DISSERTATION

POPULATION GENETICS PRIOR TO BIOLOGICAL CONTROL: *CEUTORHYNCHUS* WEEVILS PROPOSED FOR MANAGING GARLIC MUSTARD

Submitted by

Steven J. Rauth

Department of Bioagricultural Sciences and Pest Management

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, CO 80523

Summer 2009

COLORADO STATE UNIVERSITY

June 22, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY STEVEN J. RAUTH ENTITLED "POPULATION GENETICS PRIOR TO BIOLOGICAL CONTROL: *CEUTORHYNCHUS* WEEVILS PROPOSED FOR MANAGING GARLIC MUSTARD" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Michael Antolin

Lou Bjostad

Andrew Norton

Advisor: Ruth Hufbauer

Department Head: Thomas Holtzer

ABSTRACT OF DISSERTATION

POPULATION GENETICS PRIOR TO BIOLGICAL CONTROL: *CEUTORHYNCHUS* WEEVILS PROPOSED FOR MANAGING GARLIC MUSTARD

Nearly 40 years ago biological control scientists began discussing the potential importance of genetic variation in natural enemies to the successful control of their target. Despite that, research in this area is only recently increasing. In particular, few studies have explored the population genetic structure of potential biological control agents prior to release. This lack of information could be particularly critical in biological control of weeds because genetic and ecological variation among populations of herbivorous insects is common. To evaluate the hypothesized importance of genetic variation in natural enemies, first that variation must be measured.

To that end, I studied the population genetic structure of three closely related weevil species, *Ceutorhynchus alliariae*, *C. roberti*, and *C. scrobicollis*, all being considered for the biological control of garlic mustard, *Alliaria petiolata*, in North America.

My first objective was to develop a polymerase chain reaction – restriction fragment length polymorphism (PCR-RFLP) assay which could be used to identify the morphologically indistinguishable larvae. This assay was developed for use in the analysis of population genetic structure of the three species and to aid biological control

scientists in evaluating host-specificity test results where larval development was incomplete or adults failed to emerge. The resulting assay provides a relatively fast and inexpensive means of identifying otherwise indistinguishable larvae.

My second objective was to study the population genetic structure of *C. scrobicollis*, the agent most likely to be released, to evaluate whether the areas where individuals were being collected for host-specificity testing (surrounding Berlin, Germany) consisted of one or more populations, to estimate the numbers of individuals needed during host-specificity testing and later introduction to adequately represent the diversity of the population, and to evaluate dispersal potential. Results suggest that *C. scrobicollis* in the area of Berlin, Germany constitute a network of subpopulations with low but significant differentiation among sites and movement of individuals between sites. I estimated that the number of individuals that would need to be sampled to capture 90% or 99% of the genetic diversity in the Berlin area was 10 and 27, respectively. The estimated average dispersal distance based on assignment tests for *C. scrobicollis* was 28 km.

My third objective was to compare the differences in population genetic structure between *C. alliariae* and *C. roberti* to determine whether differences in genetic diversity or dispersal potential might aid in prioritizing one species over the other. These two species have similar life histories, distributions, and effects on garlic mustard. Results showed that, over a comparable region in central Europe, total gene diversity was significantly higher in *C. roberti*, though the difference between the two species was relatively small. Assignment tests suggest there is substantial gene flow among sites for

both species. Overall, the results were similar for both species, and I recommend prioritizing based on biological or methodological attributes.

Steven J. Rauth
Department of Bioagricultural Sciences and Pest Management
Colorado State University
Fort Collins, CO 80523
Summer 2009

ACKNOWLEDGEMENTS

There are many people who have influenced my life, research, and experience during my graduate career. I would like to first and foremost thank my advisor, Ruth Hufbauer, for all of the guidance and support she has provided over the years. Patience is one of her many virtues. I would also like to thank my committee members, Mike Antolin, Lou Bjostad, and Andrew Norton. They have taught me much about creative and critical thinking, and about people and the world as a whole.

There are a number of people who have contributed directly to this work, and for their help and collaboration I am very grateful: Hariet Hinz and Esther Gerber were very helpful in introducing me to this system, facilitating collections, and providing feedback on manuscripts. Oliver Bossdorf, Alecu Diaconu, Heinz Mueller-Schaerer, Norbert Maczey, Rene Sforza, and David Kikodze all helped in one way or another in collecting specimens throughout Europe.

There are also a number of friends I must thank for, at times, their critical advice, but mostly for the words of encouragement and support in helping me survive through graduate school: Scott Merrill, Amy Blair, Robin Marrs, Dan MacKinnon, Chrissy Alba-Lynn, Boris Kondratieff, and Bill Jacobi. Special thanks are also due to my parents, Joe and Carol Rauth, for being there through all the ups and downs.

Lastly, I would like to thank the Colorado Agricultural Experiment Station for funding a portion of this research, and the Departments of Biology and Bioagricultural Sciences and Pest Management for the teaching assistantships which put food on the table.

TABLE OF CONTENTS

CHAPTER 1

PCR-RFLP assay for discerning three weevil stem feeders (*Ceutorhynchus* spp.) (Col.: Curculionidae) on garlic mustard (*Alliaria petiolata*).[1]

[1]This chapter has been accepted with minor revisions to Biological Control Science and Technology as: Rauth, S.J. and Hufbauer, R.A. PCR-RFLP assays for discerning three weevil stem feeders (*Ceutorhynchus* spp.) (Col.: Curculionidae) on garlic mustard (*Alliaria petiolata*).

INTRODUCTION

The introduction of the polymerase chain reaction (PCR) (Mullis et al., 1986; Saiki et al., 1988; Saiki et al., 1985) paved the way for the development of a number of diagnostic applications in a wide variety of fields. From the food and agricultural industry (Lopez et al., 2008) to the forensic (Wells and Stevens, 2008) and biological sciences (Newton and Graham, 1997), PCR-based techniques are being used to diagnose or detect pathogens and diseases, fingerprint genomic material for investigative purposes, and identify organisms to determine taxonomic and phylogenetic relationships. For species identification, examples of common techniques used include designing species-specific primers for PCR amplification (e.g. Erlandson and Gariepy 2005), sequencing of specific genes or gene regions (e.g. DNA barcoding (Hebert et al., 2003a)), or analyzing restriction fragment length polymorphisms (RFLPs) of amplified regions (e.g. McKern and Szalanski 2007). Mutations and genomic similarity among congeners can result in problematic identifications for any of the above techniques, none-the-less, molecular identification of species can be necessary in situations where morphological identification is difficult or impossible.

Species-level identification of natural enemies is critical at all stages of biological control. Preliminary identification of candidate agents, which can take several years (Heard and Pettit, 2005), must often be performed by systematists specializing in the particular taxon. Potential agents are often closely related, or have close relatives among more generalist taxa, and differentiating between species may not be possible for some of the life stages. For example, insects often are identified to species level using characteristics of the adult stage because immature stages frequently lack morphologically distinguishing features. This can be problematic during host-specificity testing of candidate biological control agents as accurate identification is necessary to elucidate relationships between candidate agents and their hosts (Gariepy et al., 2008) and to guard against possible contamination by closely related species in laboratory or field experiments (Unruh and Woolley, 1999). Additionally, if only partial larval development occurs on alternative test plants or adult emergence rates are low, it may be difficult to confirm species identity without the morphological characteristics of adults. In such cases molecular diagnostic tools are necessary in species identification.

Here we report the development of a PCR-based assay to differentiate between three *Ceutorhynchus* weevils being considered for the biological control of garlic mustard (*Alliaria petiolata* (Bieb.) Cavara & Grande) in North America. *Ceutorhynchus alliariae* Brisout and *C. roberti* Gyllenhal feed primarily in the stem of garlic mustard while *C. scrobicollis* Nerensheimer and Wagner feeds primarily in the root crown. All three species are distributed throughout Eurasia where their host plant is native. Field collections of adult insects have shown that *C. alliariae* and *C.*

roberti co-occur at several sites in northwestern Switzerland and southwestern Germany, and that *C. alliariae* and *C. scrobicollis* co-occur at several sites in northeastern Germany (Hinz and Gerber, 2000). Additionally, all three species co-occur at two field sites (one near Berlin, Germany and one near Iaşi, Romania) (Rauth, unpublished). Caged competition experiments and rearing of field collected larvae have shown that *C. alliariae* and *C. roberti* can occur in the same host individual (Hinz and Gerber, 2001). Adults of the three species can be distinguished by differences in morphology. *Ceutorhynchus alliariae* differs from *C. roberti* and *C. scrobicollis* in having reddish tarsal segments (Hinz and Gerber, 2000). The differences between *C. roberti* and *C. scrobicollis* are less apparent and require examination under a dissecting microscope. The pronotal surface structure of *C. scrobicollis* is coarser than that of *C. roberti*, and the space between elytral intervals is larger in *C. scrobicollis* than in *C. roberti* (Dieckmann, 1972). Larvae of the three species are morphologically indistinguishable (Hinz and Gerber, 2000).

To facilitate identification of larvae of these three candidate biological control agents we developed a polymerase chain reaction – restriction fragment length polymorphism (PCR-RFLP) assay. We targeted the mitochondrial DNA (mtDNA) cytochrome oxidase subunit I (COI) for development of this assay as this region has been widely studied and shown to reliably discriminate among closely related species in most animal phyla (Hebert et al., 2003b). This assay is relatively fast and inexpensive and can be used to identify specimens when larval development is incomplete, adult emergence is problematic, or when there is concern over contamination. We evaluated the PCR-RFLP assay by comparing banding patterns

and DNA sequences from field-collected larvae of unknown identity to those obtained from positively identified adults. We focused our work on identifying an RFLP protocol of identification as opposed to a barcoding based methods, as the cost of identifying larvae by means of restriction digest is approximately $1/10^{th}$ to $1/30^{th}$ that of sequencing.

METHODS

PCR-RFLP assay development

Larvae of *Ceutorhynchus scrobicollis, C. roberti* and *C. alliariae* were collected from central European field populations and reared to adults by Drs. H. Hinz and E. Gerber (CABI Bioscience, Delemont, Switzerland), who have been performing host-specificity and efficacy tests with these weevils on *Alliaria petiolata*. Voucher specimens for *C. scrobicollis*, identified by Dr. B. Korotyaev, are housed at the Zoological Institute, St. Petersburg Russia. Voucher specimens of *C. roberti* and *C. alliariae*, identified by Lutz Behne, Deutsches Entomologisches Institut, Eberswalde, Germany, are maintained in the CABI Bioscience collection in Delemont, Switzerland. Genomic DNA was extracted from two adult individuals of each species using Qiagen DNeasy Tissue Kits (QIAGEN®, Germantown, MD, USA). A 974bp section of mtDNA encompassing approximately 850 bp of COI, tRNALeu, and a short section of COII was PCR amplified using primers C1-J-2183 (5'-CAACATTTATTTTGATTTTTTGG-3') (Simon et al., 1994) and A3113 (5'-

TGTTCT/CATTAATGGA/GGAAGC/TTCTAT-3' (Integrated DNA Technologies Inc., Coralville, IA, USA))(Brower, 1994). Amplifications were performed in 20μL reaction volumes of 1X PCR buffer (20mM Tris-HCL, pH 8.4, 50mM KCl), 2mM $MgCl_2$, 0.2mM dNTPs, 1 U Nova*Taq*™ Hot Start DNA polymerase (Novagen®, Madison, WI, USA), 1.8pmol of each primer, and 2μL of extracted genomic DNA. Genomic DNA concentrations ranged from approximately 4 to11 ng/μl. PCRs were run on a Hybaid™ PCRExpress thermal cycler (Hybaid Ltd., Ashford, Middlesex, UK) with an initial denaturation step of 95°C for 8min, then 36 cycles of 50sec at 95°C, 60sec at an annealing temperature of 53°C, 90sec at 72°C for polymerase extension, and a final 10 minute extension at 72°C. PCR products were cleaned in an exonuclease/shrimp alkaline phosphatase (USB® Corp., Cleveland, OH, USA) incubation step (37°C, 20 min., and then 94°C for 10 min to inactivate the enzymes).

We sequenced from both ends of the PCR product using BigDye® Terminator Cycle Sequencing (version 3.1) (Applied Biosystems Inc., Foster City, CA, USA) primed with the PCR primers. Sequencing reactions were cleaned using ethanol precipitation, and nucleotides were separated on an ABI 3100 capillary instrument (Applied Biosystems Inc.). Quality of forward and reverse sequences was checked by hand using SeqMan (Lasergene®, DNAStar Inc., Madison, WI, USA) and sequences were aligned. The use of Cleaver (Jarman, 2006), software for comparing restriction endonuclease digestion of DNA sequences, identified *Taq*αI (New England Biolabs® Inc., Ipswich, MA, USA) as a restriction enzyme suitable for distinguishing between the three species based on variation in restriction sites.

PCR products were digested with *Taq*αI in 20μL reaction volumes of 1x NEBuffer 3 (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol, pH 7.9 @ 25°C), 10 U *Taq*αI (in 10 mM Tris-HCl, 300 mM KCl, 1 mM Dithiothreitol, 1 mM EDTA, 500 μg/ml BSA, 50% Glycerol, pH 7.5 @ 25°C), and 8μL of PCR product (Conc. ranged from approx. 60 to 150 ng/μl). Digestions were incubated in a Hybaid™ PCRExpress thermal cycler for 5 hours at 65°C, followed by enzyme inactivation by 20min at 80°C. Fragments were run on a 1% agarose gel (80mL 1x TAE, 0.8g agarose, 5μl Ethidium Bromide) in 1x TAE buffer for 45 min at 75 volts.

We tested the robustness of the one-enzyme PCR-RFLP assay by screening and sequencing DNA extracted from 175 unidentified larvae field-collected from garlic mustard stems. Seventy-six larvae were collected from 15 sites throughout Germany, and 38 larvae were collected from 7 sites in Switzerland. We also obtained larvae from single sites in each of the following countries: Austria (n = 9), Bulgaria (n = 5), France (n = 10), Georgia (n = 20), Italy (n = 6), Romania (n = 8), and the United Kingdom (n = 3). Genomic DNA from larvae was extracted as described above. The COI region of these individuals was amplified and digested with *Taq*αI following the above protocols. To identify the collected larvae, banding patterns were compared to those obtained from the known adults. To verify digest-based identification, we sequenced the same COI region for all 175 field-collected larvae following the above protocol and sequences were compared with those from morphologically identified adults. Haplotype sequences that resulted in unique

banding patterns have been submitted to GenBank (accession numbers EU551167-71, EU551173-4).

Because our results showed that additional *Ceutorhynchus* (or closely related) species were present in our sample that were not differentiated from our target species by this assay (see below), and because a number of sequences of *Ceutorhynchus* species have been added to GenBank since the inception of this study, we performed another virtual digest using all available sequence data in Cleaver (Jarman, 2006). This data set included our sequences plus sequences for 26 species (78 seqs.) of *Ceutorhynchus* available on GenBank. We cross referenced these sequences with literature and field surveys of *Ceutorhynchus* species associated with garlic mustard (Hinz and Gerber, 1999). In addition to our target species, there are 14 other congeners reported to be associated with garlic mustard. Sequence data were available for 9 of these. All sequences were imported into Cleaver (Jarman, 2006) and the predicted digest patterns were analyzed for all enzymes. No single enzyme was found that would distinguish our three target species both from each other and all other *Ceutorhynchus* species. Results showed, however, that a double digestion with *Alu*I and *Taq*αI would differentiate our three target species from each other and from all other *Ceutorhynchus* species for which sequence was available, including the 9 also associated with garlic mustard.

We tested the ability of a double digestion with *Alu*I and *Taq*αI to differentiate our target species from each other and from the unknown species in our sample that were not differentiated with *Taq*αI. PCR products were digested with *Alu*I in 20μL reaction volumes of 1x NEBuffer 4 (50 mM potassium acetate,

20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM Dithiothreitol, pH 7.9 @ 25°C), 10 U *Alu*I (in 10 mM Tris-HCl, 100 mM KCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 200 μg/ml BSA, 50% Glycerol, pH 7.4 @ 25°C), and 8μL of PCR product. Digestions were incubated in a Hybaid™ PCRExpress thermal cycler for 5 hours at 37°C. Ten units of *Taq*αI were then added to reaction and digestions were incubated for 5 hours at 65°C, followed by enzyme inactivation by 20min at 80°C. Fragments were run on a 2.5% agarose gel (80mL 1x TAE, 0.8g agarose, 5μl Ethidium Bromide) in 1x TAE buffer for 2.5 hours at 75 volts. Banding patterns were then compared to those expected from virtual digestion of the sequences.

