Cryptic east-west divergence and molecular diagnostics for two species of silver flies (Diptera: Chamaemyiidae: Leucopis) from North America being evaluated for biological control of hemlock woolly adelgid

Nathan P. Havilla⁎, Stephen D. Gaimari, Adalgisa Caccone

⁎ Corresponding author at: USDA Forest Service, 51 Mill Pond Rd., Hamden, CT 06514, USA.
E-mail address: nphavill@fs.fed.us (N.P. Havill).

https://doi.org/10.1016/j.biocontrol.2018.02.004
Received 15 June 2017; Received in revised form 12 December 2017; Accepted 2 February 2018
Available online 03 February 2018
1049-9644/ Published by Elsevier Inc.
1. Introduction

Identification of cryptic diversity in biological control agents can avoid the wasted effort of releasing genotypes that are not effective controls, and can prevent unintended non-target impacts (Andersen and Wagner, 2016). Phylogenetic studies of biological control agents can also identify previously overlooked beneficial genotypes. This could be the situation for two species of silver fly (Diptera: Chamaemyiidae) from the Pacific Northwest, Leucopis argenticollis Zetterstedt and L. pini-perda Malloch. These two species are being evaluated as biological control agents of the hemlock woolly adelgid, Adelges tsugae Annand (Ross et al., 2011), a pest of eastern hemlock (Tsuga canadensis (L.) Carrière) and Carolina hemlock (Tsuga caroliniana Engelmann) in the eastern United States. Adelges tsugae is a complex of divergent lineages in Japan, China, Taiwan, and western North America (Havill et al., 2016). Southern Japan was found to be the source of the introduction to the eastern United States (Havill et al., 2006; Havill et al., 2016), while the lineage in western North America was found to be endemic to that region. This finding intensified the evaluation of predators, including these Leucopis species, from the West (defined hereafter as North America west of the Great Plains) as biological control agents to release in the East (defined hereafter as North America east of the Great Plains) (Havill et al., 2011).

Silver flies feed as larvae on sternorrhynchous Hemiptera such as aphids, adelgids, scale insects, and mealybugs (Gaimari, 2010). Some species specialize on adelgids, and have been released as biological control agents against the balsam woolly adelgid, Adelges piceae (Ratzburg) in North America (Schooley et al., 1981), and pine adelgids (Pineus spp.) in Hawaii (Greathead, 1995), east Africa (Mills, 1990), Chile (Zúñiga, 1985), and New Zealand (Zondag and Nuttall, 1989). Leucopis argenticollis is a widespread species, reported across North America, Europe, Japan, India, and Russia (McAlpine and Tanasijtshuk, 1972). In western North America, it is reported to feed on Adelges tsugae on Tsuga heterophylla (Rafinesque) Sargent (Kohler et al., 2008) and Pineus sp. on Pinus contorta Douglas ex Loudon (McAlpine and Tanasijtshuk, 1972). In eastern North America, it is reported to feed on Pineus strobi (Harrig) on Pinus strobus L., Pineus sp. on Pinus sylvestris L., Pineus pineoides (Cholodkovsky) on Picea sp., Pineus similis (Gillette) and Pineus sp. nr coloradensis (Gillette) on Picea glauca (Moench) Voss, and Adelges piceae on Abies balsamea (L.) Miller (McAlpine and Tanasijtshuk, 1972). Leucopis piniperda is distributed across the northern United States and Canada (Tanasijtshuk, 2002). In the West, it has been reported on Adelges tsugae on Tsuga heterophylla (Kohler et al., 2008), Adelges piceae on Abies spp., Adelges sp. on Picea glauca, and in the East on Pineus similis on Picea glauca, and Pineus strobis on Pinus strobus (Tanasijtshuk, 2002).

