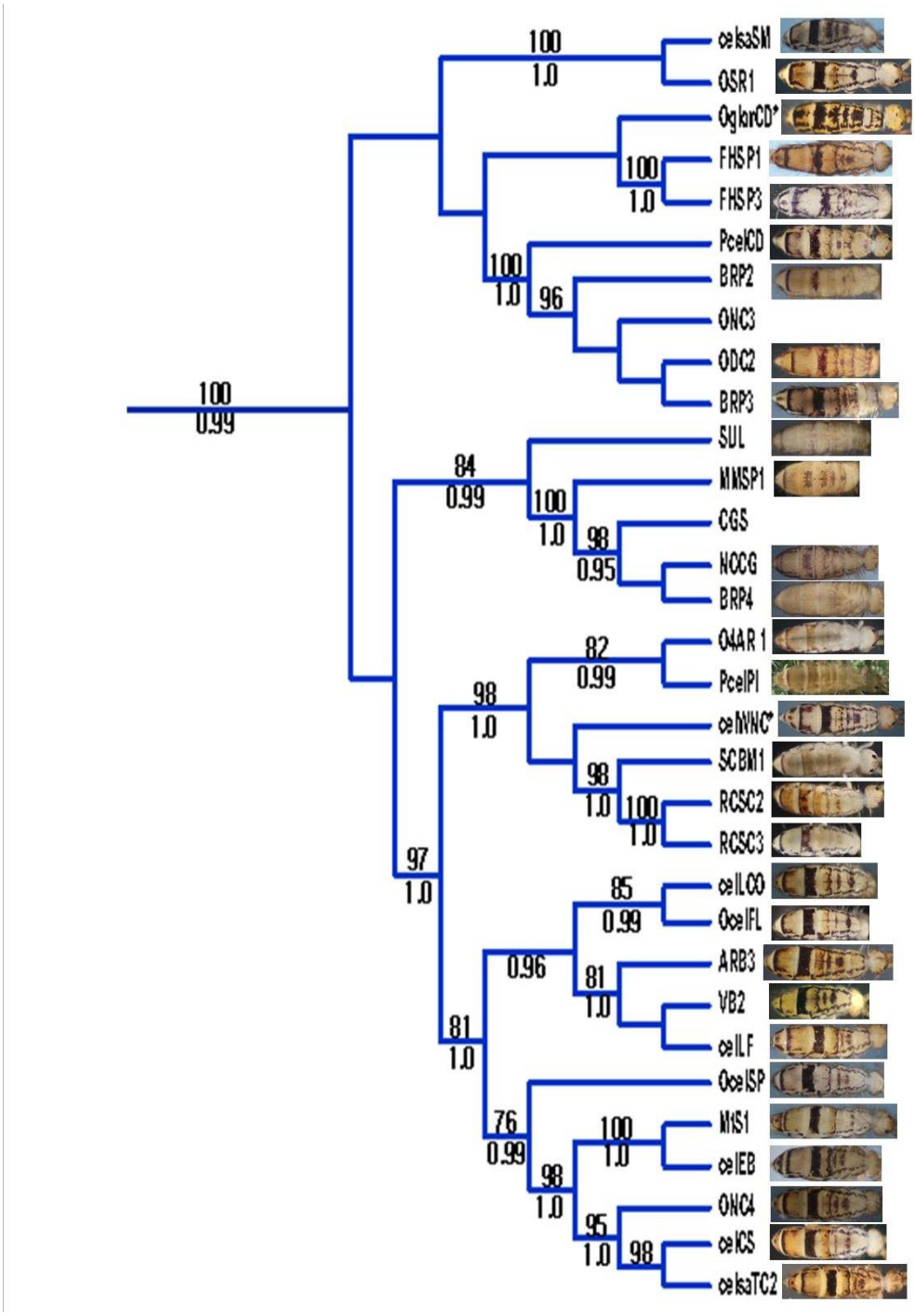

Figure 6. Total evidence tree for the *celsa* clade with dorsal pattern images.
* indicates described species from Christiansen and Bellinger, 1997.

References

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the undersestimation of global species richness. *Molecular Ecology*, **22:**5382-5396.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest2: more models, new heuristics and parallel computing. *Nature Methods*, **9**:772.

Drummond AJ, Rambaut A, (2003) BEAST v1.0.

Felderhoff KL, Bernard EB, Moulton JK (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.

Felsenstein J (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution*, **39**(4):783-791.

Frati F, Dell'Ampio E, Casasanta S, Carapelli A, Fanciulli PP (2000) Large Amounts of Genetic Divergence among Italian Species of the Genus *Orchesella* (Insecta, Collembola) and the relationships of Two New Species. *Molecular Phylogenetics an Evolution*, **17**(3):456-461.

Frati F, Simon C, Sullivan J, Swofford DL (1997) Evolution of the Mitochondrial Cytochrome Oxidase II Gene in Collembola. *Journal of Molecular Evolution,* **44**:145-158.

Guindon S, Gascual O (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, **52**(5):696-704.

Le Roux J, Wieczorek AM (2009) Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, **154**:1-17.

Luque C, Legal L, Winterton P, Mariano NA, Gers C (2011) Illustration of the Structure of Arthropod Assembleges *(Collembola* and *Lepidoptera*) in Different Forest Types: an Example in the French Pyrenees. *Diversity*, **3**: 693-711.

McGaughran A, Torricelli G, Carapelli A, Frati F, Stevens MI, Convey P Hogg ID (2010) Contrasting phylogeographical pattern for springtails reflect different evolutionary histories between the Antarctic Peninsula and continental Antarctica. *Journal of Biogeography*, **37**:107-119.

McKenna MC (1972) Eocene final separation of the Eurasian and Greenland-North American landmasses. *24th International Geological Congress, Section 7*: 275-281.

Porco D, Skarżynski D, Decaëns T, Herbert PD, Deharveng L (2013) Barcoding the Collembola of Churchill: a molecular taxonomic reassessment of species diversity in a sub-Arctic area. *Molecular Ecology Resources*, **14**(2):249-261.

Posthuma L, Hogervorst RF, Joosse ENG, Van Straalen NM (1993) Genetic Variation and Covariation for Characteristics Associated with Cadmium Tolerance in Natural Populations of the Springtail Orchesella cincta (L.). *Evolution,* **47**(2):619-631.

Potapov M, Kremenitsa A (2008) Comments on the chaetotaxy of the genus *Orchesella* (Collembola, Entomobryomorpha) with a redefinition of the '*spectabilis*' group and description of a new species of *Orchesella* from the Caucasus. *Soil Organisms*, **80**(1):99-115.

Roelofs D, Mariën J, van Straalen NM (2007) Differential gene expression profiles associated with heavy metal tolerance in the soil insect Orchesella cincta. *Insect Biochemistry and Molecular Biology,* **37**(4):287–295.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**(22):1572-1574.

Simon C, Frati F, Beckenbach AT, Crespi B, Liu H, Flook P (1994) Evolution weighting and phylogenetic utility of mitochondrial genes sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**(6):651-701.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed model. *Bioinformatics*, **22**(21): 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid Bootstrap Algorithm for the RaxML Web Servers. *Systematic Biology*, **57(**5)**:**758-771.

Timmermans MJTN, Ellers J, Mariën J, Verhoef SC, Ferwerda EB, van Straalen NM (2005) Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular Ecology*, **14**:2017-2024.

van der Wurff AWG, Gols R, Ernsting G, van Straalen NM (2005) Population genetic structure of *Orchesella cincta* (Collembola; Hexapoda) in NW Europe, as revealed by microsatellite markers. *Pedobiologica*, **49**(2005), 167-174.

CHAPTER II

PATTERN VARIATION IN *OCHESELLA CELSA* CHRISTIANSEN & TUCKER 1977 *SENSU LATO*: A RANGE OF PATTERNS, A RANGE OF SPECIES

Abstract

Specimens of *Orchesella celsa* Christiansen & Tucker 1977 *sensu lato* were collected from throughout the southeastern United States for molecular and morphological examination. Molecular examination consisted of creating phylogenetic trees using a 950 basepair fragment of the nuclear gene 5'-3' exoribonuclease II, and a 555 basepair fragment of the mitochondrial gene cytochrome oxidase subunit II (COXII). Mitochondrial and nuclear data sets were analyzed separately and combined. Strong support was observed for an upper elevation clade of species containing, among others, typotypical *O.gloriosa,*and a lower elevation species clade that contains, among one other species typotypical *O.celsa*. Within these clades strong to moderate support was observed for putative cryptic species. These entities possess unique chaetotaxonomic patterns. Based on this evidence five new species of *Orchesella* near *O.celsa* and *O.gloriosa* are proposed.

Introduction

Species of Collembola have typically been afforded large geographic ranges (Collembola of the World), which is seemingly incongruent with the biology and dispersal capability of these flightless soil arthropods. A few species (e.g. *Podura aquatica* Linnaeus, 1758) indisputably have natural circumglobal ranges, and there is some evidence suggesting that some tree climbing springtails can be carried by the wind (van der Wurff, 2005); however, relatively few species climb trees (Hopkin, 1997). Human-assisted migration by transport of horticultural, agricultural, and silvicultural goods has often been invoked to explain large geographical ranges. The presence of the Asian species *Homidia sauteri, Homidia socia, Entomobrya unostrigata* and European species, *Lepidocyrtus paradoxus, Orchesella cincta* and *Orchesella villosa* in North America provides strong evidence that this assisted migration has occurred repeatedly (Christiansen & Bellinger 1998, Maynard 1951).

