Identifying hybrids of *Peromyscus leucopus* and *P. maniculatus* in Michigan using molecular genetics

Katelyn I. Schumacher and Bradley J. Swanson
Applied Technology in Conservation Genetics Laboratory, Dept. of Biology, CMU

Abstract
This study tests for genetic evidence of hybridization between *Peromyscus leucopus* (white-footed mouse) and *P. maniculatus* (deer mouse) species using mitochondrial and microsatellite DNA techniques. The two species are generally found in different habitats (woodlands and open areas, respectively) throughout portions of their distribution in North America. In Michigan, however, both species are often captured in the same area. Similarity in appearances makes species identification in these areas of overlap often troublesome for field researchers, suggesting the species are possibly interbreeding and producing hybrids. We found that all individuals identified to species in the field were identified to the same species by mtDNA. However, 83% of individuals genotyped at two or more microsatellite loci failed to produce the same results. We hypothesize that it is the presence of null alleles in the population from which these loci were developed that accounts for much of the discrepancy between expected and observed microsatellite amplification patterns.

Introduction

The white-footed mouse (*Peromyscus leucopus*, WFM) and the deer mouse (*Peromyscus maniculatus*, DM) are two of the most widespread members of *Peromyscus* in the Great Lakes region. *P. leucopus* is a common resident of mature deciduous forests, whereas *P. maniculatus* prefers open areas (Kurta 1998). However, populations of *P. leucopus* and *P. maniculatus* overlap in their distribution throughout Michigan's Lower Peninsula and can often be captured together in the same area (Myers 1998). *P. leucopus, P. maniculatus* and their local geographic races also share several common external features – making field identification questionable (Bruseo *et al.* 1999, Reed *et al.* 2004, Rich *et al.* 1996).

Though classic interspecific mating experiments with *P. leucopus* and *P. maniculatus* produced no progeny (Dice 1933) and no instances of interbreeding have been recorded in field studies (Fitch 1963), there is still the possibility of current hybridization – compounding species identification in the field. The difficulty in species identification is especially troublesome in regions where the habitats of the two species are closely associated; mice at the interface of woods and prairies can rarely be identified unambiguously (J. Jacquot, personal communication), suggesting hybridization is occurring. Where there is grassland and woodland mosaic or interface – such as the Pierce Cedar Creek Institute Biological Field Station (PCCIBFS) and Edward Lowe Foundation (ELF) properties – no *P. maniculatus* population is completely isolated from dispersing *P. leucopus* (Kamler *et al.* 1998). However, little work has been done using modern molecular genetics to explore the question of hybridization between the two species under natural conditions, especially for those geographic races found in southwest Michigan.

Recent developments in genetic analyses and methods – such as mitochondrial and nuclear DNA sequencing or microsatellite genotyping – have fueled the expansion of molecular markers for use in population genetic studies regarding ecological, evolutionary, and conservation assessments (Manel *et al.* 2005, Selkoe and Toonen 2006). Regions of mitochondrial DNA subjected to restriction enzyme digestions (Caldwell and Novitski 1998) and species-specific primers for the COIII mitochondrial gene (Tessier *et al*. 2004) successfully differentiate the two *Peromyscus* species. However, mitochondrial

DNA is strictly maternally inherited in all mammals, thus in hybridization studies, data of only one parent can not provide genetic evidence of hybridization.

Genotyping of individuals at microsatellite loci is relatively low in cost compared to DNA sequencing, and their high mutation rate generates levels of genetic diversity needed for assessments on ecological time scales (Selkoe and Toonen 2006). Microsatellites are tandem repeats of one to six nucleotides (e.g., $[AT]_n$) found at high frequency in the nuclear genome of most taxa (Selkoe and Toonen 2006). Many microsatellite loci have been independently developed and tested in either *P. leucopus* (Schmidt 1999) or *P. maniculatus* (Mullen *et al.* 2006), but only one set of microsatellite loci developed for *P. maniculatus* (Chirhart *et al.* 2000) was cross-amplified across several other *Peromyscus* species – including *P. leucopus* (Chirhart *et al*. 2005). Unfortunately, these loci exhibited overlapping allele ranges in all but one locus – for which *P. leucopus* failed to amplify (Chirhart *et al*. 2005) – making it impossible to differentiate species using these loci. However, a set of microsatellite loci developed from the oldfield mouse, *Peromyscus polionotus subgriseus* (Prince *et al*. 2002) can be used to distinguish the two species and hybrids. Upon examination of these markers, a unique pattern of success or failure in amplification across loci for species (Prince *et al*. 2002) will be utilized for identifying *P. leucopus*, *P. maniculatus*, and hybrids of the two species (see Methods).

