

ORIGINAL ARTICLE

Landscape models for nuclear genetic diversity and genetic structure in white-footed mice (*Peromyscus leucopus*)

ZS Taylor¹ and SMG Hoffman

Dramatic changes in the North American landscape over the last 12 000 years have shaped the genomes of the small mammals, such as the white-footed mouse (*Peromyscus leucopus*), which currently inhabit the region. However, very recent interactions of populations with each other and the environment are expected to leave the most pronounced signature on rapidly evolving nuclear microsatellite loci. We analyzed landscape characteristics and microsatellite markers of *P. leucopus* populations along a transect from southern Ohio to northern Michigan, in order to evaluate hypotheses about the spatial distribution of genetic heterogeneity. Genetic diversity increased to the north and was best approximated by a single-variable model based on habitat availability within a 0.5-km radius of trapping sites. Interpopulation differentiation measured by clustering analysis was highly variable and not significantly related to latitude or habitat availability. Interpopulation differentiation measured as F_{ST} values and chord distance was correlated with the proportion of habitat intervening, but was best explained by agricultural distance and by latitude. The observed gradients in diversity and interpopulation differentiation were consistent with recent habitat availability being the major constraint on effective population size in this system, and contradicted the predictions of both the postglacial expansion and core-periphery hypotheses.

Heredity (2014) **112**, 588–595; doi:10.1038/hdy.2013.140; published online 22 January 2014

Keywords: population genetics; habitat degradation; conservation genetics; mammals; phylogeography; landscape genetics

INTRODUCTION

The documented biological effects of severe habitat destruction and fragmentation include both reduced genetic diversity and increased interpopulation differentiation in many species (Cushman *et al.*, 2012). Genetic effects of fragmentation have been well demonstrated in mammals with specific or high resource demands, such as brown bears, wolves and chamois (summarized in Zachos and Hacklaender, 2011). However, genetic differentiation can be difficult to detect, especially in animals that have large individual ranges or strong dispersal capabilities. In part, this may be due to a time lag in the response of genetic diversity to changes in habitat availability and/or population size (for example, see Anderson *et al.*, 2010; Harrison *et al.*, 2012). In short-lived animals that should show genetic effects quickly, it may also be because conventional isolation-by-distance and clustering analyses can miss differentiation that is distributed in complex, irregular or subtle ways across a landscape (Jenkins *et al.*, 2010; Selkoe *et al.*, 2010).

Primarily because of logistical limitations, population genetic analyses typically focus on individuals or populations distributed over relatively short distances, such as kilometers or tens of kilometers (for examples, see Anderson and Meikle, 2010; Harrison *et al.*, 2012). However, genetic analyses that incorporate landscape variables have been able to detect subtle genetic patterns in species that otherwise show little population genetic structure over distances of tens to hundreds of kilometers, such as marine animals (Selkoe *et al.*, 2010) and vagile terrestrial vertebrates (Robinson *et al.*, 2012), as long as the landscape features are sufficiently variable (Short Bull *et al.*, 2011).

An excellent test case for exploring the limits of landscape genetic analyses is provided by the white-footed mouse, *Peromyscus leucopus*, a common, highly vagile woodland rodent that utilizes various types of microhabitat within deciduous forests (Nupp and Swihart, 2000; Kamler and Pennock, 2004). Its range is very extensive relative to the typical dispersal distance of about 250 m for an individual mouse (Rasmussen, 1964; Krohne *et al.*, 1984), and the landscape of most of its range has been highly modified by humans in the recent past. Across a number of population genetic studies, this species has shown mostly low levels of irregularly distributed nuclear genetic structure (Mossman and Waser, 2001; Anderson and Meikle, 2010; Munshi-South and Kharchenko, 2010; Rogic *et al.*, 2013). It is thus an organism whose population genetics are very difficult to interpret without considering the effects of the present and past landscapes on its distribution (Rogic *et al.*, 2013).

Similar to many animal species in the northern part of North America, *P. leucopus* was eradicated in the Great Lakes region during the last glacial period and re-established by postglacial colonization from scattered southern refugia about 12 000 years ago (Pielou, 1992; Soltis *et al.*, 2006). Such large colonization events leave genetic signatures that can be recovered using neutral molecular markers (Soltis *et al.*, 2006). For instance, a combination of stepwise dispersal and leading-edge effects are expected to produce a decrease in genetic diversity and an increase in interpopulation divergence in the direction of colonization (Rowe *et al.*, 2006). This has in fact been observed for several Midwestern species, including *P. leucopus*, using mitochondrial DNA markers. *P. leucopus* appears to have colonized

Department of Biology, Miami University, Oxford, OH, USA

¹Current address: Department of Biology, Thomas More College, 333 Thomas More Parkway, Crestview Hills, KY 41017, USA

Correspondence: Dr ZS Taylor, Department of Biology, Thomas More College, 333 Thomas More Parkway, Crestview Hills, KY 41017, USA.

E-mail: taylorz@thomasmore.edu

Received 19 March 2013; revised 12 August 2013; accepted 30 October 2013; published online 22 January 2014

the Great Lakes region from a single refuge located to the southeast of present-day Ohio, and it shows a decline in mitochondrial haplotype diversity along an axis from southern Ohio to northern Michigan (Rowe *et al.*, 2006).

Post-glacial expansion of *P. leucopus* brought the species to its northern range limit in northern Michigan, where further advancement was likely limited by poor overwinter survivorship (Wolff, 1996; Myers *et al.*, 2009). If the range limit of this species is due to selection acting on its physiological tolerances, proximity to this range limit should result in smaller and more infrequent populations with increased turnover near the periphery (the 'core-periphery hypothesis'; Brown, 1984; Gilpin, 1991). Both the postglacial expansion and core-periphery hypotheses would predict that genetic diversity of *P. leucopus* in northern Michigan should be lower than would be expected based on the amount of habitat available (Brown *et al.*, 1995).