A visual examination of specimens of *Orchesella celsa* Christiansen & Bellinger reveals a number of different dorsal pigment patterns and colorations (Figures 7 and 8). The extent of the variability in these dorsal patterns and pigmentation has led many springtail taxonomists to regard them as population or regional differences. Yosii theorized that the pattern differences within *Lepidocyrtus* reflected a species richness that was not reflected in the chaetotaxy (Yoshii, 1989). Later, using molecular techniques, Soto-Adames (2002) corroborated this hypothesis.

Recently, reduced costs associated with the examination of genetic material have made it easier to sequence numerous specimens rapidly, allowing researchers to add this tool to their analytical repertoire. Cicconardi et al (2013) showed that the Panamanian members of the genus *Lepidocyrtus* had a high degree of genetic divergence that was not readily apparent in their morphology. The ubiquitous springtail species *Parisotoma notabilis* was recently shown by Porco et al. (2012) to be comprised of four distinct genetic lineages sufficiently different to be considered different species. Based on taxa in the vicinity of Churchill, Manitoba, Porco et al. (2013) showed by means of genetic barcoding that springtail species diversity has been seriously underestimated.

The focus of the current research was on a single but highly variable species, *Orchesella celsa* Christiansen & Tucker 1977, -described from Wilmington, North Carolina. Found mainly on the eastern seaboard of the United States, this springtail is commonly found in damp rotten stumps of *Quercus* spp., rotted wood, hardwood leaf litter, or in damp moss and lichen. This species tends to be more numerous at higher elevations, with a higher abundance on the sides of mountains and tops of ridges than in valleys. In the initial description the variability of the dorsal pigment pattern on collected specimens was mentioned, but was attributed to regional population variation. According to Christiansen & Bellinger (1998), *O.celsa* has a large geographical range. In the southern Appalachians *O.celsa* is commonly collected in Malaise traps, indicating that it is a good climber (E. C. Bernard, pers. comm.)

Very early in the DNA sequencing and analysis phase of this project, it became clear that *O.celsa,* rather than being a widespread highly variable species was in fact a complex of several species, some being morphologically separable from each other and others not so much. Besides molecular characterization via phylogenetic analysis, the other objective of this research was to morphologically describe these cryptic species.

Methods and Materials

Taxon sampling

Specimens determined as *Orchesella celsa* Christiansen and Tucker (1977) were collected from various locations in the United States. Specimens were collected from damp lichenous tree bark, or from stumps with red rot using by means of sieving and aspiration, or from leaf litter samples via Tullgren Funnel. The collected specimens were then placed in 95% ethanol and examined under a dissecting scope for initial identification. Some specimens were taken from the University of Tennessee (E. C. Bernard Research Collection) if they had been preserved only in unadulterated ethanol. Two paratype specimens (both juveniles) were borrowed from the Illinois Natural History Survey to examine their color patterns, but were not used for molecular analysis.

Photographic images were taken of their dorsal patterns and lateral markings. Once the specimens were identified as possibly being *Orchesella celsa*, one adult was then selected from each locality for genetic extraction. There were some specimens that were identified as belonging to the *O.celsa* group only after genetic analysis, due to their reduced, non-standard pigmentation.

Outgroups added to root the phylogeny included *Dicranocentrus marias* Wray and *Heteromurus major* (Moniez, 1889) Absolon, 1901, and several native North American *Orchesella* species, viz., *Orchesella gloriosa* Snider, *Orchesella flora* Chrsitiansen & Tucker, and *Orchesella annulicornis* Mills.

Molecular methods

Total genomic DNA was extracted using the ThermoScientific GeneJET Genomic DNA Extraction (ThermoScientific, Waltham, MA) kit following the manufacturer's suggested protocol, except for the final step where the gDNA was eluted in 70µL of elution buffer warmed to 52°C. Purified DNA samples were stored at -20°C. For the mitochondrial DNA, we used the forward primer used by Timmermans et al. (2005): TL2-J-3037 (5'-AATATGGCAGATTAGTGC-3') (Simon et. al, 1994) and a custom reverse primer designed by JKM (COII-R: 5'-CCACAGATTTCTGAGCATTGACC-3'), which amplify a 563 base pair fragment from the mitochondrial genome comprising the extreme 3' end of tRNA-leucine (i.e., the first 15 bases of polished final sequences) and the ca. 5' half of cytochrome oxidase II (COXII). This gene region has been used previously and successfully for phylogenetic purposes in Collembola (Frati and Carapelli, 1999; Frati et al. 2001; McGaughran et al. 2010; Luque et al. 2011).

For the nuclear DNA, we used the custom forward primers designed by J.K. Moulton: 5-3Exo2_203F (5'-CCNGGNGARGGNGARCAYAA-3'), New_Orch_Fwd (5'-GGNGARGGNGARCAYAARAT-3'), 5-3E2_Orch_P1F1 (5'-GARCAYAARATYATGGAYAA-3'), and 5-3E2_CollP1.5F (5'-GARCAYAARATYATGGAYTA-3'), Two custom reverse primers were also designed by J.K. Moulton: 5-3E2_Coll_P1R (5'-ACRTRAAYTTYGAYTCRTARTA-3'), and 5-3E2_OG_P1R (5'-GCRAAYTGNCCTTCYGGRATRAA-3'). These primers amplify an approximately 1200 base pair fragment of the gene 5'-3' exoribonuclease II. In the *O.celsa* specimens, this fragment contained two introns of highly variable genetic sequence separated by a coding exon.

Amplifications were performed in GenePro (Bioer Technology Co., Hangzhou, China) thermal cyclers, using TaKaRa Ex Taq Hotstart DNA polymerase (Takara Bio, Shiga, Japan) per the manufacturer's suggested protocol with 1.3µL template DNA, and 3µL of (7mM) of forward and reverse primer. Thermal cycling parameters were as follows: 1.5 min denaturation soak at 94°C; 5 cycles of 94°C for 30s, 52°C for 30 s, and 72°C for 45s; 15 cycles of 94°C for 30s, 47°C for 25s, and 72°C for 45s; 35 cycles of 94°C for 30s, 42°C for 20s, 72°C for 45s, 72°C soak for 3 min, and 13°C hold. PCR products were

electrophoresed in 1% agarose, excised from the gel, and purified using a QiaQuick Gel Extraction Kit (Qiagen, Valencia, CA). The reverse strand of each product was cycled sequenced in 20µL reactions using 16-fold diluted Big Dye 3.1 (Applied Biosystems, Foster City, CA). Sequencing reactions were cleaned using Centrisep columns (Prinecton Separations, Adelphia, NJ) and dried in a Centrivap Concentrator (LABCONCO, Kansas City, MO). Sequencing of samples was performed by the University of Tennessee-Knoxville Molecular Biology Resource Facility. Sequences were verified for accuracy using Sequencher 4.7 (GeneCodes, Ann Arbor, MI).

DNA alignment and phylogenetic analysis

Nuclear sequences had introns removed before a consensus sequence was made from a contig of forward and reverse strands. The nuclear and mitochondrial sequences from all collected *O. celsa* specimens and selected outgroups were assembled into a data matrix. Alignment of the sequences was straightforward, requiring no indels. Mesquite (Maddison and Maddison, 2011) was used to partition the data set into codon positions. PAUP* (Swofford, 2001) was used to calculate pairwise sequence divergence. HKY85 corrected distances were calculated because this model was selected as the best fit by jModelTest.

Bayesian and maximum likelihood analyses were performed on the nucleotides. The optimal evolutionary model was determined using jModelTest 2.1.4 (Guindon and Gascuel, 2003; Darriba et al., 2012), which selected the GTR+I+G as the best fit model for the mitochondrial data based on the Bayesian information criterion [-lnL=12893.5140;K=184; AIC=26155.0282; f(A)=0.3249; f(C)=0.2120; f(G)=0.0681; f(T)=0.3951; I=0.3900; G=0.4950].

For the nuclear and combined total evidence trees jModelTest determined that the optimal evolutionary model was TVM+I+G based on the Bayesian information criterion. The results were the same for both trees [-lnL=25360.7105; K=213; AIC=51147.4210; f(A)=0.3280; f(C)=0.2029; f(G)=0.1206; f(T)=0.3494; I=0.3570; G=0.5050].