This study uses modern molecular genetics to test the null hypothesis that *P. leucopus* (WFM) and *P. maniculatus* (DM) are not hybridizing at either of the study sites (the PCCIBFS and ELF Foundation properties). This study will be the first of its kind to test these loci (Prince *et al.* 2002) in wild populations and, along with morphometric data

collected in the field (collaboration with Becky Norris and Dr. Joseph Jacquot, Grand Valley State University), may provide empirical data on field characteristics to distinguish the two *Peromyscus* species and their hybrid.

Methods

Trapping and Tissue Collection

Field methods regarding trapping techniques and external morphology data compilation was executed as described by Becky Norris and Dr. Joseph Jacquot (2008 URGE proposal) on the PCCIBFS and ELF properties.

Tissue sampling and external tagging methodology followed guidelines and recommendations of the Animal Care and Use Committee of the American Society of Mammalogists (Gannon *et al.* 2007). A small amount of tissue (2 mm) was cut from each ear using a tissue punch and extracted with sterilized forceps. Tissue samples were stored in microcentrifuge tubes in a -20C freezer. Metallic ear tags were carefully sized to prevent snagging (Gannon *et al.* 2007). Upon completion of data compilation, sampling, and tagging, uninjured individuals were released at the site of capture; uninjured non-target species were released immediately.

DNA Extraction & Sample Preparation

Following published protocols (Qiagen 2001), DNA was extracted from one tissue sample for each individual; the remaining tissue sample was freezer-stored in case additional DNA was necessary for future genetic analyses. DNA was quantified using an Eppendorf Biophotometer (Brinkman Instruments Inc., Westbury, NY, USA) and each sample diluted to 15 ng/μL for use as a working stock template in the Polymerase Chain

Reaction (PCR), a technique which allows the amplification of a specific region of DNA that lies between two regions of known DNA sequences (in this study, cytochrome B gene and microsatellites).

Mitochondrial DNA Sequencing

Mitochondrial DNA (mtDNA) sequencing of the cytochrome B gene was performed as an initial source of genetic identification. Also, if individuals were field-identified as a possible hybrid, and later genetically-identified as a hybrid using microsatellites, the mtDNA sequencing would provide evidence for the mode of hybridization; that is, the mtDNA sequencing would show whether the hybrid's mother was *P. leucopus* or *P. maniculatus*.

We performed PCR in a 20 μL cocktail [containing 75 ng genomic DNA, 250 μM dNTPs, 0.2 μM of each primer (forward and reverse), 1X GeneChoice Standard Reaction Buffer with 1.5 mM $MgCl_2$ (CLP, San Diego, California), and 1 unit of GeneChoice Taq DNA polymerase (CLP, San Diego, California)] and amplification was conducted on an Eppendorf MasterGradient Thermocycler (Brinkman Instruments Inc., Westbury, New York). PCR product was purified using QIAquick PCR Purification Kit reagents (Qiagen, Valencia, California) and following published spin-column protocol (Qiagen 2006). Purified products were prepared for sequencing using BigDye Terminator 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California) and ethanol precipitation protocols (Applied Biosystems 2002). Once the fluorescently labeled sequencing products were generated in the BigDye reaction, high throughput sequencing was conducted via capillary electrophoresis on an ABI 3730 Genetic Analyzer equipped with

data collection and sequence analysis software (Applied Biosystems, Foster City, California).

mtDNA Species Identification

We ran sequences through the National Center for Biotechnology Information (NCBI) online version of the program Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990). BLAST (http://blast.ncbi.nlm.nih.gov) compares imputed sequences to those submitted to the National Institute of Health's annotated and publically available genetic sequence database, GenBank, and calculates the statistical significance of matches (Benson *et al.* 2008). GenBank currently holds over 82 million sequence records for more than 260,000 named organisms and is commonly used for species identification studies (Benson *et al.* 2008).