RESULTS & DISCUSSION

PCR-RFLP assay evaluation

As predicted from the sequence data of previously-identified adults, restriction digest of the mtDNA COI region with *Taq*αI allowed differentiation between the three morphologically indistinguishable larvae of *C. alliariae, C. roberti,* and *C. scrobicollis*, similarily to what was observed for adults. All *C. alliariae* larvae digested showed one of two banding patterns, all *C. roberti* larvae showed a single uncut fragment, and all *C. scrobicollis* larvae exhibited a single banding pattern (Table 1; Figure 1). All restriction digest based identifications were confirmed by

comparing larval sequences with sequences from morphologically identified adults. Fifty-seven *C. alliariae* were identified from samples collected in Germany, Switzerland, Austria, Italy, Romania, France, and the United Kingdom. Sixty *C. roberti* were identified from collections made in Germany, Switzerland, Bulgaria, and Romania. Forty-three *C. scrobicollis* were identified in field collections from Germany, Romania, and Georgia (Table 1; Figure 1).

Sequencing of the COI region of all field-collected larvae suggested that three additional taxa were collected. One of these taxa occurred only in samples from Georgia (n = 13) and exhibited a distinct banding pattern when digested with *Taq*αI (Table 1). Sequence data from these 13 individuals showed a 6.2% average difference from *C. roberti*, the most similar of the targeted taxa. The other two taxa exhibited banding patterns that matched our target species. A single individual collected from central Germany showed the same banding pattern observed for most *C. alliariae* in spite of an 11.6% difference in sequence. The PCR product of another individual collected from Bulgaria was, like for *C. roberti,* not cut by *Taq*αI. Sequence data from this individual was 11% different on average from *C. alliariae*, the most similar of the target species. BLAST (Basic Local Alignment and Search Tool, http://www.ncbi.nlm.nih.gov/blast/ Blast.cgi) searches suggest that these three taxa were members of the genus *Ceutorhynchus* or belonged to closely-related genera. The percent sequence differences between candidate agents (9.4-17.0%), between unknown taxa and candidate agents (6.2-17.8%), and between the three unknown taxa (12.7-17.8%) was similar to the average among species differences (12.9%) observed from a GenBank survey of 28 Ceutorhynchus COI sequences. While

percent sequence divergence is not a definitive tool for differentiating among species (Cognato, 2006), the differences observed suggest that each of the three unknown taxa represents a different species. This is also supported by the average within species (0.9%) percent difference for 14 Ceutorhynchus species surveyed from GenBank. There was no evidence that the PCR primers used here amplified DNA from parasitoids of these species. Specific attempts directed at amplifying mtDNA from the larval ectoparasitoids found in our samples using C1-J-2183 and A3113 resulted in no PCR product.

Double digestion with both *Alu*I and *Taq*αI allowed differentiation of *C. alliariae, C. roberti,* and *C. scrobicollis* from the unknown species in our sample (Figure 2), and observed banding patterns matched those expected based on sequence data (Table 2).

The PCR-RFLP assays presented here have been used to rapidly differentiate larvae of the three proposed candidate agents during quarantine and native range host-specificity studies (Gerber et al., unpublished data). While it is possible that undetected intraspecific variation at the restriction sites may exist, our relatively large sample sizes collected throughout Europe and our evaluation of *Ceutorhynchus* COI sequences from GenBank, suggest these PCR-RFLP assays are fairly robust and may be reliably used in identification of *C. alliariae, C. roberti,* and *C. scrobicollis* larvae.

REFERENCES

Brower, A.V.Z., 1994. Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). Molecular Phylogenetics and Evolution 3, 159-174.

Cognato, A.I., 2006. Standard percent DNA sequence difference for insects does not predict species boundaries. Journal of Economic Entomology 99, 1037-1045.

Dieckmann, L., 1972. Beiträge zur Insektenfauna der DDR. Coleoptera – Curculionidae: Ceutorhynchinae. Beitraege zur Entomologie 22, 3-128.

Erlandson, M.A., Gariepy, T.D., 2005. Tricks of the trade: developing species-specific PCR primers for insect identification. Bulletin of the Entomological Society of Canada 37, 76-82.

Gariepy, T., Kuhlmann, U., Gillott, C., Erlandson, M., 2008. A large-scale comparison of conventional and molecular methods for the evaluation of host-parasitoid associations in non-target risk-assessment studies. Journal of Applied Ecology 45, 708-715.

Heard, T.A., Pettit, W., 2005. A review of the surveys for natural enemies of *Mimosa pigra*: what does it tell us about surveys for broadly distributed hosts? Biological Control 34, 247-254.

Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003a. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B-Biological Sciences 270, 313-321.

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

Hinz, H., Gerber, E., Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande, Annual Report for 1998. CABI Bioscience, Delemont, Switzerland, 1999.

Hinz, H., Gerber, E., Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande, Annual Report for 2000. CABI Bioscience, Delemont, Switzerland, 2001.

Hinz, H.L., Gerber, E., Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande. Annual Report for 1999. CABI Bioscience, Switzerland, Delemont, 2000.

Jarman, S.N., 2006. Cleaver: software for identifying taxon specific restriction endonuclease recognition sites. Bioinformatics 22, 2160-2161.

Lopez, M.M., Llop, P., Olmos, A., Marco-Noales, E., Cambra, M., Bertolini, E., 2008. Are molecular tools solving the challenges posed by detection of plant pathogenic bacteria and viruses? Current Issues in Molecular Biology 11, 13-45.

McKern, J.A., Szalanski, A.L., 2007. Molecular diagnostics of economically important clearwing moths (Lepidoptera : Sesiidae). Florida Entomologist 90, 475-479.

Mullis, K.B., Faloona, F., Scharf, S., Saiki, R.K., Horn, G.T., Erlich, H.A., Specific enzymatic amplification of DNA invitro - The polymerase chain - reaction, Cold Spring Harbor Symposia on Quantitative Biology, Vol. 51, 1986, pp. 263-273.

Newton, C.R., Graham, A., 1997. PCR. BIOS Scientific Publishers, Oxford.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487-91.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, 1350-4.

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

Unruh, T.R., Woolley, J.B., 1999. Molecular methods in classical biological control. In: Bellows, T., Fisher, T.W., Eds.), Handbook of Biological Control. Academic Press, San Diego, pp. 57-85.

Wells, J.D., Stevens, J.R., 2008. Application of DNA-based methods in forensic entomology. Annual Review of Entomology 53, 103-120.

Table 1. Estimated fragment lengths generated with digestion by restriction endonuclease *Taq*αI. Country codes are: AT (Austria), BG (Bulgaria), CH (Switzerland), DE (Germany), FR (France), GE (Georgia), IT (Italy) RO (Romania), UK (United Kingdom). Numbers of individuals from each locale are given in parentheses after country codes. Unk. = Unknown.

Species	Location(n)	Fragment Lengths
C. alliariae	CH(6),IT(6),AT(9),DE(20),FR(10),RO(1),UK(3)	742,232
	CH(2)	616,232,126
C. roberti	CH(30),DE(23),RO(3),BG(4)	974
C. scrobicollis	DE(32),RO(4),GE(7)	616,358
Unk. Sp. 1	DE(1)	742,232
Unk. Sp. 2	BG(1)	974
Unk. Sp. 3	GE(13)	776,198

Table 2. Estimated fragment lengths generated with double digestion by restriction endonucleases *Alu*I and *Taq*αI for differentiation *Ceutorhynchus alliariae, C. roberti,* and *C. scrobicollis* and three unknown species (Unk. Sp.) Country codes follow Table 1. Unk. = Unknown.

Species	Location(n)	Fragment Lengths
C. alliariae	CH(6),IT(6),AT(9),DE(20),FR(10),RO(1),UK(3)	476,252,232,14
	CH(2)	476,232,140,112,14
C. roberti	CH(28),DE(22),RO(3),BG(4)	293,252,158,96,88,87
	CH(2),DE(1)	348,293,158,88,87
C. scrobicollis	DE(31),RO(2)	323,270,246,112,23
	RO(2)	270,246,188,158,112
	GE(7), DE(1)	270,246,165,158,112,23
Unk. Sp. 1	DE(1)	270,232,170,165,123,14
Unk. Sp. 2	BG(1)	458,270,246
Unk. Sp. 3	GE(13)	435,293,158,48,40

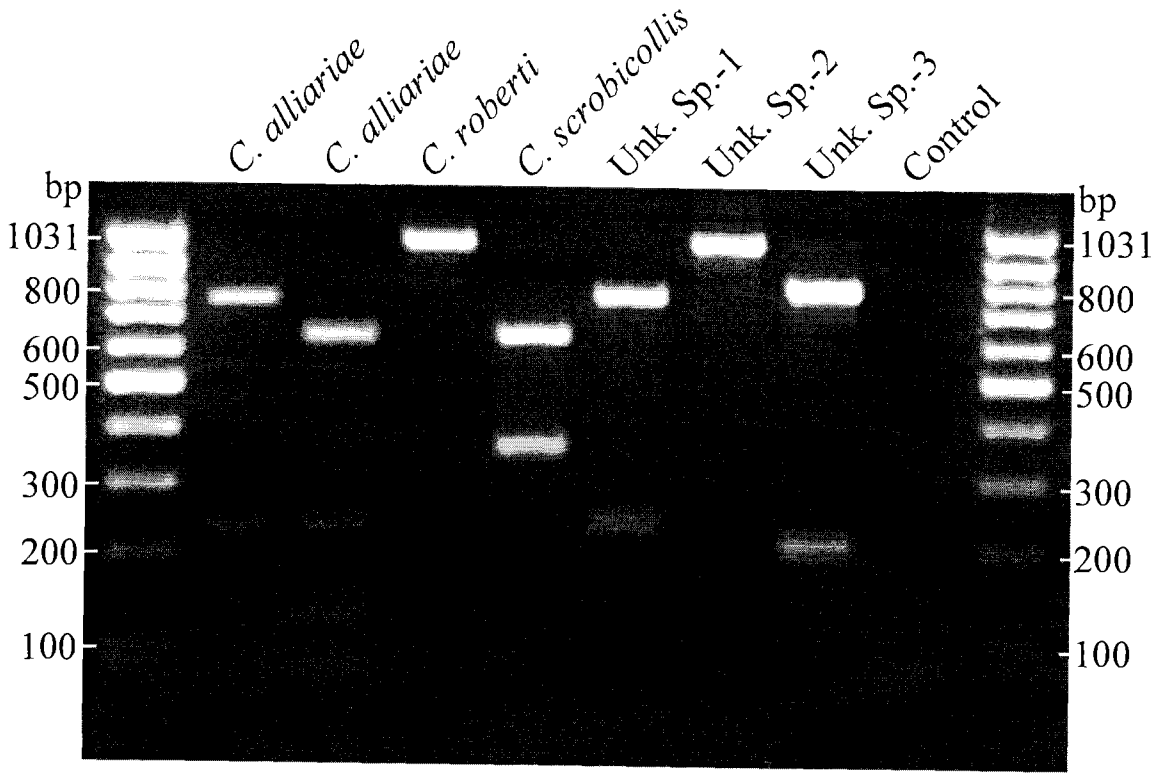

Figure 1. *Taq*αI digest of *Ceutorhynchus alliariae, C. roberti,* and *C. scrobicollis* and three unknown species (Unk. Sp.) run on a 2% agarose gel for 1.75 hours at 75 volts. Ladder is Low Range Mass Ruler (Fermentas).

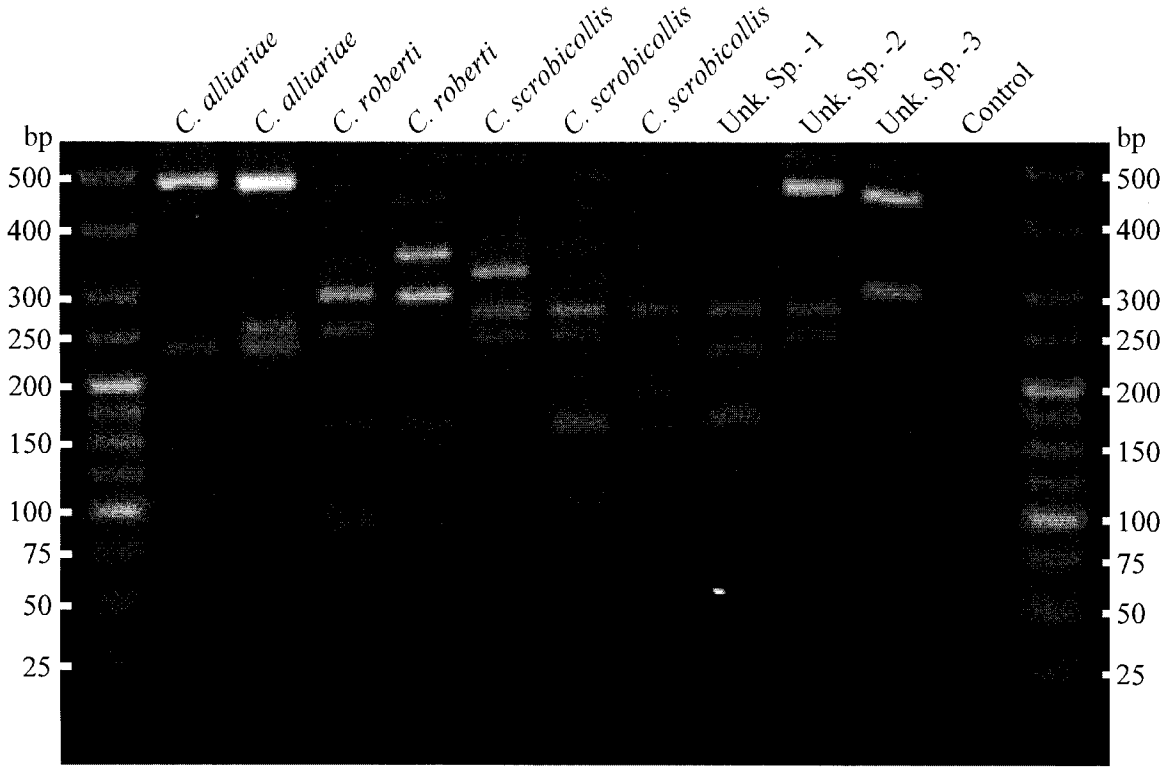

Figure 2. Double digestion with *Alu*I and *Taq*αI of *Ceutorhynchus alliariae, C. roberti,* and *C. scrobicollis* and three unknown species (Unk. Sp.) run on a 2.5% agarose gel for 2.5 hours at 75 volts. Ladder is HyperLadder V (Bioline).

CHAPTER 2

The benefits of pre-release population genetics: A case study using *Ceutorhynchus scrobicollis*, a candidate agent of garlic mustard.

INTRODUCTION

Classical biological control, the introduction of natural enemies of an invasive pest from its area of origin to the invaded range, is sometimes the only available option for managing widespread invasive weeds. Program success rates are reasonably high (~50-80%), however, successful control by individual agents is lower (~20%)(McFadyen, 1998). Concern over non-target and indirect effects has led to suggestions of methods to refine the weed biological control process, including pre-release evaluations of agent effectiveness and reductions in the number of agents released (Louda et al., 2003; Pearson and Callaway, 2005; Simberloff and Stiling, 1996). While many factors influence whether introduced agents establish and achieve control, one factor long thought to have the potential to affect the outcome of biological control programs is the amount of genetic variation within the biological control agent (Messenger and van den Bosch, 1971; Messenger et al., 1976; Myers and Sabath, 1981). Unruh and Woolley (1999) reviewed the role of molecular methods in classical biological control and discuss its potential value in helping to clarify the many cases of subspecies or geographic races encountered in biological control (DeBach, 1969; Rosen, 1986). While it is becoming more routine, we are currently aware of only five studies that have characterized genetic variation within or among populations of weed

biological control agents prior to their release (Briese et al., 1996; Goolsby et al., 2006a; Manguin et al., 1993; Nowierski et al., 1996; Olivieri et al., 2008). This paucity of research is surprising, particularly given that specialized natural enemies are a critical limiting resource in weed biological control (Goolsby et al., 2006b). Basic population genetic data can help maximize the potential of the few agents available.