Best-fit models were implemented in Bayesian analyses using MrBayes 3.2.2 (Ronquist and Huelsenbeck, 2003). Every Markov chain in the Bayesian search was started from a random tree and set to compute 1×10^7 generations, sampling every 1000^{th} one from the chain, resulting in a total of 1000 trees. Three hot chains and 1 cold chain were run simultaneously, with pre-stationarity trees discarded as burn-in. Each simulation was run twice. Default settings for the priors were used, and the base frequencies were estimated from the data. Tracer 1.6 (Rambaut et al., 2013) was used to parse and combine the log files, determine at which point the Markov Chain Monte Carlo (MCMC) began to sample from the stationarity distribution, and to check that effective sample sizes (ESSs) were sufficient for all parameters. To reduce the probability of convergence on local optima, multiple starting points for each chain were used. Maximum likelihood analysis was performed using RAxML-HPC2 (Stamatakais, 2006; Stamatakis et al., 2008), as implemented in CIPRES-XSEDE (Miller et al.,

2010). Analyses were conducted using the evolutionary model GTRGAMMAI (GTR+I+G) as well as the default model (GTRCAT), each with 1,000 bootstrap replicates, with the data partitioned by codon position and gene. There were no discernible differences between approaches. Bayesian posterior probabilities and nonparametric bootstrap proportions were used to assess node support (Felsenstein, 1985).

Morphological methods

Individuals from the same collection localities of the *O.celsa* specimens examined molecularly were cleared and slide mounted for chaetotaxic pattern determination. Five small glass dishes were filled with the following solutions: 95% ethanol, distilled water, 10% potassium hydroxide (KOH), distilled water, and the final dish filled only halfway with distilled water.

Each specimen was put into the dish with 95% ethanol and dissected with a fine-tipped forceps and a cornea knife. The head, furcula, and metathoracic leg were severed and put into distilled water. After the prothoracic and mesothoracic legs were removed from the thorax and discarded, the thorax and abdomen were placed in distilled water. The various body parts remained in the distilled water until all of the ethanol had been exchanged with water, approximately 15 minutes.

The body parts were then carefully transferred to the dish containing the KOH solution where the clearing of the integument occurred. If possible, the gut contents of the specimen were carefully removed from the body by either splitting the specimen ventrally, or by carefully squeezing the digestive tract out of the body. After three minutes in this solution, the dorsal surface of the body and the head were carefully shaved with the cornea knife to remove the macro-setae from the body for easier identification.

After approximately 10 minutes in the KOH solution the body parts were transferred to a dish filled with distilled water where they were immersed for at least 15 minutes. All of the body parts were then transferred to the dish filled halfway with distilled water. Drops of 95% ethanol were added to the distilled water until the dish was full. The liquid was then carefully removed until half of it remained and more drops of 95% ethanol could be added. This was repeated two more times until the liquid in the dish was approximately 95% ethanol.

The body parts were then transferred to a covered dish containing Andre's Solution I. The parts were incubated for at least 10 minutes to finish clearing. Once that incubation was done, two drops of Hoyer's Medium were placed on a clean glass slide. The head, furcula, and metathoracic leg were carefully placed in one drop. The head was positioned dorsal side up. The body was placed in the other drop, also dorsal side up. A cover slip was placed on each drop and carefully flattened.

The slides were then labeled and incubated in a 50°C incubator for 48 hours to harden.

The slides were viewed on a compound microscope at 200 magnification and the morphological sketches were made with the aid of a drawing tube. Identification of bothriotrichia and pseudopores was done under oil immersion at 600 magnification.

Results

Molecular results

Strong support was obtained for an *O.celsa* clade for all three resulting trees (BP=100, PP=1.0 for all three inferences). This clade was comprised of several strongly supported subgroups, with relationships within groups differing by gene and analysis. Contained within this group are two named species, *O. gloriosa* and *O.celsa sensu stricto*.

Mitochondrial analyses.

There was moderate support for two different groups in the *O.celsa* clade, with one group (BP=73) composed of OTUs found more consistently at higher elevations, and the other group (BP= 86, PP=1.0) comprised of OTUs found more consistently at lower elevations. Within the upper elevation group there are two major divisions. There was strong support for one group of three closely related specimens (MMSP1, BRP4, NCCG) (BP=97, PP=1.0), with a sister group relationship to SUL (BP=93, PP=1.0). The other grouping of the mountain clade had moderate to strong support (BS=73, PP=1.0) for the following relationships: FHSP1 and FHSP3 as sister group to three different strongly supported groups- SMcel and OSR1OB (BS= 97, PP=0.99), OglorCD and PcelCD, BRP2, ODC2, and BRP3 (BS= 100, PP=1.0).

Within the lowland group, there was strong support for one clade (BS=98, PP=1.0) containing *O.celsa sensu stricto* (SCBM1, celWNC, O4AR1, Pcel1, RCSC2, and RCSC3), and weak support for a second clade (BS=59, PP=0.82) containing (CelLCO, OcelFL, celLF, ARB3, VB2, OcelSP, OcelCS, OcelTC1, ONC4, celEB, and MtS1).

Nuclear analyses

There was strong support for the *O.celsa* clade (BS=100, PP=1.0), but varying levels of support within the group. The lowland group was recovered but with no real statistical support, and the mountain group was recovered as a paraphyletic laddering with taxa.

Nuclear data indicated that seven strongly supported groupings were present: FHSP1 and FHSP2 (BS= 100, PP=1.0), PcelCD, ODC2 and ONC3 (BS=100, PP=1.0), celsaSM and OSR1 (BS= 96, PP=1.0), MMSP1 and NCCG (BS=100, PP=1.0), celsaTC2 and ONC4 (BS = 100, PP = 0.97), O4AR1, Pcel1,

SCBM1, RCSC2, RCSC3 (BS=97, PP=1.0), and celWNC, ARB3, OcelSP, VB2, celLF, celLCO, ocelFL (BS=70, PP=0.94).

Other than the position of celWNC within the lowland group the two major lowland group clades are recovered the same as in the mitochondrial tree.

Total evidence (combined data)

The total evidence tree mirrors the mitochondrial tree with the support of the upper elevation and lower elevation groups and the subgroups contained within.

Discussion

The two subgroups of the *O. celsa* clade tend to split along elevation preferences, with the clade containing *O. gloriosa* preferring peaks over 400 meters, and in the case of specimens OglorCD and PcelCD elevations over 2,000 meters. The other clade, containing *O. celsa sensu stricto* (celWNC) is more readily found at lower elevations, typically those under 600 meters. This suggests that there could be some overlap in the ranges of this group, but this has not been observed in the field.

Orchesella gloriosa, collected from mossy seeps at the type locality (Clingmans Dome), is more closely related to specimen PcelCD collected from lichen- covered trees on Clingmans Dome. This relationship suggests that *O.gloriosa* speciated from the other high elevation *celsa*-types by exploiting a unique niche on Clingmans Dome. If the species concept of *O.gloriosa* is valid, then the rest of the upper elevation groups must be similarly defined.

Other OTUs in the higher elevation group seem to group along connected mountain ridges: MMSP1, BRP4, and NCCG1, and PcelCD, BRP2, ODC2, and BRP3. Specimens FHSP1 and FHSP3 may have closely related groups if other mountains in the area were sampled. Specimens SMcel and OSR1 are more distantly related to each other than the OTUs of other groups, but they are from approximately 120 kilometers apart. Additionally, the urbanization of the areas around Sharp's Ridge Park where OSR1 was collected have effectively isolated this population.

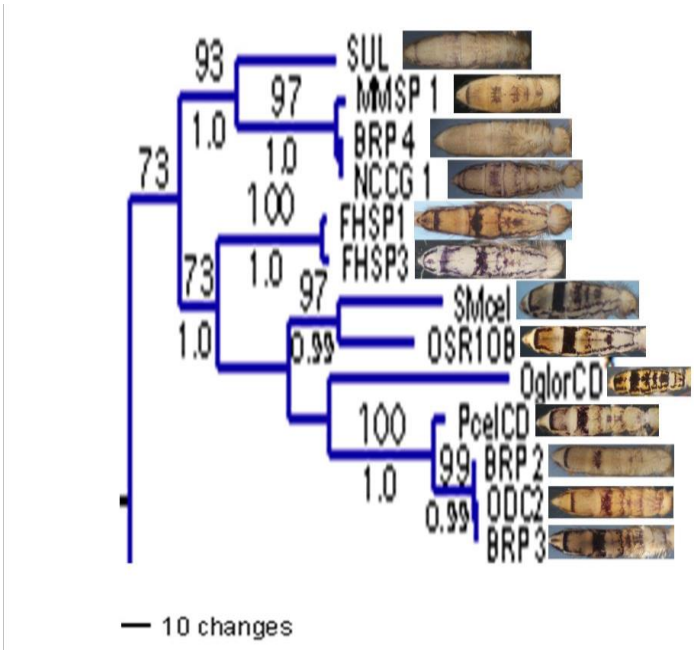

Figure 7. Portion of the mitochondrial tree showing the upper elevation group of the *celsa* clade with dorsal pattern images.

The lower elevation group containing SCBM1, celWNC (*O.celsa sensu stricto),* O4AR1, Pcel1, RCSC2, and RCSC3 were collected from the Piedmont and coastal plain of the Carolinas on the eastern side of the Smoky Mountains. Images of Florida specimens on BugGuide (http://bugguide.net/node/view/610760/bgimage) appear to be assignable to this species:

The other lower elevation group containing celLCO, OcelFL, celLF, ARB3, VB2, OcelSP, OcelCS, OcelTC1, ONC4, celEB, and MtS1 are mainly from lower elevations on the western side of the Smoky Mountains.