Microsatellite Genotyping

We chose five of the eleven developed (Prince *et al.* 2002) microsatellite loci that allow for species and hybrid identification; this included four diagnostic loci and one universal locus. For each locus, we performed PCR in a 10 μL cocktail [containing 37.5 ng genomic DNA, 250 μM dNTPs, 0.2 μM of each primer (forward and reverse), 1X GeneChoice Standard Reaction Buffer with 1.5 – 2.5 mM $MgCl_2$ (CLP, San Diego, California), and 1 unit of GeneChoice Taq DNA polymerase (CLP, San Diego, California)] and amplification was conducted on an Eppendorf MasterGradient Thermocycler (Brinkman Instruments Inc., Westbury, New York). Following amplification, all samples were visualized via capillary electrophoresis on an ABI 310 Genetic Analyzer with GeneScan Analysis 3.1.2 (Applied Biosystems, Foster City, California) software. Alleles for

genotype construction were determined using Genotyper 2.0 software (Applied Biosystems, Foster City, California).

Microsatellite Species Identification

Identification of species and possible hybrids was based upon the composite amplification success pattern for each individual's complete genotype (Table 1). Since these microsatellites are from nuclear DNA, each individual will have two alleles – one from each parent. These alleles can either be of different size (heterozygote) or the same size (homozygote). Amplification of only one of these two alleles would resemble a homozygote. For example, a *P. leucopus* individual would amplify two alleles at loci PO-9, PO-26, and PO-40 (but no successful amplification at PO-105 and PO-21). A *P. maniculatus* individual would amplify two alleles at loci PO-105, PO-21, PO-9 (but no successful amplification at PO-40 and PO-26). A hybrid individual would amplify two alleles at locus PO-9 but only one allele at loci PO-40, PO-26, PO-105, and PO-21. Each locus where there should be amplification of two alleles, individuals could be either homozygous (two copies of same allele) or heterozygous (each allele is different).

Results & Discussion

We conducted 4,034 trap nights over the course of this study and captured a total of 288 *Peromyscus*, including 116 unique individuals (Table 2). We sequenced mtDNA for all adults (N=23) and a subset of juveniles (N=5); this sampling included all field identified *P. maniculatus* (N=4) and questionable field identified individuals (possible hybrids, N=2). Using a 96-99% threshold in BLAST (Altschul *et al.* 1990), all individuals

identified to species in the field were identified to the same species and possible hybrids were identified as *P. leucopus* by mtDNA (Table 3).

Microsatellite primer optimization was conducted for four of the five chosen microsatellite loci; locus PO-21 was excluded due to unavailability and time constraints. All remaining loci were amplified following published protocols (Prince *et al.* 2002), however samples failed to amplify in the expected pattern as based on field identification (as shown earlier in Table 1). Consecutive runs included annealing temperature and $MgCl_2$ concentration gradients (two variables that affect probability of successful amplification), yet successive reactions (nearly 150 individual reactions) failed to produce amplification patterns as published (Prince *et al.* 2002). For example, samples 386 and 406 were analyzed at all four microsatellite loci, and still neither individual's composite genotype fully matched the expected genotype for *P. leucopus*, *P. maniculatus* or a hybrid (Table 3, also see Table 1). Similar patterns were seen for other individuals that were genotyped for at least one locus (Table 4) and as a whole, these results did not agree with our mtDNA sequencing species identification.

Though mtDNA sequencing only accounts for maternal genetic inheritance, our mtDNA identification still provides a 96-99% probability that an individual is one of the two species or a hybrid; that is, an individual identified as *P. leucopus* using mtDNA could actually be a hybrid, but could not be a *P. maniculatus*. Thus, when considering the nuclear *Peromyscus* microsatellites compared to the mtDNA identification, an individual's composite microsatellite genotype must either match mtDNA identification or reveal that the individual is a hybrid. However, 83% of individuals genotyped at two or

more microsatellite loci (and 56% of individuals genotyped at only one locus) failed to produce either of these two results (see Table 4). The high proportion of mismatched microsatellite species identification led us to question the application of these loci to our study. We suggest that available information regarding the development of primers for these microsatellite loci may reveal why the published amplification pattern does not apply to studies like ours that look at the genetics of natural populations.