The postglacial landscape of the Great Lakes region has continued to change over the last several 100 years. Habitat conversion for farming and human settlement has removed the vast majority of the heavy forest cover that existed there before European colonization. For *P. leucopus* populations in Ohio and southern Michigan, the scarcity of intact forest habitat in these heavily farmed regions should have resulted in genetic isolation and rapid drift (Kyle and Strobeck, 2001; Cushman *et al.*, 2012). In contrast, mice should exhibit higher diversity and less differentiation in areas such as northern Michigan, where much of the land has been allowed to return to forest after extensive logging and the large-scale abandonment of unproductive farms in the early 1900s (Botti and Moore, 2006). These recent changes would thus be expected to cause a pattern of greater diversity in northern populations, exactly opposite to the pattern expected if postglacial colonization or core-periphery effects were the major determinants of present-day population structure.

We examined microsatellite genetic diversity and analyzed pre-classified land cover data along a 700-km transect from southwestern Ohio to northern Michigan, to answer two main questions about the relative influence of historic events versus current landscape on the genetic structure of *Peromyscus leucopus* in the upper Midwest. Question 1: we asked whether changes in *P. leucopus* nuclear genetic structure and diversity are more consistent with being determined by postglacial expansion from southern refugia, by core-periphery effects,

by habitat availability or by isolation-by-distance. Genetic diversity should decrease and interpopulation divergence should increase to the north if postglacial colonization and/or core-periphery effects are key determinants, or to the south if habitat availability is the primary determinant (Kyle and Strobeck, 2001; Rowe *et al.*, 2006; Soltis *et al.*, 2006; Cushman *et al.*, 2012); interpopulation differentiation should be highest at middle latitudes if only isolation-by-distance is a factor in population structure. To address this question, we characterized populations for genetic diversity, interpopulation differentiation and isolation-by-distance, and we obtained indices of habitat distribution in the vicinity of each study site at three different spatial scales. Question 2: we asked whether interpopulation differentiation is best explained by the nature of the landscape separating populations. We evaluated landcover-based measurements using an information theoretical approach to find the best explanatory model for nuclear genetic distances between populations.

MATERIALS AND METHODS

Population genetic structure

Sampling and genotyping. Tissue samples were obtained from 12 *P. leucopus* populations in Michigan and Ohio (Figure 1b; exact sampling locations are shown in Supplementary Table S1), from 2002 to 2009. We trapped eight of these populations directly, using non-lethal traps (HB Sherman Traps, Tallahassee, FL, USA) baited with whole oats that were placed in the evening and checked the following morning. A small piece of ear tissue was removed from each animal for DNA extraction before it was released. Animals were handled in accordance with guidelines established by the American Society of Mammalogists (Gannon *et al.*, 2007), using a protocol approved by the Institutional Animal Care and Use Committee of Miami University. DNA samples from the Jericho, Bachelor and Richardson populations in southwestern Ohio (Anderson and Meikle, 2010) were obtained from C Anderson (Miami University). DNA samples from the Carter Woods population in northwestern Ohio were obtained from S Vessey (Bowling Green State University).

We isolated DNA from ear tissue using the e.Z.N.A. Tissue DNA Kit (Omega Bio-Tek, Norcross, GA, USA). We selected microsatellite primer sets from the literature and amplified the markers by PCR using Promega Go-Taq Flexi Polymerase system (Promega Corporation, Madison, WI, USA) under empirically determined conditions (Supplementary Table S2). Typical reactions contained 20 ng template DNA, 1.5–3.0 mM MgCl₂ and 0.2 mM each dNTP in 15 µl total volume. The PCR cycle consisted of 94 °C for 2 min, 40 cycles of 30

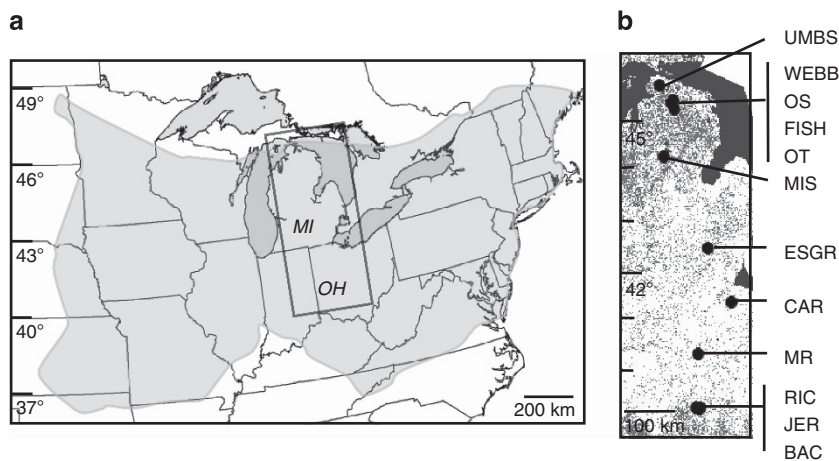


Figure 1 Study area and trapping sites for *Peromyscus leucopus noveboracensis*. (a) Northeastern United States and southeastern Canada, with the range of *P. l. noveboracensis* shown in grey (other *P. leucopus* subspecies are found to the west and south), after Hall (1981); inset highlights study area. OH, Ohio; MI, Michigan. (b) National land cover data set showing forested (gray) and unforested regions in MI and OH, with trapping sites indicated by black circles. Approximate latitude is given in 1° increments.

at 94 °C, 30 s at 50–63.3 °C and 1 min at 72 °C, 5 min at 72 °C. Forward primers labeled with the G5 dye set were obtained from Applied Biosystems (Foster City, CA, USA) and products were run on an Applied Biosystems 3730 DNA Analyzer with the 600LIZ internal size standard (Applied Biosystems). Product peaks were identified manually, using Peak Scanner v1.0 software (Applied Biosystems). Samples that produced ambiguous or negative results on a first attempt were repeated; genotypes that remained ambiguous were repeated until identical duplicate genotypes were obtained. Samples producing consistently ambiguous or negative genotypes after three repetitions were treated as null at that locus. We determined an error rate for genotyping by independently amplifying and genotyping 10% of the samples a second time for each locus; null results attributable to poor amplification were not included in error rate estimates. Any identified errors were corrected by repeated genotyping, as described above.