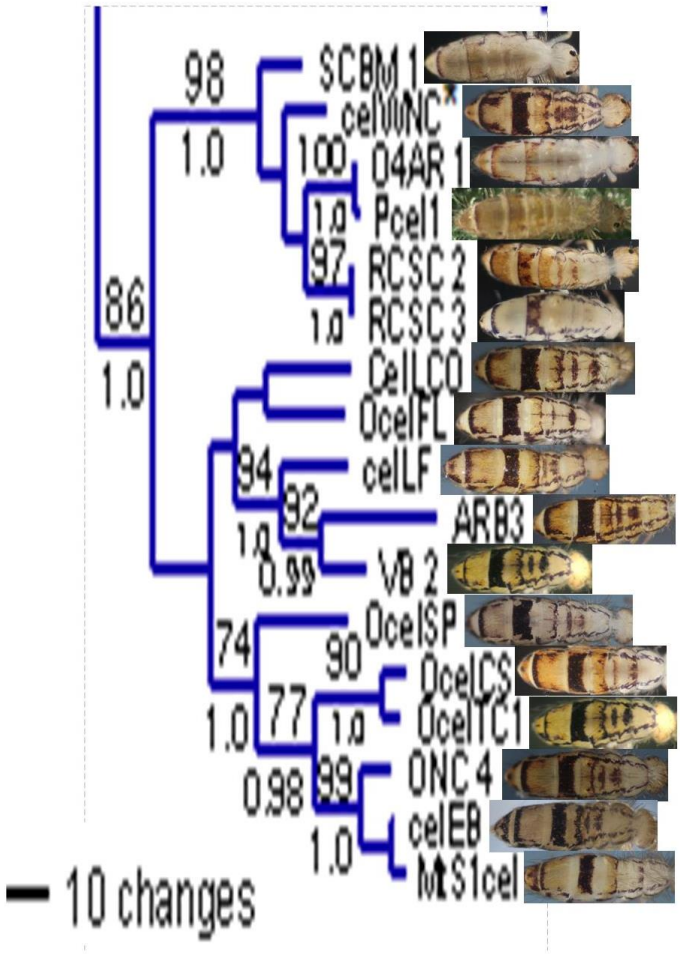

Figure 8. Section of the mitochondrial tree showing the lower elevation group of the *celsa* clade with dorsal images.

The chaetotaxic descriptions of the proposed new species is based off of the scheme used in Potapov (2008) in his description new species of Old World *Orchesella*. Specimens of *O. celsa* from the type locality of New Hanover, North Carolina were collected, and slide mounted. Chaetotaxic sketches were made from those specimens, and these sketches were compared to similar chaetotaxic sketches made from slide-mounted specimens representing the proposed new species to elucidate any potential meaningful differences.

The main differences in chaetotaxic patterns can be found in the number and arrangement of the macrosetal pores of the second, third, and fourth abdominal segments.

***Orchesella celsa* Christiansen & Tucker, 1977**

Figure 9. Chaetotaxic scheme of *Orchesella celsa sensu stricto*.

Table 4. Collection data for specimens of *Orchesella celsa sensu stricto*.

Specimen Designation	Collection Data
celWNC* (*Orchesella celsa*)	NC: New Hanover Co., near Wilmington
O4AR1	SC: Lancaster Co., 40-Acre Rock WMA
OcelSP	SC: Lexington Co., Shealy's Pond Heritage Preserve
Pcel1	SC: Georgetown Co., Pawley's Island
RCSC2	SC: Chesterfield Co., Near Rocky Creek
RCSC3	SC: Chesterfield Co., Near Rocky Creek
SCBM1	SC: Georgetown Co., Black Mingo Creek

Orchesella celsa sensu stricto, has on the second abdominal segment six to seven macrosetae in A1, six macrosetae in A2 arranged in a pattern of two-three-one. On the third abdominal segment, it has A3+A4 with three macrosetae, and A5 with two macrosetae. The fourth abdominal segment has no macrosetae in A6, so it is not shown. A7 has three macrosetae, A7+A8 has two macrosetae and one pseudopore, and A10 has one macrosetae.

***Orchesella baeni* n. sp.**

Orchesella baeni is formed by the specimens SUL, MMSP1, BRP4, and NCCG1. It is named after Jim Baen, a gentleman who kept me interested in science throughout my formative years.

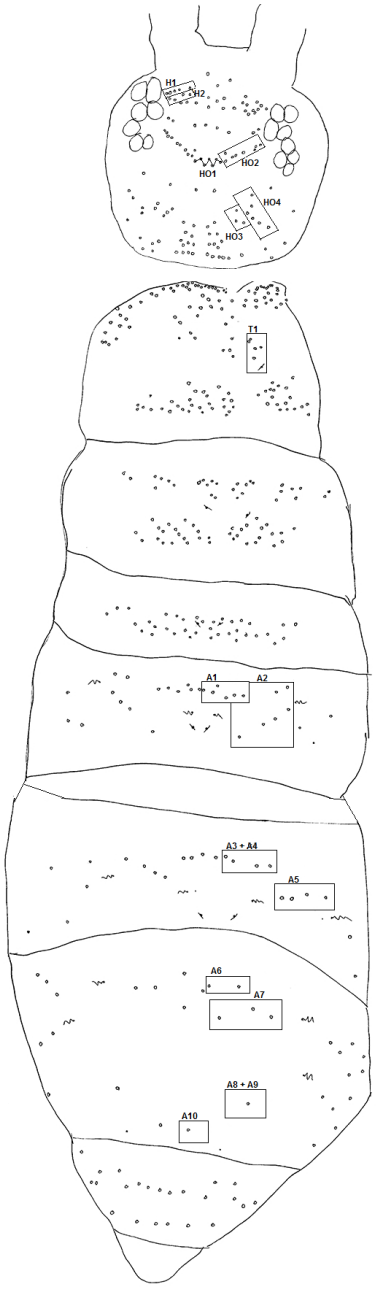

Figure 10. Chaetotaxic scheme of proposed species *Orchesella baeni*.

Table 5. Collection data for specimens of *Orchesella baeni* n. sp.

Specimen Designation	Collection Data
BRP4	NC: Buncombe Co., Blue Ridge Parkway
MMSP1	NC: Yancey Co., Mt. Mitchell State Park
NCCG1	NC: Buncombe Co., Blue Ridge Parkway
SUL	TN: Sullivan Co., Cherokee National Forest

Abdominal segment 2 of *Orchesella baeni* n.sp. has A1 with six macrosetae, and A2 with six macrosetae arranged in a pattern of two-three-one. On the third abdominal segment, this species has A3 + A4 with four macrosetae, and A5 has four macrosetae. On the fourth abdominal segment, A6 has two macrosetae, A7 has three macrosetae, A8 + A9 has one to two macrosetae, and A10 has one macrosetae.

***Orchesella jotunensis* n. sp.**

Orchesella jotunensis n. sp. is formed by the two specimens: FSHP1 and FSHP3. The species name comes from the Norse name “Jotunheim” for the land where the frost giants dwelled. The specimens were collected from Frozen Head State Park in Morgan County, Tennessee.

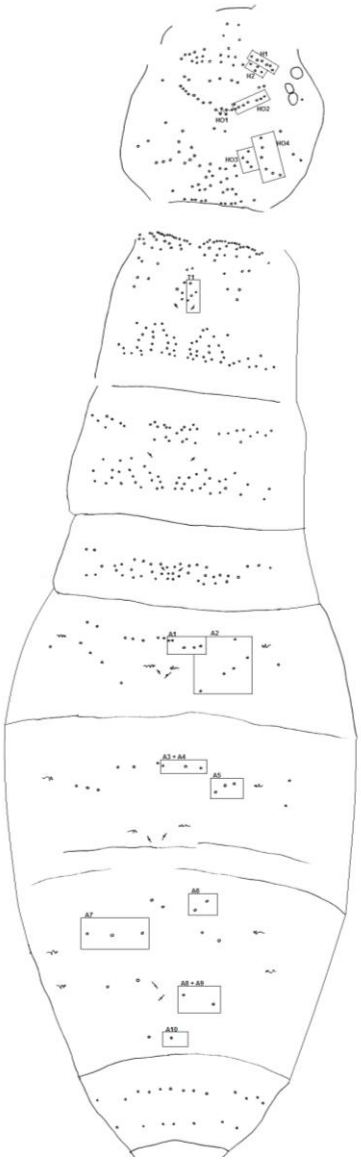

Figure 11. Chaetotaxic scheme of proposed species *Orchesella jotunensis*.

Table 6. Collection data for specimens of *Orchesella jotunensis* n. sp.

Specimen Designation	Collection Data
FHSP1	TN: Morgan Co., Frozen Head State Park, Lookout Tower Trail
FHSP3	TN: Morgan Co., Frozen Head State Park, Lookout Tower Trail

In A1, *Orchesella jotunensis* n. sp. has five macrosetae, and A6 has five macrosetae. On the third abdominal segment, A3 + A4 has three macrosetae and A5 has three macrosetae. This species has A6 present with two macrosetae, A7 has three macrosetae, A8 + A9 has two macrosetae, and A10 has a single macrosetae.

***Orchesella luteusnota* n. sp.**

Specimens SMcel and OSR1 form the proposed species *Orchesella luteusnota*. The species name comes from the Latin “luteus”, meaning “yellowy” or “saffron-colored”, and from the Greek “nota”, meaning “back”, or “rear”. This name is in reference to the orange or saffron-colored fourth abdominal segment in the preserved specimens.