The developers of these primers (Prince *et al.* 2002) used DNA from the *Peromyscus Genetic Stock Center* (PGSC) for testing loci amplification patterns across eight *Peromyscus* taxa, including *P. leucopus* and *P. maniculatus*, for microsatellite loci developed from the oldfield mouse (*P. polionotus subgriseus*). The PGSC maintains *Peromyscus* laboratory closed breeding colonies of "wild-type" and "mutant" animals, as well as provides these specimens to academic or scientific institutions. The *P. maniculatus* colony derives from 40 wild-caught individuals (near Ann Arbor, MI) and has been bred in captivity since 1948; similarly, the *P. leucopus* colony has been bred in captivity and descends from 38 wild-caught individuals that were captured between 1982 and 1985 (near Linville, NC) (*Available Peromyscus*, http://stkctr.biol.sc.edu).

Our results show that wild-caught individuals do not show the same amplification pattern across loci as published for these closed PGSC populations. It is apparent that the closed breeding colonies have genetically diverged from natural *Peromyscus* populations over the 20-60 years of genetic isolation (no new influx of alleles) they have experienced. This divergence is likely expressed as an accumulation of mutations in the primer-binding region that flanks the actual microsatellite DNA sequence.

If a mutation occurs in the primer annealing site, this can result in failure to amplify one or both of the alleles present at that particular locus. An individual that possesses a single "null" allele at a locus will appear homozygous whereas if an individual has two null alleles, there will be no amplification. We hypothesize that the presence of these null alleles (and their likely accumulation, or high frequency, in the closed colony gene pool due to inbreeding) accounts for the "no amplification" in the published reference; therefore, any loci that did not amplify in closed strains may amplify in wild populations.

This discovery of uncertainty in wild population amplification has serious repercussions not only for the utilization of these loci in *P. maniculatus* and *P. leucopus*, but also in the six other *Peromyscus* taxa for which the loci were cross-amplified. The authors who developed the primers stated that "placing these as well as other *Peromyscus* markers on a *P. maniculatus* map, currently underway, will also enable genetic approaches to problems involving the abundant and widely distributed *P. leucopus*" (Prince *et al.* 2002); our results obviously suggest otherwise. We have obtained samples of both species through the PGSC and plan to incorporate these with our own to conduct further testing of these loci through reverse-sequencing of the primer annealing region. We then plan to publish our findings in the same source as the developers, *Molecular Ecology Notes*.

Table 1. Characterization of expected amplification patterns for *P. leucopus, P. maniculatus,* and possible hybrids collected during this study. Note that hybrids will require appearing homozygous (+/-) at 4 loci (PO-40, PO-26, PO-105, PO-21). Where two alleles should amplify (+/+), individuals may be heterozygous or homozygous at that locus.

	PO-40	PO-26	PO-105	PO-21	PO-9
P. leucopus	+/+	+/+	-/-	-/-	+/+
P. maniculatus	-/-	-/-	+/+	+/+	+/+
Hybrid	+/-	+/-	+/-	+/-	+/+

Table 2. Number of trap nights, *Peromyscus* captures, and number of unique individuals captured over the course of this study, including records by site and totals. Species identification for number of captures is based on field identification (WFM = white-footed mice (*P. leucopus*), DM = deer mice (*P. maniculatus*), "hybrids" = questionable species identification). Note that only WFM were captured at PCCIBFS.

	PCCIBFS	ELF	Total
Trap nights	2449	1585	4034
Peromyscus captures	173	115	288
Unique WFM	65	45	110
Unique DM	0	4	4
Unique "hybrids"	0	2	2

Table 3. Complete microsatellite genotyping results for individuals 386 (field and mtDNA identified as *P. leucopus*) and 460 (field and mtDNA identified as *P. maniculatus*), illustrating that amplification patterns – and their subsequent relation to species identification as based on Prince *et al.* (2002) – were not consistent across loci. Each individual's composite genotype included identification as *P. leucopus* (W), *P. maniculatus* (D), or a hybrid (B); note that 460 matched none (N) of these identification possibilities at locus PO-09.