Basic population genetic analyses. We estimated classical population genetic parameters in GenePop version 4.0 (Rousset, 2008) unless otherwise noted. Data were evaluated for linkage disequilibrium and deviation from Hardy–Weinberg equilibrium (HWE) using a log-likelihood ratio test in GenePop version 4.0 (Rousset, 2008). Significance of results was determined after applying the B–H false discovery rate correction for multiple comparisons to retain detection power while conducting many comparisons (Narum, 2006). The possibility of null alleles, which can lead to failed genotype determination and heterozygote deficiency, was examined using MicroChecker version 2.2.3 (Van Oosterhout *et al.*, 2004) to analyze homozygote excess over the various allele size classes. When null alleles were found to be present at frequencies of over 5%, allele frequencies were adjusted for some analyses using an estimator based on frequencies of heterozygote deficiency and null homozygotes (Brookfield, 1996). This method corrects for bias in allele frequency estimates due to null alleles, by assuming HWE (Van Oosterhout *et al.*, 2004). Using this approach does risk artificially masking some real heterozygote deficiencies found in non-HWE populations, but because null alleles have previously been reported for this species (Anderson and Meikle, 2010; Rogic *et al.*, 2013), we considered underestimates of heterozygosity a greater concern.

Genetic diversity. We measured genetic diversity using allelic richness (A_R), the average number of alleles per locus observed in a population, and also using observed and expected heterozygosity (H_O , H_E). We standardized A_R for sample size using the rarefaction method of HP-Rare version 1.0 (Kalinowski, 2006). H_O and H_E were calculated in Arlequin v3.11 (Excoffier *et al.*, 2005).

Interpopulation differentiation. We analyzed genetic structure using the Bayesian program Structure v2.3.3 (Pritchard *et al.*, 2000) to assign individuals to populations based on the assumption of characteristic allele frequencies and HWE in each population (Pritchard *et al.*, 2000). To determine the number of genetically relevant populations, we performed three runs of Structure at each value of K from 1 to 10, with 10^6 repetitions in each run and a burn-in of 100 000 repetitions. Although the Structure algorithm is known to be problematic when allele frequencies vary continuously due to isolation by distance or high levels of gene flow (Pritchard *et al.*, 2000; Evanno *et al.*, 2005), we used it in conjunction with other methods described below as a broad descriptor of levels of differentiation among our study sites. We determined the number of clusters using the ΔK metric (Evanno *et al.*, 2005), which is based on the second-order rate of change of the posterior probability of the data ($\ln P(D)$) and has been shown to perform better than posterior probability for diagnosing structure under some circumstances (Evanno *et al.*, 2005). We then fixed the value of K based on the mode of ΔK , performed 10 longer runs of 1.5×10^6 generations with the fixed value of $K=2$ and compiled these runs using the program CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007) to produce consensus cluster assignments.

We also measured interpopulation differentiation using F_{ST} and the Cavalli-Sforza and Edwards (1967) chord distance (D_C), two widely used and distinct differentiation metrics for microsatellite data. We estimated F_{ST} as θ , which was calculated from sample-size-weighted variation in allele frequencies in GenePop v4.0 (Weir and Cockerham, 1984; Rousset, 2008). We calculated the Cavalli-Sforza and Edwards chord distance (D_C) in Fstat v2.9.3.2 (Goudet,

1995). We examined correlation between geographical distance and both measures of genetic distance using Mantel tests (Mantel, 1967).

Landscape models of genetic diversity

Landscape characteristics. Landscape analyses were performed in ArcGIS version 9 with the Spatial Analyst extension (ESRI, Redlands, CA, USA). The 2001 national land cover data set (Homer *et al.*, 2004) was obtained for our study area in Michigan and Ohio from the National Map Seamless Server (<http://seamless.usgs.gov>). The national land cover data set classifies the land cover of 30×30 m pixels into 1 of 16 different cover classes (exclusive of Alaska and coastal areas not included in this study). As *P. leucopus* is known to prefer hardwood forest habitat (Parren and Capen, 1985; Cummings and Vessey, 1994; Kamler and Pennock, 2004), we considered categories 41 (deciduous forest) and 43 (mixed forest) to be potential mouse habitat, and other classes to be non-habitat, as described below. However, white-footed mice can utilize other habitat types as sinks, or utilize forest fragments within other types; thus, our methods may underestimate the true availability of habitat for this species.

We calculated the habitat abundance and distribution in the neighborhood of our trapping locations at three scales likely to impact the genetics of local *P. leucopus*. We defined our smallest scale based on Rasmussen's (1964) estimate that the effective population size for mice is approximated by the number of adults in a circle of radius $2s$, where s is the s.d. of the distribution of female dispersal distances between birth and breeding from a study, and thus an estimate of typical dispersal distances. For s , we used a value of 250 m, the midpoint of estimates of typical dispersal distances for *Peromyscus* spp. females from several studies, which range from 200 to 300 m (Rasmussen, 1964; Krohne *et al.*, 1984; Neigel *et al.*, 1991). Therefore, we measured habitat composition within a radius of 0.5, 5 or 50 km. At each spatial scale, we calculated five estimates of habitat abundance/distribution using ArcGIS v9.3 with the Spatial Analyst extension: area of forest habitat coverage; area: perimeter ratio (calculated as total habitat area divided by total habitat perimeter length); average patch size; patch density (number of patches per total area); and average nearest-neighbor distance between patches.