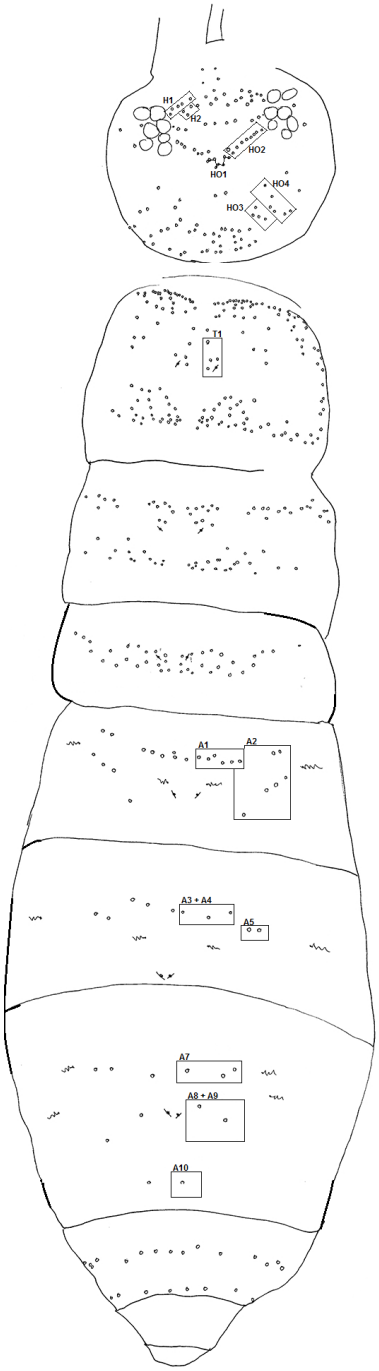

Figure 12. Chaetotaxic scheme of the proposed cryptic species *Orchesella luteusnota.*

Table 7. Collection data for specimens of *Orchesella luteusnota* n. sp.

Specimen Designation	Collection Data
OSR1OB	TN: Knox Co., Sharp's Ridge Park
SMcel	TN: Cooke Co., Cherokee National Forest

Based on the dorsal chaetotaxy, this species is similar to *O. celsa sensu stricto*; however, genetically it is affiliated with the higher elevation group. The color pattern of this group, particularly those specimens collected at Sharp's Ridge Park, is very distinct with the fourth abdominal segment having a dark U-shaped border and a distinctly orange interior. A more in-depth examination may uncover additional morphological differences.

***Orchesella diabolica* n. sp.**

Speciemens PcelCD, BRP2, ODC2, and BRP3 form the proposed species *Orchesella diabolica*. The species name is in reference to Devil's Courthouse, a mountain peak in the Appalachian Mountains in Transylvania County, North Carolina where a large population of specimens was collected.

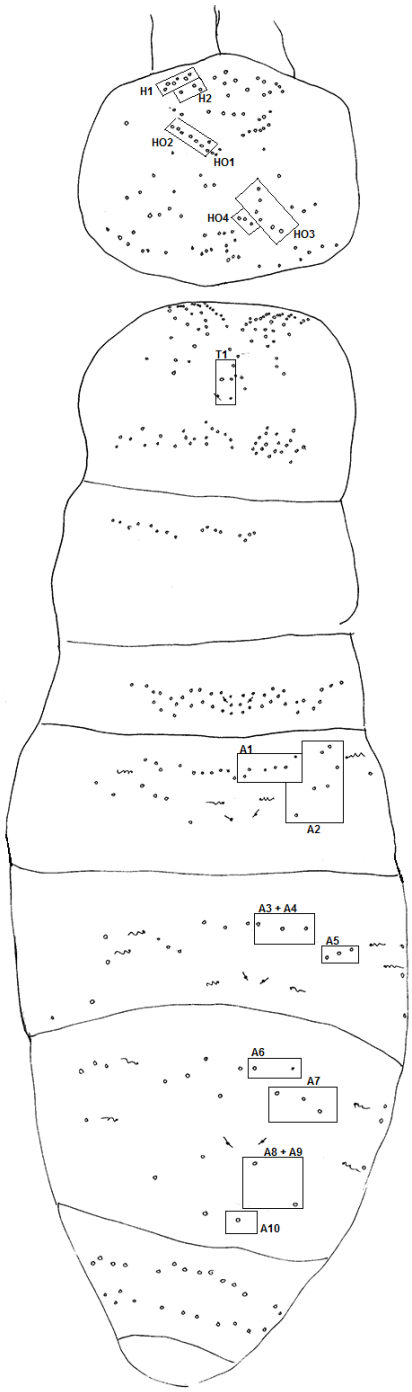

Figure 13. Chaetotaxic scheme of proposed species *Orchesella diabolica*.

Table 8. Collection data for specimens of *Orchesella diabolica* n. sp.

Specimen Designation	Collection Data
BRP2	NC: Jackson Co., Blue Ridge Parkway, Fork Ridge Overlook
BRP3	NC: Jackson Co., Blue Ridge Parkway, Fork Ridge Overlook
ODC2	NC: Transylvania Co., Devil's Courthouse
PcelCD	NC: Swain Co., Clingmans Dome, Clingmans Dome Trail

In *Orchesella diabolica* n. sp., A1 has six macrosetae and A2 has six macrosetae on the second abdominal segment in a two-three-one pattern (though about half the specimens show a three-three-one pattern). A3 + A4 has three macrosetae, and A5 has three macrosetae on the third abdominal segment. On the fourth abdominal segment, A6 has two macrosetae, A7 has three macrosetae, A8 + A9 has two macrosetae, and A10 has a single macrosetae.

***Orchesella abramsi* n. sp.**

The final proposed new species, *Orchesella abramsi*, is comprised of the lower elevation specimens CelLCO, OcelFL, celLF, ARB3, VB2, OcelSP, OcelCS, OcelTC1, ONC4, celEB, and MtS1cel. This species is named after Peter Abrams, a gentleman who has given me a great deal of support and inspiration throughout my career in entomology.

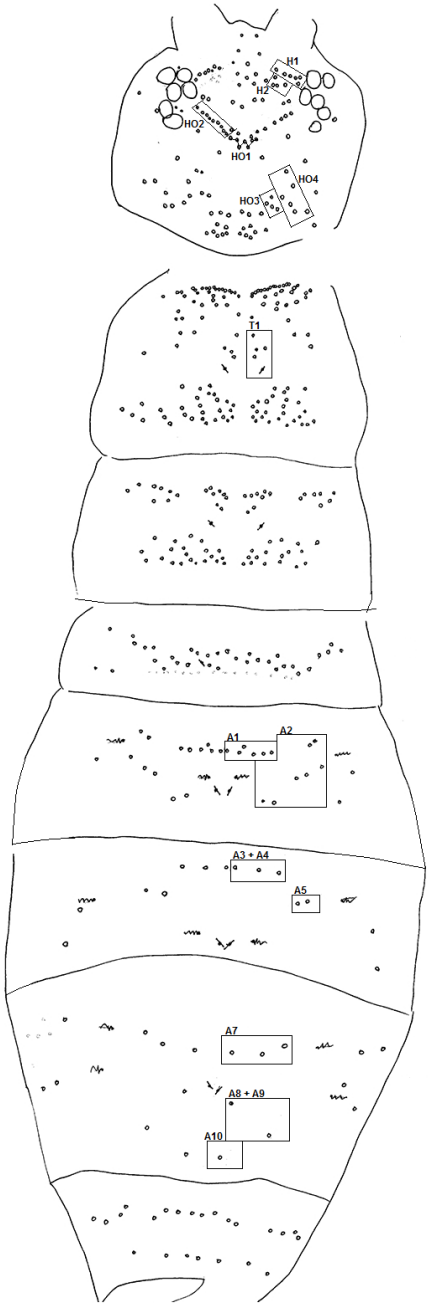

Figure 14. Chaetotaxic scheme of proposed species *Orchesella abramsi*.

Table 9. Collection data for specimens of *Orchesella abramsi* n. sp.

Specimen Designation	Collection Data
ARB3	TN: Anderson Co., UT Arboretum
celEB	NC: Henderson Co., Blue Ridge Parkway
celLCO	AL: Cleburne Co., Choccolocco Forest
celLF	GA: Walker Co., Lula Lake Land Trust
MtS1cel	NC: Haywood Co. , Mt. Sterling
OcelFL	FL: Liberty Co., Torreya State Park
OcelTC1	TN: Sevier Co., GSMP, Twin Creeks Pavilion
OcelCS	TN: Monroe Co., Indian Boundary Campground
ONC4	NC: Graham Co., SR 143 near Joyce Kilmer Memorial Forest
VB2	TN: VanBuren Co.

Based on the chaetotaxy, this proposed species is very similar to *Orchesella celsa sensu stricto*. Genetically, however, they are different. For now it is best this species be considered a weakly differentiated cryptic species near *O. celsa s. str.*

References

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

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, **22**:5382-5396.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest2: more models, new heuristics and parallel computing. *Nature Methods,* **9**:772.

Drummond AJ, Rambaut A, (2003) BEAST v1.0.

