Sexual imprinting and speciation in two *Peromyscus* species

E.K. Delaney[*] and H.E. Hoekstra

Department of Organismic and Evolutionary Biology, Department of Molecular and Cellular Biology, Museum of Comparative Zoology, Howard Hughes Medical Institute, Cambridge MA 02138 USA Email: hoekstra@oeb.harvard.edu

**Running title:** Sexual isolation in *Peromyscus* mice

**Key words**: deer mice, mate choice, mosaic sympatry, reproductive isolation, sexual isolation

**Author contributions**: EKD and HEH conceived the project. EKD collected and analyzed the data. EKD and HEH wrote the paper.

**Acknowledgements**: We thank members of the Hoekstra lab who helped with field work: A. Chiu, G. Gonçalves, V. Domingues, E. zu Ermgassen, H. Fisher, E. Jacobs-Palmer, E. Kingsley, K. Lin, C. Linnen, H. Metz, B. Peterson, and J. Weber; we also thank E. Pivorun, R. Sikes, and their students for assistance. J. Chupasko and M. Omura from the Harvard Museum of Comparative Zoology Mammal Department helped with specimen curation. Nokuse Plantation granted us permission to trap on their property. The following museums donated tissue samples: Harvard Museum of Comparative Zoology, Florida Museum of Natural History, Oklahoma State University Collection of Vertebrates, Oklahoma Collection of Genomic Resources, and the Museum of Texas Tech University. We also wish to thank N. Delaney, K. Ferris, J. Mallet and our reviewers for helpful discussions and comments on this manuscript. This work was funded by: NSF GRFP, NSF DDIG, and grants from the American Society of Mammalogists, Animal Behavior Society, and the Harvard University Mind, Brain & Behavior Initiative to EKD. HEH is an Investigator of the Howard Hughes Medical Institute. The authors have no conflicts of interest.

**Data archival**: Short read data have been deposited in GenBank (SRP123258).

[*]Current address: Department of Evolution and Ecology, University of California, Davis CA 95616 USA

**Abstract**

Sexual isolation, a reproductive barrier, can prevent interbreeding between diverging populations or species. Sexual isolation can have a clear genetic basis; however, it may also result from learned mate preferences that form via sexual imprinting. Here, we demonstrate that two sympatric species of mice—the white-footed mouse (*Peromyscus leucopus*) and its sister species, the cotton mouse (*P. gossypinus*)—hybridize only rarely in the wild despite co-occurrence in the same habitat and lack of any measurable intrinsic postzygotic barriers in laboratory crosses. We present evidence that strong conspecific mating preferences in each species result in significant sexual isolation. We find that these preferences are learned in at least one species: *P. gossypinus* sexually imprints on its parents, but in *P. leucopus,* additional factors influence mating preferences. Our study demonstrates that sexual imprinting contributes to reproductive isolation that reduces hybridization between otherwise interfertile species, supporting the role for learning in mammalian speciation.

## INTRODUCTION

Sexual isolation, where sexual interactions such as divergent mating preferences or courtship behaviors reduce interbreeding, is a prevalent premating reproductive barrier that may facilitate speciation. Relative to some intrinsic postzygotic reproductive barriers, sexual isolation can accumulate rapidly among young allopatric (e.g. Mendelson 2003) and sympatric species (e.g. Coyne and Orr 1989, 1997), and it often acts as a major reproductive barrier among incipient sympatric species pairs (Coyne and Orr 1997; Noor 1997; Ramsey et al. 2003; Boughman et al. 2005; Nosil 2007; Matsubayashi and Katakura 2009). In several cases, sexual isolation is the sole reproductive barrier preventing hybridization between sympatric species, indicating that sexual isolation alone can be strong enough to reduce hybridization and thereby maintain genetic differentiation (e.g. Seehausen 1997; Fisher et al. 2006). Yet, despite the role that sexual isolation can play in instigating or maintaining reproductive isolation among species, its mechanistic basis—whether mating preference is genetic or learned—is often unknown.

Sexual isolation can evolve when mating traits and preferences are genetically encoded. If polymorphisms exist at a mating-trait locus and a preference locus, divergent alleles can co-evolve and fix between a pair of populations causing assortative mating. This scenario is known as a "two-allele mechanism" of reproductive isolation because two alleles must be present at both the mating-trait and preference loci (Felsenstein 1981). With the exception of a single pleiotropic trait/preference locus (Smadja and Butlin 2011), sexual isolation formed by the two-allele mechanism will break down due to recombination between the separate trait and preference loci unless strong selection, weak gene flow, or a high degree of linkage disequilibrium exists (Felsenstein 1981).

36 Sexual isolation can also evolve without genetically encoded preferences. Under a "one-
37 allele mechanism" of reproductive isolation, a single allele yields assortative mating—for
38 example, because of self-referent matching, mechanical assortment, or philopatry (Kopp et al. *in*
39 *press*). Sexual imprinting, a process in which offspring learn to prefer familial traits at a young
40 age (i.e. those of a mother, father, or siblings), has been considered an "one-allele mechanism"
41 (Verzijden et al. 2012a) because populations that diverge in a sexually imprinted mating trait can
42 mate assortatively thus leading to sexual isolation. Mechanisms such as sexual imprinting are
43 arguably more efficient at establishing reproductive isolation than the above-mentioned two-
44 allele mechanisms because they are immune to genetic recombination: separate preference
45 alleles do not need to be associated with polymorphisms in mating-trait alleles to produce
46 assortative mating (Felsenstein 1981; Smadja and Butlin 2011). Moreover, several theoretical
47 models have shown that learned mating preferences will maintain sexual isolation much longer
48 in populations experiencing gene flow than if mating preferences had a genetic basis because
49 sexual imprinting lowers the amount of divergent natural selection needed to isolate groups
50 (Laland 1994; Verzijden et al. 2005). Sexual imprinting may also boost reproductive isolation
51 through reinforcement (Servedio et al. 2009) or by driving divergence in mating traits. If
52 offspring develop preferences for more extreme versions of the traits on which they have
53 sexually imprinted, peak shift can occur (ten Cate and Rowe 2007), which can in turn drive
54 mating-trait evolution (ten Cate et al. 2006) and promote adaptive radiation (Gilman and Kozak
55 2015).
56 While sexual imprinting has long been recognized as a phenomenon that occurs within
57 species, its potential impact on speciation has become better appreciated only over the last two
58 decades (Irwin and Price 1999). It is a phenomenon that occurs in species with parental care, and

59 has now been documented in over 15 orders of birds (ten Cate and Vos 1999) as well as some
60 mammals (Kendrick et al. 1998; Montero et al. 2013) and fish (Verzijden and ten Cate 2007;
61 Kozak and Boughman 2009; Verzijden and Rosenthal 2011). A few empirical studies have
62 explicitly tested for a connection between sexual imprinting and sexual isolation between closely
63 related populations or species. For example, benthic and limnetic sticklebacks sexually imprint
64 on paternal traits under ecologically divergent selection, which results in significant sexual
65 isolation between the two morphs (Kozak et al. 2011). Other studies in cichlids (Verzijden and
66 ten Cate 2007), tits (Slagsvold et al. 2002), and Darwin's finches (Grant and Grant 1997) have
67 demonstrated that sexual imprinting can maintain sexual isolation. Therefore, sexual imprinting
68 seems to be an important, but underexplored, avenue to speciation.

69 Here we assess the role of sexual imprinting in generating reproductive isolation
70 between two mammalian species, the white-footed mouse (*Peromyscus leucopus*) and its sister
71 species, the cotton mouse (*P. gossypinus*), which diverged in allopatry during the Pleistocene
72 (Blair 1950). *P. leucopus* is distributed across the Midwest and eastern United States, whereas *P.*
73 *gossypinus* is restricted to the Southeast (Figure 1); their ranges overlap in the Gulf Coast states,
74 from Texas to Virginia. These species show some level of sexual isolation: when allopatric or
75 sympatric *P. leucopus* and *P. gossypinus* are placed in large arenas, both species mate with
76 conspecifics (Bradshaw 1965, 1968). While assortative mating in laboratory studies is potentially
77 strong, there is mixed evidence as to whether it is strong enough to prevent hybridization in wild
78 sympatric populations (Howell 1921; Dice 1940; McCarley 1954a; Price and Kennedy 1980;
79 Robbins et al. 1985; Barko and Feldhamer 2002).

80 In this study, we used genomic data to first assess hybridization in the wild and found
81 that the two species remain genetically distinct in sympatry despite rare hybridization events. We

82 then measured the degree of sexual isolation between *P. leucopus* and *P. gossypinus* in the lab,
83 and tested if it had a learned or genetic basis. Our results show that sexual imprinting produces
84 strong sexual isolation, and suggest that learning disproportionately contributes to the total
85 reproductive isolation we observed between two interfertile, sympatric sister species.
86
87 **METHODS**
88 **Study species**
89 *Peromyscus leucopus* and *P. gossypinus* are sister species that are thought to have
90 diverged during the Pleistocene over the last 2 million years (Blair 1950; Platt et al. 2015).
91 Fossils of *P. gossypinus* have been found in Florida and Texas (Wolfe and Linzey 1977), and *P.*
92 *leucopus* fossils have been found between Texas and Pennsylvania, and as far west as Missouri
93 (Lackey et al. 1985)—mirroring the current ranges of both species (Figure 1). The average
94 genetic distance (*D*; Nei 1972), a proxy for divergence time, between *P. leucopus* and *P.*
95 *gossypinus* is estimated to be 0.178 (Zimmerman et al. 1978). This estimate is lower than that of
96 well-differentiated *Peromyscus* species (*D* = 0.334-0.431; Zimmerman et al. 1978), suggesting
97 that *P. leucopus* and *P. gossypinus* are at an intermediate stage of speciation.
98
99 *Wild samples*
100 During April 2008 and January-February of 2010 and 2011, we collected 238 mice from
101 ten allopatric locations and 12 sympatric locations in the central and eastern United States
102 (Figure 1). At each location, we placed up to 300 Sherman traps every 20 feet in transects of 50
103 traps per line. From each mouse captured, we took liver or tail tissue and stored tissues in 100%
104 ethanol for subsequent DNA extraction. We augmented our own sampling with tissues from

museum specimens at the Harvard Museum of Comparative Zoology, Florida Museum of Natural History, Oklahoma State University Collection of Vertebrates, Sam Noble Museum Oklahoma Collection of Genomic Resources, and the Museum of Texas Tech University Genetic Resources Collection. Collecting locations and sample sizes for all animals included in this study are provided in Supplemental Table 1.

*Lab strains*

We obtained *P. leucopus* animals from the *Peromyscus* Genetic Stock Center (University of South Carolina). The *P. leucopus* stock was established with 38 founders caught between 1982-1985 from Avery County, North Carolina. In 2009, we established a stock of *P. gossypinus* animals from 18 founders caught in Jackson and Washington Counties, Florida. Both stocks were derived from allopatric sites, in which only one of the two species was present. In captivity, breeding colonies have been deliberately outbred to minimize inbreeding and preserve genetic diversity.

All animals were housed in standard mouse cages in either mated pairs (one female and one male) or in same sex cages with a maximum of five adults. Offspring were weaned into same sex cages 23 days after birth. We set the light cycle to 14 hours of light and 10 hours of dark and maintained a room temperature between 70 and 77 degrees Fahrenheit. All mice were fed a regular Purina diet (Purina Iso Pro 5P76) *ad libitum*.

In addition to maintaining these two species, we also bred hybrids in the laboratory. First generation (F1) hybrids were generated from both *P. gossypinus* female x *P. leucopus* male matings as well as the reciprocal cross. These F1 hybrids were then backcrossed to either *P. gossypinus* or *P. leucopus*.

128

129 **Detection of hybrids in sympatric populations**

130 *ddRADseq library construction and genotyping*

131 We extracted genomic DNA from 374 wild-caught individuals and two lab-raised hybrids
132 using an Autogen kit and AutoGenprep 965 instrument. We prepared double digest restriction-
133 associated DNA tag (ddRAD) libraries from each individual following the protocol described in
134 Peterson et al. (2012). Briefly, we digested 100-200 ng of DNA from every individual with two
135 restriction enzymes, EcoRI-HF and MspI (New England Biolabs), and purified the reactions with
136 AMPure XP beads (Beckman Coulter Genomics). After quantifying the cleaned and digested
137 product on a spectrophotometer plate reader (SpectraMax Gemini XS Plate Reader), we ligated
138 approximately 50 ng of digested DNA to uniquely barcoded EcoRI adapters and MspI adapters
139 in a 40 µl reaction volume with T4 DNA ligase (New England Biolabs). We pooled equal
140 amounts of 32-48 ligated samples and used two rounds of AMPure XP bead purification to
141 reduce the total pooled volume to 30 µl. We loaded each ligation pool onto a 2% agarose Pippin
142 Prep cassette (Sage Science) and selected fragments with a size of 300 ± 35 bp. We ran five
143 replicate Phusion PCRs according to the Finnzymes kit directions (Thermo Fisher Scientific) for
144 12 cycles with 5 µl of eluted Pippin Prep product as template. Each PCR was indexed using a
145 unique reverse primer (primer and index sequences from Peterson et al. 2012). Following PCR,
146 we pooled all replicate reactions and purified them with AMPure XP beads to concentrate each
147 ddRAD library. We multiplexed ddRAD libraries in equimolar ratios and sequenced 32-48
148 individuals per lane on the Genome Analyzer II or multiple sets of 48 individuals on the
149 HiSeq2000 across 9 total lanes on 7 flowcells. All reads were single end and ranged between 37-
150 47 bp.