Statistical analyses. We used ordinary least squares simultaneous autoregression (SAR) within an information theoretical framework to identify the model that best explains the relationship between landscape variables and genetic diversity measurements. SAR is a form of spatial regression that incorporates autocorrelation in the residuals of a generalized least squares model (Rangel *et al.*, 2006). Before modeling, we tested for colinearity among landscape variables by using the modified t -test of Dutilleul (1993) to evaluate Pearson's product-moment correlations in the context of autocorrelated spatial data, and did not combine variables that were correlated beyond a cutoff value of $r > |0.6|$ in the same model. We also evaluated each landscape predictor variable for the assumption of normality and used the natural logarithms of the variables for model building when necessary. We performed the modified t -test in PASSAGE 1.1 (Rosenberg, 2001). We produced a set of candidate models by choosing the variables most likely to affect population connectivity in terms of habitat availability (habitat area), patch size or shape (area: perimeter ratio), and patch organization (nearest-neighbor distance). We then examined all 12 of the possible 1-, 2- or 3-variable combination models (for each genetic diversity parameter) that did not combine highly correlated predictor variables. We then ranked models based on AICc, an adjustment of Akaike's Information Criterion (AIC) for small sample sizes (Hurvich and Tsai, 1989). AIC estimates the distance between current and ideal models, and includes a penalty for the incorporation of additional parameters; the model with the lowest AIC value provides the best approximation of the true model (Burnham and Anderson, 1998). We performed SAR and AICc calculations in SAM v4.0 (Rangel *et al.*, 2006).

Landscape models of genetic differentiation

Landscape distances. We quantified landscape distances likely to affect population differentiation in several ways. We chose to use linear landscape distances because (1) over distances of hundreds of kilometers, many dispersal routes are possible, and least-cost distances are therefore more likely to be

artifacts of the algorithm and cost weighting used than routes used by actual mice; (2) with nearly 3×10^8 pixels in our geographical data set, optimization of cost weights is computationally difficult without smoothing that would obscure the small amount of heterogeneity in much of the region. Euclidean distances between trapping sites were calculated using the Geographical Distance Matrix Generator (Ersts, 2010). We also produced land cover-specific distances by quantifying the land cover between each site with the 30 m national land cover data set (Homer *et al.*, 2004) in ArcGIS v9.3 and its Spatial Analyst extension. We used the zonal histogram of a straight-line polyline shapefile drawn between each trapping site to count the number of 30 m pixels of each land cover type that intersect a line drawn between each trapping site pair. We then compiled the 15 recognized land cover classes found in our study area into 5 summary distances, as follows: potential mouse habitat (sum of pixels of land cover classes 41 and 43), open water (class 11), natural non-habitat (classes 31, 42, 52, 71, 90 and 95), agricultural lands (81 and 82) and developed lands (21, 22, 23 and 24). A complete description of the national land cover data set land cover classes can be found at http://www.mrlc.gov/nlcd_definitions.php. As all of these distance values are expected to increase with geographical distance, we also used the habitat:distance ratio (potential habitat distance/total distance) as a distance-independent measure of the environment of each trapping site pair.

Correlation analyses. We analyzed the relationship between assignment to Bayesian clusters and local landscape statistics using SAR implemented in SAM v4.0 (Rangel *et al.*, 2006). To determine whether genetic clustering could be responding to the same landscape characteristics as is genetic diversity, we compared the average proportion of the genome per population assigned to Cluster 1 (as determined by Structure) with the most important variables from genetic diversity analyses: latitude and habitat availability at 0.5 km.

We used two complementary methods to identify landscape variables that were related to genetic distances. First, we used Mantel tests (Mantel, 1967) and partial Mantel tests to identify landscape distances correlated with interpopulation differentiation. For the independent variables, we tested each of the five summary land cover distances individually and in combination, as well as the proportion of intervening distance that was habitat and the latitudinal midpoint of each location pair for a total set of 17 candidate models. For the γ -matrix, we used both D_C from Fstat and the F_{ST} genetic distance matrix produced by GenePop, linearized with respect to genetic differentiation ($F_{ST}/(1-F_{ST})$) (Slatkin, 1995). We performed Mantel and partial Mantel tests in Fstat v2.9.3.2. (Goudet, 1995), and evaluated the significance of results after 10 000 permutations, after applying the B–H false discovery rate correction (Narum, 2006). We use the term r_M below to distinguish statistics produced by Mantel tests from those produced by regression.

To complement the Mantel tests, we used linear mixed models within an information theoretical framework. We used the approach of Selkoe *et al.* (2010) and procedures described by Zuur *et al.* (2009) to model the influence of landscape distances on genetic distances, while incorporating site effects and the non-independence of data points. We fit all models using maximum likelihood and a Gaussian response model. We added two nominal variables, indicating the sites represented by each distance measurement, and used these random-effects variables in our models, along with the fixed effects represented by the landscape distance measurements. We tested a total of 17 candidate models, including the five summary land cover distances in single- or two-variable models, as well as the proportion of intervening distance that was habitat and the latitudinal midpoint of each pair of locations. We constructed linear mixed models and AIC calculations using the lme4 package (Bates *et al.*, 2012) in the program R v2.15.1 (R Development Core Team, 2008). We calculated AICc according to the formula: $AICc = AIC + 2(k+1)/(n-k-2)$, where k is the number of parameters and n is the number of observations (Hurvich and Tsai, 1989). In addition to r^2 and AICc, we provide the deviance (D) as a measure of fit, where $D = -2$ (log likelihood).