Frati F, Carapelli A (1999) An assessment of the value of nuclear and mitochondrial genes in elucidating the origin and evolution of *Isotoma klovstadi* Carpenter (Insecta, Collembola). *Antarctic Science*, **11**:160-174.

Frati F, Spinsanti G, Dallai R (2001) Genetic variation of mtCOII gene sequences in the collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population differentiation. *Polar Biology*, **24**:934-940.

Guindon S, Gascuel O (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, **52**(5):696-940.

Hopkin, SP (1997) Biology of the Springtails (Insecta: Collembola). Oxford University Press.

Luque C, Legal L, Winterton P, Mariano NA, Gers C (2011) Illustrations of the Structure of Arthropod Assemblages *(Collembola* and *Lepidoptera)* in Different Forest Types: An Example in the French Pyrenees. *Diversity*, **3**:693-711.

Maddison WP, Maddison DR (2011) Mesquite 2.75: a modular system for evolutionary analysis.

Maynard EA (1951) A Monograph of the Collembola or the Springtail Insects of New York State. Comstock Publishing CO. Ithaca, New York.

McGaughran A. Torricelli G, Carapelli A, Frati F, Stevens MI, Convey P, Hogg ID (2010) Contrasting phylogeographical patterns for springtails reflect different evolutionary histories between the Antarctic Peninsula and continental Antarctica. *Journal of Biogeography*, **37**:107-119.

Porco D, Potapov M, Bedos A, Busmachiu G, Weiner WM, Hamra-Kroua S, Deharveng L (2012) Cryptic Diversity in the Ubiquist Species *Parisitoma notabilis* (Collembola, Isotomidae): A Long-Used Chimeric Species? *PLoS ONE*, **7**(9).

Porco D, Skarzynski D, Decaens T, Herbert PDN, Deharveng L (2013) Barcoding the Collembola of Churchill: a molecular taxonomic reassessment of species diversity in a sub-Arctic area. *Molecular Ecology Resourcs*, **14**(2): 249-261.

Potapov M, Kremenitsa A (2008) Comments on the chaetotaxy of the genus *Orchesella* (Collembola, Entomobryomorpha) with a redefinition of the '*spectabilis*' group and description of a new species of *Orchesella* from the Caucasus. *Soil Orginisms*, **80**(1):99-115.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**(22):1572-1574.

Swofford D, (2002) (PAUP*)Phylogenetic Analysis Using Parsimony (*and other methods) V4.0. Sinaurer Associates, Sunderland.

Szeptycki A, (1979) Chaetotaxy of the *Entomobryidae* and its Phylogenetic Significance. Morpho-systematic studies of Collembola. –IV. Polska Akademia Nauk, Zaklad Zoologii Systematycznej Doswiadczalnej.

Yoshii R, (1989) On some Collembola of New Caledonia, with notes on the "colour pattern species". Contr. Biol. Lab. Kyoto University. **27**(1989): 233-259.

CHAPTER III

POPULATIONS OF THE EUROPEAN SPRINGTAILS *ORCHESELLA CINCTA* (L.) AND *ORCHESELLA VILLOSA* (L.) IN NORTH AMERICA

Abstract

North American specimens of the European invasive springtail *Orchesella cincta* were compared to several published European haplotypes in a phylogenetic framework using likelihood methods based on a portion of cytochrome oxidase II. Our analyses provide direct evidence of at least two distinct introductions of this invasive to North America from different regions of Europe. Additional introduction events cannot be ruled out because detection is limited by extremely low sequence divergence among populations inhabiting different regions of the continent. *Orchesella villosa*, another invasive from Europe, is another candidate for multiple introductions. Herein we include the CoxII sequence from single specimens of *O. villosa* from Maine and Oregon. Although these two specimens are identical in sequence, they differ from a published sequence from a European specimen by 15%, indicating significant undocumented genetic variation in the natal range of *O. villosa*. Additional sampling of Nearctic populations of *O. villosa* might reveal the same situation reported herein for *O. cincta*.

Introduction

The hexapod class Collembola consists of small, primitively wingless hexapods with an ancient worldwide distribution (Hopkins 1997). Commonly known as springtails, they primarily feed on lichens and fungal hyphae, where they are an important part of the soil formation process (Hopkins 1997). The springtail genus *Orchesella* has a Holarctic distribution, with members found in North America, Europe, and Western Asia (Frati et al., 1992; Stach, 1960). Of the Old World members, two species have been accidentally introduced to North America: *O. villosa* (L.) and *O. cincta* (L.). It is thought that both species were originally introduced to North America in the ballast of early European ships, or when agricultural plants or products were transported to the New World (Maynard, 1951); probably both species have been introduced repeatedly since then. Both *O. cincta* and *O. villosa* have since become established and widespread but spotty in North America (Christiansen & Bellinger, 1998), most often in disturbed habitats of the northern United States and southern Canada along coasts and near seaports. *Orchesella cincta* also has been reported from Portland, Oregon (Scott, 1942), where native Nearctic congeners apparently are absent (Christiansen & Bellinger, 1998).

Timmermans et al. (2005) used the mitochondrial gene, cytochrome oxidase II (CoxII), to study the genetic relationships among different populations of *O. cincta* in Europe. Their findings suggested there were three main

population groups within the native range of *O. cincta* (northwestern Europe, central Europe, Italy). The objective of this research was to use the data of Timmermans et al. to determine whether the presence of *O. cincta* populations in the Nearctic Region, is due to single or multiple introductions. To meet this objective we obtained sequences of cytochrome oxidase II corresponding to the fragment used by Timmermans et al. (2005) from several Nearctic populations of *O. cincta,* added these sequences to their data, and conducted phylogenetic analyses. A similar but less involved analysis was conducted on two specimens of *O. villosa*, one each from Maine and Oregon.

Methods and Materials

Taxon sampling

Specimens of *Orchesella cincta* were collected from Wisconsin, Minnesota (two sites), Maryland, Michigan, Oregon, and Washington. Specimens of *O. villosa* were collected from Maine and Oregon. Most specimens were aspirated from under the bark of decaying logs, collected by sifting leaf litter, or beaten from lichen-covered branches; the Maryland specimen of *O. cincta*, or extracted from leaf litter samples by means of a Tullgren funnel. Voucher specimens and skins are deposited at the University of Tennessee (E. C. Bernard Research Collection). When multiple specimens were collected (Minnesota, Wisconsin, Michigan, Oregon), the largest individual was photographed prior to total genomic DNA extraction for vouchering purposes. Outgroups added to root the phylogeny included *Dicranocentrus marias* Wray and *Heteromurus major* (Moniez), 1889) Absolon, 1901, several Nearctic *Orchesella* species, viz., *Orchesella hexfasciata* Harvey, *Orchesella gloriosa* Snider, and *Orchesella annulicornis* Mills, and the closely related European species *O. villosa* (represented by the Maine and Oregon specimens). Collection information for all specimens is given in Table 10.

Laboratory methods

Total genomic DNA was extracted using the ThermoScientific GeneJET Genomic DNA Extraction (ThermoScientific, Waltham, MA) kit following the manufacturer's suggested protocol, except for the final step where the gDNA was eluted in 70µL of elution buffer warmed to 52°C. Purified DNA samples were stored at -20°C. We used the forward primer used by Timmermans et al. (2005): TL2-J-3037 (5'-AATATGGCAGATTAGTGC-3') (Simon et. al, 1994) and a custom reverse primer designed by JKM (COII-R: 5'-CCACAGATTTCTGAGCATTGACC-3'), which amplify a 563 base pair fragment from the mitochondrial genome comprising the extreme 3' end of tRNA-leucine (i.e., the first 15 bases of polished final sequences) and the ca. 5' half of cytochrome oxidase II (COXII). This gene

region has been used previously and successfully for phylogenetic purposes in Collembola (Frati and Carapelli, 1999; Frati et al. 2001; McGaughran et al. 2010; Luque et al. 2011).

Table 10. Sources of North American specimens used in this study.