Population genetic data gathered across multiple sites can help determine the spatial extent of populations of candidate biological control agents and the degree of differentiation among them. This knowledge can inform decisions about where to collect agents for host-specificity testing and later introductions. Highly specialized phytophagous insects can exhibit substantial genetic differentiation among populations and local adaptation within limited geographical areas (Komatsu and Akimoto, 1995; Mopper, 1996; Mopper et al., 1995). Traits such as host-preference and performance, efficacy on different genotypes of the host, and temperature requirements have been shown to vary among populations of some arthropods (Goolsby et al., 2006a; Ireson et al., 2008; Milbrath and DeLoach, 2006; Rajpurohit et al., 2008). Information on the population genetic structure of candidate agents enables researchers to collect and test with an understanding of whether they are working with one or more populations. In addition, informed choices can be made on whether to test distinct populations for variation in ecologically relevant traits. Although genetic variation in molecular loci among populations is not a surrogate for direct assessment of ecologically important

traits, it may be correlated with variation in ecologically relevant traits (Merila and Crnokrak, 2001), and can thus help guide research.

Population genetic data gathered across multiple sites can also provide researchers with estimates of agent dispersal ability, a trait that should be incorporated into the development of release and re-distribution strategies (Jonsen et al., 2007). Mark and recapture experiments are the traditional, direct method of estimating dispersal. However, such data are difficult to obtain over large scales, and rare long-distance dispersal events are seldom observed (Peterson and Denno, 1998). Data on the genetic structure of populations can provide an alternative means of assessing the movement of individuals among populations (Berry et al., 2004).

Population genetic data gathered within sites or populations can provide a measure of the diversity of the population being studied. This information can be used to estimate the number of individuals that should be collected for host-specificity testing to ensure that the diversity of the population is encompassed. Preference for a host and performance on a host are physiologically complex traits (Miller and Strickler, 1984; Sheppard et al., 2005), and are likely influenced by many interacting genes (Forister et al., 2007). Thus, adequate representation of the genetic diversity of a population in host-specificity testing can help assure the accuracy of those results.

Within-population genetic data also provide the information necessary to maintain high genetic diversity through the introduction process. Genetic diversity of introduced agents is hypothesized to be important in enabling introduced

agents to adapt in the novel environment (Messenger and van den Bosch, 1971; Messenger et al., 1976; Myers and Sabath, 1981; Phillips et al., 2008). Though we are aware of no data from biological control agents addressing this hypothesis directly, a recent meta-analysis showed that population genetic diversity is strongly correlated with fitness (Reed and Frankham, 2003). Data on within-population genetic diversity can be used to estimate the number of individuals that should be collected in order to prevent a loss of diversity upon introduction.

Here we present pre-release data on genetic variation of *Ceutorhynchus scrobicollis* Nerensheimer & Wagner (Curculionidae), a candidate agent for the biological control of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande (Brassicaceae). Using mitochondrial DNA sequence data and amplified fragment length polymorphisms, our specific aims were to 1) evaluate whether the area in Europe where *C. scrobicollis* is being collected for host-specificity testing constitute one or more populations, 2) assess the levels of genetic diversity within that area, and use these measurements to estimate the numbers of individuals needed during host-specificity testing and later introduction to adequately represent the diversity of the population, and 3) evaluate the dispersal potential of the candidate agent. We further determined intraspecific genetic variation over a broader range by including populations from more distant geographic regions in Europe.

METHODS

Study System. Garlic mustard (*Alliaria petiolata*) is a strictly biennial herb native to Eurasia. Garlic mustard spends its first year as a rosette and develops one to several flowering stems the following spring. Primary habitats in both the native and introduced range include forest understory and moist, semi-shaded habitats. Thought to have been originally introduced for medicinal use and as a green vegetable (Cavers et al., 1979), garlic mustard has expanded its range exponentially since it was first recorded in 1868 on Long Island, NY (Nuzzo, 1993). Garlic mustard now occurs in 37 states and 4 Canadian provinces (Plants National Database: http://plants.usda.gov/index.html) and is directly reduces of native plant diversity (Stinson et al., 2007). Two studies of genetic variation indicate that North American populations of *A. petiolata* are slightly less diverse than populations in the native range (Durka et al., 2005; Meekins et al., 2001). However, significant variation among populations was present and, when compared to the probable source regions, no bottleneck was evident (Durka et al., 2005). High allelic diversity in the introduced range strongly suggests multiple introductions of *A. petiolata* to North America, most likely from the British Isles, and northern and central Europe (Durka et al., 2005; Meekins et al., 2001).

Ceutorhynchus scrobicollis is a univoltine weevil proposed for the control of garlic mustard in the United States. *Ceutorhynchus scrobicollis* is recorded from Germany (Dieckmann, 1972), Italy (Abazzi and Osella, 1992), eastern France (Schott, 2000), Austria, Hungary (pers. comm. O. Merkl), and the Czech

Republic (pers. comm. C. Brandstetter). The northernmost records are from Northern Germany and Poland, while the easternmost records are from the eastern Caucasus region (E. Colonelli and B. A. Korotyaev, pers. comm.). Adults aestivate over the summer and lay eggs into rosettes from mid-September through March (Blossey et al., 2001; Gerber et al., 2009). Larvae develop primarily in the root crowns through winter and early spring until leaving the plants to pupate in the soil. Adults emerge during May and June.

Study Approach

Collections. Sampling of *C. scrobicollis* concentrated mainly on the region around Berlin, Germany (Table 1 & Fig 1) where *C. scrobicollis* has been collected for host-specificity testing. To provide perspective on the population genetic variation in the greater Berlin area, we sampled an additional site near Klieken at approximately 100 km distance from Berlin (site 19, Fig 1) and collaborators collected from one site each in Romania (Iasi) and Georgia (T'bilisi). Collections were made in April 2003 for all German sites, and in April 2004 for Romania and Georgia. To obtain *C. scrobicollis* larvae for analyses, garlic mustard plants were collected from each site at one meter intervals along a transect across the long axis of the population. Plants were dissected and all larvae placed immediately in 100% ethanol to preserve DNA. Genomic DNA was extracted from a single larva per plant using DNeasy Tissue Kits (QIAGEN®, Germantown, MD, USA) following the manufacturer's instructions. DNA was eluted into 150μl of the supplied buffer and stored at -20°C. *Ceutorhynchus scrobicollis* larvae were

identified using a polymerase chain reaction – restriction fragment length polymorphism (PCR-RFLP) (Rauth and Hufbauer, submitted, Ch. 1).

MtDNA sequencing. We sequenced a portion of the mitochondrial DNA (mtDNA) encompassing approximately 850 bp of cytochrome oxidase subunit one (COI), tRNALeu, and a short section of COII to explore intraspecific diversity and the evolutionary relationships among collection sites. Between 1 and 7 individuals from each collection site (avg. 3.6 per site) were sequenced to examine the intraspecific variation of the COI region. The sequencing protocol is described in detail in Rauth and Hufbauer (submitted, Ch. 1).

Amplified fragment length polymorphisms. Amplified fragment length polymorphisms (AFLPs) were used to assess genetic variation among sites where *C. scrobicollis* was collected. We followed the protocol of Vos et al. (1995) with modifications. The adapters and primer cores used followed Hawthorne (2001). The enzymes were *Eco*RI and *Pst*I (New England Biolabs®, Germantown, MD, USA), two six-base cutters. Four selective primer pair combinations were used during the selective amplification PCR: (* denotes the 5' end was labeled with 6-FAM™, ^ denotes the 5' end was labeled with VIC™) **Eco*RI-AGTG/*Pst*I-GAC; ^*Eco*RI-CGAT/*Pst*I-GAC; ^*Eco*RI-GACT/*Pst*I-AGT; **Eco*RI-TGAG/*Pst*I-AGT.

Genomic DNA was digested and adapters were ligated to the resulting fragments in a single two hour incubation step at 37°C. Digestion and ligation were performed in 20μL reaction volumes of 0.5x NEBuffer 3 (50 mM Tris-HCl,

100 mM NaCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol, pH 7.9 @ 25°C), 0.5x T4 DNA Ligase Reaction Buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM Dithiothreitol, 25 µg/ml BSA, pH 7.5 @ 25°C), 0.5µg BSA, 4.5pmol of each adapter, 5 U *Eco*RI (in 300 mM NaCl, 10 mM 2-Mercaptoethanol, 0.1 mM EDTA, 200 µg/ml BSA, 50% Glycerol, 0.15% Triton X-100, pH 7.5 @ 25°C), 5 U *Pst*I (in 10 mM Tris-HCl, 200 mM NaCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 200 µg/ml BSA, 50% Glycerol, 0.15% Triton X-100, pH 7.4 @ 25°C), 66.8 U T4 DNA ligase (in 10 mM Tris-HCl, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, 50% Glycerol, pH 7.4 @ 25°C), and 8µL genomic DNA. After digestion and ligation, the product was diluted with 100µL $TE_{0.1}$. For the pre-amplification PCR *Eco*RI and *Pst*I primers with three selective nucleotides were used (e.g. 5' *Eco*RI-AGT/*Pst*I-GAC). The pre-amplification step was performed in 10µL reaction volumes of 1x PCR buffer (20mM Tris-HCL, pH 8.4, 50mM KCl), 2mM $MgCl_2$, 0.2mM dNTPs, 1 U *Taq* polymerase, 1.8pmol of each primer, and 1µL of diluted digestion/ligation product. Pre-amplification PCRs were run on a Mastercycler gradient (Eppendorf AG, Westbury, NY, USA) thermal cycler starting at 72°C for 2min, followed by 20 cycles of 20s at 94°C, 30s at an annealing temperature of 56°C, 2min at 72°C for polymerase extension, and a final 30 minutes at 60°C. After pre-amplification PCR product was diluted with 140µL $TE_{0.1}$. The selective amplification step was performed with the same reaction constituents as for the pre-amplification except for the change of *Eco*RI primers. After an initial 2min denaturation period, reactions underwent 30 cycles of 20s at 94°C, 30s at 66°C (decreasing by 1°C each cycle until at 56°C), 2min at

72°C, with a final 30min at 60°C. One microliter of selective amplification product was then added to 10μL HiDi-Formamide and 0.13μL GeneScan™ LIZ® 500 size standard (Applied Biosystems Inc., Foster City, CA, USA). The product was then denatured at 95°C for 3 minutes and placed immediately on ice prior to being run on an ABI 3100 capillary DNA sequencer (Applied Biosystems Inc.). Electropherograms were visualized using GeneMarker ver.1.51 (SoftGenetics LLC, State College, PA, USA). Markers were scored individually by hand for presence/absence of a peak. Only polymorphic loci that could be scored unambiguously across the data set were used in analyses. Scoring was performed without reference to the identity of the individual or site of origin. Only markers with frequencies greater than 5% or less than 95% were used in analyses.

Analyses

MtDNA sequence analysis. To visualize *C. scrobicollis* haplotype diversity and distribution of haplotypes among collection locations, we constructed a maximum parsimony haplotype network using the statistical algorithms of Templeton et al. (1992) implemented in the program TCS 1.2.1 (Clement et al., 2000). For the samples from Germany, haplotype diversity, nucleotide diversity (π), the average number of pairwise nucleotide differences (k), and theta (θ) per site, which estimates the number of differences per nucleotide expected between two randomly chosen alleles under mutation-drift equilibrium, were calculated in DNASP v. 4.50.2 (Rozas et al., 2003).

AFLP analyses. We tested the independence of markers by calculating their pairwise linkage index $D_{A,B} = 1/n\sum i|V(A,i) - V(B,i)|$. $V(A,i)$ is the allele value for marker A of individual i, $V(B,i)$ is the allele value for marker B of individual i, and n is the total number of individuals analyzed. D-values were calculated using the software DDM version 0.1 (P. Berthier, personal communication).

Allele frequencies, which are used to generate many common population genetic statistics, must be estimated for dominant AFLP data. This is because of the presence/absence nature of the markers, which prevents distinguishing individuals that are homozygous for an allele from heterozygous individuals. We assumed Hardy-Weinberg equilibrium and then used Zhivotovsky's (1999) Bayesian method with a non-uniform prior distribution as implemented in AFLPsurv 1.0 (Vekemans et al., 2002) to estimate allele frequencies. To evaluate the population structure of *C. scrobicollis,* estimated allele frequencies were used to calculate genetic diversity and Nei's (1972) genetic distance following the approach of Lynch and Milligan (1994), as implemented in AFLPsurv 1.0. Significance of genetic distances among pairs of sites was determined by bootstrapping over loci. Confidence intervals were based on bootstrap percentiles generated from 10000 permutations. To estimate the number of individuals from the Berlin area that would need to be collected for host-specificity testing and later introductions so that 90-99% (of the alleles present at greater than 5%) were captured in the sampling, we randomly sampled one to n individuals from the data set and calculated the proportion of alleles captured. We performed this estimate for the Berlin area because this is where individuals have been collected for host-

specificity testing. Our sample size for this area was 178 individuals. We performed 100 permutations and calculated the average and standard error for these estimates in an excel spreadsheet.

As a means of visualizing the genetic differentiation among collection sites, a principal coordinates analysis was performed in GENAlEx 6.2 (Peakall and Smouse, 2006) with Nei's pairwise genetic distances. This analysis allows us to reduce the multidimensional data set of all population pairwise genetic distances down to two dimensions while still capturing most of the variation among the pairwise relationships. To examine how genetic differentiation among collection sites corresponded to the spatial relationship of those sites, we assessed whether populations were isolated by distance, that is, whether genetic distance increased with increasing geographic distance. Isolation by distance was examined at three spatial scales: for all sites, restricting the analysis to sites within Germany, and further restricting the analysis to the sites within the Berlin focal area. We implemented a Mantel test using IBD ver. 1.52 (Bohonak, 2002).

To gain a different perspective on population structure and evaluate the likelihood that sampled individuals had dispersed from other locations, population assignment tests were performed using AFLPOP 1.1 (Duchesne and Bernatchez, 2002). This assignment method was adapted from (Paetkau et al., 1995) and computes the likelihood that an individual with a given genotype be found in each of the candidate populations based on the allele frequencies of the different populations. We used the re-allocation (leave-one-out) procedure with the allocation minimal log-likelihood difference set to one. With this procedure an

individual is removed from its collection site, allele frequencies are calculated for all collection sites, then the likelihood that the individual came from each of the sites is computed. With the allocation minimal log-likelihood difference set to one, for an individual to be assigned to any collection site, it had to be ten times more likely to have come from that site than any other site. By calculating the distances between sites where individuals were collected and where they were reassigned we were able to evaluate dispersal. Comparison of dispersal estimates from assignment tests with long-term mark-and-recapture records has shown that assignment tests can detect dispersal accurately even when there is relatively little population genetic structure (e.g. low F_{ST} or Nei's D) (Berry et al., 2004). We chose not to use $F_{ST}=1/(4Nm+1)$ (e.g. Slatkin 1987) to estimate dispersal from our molecular data because our study system likely violates the assumptions inherent to that relationship (Whitlock and McCauley, 1999).

RESULTS

Haplotype Network and Sequence Diversity. We obtained 592bp of sequence from of each 43 individuals. Six haplotypes of the mtDNA cytochrome oxidase I gene were found from the 43 individuals sequenced (Accession numbers EU551169, EU551173, EU852398-852401). Haplotype A, the most common haplotype, was found in 25 individuals from each of the German sites except site 17 (Fig. 2). Haplotype C was found in individuals from four sites (10, 11, 16, and

17), in the central and western areas of Berlin. Haplotype D was found in one individual each from site 17 and 11. Haplotypes B and F were only found in Romania. Only a single haplotype (E) was found from the seven individuals sequenced from Georgia. The average number of nucleotide differences (k) among all sequences was 6.7. The maximum number of nucleotide differences (18 SNPs) was found between haplotypes D, in Berlin, and E, in Georgia.