Both of these Leucopis species have been the subject of biological control evaluations since Kohler et al. (2008) found that they were among the most abundant predators on A. tsugae in the Pacific Northwest. Collectively, these species had peaks in abundance coinciding with the two periods of adelgid oviposition (Kohler et al., 2008; Kohler et al., 2016; Grubin et al., 2011), and there was a strong positive correlation between these Leucopis species and adelgid abundance (Kohler et al., 2008). No-choice laboratory feeding assays with Leucopis larvae collected from adelgid-infested Tsuga heterophylla found that survival was higher on hemlock woolly adelgid than on four other alternate adelgid prey species (Grubin et al., 2011). Some Leucopis larvae developed to the adult stage on A. tsugae as well as on each of the other adelgid species, suggesting that western Leucopis spp. prefer hemlock adelgids, but may not be strictly specialized. Motley et al. (2017) reported that both Leucopis species can feed and develop on the Japanese A. tsugae lineage that was introduced to the eastern United States as successfully as they did on the western North American adelgid lineage. They also completed caged field releases on adelgid-infested Tsuga canadensis in the eastern United States (New York and Tennessee) and found that both of the Leucopis species could complete development and reproduce on this prey in field conditions.

Here, we report phylogenetic analyses using DNA sequence data from one mitochondrial and two nuclear genes comparing 606 specimens of Leucopis argenticollis and L. piniperda collected across North America on different adelgid prey species to examine variation associated with geographic origin and prey association. We also developed and tested a PCR-RFLP assay based on the mitochondrial cytochrome oxidase subunit I gene (COI) gene as an inexpensive genotyping method to distinguish the two species and the major lineages within each species. These assays will assist with documenting establishment and impact following their release for biological control, due to the difficulty of morphologically distinguishing the different species and lineages in the immature stages.

2. Methods

A total of 296 L. argenticollis and 310 L. piniperda specimens were analyzed for this study (Supplementary Materials, Table S1). This included field collected samples, plus 108 L. argenticollis and 116 L. pini-perda that were first-generation reared samples recovered from a field enclosure study (Motley et al., 2017). All of the samples of both species from west of the Great Plains (Washington and Idaho) were from A. tsugae on T. heterophylla, with the exception of one L. piniperda from Saskatchewan from Pineus strobi on Picea sp., and one sample of each Leucopis species collected in Colorado from Pineus coloradensis on Picea engelmannii. The L. argenticollis samples from east of the Great Plains (Connecticut, Pennsylvania, and Minnesota) were all from Pineus strobi on Pinus strobus. The L. piniperda samples from east of the Great Plains (Connecticut, New York, Pennsylvania, Maryland, Virginia, and West Virginia) were from Pineus strobi on Pinus strobus, Pineus boereri Annand on Pinus densiflora Siebold & Zucc., Pineus pini (Macquart) on Pinus sylvestris, or Pineus sp. on Picea glauca.

DNA was extracted using the DNA IQ Extraction Kit (Promega, Madison, WI, USA) or the Mag-Bind Blood & Tissue kit (Omega Bio-Tek, Norcross, GA, USA). For adult flies, the thorax was removed and ground with a pestle for extraction, and the rest of the fly was retained as a voucher. For larvae or puparia, a small slit was cut into the side of the specimen with a scalpel, then it was incubated with proteinase K for at least an hour, and spun in a microcentrifuge tube to force out the body contents for extraction. The cuticle was removed and slide-mounted as a voucher. Vouchers were deposited at the California State Collection of Arthropods (CSCA) or the Yale Peabody Museum (YPM) (Supplementary Materials, Table S1).

For all samples, the standard DNA barcoding region on the 5’ end of the mitochondrial COI gene was amplified using the primers LepF1 and LepR1 (Hebert et al., 2004). Two nuclear genes were also sequenced for a subset of seven to fifteen samples per region (eastern and western North America) in each species. A portion of the nuclear triose phosphate isomerase (TPI) gene was amplified using primers 111Fb and Cham275R (GCCCAAGCGGGCTYGTAGGC), the latter primer modified from 275R (Bertone et al., 2008). A portion of the nuclear carbamoylphosphate synthetase domain of the CAD (rudimentary) gene was amplified using the primers 787F and Cham1098R (TTTAGTACGTGCC TCCCAT), the latter primer modified from 1098R (Mouton and Wiegmann, 2004). Sequencing reactions were performed using the BigDye Terminator kit (Applied Biosystems, Foster City, CA, USA) and analyzed on an Applied Biosystems 3730 automated sequencer at the DNA Analysis Facility on Science Hill at Yale University (New Haven, CT, USA).