Taxa	Country	State	County	GB Accession #
Heteromerus major	USA	Oregon	Lane	TBD
Dicranocentrus marias	Puerto Rico	N/A	Isabela	TBD
Orchesella gloriosa	USA	Tennessee	Sevier	TBD
Orchesella annulicornis	USA	Illinois	Will	TBD
Orchesella hexfasciata	USA	Maine	Hancock	TBD
Orchesella villosa USA: Oregon	USA	Oregon	Lane	TBD
Orchesella villosa USA: Maine	USA	Maine	Hancock	TBD
Orchesella cincta USA: Michigan	USA	Michigan	Alger	TBD
Orchesella cincta USA: Oregon	USA	Oregon	Lane	TBD
Orchesella cincta USA: Washington	USA	Washington	Clallam	TBD
Orchesella cincta USA: Minnesota, TC[a]	USA	Minnesota	Hennepin	TBD
Orchesella cincta USA: Wisconsin	USA	Wisconsin	Taylor	TBD
Orchesella cincta USA: Maryland	USA	Maryland	Anne Arundel	TBD
Orchesella cincta USA: Minnesota, LISP[b]	USA	Minnesota	Clearwater	TBD

[a] TC = Twin Cities, i.e., Minneapolis-St. Paul environs
[b] LISP = Lake Itasca State Park

Amplifications were performed in GenePro (Bioer Technology Co., Hangzhou, China) thermal cyclers, using TaKaRa Ex Taq Hotstart DNA polymerase (Takara Bio, Shiga, Japan) per the manufacturer's suggested

protocol with 1.3µL template DNA, and 3µL of (7mM) of forward and reverse primer. Thermal cycling parameters were as follows: 1.5 min denaturation soak at 94°C; 5 cycles of 94°C for 30s, 52°C for 30 s, and 72°C for 45s; 15 cycles of 94°C for 30s, 47°C for 25s, and 72°C for 45s; 35 cycles of 94°C for 30s, 42°C for 20s, 72°C for 45s, 72°C soak for 3 min, and 13°C hold. PCR products were electrophoresed in 1% agarose, excised from the gel, and purified using a QiaQuick Gel Extraction Kit (Qiagen, Valencia, CA). The reverse strand of each product was cycled sequenced in 20µL reactions using 16-fold diluted Big Dye 3.1 (Applied Biosystems, Foster City, CA). Sequencing reactions were cleaned using Centrisep columns (Prinecton Separations, Adelphia, NJ) and dried in a Centrivap Concentrator (LABCONCO, Kansas City, MO). Sequencing of samples was performed by the University of Tennessee-Knoxville Molecular Biology Resource Facility. Sequences were verified for accuracy using Sequencher 4.7 (GeneCodes, Ann Arbor, MI).

DNA alignment and phylogenetic analysis

Sequences from all European populations of *O. cincta* reported in Timmerman et. al (2005) were added to our data matrix of Nearctic *O. cincta* populations plus selected outgroups. Alignment of the sequences was straightforward, requiring no indels. Mesquite (Maddison and Maddison, 2011) was used to partition the data set into codon positions. PAUP* (Swofford, 2001) was used to calculate pairwise sequence divergence. HKY85 corrected distances were calculated because this model was selected as the best fit by jModelTest (see Figure 1).

Bayesian and maximum likelihood analyses were performed on the nucleotides. The optimal evolutionary model was determined using jModelTest 2.1.4 (Guindon and Gascuel, 2003; Darriba et al., 2012), which selected the HKY+I+G as the best fit model based on the Bayesian information criterion [-lnL=3885.2396;K=54; BIC=8112.4764; f(A)=0.3358; f(C)=0.1979; f(G)=0.1087; f(T)=0.3577; Ti/tv=4.7769;I=0.4620; G=1.018].

Best-fit models were implemented in Bayesian analyses using MrBayes 3.2.2 (Ronquist and Huelsenbeck, 2003). Every Markov chain in the Bayesian search was started from a random tree and set to compute 1×10^7 generations, sampling every 1000^{th} one from the chain, resulting in a total of 1000 trees. Three hot chains and 1 cold chain were run simultaneously, with pre-stationarity trees discarded as burn-in. Each simulation was run twice. Default settings for the priors were used, and the base frequencies were estimated from the data. Tracer 1.6 (Rambaut et al., 2013) was used to parse and combine the log files, determine at which point the Markov Chain Monte Carlo (MCMC) began to sample from the stationarity distribution, and to check that effective sample sizes (ESSs) were sufficient for all parameters. To reduce the probability of convergence on local optima, multiple starting points for each chain were used. Maximum likelihood analysis was performed using RAxML-HPC2 (Stamatakais,

2006; Stamatakis et al., 2008), as implemented in CIPRES-XSEDE (Miller et al., 2010). Analyses were conducted using the evolutionary model GTRGAMMAI (GTR+I+G) as well as the default model (GTRCAT), each with 1,000 bootstrap replicates, with the data partitioned by codon position and gene. There were no discernible differences between approaches. Bayesian posterior probabilities and nonparametric bootstrap proportions were used to assess node support (Felsenstein, 1985).

Results

mtDNA data set properties

The analyzed data matrix was comprised of 25 taxa, including 2 distal outgroups, 4 proximal outgroups from the Nearctic region, 2 proximal outgroups from Europe (1 specimen of *O*. *villosa* from Maine that was identical in sequence to a specimen taken in Oregon and 1 from Europe (Carapelli et al., 2007), 7 Nearctic individuals of *O. cincta*, and 11 specimens of *O. cincta* from Europe (Timmermanns et al., 2005). CoxII sequences for the newly acquired North American samples are deposited in GenBank (Table 10).

Percent pairwise divergence between distal outgroups ranged from 30.4 (*Heteromurus major* and *Dicranocentrus maria*) to 38.9 (*Heteromurus major* and *O*. *gloriosa*). Percent pairwise divergence between *O. villosa* and *O. cincta* populations ranged from 21.3 (*O*. *villosa* Maine and *O. cincta* Oregon) to 28.7 (*O*. *villosa* Europe and Italy 3). Percent pairwise divergence within populations of *O. cincta* ranged from 0.0% (12 comparisons) to 19.5 % (Italy 1 and Italy 3). Pairwise divergence between USA and European specimens of *O. villosa* was 15.2%.

The North American specimens of *O. cincta* from Oregon, Washington, and Minnesota were identical to published sequences from specimens examined by Timmermans et al. (2005) from northwestern Europe (NW Europe 1 and NW Europe 6). Our Michigan *O. cincta* specimen differed from the Timmermans et al. (2005) Italy 5 specimen by 1.1% divergence.

Results of phylogenetic analyses

With *Heteromurus major* selected as the root, *Dicranocentrus marias* was the sister group to *Orchesella* spp. The three Nearctic representatives of *Orchesella* were strongly supported (BS = 97, PP = 1.0) as a monophyletic group and formed the sister group to a weakly supported *O. villosa* plus *O. cincta* clade. Monophyly of *O. cincta* was strongly supported (BS = 100, PP = 1.0). Within *O. cincta*, Italy 1 formed the sister group to all remaining sampled populations, although with weak support. A strongly supported (BS = 91, PP = 1.0) clade of three Italian *O. cincta* populations (Italy 3, Italy 4, and Italy 6) formed the sister group to Italy 5 plus Michigan, which were strongly supported as sister taxa (BS

= 100, PP = 1.0) and in turn were sister to a well-supported (BS=94, PP=0.99) clade comprised of all remaining sampled *O. cincta* specimens.

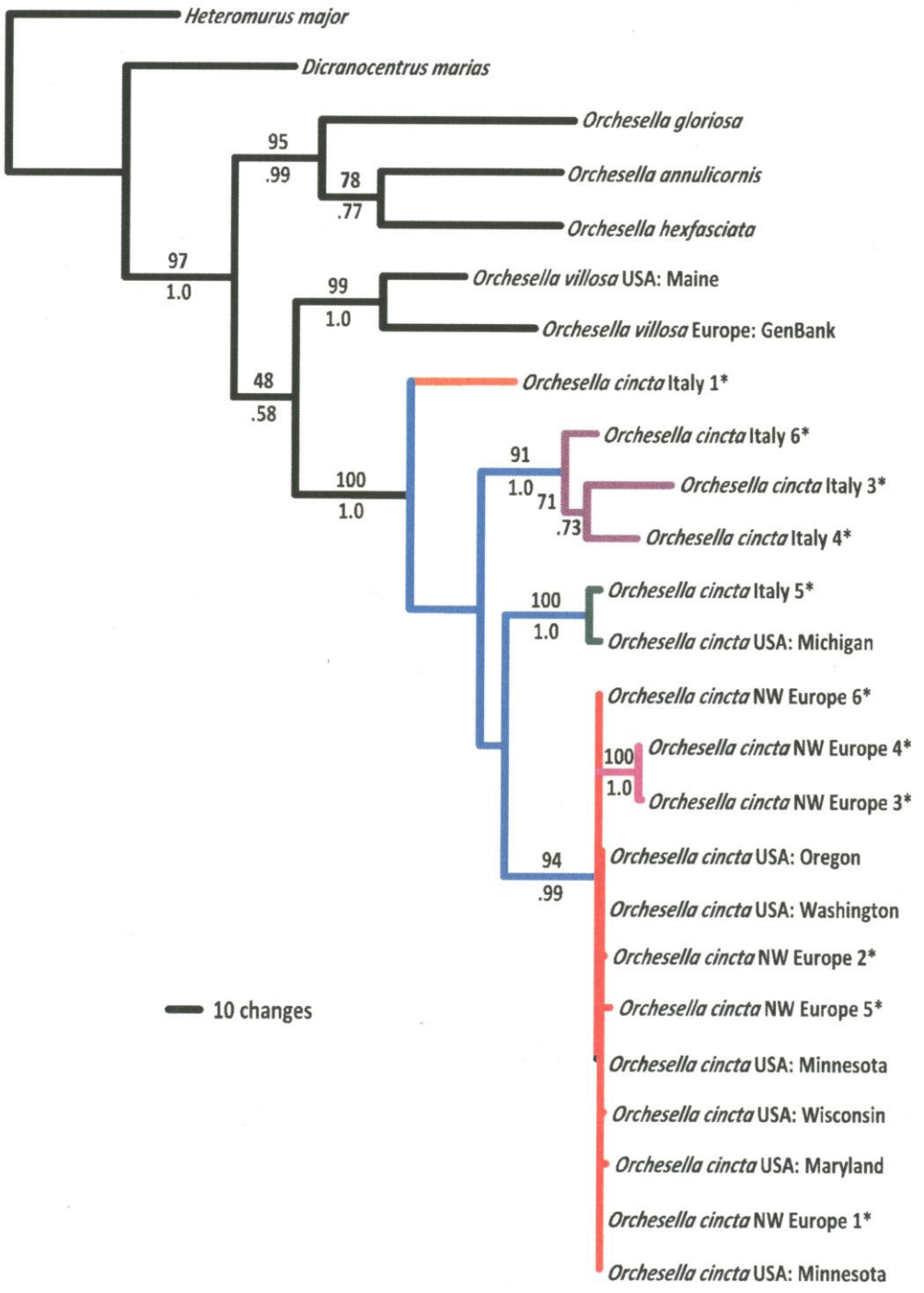

Figure 15. COXII tree with bootstrap values and RaxML values.