Within the Berlin area (sites 10-18) haplotype diversity (H_d (±SD)) was moderate at 0.43 (±0.10), nucleotide diversity (π (±SD)) was 0.004 (±0.001), the average number of nucleotide differences (k) was 2.54, and theta (θ) per site was 0.003.

Genetic Diversity in AFLP. We scored 73 polymorphic loci for 260 individuals from 11 populations. The mean pairwise linkage index was 0.41 and no marker pairs had *D*-values lower than 0.05 or higher than 0.95. These values indicate that all markers were unlinked and, therefore, all were retained for further analyses.

Over all sites, total gene diversity, H_T, was 0.29, with 83% (H_W = 0.24) of the total diversity being found within sample sites and the remaining 17% (H_S = 0.05) due to genetic differentiation among sample sites. Within the Berlin area, total gene diversity, H_T, was 0.32, with 94% (H_W = 0.30) of the total diversity being found within sample sites and the remaining 6% (H_S = 0.02) due to genetic differentiation among sample sites. Among population variation was greater when all sites were analyzed than when just the Berlin area sites were analyzed (Overall sites: F_{st} = 0.170; Berlin area only: F_{st} = 0.057). The estimate of the

number of individuals that would need to be sampled to capture 90% of the genetic diversity of common alleles (> 5% presence) in the Berlin area was 10 (± 0.004) individuals. To capture 99% of the diversity in the Berlin area at least 27 (± 0.001) individuals should be collected (Fig. 3).

Population Structure. The principal coordinates analysis calculated from Nei's genetic distances (Appendix A) showed that most of the German populations clustered together. Site 19 (~100 km SW of central Berlin), but also site 16 (within Berlin area) did not cluster as tightly, in spite of their geographic proximity to the other sites (Fig. 4, Appendix A). Romanian and Georgian samples were also genetically distinct from all other sites and each other. The first two axes captured 82% (63% and 19%, respectively) of the variation in Nei's genetic distances among the sample locations.

A Mantel test to determine whether Nei's genetic distances between all pairs of sites were correlated with geographic distances between sites revealed significant isolation by distance ($Z = 5417$, $r = 0.95$, $P < 0.001$, Fig. 5a). When excluding the two most remote sites in Romania and Georgia, the Mantel test of all German sites (max. distance between sites 117 km) still showed significant isolation by distance ($Z = 70$, $r = 0.70$, $P < 0.005$, Fig. 5b). However, within the Berlin area alone, sites were not significantly isolated by distance ($Z = 27$, $r = 0.18$, $P = 0.20$, Fig. 5c).

Assignment tests further supported the relationships among sites observed via the principal coordinates analysis. Most Berlin sites had either a large

proportion of individuals that could not be reassigned with confidence or 1-3 individual(s) that were reassigned to other sites (Table 2). Site 16, from the Berlin area, stands out in that the majority of individuals (72%) were reassigned back to site 16 and no individuals were assigned to any other Berlin area sites. Site 19, the one German site outside the Berlin area, also had a large proportion of individuals that reassigned back to their collection site, and one individual that was reassigned to site 16, the nearest Berlin area site (65 km). Overall, nine of the 195 individuals from Germany (4.6%) were reassigned to sites other than their collection site (Table 3, Figure 6) and the average dispersal distance estimated from these reassigned individuals was 28 km (Table 3). Focusing only on the 107 individuals that had been assigned to sites, 8.4% were reassigned to one other than their collection site. One hundred percent of individuals from both Georgia and Romania were reallocated to their original collection site. The relatively large proportion of individuals not assigned within Germany is a pattern also seen in other studies (e.g. Conord et al., 2006; Raffl et al., 2006). It may be that they were from local sites not sampled as part of this study, or were 2nd or later generation offspring from recently dispersed individuals who had mated with individuals originating at the site, and thus thwarting our ability to assign them to any site.

DISCUSSION

In this paper we report the population genetic structure and diversity data for *C. scrobicollis*, a candidate for the biological control of garlic mustard, and discussed how this type of information has the potential to help improve establishment, assure safety, direct release strategies and increase efficacy of this weevil.

Number of populations. Results suggest that *C. scrobicollis* in the vicinity of Berlin, Germany, the region from which individuals have been collected for host-specificity testing, constitute a single population consisting of many sub-populations, with gene flow occurring between sub-populations. Individuals are dispersing in and around Berlin, however some sites show enough differentiation that migrants can be detected.

Mitochondrial DNA sequence data indicate that the sample sites share a recent genetic history. The most common COI haplotype was found at all but one of the sites in Germany (site 17). This exception may be due to sample size or a founder effect. Site 17, which was a relatively small stand of garlic mustard, was only about one kilometer from a large, heavily infested stand (site 18) where the common haplotype was found. Our AFLP results further indicate that weevils from most sites in the Berlin area are closely related. Individuals from one site in this area (site 16) appear to be more closely related to samples from site 19 which was 65 km farther to the southwest (Appendix A). The lack of significant

isolation by distance among Berlin area sites suggests that, at this scale, genetic differentiation is not due to geographic distance. Assignment tests suggested that migration among patches of garlic mustard in the Berlin area is common, however, in concordance with the principal coordinates analysis, no individuals from site 16 were reassigned to any other collection sites. These results suggest that migration of *C. scrobicollis* is occurring among the patches of garlic mustard sampled in the Berlin area, but also that individuals from site 16 are somewhat distinct. There are two main possibilities that might explain the genetic similarity between site 16 and the more distant site 19; site 16 may have been recently founded by individuals from site 19 or close relatives, or there may be close continuity of garlic mustard stands between these two collection sites. That both sites 16 and 19 are differentiated from the other collection sites in the Berlin area shows that population differentiation can occur over relatively short distances. This local differentiation suggests that care should be taken to restrict collections of individuals for introduction to the area around Berlin from which individuals were collected for host-specificity testing.

Capturing Genetic Diversity. The number of individuals used in host-specificity testing is often determined by the logistics of collecting individuals for testing and the spatial and temporal limitations on experiment size. Incorporating information on the genetic diversity of the agent population used in host-specificity testing would provide increased confidence in the testing results. To the best of our knowledge, this has not been done so far. Using AFLP data, we estimated that

only 10 or 27 individuals need to be sampled to capture 90 or 99%, respectively, of the common alleles (> 5%) of *C. scrobicollis* populations in Germany. Over the past ten years more than 4,100 (50-1,100 per year) individuals have been sampled and incorporated in the rearing colony used to conduct host specificity tests (E. Gerber and H. Hinz, unpublished data). These numbers dramatically exceed the estimated number of individuals needed to capture most of the genetic diversity of *C. scrobicollis* populations in the sampling area. We are therefore confident that the results of host-specificity tests adequately represent the potential host range of the candidate agent population proposed for field release. Furthermore, if adaptation to the novel environment is important for successful biological control then the knowledge gained in this study can also be used to facilitate maintenance of diversity through the quarantine and introduction process by providing researchers with a benchmark against which the genetic diversity of the introduced population can be compared.

Dispersal. Information on potential dispersal distances may be important in planning release strategies. Our AFLP data suggest that while a good proportion of *C. scrobicollis* remain in their natal sites, they can disperse. Assignment test-based dispersal distance estimations among our sites averaged 28 km, with a maximum of 65 km (Table 3). The percent of individuals assigned to sites other than their collection site was small (4.6% when considering all individuals, and 8.4% when focusing just on those with confident assignment), but the distances traveled were substantial. This result clearly indicates that *C. scrobicollis* can

disperse by flight, even though in all of the laboratory and field research work with this weevil no flight has been observed (E. Gerber pers. comm.). Reports on the range expansion of the congener *C. obstrictus*, which is similar in many biological and life history traits, suggest that it is capable of dispersing approximately 55 km a year (Dosdall et al., 2002). The maximum dispersal inferred in this study (65 km) was similar to that observed for *C. obstrictus*, however, due to our limited sampling locations, we cannot say what the upper most range of potential dispersal may be. It should be noted that adults of *C. scrobicollis* may live for periods of three years or more (Gerber et al., 2003), and thus movement throughout the life of an adult has the potential to be considerably greater than we observed. Lack of shared COI haplotypes between Germany and Romania and Georgia and between Romania and Georgia, however, suggests that dispersal over much greater distances is very rare if it occurs at all. While dispersal can be a double edged sword in biological control (see Jonsen et al., 2007), our data suggest that if *C. scrobicollis* establishes in North America, then we would expect it to be able to disperse to novel patches of garlic mustard within the eastern U.S. where its host is widespread. However, in the western Midwest and the Northwest, where infrequent collections indicate the plant may be a sporadic rather than common component of the regional flora, and long distances without garlic mustard patches are frequent, additional releases targeting isolated areas will likely be needed.

Exploration of intraspecific variation. It was suggested over 20 years ago (Sands and Harley, 1981) that the success rates of biological control may be improved by investigating intraspecific genetic variation of potential agents. This appears to be true for two recent biological control programs. In one, different populations (ecotypes) of *Diorhabda elongata* Brulle, introduced for the control of salt cedar (*Tamarix* spp.), exhibit variation in photoperiod induced diapause (DeLoach et al., 2004). The different ecotypes are being tested to determine if they can establish at the lower latitude infestations in North America (Bean et al., 2007). In the second example, variation among populations in host use is being expoited for biological control. Different populations of the mite *Floracarus perrepae* Knihinicki & Boczek show varying ability to develop on and damage different introduced genotypes of the invasive fern *Lygodium microphyllum* (Cav.) R. Br. in Florida, US (Goolsby et al., 2006a). Searching for such intraspecific variation in ecologically important traits (for example tolerance to low winter temperature) might be of interest if the *C. scrobicollis* population from the Berlin area does poorly in North America. Our mitochondrial and AFLP data suggests that the individuals collected from Romania and Georgia are differentiated from the German individuals and from each other. Given this variation in (presumably) neutral markers, it may be worthwhile investigating whether parallel differences exist in ecologically relevant traits such as temperature tolerance.

Applying the tools of population genetics to biological control agents prior to release has great potential to benefit biological control. It gives us confidence that the individuals being used in host-specificity tests are from one population,

which leads to reduced risk. It enables us to make clear predictions about the minimal numbers of individuals needed for introduction to maintain diversity, which may lead to greater rates of establishment and success. It gives us an indication of dispersal potential, which may facilitate design of future release strategies. And ultimately, pre-release population genetic studies such as this one will give us baseline data with which we can evaluate the importance of within and among population genetic variation to the safety and success of biological control.

REFERENCES

Abazzi, P., Osella, G., 1992. Elenco sistematico-faunistico degli Anthribidae, Rhinomaceridae, Attelabidae, Apionidae, Brenthidae, Curculionidae italiani (Insecta, Coleoptera, Curculionoidea). Redia 75, 267-414.

Bean, D.W., Dudley, T.L., Keller, J.C., 2007. Seasonal timing of diapause induction limits the effective range of *Diorhabda elongata* deserticola (Coleoptera : Chrysomelidae) as a biological control agent for tamarisk (Tamarix spp.). Environmental Entomology 36, 15-25.

Berry, O., Tocher, M.D., Sarre, S.D., 2004. Can assignment tests measure dispersal? Molecular Ecology 13, 551-561.

Blossey, B., Nuzzo, V., Hinz, H., Gerber, E., 2001. Developing biological control of *Alliaria petiolata* (M. Bieb.) Cavara and Grande (garlic mustard). Natural Areas Journal 21, 357-367.

Bohonak, A.J., 2002. IBD (isolation by distance): A program for analyses of isolation by distance. Journal of Heredity 93, 153-154.

Briese, D.T., Espiau, C., Pouchot-Lermans, A., 1996. Micro-evolution in the weevil genus *Larinus*: The formation of host biotypes and speciation. Molecular Ecology 5, 531-545.

Cavers, P.B., Heagy, M.I., Kokron, R.F., 1979. Biology of Canadian Weeds .35. *Alliaria-petiolata* (M Bieb) Cavara and Grande. Canadian Journal of Plant Science 59, 217-229.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Conord, C., Lemperiere, G., Taberlet, P., Despres, L., 2006. Genetic structure of the forest pest *Hylobius abietis* on conifer plantations at different spatial scales in Europe. Heredity 97, 46-55.

DeBach, P., 1969. Uniparental, sibling and semi-species in relation to taxonomyand biological control. Israel Journal of Entomology 4, 11-28.

DeLoach, C.J., Carruthers, R., Dudley, T., Eberts, D., Kazmer, D., Knutson, A., Bean, D., Knight, J., Lewis, P., Tracy, J., Herr, J., Abbot, G., Prestwich, S., Adams, G., Mityaev, I., Jashenko, R., Li, B., Sobhian, R., Kirk, A., Robbins, T., Del Fosse, E., First results for control of saltcedar (*Tamarix* spp.) in the open field in the western United States. In: Cullen, J.M., Briese, D.T., Kriticos, D.J., Lonsdale, W.M., Morin, L., Scott, J.K., Eds.), XI International Symposium on the Biological Control of Weeds. CSIRO, Canberra, Australia, 2004, pp. 505-513.

Dieckmann, L., 1972. Beiträge zur Insektenfauna der DDR: Coleoptera - Curculionidae: Ceutorhynchinae. Beiträge zur Entomologie 22, 3-128.

Dosdall, L.M., Weiss, R.M., Olfert, O., Carcamo, H.A., 2002. Temporal and geographical distribution patterns of cabbage seedpod weevil (Coleoptera : Curculionidae) in canola. Canadian Entomologist 134, 403-418.

Duchesne, P., Bernatchez, L., 2002. AFLPOP: a computer program for simulated and real population allocation, based on AFLP data. Molecular Ecology Notes 2, 380-383.

Durka, W., Bossdorf, O., Prati, D., Auge, H., 2005. Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Molecular Ecology 14, 1697-1706.

Forister, M.L., Ehmer, A.G., Futuyma, D.J., 2007. The genetic architecture of a niche: variation and covariation in host use traits in the Colorado potato beetle. Journal of Evolutionary Biology 20, 985-996.

Gerber, E., Cortat, G., Hinz, H.L., Blossey, B., Katovich, E., Skinner, L., 2009. Biology and host specificity of *Ceutorhynchus scrobicollis* (Curculionidae; Coleoptera), a root-crown mining weevil proposed as biological control agent against *Alliaria petiolata* in North America. Biocontrol Science and Technology 19, 117-138.

Gerber, E., Hinz, H., Schat, M., Cortat, G., Guazzone, N., Lovis, L., Biological control of garlic mustard, *Alliaria petiolata*, Annual Report 2002. Unpublished Report. CABI Bioscience Switzerland Centre, Delemont, Switzerland, 2003.

Goolsby, J.A., De Barro, P.J., Makinson, J.R., Pemberton, R.W., Hartley, D.M., Frohlich, D.R., 2006a. Matching the origin of an invasive weed for selection of a herbivore haplotype for a biological control programme. Molecular Ecology 15, 287-297.

Goolsby, J.A., van Klinken, R.D., Palmer, W.A., 2006b. Maximising the contribution of native-range studies towards the identification and prioritisation of weed biocontrol agents. Australian Journal of Entomology 45, 276-286.

Hawthorne, D.J., 2001. AFLP-based genetic linkage map of the Colorado potato beetle *Leptinotarsa decemlineata*: Sex chromosomes and a pyrethroid-resistance candidate gene. Genetics 158, 695-700.

Ireson, J.E., Holloway, R.J., Chatterton, W.S., 2008. The influence of host plant genotype on variation in population densities of the gorse thrips, *Sericothrips staphylinus* (Thysanoptera: Thripidae), and its consideration in relation to release strategies. Biocontrol Science and Technology 18, 949-955.

Jonsen, I.D., Bourchier, R.S., Roland, J., 2007. Influence of dispersal, stochasticity, and an Allee effect on the persistence of weed biocontrol introductions. Ecological Modelling 203, 521-526.