Sequences from each of the three genes were aligned using Geneious 10.0.5 (Kearse et al., 2012). A network of COI haplotypes was reconstructed using the statistical parsimony method of Templeton et al. (1992), using the software TCS 1.21 (Clement et al., 2000) with a 90% connection limit. Phylogenetic analyses were performed for each gene separately and for a combined, concatenated data set using MrBayes 3.2.6 (Ronquist and Huelsenbeck, 2003) with default priors, four
incrementally heated Markov chains, and two concurrent runs of 1,000,000 generations sampled every 1000 generations. The first 25% trees were discarded. Separate, unlinked partitions were used with substitution models GTR + I + G for COI and CAD, and HKY + I + G for TPI, as determined using PartitionFinder 2.1.1 (Lanfear et al., 2012). Lipoleucopis precox de Meijere and Anchioleucopis geniculata (Zetterstedt) were used as outgroups. A specimen of Leucopis arcticae Aldrich was included as a closely related species found only in western North America that is also a predator of adelgids, but which is not currently being evaluated as a potential biological control agent. Vouchers for these three samples are deposited at the California State Collection of Arthropods (accession numbers 09E474, 09E465, and 09E481, respectively). GenBank accession numbers for the sequences generated for this study are: COI: MF468329 to MF468936; TPI: MF468937 to MF468984; CAD: MF468985 to MF469032 (Supplementary Materials, Table 1).

PCR-RFLP assays were designed using the COI sequences generated in this study to distinguish L. argenticollis from L. piniperda and the major lineages within each species. The alignment of 606 COI sequences were examined in Geneious, and primers were designed using Primer3 (Untergasser et al., 2012) in conserved regions to amplify shorter sections of the gene that flanked potentially useful restriction sites. Primers LeucoShortF1 (TCAATTATTATAAGTATTTGAACG) with LeucoShortR (GGTATTCGATCAAAATTAATTCC) amplify a 393 base pair (bp) fragment and LeucoShortF2 (AGTTTTTGAATTTACCTCCC CCATCA) with LeucoShortR amplify a 245 bp fragment (Fig. 1). The “Find Restriction Sites” feature in Geneious was used to map enzyme cleavage sites for that were fixed in each group. Enzyme Rsal was tested to distinguish the species, BfaI was tested to distinguish eastern versus western L. argenticollis, and FauI was tested to distinguish eastern versus western L. piniperda. Samples representing all unique COI haplotypes found in this study (N = 84) were tested to confirm that the banding patterns produced by these enzymes were distinguishable and fixed within groups. DNA was amplified using the newly-designed primers in 30 μL reactions. Since BfaI does not efficiently cleave unpurified PCR products, the products were purified using Sera-Mag SpeedBeads (GE Healthcare Life Sciences, Pittsburgh, PA, USA) prior to digestion. Ten microliters of purified PCR product were included in 15 μL RFLP reactions containing 0.5 μL of restriction enzyme and 1.5 μL of Cutsmart Buffer (New England BioLabs, Ipswich, MA, USA). Reactions were incubated for one hour at 55 °C for Faul, and at 37 °C for the other enzymes, then heat-inactivated at 80 °C for 20 min. Reactions were also performed with Rsal, BfaI, and Faul to distinguish all four groups. For these reactions, Rsal and BfaI were incubated together at 37 °C for one hour, then Faul was added and incubated at 55 °C for an additional hour. Banding patterns were imaged following gel electrophoresis in 1.5% agarose gels.

3. Results

Amplification of COI, TPI, and CAD resulted in sequences that were 658, 438, and 798 bp long, respectively. There were no insertions or deletions in the alignments of all three genes.