* indicates data from Timmermanns et. al 2005 paper

Discussion

With one exception, all sampled North American *O.cincta* specimens matched up closely with an assemblage of weakly or undifferentiated populations found in northwestern Europe. Immigration patterns of cultural groups from this region of Europe into the U.S. (Anderson and Blanck, 2012) would predict such an occurrence. Migration of humans and intercontinental trade of agricultural and horticultural goods from the Upper Midwest to the West Coast are the most likely explanation for western movement of newly established northwestern European *O. cincta* populations in the Nearctic. The one exception among the Midwestern USA populations was the specimen collected from Michigan, which paired strongly with a specimen Timmermans et al. (2005) examined from Italy (Italy 5). International shipping through the Saint Lawrence River/Great Lakes system to the large ports in Michigan could explain how this population was introduced. The presence of two distinct populations of *O. cincta* in the United States clearly indicates this species has been introduced multiple times from different parts of its native range.

The considerable genetic distance (i.e, 75 base pairs, 15.2% pairwise divergence) between the *O. villosa* specimens collected in Maine and Oregon, which were identical in sequence, and the specimen from which the published *O. villosa* mitochondrial genome (GB# NC_010534) was generated, is on par with observations made by Frati et al. (2000), who noted large amounts of differentiation among native populations of this species. A thorough phylogeographic study of this species in its native established ranges might yield data similar to what we present here for *O. cincta*.

The results reported here do not agree fully with the conclusions of Porco et al. (2013), whose analysis of barcodes for Canadian *O. cincta* and *O. villosa* vs. European specimens suggested similar genetic structure of European and North American populations due to massive and multiple introductions. Instead, it appears that both species have been introduced repeatedly but different haplotypes occur only at widely scattered sites.

Orchesella cincta is common throughout Europe (Stach, 1960) and sometimes more abundant in forest than any other springtail species (van Straalen, 1989). The potential effects of *O. cincta* in North America on native springtail populations are unknown but not likely to be of consequence, as the species has been present on the continent for at least 140 years without becoming widespread. Other much more recently introduced species (e.g., *Homidia socia* Denis *and Lepidocyrtus paradoxus* Uzel) now occur over much of the eastern half of temperate North America (see records in Christiansen & Bellinger, 1998). Packard (1873) noted *O. cincta* (as *O. flavopicta*) from Massachusetts. Since then it has been verified from U.S. localities stated in this paper as well as from coastal Maine and Vermont (Christiansen & Bellinger, 1998), and from the Canadian provinces of New Brunswick, Newfoundland, Nova Scotia and Ontario (Porco et al., 2013). Its apparent preference for cooler

regions, its large size and ease of identification make it a potentially good indicator species for assessment of climatic warming.

References

Anderson PJ, Blanck D (2012) Norwegians and Swedes in the United States. Minnesota Historical Society Press. St. Paul, Minnesota.

Carapelli A, Lió P, Nardi F, van der Wath E, Frati F (2007) Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *Evolutionary Biology* **7** (Suppl 2), S8.

Christiensen K, Bellinger P (1998) The Collembola of North America, north of the Rio Grande: A taxonomic analysis. Grinnell College. Grinnell, Iowa.

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology,* **22**:5382-5396.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModeltTest2: more models, new heuristics and parallel computing. *Nature Methods,* **9**:772.

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution,* **39**(4):783-791.

Frati F, Carapelli A (1999) An assessment of the value of nuclear and mitochondrial genes in elucidating the origin and evolution of *Isotoma klovstadi* Carpenter (Insecta, Collembola). *Antarctic Science,* **11**:160-174.

Frati F, Dell'Ampio E, Casasanta S, Carapelli A, Fanciulli PP (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.

Frati F, Faniciulli PP, Dallai R (1992) Genetic diversity and taxonomy in soil-dwelling insects: The genus *Orchesella*. *Journal of Heredity,* **83**:275-281.

Frati F, Spinsanti G, Dallai R (2001) Genetic variation of mtCOII gene sequences in the collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population differentiation. *Polar Biology,* **24**:934-940.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood. *Systematic Biology,* **52**(5):696-704.

Hopkin, SP (1997) Biology of the springtails (Insecta: Collembola). Oxford University Press.

Luque C, Legal L, Winterton P,Mariano NA, Gers C (2011) Illustration of the structure of arthropod assemblages (Collembola and Lepidoptera) in different forest types:An example in the French Pyrenees. *Diversity,* **3**: 693-711.

Maddison, WP, Maddison DR (2011) Mesquite: a modular system for evolutionary analysis. Version 2.72 <http://mesquiteproject.org>.

Maynard EA (1951) A Monograph of the Collembola or springtail insects of New York State. Comstock Publishing Co. Ithaca, New York.

McGaughran A, Torricelli1 G, Carapelli A, Frati F, Stevens MI, Convey P, Hogg ID (2010) Contrasting phylogeographical patterns for springtails reflect different evolutionary histories between the Antarctic Peninsula and continental Antarctica. *Journal of Biogeography,* **37**:107-119.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in Proceedings of the gateway Computing Environments Workshop (GCE), 14 NOV 2010, New Orleans, LA, pp.1-8.

Packard, A.S. (1873) Synopsis of the Thysanura of Essex County, Mass., with descriptions of a few extralimital forms. *Annual Report of the Trustees of the Peabody Academy of Science* **1873**:23-51.

Porco, D., Decaëns, T., Deharveng, L., James, S.W., Skarzyinski, D., Erséus, C., Butt, K.R., Richard, B. and Hebert, P.D.N. (2013) Biological invasions in soil: DNA barcoding as a monitoring tool in a multiple taxa survey targeting European earthworms and springtails in North America. *Biological Invasions,* **15:**899-910 (Supplemental data, Table S1).

Rambaut A, Suchard M, Drummond AJ (2013) Tracer v1.6, available from http://tree.bio.ed.ac.uk/software/tracer/.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3:Bayesian phylogenetic inference under mixed models. *Bioinformatics,* **19**(22):1572-1574.

Simon C, Frati F, Beckenbach AT, 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):651-701.

Stach, J, (1960) The apterygotan fauna of Poland in relation to the world fauna of this group of insects. Tribe: Orchesellini. Pastwowe Wydawnictwo Naukowe, Kraków.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**(21): 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web Servers. *Systematic Biology,* **57**(5):758-771.

Swofford, DL (2001) PAUP*: Phylogenetic analysis using Parsimony (and Other Methods). Version 4.0b10. Sinauer, Sunderland, MA.

Timmermanns MJTN, Ellers J, Mariën J, Verhoef SC, Ferwerda EB, van Straalen NM (2005) Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular Ecology,* **14**:2017-2024.

Van Straalen, N.M. (1989) Production and biomass turnover in two populations of forest floor Collembola. *Netherlands Journal of Zoology*, **39**:156–168.

CONCLUSION

Even after the work done in this project, there remains a great deal of research to be done on the genus *Orchesella*. The disconnect between dorsal patterning and relatedness in the *hexfasciata-flora* clade could prove challenging for future investigations into strongly supported morphological species descriptions. While the Texas clade provides the opportunity to expand collecting west of the Mississippi to flesh out species relatedness. Finally, the work on *O. celsa sensu latu*, shows the value of combining molecular and morphological techniques to discover new species within the existing framework.

The importance of understanding the biodiversity of the soil ecology cannot be understated. This is where the productive and fertile substrates that plants depend on are formed, and which our modern agriculture depends on. By understanding and identifying the organisms that exist in this soil formation process, we can better understand how to preserve them for continued plant health.

VITA

Born in 1980, Catherine L. Smith has been fascinated with insects from an early age. As an undergraduate at the University of Wisconsin-Madison, she earned a Bachelor of Science in Entomology and a Bachelor of Science in Agronomy. As a student hourly, she gained genetic experience taking care of a butterfly colony for an Evolutionary Development lab. That same lab trained her as a lab tech after graduation.

She was hired by the United States Department of Agriculture-Agricultural Research Service as a Biological Laboratory Technician in Stoneville, MS where she kept numerous insect colonies and assisted in research on insect molecular genetics. After six years in the Mississippi Delta, she left the USDA and worked in the Pack Department at an Amazon warehouse until she was accepted into the Entomology and Plant Pathology Graduate Program at the University of Tennessee-Knoxville.