Komatsu, T., Akimoto, S., 1995. Genetic Differentiation as a Result of Adaptation to the Phenologies of Individual Host Trees in the Galling Aphid *Kaltenbachiella japonica*. Ecological Entomology 20, 33-42.

Louda, S.M., Pemberton, R.W., Johnson, M.T., Follett, P.A., 2003. Nontarget effects - The Achilles' Heel of biological control? Retrospective analyses to reduce risk associated with biocontrol introductions. Annual Review of Entomology 48, 365-396.

Lynch, M., Milligan, B.G., 1994. Analysis of population genetic-structure with RAPD markers. Molecular Ecology 3, 91-99.

Manguin, S., White, R., Blossey, B., Hight, S.D., 1993. Genetics, Taxonomy, and Ecology of Certain Species of *Galerucella* (Coleoptera, Chrysomelidae). Annals of the Entomological Society of America 86, 397-410.

McFadyen, R.E.C., 1998. Biological control of weeds. Annual Review of Entomology 43, 369-393.

Meekins, J.F., Ballard, H.E., McCarthy, B.C., 2001. Genetic variation and molecular biogeography of a North American invasive plant species (*Alliaria petiolata*, Brassicaceae). International Journal of Plant Sciences 162, 161-169.

Merila, J., Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology 14, 892-903.

Messenger, P.S., van den Bosch, R., 1971. The adaptability of introduced biological agents. In: Huffaker, C.B., (Ed.), Biological Control. Plenum Press, New York, pp. 68-92.

Messenger, P.S., Wilson, F., Whitten, M.J., 1976. Variation, fitness, and adaptability of natural enemies. In: Huffaker, C.B., Messenger, P.S., Eds., Theory and practice of biological control. Academic Press, New York, pp. 209-231.

Milbrath, L.R., DeLoach, C.J., 2006. Host specificity of different populations of the leaf beetle *Diorhabda elongata* (Coleoptera : Chrysomelidae), a biological control agent of saltcedar (*Tamarix* spp.). Biological Control 36, 32-48.

Miller, J.R., Strickler, K., 1984. Plant-herbivore relationships: Finding and accepting host plants. In: Bell, W.J., Carde, R.T., Eds., Chemical Ecology of Insects. Chapman & Hall, London, pp. 127-157.

Mopper, S., 1996. Adaptive genetic structure in phytophagous insect populations. Trends in Ecology & Evolution 11, 235-238.

Mopper, S., Beck, M., Simberloff, D., Stiling, P., 1995. Local adaptation and agents of selection in a mobile insect. Evolution 49, 810-815.

Myers, J.H., Sabath, M.D., Genetic and phenotypic variability, genetic variance, and the success of establishment of insect introductions for the biological control of weeds. In: Del Fosse, E.(Ed.), Proceedings of the Fifth international Symposium on Biological Control of Weeds. Commonwealth Scientific and Industrial Research Organization, Brisbane, Australia, 1981, pp. 649.

Nei, M., 1972. Genetic Distance between Populations. American Naturalist 106, 283-292.

Nowierski, R.M., McDermott, G.J., Bunnell, J.E., Fitzgerald, B.C., Zeng, Z., 1996. Isozyme analysis of *Aphthona* species (Coleoptera: Chrysomelidae) associated with different *Euphorbia* species (Euphorbiaceae) and environmental types in Europe. Annals of the Entomological Society of America 89, 858-868.

Nuzzo, V., 1993. Distribution and spread of the invasive biennial *Alliaria petiolata* (garlic mustard) in North America. In: McKnight, B.L., (Ed.), Biological Pollution: control and impact of invasive exotic species. Indiana Academy of Science, Indianapolis, pp. 115-124.

Olivieri, I., Singer, M.C., Magalhaes, S., Courtiol, A., Dubois, Y., Carbonell, D., Justy, F., Beldade, P., Parmesan, C., Michalakis, Y., 2008. Genetic, ecological, behavioral and geographic differentiation of populations in a thistle weevil: implications for speciation and biocontrol. Evolutionary Applications 1, 112-128.

Paetkau, D., Calvert, W., Stirling, I., Strobeck, C., 1995. Microsatellite analysis of population-structure in canadian polar bears. Molecular Ecology 4, 347-354.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Pearson, D.E., Callaway, R.M., 2005. Indirect nontarget effects of host-specific biological control agents: Implications for biological control. Biological Control 35, 288-298.

Peterson, M.A., Denno, R.F., 1998. Life history strategies and the genetic structure of phytophagous insect populations. In: Mopper, S., Strauss, S., Eds., Genetic structure and local adaptation in natural insect populations: Effects of ecology, life history, and behavior. Chapman & Hall, New York, pp. 263-322.

Phillips, C.B., Baird, D.B., Iline, II, McNeill, M.R., Proffitt, J.R., Goldson, S.L., Kean, J.M., 2008. East meets west: adaptive evolution of an insect introduced for biological control. Journal of Applied Ecology 45, 948-956.

Raffl, C., Schonswetter, P., Erschbamer, B., 2006. 'Sax-sess' - genetics of primary succession in a pioneer species on two parallel glacier forelands. Molecular Ecology 15, 2433-2440.

Rajpurohit, S., Parkash, R., Ramniwas, S., Singh, S., 2008. Variations in body melanisation, ovariole number and fecundity in highland and lowland populations of *Drosophila melanogaster* from the Indian subcontinent. Insect Science 15, 553-561.

Reed, D.H., Frankham, R., 2003. Correlation between fitness and genetic diversity. Conservation Biology 17, 230-237.

Rosen, D., 1986. The role of taxonomy in effective biological control programs. Agriculture Ecosystems & Environment 15, 121-129.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Sands, D.P.A., Harley, K.L.S., Importance of geographic variation in agents selected for biological control of weeds. In: Del Fosse, E.S., (Ed.), Proceedings of the Fifth International Symposium on Biological Control of Weeds. Commonwealth Scientific and Industrial Research Organization, Brisbane, Australia, 1981, pp. 649.

Schott, C., 2000. Catalogue et atlas des Coléoptères d'Alsace. Musee Zoologique de l'Université et de la Ville de Strasbourg, Strasbourg, F.

Sheppard, A.W., van Klinken, R.D., Heard, T.A., 2005. Scientific advances in the analysis of direct risks of weed biological control agents to nontarget plants. Biological Control 35, 215-226.

Simberloff, D., Stiling, P., 1996. Risks of species introduced for biological control. Biological Conservation 78, 185-192.

Slatkin, M., 1987. Gene Flow and the Geographic Structure of Natural Populations. Science 236, 787-792.

Stinson, K., Kaufman, S., Durbin, L., Lowenstein, F., 2007. Impacts of garlic mustard invasion on a forest understory community. Northeastern Naturalist 14, 73-88.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. Cladogram Estimation. Genetics 132, 619-633.

Unruh, T.R., Woolley, J.B., 1999. Molecular methods in classical biological control. In: Bellows, T., Fisher, T.W., Eds., Handbook of Biological Control. Academic Press, San Diego, pp. 57-85.

Vekemans, X., Beauwens, T., Lemaire, M., Roldan-Ruiz, I., 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology 11, 139-151.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP - a new

technique for DNA-fingerprinting. Nucleic Acids Research 23, 4407-4414.

Whitlock, M.C., McCauley, D.E., 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. Heredity 82, 117-125.

Zhivotovsky, L.A., 1999. Estimating population structure in diploids with multilocus dominant DNA markers. Molecular Ecology 8, 907-913.

Table 1. Sites sampled for *Ceutorhynchus scrobicollis* in Europe, number of individuals used for AFLP and COI sequence analyses, and the different haplotypes recorded at each site. Sites 10-19 are German (DE) sites.

				AFLP	COI Sequence	
Site	**Location**	**Latitude**	**Longitude**	**n**	**n**	**Haplotypes**
10	W Berlin	52.4672	12.9271	0	2	A,C
11	NW Berlin	52.6140	13.0056	26	8	A,C,D
12	NW Berlin	52.6777	13.1334	13	3	A
13	N Berlin	52.6786	13.3757	21	1	A
14	E Berlin	52.4736	13.7598	10	2	A
15	E Berlin	52.4116	13.8256	33	2	A
16	S Potsdam	52.2989	13.0250	29	3	A,C
17	SW Berlin	52.4364	13.1830	19	2	C,D
18	SW Berlin	52.4290	13.1920	27	4	A
19	W Klieken	51.8932	12.3355	17	5	A
Romania (RO)	near Iasi	47.17	27.58	54	4	B,F
Georgia (GE)	near T'bilisi	41.7	44.8	11	7	E

Table 2. Results of re-allocation assignment tests. Left column for each site (LL1) gives the number of individuals and where they were assigned to with the minimal log-likelihood difference set to one (i.e. where individuals were 10 times more likely to have come from). The center column for each site (%) shows the percentage of the total site sample (n) that was reallocated to each site. For example, of the 26 individuals from Site 11, 73% (19 inds.) were ten times more likely to have come from Site 11, where they were collected, than from any other site. One individual (4%) was ten times more likely to have come from Site 13, and 6 individuals (23%) where not reallocated because they were not 10 times more likely to have come from any particular site relative to the others.

	From site:																					
	11		12		13		14		15		16		17		18		19		RO		GE	
Reallocated to site:	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%
11	19	73																				
12			9	69																		
13	1	4			11	52																
14							1	10	3	9					1	6						
15									7	21												
16											21	72					1	6				
17									1	3			6	32								
18			1	8	1	5									10	6						
19																	14	82				
RO																			54	100		
GE																					11	100
None	6	23	3	23	9	43	9	90	22	67	8	28	13	68	16	89	2	12	0	0	0	0

Table 3. Pairs of sites linked by migration inferred from to assignment tests. The first site number is the source, the second is the recipient. The number of individuals assigned to sites other than where they were collected and the minimum distance to move between sites are indicated.

Site pair	Number of individuals	Distance (km)
11-13	1	26
12-18	1	28
13-18	1	30
14-18	1	39
15-14	3	8
15-17	1	44
16-19	1	65
	Mean distance	28.4

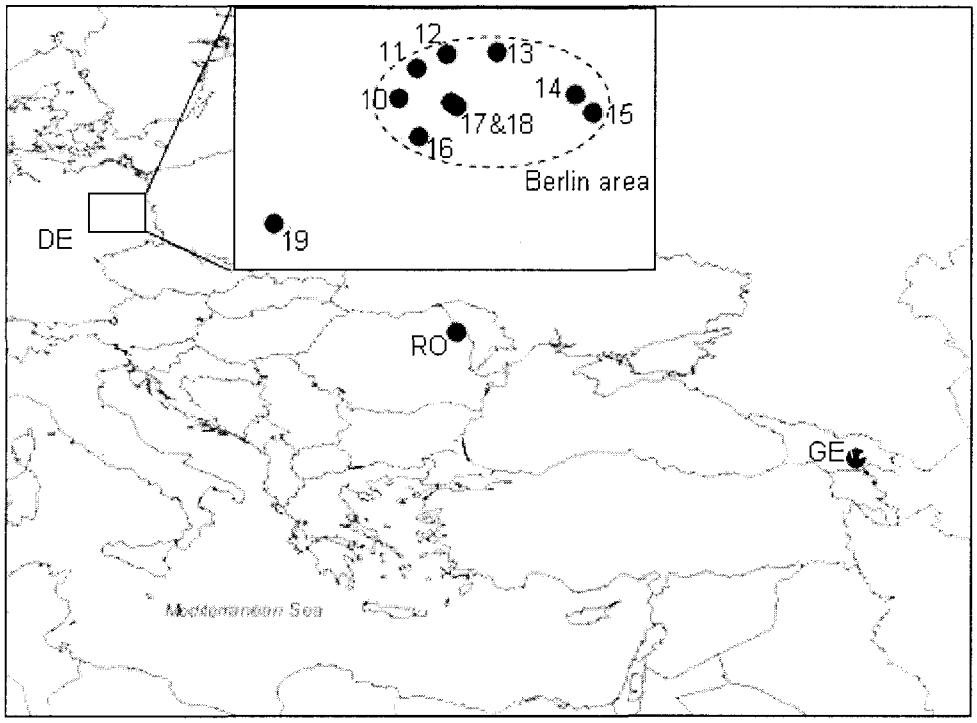

Figure 1. Map of the *Ceutorhynchus scrobicollis* collection sites showing the focal area around Berlin, Germany (DE) where primary collections for host-specificity tests were made, and sites from Romania (RO) and Georgia (GE).

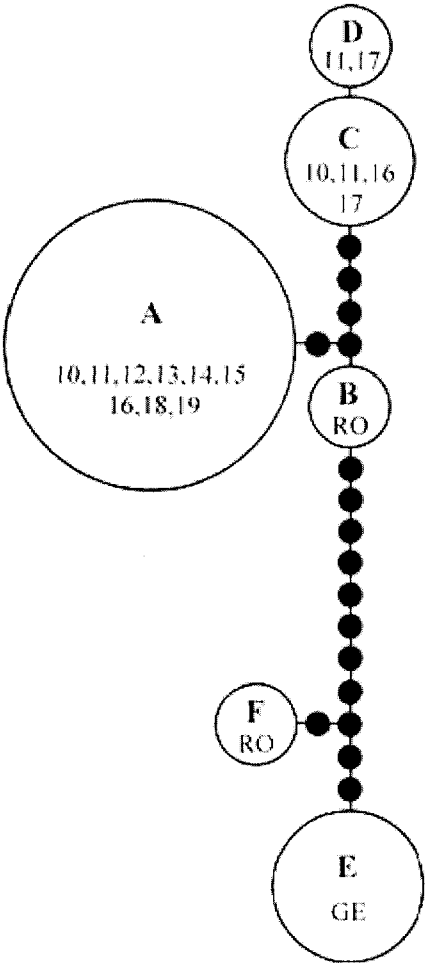

Figure 2. Haplotype network for *Ceutorhynchus scrobicollis* individuals analyzed showing the relationship of the 6 different haplotypes of the 592 bp mtDNA COI region. Haplotypes, labeled A through F, are designated by open circles. The relative sizes of the circles indicate the relative frequency of the haplotype in the sample. Small black circles indicate inferred intermediate haplotypes that were not found in the sample. Lines between circles indicate mutations.

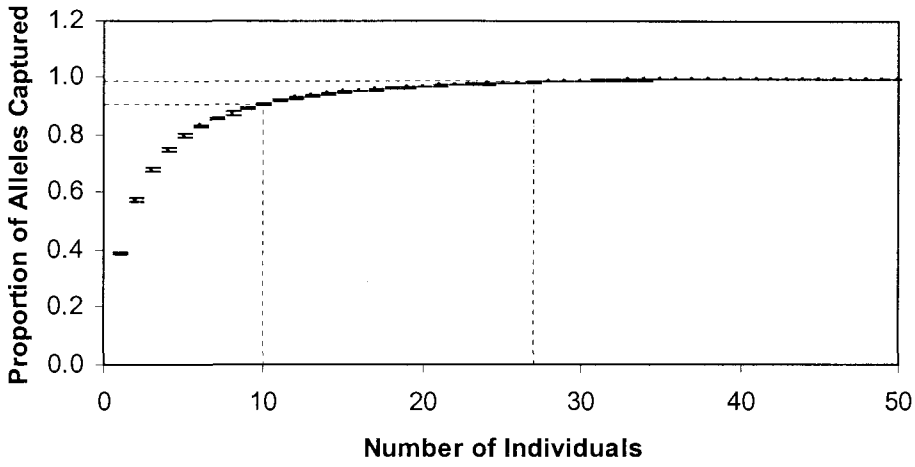

Figure 3. Relationship between the number of individuals randomly sampled and the proportion of alleles captured for *Ceutorhynchus scrobicollis* in the vicinity of Berlin, Germany (sites 10-18). Data shown are averages ± SE from 100 permutations. Dashed lines indicate the number of *Ceutorhynchus scrobicollis* individuals that need to be sampled to capture 90 and 99% of alleles present in at least five percent of individuals sampled.