151 We demultiplexed reads and aligned them by sample to a draft genome sequence of
152 *Peromyscus maniculatus* (NCBI: GCA_000500345.1) with STAMPY run in hybrid mode using
153 the BWA mem algorithm with default parameters (Lunter and Goodson 2011). We identified and
154 removed adapter sequences with Picard-tools 1.100 (http://broadinstitute.github.io/picard). We
155 realigned potential indels with the Genome Analysis Tool Kit v. 3.2-2 (GATK) IndelRealigner
156 (McKenna et al. 2010) and performed SNP discovery across all samples simultaneously using
157 the GATK UnifiedGenotyper (DePristo et al. 2011). We filtered alignments, keeping regions
158 with 100 or more total reads and an average base quality greater than 20. We retained biallelic
159 SNPs with a minimum mapping quality of 30 that were present in at least 90% of our individuals
160 at a depth of 10 or greater. To reduce linkage among SNPs in our dataset, we identified
161 "clusters" of SNPs within 100 bp of each other and more than 100 bp from another SNP, and we
162 randomly selected one SNP per cluster. Our final dataset contained 3,707 SNPs and 316 mice
163 that had over 90% of genotypes present at these SNPs (Supplemental Table 1). On average, each
164 individual had calls for 3,607 SNPs with an average a depth of coverage of 18.6. Of these mice,
165 we considered 71 to be of known ancestry: 20 *P. leucopus* were caught at allopatric sites or lab-
166 raised, 49 *P. gossypinus* were caught in allopatric sites or lab-raised, and two individuals were
167 lab-reared hybrids from our colonies. The remaining 245 individuals were of unknown ancestry
168 and collected in the predicted sympatric range.
169 Short read data were deposited in GenBank (accession number: SRP123258).
170
171 *Identification of hybrids*
172 We first used a model-free genetic principal component analysis (PCA) to evaluate
173 admixture between *P. leucopus* and *P. gossypinus*. We implemented genetic PCA using smartpca

174 from the Eigensoft v.6.0.1 package (Patterson et al. 2006) and output the first ten principal
175 components (PCs). After excluding outlier individuals and SNPs, our final dataset contained 288
176 individuals and 2,528 SNPs. We included individuals with known ancestry (i.e. from allopatric
177 sites in their range or taxonomically verified museum specimens) to identify PC values
178 corresponding to each species and identified hybrids as individuals with intermediate values
179 along the first principal component (McVean 2009). We assessed PC significance using Tracy-
180 Widom statistics (Patterson et al. 2006) implemented using twstats in Eigensoft v.6.0.1.
181 In a complementary model-based analysis, we used the Bayesian admixture model in
182 Structure v.2.3.4 (Pritchard et al. 2000) to assign individual coefficients of membership to
183 discrete clusters. We ran Structure with a burn-in period of 50,000 MCMC iterations, followed
184 by 50,000 iterations, and estimated membership coefficients in five replicate runs for cluster
185 sizes (*K*) ranging between 1 and 10. We used the Evanno method (Evanno et al. 2005)
186 implemented in Structure Harvester (Earl and VonHoldt 2011) to determine the most likely
187 number of clusters. We then used the full search algorithm in CLUMPP v.1.1.2 (Jakobsson and
188 Rosenberg 2007) to estimate individual membership coefficients for all 316 individuals in our
189 dataset across the replicate Structure runs. We considered individuals to be putative hybrids if
190 they had >10% membership to a second cluster. To visualize our date, we used distruct v.1.1
191 (Rosenberg 2004).
192
193 **Measurement of sexual isolation between species**
194 Using our laboratory *P. leucopus* and *P. gossypinus* stocks, we first tested for intrinsic
195 postzygotic isolation and estimated sexual isolation without mate choice. We then compared our
196 sexual isolation estimate from no-choice assays to those with mate choice to quantify the

contribution of mating preferences to reproductive isolation between *P. leucopus* and *P. gossypinus*.

*Intrinsic postzygotic isolation and sexual isolation without choice*

We tested for intrinsic postzygotic sexual isolation and sexual isolation between lab-raised *P. leucopus* and *P. gossypinus* using no-choice trials. We set up 20 crosses for each conspecific and heterospecific pairing: L♀ x L♂, G♀ x G♂, L♀ x G♂, and G♀ x L♂ (in which "L" represents *P. leucopus* and "G" represents *P. gossypinus*). When F1 offspring were produced, we used these mice in additional no-choice trials in backcross mating pairs: F1♀ x L♂, F1♀ x G♂, L♀ x F1♂, and G♀ x F1♂. We avoided any sib-sib or sib-parent pairings.

We set up mating pairs by adding a sexually receptive virgin female to the cage of a virgin, sexually mature male. We determined female sexual receptivity through vaginal lavage and considered a female to be receptive between proestrus and estrus stages. We gave pairs 60 days to produce a litter, which is approximately 12 estrous cycles (mean estrous cycle length for both species is 5-6 days; Dewsbury et al. 1977) or opportunities for successful reproduction. We considered the production of offspring as a successful mating event and inferred the latency to the first successful mating by subtracting the average gestation period—23 days in both species (Pournelle 1952; Wolfe and Linzey 1977; Lackey et al. 1985)—from the total number of days until a litter was born. Although our metric for mating success is conservative because it is confounded with any fertility differences that might exist among individuals or between the species, our assay nonetheless captures hybridization between these species.

We first used the no-choice assays to test hybrid viability and fertility in our laboratory strains of *P. leucopus* and *P. gossypinus*. We scored offspring survival to reproductive age in

220 heterospecific crosses (L♀ x G♂, G♀ x L♂), and then used these F1 hybrids in backcrosses to
221 look for evidence of reduced fertility relative to conspecific crosses. To compare the proportion
222 of successful mating events between conspecific and heterospecific crosses, we used a logistic
223 regression to quantify the effects of the female species, male species, or the interaction between
224 female and male species. We then selected the best-fit model using backward stepwise selection
225 based on the lowest Akaike Information Criterion (AIC). We compared the 95% confidence
226 intervals for the mean mating success among backcross pairs (F1♀ x L♂, F1♀ x G♂, L♀ x F1♂,
227 G♀ x F1♂) to those of conspecific crosses. Together, these no-choice data provide an estimate of
228 hybrid viability and relative fertility.

229 We next tested for differences in mating latency between conspecific, heterospecific, and
230 backcross mating pairs using a non-parametric Kruskal-Wallis rank sum test followed by
231 pairwise Wilcoxon tests with Bonferroni-corrected p-values. To quantify sexual isolation, we
232 counted the number of successful mating events to estimate a isolation index, $I_{PSI}$ (Rolán-Alvarez
233 and Caballero 2000), which compares observed to expected mating events (assuming random
234 mating among individuals) among conspecific and heterospecific pairs. This index ranges from -
235 1 (all mating occurred between species) to +1 (all mating occurred within species), with a value
236 of 0 indicating equal mating among pair types. We used the number of conspecific and
237 heterospecific pairs that produced litters to estimate $I_{PSI}$ in JMATING v.1.0.8 (Carvajal-
238 Rodriguez and Rolan-Alvarez 2006). We used 10,000 bootstrap replicates to estimate the sexual
239 isolation indices, their standard deviation, and to test the hypothesis that our estimates of the
240 sexual isolation index deviated significantly from a null hypothesis of random mating.
241
242 *Sexual isolation with choice*

243 We contrasted our estimate of the sexual isolation index ($I_{PSI}$) from no-choice assays to
244 the sexual isolation index estimated from two-way choice assays. We measured conspecific
245 mating preferences in a two-way electronically-controlled gated mate choice apparatus that
246 consisted of three collinear rat cages, with each pair of cages separated by two RFID antennae
247 and gates (FBI Science Gmbh; Figure 3A). Each pair of gates was programmed to allow passage
248 depending on the identity of the mouse. Specifically, for each trial we implanted three mice with
249 small transponders (1.4 mm x 9 mm, ISO FDX-B, Planet ID Gmbh) in the interscapular region
250 using a sterile hypodermic implanter and programmed the gates to allow the designated
251 "chooser" mouse (i.e. the individual whose preference we tested) to pass freely through all cages
252 while constraining each "stimulus" mouse to the left or right cage, respectively.
253 With this apparatus, we tested mate preferences of males and females of each species for
254 conspecific and heterospecific stimuli of the opposite sex. We allowed the chooser mouse—
255 either a sexually receptive virgin female (in proestrus or estrus as determined by vaginal lavage)
256 or a sexually mature virgin male—to acclimate to the apparatus for one day, adding food, water,
257 used nesting material, and a hut from each stimulus mouse's colony housing cage to the flanking
258 cages of the apparatus. Approximately 24 hours later, we returned the chooser mouse to the
259 center cage if it had not already nested there, closed all gates, and added stimulus mice to the two
260 flanking cages to allow them two to four hours to acclimate to their new environment. At lights
261 out (4:00 pm; 14:10 hour light:dark cycle), we re-opened the gates and recorded RFID readings
262 at all antennae as well as webcam video streams from each flanking cage for two nights (~44
263 hours; camera model: DLINK DCS-942L). Each chooser mouse was tested once.
264 At the end of each trial, we parsed a log file of RFID readings and calculated chooser
265 preference for a stimulus as the proportion of time spent with that stimulus divided by the time

spent with both stimuli. We analyzed only trials in which the chooser mouse investigated both cages during the acclimation, the chooser mouse spent at least 10 minutes investigating one stimulus during the trial, and both stimuli mice were in their cages at least 75% of the trial period (we discarded 15% of trials that did not meet these criteria).

We compared the preferences of 8-11 adults (at 9-14 weeks of age) of each species and sex for conspecific and heterospecific stimuli of the opposite sex. For female-choice trials, we tested virgin female preferences for either: (1) pairs of sexually experienced males that had successfully sired offspring with a conspecific female prior to use in the two-way choice trials (*P. leucopus*, $N = 5$ trials; *P. gossypinus*, $N = 7$ trials), or (2) pairs of virgin males as stimuli (*P. leucopus*, $N = 6$ trials; *P. gossypinus*, $N = 4$ trials). Because we did not detect a significant difference in female preference based on male stimulus sexual experience (two-sided Wilcoxon rank sum test, *P. leucopus* females: $W = 15$, $p = 1$; *P. gossypinus* females: $W = 9$, $p = 0.41$), we combined female preference data from trials with sexually experienced and virgin male stimuli. For male-choice trials, we used only virgin females as stimuli.

We estimated $I_{PSI}$ for each sex separately in JMATING v.1.0.8 (Carvajal-Rodriguez and Rolan-Alvarez 2006) because behavior of the stimuli may not be similar across male- and female-choice trials. We estimated $I_{PSI}$ by considering the chooser and its most preferred stimulus as a "mated" pair; when we observed no mating, we replaced zero values with a 1 to allow for bootstrapping with resampling. We used 10,000 bootstrap replicates to estimate the isolation indices and test for deviation from random mating ($I_{PSI} = 0$).

**Testing for sexual imprinting**

288 To determine whether conspecific mating preferences are learned in the nest, we
289 measured the preferences of mice from each species after they had been cross-fostered—raised
290 from birth until weaning—by parents of the opposite species. We swapped whole litters at birth
291 between breeding pairs of *P. leucopus* and *P. gossypinus*, reducing litters to the same number of
292 offspring if litters differed in number of pups. All cross-fostering attempts were successful,
293 indicating that parents readily attended to unrelated offspring. We allowed cross-fostered
294 offspring to remain with their foster parents until weaning (23 days after birth), when we
295 separated offspring into same sex cages; this matches the life cycle of all other mice in our study.
296 As a control, we also cross-fostered offspring within species (i.e. swapped litters between
297 conspecific families) to partition the effects of litter transfer and foster parent species on mating
298 preference. Although there is mixed (or incomplete) information for whether fathers contribute
299 parental care in *P. leucopus* and *P. gossypinus* (Hartung and Dewsbury 1979; Schug et al. 1992),
300 we cross-fostered offspring to both parents because we maintained male-female breeding pairs in
301 our laboratory colonies of *P. leucopus* and *P. gossypinus* and aimed to compare preferences of
302 mice from cross-fostered and non-cross-fostered trials.