RESULTS

Population genetic structure

Sampling and genotyping. One hundred and thirty-seven individuals from 12 populations were trapped and sampled, including 5–20

Table 1 Standard molecular diversity indexes for populations of *Peromyscus leucopus*, listed South to North

Population	N	A	A_R	H_E	H_O
BAC	7	6.5	6.5	0.84	0.65
JER	12	7.9	6.71	0.86	0.68
RIC	12	7.4	6.13	0.84	0.62
MR	20	9.3	6.57	0.85	0.66
CAR	13	7.4	6.32	0.84	0.80
ESGR	13	9.7	7.38	0.88	0.71
MIS	11	8.3	7.50	0.87	0.67
OT	12	10.4	8.15	0.91	0.72
FISH	11	8.8	7.38	0.88	0.80
OS	10	9.0	7.81	0.90	0.73
WEBB	5	5.6	ND	0.89	0.60
UMBS	8	7.7	7.38	0.88	0.74

Abbreviations: A, allelic richness; A_R , allelic richness corrected by rarefaction for a sample size of 7; BAC, Bachelor Woods, OH; CAR, Carter Woods, OH; ESGR, George Reserve, MI; FISH, Fisherman Road, MI; H_E , expected heterozygosity; H_O , observed heterozygosity; JER, Jericho, OH; MIS, Missaukee County, MI; MR, Mercer Co., OH; N, sample size; ND, not determined; OS, Osmun Road, MI; OT, Otsego County, MI; RIC, Richardson, OH; UMBS, UMBS, MI; WEBB, Webb Road, MI. Allelic richness was not determined for the BAC or WEBB populations.

animals per population and at least 10 animals in 9 of the 12 populations (Table 1). Mice were genotyped at 10 loci (Supplementary Table S2), all of which were polymorphic in all populations. The overall rate for incorrect genotyping, measured from random sample repetition, was 0.7%.

Basic population genetic analyses. We detected null alleles at all loci, except PLGT-62. In most cases, detected frequencies were low and allele frequencies were corrected before further analyses. For loci PO-9 and PO-26, the frequency of null alleles exceeded 20%; we therefore excluded these loci from all analyses, except allelic richness. After correction for multiple comparisons with the B–H method, we detected significant deviations from linkage disequilibrium in 1 of 336 locus–locus/population comparisons (Bonferroni or false discovery rate methods) and from HWE in 14 of 96 population–locus comparisons. The 14 HWE deviations involved 9 populations and 5 loci. The MIS population was out of HWE at four loci; no other population deviated at more than two loci. HWE deviations were concentrated at loci PO68 (in four populations), Pml-04 (four populations) and PO35 (three populations).

Genetic diversity. Genetic diversity was highest in northern Michigan (site Otsego County: $A_R = 8.15$; $H_E = 0.91$) and lowest in southern Ohio (site Richardson: $A_R = 6.13$; $H_E = 0.84$; Table 1). Allelic richness (A_R) and expected heterozygosity were significantly positively correlated with each other (linear regression, $r = 0.99$; $P < 0.0001$). A_R was not significantly correlated with observed heterozygosity ($r = 0.10$; $P = 0.79$), possibly reflecting the presence of null alleles.

Interpopulation differentiation. In the Structure genetic clustering analyses, the mode of the ΔK metric for all values of K from 1 to 10 was at $K = 2$ (Table 2). Over 10 runs at $K = 2$ ($\ln P(D) = -5351.20$; $\text{Var}(\ln P(D)) = 234.62$), most individuals from the northern trapping sites were assigned to Cluster 1, whereas assignment of individuals from other sites was split between the two clusters and showed no obvious spatial pattern (Figure 2).

Analyses of these populations using F_{ST} revealed modest levels of differentiation. Overall levels of differentiation were low but significant ($F_{ST} = 0.024$; $P < 0.0001$). The lowest pairwise F_{ST} values were

Table 2 Clustering parameters of *Peromyscus leucopus* samples determined by Structure analyses

K	Ln P(D)	Var [Ln P(D)]	ΔK
1	-5490.4	77.2	NA
2	-5351.2	234.6	9.02
3	-5350.16	553.8	4.47
4	-5243.9	603.0	3.53
5	-5224.4	755.9	1.57
6	-5247.9	955.3	1.21
7	-5234.0	1064.7	3.83
8	-5346.1	1353.3	3.58
9	-5326.2	1461.9	1.62
10	-5368.7	1640	1.05

Abbreviation: NA, not applicable.

The average values from three STRUCTURE runs at each value of K are given for the log probability of data (Ln P(D)), variance (Var [Ln P(D)]) and the ΔK statistic recommended by Evanno *et al.*, (2005).

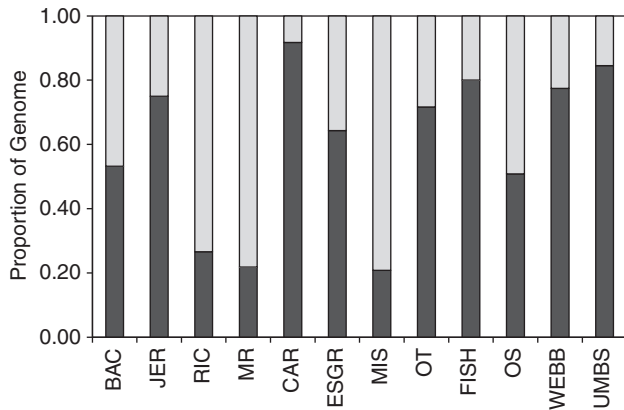


Figure 2 Clustering analysis of *Peromyscus leucopus* from sites in Michigan and Ohio. Structure results for $K=2$ (Ln P(D) = -5351.2; Var [Ln P(D)] = 234.6; $\Delta K=9.02$), showing the total proportion of individual genotypes from each site assigned to Cluster 1 (black) or to Cluster 2 (gray).

between pairs of northern trapping sites such as Otsego and Webb Road ($F_{ST} = -0.0046$; $P=0.43$), whereas the highest were between more southern sites such as Jericho and Carter Woods ($F_{ST}=0.087$; $P<0.0001$). Mantel tests for D_C values showed isolation-by-distance ($r_M=0.55$; $P=0.0001$), whereas F_{ST} values were near significance for correlation with geographical distance ($r_M=0.40$; $P=0.06$).

Landscape models of genetic diversity

Allelic richness was correlated with several measures of habitat availability and with latitude (SAR, $r^2=0.67$; $P=0.002$). The selected model for allelic richness was habitat area at the 0.5 km scale (Figure 3a, Table 3; SAR, $r^2=0.70$; $AICc=23.03$; $P=0.001$); at the larger scales, many of the models were indistinguishable using $AICc$ (Table 3), due to correlations among the predictor variables (that is, predictor variables were more similar at larger scales, because most variation occurs over short distances). A number of the models incorporating nearest-neighbor distance and area provided good descriptions of the data for expected heterozygosity (H_E): the 0.5- and 5-km area models were indistinguishable (Figure 3b). None of the models for observed heterozygosity (H_O) were significant.