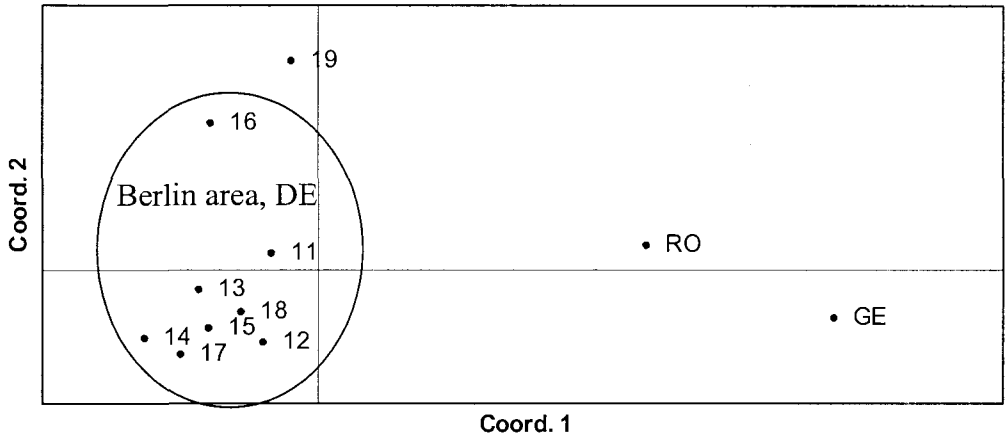

Figure 4. Principal coordinates analysis based on Nei's (1972) pairwise genetic distance measures among all sites. This two dimensional plot captures 82% of the variation from the multidimensional space generated by comparing genetic distances among all pairwise combinations of sites. Axis 1 captured 63% of the variation and axis 2 captured 19%. Sites 11-18 are from the Berlin, DE area (oval), and site 19 is ~100 km to the southwest of central Berlin. RO denotes Romania and GE denotes Georgia.

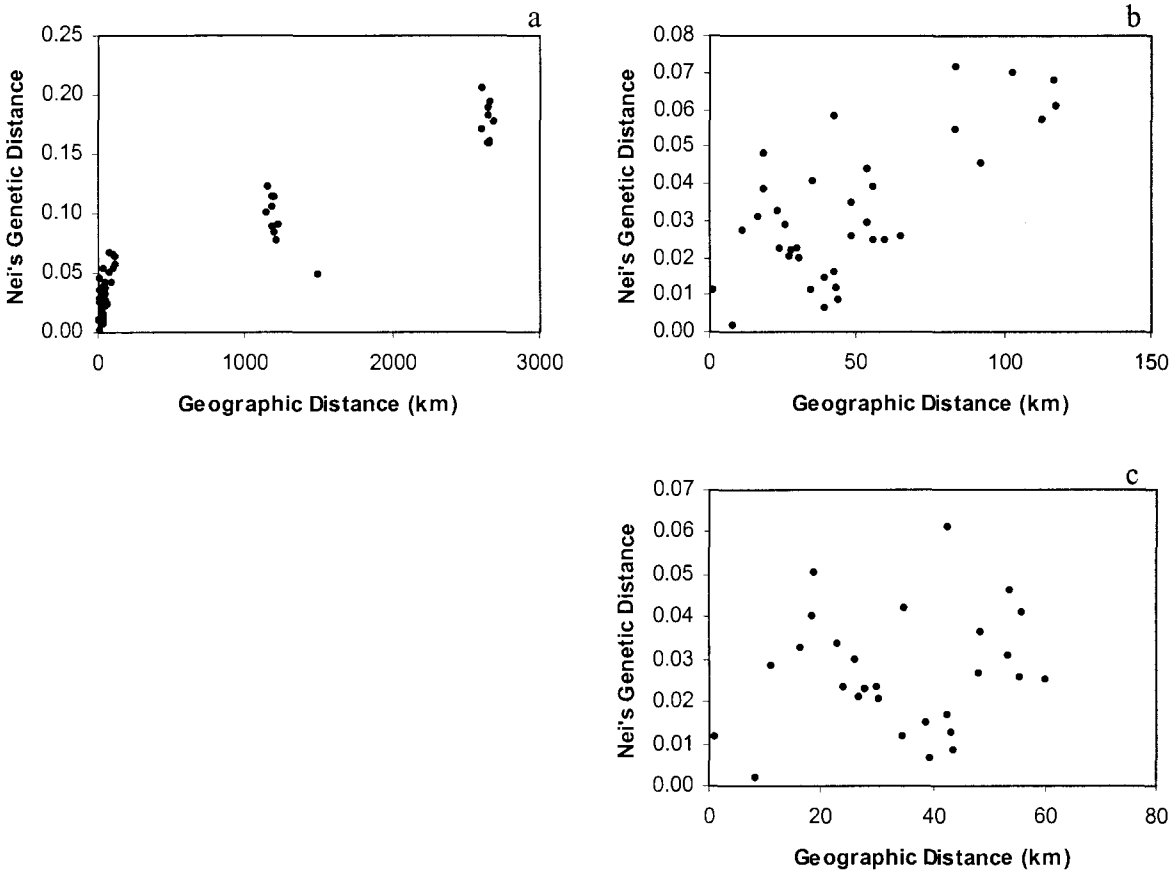

Figure 5. The relationships between geographic and genetic distance of *C. scrobicollis* at three spatial scales. When all sites are included (a) or only sites within Germany (b), there is significant isolation by distance. When the analysis is restricted to sites within the Berlin area only (c), there is no significant isolation by distance. See text for statistical details.

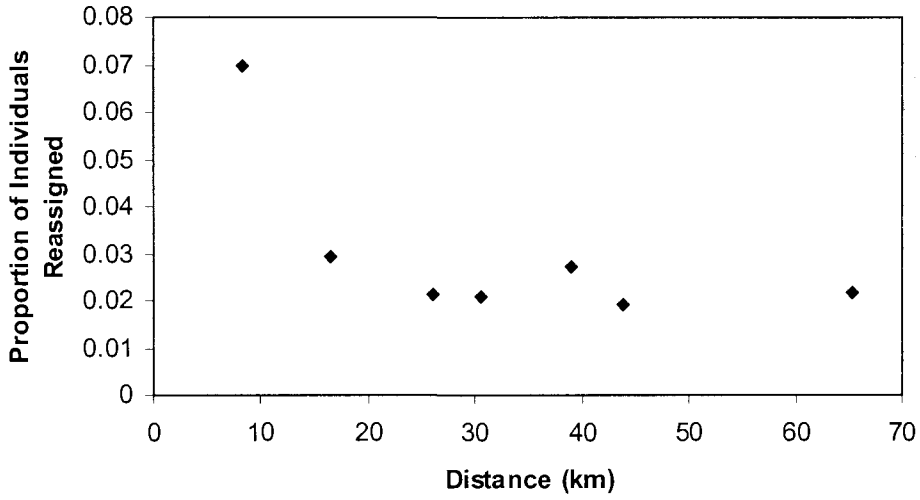

Figure 6. The relationship between dispersal events inferred from assignment tests and the distance between sites.

APPENDIX

Appendix A. Nei's pairwise genetic distances after Lynch and Milligan (1994) are above the diagonal. Pairwise geographic distance (km) are below the diagonal. Genetic distances are significant ($P<0.05$) unless otherwise denoted (n.s.). Sites 11-19 are from Germany.

	Site 11	Site 12	Site 13	Site 14	Site 15	Site 16	Site 17	Site 18	Site 19	Romania	Georgia
Site 11	—	0.025	0.027	0.028	0.023	0.038	0.030	0.021	0.042	0.077	0.159
Site 12	11	—	0.028	0.024	0.023	0.054	0.019	0.020	0.064	0.083	0.160
Site 13	26	16	—	0.010	0.015	0.032	0.021	0.018	0.053	0.105	0.181
Site 14	53	48	35	—	0.002 n.s.	0.041	0.006	0.014	0.063	0.123	0.205
Site 15	60	56	43	8	—	0.036	0.008	0.011	0.056	0.100	0.171
Site 16	35	43	48	54	56	—	0.045	0.036	0.024	0.113	0.193
Site 17	23	27	30	39	44	19	—	0.010	0.066	0.114	0.188
Site 18	24	28	30	39	43	18	1	—	0.050	0.089	0.158
Site 19	92	103	113	117	117	65	84	84	—	0.090	0.176
Romania	1207	1202	1187	1154	1147	1191	1188	1187	1219	—	0.049
Georgia	2669	2663	2647	2616	2610	2660	2653	2652	2696	1495	—

CHAPTER 3

Comparative population genetics of *Ceutorhynchus alliaria* and *C. roberti*, two candidate biological control agents of garlic mustard.

INTRODUCTION

When developing new biological control programs, scientists can be faced with choosing between two or more candidate agents with similar life histories, host-ranges, and effects on the target pest. In such cases, other factors, like distribution or ease of rearing may play a role in determining which to petition for introduction. Rarely is information on the population genetics of candidate agents known prior to release. However, such information can potentially facilitate choices between agents by providing information on genetic diversity within populations and dispersal between populations.

Genetic diversity is hypothesized to play a role in the ability of natural enemies to adapt to the novel environment of the introduced range (Messenger and van den Bosch, 1971; Messenger et al., 1976; Myers and Sabath, 1981). While few studies have explored natural enemy genetic diversity pre-release, and no studies, to our knowledge, have tested the effect of different levels of genetic diversity on agent establishment and success, a meta-analysis has shown that population genetic diversity is significantly correlated with fitness (Reed and Frankham, 2003).

Furthermore, high genetic variation is associated with increased population growth rates and higher probability of successfully founding of new populations (Hanski and Saccheri, 2006; Kephart, 2004; Newman and Pilson, 1997; Reed et al., 2007; Vergeer et al., 2005). Thus, knowing which species have higher genetic diversity may aid in choosing between candidate agents that are otherwise biologically similar.

Dispersal ability, while difficult to measure with traditional methods, is also useful to know when choosing among candidate agents. Studying the population structure of candidate agents can provide insight into their comparative dispersal abilities. Theoretically, high rates of dispersal can increase Allee effects following release (e.g. Robinet et al., 2008) but also may speed the geographic spread of biological control agents. Data illustrating that increasing the number of individuals released contributes to successful establishment (Drake et al., 2005; Drake and Lodge, 2006; Grevstad, 1999a; Grevstad, 1999b; Memmott et al., 2005) suggest that Allee effects may be common in biological control introductions. In a system where Allee effects are likely to be strong (e.g. agents feed gregariously to overcome host defenses), then an agent with lower dispersal abilities might be advantageous. However, if Allee effects are not likely to be particularly strong (e.g. the agents are easy to rear and release in large numbers), then those agents with high dispersal abilities should be favored, as subsequent to establishment redistribution efforts would not need to be as extensive.

Here we present data on the population genetic structure of *Ceutorhynchus alliariae* Brisout and *C. roberti* Gyllenhal (Curculionidae), two insects being considered for the biological control of garlic mustard, *Alliaria petiolata* (Bieb)

Cavara and Grande. Both weevils have similar life histories and effects on their host, and their geographical distributions overlap (Gerber et al., 2008), making choosing between them when petitioning for release difficult. Our objective was to determine the within and among population genetic diversity and dispersal potential of these two species and to determine whether this information might help to prioritize one of the agents.

Adults of both species lay eggs in the stems or petioles of garlic mustard plants as the plants begin bolting in the spring (Blossey et al., 2001). Larvae of both species are endophagous, feeding primarily in the stems of their host. At the completion of larval development in late May both species leave the host to pupate in the soil. Adults of both species begin emerging in June and emergence continues through August. Both species are highly host specific. Both species have been found occupying the same garlic mustard stem in the field (Hinz and Gerber, 2001). Studies of larval competition between the two species have shown that while the number of larvae per stem affects survival (with reduced survival at high densities), the species composition of the larvae in the stem does not (Gerber et al., 2008).

METHODS

Collections. Sampling of *Ceutorhynchus alliariae* and *C. roberti* was performed in the spring of 2003 and concentrated mainly around Switzerland (CH) and Germany (DE) where most of the collections for garlic mustard biological control agents have been made (Table 1). Specifically, host-specificity testing for both *C. roberti* and *C.*

alliariae has primarily involved the use of individuals from around the Delemont, CH area and southern Germany, and for *C. alliariae*, around Berlin, DE as well (H. Hinz and E. Gerber pers. comm.). Additional collections were made in Austria (AT) and northern Italy (IT), and from collaborators we received specimens in 2004 from Bulgaria (BG), France (FR), Romania (RO), and the United Kingdom (UK). For all collections, garlic mustard plants were sampled at one meter intervals along a transect through the long axis of the population. Larvae were dissected from the stems and placed immediately in 100% ethanol to preserve DNA. Genomic DNA was extracted from a single larva per plant using DNeasy Tissue Kits (Qiagen®, Germantown, MD, USA) following the manufacturer's instructions. Larvae of *C. alliariae* and *C. roberti* were identified using a polymerase chain reaction – restriction fragment length polymorphism (PCR-RFLP) technique described in Rauth and Hufbauer (submitted, Ch. 1).

MtDNA sequencing. We sequenced a portion of mtDNA encompassing approximately 850 bp of COI, tRNALeu, and a short section of COII to explore intraspecific diversity and among site variation in this region for *C. alliariae* and *C. roberti*. Between 1 and 10 individuals from each site were sequenced. The sequencing protocol is described in detail in Rauth and Hufbauer (submitted, Ch. 1).

Amplified fragment length polymorphisms. Amplified fragment length polymorphisms (AFLPs) were used to assess genetic variation among sites where *C. scrobicollis* was collected. We followed the protocol of Vos et al. (1995) with

modifications as outlined below. Adapters and primer cores followed Hawthorne (2001). We used two enzymes, each with a six base pair restriction site: *Eco*RI and *Pst*I (New England Biolabs®, Germantown, MD, USA). Four selective primer pair combinations were used during the selective amplification PCR: (**Eco*RI-AGTG/*Pst*I-GAC; ^*Eco*RI-CGAT/*Pst*I-GAC; ^*Eco*RI-GACT/*Pst*I-AGT; **Eco*RI-TGAG/*Pst*I-AGT, where * denotes the 5' end was labeled with 6-FAM™, ^ denotes the 5' end was labeled with VIC™).

Genomic DNA was digested and adapters were ligated to the resulting fragments in a single two hour incubation step at 37°C. Digestion and ligation were performed in 20μL reaction volumes of 0.5x NEBuffer 3 (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol, pH 7.9 @ 25°C), 0.5x T4 DNA Ligase Reaction Buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM Dithiothreitol, 25 μg/ml BSA, pH 7.5 @ 25°C), 0.5μg BSA, 4.5pmol of each adapter, 5 U EcoRI (in 300 mM NaCl, 10 mM 2-Mercaptoethanol, 0.1 mM EDTA, 200 μg/ml BSA, 50% Glycerol, 0.15% Triton X-100, pH 7.5 @ 25°C), 5 U PstI (in 10 mM Tris-HCl, 200 mM NaCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 200 μg/ml BSA, 50% Glycerol, 0.15% Triton X-100, pH 7.4 @ 25°C), 66.8 U T4 DNA ligase (in 10 mM Tris-HCl, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, 50% Glycerol, pH 7.4 @ 25°C), and 8μL genomic DNA. The product from digestion and ligation was diluted with 100μL $TE_{0.1}$. For the pre-amplification PCR *Eco*RI and *Pst*I primers with three selective nucleotides were used (e.g. 5' *Eco*RI-AGT/*Pst*I-GAC). The pre-amplification step was performed in 10μL reaction volumes of 1x PCR buffer (20mM Tris-HCL, pH 8.4, 50mM KCl), 2mM $MgCl_2$, 0.2mM dNTPs, 1

U *Taq* polymerase, 1.8pmol of each primer, and 1μL of diluted digestion/ligation product. Pre-amplification PCRs were run on a Mastercycler gradient (Eppendorf AG, Westbury, NY, USA) thermal cycler starting at 72°C for 2min, followed by 20 cycles of 20s at 94°C, 30s at an annealing temperature of 56°C, 2min at 72°C for polymerase extension, and a final 30 minutes at 60°C. After pre-amplification, PCR product was diluted with 140μL $TE_{0.1}$. The selective amplification step was performed with the same reaction constituents as for the pre-amplification except for the change of *Eco*RI primers. After an initial 2min denaturation period, reactions underwent 30 cycles of 20s at 94°C, 30s at 66°C (decreasing by 1°C each cycle until at 56°C), 2min at 72°C, with a final 30min at 60°C. One microliter of selective amplification product was then added to 10μL HiDi-Formamide and 0.13μL GeneScan™ LIZ® 500 size standard (Applied Biosystems Inc., Foster City, CA, US). Product was then denatured at 95°C for 3 minutes and placed immediately on ice prior to being run on an ABI 3100 capillary instrument (Applied Biosystems Inc.). Electropherograms were visualized using GeneMarker ver.1.51 (SoftGenetics LLC, State College, PA, USA). Loci were scored individually by hand for presence/absence of a peak. Only polymorphic loci that could be scored unambiguously across the data set were used in analyses. Scoring was performed without reference to the identity of the individual or site of origin. Only loci found in more than 5% or less than 95% of the individuals were used in analyses.