	PO-09	PO-26	PO-40	PO-105
386	+/? W, D, H	+/+ W	-/- D	+/? D, H
460	-/- N	-/- D	+/? W, H	+/+ D

Table 4. Field identification versus genetic identification for expected (A) and observed (B) results. Microsatellite expected results by species (WFM=white-footed mouse, *P. leucopus*; DM=deer mouse, *P. maniculatus*, HYB = hybrid) are as based on published data (Prince *et al.* 2002, Table 1); (+/+) = heterozygote, (+/?) = homozygote (two copies of same allele) or hybrid genotype (amplification of only one allele). Observed results include individual identification (ID#), status (AD=adult, JV=juvenile), capture site (PCCI or ELF), as well as field and genetic identification (see above). Captured individuals identified as possible hybrids were identified as WFM by mtDNA. No microsatellite loci followed the expected amplification pattern.

A

SPECIES	Field	mtDNA	PO-09	PO-26	PO-40	PO-105
WFM	WFM	WFM	+/+ or +/?	+/+ or +/?	+/+ or +/?	-/-
DM	DM	DM	+/+ or +/?	-/-	-/-	+/+ or +/?
HYB	WFM or DM	WFM or DM	+/+ or +/?	+/-	+/-	+/-

B

ID#	STATUS	SITE	Field	mtDNA	PO-09	PO-26	PO-40	PO-105
151	AD	PCCI	WFM	WFM	*	*	*	+/+
173	AD	PCCI	WFM	WFM	*	*	*	+/+
326	AD	PCCI	WFM	WFM	*	*	*	+/+
328	AD	PCCI	WFM	WFM	*	*	*	+/+
331	AD	PCCI	WFM	WFM	*	*	*	-/-
332	AD	PCCI	WFM	WFM	*	*	*	+/+
333	JV	PCCI	WFM	WFM	*	*	*	+/?
334	AD	PCCI	WFM	WFM	*	*	*	+/?
336	AD	PCCI	WFM	WFM	*	*	*	+/+
341	AD	PCCI	WFM	WFM	*	*	*	+/+
344	AD	PCCI	WFM	WFM	*	*	*	+/?
376	JV	PCCI	WFM	WFM	-/-	*	-/-	+/+
386	JV	ELF	WFM	WFM	+/?	+/+	-/-	+/?
392	AD	PCCI	WFM	WFM	+/?	*	*	+/+
393	AD	ELF	WFM	WFM	*	*	*	-/-
396	AD	ELF	WFM	WFM	*	+/?	*	-/-
399	AD	ELF	WFM	WFM	*	*	*	+/?
402	AD	ELF	DM	DM	+/+	+/+	*	+/+
403	JV	ELF	DM	DM	+/?	+/+	*	+/+
419	AD	ELF	WFM	WFM	-/-	+/?	*	+/+
452	AD	ELF	WFM	WFM	*	*	*	+/+
454	AD	ELF	WFM	WFM	*	*	*	+/+
456	AD	ELF	WFM	WFM	*	*	*	+/?
460	JV	ELF	DM	DM	-/-	-/-	+/?	+/+
479	AD	ELF	HYB	WFM	*	+/?	*	-/-
480	AD	ELF	HYB	WFM	*	+/?	*	+/+
491	AD	ELF	WFM	WFM	*	+/+	*	+/+
494	AD	ELF	WFM	WFM	*	-/-	*	+/+

Acknowledgments

Project funding and researcher stipends were provided by the Pierce Cedar Creek Institute Undergraduate Research Grant for the Environment (URGE) program. We sincerely thank the Pierce Cedar Creek Institute Biological Field Station and Edward Lowe Foundation staffs for their generous accommodations, logistic field support, and continued hospitality. Likewise, we thank Becky Norris and Dr. Joseph Jacquot (Grand Valley State University) for the enthusiasm and dedication they brought to this collaborative project.

Literature Cited

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology **215**: 403-410.

Applied Biosystems. 2002. BigDye terminator v3.1. cycle sequencing kit protocol. Applied Biosystems: Foster City, California.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, D. L. Wheeler. 2008. Nucleic Acids Research **36** (Database issue): 25-30.