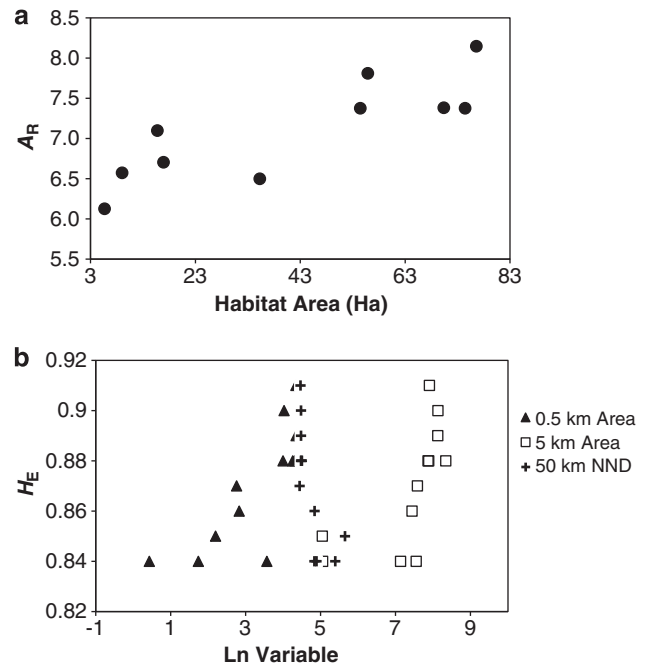


Figure 3 Models of genetic diversity for *Peromyscus leucopus*. (a) Variation in allelic richness (A_R) and habitat availability (total habitat area, in hectares) in Michigan and Ohio, at 0.5 km. (b) Best model for expected heterozygosity at each spatial scale. The natural log of each landscape variable (Ln Variable) is shown on the horizontal axis. Variables shown are total habitat area within a circular buffer of radius 0.5 (filled triangles) or 5 km (open squares), in hectares, and nearest-neighbor distance within 50 km (crosses), in meters.

Table 3 Model selection for regression of *Peromyscus leucopus* genetic diversity on landscape variables

Parameter	Scale (km)	Variable	r	r ²	AICc	$\Delta AICc$
A_R	0.5	Area	0.835	0.698	23.03	0
		APR	0.734	0.539	27.67	4.64
	50	LnNND	0.728	0.530	27.88	4.85
		APR	0.697	0.486	28.87	5.83
H_E	0.5	Area	0.770	0.593	-52.40	0
	5	Area	0.731	0.535	-50.81	1.59
		LnNND	0.717	0.514	-50.29	2.11
50	LnNND	0.697	0.485	-49.59	2.81	

Abbreviations: AIC, Akaike's Information Criterion; APR, area:perimeter ratio; LnNND, natural logarithm of average nearest neighbor distance. Regression parameters and AIC values are given for the best model for each parameter at each spatial scale; models within 2 $AICc$ units are indistinguishable.

Landscape models of genetic differentiation

The five northernmost populations that were identified by Structure as consisting of Cluster 1 mice are located in the most heavily forested portion of the study area, where agricultural lands are almost entirely absent (average agricultural distance = 0.79 km; s.d. = 0.65 km). However, spatial autoregression did not identify a significant correlation of membership in the Structure Cluster 1 with either habitat availability at 0.5 km (SAR, $r^2=0.31$; $P=0.141$) or latitude (SAR, $r^2=0.048$; $P=0.203$), due to the high scatter in Cluster 1 membership among the southern sites (Figure 2).

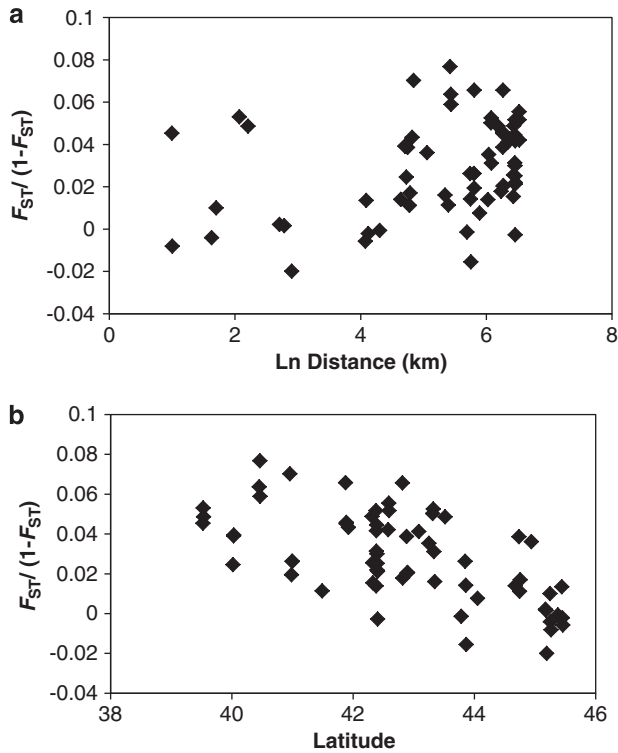


Figure 4 Interpopulation differentiation for populations of *Peromyscus leucopus*. (a) Linearized F_{ST} values compared with the natural log of the Euclidean distance between sites, $r_M=0.40$; $P=0.06$. (b) Comparison of F_{ST} to midpoint latitude; linear mixed model (LMM): $AICc = -369.6$; $r^2 = 0.65$; $D = -381$.