303 We tested the mating preferences of all cross-fostered mice in the two-way gated choice
304 assay described above. We predicted that if young mice sexually imprint on their parents, cross-
305 fostered mice raised with the opposite species should prefer heterospecific stimuli and exhibit a
306 weaker preference for conspecifics compared to individuals raised by their biological parents or
307 other unrelated conspecific parents. We evaluated the effects of chooser sex and cross-fostering
308 treatment on preferences for *P. leucopus* in each species separately using linear modeling after
309 applying an arcsin transformation to the proportion of time spent with *P. leucopus*. To test for the
310 possibility that the sexes within each species might react differently to cross-fostering, we

311 considered models with and without an interaction between chooser sex and cross-fostering
312 treatment and selected the best-fit models using backward stepwise selection based on the lowest
313 AIC. We compared mean estimated preferences using two-sided t-tests with Bonferroni-
314 corrected p-values.
315
316 **Assessment of two-way choice assay**
317 We confirmed that our two-way mate choice assay accurately predicts mating preference
318 by measuring whether the most preferred stimulus corresponded to mating events in a subset of
319 trials in which mating occurred. We identified trials with successful mating events by either the
320 presence of sperm in a female reproductive tract at the end of a trial or the birth of a litter three
321 weeks later. If a female-choice trial resulted in offspring, we determined the identity of the father
322 by genotyping both the male stimuli and the pups at two to three microsatellite markers (loci 14,
323 35, and 80 from Weber et al. 2010) following the protocol described in Weber et al. 2010 ($N$ =
324 15 trials) or screening video data for copulation events ($N$ = 5 trials). We tested whether the most
325 preferred individual (as determined by the greatest proportion of association time) predicted
326 mating success using a linear regression. We applied an arcsin transformation to association time
327 proportions. This analysis allowed us to determine that association time is an accurate predictor
328 of mating, and thus reflects mating preference.
329
330 **RESULTS**
331 **Hybridization is rare in sympatric populations**
332 Using thousands of markers across the genome summarized in a genetic PCA, we tested
333 for evidence of hybridization between *P. leucopus* and *P. gossypinus* in sympatric populations.

334 We estimated ten principal components (PCs) and removed 28 outlier individuals that exceeded
335 six standard deviations for one of the PCs. Six of the ten PCs were significant by Tracy-Widom
336 statistics with the following eigenvalues: (1) 37.855, (2) 4.352, (3) 3.627, (4) 3.161, (5) 3.054,
337 and (6) 2.941. Based on clustering with known allopatric and previously identified *P. leucopus*
338 and *P. gossypinus* specimens, PC1 clearly separates *P. leucopus* (negative values) and *P.*
339 *gossypinus* (positive values) (Figure 1A). As expected, a control lab-generated F1 hybrid falls at
340 the midpoint along PC1 and a lab backcross mouse (F1 x *P. gossypinus*) falls halfway between
341 the F1 hybrid and the mean value of *P. gossypinus* values (Figure 1A). Of the remaining
342 sympatric mice we collected (i.e. samples not identified as outliers), all could be easily assigned
343 to either the *P. leucopus* or *P. gossypinus* species, with only two exceptions: two mice (EHK566
344 and EHK572) from Big Lake Wildlife Management Area, Louisiana had intermediate values
345 along PC1 (Figure 1A). These admixed individuals showed greater *P. leucopus* ancestry, similar
346 to a F1 backcross or advanced backcross to *P. leucopus*.
347 The second PC revealed two genetically distinct *P. gossypinus* subgroups. These likely
348 reflect genetic differences between *P. gossypinus* subspecies, *P. gossypinus gossypinus* and *P.*
349 *gossypinus megacephalus*. Specifically, higher PC2 values corresponded to mice caught east of
350 the Mississippi river—which are more likely to be *P. g. gossypinus*—whereas lower PC2 values
351 corresponded to mice caught west of the river—which are more likely to be *P. g. megacephalus*
352 (Wolfe and Linzey 1977). The Mississippi river is a known biogeographic barrier for many
353 species (Soltis et al. 2006), and our data suggest that this may also be the case for *P. gossypinus*.
354 Only one individual from the Tunica Hills wildlife management area population in Louisiana
355 failed to fit this pattern (Figure 1A): this individual was collected east of the Mississippi river but
356 it clustered with individuals from the western group. We did not find any evidence to suggest a

357 similar barrier to gene flow in *P. leucopus*, but we also did not have the equivalent population-
358 level sampling on both sides of the river. The remaining four PCs (3, 4, 5, and 6) identified
359 population structure within *P. leucopus* (Supplemental Figure 1).
360 We also estimated the optimal number of clusters in our dataset using a Bayesian
361 admixture model in Structure. This analysis provided parallel results to our genetic PCA results:
362 two clusters ($K = 2$) were identified in our data corresponding to *P. leucopus* and *P. gossypinus*
363 (Figure 1B) according to the Evanno method. Unlike genetic PCA, Structure estimated cluster
364 coefficients for all individuals in our analysis (i.e. Structure included 28 individuals that were
365 removed as outliers in the genetic PCA). We used the average individual ancestry assignments
366 across five replicate runs to identify potential hybrid individuals; in addition to the two potential
367 hybrids identified in genetic PCA, three additional individuals (MCZ68799, MCZ68800, and
368 EHK144) had ancestry proportions that were 83-90% *P. leucopus* and 10-17% *P. gossypinus*.
369 Two of these individuals were from Nannie M. Stringfellow Wildlife Management Area, Texas
370 (Figure 1C, site 13) and one was from Hart Creek, Georgia (Figure 1C, site 20).
371
372 **<u>*P. leucopus* and *P. gossypinus* co-occur in mosaic sympatry</u>**
373 Using cluster assignments based on the genetic PCA, eight of 14 sites where the species'
374 ranges overlap contained both species (Figure 1B, C). The other six sites contained only a single
375 species, highlighting the patchy distribution of both species within their broadly sympatric range
376 from Texas and Virginia.
377
378 **<u>No evidence for intrinsic postzygotic isolation</u>**

379 Previous studies suggested that there is no measurable intrinsic postzygotic isolation in
380 laboratory crosses of *P. leucopus* and *P. gossypinus* (Dice 1937). We confirmed this result in our
381 independent lines (i.e. different spatial and temporal origin) of these two species. We first
382 measured reproductive success within and between species in no-choice assays. Mating success
383 was determined largely by the female (logistic regression: $\beta = 1.25$, SE = 0.47, $p = 0.008$;
384 Supplemental Table 2), with *P. leucopus* females showing greater mean mating success than *P.*
385 *gossypinus* (Supplemental Figure 2). Importantly, this means that *P. leucopus* females had
386 greater reproductive success with *P. gossypinus* males (12/20 pairs had offspring) than the
387 reciprocal cross between *P. gossypinus* females and *P. leucopus* males (6/20 pairs had offspring),
388 indicating some asymmetry in mate preferences, copulation attempts, or female fertility.
389 Successful heterospecific crosses confirmed the ability to produce viable F1 hybrids, which
390 survive until reproductive age. In addition, we compared the mating successes of backcrosses to
391 conspecific and heterospecific mates. We found that F1 hybrids are as fertile in backcrosses (i.e.
392 had similar frequency of litter production) as either conspecific or heterospecific crosses, and
393 that all backcross offspring are also viable (Supplemental Figure 2).
394
395 **Mate choice leads to sexual isolation**
396 We next examined whether mating preferences lead to sexual isolation between the
397 species in a laboratory environment. In no-choice assays, heterospecific pairs hybridized and
398 produced viable offspring (Supplemental Table 3), indicating no measurable sexual isolation in
399 the absence of mate choice ($I_{PSI} = 0.00$, $SD = 0.19$, $p = 0.960$). However, conspecific,
400 heterospecific, and backcross mating pairs had significantly different latencies to produce
401 offspring (Figure 2; Kruskal-Wallis: $\chi^2 = 6.7626$, df = 2, $p = 0.034$). Pairwise comparisons

402 between mating pairs revealed significance differences in latency to mating only between
403 conspecific and heterospecific mating pairs (W = 69, $p_{\text{Bonferroni}} = 0.010$), but not between
404 conspecific and backcross mating pairs (W = 130, $p_{\text{Bonferroni}} = 0.949$) or between heterospecific
405 and backcross mating pairs (W = 188.5, $p_{\text{Bonferroni}} = 1$). Heterospecific pairs took an average of
406 5.4 days longer to produce litters than conspecific pairs, indicative of either delayed
407 heterospecific mating or longer hybrid gestation times. This delay is roughly equivalent to one
408 estrus cycle in *Peromyscus* (Dewsbury et al. 1977). No significant differences were detected
409 between the two conspecific pair types, L♀ x L♂ or G♀ x G♂ (W = 53, $p_{\text{Bonferroni}} = 0.238$), or
410 between the two heterospecific pair types, L♀ x G♂, and G♀ x L♂ (W = 25, $p_{\text{Bonferroni}} = 0.645$).
411 By contrast, we detected significant sexual isolation between the species in two-way
412 choice assays (Supplemental Table 3). Sexual isolation estimates were similar in female- and
413 male-choice trials: *P. leucopus* and *P. gossypinus* females strongly preferred conspecific mates
414 (Figure 3B; $I_{PSI} = 0.75$, $SD = 0.14$, $p < 0.01$) as did *P. leucopus* and *P. gossypinus* males (Figure
415 3B; $I_{PSI} = 0.75$, $SD = 0.15$, $p < 0.01$). More generally, there were strong preferences for
416 conspecific mates in both species, regardless of sex.
417
418 **<u>Sexual imprinting contributes to sexual isolation in at least one species</u>**
419 We then investigated whether mating preferences in these species had a learned or
420 genetic basis using a series of cross-fostering experiments. We found that cross-fostering had
421 different effects on mating preference in the two focal species. In *P. leucopus*, mating preference
422 was best predicted by a full model with cross-fostering, sex, and their interaction ($F = 5.09$ on 3
423 and 25 df, $p = 0.007$); a reduced model was not selected by AIC (Supplemental Table 4). When
424 raised with their own parents, *P. leucopus* of both sexes preferred *P. leucopus* stimuli (Figure

425 3B; estimated proportion of female time spent with *P. leucopus* = 0.689; estimated proportion of
426 male time spent with *P. leucopus* = 0.959). *P. leucopus* males that were cross-fostered
427 significantly changed their preference (Figure 3C; estimated proportion of cross-fostered male
428 time spent with *P. leucopus* = 0.184; t = -3.853, $p_{\text{Bonferroni}} = 0.003$), whereas cross-fostering did
429 not significantly change female preference (Figure 3C; estimated proportion of cross-fostered
430 female time spent with *P. leucopus* = 0.764; t = 0.390, $p_{\text{Bonferroni}} = 1$). Thus, *P. leucopus* females
431 always preferred *P. leucopus* to *P. gossypinus* mates, whereas a male spent more time with the
432 species with which it was raised.

433 In *P. gossypinus*, mating preference was best predicted by a reduced model
434 (Supplemental Table 5) with a significant cross-fostering term but no significant sex effects or
435 interactions between cross-fostering and sex ($F = 51.31$ on 1 and 33 df, $p < 0.001$). When raised
436 with their own parents, *P. gossypinus* of both sexes preferred *P. gossypinus* stimuli (Figure 3B;
437 estimated preference for *P. leucopus* = 0.069), whereas *P. gossypinus* raised with *P. leucopus*
438 preferred *P. leucopus* stimuli (Figure 3C; estimated preference for *P. leucopus* = 0.781).

439 To confirm that cross-fostering affect was caused by the foster parent species and not due
440 to transferring litters, we collected an additional control dataset for *P. gossypinus.* We cross-
441 fostered *P. gossypinus* to unrelated *P. gossypinus* foster parents (females: $N = 4$, males: $N = 7$)
442 and found that foster species, and not the transfer itself, affected *P. gossypinus* preferences
443 (Supplemental Figure 3). Pairwise t-tests on arcsin-transformed proportion of time spent with *P.*
444 *leucopus* revealed no significant differences between *P. gossypinus* raised with their own parents
445 or unrelated conspecific parents ($t = -0.72$, df = 15.38, $p_{\text{Bonferroni}} = 1$).