MtDNA sequence analysis. To visualize the relationship and diversity among haplotypes for *C. alliariae* and *C. roberti*, we constructed a maximum parsimony

haplotype network using the statistical algorithms of Templeton et al. (1992) implemented in the program TCS 1.2.1 (Clement et al., 2000). To further explore the differences between the two species we also calculated haplotype diversity, nucleotide diversity (π), and the average number of pairwise nucleotide differences (k) using the software DNASP v. 4.50.2 (Rozas et al., 2003). Theta (θ) per site, which estimates the number of differences per nucleotide expected between two randomly chosen alleles under mutation-drift equilibrium, was also calculated.

AFLP analyses. We tested the independence of markers by calculating their pairwise linkage index $D_{A,B} = \frac{\sum i|V(A,i) - V(B,i)|}{n}$. $V(A,i)$ is the allele value for marker A of individual i, $V(B,i)$ is the allele value for marker B of individual i, and n is the total number of individuals analyzed. D-values were calculated using the software DDM version 0.1 (P. Berthier, available upon request).

Allele frequencies, which are used to generate many common population genetic statistics, must be estimated for dominant AFLP data. This is because of the dominant nature of the markers, which prevents distinguishing individuals that are homozygous for an allele from heterozygous individuals. We assumed Hardy-Weinberg equilibrium and then used Zhivotovsky's Bayesian method (Zhivotovsky, 1999) with a non-uniform prior distribution as implemented in AFLPsurv 1.0 (Vekemans et al., 2002) to estimate allele frequencies. To evaluate the population structure of *C. alliariae* and *C. roberti,* estimated allele frequencies were used to calculate genetic diversity and Nei's (1972) genetic distance following the approach

of Lynch and Milligan (1994), as implemented in AFLPsurv 1.0. Significance of genetic distances among pairs of sites was determined by bootstrapping over loci. Confidence intervals were based on bootstrap percentiles generated from 10,000 permutations. We evaluated whether sites were isolated by distance, using a Mantel test as implemented in the software IBD ver. 1.52 (Bohonak, 2002). For *C. alliariae* we examined both the full data set and a subset of the data that was geographically comparable to the *C. roberti* data (samples from CH and DE only).

To evaluate the likelihood that sampled individuals had dispersed from other locations, population assignment tests were performed using AFLPOP 1.1 (Duchesne and Bernatchez, 2002). This assignment method computes the likelihood that an individual with a given genotype be found in each of the candidate populations based on the allele frequencies of the different populations. We used the re-allocation (leave-one-out) procedure with the allocation minimal log-likelihood difference set to one. With this procedure an individual is removed from its collection site, allele frequencies are calculated for all collection sites, and then the likelihood that the individual came from each of the sites is computed. With the allocation log-likelihood minimal difference set to one, for an individual to be assigned to any collection site, it had to be ten times more likely to have come from that site than any other site. By calculating the distances between sites where individuals were collected and where they were reassigned we were able to evaluate dispersal. Comparison of dispersal estimates from assignment tests with long-term mark-and-recapture records has shown that assignment tests can detect dispersal accurately even when there is relatively little population genetic structure (Berry et al., 2004). Because the

individuals sampled are larvae, and cannot themselves disperse between sites, the dispersal we infer is due to dispersal of parents. Thus we are more likely to detect dispersal events in offspring from females that mated prior to dispersal than in offspring from dispersing males or females that mated after dispersing.

RESULTS

Haplotype Networks and Sequence Diversities. We obtained sequence from 57 individuals (28 from CH and DE area only) of *C. alliariae* and 60 individuals (53 from CH and DE area only) of *C. roberti*. Seven haplotypes were found for *C. alliariae* (Accession numbers EU551168, EU551171, EU852377 - 852381) and 18 haplotypes were found for *C. roberti* (Accession numbers EU551167, EU551170, EU852382 - 852397). None of the haplotypes were shared between the two species. For *C. alliariae*, haplotype A was the most common haplotype and was found in 45 individuals representing all sites except for site 13 (Table 1, Figure 1). Haplotype B was found in one individual each from sites 14 and 16. Haplotypes D, E, F and G differed from haplotype A by a single nucleotide and were found in individuals from sites 5, 15, 10, and AT, respectively. Haplotype C differed from A by two nucleotides and was found in a single individual from site 13. In *C. roberti*, haplotype H was the most common haplotype and was found in 27 individuals from sites 1, 2, 3, 4, 5, 6, 7,

21, 22, 23, and 24. Haplotypes K, from sites 1 and 24, and P, from sites 1, 5, and 23, were the next most common haplotypes and were found in six individuals each. All other haplotypes were found in 4 or fewer individuals. For the full data sets, the average number of nucleotide differences (k) was 0.44 for *C. alliariae* and 2.33 for *C. roberti* (Table 2). For the geographically comparable data sets encompassing the region (CH & DE) where collections have been made for host specificity testing and where both species were found, the average number of nucleotide differences was 0.66 and 2.04 for *C. alliariae* and *C. roberti*, respectively. For both the full and comparable datasets, haplotype diversity, nucleotide diversity, and theta per site were all higher in *C. roberti* than *C. alliariae*, as well (Table 2).

AFLP results. Using four primer combinations we scored 80 polymorphic loci for 183 individuals from 10 populations of *C. alliariae* and 47 polymorphic loci for 174 individuals from 9 populations of *C. roberti*. The mean pairwise linkage index was 0.48 and 0.50 for *C. alliariae* and *C. roberti*, respectively. No marker pairs had *D*-values lower than 0.1 or higher than 0.9 for either species. These values indicate that all markers were unlinked and, therefore, all were retained for further analyses.

Over all sites, total gene diversity (H_T) was 0.35 for *C. alliariae* and 0.40 for *C. roberti*. In *C. alliariae,* 88% ($H_W = 0.31 \pm 0.02$) of the total diversity was found within sample sites and 12% was due to genetic differentiation among sample sites ($F_{ST} = 0.12$). In *C. roberti,* 99% ($H_W = 0.39 \pm 0.00$) of the total diversity was found within sample sites and only 1% was due to genetic differentiation among sample sites ($F_{ST} = 0.01$). When we again paired down the *C. alliariae* data set to include just

Swiss and German sites, like what we had for *C. roberti*, total gene diversity was 0.36, 99% ($H_W = 0.35 \pm 0.00$) of which was found within sample sites. Total gene diversity in the region of Switzerland and Germany was significantly higher in *C. roberti* than in *C. alliariae* ($t_{12} = -7.1$, $p < 0.001$).

Among-population differentiation was low for both *C. alliariae* and *C. roberti*. Nei's genetic distances ranged from 0.000 between sites 14 and 15, to 0.021 between sites 1 and 20, in the region of CH and DE, or to 0.166 between sites FR and AT in the full *C. alliariae* data set (Table 3). For *C. roberti*, Nei's genetic distances ranged from 0.000 between sites 5 and 6, and 22 and 24, to 0.028 between sites 21 and 23. For both species, less than 45% of the pairwise comparisons of genetic distances between sites were significant in the region of CH and DE.

Mantel tests to determine whether genetic distances between sites were correlated with geographic distances between sites revealed that neither *C. alliariae* (both full and CH and DE only data sets) nor *C. roberti* were isolated by distance (*C. alliariae*: $Z = 1766$, $r = 0.23$, $P = 0.21$; *C. alliariae* CH and DE only: $Z = 49$, $r = 0.05$, $P = 0.39$; *C. roberti*: $Z = 57$, $r = 0.22$, $P = 0.16$; Figure 2).

Assignment test results indicated that both *C. alliariae* and *C. roberti* are highly vagile within the region of Switzerland (CH) and Germany (DE) (Table 4). The mean percentage of individuals reassigned to their collection site in CH and DE was 28% for *C. alliariae* and 7% for *C. roberti*. The mean percentage of individuals that were not assigned to any site in this region was 55% for *C. alliariae* and 81% for *C. roberti*. However, for *C. alliariae* outside the region of CH and DE the mean

percentage of individuals reassigned to their collection site was 99%. These sites included Italy (IT), Austria (AT), France (FR), and United Kingdom (UK).

DISCUSSION

C. alliariae and *C. roberti* are similar ecologically, and we collected data to evaluate whether those similarities are reflected in population genetic data, or whether there are differences that could be used to aid in prioritizing one over the other for biological control. We found that the species are quite similar in their levels of population genetic variation and their dispersal potential. Overall both our AFLP and mtDNA sequence data suggest that *C. roberti* harbors slightly greater genetic diversity than *C. alliariae*.

Most of the genetic diversity observed in both species was found within populations. Similar patterns have been found in other curculionids (Briese et al., 1996; Conord et al., 2006; Kerdelhue et al., 2002; Li et al., 2008). Such large proportions of genetic variation within populations can be associated with recent range expansions. For example, such patterns were found in the cotton boll weevil (*Anthonomus grandis*) in North America (Kim and Sappington, 2004) and in a the large pine weevil (*Hylobius abietis*) a European forest pest (Conord et al., 2006). Garlic mustard was first recorded in N.A. in 1868 (Nuzzo, 1993), yet molecular analysis of plants from both the native and introduced range show similar levels of population structure in both continents (Durka et al., 2005). This suggests the

possibility that garlic mustard may have recently expanded its range in Europe. If this is the case, it could account for the pattern of high within-population genetic diversity observed in *C. alliariae* and *C. roberti*.

Ceutorhynchus alliariae and *C. roberti* also show similar, and considerable, dispersal capabilities. Among population differentiation in *C. roberti* was nearly identical to that observed in the comparable *C. alliariae* data set, and although F_{ST} for the full *C. alliariae* data set was higher, levels of population differentiation were still low. This may reflect high migration rates among populations, either natural or facilitated by human transport. Distances between collection sites and sites where individuals were reassigned suggest that, based on sampled individuals, *C. alliariae* and *C. roberti* are dispersing an average of 341 km and 107 km, respectively. While individuals of both species may live for 3 years, and dispersal estimates of other curculionids have suggested average dispersals of 55 km yr^{-1} (*Ceutorhynchus obstrictus* (Dosdall et al., 2002)) and 95 km yr^{-1} (*Anthonomus grandis* (Culin et al., 1990)). There are three reasons that suggest these large distances may not be accurate estimates. First, as stated above, if garlic mustard populations are still expanding, then it is likely that both *C. allairiae* and *C. roberti*, like their host, are not in demographic equilibrium. Second, relatively low sample sizes for some of the sites in conjunction with weak among population differentiation may result in poor estimation of allele frequencies at sites and, consequently, inaccurate reassignments. Lastly, individuals of both species have been collected and transported around central Europe for host-specificity testing. Because of this we can not rule out the possibility that the low among population differentiation and high estimated dispersal are due to common

garden escapees. It is worth noting, however, that garlic mustard, though both biennial and immobile, was estimated to have expanded its range in N.A. at a rate of 366 km^2/yr between 1868 and 1929 and 6400 km^2/yr between 1950 and 1991 (Nuzzo, 1993).

Genetic distance results from our more widely distributed samples of *C. alliariae* further suggest an important note about the dispersal capability of this species: mountains may constitute a more considerable geographic barrier to gene flow than water. The Italian site was only 57 km from the Austrian site, yet the genetic distance between these two sites was more than double what was seen between United Kingdom and Swiss and German sites. This may be an artifact of a founder effect at the Italian site or it may be due to the fact that this site was isolated from all others by the Alps. We can not say from our data whether mountains may also pose as barriers to dispersal for *C. roberti*.

Overall, both *C. alliariae* and *C. roberti* exhibit remarkably similar population genetic diversities and structure. The greater haplotypic and within-population diversity in *C. roberti* may indicate that it has remained at larger population sizes for longer than *C. alliariae*, or it may be that *C. alliariae* has undergone a more recent bottleneck followed by a range expansion. Both species appear to exhibit high dispersal and low among population variation. It is possible that the similarities in life histories, host-specificities, and competitive abilities of these two congeners play a significant role in structuring populations. If genetic diversity is important to the successful establishment of introduced agents, then it is possible that *C. roberti* could have a slight advantage over *C. alliariae*. However, given the minimal difference

between the two, we can not recommend prioritizing one over the other based on these data.

REFERENCES

Berry, O., Tocher, M.D., Sarre, S.D., 2004. Can assignment tests measure dispersal? Molecular Ecology 13, 551-561.

Blossey, B., Nuzzo, V., Hinz, H., Gerber, E., 2001. Developing biological control of *Alliaria petiolata* (M. Bieb.) Cavara and Grande (garlic mustard). Natural Areas Journal 21, 357-367.

Bohonak, A.J., 2002. IBD (isolation by distance): A program for analyses of isolation by distance. Journal of Heredity 93, 153-154.

Briese, D.T., Espiau, C., Pouchot-Lermans, A., 1996. Micro-evolution in the weevil genus *Larinus*: The formation of host biotypes and speciation. Molecular Ecology 5, 531-545.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Conord, C., Lemperiere, G., Taberlet, P., Despres, L., 2006. Genetic structure of the forest pest *Hylobius abietis* on conifer plantations at different spatial scales in Europe. Heredity 97, 46-55.

Culin, J., Brown, S., Rogers, J., Scarborough, D., Swift, A., Cotterill, B., Kovach, J., 1990. A simulation-model examining boll-weevil dispersal - historical and current situations. Environmental Entomology 19, 195-208.

Dosdall, L.M., Weiss, R.M., Olfert, O., Carcamo, H.A., 2002. Temporal and geographical distribution patterns of cabbage seedpod weevil

(Coleoptera : Curculionidae) in canola. Canadian Entomologist 134, 403-418.

Drake, J.M., Baggenstos, P., Lodge, D.M., 2005. Propagule pressure and persistence in experimental populations. Biology Letters 1, 480-483.

Drake, J.M., Lodge, D.M., 2006. Allee effects, propagule pressure and the probability of establishment: Risk analysis for biological invasions. Biological Invasions 8, 365-375.

Duchesne, P., Bernatchez, L., 2002. AFLPOP: a computer program for simulated and real population allocation, based on AFLP data. Molecular Ecology Notes 2, 380-383.

Durka, W., Bossdorf, O., Prati, D., Auge, H., 2005. Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Molecular Ecology 14, 1697-1706.

Gerber, E., Hinz, H.L., Blossey, B., 2008. Pre-release impact assessment of two stem-boring weevils proposed as biological control agents for *Alliaria petiolata*. Biological Control 45, 360-367.

Grevstad, F.S., 1999a. Experimental invasions using biological control introductions: the influence of release size on the chance of population establishment. Biological Invasions 1, 313-323.

Grevstad, F.S., 1999b. Factors influencing the chance of population establishment: Implications for release strategies in biocontrol. Ecological Applications 9, 1439-1447.

Hanski, I., Saccheri, I., 2006. Molecular-level variation affects population growth in a butterfly metapopulation. Plos Biology 4, 719-726.

Hawthorne, D.J., 2001. AFLP-based genetic linkage map of the Colorado potato beetle Leptinotarsa decemlineata: Sex chromosomes and a pyrethroid-resistance candidate gene. Genetics 158, 695-700.

Hinz, H., Gerber, E., Investigations on potential biological control agents of garlic mustard, *Alliaria petiolata* (Bieb.) Cavara & Grande, Annual Report for 2000. CABI Bioscience, Delemont, Switzerland, 2001.