Table 4 Model selection for linear mixed models of *Peromyscus leucopus* genetic and landscape distances

Dependent	Independent	AICc	$\Delta AICc$	r^2	D	Coeff.
F_{ST}	Midpoint latitude	-369.6	0	0.65	-381	-0.0091
	Agricultural distance	-366.2	3.4	0.23	-377.6	0.0044
	Water distance	-361.4	8.2	0.03	-372.8	-0.0017
	Non-hab. nat. areas	-361.2	8.4	0.14	-372.6	-0.0022
	Fraction hab. interven.	-360.3	9.3	0.61	-371.7	-0.0061
D_C	Agricultural distance	-251.2	0	0.26	-261.9	0.010
	Midpoint latitude	-246.5	5.4	0.62	-256.5	-0.016
	Fraction hab. interven.	-245.3	6.6	0.62	-255.3	-0.017
	Water distance	-244.3	6.9	0.05	-255.0	0.0046
	Habitat distance	-243.6	7.6	0.38	-254.3	0.0033

$\Delta AICc$ values indicate the change from the lowest AICc value. Fit statistics r^2 and deviation (D) are also shown for each model along with the estimated coefficient.

In landscape distance modeling with Mantel tests, F_{ST} and D_C were positively correlated to each other ($r_M = 0.46$; $P < 0.0001$). After error rate correction, F_{ST} was significantly correlated with agricultural distance and developed land distance, as well as with proportion of habitat intervening ($r_M = -0.65$; $P < 0.0001$) and midpoint latitude (Figure 4b; $r_M = 0.49$; $P = 0.0003$). D_C was correlated with the proportion of habitat intervening ($r_M = 0.35$; $P = 0.0032$), with midpoint latitude ($r_M = 0.74$; $P < 0.0001$) and with all landscape distances ($0.31 < r_M < 0.42$; $P < 0.0375$), except overwater distances ($r_M = 0.19$; $P = 0.1161$).

In linear mixed modeling, the single-variable models based on agricultural distance, latitude and fraction of habitat intervening were the best descriptors for both of the measures of genetic distance (Table 4): midpoint latitude was best for F_{ST} ($r^2 = 0.65$; $D = -381$; $AICc = -369.6$) and agricultural distance was best for D_C ($r^2 = 0.26$; $D = -261.9$; $AICc = -251.24$). All multivariable models had higher $\Delta AICc$ values (that is, were less descriptive) than the best single-variable models.

DISCUSSION

Genetic change in *P. leucopus*

In our study of Great Lakes white-footed mice (*P. leucopus*), changes in genetic diversity, cluster assignments and interpopulation differentiation were associated with spatial variation in habitat availability. In all cases, correlation is in the direction that would be expected if habitat availability were a major constraint on genetic drift in this species (Hutchison and Templeton, 1999), that is, with lower diversity and higher differentiation to the south. Our results indicate that neither postglacial history nor core-periphery effects determine the genetic structure of *P. leucopus* in this region, because both hypotheses would predict changes in diversity and differentiation with the opposite directionality. In the case of postglacial expansion, increasing differentiation and decreased diversity to the north would be expected as a result of stepwise dispersal and colonization. The fact that this was seen with mitochondrial DNA (Rowe *et al.*, 2006) but not with microsatellite data may be because 12 000 years is too long a span for the genetic signal to be obtainable from microsatellites in a rodent species with large populations and short generation times. In contrast, the core-periphery hypothesis, which predicts the same pattern of higher differentiation and lower diversity to the north as the result of populations at the periphery of the species' range being smaller and rarer than core populations (Brown *et al.*, 1995; Cushman *et al.*, 2012), is clearly rejected by our data.

Drift and landscape genetic structure in *P. leucopus*

Most prior studies of *P. leucopus* have shown significant levels of interpopulation differentiation at the local level, but no isolation-by-distance or spatially consistent distribution of genetic structure (Mossman and Waser, 2001; Anderson and Meikle, 2010; Munshi-South and Kharchenko, 2010). If these patterns were applied uniformly across our study area, we would not expect the gradients in genetic diversity and differentiation described here. However, our results for the southern end of our study are consistent with results from other studies of *P. leucopus* in Indiana and Ohio (Mossman and Waser, 2001; Anderson and Meikle, 2010): both found relatively moderate but extremely variable levels of differentiation in heavily fragmented landscapes. Specific values for differentiation are also consistent across studies: the overall F_{ST} of 0.024 reported here is comparable to values of 0.041 (Anderson and Meikle, 2010), 0.033 (Mossman and Waser, 2001) and pairwise $F_{ST} \leq 0.054$ (Rogic *et al.*, 2013) found in comparable prior studies; a higher value of 0.071 has been reported among populations embedded in the dense urban landscape of New York City (Munshi-South and Kharchenko, 2010).

Although we observed changes in genetic differentiation related to several measures of landscape distance, only agricultural distance stood out as a potential driver of differentiation. This is important because data on use of agricultural areas by *P. leucopus* are equivocal. Demographic studies have shown that white-footed mice travel in a limited fashion through agricultural fields (Cummings and Vessey, 1994; Krohne and Hoch, 1999; Rizkalla and Swihart, 2007). However, genetic studies conducted at a local scale have universally concluded

that agricultural landscapes have little effect on gene flow in this species (Mossman and Waser, 2001; Anderson and Meikle, 2010; Rogic *et al.*, 2013), whereas rivers, highways and urban development can have stronger effects (Munshi-South, 2012; Rogic *et al.*, 2013). Our results suggest a subtle effect of agriculture on dispersal that is too minor to be detected at a small scale, possibly because mice in undetected remnant habitat fragments within the agricultural matrix mediate low levels of gene flow.