446 To examine the effects of sexual imprinting on sexual isolation, we calculated the sexual
447 isolation index ($I_{PSI}$) assuming the most preferred stimulus from each heterospecific cross-

448 fostered trial (Figure 3C) as a "successful mating". Cross-fostering eliminated sexual isolation in
449 female-choice trials ($I_{PSI}$ = 0.25, *SD* = 0.34, *p* = 0.57) and male-choice trials ($I_{PSI}$ = -0.29, *SD* =
450 0.42, *p* = 0.32). Thus, our cross-fostering results confirm that sexual isolation between *P.*
451 *leucopus* and *P. gossypinus* is the result of sexual imprinting.
452
453 **Two-way choice test accurately measures preference**
454 To confirm that the time spent with a stimulus mouse was an accurate predictor of mate
455 preference and hence mate choice, we recorded 20 mating events in our two-way choice assays:
456 12 mating events occurred in trials where choosers were raised with their own parents and 8
457 occurred in trials where choosers were raised with heterospecific foster parents. In 19 out of 20
458 trials, choosers mated with the stimulus individual with whom they spent the most time (Figure
459 4). Mating outcome (with conspecific or heterospecific stimulus) was predicted by the proportion
460 of time spent with the conspecific stimulus (logistic regression: $\beta$ = 10.06, *SE* = 4.86, *p* = 0.04),
461 indicating that our two-way choice assay accurately detects mating preferences.
462
463 **DISCUSSION**
464 Sexual imprinting can be a powerful generator of sexual isolation because it quickly and
465 effectively associates preferences with traits in populations. Furthermore, sexual imprinting has
466 been documented in a diversity of taxa—e.g. birds, fish, mammals, amphibians, and insects—
467 suggesting it may be a broadly important driver of speciation (Immelmann 1975). Our study
468 shows that that sexually imprinted mate-choice has likely contributed to and maintained strong
469 sexual reproductive isolation between a pair of mammalian sister species.
470

471 **Rare hybridization in sympatry indicates a high degree of reproductive isolation**

472 To test the strength of reproductive isolation between *P. leucopus* and *P. gossypinus* in
473 nature, we first collected mice from across their ranges and used genomic data to test for
474 hybridization between these species in sympatry. Classic studies by mammalogists in the mid
475 1900's reported conflicting results as to the extent of interspecific hybridization in sympatric
476 populations. In Louisiana, Alabama, and southern Illinois, Howell (1921), McCarley (1954a),
477 and later Barko and Feldhamer (2002) identified a few intermediate individuals resembling
478 hybrids based on morphology and allozyme genotypes. By contrast, Dice (1940) found no
479 evidence of morphological intermediates in his studies in Virginia. Thus, the degree of
480 hybridization if any between these two species in the wild has been contested historically.

481 In total, our analyses identified only five potential wild hybrids out of 245 mice that were
482 collected from locales where the species' ranges overlap (Figure 1C). Two hybrids were
483 identified in both genetic analyses (genetic PCA and Structure) and three identified by Structure
484 alone; all had greater proportions of *P. leucopus* ancestry. Thus, we found that approximately 2%
485 of individuals were admixed. Interestingly, the five hybrids we identified occurred in locations
486 where *P. gossypinus* was the rarer species, providing one explanation as to why they likely
487 backcrossed to *P. leucopus*. Nonetheless, both model-free and model-based clustering methods
488 showed that the vast majority of mice in our study clustered into two discrete groups, one for
489 each species, regardless of population. Our genomic analysis thus suggests that, despite rare
490 hybrids, *P. leucopus* and *P. gossypinus* remain genetically distinct in nature.

491 Our genomic data, which allowed us to confidently assign individuals to species, also
492 revealed that *P. leucopus* and *P. gossypinus* are distributed in a mosaic sympatry, with several
493 sites containing only one species (six of 14 sampling sites). This patchiness could be driven by

494 differences in microhabitat use: *P. leucopus* often occupy upland habitat and use more arboreal
495 nest sites while *P. gossypinus* often occupy swamps and bottomland habitat and use more ground
496 nest sites when they co-occur (McCarley 1954b, 1963; Taylor and McCarley 1963). However,
497 these habitat differences are not enough to exclude contact in sympatry because both species can
498 be trapped in the same patch of forest, especially where these habitat types abut (Dice 1940;
499 Calhoun 1941; Price and Kennedy 1980; Roehrs et al. 2012). In fact, we often caught both
500 species in the same trap line, indicating that the species overlap within each other's cruising
501 ranges. Similarly, there do not appear to be any significant differences in breeding seasons: the
502 two species have overlapping peak reproductive activities in the winter months, but adults from
503 both species can also be caught in reproductive condition throughout the year in Texas,
504 Louisiana, and Alabama (Pournelle 1952; McCarley 1954c; Wolfe and Linzey 1977). Thus, the
505 distributions, habitat preferences, and breeding seasons are unlikely to form complete or even
506 strong reproductive barriers, suggesting that behavioral differences may be an important
507 contributor to the level of reproductive isolation we observed in the wild.
508
509 **Learned sexual isolation in *P. leucopus* and *P. gossypinus***
510 As previous studies suggested that mating preferences might explain the lack of
511 hybridization in the wild, we tested for evidence of sexual isolation. Using no-choice and choice
512 assays to examine *P. leucopus* and *P. gossypinus* mating preferences, we found that conspecific
513 preferences form a significant sexual barrier between the two species. Without a choice of mates,
514 *P. leucopus* and *P. gossypinus* did not show significant sexual isolation, although there was an
515 increase in latency to mate in heterospecific crosses relative to conspecific crosses. However,
516 when given a choice of mates, the species mated assortatively, and we estimated the average

517 sexual isolation index ($I_{PSI}$) between the species to be 0.651. While sexual isolation is high, it is
518 not yet complete ($I_{PSI} < 1$) between these species. However, the amount of sexual isolation we
519 have observed is far greater than what has been detected among cactophilic ($I_{PSI} = 0.12$; Etges
520 and Tripodi 2008) or Caribbean *Drosophila* ($I_{PSI} = 0.159$-$0.282$; Yukilevich and True 2008),
521 walking stick insect populations ($I_{PSI} = 0.24$-$0.53$; Nosil et al. 2013), or gold and normal
522 Nicaraguan cichlid color morphs ($I_{PSI} = 0.39$ and $0.86$; Elmer et al. 2009), placing *P. leucopus*
523 and *P. gossypinus* quite far along a speciation continuum.

524 Using cross-fostering experiments, we found that conspecific mating preferences were
525 largely determined by sexual imprinting. This result implies that sexual isolation, a primary
526 reproductive barrier between sympatric, interfertile populations of *P. leucopus* and *P.*
527 *gossypinus*, is mostly due to learning. This work also implies that there are informative cues that
528 the species reliably use to distinguish between *P. leucopus* from *P. gossypinus* (but we do not yet
529 know if these signals are chemical, audial, or visual). Our work suggests that mammalian species
530 that sexually imprint might therefore be poised to form strong reproductive barriers at earlier
531 stages in the speciation process that enable sympatry without rampant hybridization. In fact,
532 other species of *Peromyscus* are also affected by cross-fostering (Carter and Brand 1986; Bester-
533 Meredith and Marler 2001), raising the possibility that their speciation trajectories could have
534 similarly been affected by learned mating preferences.

535 Intriguingly, our cross-fostering studies also revealed that the degree of imprinting
536 differed by species and sex. We found that both male and female *P. gossypinus* strongly sexually
537 imprinted on their foster parent species. By contrast, we found that *P. leucopus* also sexually
538 imprint on parents, although only weakly. Some *P. leucopus* males had a reduced preference for
539 conspecifics when raised with heterospecific parents, whereas all *P. leucopus* females appeared

540 unaffected by cross-fostering. *P. leucopus* showed a similar sexual difference in a study that
541 examined preferences for soiled bedding after cross-fostering to grasshopper mice, *Onychomys*
542 *torridus* (McCarty and Southwick 1977): although both male and female *P. leucopus* raised with
543 *O. torridus* parents had decreased preference for conspecific soiled bedding, the effect was more
544 dramatic in males than females. Thus, both *P. leucopus* and *P. gossypinus* appear to learn mating
545 preferences, but the degree of sexual imprinting varies between the two species, and between the
546 sexes in *P. leucopus*.
547
548 **Interspecific and sex-biased differences in sexual imprinting**
549 While *P. gossypinus* males and females form strong conspecific mating preferences
550 through sexual imprinting, only males of its sister species, *P. leucopus*, appear to sexually
551 imprint. Such asymmetric effects of sexual imprinting on congeneric species may not be usual.
552 For example learning affects mating preferences asymmetrically in congeneric tits (Slagsvold et
553 al. 2002) and swordtails (Verzijden et al. 2012b). What might cause this variation in learning
554 between *Peromyscus leucopus* and *P. gossypinus,* and why are preferences in *P. leucopus*
555 females robust to sexual imprinting?
556 One possibility is that conspecific mating preferences are innate and genetically
557 controlled in *P. leucopus* females due to reinforcement with *Peromyscus maniculatus*, a
558 sympatric species whose geographic range largely overlaps with *P. leucopus* (Hall 1981).
559 Because hybrids between *P. leucopus* and *P. manciulatus* are inviable (Maddock and Dawson
560 1974), natural selection could have reinforced the canalization of conspecific mating preferences
561 in *P. leucopus* females if they incur high costs from heterospecific mating. Innate genetic
562 conspecific mating preferences in *P. leucopus* females would suggest that the hybrids we

563 detected are more likely to been progeny from crosses between *P. leucopus* males with *P.*
564 *gossypinus* females.

565 Alternatively, *P. leucopus* may sexually imprint on parents but modify their preferences
566 after interactions with conspecifics and heterospecifics. In our study, male *P. leucopus* stimuli
567 may direct more copulatory behavior toward *P. leucopus* females, whereas male *P. gossypinus*
568 stimuli may be more antagonistic, thereby causing females to reverse learned preferences for
569 heterospecifics. Such preference reversals following cross-fostering have been observed in other
570 species (Rosenthal 2017). For example, a study of the effects of cross-fostering between sheep
571 and goats found that females raised with heterospecific foster parents initially preferred
572 heterospecific males, but later preferred conspecifics after a year of socialization (Kendrick et al.
573 1998); in contrast, males continued to prefer mates of their foster parent species. Similarly,
574 female zebra finches cross-fostered with Bengalese foster parents spent more time with
575 Bengalese males but directed more sexually receptive tail quivering behavior to conspecific
576 males who sang more vigorously and frequently (ten Cate and Mug 1984). If mating preferences
577 in *P. leucopus* females are indeed learned but susceptible to adult social interactions, mating
578 attempts by *P. gossypinus* males might account for the few hybrids we observed in our study.

579 Finally, the species and sexes could differ in their sexual imprinting sets. Imprinting on
580 fathers is more likely to evolve than imprinting on mothers (Tramm and Servedio 2008) and
581 could potentially occur in *Peromyscus,* as it does in *Mus* (Montero et al. 2013), if males associate
582 with juvenile offspring. Should the few hybrids we discovered be primarily produced from one
583 type of heterospecific cross, imprinting on either mothers or fathers would lead to biased
584 introgression. In addition, imprinting on siblings is also possible given that we cross-fostered
585 whole litters to male-female pairs. Thus, the own-species bias in *P. leucopus* females but not *P.*

586 *gossypinus* might also be the result of imprinting on siblings. Future experiments could
587 experimentally test for the imprinting set, and even specific cues involved, determining if and
588 how they differ between species and sexes.
589
590 **Reproductive isolation in sympatry**
591 Sexual imprinting could be even stronger between *P. leucopus* and *P. gossypinus* than
592 what we have measured in the lab if it were reinforced in sympatric populations (Irwin and Price
593 1999; Servedio et al. 2009). Although we did not find evidence of hybrid inviability or sterility in
594 the laboratory using allopatric stocks, the degree of hybrid fertility could vary in severity in
595 natural hybrid zones (e.g. Turner et al. 2011). Additionally, extrinsic postzygotic barriers, such
596 as behavioral sterility, may create an opportunity for reinforcement. Previous work found that *P.*
597 *leucopus* and *P. gossypinus* reciprocal hybrids initiated copulation less frequently than either *P.*
598 *leucopus* or *P. gossypinus* despite having similar copulatory behaviors (Lovecky et al. 1979).
599 Hybrids also differed in exploratory behavior compared to either parental species (Wilson et al.
600 1976), which may reduce hybrid fitness. Finally, hybrids might be behaviorally sterile if they
601 have intermediate mating traits. For example, hybrids between *M. m. musculus* and *M. m.*
602 *domesticus* have intermediate urinary signals that are selected against by each subspecies (Latour
603 et al. 2014). That we have found moderate sexual isolation in our allopatric lab stocks implies
604 that learning could be selected and strengthened in sympatry if it reduced the production of
605 behaviorally unfit hybrids. The potential for behaviorally-induced reinforcement, coupled with
606 the fact that moderate sexual imprinting induces sexual isolation in our lab stocks, could boost
607 reproductive isolation in sympatry and help explain the paucity of hybrids we have observed in
608 our study.