The COI network of 606 COI sequences (Fig. 2) had 84 haplotypes arranged in four clusters (Clusters 1–4) that were not connected using a 90% connection limit of 16 mutation steps. Cluster 1 included 274 L. argenticollis individuals collected from Washington, Idaho, and Colorado assigned to 33 haplotypes. Cluster 2 included 21 L. argenticollis individuals collected from Connecticut, Pennsylvania, and Minnesota,
assigned to 10 haplotypes. Cluster 3 included 267 L. piniperda individuals collected from Washington, Saskatchewan, Colorado, and Minnesota assigned to 34 haplotypes. Cluster 4 included 43 L. piniperda individuals from Connecticut, New York, Pennsylvania, Maryland, Virginia, and West Virginia, assigned to seven haplotypes. Thus, western and eastern samples of L. argenticollis were in Cluster 1 and 2, respectively, and western and eastern samples of L. piniperda were in Clusters 3 and 4, respectively. The notable exception to this pattern was a larva of L. piniperda collected in Minnesota from Picea glauca which was in Cluster 3 with the western samples.

The mean pairwise COI p-distance between L. argenticollis and L. piniperda samples was 0.0821 (min: 0.0600; max: 0.0946), between the L. argenticollis western and eastern clusters was 0.0485 (min: 0.0381; max: 0.0548), and between the L. piniperda western and eastern clusters was 0.0497 (min: 0.0397; max: 0.0612). The mean pairwise p-distance for samples within the western L. argenticollis cluster was 0.0018, within the eastern L. argenticollis cluster was 0.0044, within the western L. piniperda cluster was 0.0061, and within the eastern L. piniperda cluster was 0.0006. For TPI, the mean pairwise P-distance between the species was 0.057 (min: 0.0456; max: 0.0706), within L. argenticollis was 0.014 (min: 0.000; max: 0.023), and within L. piniperda was 0.013 (min: 0.000; max: 0.021). For CAD, the mean pairwise P-distance between the species 0.093 (min: 0.0866; max: 0.102), within L. argenticollis was 0.015 (min: 0.000; max: 0.038) and within L. piniperda was 0.0005 (min: 0.000; max: 0.010).

The concatenated data set for phylogenetic analysis using all three genes was 1894 bp long and included 28 L. argenticollis samples (15 eastern, 13 western), and 17 L. piniperda samples (7 eastern, 10 western). The 50% majority rule consensus Bayesian tree from the combined data set is shown in Fig. 3. The sample of L. atricapillus was sister to L. piniperda, but with low posterior probability (68%). As in the COI haplotype network, there were separate eastern and western clades within each species, with the exception of the L. piniperda sample (CSCA#09E475) collected from Picea in Minnesota. Analysis of each gene separately yielded phylogenies consistent with each other but with lower clade support than the consensus tree (Supplementary Material, Figs. S1–S3).

The PCR-RFLP assays using short fragments of the COI gene successfully distinguished L. argenticollis from L. piniperda, and eastern versus western lineages within each species (Figs. 3 and 4). For the 245 bp PCR product amplified with primers LeucoShort2 and LeucoShortR, the enzyme Rsal can distinguish the species because digestion results in fragments of 176 and 69 bp for L. argenticollis and does not cleave L. piniperda (Fig. 4A). Digestion with BfaI can distinguish eastern from western L. argenticollis by producing fragments for the former of 177 and 70 bp and not cleaving the other groups (Fig. 4C). Digestion with FauI can distinguish eastern from western L. piniperda by producing fragments for the former of 136 and 111 bp, and not cleaving the other groups (Fig. 4D). Digestion of the 245 bp PCR product with all three enzymes resulted in fragments of 108, 70, and 69 bp for eastern L. argenticollis, 176 and 69 bp for western L. argenticollis, 136 and 111 for eastern L. piniperda, and did not cleave western L. piniperda, thus distinguishing all four groups (Fig. 4E). The 393 bp PCR product, amplified with primers LeucoShort1 and LeucoShortR, incorporated an additional cut site for Rsal in western L. piniperda compared to the 245 bp fragment, such that digestion resulted in fragments of 176, 121, and 96 bp for all L. argenticollis, fragments of 297 and 96 bp for western L. piniperda, and did not cleave eastern L. piniperda (Fig. 4B). Digestion of the 393 bp PCR product with FauI would also effectively distinguish eastern and western L. piniperda using the same cut site as the 245 bp PCR product (Fig. 1; gel image not shown). Digestion of the 393 bp PCR product with BfaI cannot reliably distinguish eastern from western L. argenticollis because there is an additional cut site that is variable in western L. piniperda that would produce a banding pattern similar to eastern L. argenticollis (Fig. 1).