Kephart, S.R., 2004. Inbreeding and reintroduction: Progeny success in rare *Silene* populations of varied density. Conservation Genetics 5, 49-61.

Kerdelhue, C., Roux-Morabito, G., Forichon, J., Chambon, J.M., Robert, E., Lieutier, F., 2002. Population genetic structure of *Tomicus piniperda* L. (Curculionidae : Scolytinae) on different pine species and validation of T-destruens (Woll.). Molecular Ecology 11, 483-494.

Kim, K.S., Sappington, T.W., 2004. Boll weevil (*Anthonomus grandis* Boheman) (Coleoptera : Curculionidae) dispersal in the southern United States: Evidence from mitochondrial DNA variation. Environmental Entomology 33, 457-470.

Li, S.J., Sun, L., Oseto, C.Y., Ferris, V.R., 2008. Population structure in gray sunflower seed weevil (Coleoptera : Curculionidae), based on analyses of mitochondrial DNA and microsatellites. Annals of the Entomological Society of America 101, 204-214.

Lynch, M., Milligan, B.G., 1994. Analysis of population genetic-structure with RAPD markers. Molecular Ecology 3, 91-99.

Memmott, J., Craze, P.G., Harman, H.M., Syrett, P., Fowler, S.V., 2005. The effect of propagule size on the invasion of an alien insect. Journal of Animal Ecology 74, 50-62.

Messenger, P.S., van den Bosch, R., 1971. The adaptability of introduced biological agents. In: Huffaker, C.B., (Ed.), Biological Control. Plenum Press, New York, pp. 68-92.

Messenger, P.S., Wilson, F., Whitten, M.J., 1976. Variation, fitness, and adaptability of natural enemies. In: Huffaker, C.B., Messenger, P.S., Eds., Theory and practice of biological control. Academic Press, New York, pp. 209-231.

Myers, J.H., Sabath, M.D., Genetic and phenotypic variability, genetic variance, and the success of establishment of insect introductions for

the biological control of weeds. In: Del Fosse, E., (Ed.), Proceedings of the Fifth international Symposium on Biological Control of Weeds. Commonwealth Scientific and Industrial Research Organization, Brisbane, Australia, 1981, pp. 649.

Nei, M., 1972. Genetic Distance between Populations. American Naturalist 106, 283-292.

Newman, D., Pilson, D., 1997. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. Evolution 51, 354-362.

Nuzzo, V., 1993. Distribution and spread of the invasive biennial *Alliaria petiolata* (garlic mustard) in North America. In: McKnight, B.L., (Ed.), Biological Pollution: control and impact of invasive exotic species. Indiana Academy of Science, Indianapolis, pp. 115-124.

Reed, D.H., Frankham, R., 2003. Correlation between fitness and genetic diversity. Conservation Biology 17, 230-237.

Reed, D.H., Nicholas, A.C., Stratton, G.E., 2007. Genetic quality of individuals impacts population dynamics. Animal Conservation 10, 275-283.

Robinet, C., Lance, D.R., Thorpe, K.W., Onufrieva, K.S., Tobin, P.C., Liebhold, A.M., 2008. Dispersion in time and space affect mating success and Allee effects in invading gypsy moth populations. Journal of Animal Ecology 77, 966-973.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. Cladogram Estimation. Genetics 132, 619-633.

Vekemans, X., Beauwens, T., Lemaire, M., Roldan-Ruiz, I., 2002. Data from amplified fragment length polymorphism (AFLP) markers show

indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology 11, 139-151.

Vergeer, P., van den Berg, L.J.L., Roelofs, J.G.M., Ouborg, N.J., 2005. Single-family versus multi-family introductions. Plant Biology 7, 509-515.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP - a new technique for DNA-fingerprinting. Nucleic Acids Research 23, 4407-4414.

Zhivotovsky, L.A., 1999. Estimating population structure in diploids with multilocus dominant DNA markers. Molecular Ecology 8, 907-913.

Table 1. Sites sampled for *Ceutorhynchus alliariae* and *C. roberti* in Europe, number of individuals used for AFLP and COI sequence analyses, and the different haplotypes recorded at each site.

				C. alliariae			*C. roberti*		
				AFLP	COI Sequence		AFLP	COI Sequence	
Site	Location	Latitude	Longitude	n	n	Haplotypes	n	n	Haplotypes
1	SW Delemont, CH	47.2734	7.0348	9	1	A	11	5	H,J,K
2	S Delemont, CH	47.2422	7.2823		1	A		5	H,J,O
3	NE Delemont, CH	47.4538	7.5992					3	H,L
4	NE Delemont, CH	47.4434	7.5462		1	A	23	4	H,J,L
5	NE Delemont, CH	47.4254	7.5093	9	4	A,D	12	6	H,O,P
6	NE Delemont, CH	47.4022	7.4767		1	A	21	3	H
7	Delemont, CH	47.3708	7.3611				28	4	H,J,P
10	W Berlin, DE	52.4672	12.9297		3	A,F			
11	NW Berlin, DE	52.6140	13.0056		1	A		1	N
13	N Berlin, DE	52.6786	13.3757		1	C			
14	E Berlin, DE	52.4736	13.7598	15	7	A,B			
15	E Berlin, DE	52.4116	13.8256	12	3	A,E			
16	S Potsdam, DE	52.2989	13.0250		2	A,B			
19	W Klieken, DE	51.8932	12.3355	17	1	A			
20	Triptis, DE	50.7368	11.8519	17	2	A			
21	Nurnberg, DE	49.6175	11.4076				32	7	H,I,T
22	Heilbronn, DE	49.2028	9.6855				18	4	H,M,Q
23	Neuenburg am Rhein, DE	47.8249	7.5636				11	3	H,P,R
24	Neuenburg am Rhein, DE	47.8289	7.5660				18	8	H,K
IT	Artegna, IT	46.2419	13.1388	12	6	A			
AT	Kerschbaum, AT	46.7552	13.2035	37	9	A,G			
BG	Montpellier, FR	43.6085	3.8792	25	10	A			
FR	Iasi, RO	47.1667	27.58		1	A		3	V,Y,W
RO	Sunningdale, UK	51.4	-0.65	30	3	A			
UK	Sofia, BG	42.7105	23.3238					4	S,U,X,N

Table 2. Number of sequences and haplotypes, haplotype and nucleotide diversities, and Theta per site for both the full and comparable data sets for *Ceutorhynchus alliariae* and *C. roberti*. Values are based on a 592 bp section of the mtDNA COI. Standard deviations given in parentheses.

	C.alliariae		*C. roberti*	
	Full	CH & DE only	Full	CH & DE only
Number of Sequences	57	28	60	53
Number of Haplotypes	7	6	18	12
Haplotype Diversity	0.373 (±0.080)	0.529 (±0.105)	0.780 (±0.052)	0.718 (±0.061)
Nucleotide Diversity	0.001 (±0.000)	0.001 (±0.000)	0.004 (±0.001)	0.004 (±0.001)
Theta per site	0.002 (±0.001)	0.002 (±0.001)	0.007 (±0.002)	0.005 (±0.002)
Avg. Num. Nuc. Diffs.	0.439	0.656	2.331	2.042

Table 3. Nei's pairwise genetic distances after Lynch and Milligan (1994) are above the diagonal. Pairwise geographic distance (km) are below the diagonal. Genetic distances are significant (P<0.05) unless otherwise denoted (n.s.).

C. alliariae

	1	5	14	15	19	20	IT	AT	FR	UK
1	—	0.002 n.s.	0.007 n.s.	0.005 n.s.	0.004 n.s.	0.021	0.116	0.073	0.027	0.046
5	40	—	0.004 n.s.	0.006 n.s.	0.007 n.s.	0.014	0.119	0.068	0.036	0.044
14	752	717	—	0.000 n.s.	0.006	0.015	0.122	0.061	0.037	0.055
15	750	715	8	—	0.004 n.s.	0.013	0.124	0.064	0.037	0.047
19	640	606	117	117	—	0.014	0.110	0.079	0.044	0.047
20	521	485	234	231	133	—	0.140	0.063	0.048	0.045
IT	479	448	694	688	631	509	—	0.151	0.166	0.156
AT	471	437	637	631	575	454	57	—	0.113	0.129
FR	476	510	1227	1225	1116	995	785	810	—	0.060
UK	719	735	991	995	894	873	1157	1129	930	—

C. roberti

	1	4	5	6	7	21	22	23	24
1	—	0.008 n.s.	0.006 n.s.	0.011 n.s.	0.012	0.001 n.s.	0.009 n.s.	0.005 n.s.	0.000 n.s.
4	43	—	0.001 n.s.	0.002 n.s.	0.005 n.s.	0.011	0.008	0.024	0.008
5	40	3	—	0.000 n.s.	0.010	0.007 n.s.	0.010 n.s.	0.015	0.010 n.s.
6	36	7	4	—	0.003 n.s.	0.011	0.015	0.017	0.013
7	27	16	13	9	—	0.020	0.014	0.017	0.006 n.s.
21	415	373	377	380	389	—	0.002 n.s.	0.028	0.004 n.s.
22	291	252	255	258	267	133	—	0.022	0.000 n.s.
23	73	42	45	47	53	345	219	—	0.008 n.s.
24	74	43	45	48	53	345	218	0.5	—

Table 4. Results of re-allocation assignment tests for *Ceutorhynchus alliariae* and *C. roberti*. Left column for each site (LL1) gives the number of individuals and where they were assigned to with the minimal log-likelihood difference set to 1 (i.e. where individuals were 10 times more likely to have come from). The center column for each site (%) shows the percentage of the total site sample (n) that was reallocated to each site. For example, of the 17 individuals from Site 19, 24% (4 inds.) were ten times more likely to have come from Site 19, where they were collected, than from any other site. One individual (6%) was ten times more likely to have come from Site 14, and 12 individuals (71%) where not reallocated because they were not 10 times more likely to have come from any particular site relative to the others.

C. alliariae

	From site:																					
	1		5		14		15		19		20		IT		AT		FR		UK	
Reallocated to:	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%
1	2	22			1	7														
5	1	11																		
14			1	11	5	33	2	17	1	6	1	6								
15	1	11	1	11	1	7	3	25												
19			1	11					4	24	1	6								
20											11	65								
IT													12	100						
AT															37	100				
FR																	24	96		
UK																			30	100
None	5	56	6	67	8	53	7	58	12	71	4	24					1	4		

Table 3 cont'd.

C. roberti

	From site:																			
	1		4		5		6		7		21		22		23		24	
Reallocated to site:	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%	LL1	%
1	1	9									1	3	1	6			2	11
4			1	4														
5			1	4			1	5										
6			2	9	1	8	1	5										
7			1	4			1	5	1	4							1	6
21	1	9									5	16						
22											2	6	2	11			1	6
23	1	9	1	4											2	18	1	6
24																		
None	8	73	17	74	11	92	18	86	27	96	24	75	15	83	9	82	13	72

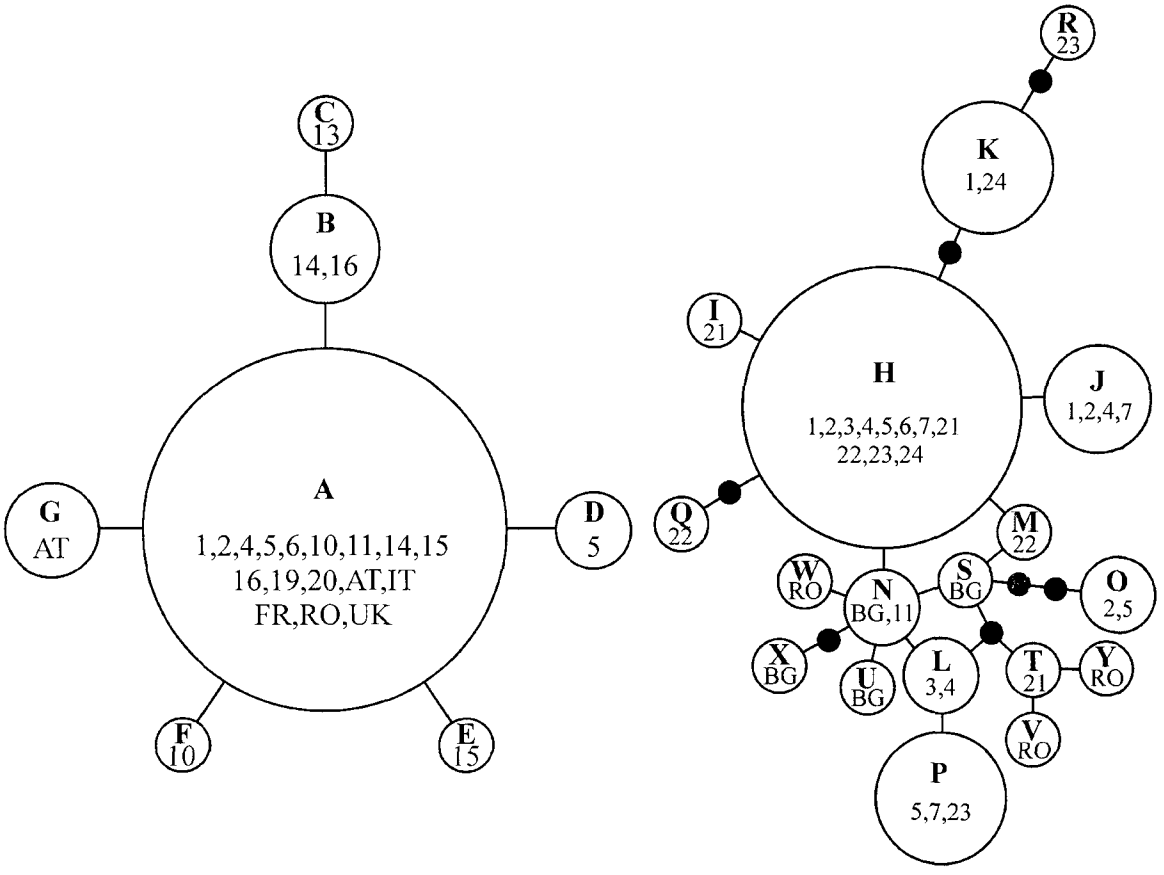

Figure 1. Haplotype networks showing the relationship of the different haplotypes for a 592 bp section of mtDNA COI region for *Ceutorhynchus alliariae* (left) and *C. roberti* (right). Haplotypes, designated with letters, are denoted by open circles. The relative sizes of the circles indicate the relative frequency of the haplotype in the samples. Small black circles indicate inferred intermediate haplotypes that were not found in the samples. Lines between circles indicate mutations.

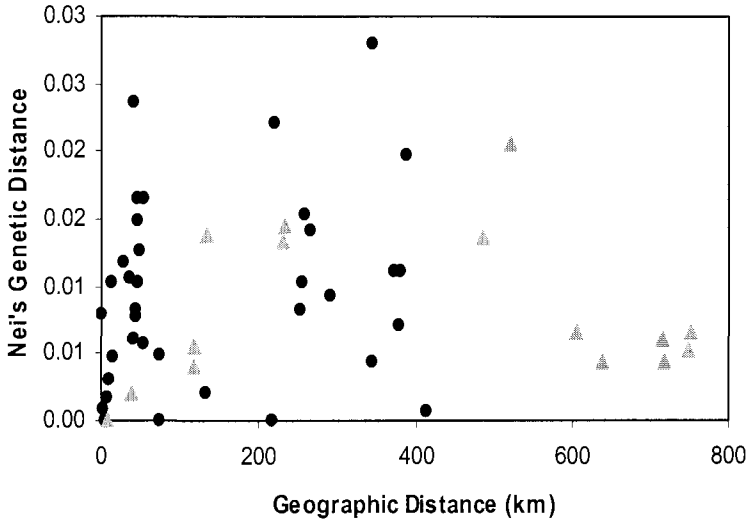

Figure 2. The relationships between geographic and genetic distance of *Ceutorhynchus alliariae* and *C. roberti* for the Swiss (CH) and German (DE) populations. Grey triangles represent population pairs for *C. alliariae*, and black circles represent population pairs for *C. roberti*. There is no significant isolation by distance for either species at this scale. See text for statistical details.