## CONCLUSION

Our study supports an emerging view that sexual imprinting may be vital to the generation and maintenance of sexual reproductive barriers. Pending divergence in an imprintable trait, a species that learns mating preferences may develop significant sexual isolation that might mitigate the homogenizing effects of hybridization. Our demonstration of sexual imprinting in *Peromyscus gossypinus* and *P. leucopus*, sympatric sister species that have few other measurable reproductive barriers between them, suggests that sexual imprinting may be an important contributor to their overall reproductive isolation. However, it is notable that the strength of imprinting differs between the species, and in one species, is largely sex-specific. Nonetheless, sexual imprinting could sculpt reproductive isolation in subspecies (e.g. benthic and limnetic sticklebacks) undergoing initial morphological and behavioral divergence, or help preserve reproductive isolation between already divergent species, as we see in *P. leucopus* and *P. gossypinus*. Examining the role of sexual imprinting in similar cases of speciation driven by sexual reproductive barriers will continue to expand our understanding of the role of behavior in speciation.

## References

Barko, V. A., and G. A. Feldhamer. 2002. Cotton mice (*Peromyscus gossypinus*) in southern Illinois: evidence for hybridization with white-footed mice (*Peromyscus leucopus*). Am. Midl. Nat. 147:109–115.

Bester-Meredith, J. K., and C. A. Marler. 2001. Vasopressin and aggression in cross-fostered California mice (*Peromyscus californicus*) and white-footed mice (*Peromyscus leucopus*).

632 Horm. Behav. 40:51–64.

633 Blair, W. F. 1950. Ecological factors in speciation of *Peromyscus*. Evolution 4:253–275.

634 Boughman, J. W., H. D. Rundle, and D. Schluter. 2005. Parallel evolution of sexual isolation in
635 sticklebacks. Evolution 59:361–373.

636 Bradshaw, W. N. 1965. Species discrimination in the *Peromyscus leucopus* group of mice. Texas
637 J. Sci. 17:278–293.

638 Bradshaw, W. N. 1968. Progeny from experimental mating tests with mice of the *Peromyscus*
639 *leucopus* group. J. Mammal. 49:475–480.

640 Calhoun, J. B. 1941. Distribution and food habits of mammals in the vicinity of the Reelfoot
641 Lake Biological Station. J. Tennessee Acad. Sci. 16:177–187.

642 Carter, R. L., and L. R. Brand. 1986. Species recognition in wild-caught, laboratory-reared and
643 cross-fostered *Peromyscus californicus* and *Peromyscus eremicus* (Rodentia, Cricetidae).
644 Anim. Behav. 34:998–1006.

645 Carvajal-Rodriguez, A., and E. Rolan-Alvarez. 2006. JMATING: a software for the analysis of
646 sexual selection and sexual isolation effects from mating frequency data. BMC Evol. Biol.
647 6:40.

648 Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution 43:362–381.

649 Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution
650 51:295–303.

651 DePristo, M. A., E. Banks, R. Poplin, K. V Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis,
652 G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y.
653 Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework
654 for variation discovery and genotyping using next-generation DNA sequencing data. Nat.

655 Genet. 43:491–498.
656 Dewsbury, D. A., D. Q. Estep, and D. L. Lanier. 1977. Estrous cycles of nine species of muroid
657 rodents. J. Mammal. 58:89–92.
658 Dice, L. R. 1937. Fertility relations in the *Peromyscus leucopus* group of mice. Contrib. Lab.
659 Vert. Genet. 4:1–3.
660 Dice, L. R. 1940. Relationships between the wood-mouse and the cotton-mouse in eastern
661 Virginia. J. Mammal. 21:14–23.
662 Earl, D. A., and B. M. VonHoldt. 2011. STRUCTURE HARVESTER: a website and program
663 for visualizing STRUCTURE output and implementing the Evanno method. Conserv.
664 Genet. Resour. 4:359–361.
665 Elmer, K. R., T. K. Lehtonen, and A. Meyer. 2009. Color assortative mating contributes to
666 sympatric divergence of neotropical cichlid fish. Evolution 63:2750–2757.
667 Etges, W. J., and D. A. Tripodi. 2008. Premating isolation is determined by larval rearing
668 substrates in cactophilic *Drosophila mojavensis*. VIII. Mating success mediated by
669 epicuticular hydrocarbons within and between isolated populations. J. Evol. Biol. 21:1641–
670 1652.
671 Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals
672 using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611–2620.
673 Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of
674 animals? Evolution 35:124–138.
675 Fisher, H. S., B. B. M. Wong, and G. G. Rosenthal. 2006. Alteration of the chemical
676 environment disrupts communication in a freshwater fish. Proc. R. Soc. B. 273:1187–1193.
677 Gilman, R. T., and G. M. Kozak. 2015. Learning to speciate: the biased learning of mate

678 preferences promotes adaptive radiation. Evolution 69:3004–3012.

679 Grant, P. R., and B. R. Grant. 1997. Hybridization, sexual imprinting, and mate choice. Am. Nat.
680 149:1–28.

681 Hall, E. R. 1981. The mammals of North America. Second edition. John Wiley & Sons, New
682 York.

683 Hall, E. R., and K. R. Kelson. 1959. The mammals of North America. Ronald Press Company,
684 New York.

685 Hartung, T. G., and D. A. Dewsbury. 1979. Paternal behavior in six species of muroid rodents.
686 Behav. Neural Biol. 26:466–478.

687 Howell, A. H. 1921. A biological survey of Alabama. North Am. Fauna 45:1–89.

688 Immelmann, K. 1975. Ecological significance of imprinting and early learning. Annu. Rev. Ecol.
689 Evol. Syst. 6:15–37.

690 Irwin, D. E., and T. Price. 1999. Sexual imprinting, learning and speciation. Heredity 82:347–
691 354.

692 Jakobsson, M., and N. a Rosenberg. 2007. CLUMPP: a cluster matching and permutation
693 program for dealing with label switching and multimodality in analysis of population
694 structure. Bioinformatics 23:1801–1806.

695 Kendrick, K. M., M. R. Hinton, K. Atkins, M. A. Haupt, and J. D. Skinner. 1998. Mothers
696 determine sexual preferences. Nature 395:229–230.

697 Kopp, M., R. J. Safran, M. R. Servedio, R. L. Rodr, T. C. Mendelson, M. E. Hauber, E. C.
698 Scordato, C. N. Balakrishnan, L. B. Symes, D. M. Zonana, and G. S. Van Doorn. n.d.
699 Mechanisms of assortative mating in speciation with gene flow: connecting theory and
700 empirical research. Am. Nat. *in press*.

701 Kozak, G. M., and J. W. Boughman. 2009. Learned conspecific mate preference in a species pair
702 of sticklebacks. Behav. Ecol. 20:1282–1288.
703 Kozak, G. M., M. L. Head, and J. W. Boughman. 2011. Sexual imprinting on ecologically
704 divergent traits leads to sexual isolation in sticklebacks. Proc. R. Soc. B. 278:2604–2610.
705 Lackey, J. A., D. G. Huckaby, and B. G. Ormiston. 1985. *Peromyscus leucopus*. Mamm. Species
706 247:1–10.
707 Laland, K. N. 1994. On the evolutionary consequences of sexual imprinting. Evolution 48:477–
708 489.
709 Latour, Y., M. Perriat-Sanguinet, P. Caminade, P. Boursot, C. M. Smadja, and G. Ganem. 2014.
710 Sexual selection against natural hybrids may contribute to reinforcement in a house mouse
711 hybrid zone. Proc. R. Soc. B. 281:20132733.
712 Lovecky, D. V., D. Q. Estep, and D. A. Dewsbury. 1979. Copulatory behavior of cotton mice
713 (*Peromyscus gossypinus*) and their reciprocal hybrids with white-footed mice (*P. leucopus*).
714 Anim. Behav. 27:371–375.
715 Lunter, G., and M. Goodson. 2011. Stampy: a statistical algorithm for sensitive and fast mapping
716 of Illumina sequence reads. Genome Res. 21:936–939.
717 Maddock, M. B., and W. D. Dawson. 1974. Artificial insemination of deermice (*Peromyscus*
718 *maniculatus*) with sperm from other rodent species. J. Embryol. Exp. Morphol. 31:621–634.
719 Matsubayashi, K. W., and H. Katakura. 2009. Contribution of multiple isolating barriers to
720 reproductive isolation between a pair of phytophagous ladybird beetles. Evolution 63:2563–
721 2580.
722 McCarley, W. H. 1954a. Natural hybridization in the *Peromyscus leucopus* species group of
723 mice. Evolution 8:314–323.

McCarley, W. H. 1954b. The ecological distribution of the *Peromyscus leucopus* species group in eastern Texas. Ecology 35:375–379.

McCarley, W. H. 1954c. Fluctuations and structure of *Peromyscus gossypinus* populations in eastern Texas. J. Mammal. 35:526–532.

McCarley, W. H. 1963. Distributional relationships of sympatric populations of *Peromyscus leucopus* and *P. gossypinus*. Ecology 44:784–788.

Mcvean, G. 2009. A genealogical interpretation of principal components analysis. PLoS Genet. 5:e1000686.

McCarty, R., and C. H. Southwick. 1977a. Cross-species fostering: effects on the olfactory preference of *Onychomys torridus* and *Peromyscus leucopus*. Behav. Biol. 19:255–260.

McCarty, R., and C. H. Southwick. 1977b. Patterns of parental care in two Cricetid rodents, *Onychomys torridus* and *Peromyscus leucopus*. Anim. Behav. 25:945–948.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.

Mendelson, T. C. 2003. Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: *Etheostoma*). Evolution 57:317–327.

Montero, I., M. Teschke, and D. Tautz. 2013. Paternal imprinting of mating preferences between natural populations of house mice (*Mus musculus domesticus*). Mol. Ecol. 22:2549–2562.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283–292.

Noor, M. A. F. 1997. How often does sympatry affect sexual isolation in *Drosophila*? Am. Nat. 149:1156–1163.

747 Nosil, P. 2007. Divergent host plant adaptation and reproductive isolation between ecotypes of
748 *Timema cristinae* walking sticks. Am. Nat. 169:151–162.
749 Nosil, P., R. Riesch, and M. Muschick. 2013. Climate affects geographic variation in host-plant
750 but not mating preferences of *Timema cristinae* stick-insect populations. Evol. Ecol. Res.
751 15:1–16.
752 Patterson, N., A. L. Price, and D. Reich. 2006. Population structure and eigenanalysis. PLoS
753 Genet. 2:2074–2093.
754 Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest
755 RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and
756 non-model species. PLoS One 7:e37135.
757 Platt, R. N., B. R. Amman, M. S. Keith, C. W. Thompson, and R. D. Bradley. 2015. What Is
758 Peromyscus? Evidence from nuclear and mitochondrial DNA sequences suggests the need
759 for a new classification. J. Mammal. 96:708–719.
760 Pournelle, G. H. 1952. Reproduction and early post-natal development of the cotton mouse,
761 *Peromyscus gossypinus*. J. Mammal. 33:1–20.
762 Price, P. K., and M. L. Kennedy. 1980. Genic relationships in the white-footed mouse,
763 *Peromyscus leucopus*, and the cotton mouse, *Peromyscus gossypinus*. Am. Midl. Nat.
764 103:73–82.
765 Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using
766 multilocus genotype data. Genetics 155:945–959.
767 Ramsey, J., H. D. Bradshaw, and D. W. Schemske. 2003. Components of reproductive isolation
768 between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). Evolution
769 57:1520–1534.

Robbins, L. W., M. H. Smith, M. C. Wooten, and R. K. Selander. 1985. Biochemical polymorphism and its relationship to chromosomal and morphological variation in *Peromyscus leucopus* and *Peromyscus gossypinus*. J. Mammal. 66:498–510.

Roehrs, Z. P., J. B. Lack, C. E. S. Jr., C. J. Seiden, R. Bastarache, W. D. Arbour, D. M. L. Jr., and R. A. Van Den Bussche. 2012. Mammals of Red Slough Wildlife Management Area, with comments on McCurtain county, Oklahoma. Texas Tech Univ. Occas. Pap. 309:1-24.

Rolán-Alvarez, E., and A. Caballero. 2000. Estimating sexual selection and sexual isolation effects from mating frequencies. Evolution 54:30–36.

Rosenberg, N. A. 2004. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4:137–138.

Rosenthal, G. G. 2017. Mate choice: the evolution of sexual decision making from microbes to humans. Princeton University Press, Princeton.

Schug, M. D., S. H. Vessey, and E. M. Underwood. 1992. Paternal behavior in a natural population of mice (*Peromyscus leucopus*). Am. Midl. Nat. 127:373–380.

Seehausen, O., J. M. van A. Jacques, and F. Witte. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. Science 277:1808–1811.

Servedio, M. R., S. A. Sæther, and G.-P. Sætre. 2009. Reinforcement and learning. Evol. Ecol. 23:109–123.

Slagsvold, T., B. T. Hansen, L. E. Johannessen, and J. T. Lifjeld. 2002. Mate choice and imprinting in birds studied by cross-fostering in the wild. Proc. R. Soc. B. 269:1449–1455.

Smadja, C. M., and R. K. Butlin. 2011. A framework for comparing processes of speciation in the presence of gene flow. Mol. Ecol. 20:5123–5140.

Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative

793 phylogeography of unglaciated eastern North America. Mol. Ecol. 15:4261–4293.
794 Taylor, R. J., and H. McCarley. 1963. Vertical distribution of *Peromyscus leucopus* and *P.*
795 *gossypinus* under experimental conditions. Southwest. Assoc. Nat. 8:107–115.
796 ten Cate, C., and G. Mug. 1984. The development of mate choice in zebra finch females.
797 Behaviour 90:125–150.
798 ten Cate, C., and C. Rowe. 2007. Biases in signal evolution: learning makes a difference. Trends
799 Ecol. Evol. 22:380–387.
800 ten Cate, C., M. N. Verzijden, and E. Etman. 2006. Sexual imprinting can induce sexual
801 preferences for exaggerated parental traits. Curr. Biol. 16:1128–1132.
802 ten Cate, C., and D. R. Vos. 1999. Sexual imprinting and evolutionary processes in birds: a
803 reassessment. Pp. 1–31 in P. J. Slater, J. S. Rosenblatt, C. T. Snowdon, and T. J. Roper, eds.
804 Mate Choice. Academic Press.
805 Tramm, N. A., and M. R. Servedio. 2008. Evolution of mate-choice imprinting: competing
806 strategies. Evolution 62:1991–2003.
807 Turner, L. M., D. J. Schwahn, and B. Harr. 2011. Reduced male fertility is common but highly
808 variable in form and severity in a natural mouse hybrid zone. Evolution 66:443–458.
809 Verzijden, M. N., and C. ten Cate. 2007. Early learning influences species assortative mating
810 preferences in Lake Victoria cichlid fish. Biol. Lett. 3:134–136.
811 Verzijden, M. N., and G. G. Rosenthal. 2011. Effects of sensory modality on learned mate
812 preferences in female swordtails. Anim. Behav. 82:557–562.
813 Verzijden, M. N., R. F. Lachlan, and M. R. Servedio. 2005. Female mate-choice behavior and
814 sympatric speciation. Evolution 59:2097–2108.
815 Verzijden, M. N., C. Cate, M. R. Servedio, G. M. Kozak, J. W. Boughman, and E. I. Svensson.

816 2012a. The impact of learning on sexual selection and speciation. Trends Ecol. Evol.
817 27:511–519.
818 Verzijden, M. N., Z. W. Culumber, and G. G. Rosenthal. 2012b. Opposite effects of learning
819 cause asymmetric mate preferences in hybridizing species. Behav. Ecol. 23:1133–1139.
820 Vos, D. R. 1995. Sexual imprinting in zebra-finch females: do females develop a preference for
821 males that look like their father? Ethology 99:252–262.
822 Weber, J. N., M. B. Peters, O. V Tsyusko, C. R. Linnen, C. Hagen, N. A. Schable, T. D.
823 Tuberville, A. M. McKee, S. L. Lance, K. L. Jones, H. S. Fisher, M. J. Dewey, H. E.
824 Hoekstra, and T. C. Glenn. 2010. Five hundred microsatellite loci for *Peromyscus*. Conserv.
825 Genet. 11:1243–1246.
826 Wilson, R. C., T. Vacek, D. L. Lanier, and D. A. Dewsbury. 1976. Open-field behavior in
827 muroid rodents. Behav. Biol. 506:495–506.
828 Wolfe, J. L., and A. V Linzey. 1977. *Peromyscus gossypinus*. Mamm. Species 70:1–5.
829 Yukilevich, R., and J. R. True. 2008. Incipient sexual isolation among cosmopolitan *Drosophila*
830 *melanogaster* populations. Evolution 62:2112–2121.
831 Zimmerman, E. G., C. W. Kilpatrick, and B. J. Hart. 1978. The genetics of speciation in the
832 rodent genus *Peromyscus*. Evolution 32:565–579.
833

834 **Figure Legends**
835
836 **Figure 1.** Hybridization is rare between sympatric *P. leucopus* and *P. gossypinus* mice. (**A**)
837 Genetic PCA discriminates between species. The first PC strongly separates species based on
838 known *P. leucopus* (green dots) and *P. gossypinus* (blue dots) mice. The second PC detects
839 population structure within *P. gossypinus* that largely corresponds to mice collected east (higher
840 values) and west (lower values) of the Mississippi river. Known lab-generated F1 and backcross
841 (F1 x *P. gossypinus*) hybrids (cyan dots) fall intermediate along PC1. Mice collected from
842 sympatry (grey dots) cluster discretely with *P. leucopus* or *P. gossypinus* with the exception of
843 two mice that may be hybrids (arrows), but showing greater *P. leucopus* ancestry. (**B**) A
844 Bayesian admixture model implemented in Structure also supports the partitioning of allopatric
845 and sympatric mice into two clusters corresponding to *P. leucopus* (green) and *P. gossypinus*
846 (blue). Individuals are represented by vertical bars showing their estimated ancestry proportions
847 from each species. Note that Structure assigned ancestry in 28 individuals that were discarded as
848 outliers in the genetic PCA. Populations are labeled numerically (see **C** and Supplemental Table
849 1 for locality information). Structure identified the same two individuals from site 17 as hybrids,
850 but also indicated that three individuals from sites 13 and 20 may also be hybrids (arrows);
851 however these individuals were discarded as outliers in the genetic PCA. (**C**) Range map of the
852 two species: *P. leucopus* (green) and *P. gossypinus* (blue) adapted from (Hall and Kelson 1959;
853 Hall 1981), showing areas of allopatry and sympatry. Pie diagrams show collecting locations and
854 frequencies of each species scaled in size to represent the number of mice sampled at each site.
855 For more information, see Supplemental Table 1. Mice were classified as *P. leucopus* (green), *P.*
856 *gossypinus* (blue), or potential hybrids (cyan) based on the genetic PCA (shown in **A**) and
857 Structure analysis (shown in **B**).
858
859 **Figure 2**. Latency to mating between *P. leucopus* (L), *P. gossypinus* (G) and their hybrids (F1).
860 Estimated days since copulation are shown for conspecific, heterospecific, and backcross mating
861 pairs that produced offspring (sample size in parentheses) in no-choice assays. F1 hybrids were
862 generated with both LxG and GxL crosses. In all pairs, the female is listed first. ** $p = 0.01$.
863
864 **Figure 3.** Mating preferences in two-way choice trials. (**A**) Photograph of the mate-choice
865 apparatus. Center chamber is connected to two test chambers, each housing a "stimulus" animal,
866 separated by gated doors activated by only the "chooser" animal. (**B**) Mating preferences for
867 mice raised by their own parents. *P. leucopus* spent greater time with *P. leucopus* stimuli than
868 both *P. gossypinus* sexes. (**C**) Mating preferences for mice raised by heterospecific foster
869 parents. *P. leucopus* males were significantly affected by cross-fostering ($p = 0.004$), whereas *P.*
870 *leucopus* females were not. Both *P. gossypinus* sexes spent significantly more time with the
871 heterospecific stimulus than when raised by their own parents ($p < 0.001$).
872
873 **Figure 4.** The proportion of time spent with a stimulus predicts mating outcome in trials when
874 mating occurred. Mating occurred in 12 trials when choosers were raised with their own parents
875 (gray triangles) and 8 trials in where choosers were raised with heterospecific parents (black
876 dots). Dotted line indicates the predicted probability for mate choice (conspecific versus
877 heterospecific) given the proportion of time a chooser spent with a conspecific individual, which
878 strongly predicts the mating partner ($p = 0.038$). With the exception of one *P. leucopus* female
879 raised with her own parents, all mice spent more time with their preferred mate.

Figure 1

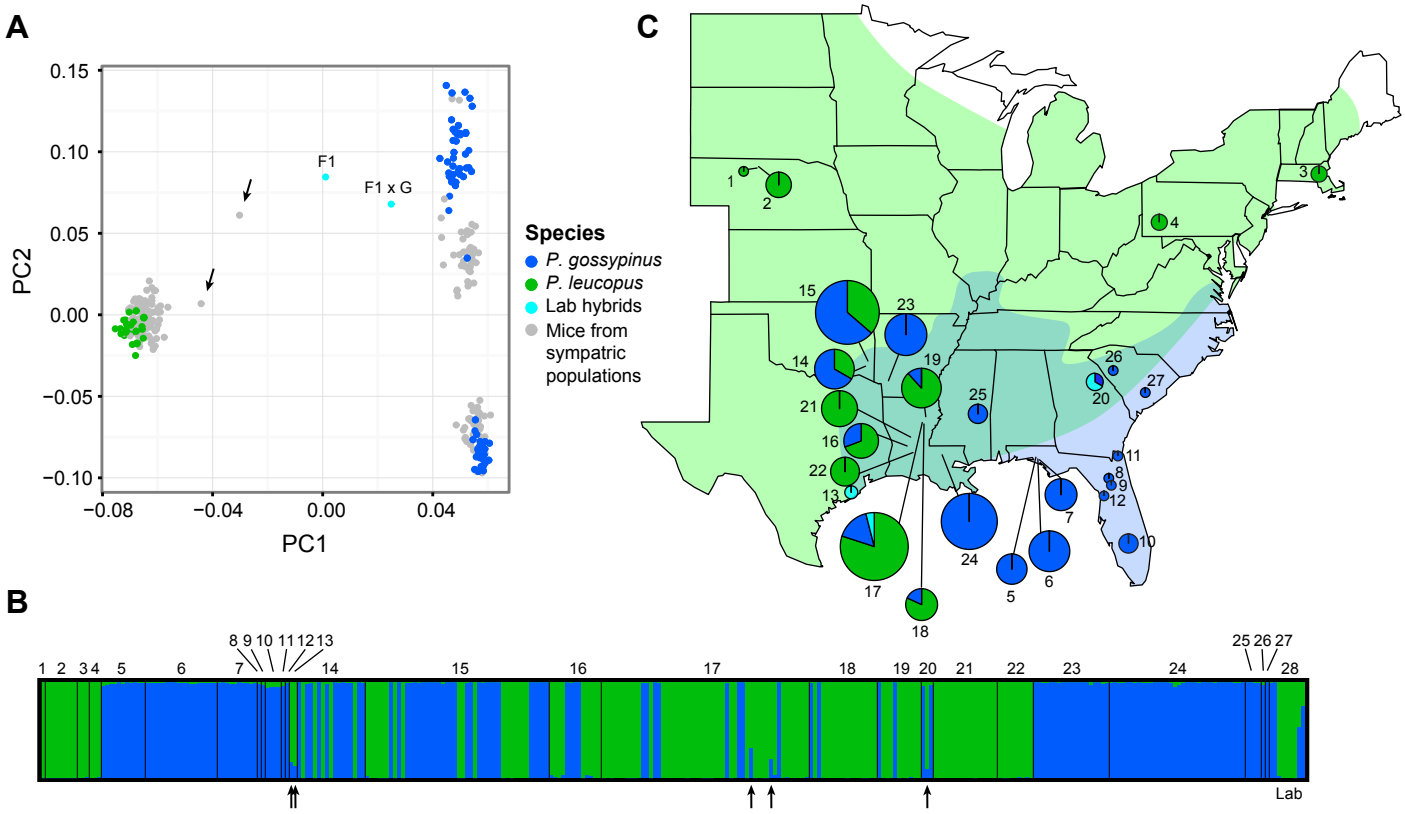

A
0.15
0.10
0.05
0.00
−0.05
−0.10
PC2
−0.08
−0.04
0.00
0.04
PC1
F1
F1 x G
Species
P. gossypinus
P. leucopus
Lab hybrids
Mice from sympatric populations
C
B
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Lab