4. Discussion

This study revealed divergent lineages of L. argenticollis and L. piniperda feeding on different adelgid prey species in eastern versus western North America. Evidence for distinct lineages include COI sequence divergence between interspecific lineages of more than 3.81% within L. argenticollis and 3.97% within L. piniperda, and confirmation with phylogenetic analysis using sequence data from COI plus two nuclear genes that also shows the same pattern. Inclusion of the nuclear genes ruled out past hybridization events as the cause of divergence in the maternally inherited mitochondrial COI gene because they all showed consistent topologies. A single sample of L. piniperda collected in the East (Minnesota) that was a member of the western lineage was an exception to the general pattern. In contrast, all of the L. argenticollis samples collected in Minnesota clustered with samples from the East. Western hemlock and the western lineage of hemlock woolly adelgid do not extend east to the range of eastern white pine and pine bark adelgid, so the geographic ranges of the different prey species do not overlap. Together, this could indicate that the range of western L. piniperda extends west to Minnesota while the eastern lineage does not reach that far east, or that eastern and western lineages of L. piniperda co-exist in central North America. Additional sampling would clarify this.

Besides molecular evidence, there is also some morphological evidence for regional variation within one of the species, L. piniperda. We found that the color of pruinosity (dusty-looking coating) on the lunule, face, and parafacial are variable in this species relative to geography (Fig. 5). Specimens from eastern North America (New Brunswick, Quebec, Connecticut, Illinois, Maryland, Minnesota, North Carolina, Ohio, Rhode Island, and Virginia) all have uniformly silvery grey pruinosity (Fig. 5C), while those in western North America (Alberta, British Columbia, California, Colorado, Idaho, New Mexico, Oregon, Washington, and Wyoming) all have uniformly brown to black pruinosity (Fig. 5A). Only specimens from Ontario have a true mixed state, with the lunule and face being brown to black pruinose but the parafacial silvery grey pruinose (Fig. 5B). For L. argenticollis, although McAlpine and Tanasijtshuk (1972) noted variation among specimens from different localities and prey or host plants, there was no distinct geographical pattern, with variation between close sites being often as great as between distant sites. Further examination of morphological variation within the species is necessary to determine if there are characters that distinguish the lineages that we describe here.

While our results show a clear pattern of genetic differentiation in these species, more work will be needed to explain how this evolved. Unfortunately, the roles of geographic isolation versus prey specialization in promoting and maintaining this pattern may be confounded in these species because the ranges of the different adelgid prey and host tree species do not overlap. In western North America, the lineage of hemlock woolly adelgid that feeds on the two endemic western hemlock species may have arrived there prior to the last glacial period 40,000 years ago, while in the east, a different lineage of hemlock adelgid that was introduced from Japan in recent times feeds on the two endemic eastern hemlock species (Havill et al., 2016). Leucopsis argenticollis and L. piniperda can be found in high abundance on hemlock woolly adelgids in the West (Kohler et al., 2008). In contrast, only one (Wallace and Hain, 2000) of the many surveyed adelgid-infested hemlock trees for natural enemies in the East reported Leucopsis (e.g. Montgomery and Lyon, 1996, Mauel et al., 2008, Jones et al., 2014). In that study, the species of Leucopsis is not indicated, as only larvae were reported. Conversely, the pine bark adelgid, Piceus strobi, and its host, Pinus strobus, are native to eastern North America, where the eastern lineages of Leucopsis readily feed on them, but we have not recovered these Leucopsis species on related pine adelgids in the West, to date. It therefore seems likely that the western Leucopsis lineages have evolved to feed on hemlock adelgids while the eastern lineages have evolved separately to feed on pine adelgids in response to different resources available in the different regions.
Since there are no morphological characters to separate the eastern versus western lineages within each species (with the exception of regional variation in *L. piniperda* adults), the PCR-RFLP assays that we developed will be valuable for distinguishing each group following release of western *Leucopis* in the East. It should be noted however, that since the assays use a fragment of a mitochondrial gene, which is maternally inherited, they would not be able to detect hybrids between eastern and western lineages of each species. Future work should address the prospect of hybridization between eastern and western lineages following field release, and how this might affect their biological control efficacy. Assessing this hybridization will require developing nuclear markers such as microsatellites or single nucleotide polymorphisms (SNPs). There were multiple SNPs that were fixed within lineages in the two nuclear genes that we examined, but additional samples would need to be examined to determine whether they would be reliable diagnostic markers.