Bruseo, J. A., S. H. Vessey, and J. S. Graham. 1999. Discrimination between *Peromyscus leucopus noveboracensis* and *Peromyscus maniculatus nubiterrae* in the field. Acta Theriologica, **44**: 151 – 160.

Caldwell, L. and C. E. Novitski. 1998. DNA analysis for distinguishing among Michigan mice of the genus *Peromyscus*. Michigan Academician, **30**: 183 – 189.

Chirhart, S. E., R. L. Honeycutt, and I. F. Greenbaum. 2000. Microsatellite markers for

the deer mouse *Peromyscus maniculatus*. Molecular Ecology **9**: 1661 – 1686.

Chirhart, S. E., R. L. Honeycutt, and I. F. Greenbaum. 2005. Microsatellite variation and evolution in the *Peromyscus maniculatus* species group. Molecular Phylogenetics and Evolution, **34**: 408 – 415.

Dice, L. R. 1933. Fertility relationships between some of the species and subspecies of mice in the genus *Peromyscus*. Journal of Mammalogy, **14**: 298 – 305.

Fitch, J. H. 1963. A comparative behavioral study of *Peromyscus leucopus* and *Peromyscus maniculatus*. Transactions of the Kansas Academy of Science, **66**: 160 – 164.

Gannon, W. L., R. S. Sikes, and the Animal Care and Use Committee of the American Society of Mammalogists. 2007. Guidelines of the American Society of Mammalogists for the use of wild mammals in research. Journal of Mammalogy, **88**: 809 – 823.

Kamler, J. F., D. S. Pennock, C. Welch, and R. J. Pierotti. 1998. Variation in morphological characteristics of the white-footed mouse (*Peromyscus leucopus*) and the deer mouse (*P. maniculatus*) under allotopic and syntopic conditions. American Midland Naturalist, **140**: 170 – 179.

Kurta, A. 1998. Mammals of the Great Lakes region – Revised edition. The University of Michigan Press: Ann Arbor.

Manel, S., O. E. Gaggiotti, and R. S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution, **20**: 136 – 142.

Mullen, L. M., R. J. Hirschmann, K. L. Prince, T. C. Glenn, M. J. Dewey, and H. E. Hoekstra. 2006. Sixty polymorphic microsatellite markers for the oldfield mouse developed in *Peromyscus polionotus* and *Peromyscus maniculatus*. Molecular Ecology Notes, **6**: 36 – 40.

Myers, E. R. 1998. Aggression between two woodland-dwelling mice of the genus *Peromyscus*. Thesis, Central Michigan University.

Prince, K. L., T. C. Glenn, and M. J. Dewey. 2002. Cross-species amplification among peromyscines of new microsatellite DNA loci from the oldfield mouse (*Peromyscus polionotus subgriseus*). Molecular Ecology Notes, **2**: 133 – 136.

Qiagen. 2001. DNeasy tissue kit handbook. Qiagen: Valencia, California, USA.

Qiagen. 2006. QIAquick spin handbook for PCR purification kit. Qiagen: Valencia, California, USA.

Rich, S. M., C. W. Kilpatrick, J. L. Shippee, and K. L. Crowell. 1996. Morphological differentiation and identification of *Peromyscus leucopus* and *P. maniculatus* in northeastern North America. Journal of Mammalogy, **77**: 985 – 991.

Reed, A. W., P. K. Kennedy, M. L. Beck, and M. L. Kennedy. 2004. Using morphological characters to identify *Peromyscus* in sympatry. American Midland Naturalist, **151**: 190 – 195.

Schmidt, C. A. 1999. Variation and congruence of microsatellite markers for *Peromyscus leucopus*. Journal of Mammalogy, **80**: 522 – 529.

Selkoe, K. A. and R. J. Toonen. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters, **9**: 615 – 629.

Tessier, N., S. Noël, and F. J. Lapointe. 2004. A new method to discriminate the deer mouse (*Peromyscus maniculatus*) from the white-footed mouse (*Peromyscus leucopus*) using species-specific primers in multiplex PCR. Canadian Journal of Zoology, **82**: 1832 – 1835.