Figure 2

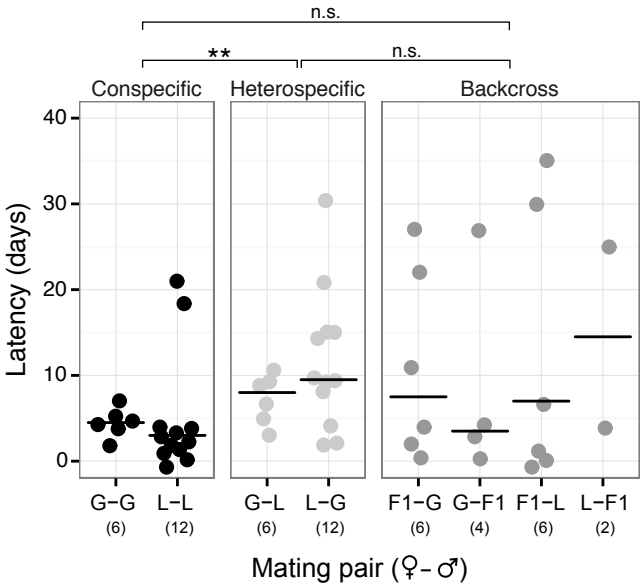
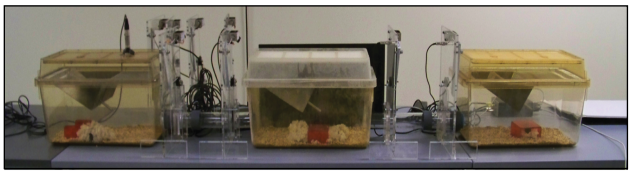
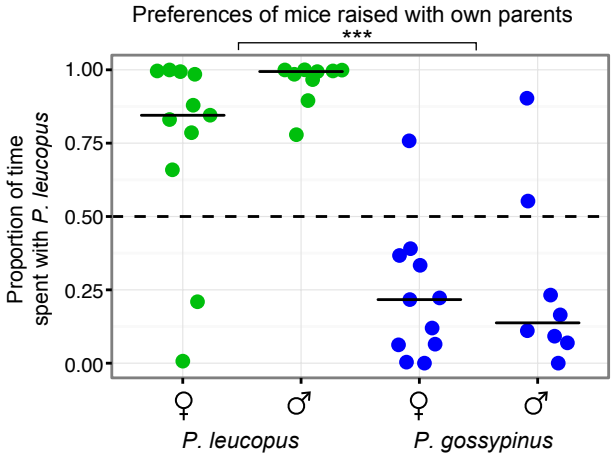

n.s.
**
n.s.
Conspecific
Heterospecific
Backcross
Latency (days)
40
30
20
10
0
G−G
L−L
G−L
L−G
F1−G
G−F1
F1−L
L−F1
(6)
(12)
(6)
(12)
(6)
(4)
(6)
(2)
Mating pair (♀-♂)

Figure 3

**A**



**B**


Preferences of mice raised with own parents
***
Proportion of time spent with P. leucopus
1.00
0.75
0.50
0.25
0.00
♀
♂
♀
♂
P. leucopus
P. gossypinus

**C**

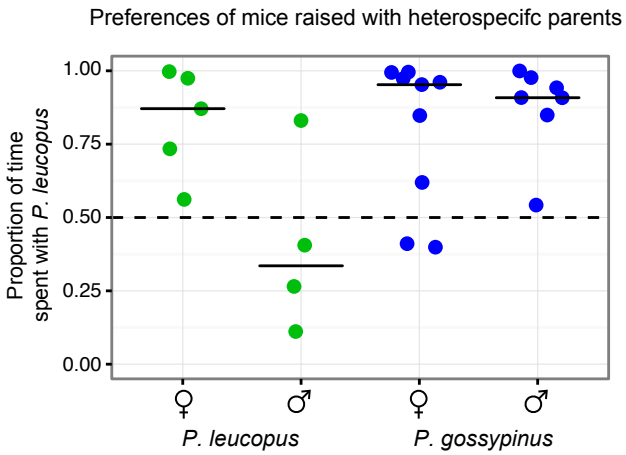
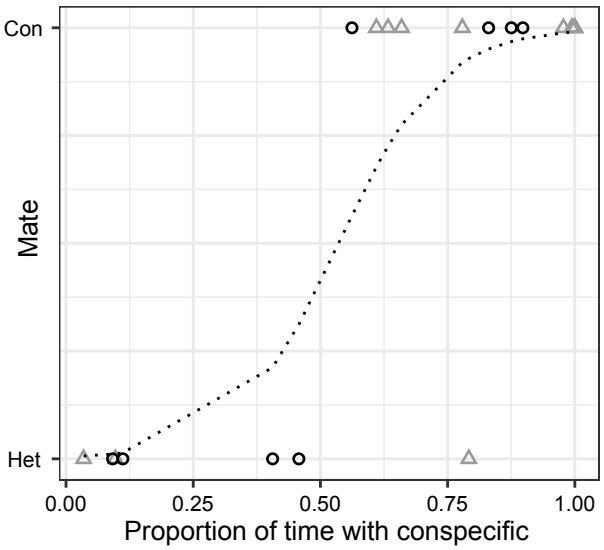
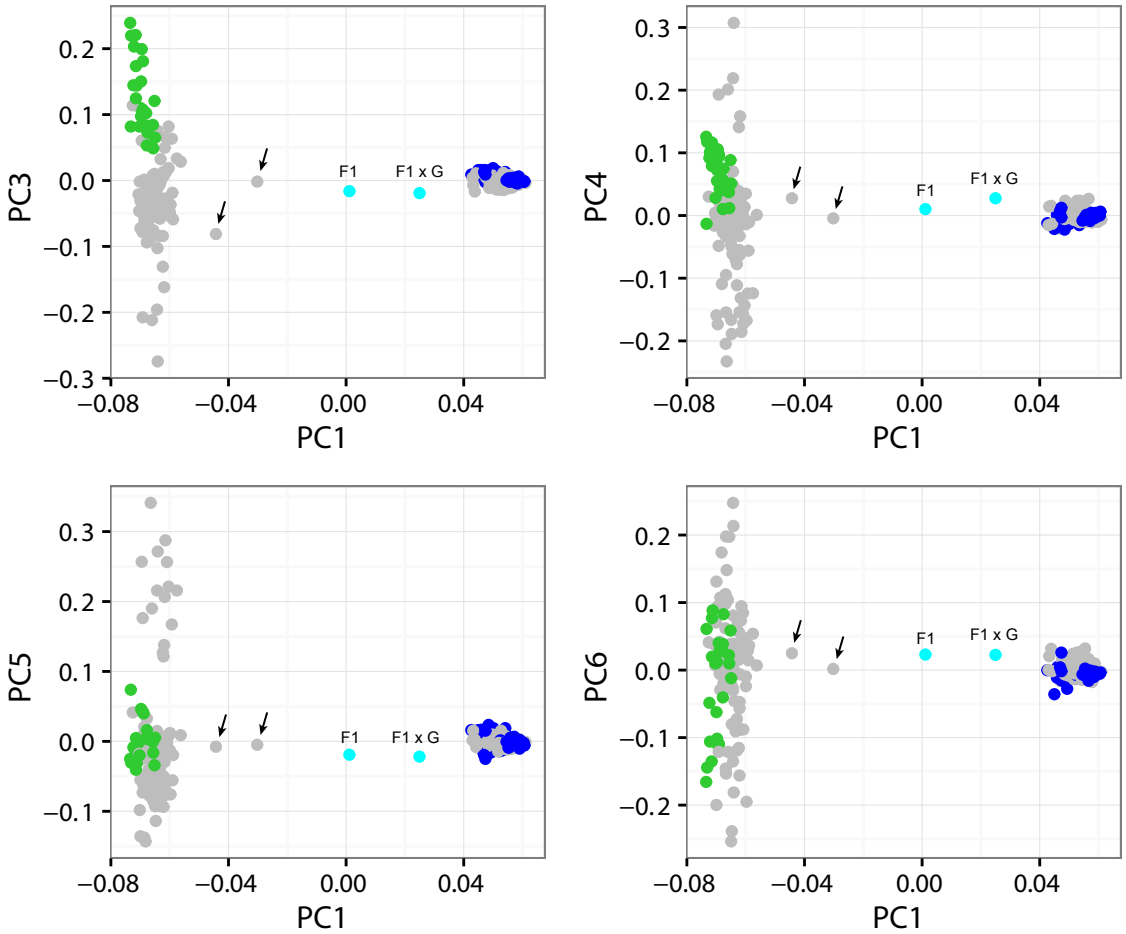
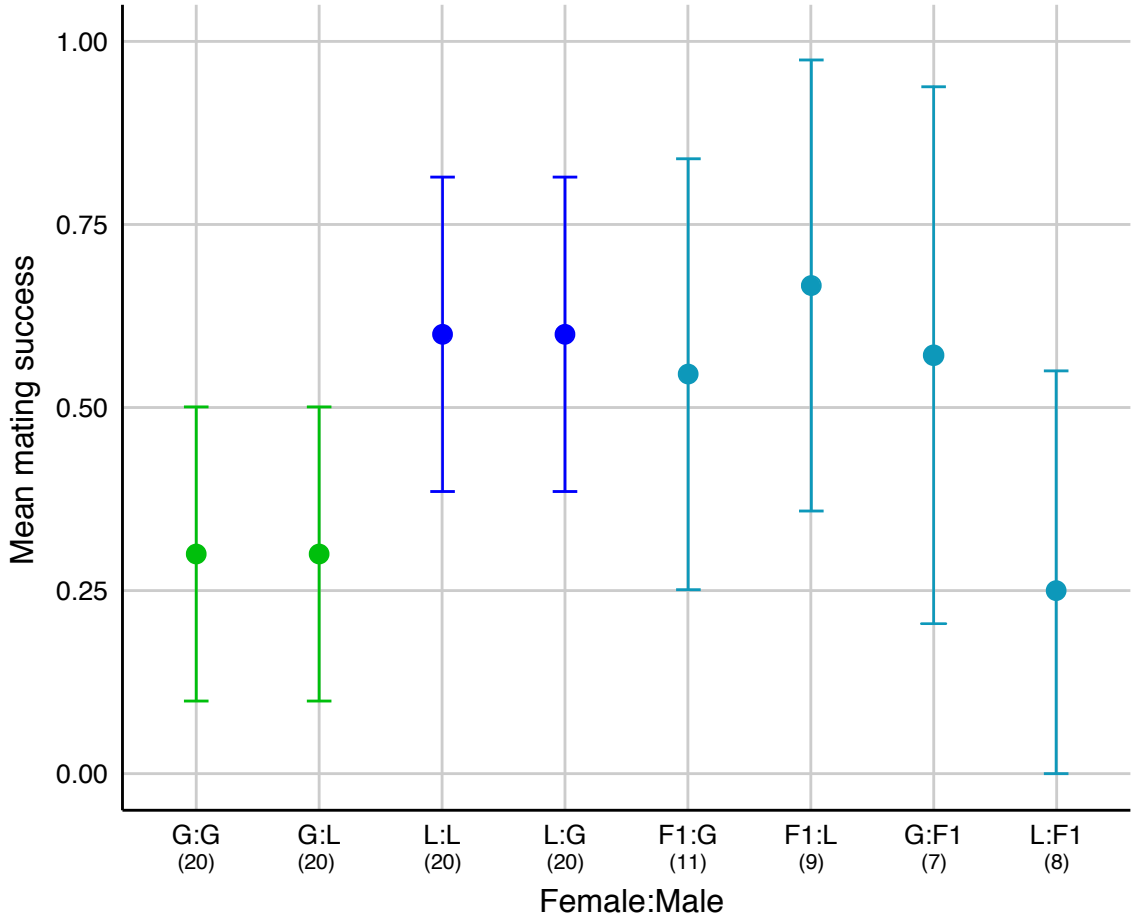

Preferences of mice raised with heterospecifc parents
Proportion of time spent with P. leucopus
1.00
0.75
0.50
0.25
0.00
♀
♂
♀
♂
P. leucopus
P. gossypinus

Figure 4

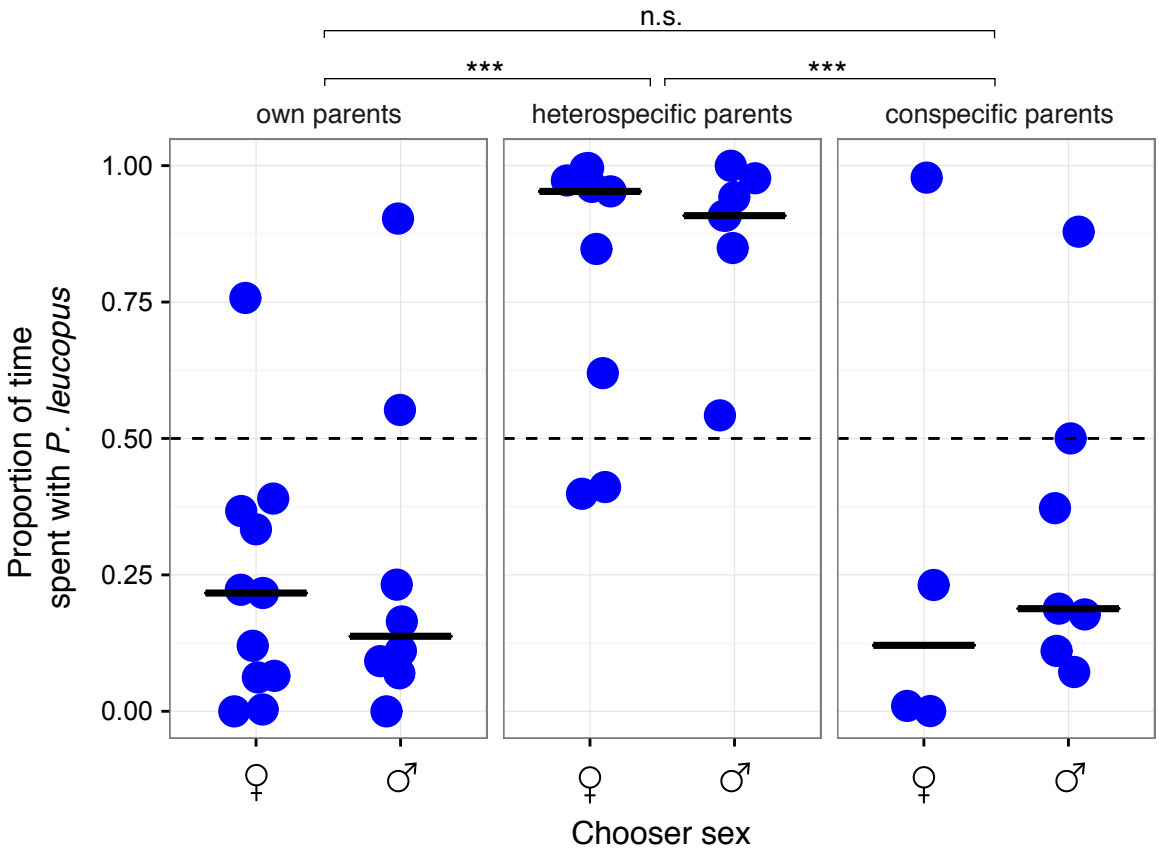

Con
Mate
Het
0.00
0.25
0.50
0.75
1.00
Proportion of time with conspecific




**Supplemental Figure 1.** The first principal component (PC) separates the species *P. leucopus*, *P. gossypinus* and their possible hybrids. Mice collected from sympatry (gray dots) cluster discretely with either known allopatric *P. leucopus* (green dots) or known *P. gossypinus* (blue dots) with the exception of two individuals that may be hybrids (arrows) but show more *P. leucopus* ancestry. The possible hybrids fall intermediate along PC1, similar to known lab-generated F1 and backcross (F1 x *P. gossypinus*) hybrids (cyan dots). PCs 3, 4, 5, and 6 identified population structure within *P. leucopus*.