There is a recent example of another biological control agent used against hemlock woolly adelgid that has many similarities to these *Leucopis* species. A western North American adelgid predator, *Laricobius nigrinus* Fender (Coleoptera: Derodontidae), was released in the East for hemlock woolly adelgid biological control where it subsequently hybridized with a closely related eastern species, *L. rubidus* LeConte (Havill et al., 2012). These two *Laricobius* species were found in laboratory assays to have different feeding preference (Zilahi-Balogh et al., 2002; Zilahi-Balogh et al., 2005) and are preferentially attracted to different prey host trees (Wallin et al., 2011;Arsenault et al., 2015), such that *L. nigrinus* prefers hemlock woolly adelgid and hemlock while *L. rubidus* prefers pine bark adelgids and pine. Several follow-up field studies assessed the hybridization rate between these species (Jones et al., 2014, Fischer et al., 2015, Mayfield et al., 2015, Wiggins et al., 2016). Together, these studies report a stable rate of hybridization of 11–13% in different locations and over time, with both species maintaining their genetic integrity in release sites. The different prey and habitat preferences may be what is maintaining isolation between these *Laricobius* species. Given an apparently similar pattern of prey preferences in eastern versus western lineages of these *Leucopis* species, we might expect to see a similar result if they can successfully hybridize.

In conclusion, DNA sequence data from mitochondrial and nuclear genes uncovered previously unrecognized cryptic North American diversity in both *L. argenticollis* and *L. piniperda* that appears to be strongly associated with geography and prey preference. It is remarkable that both of these species show this same unexpected pattern. Given the widespread occurrence and abundance of the hemlock-associated *Leucopis* genotypes in the Pacific Northwest (Kohler et al., 2008), and the absence of these genotypes in the East, the western genotypes could be strong prospects for enhancing biological control of hemlock woolly adelgid. Future research directions could use additional markers and wider sampling to reconstruct the ecological and biogeographic forces of...
that led to this pattern. The RFLP assays that we developed are inexpensive and reliable ways to track establishment and impact of these species following release, but will need to be complemented with additional markers to assess potential hybridization between eastern and western lineages.

Acknowledgments

We thank DeAdra Newman and Joanne Klein for help in the laboratory, and Jana Albers, Arielle Arsenault, Rachel Brooks, Melissa Fisher, Sarah Grubin, Amalia Havill, Melody Keena, Glenn Kohler, Sandy Liebhold, Richard McDonald, Michael Montgomery, Maya Nehme, Darrell Ross, Robert Tiplady, Kimberly Wallin, and Mark Whitmore for helping to collect samples. This research was supported by USDA Forest Service grants #09-CA-11420004-156 and 12-JV-11242303-096. We would also like to thank Andrea Gloria-Soria, John Podgwaite, and three anonymous reviewers whose suggestions greatly improved this manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biocontrol.2018.02.004.

Fig. 4. Agarose gel electrophoresis of PCR-RFLP reactions of cytochrome oxidase subunit I (COI) 245 bp or 393 bp PCR products for representative L. argenticollis and L. piniperda. (A) RsaI-245 bp, (B) RsaI-393 bp, (C) BfaI, (D) FauI, (E) RsaI + BfaI + FauI. The first lane in each gel is 100 bp ladder (New England Biolabs), and the second lane is undigested PCR product.

Fig. 5. Frontal views of the heads of Leucopis piniperda from western North America (A) and eastern North America (C) displaying the differences in pruinosity on the face, lunule and parafacial, and the mixed state in specimens from Ontario, Canada (B). Abbreviations: f = face, l = lunule, p = parafacial.

References

Havill, N., Montgomery, M., Keena, M., 2011. Chapter 1: Hemlock woolly adelgid and its...