**Supplemental Figure 2.** Mean proportion of mating successes, defined as the production of a litter, in no-choice trials with *P. leucopus*, *P. gossypinus*, and F1 hybrids (from L:G and G:L crosses). Proportions (dots) and their 95% confidence intervals are plotted for each cross (sample size in parentheses). Confidence intervals overlap for all cross types, indicating that hybrids do not suffer reduced mating success in backcrosses compared to conspecific crosses (G:G or L:L).




**Supplemental Figure 3.** *Peromyscus gossypinus* mating preferences for mice raised with their own parents, heterospecific parents, or unrelated conspecific parents. *P. gossypinus* raised with heterospecific parents differed signifcantly from mice raised with their own parents (t = -7.04, df = 28.89, $p_{Bonferroni} < 0.01$) and mice raised with conspecific parents (t = 4.31, df = 18.90, $p_{Bonferroni} < 0.01$), but *P. gossypinus* raised with their own parents did not significantly differ from mice raised with unrelated conspecific parents (t = -0.72, df = 15.38, $p_{Bonferroni} = 1$).

**Supplemental Table 1.** Trapping locations and specimen information.

| Site number | Site name | County/ Parish | State | Latitude/Longitude | Sample size | Specimen number |
|---|---|---|---|---|---|---|
| 1 | Valentine Fish Hatchery | Cherry | NE | 42.890533, -100.525666 | 1 | MCZ: 66476 |
| 2 | Borman Bridge Wildlife Management Area | Cherry | NE | 42.851566, -100.520633 | 8 | MCZ: 66485, 66491, 66492, 66493, 66494, 66497, 66502, 66607 |
| 3 | Cambridge | Middlesex | MA | 42.383328, -71.115940 | 3 | EHK: 1FMA, 2MMA<br>MCZ: 63293 |
| 4 | Powdermill Nature Reserve | Westmoreland | PA | 40.160940, -79.271797 | 3 | EPK: 087, 088, 096 |
| 5 | Lafayette Creek Wildlife Management Area | Walton | FL | 30.527960, -86.052775 | 11 | MCZ: 68628, 68629, 68630, 68631, 68632, 68633, 68634, 68635, 68636, 68638, 68640 |
| 6 | Nokuse Plantation | Walton | FL | 30.452233, -85.950733 | 18 | EHK: 203, 204, 205, 207, 213, 214, 215<br>MCZ: 68652, 68653, 68654, 68655, 68656, 68657, 68658, 68659, 68660, 68661, 68662 |
| 7 | Pine Log State Forest | Washington | FL | 30.421503, -85.869132 | 10 | MCZ: 68641, 68642, 68643, 68644, 68645, 68646, 68647, 68648, 68649, 68650 |
| 8 | Gainesville | Alachua | FL | 29.649743, -82.253675 | 1 | TK: 157309 |
| 9 | Paynes Prairie Preserve State Park | Alachua | FL | 29.529491, -82.297902 | 1 | MCZ: 68621 |
| 10 | N/A | Highlands | FL | N/A | 4 | FMNH: 31436, 31437, 31441, 31442 |
| 11 | N/A | Nassau | FL | N/A | 1 | TK: 157304 |
| 12 | N/A | Levy | FL | N/A | 1 | TK: 157338 |
| 13 | Nannie M. Stringfellow Wildlife Management Area | Brazoria | TX | 28.964713, -95.615675 | 2 | MCZ: 68799, 68800 |
| 14 | Red Slough Wildlife Management Area | McCurtain | OK | 33.75314, -94.65902;<br>33.75068, -94.64888;<br>33.72513, -94.70072;<br>33.71331, -94.60746;<br>33.70822, -94.63638;<br>33.73742, -94.65853 | 17 | OK: 11702, 11705, 11709;<br>11718, 11720, 11722, 11723, 11724, 11725;<br>11764;<br>11806, 11808, 11810;<br>11813, 11821;<br>11826, 11827 |
| 15 | Honobia | Le Flore | OK | 34.53903, -94.88821 | 46 | OCGR: 5240, 5333, 5334, 5409, 5469, 5569, 7857, 7858, 7967, 7969, 8589, 8677, 8678, 8704, 8705, 8726, 8727, 8728, 8761, 8797, 8935, 8936, 9396, 9398, 9399, 9444, 9445, 9946, 9506, 9507, 9508, 9509, 9511, 9512, 9515, 9516, 9519, 9522, 9523, 9524, 9525, 9534, 9535, 9536, 9537, 9538 |
| 16 | Alexander State Forest | Rapides | LA | 31.157083, -92.469283;<br>31.170700, -92.472567;<br>31.162867, -92.470967 | 13 | MCZ: 68663, 68664, 68665, 68666;<br>68668, 68670;<br>68671, 68672, 68673, 68674, 68676, 68677, 68678 |
| 17 | Big Lake Wildlife Management Area | Tensas | LA | 32.104517, -91.497400;<br>32.160017, -91.487133 | 52 | EHK: 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548;<br>549, 550, 551, 552, 553, 554, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 571, 572, 573, 574, 575, 664, |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | 665, 666, 669, 698<br>MCZ: 68815 |
| 18 | Buckhorn Wildlife Management Area | Tensas | LA | 32.057367, -91.414000 | 17 | EHK: 510, 511, 512, 513, 514, 515, 516, 517, 518, 520, 521, 522, 523, 524, 525, 526, 527 |
| 19 | Tensas River National Wildlife Refuge | Madison | LA | 32.349658, -91.376399 | 11 | EHK: 689, 704, 707, 711, 712, 713, 714, 715, 717, 724<br>MCZ: 68791 |
| 20 | Hart Creek | McDuffie | GA | 33.593730, -82.558806 | 3 | EHK: 140, 144, 152 |
| 21 | Dewey W. Wills Wildlife Management Area | La Salle | LA | 31.495000, -92.011717;<br>31.508417, -92.032200 | 16 | MCZ: 68689, 68690, 68691, 68692, 68693, 68694, 68695, 68696, 68697, 68698;<br>68700, 68701, 68702, 68703, 68704, 68705 |
| 22 | Sherburne Wildlife Management Area | Pointe Coupee | LA | 30.499455, -91.685097 | 9 | MCZ: 68679, 68680, 68681, 68682, 68683, 68684, 68685, 68686, 68688 |
| 23 | Sikes Family Farm | Little River | AR | 33.822863, -94.293169 | 19 | EHK: 596, 601, 605, 608, 614, 615, 636, 637, 638, 648, 649, 650<br>MCZ: 68802, 68803, 68808, 68809, 68811, 68813, 68814 |
| 24 | Tunica Hills Wildlife Management Area | West Feliciana | LA | 30.937317, -91.508817 | 34 | EHK: 309, 327<br>MCZ: 68715, 68716, 68718, 68729, 68731, 68733, 68734, 68735, 68736, 68738, 68739, 68740, 68741, 68742, 68743, 68744, 68746, 68748, 68751, 68752, 68753, 68754, 68755, 68756, 68757, 68758, 68759, 68762, 68763, 68764, 68765, 68766 |
| 25 | Mississippi State University Coastal Plain Branch Experiment Station | Newton | MS | 32.334130, -89.077485 | 4 | EHK: 469, 470, 471, 474 |
| 26 | Ninety Six National Historic Site | Greenwood | SC | 34.138382, -82.023579 | 1 | EHK: 49 |
| 27 | Francis Beidler Forest | Dorchester | SC | 33.221088, -80.350820 | 1 | EHK: 92 |
| 28 | Lab colony | - | - | - | 9 | F025, F028 (*P. gossypinus*)<br>LLF045, LLM058, LLM061, LLF079, LLF081 (*P. leucopus*)<br>M074x (F1 hybrid)<br>Mbxunp1 (F1 x *P. gossypinus*) |

EHK = Tail tip tissue collected by Emily K. Delaney
MCZ = Museum of Comparative Zoology, Harvard University
FMNH = Florida Museum of Natural History
OK = Collection of Vertebrates, Oklahoma State University
OCGR = Oklahoma Collection of Genomic Resources, Sam Noble Museum of Natural History, Oklahoma University
TK = Collection of Genomic Resources, Museum of Texas Tech University, Texas Tech University

**Supplemental Table 2.** Linear models for mating success in conspecific and heterospecific no-choice trials.

| Model | Log-likelihood | Number of Parameters | AIC | ΔAIC |
|---|---|---|---|---|
| Mating success* ~ Female species + Male species + Female species×Male species | -51.355 | 4 | 110.710 | - |
| Mating success ~ Female species + Male species | -51.355 | 3 | 108.710 | 2.000 |
| Mating success ~ Female species | -51.355 | 2 | 106.710 | 2.000 |

*Mating success is defined as a confirmed copulation event (either the birth of a litter or presence of sperm).

**Supplemental Table 3.** Mating results for no-choice, female-choice, and male-choice trials with non-cross-fostered and cross-fostered mice.

| Trial type (Female x Male) | Matings | Number of trials |
|---|---|---|
| No-choice | | |
| G x G | 6 | 20 |
| G x L | 6 | 20 |
| L x G | 12 | 20 |
| L x L | 12 | 20 |
| | | |
| Female-choice* | | |
| G x G | 10 | 11 |
| G x L | 1 | 11 |
| L x G | 2 | 11 |
| L x L | 9 | 11 |
| | | |
| Male-choice* | | |
| G x G | 6 | 8 |
| G x L | 2 | 8 |
| L x G | 0 | 9 |
| L x L | 9 | 9 |
| | | |
| Female-choice (cross-fostered)* | | |
| G x G | 2 | 9 |
| G x L | 7 | 9 |
| L x G | 0 | 5 |
| L x L | 5 | 5 |
| | | |
| Male-choice (cross-fostered)* | | |
| G x G | 0 | 7 |
| G x L | 7 | 7 |
| L x G | 1 | 4 |
| L x L | 3 | 4 |

*Matings are inferred from female- and male-choice trials based on proportion of time spent with each stimulus (see Figures 3 and 4). The stimulus individual with whom the chooser spent more time was considered a mate.

**Supplemental Table 4.** Linear models for *Peromyscus leucopus* preference.

| Model | Log-likelihood | Number of Parameters | AIC | ΔAIC |
|---|---|---|---|---|
| Preference* ~ Sex + CrossFoster + Sex×CrossFoster | -12.310 | 5 | 34.620 | - |

*Preference is defined as the arcsin-transformed proportion of time spent with *P. leucopus.*

**Supplemental Table 5.** Linear models for *Peromyscus gossypinus* preference.

| Model | Log-likelihood | Number of Parameters | AIC | Delta AIC |
|---|---|---|---|---|
| Preference* ~ Sex + CrossFoster + Sex×CrossFoster | -10.238 | 5 | 30.477 | 0 |
| Preference ~ ChooserSex + CrossFoster | -10.261 | 4 | 28.523 | 1.954 |
| Preference ~ CrossFoster | -10.502 | 3 | 27.004 | 1.519 |

*Preference is defined as the arcsin-transformed proportion of time spent with *P. leucopus.*