Our results from both clustering analyses and distance model selection indicated that populations surrounded by unsuitable habitat (that is, those in the south) tend to be more different from each other than are those surrounded by suitable habitat (in the north), regardless of their distance from each other (Figure 4). As with the observed gradient in genetic diversity, these patterns could result either from accelerated drift and rapid differentiation in small populations whose sizes are limited by habitat availability or from resistance to migration posed by the agricultural landscape. In either case, our data strongly support a role for genetic drift in shaping the genetics of *P. leucopus* across this region. Together with the approximately tenfold scatter we observed in the F_{ST} -distance curve (Figure 4a), these patterns resemble Case III described by Hutchison and Templeton (1999), which is predicted by a stepping-stone model of genetic structure with a regional lack of dispersal-drift equilibrium, and with a larger effect due to drift than to gene flow. This pattern is often seen when a species has limited dispersal ability, as in island populations of Galápagos lava lizards (*Microlophus albemarlensis*; Jordan and Snell, 2008), or when populations are geographically isolated, as in island populations of Aegean wall lizards (*Podarcis erhardii*; Hurston *et al.*, 2009). The fact that the overall level of differentiation in *P. leucopus* is much lower than is typical of island populations could be due to the recent fragmentation of its habitat, or to a superior dispersal ability of these mice.

CONCLUSION

Analyses of the landscape variables that shape population-level processes such as population size and gene flow provide attractive tools for explaining the distribution of natural genetic variation. For the present study, we used a geographical data set of $\sim 2.4 \times 10^8$ pixels to examine landscape genetic processes over 220 000 km², along a transect that lacks major topographical variation or obvious barriers to the dispersal of terrestrial mammals. On this scale, we assume that neither the locations of individual animals nor minor local variation in habitat distribution are likely to have noticeable effects. Therefore, it is reasonable to use a population approach and aggregate landscape characteristics instead of using individual animal locations and dispersal pathways. Our finding of significant relationships between habitat availability and genetic diversity/differentiation is not surprising, given similar demonstrations of landscape genetic effects in the literature (see Storfer *et al.*, 2010); however, if subsets of our data representing local groups of populations were taken separately, the high scatter of data values would resemble those of other studies of this species (Mossman and Waser, 2001; Anderson and Meikle, 2010). In other words, large distances in landscape studies are extremely helpful in providing the context necessary to observe and understand the existing patterns (Cushman *et al.*, 2012) in a way that more localized studies cannot. Population and landscape genetics would be greatly helped by a move towards larger-scale studies (genetically and geographically), which would in turn be helped by the development and use of higher-throughput markers and analyses, and by efficient software methods capable of handling the massive quantities of data required for landscape analyses.

Collectively, the results from this study provide a useful test of predictions about the response of populations to rapid habitat changes across large distances, a topic of increasing importance in a world beset by large-scale climate shifts. The residual effects imposed by the relatively recent glaciation of the Great Lakes region that were shown by mitochondrial DNA (Rowe *et al.*, 2006) are not shown by microsatellite markers for *P. leucopus*. Instead, nuclear genetic diversity increases to the north, as predicted by the local availability of habitat, whereas interpopulation differentiation reflects both latitude and distance between suitable habitat patches. The discordance between mitochondrial and nuclear data, with the latter primarily reflecting very recent habitat fragmentation, shows that the effects of ongoing habitat loss can be separated out from historical influences even when genetic differentiation is low. By using a landscape genetics approach on a broad spatial scale, we were able to distinguish the effects of habitat availability even when the genetic signal was quite weak on a local scale. We suggest that using an increased spatial scale and including the examination of landscape effects can elucidate otherwise obscure population genetic patterns for many other species.

DATA ARCHIVING

Genotype and landscape data available from the Dryad Digital Repository: doi:10.5061/dryad.j272c.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank C Anderson, D Meikle and S Vessey for contributing DNA samples to this study; R Abbitt for assistance with GIS methodologies; the ES George Reserve, University of Michigan Biological Station and the staff of Pigeon River State Forest for permission to conduct research and for additional assistance; D Berg for many conversations and suggestions; P Myers for helpful conversations and field contributions; T Crist and three anonymous reviewers for comments on early versions of this manuscript; R Moscarella, M Steinwald, the Field Ecology class from the University of Michigan and many others for assistance with field or laboratory work; The Miami University Zoology Summer Field Research Workshop and a Department of Zoology Dissertation Scholarship to ZT for funding.

-
- Anderson CD, Epperson BK, Fortin MJ, Holderegger R, James PMA, Rosenberg MS *et al.* (2010). Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol* **19**: 3565–3575.
- Anderson CS, Meikle DB (2010). Genetic estimates of immigration and emigration rates in relation to population density and forest patch area in *Peromyscus leucopus*. *Conserv Genet* **11**: 1593–1605.
- Bates D, Maechler M, Bolker B (2012). Package 'lme4'. Available at <http://cran.stat.sfu.ca/web/packages/lme4/lme4.pdf>
- Botti WB, Moore MD (2006). *Michigan's State Forests: a Century of Stewardship*. Michigan State University Press: East Lansing.
- Brookfield JFY (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* **5**: 453–455.
- Brown JH (1984). On the relationship between abundance and distribution of species. *Am Nat* **124**: 255–279.
- Brown JH, Mehlman DW, Stevens GC (1995). Spatial variation in abundance. *Ecology* **76**: 2028–2043.
- Burnham KP, Anderson DR (1998). *Model selection and Inference: A Practical Information-Theoretic Approach*. Springer: New York.
- Cavalli-Sforza LL, Edwards AWF (1967). Phylogenetic analysis models and estimation procedures. *Am J Hum Genet* **19**: 233–257.
- Cummings JR, Vessey SH (1994). Agricultural influences on movement patterns of white-footed mice (*Peromyscus leucopus*). *Am Midl Nat* **132**: 209–218.
- Cushman SA, Shirk A, Landguth EL (2012). Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landsc Ecol* **27**: 369–380.

- Dutilleul P (1993). Modifying the t-test for assessing the correlation between 2 spatial processes. *Biometrics* **49**: 305–314.
- Ersts PJ (2010). Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.
- Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.
- Gannon WL, Sikes RS, Comm ACU (2007). Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *J Mammal* **88**: 809–823.
- Gilpin M (1991). The genetic effective size of a metapopulation. *Biol J Linn Soc Lond* **42**: 165–175.
- Goudet J (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *J Hered* **86**: 485–486.
- Hall ER (1981). *The Mammals of North America*. John Wiley & Sons: New York.
- Harrison KA, Pavlova AP, Amos JN, Takeuchi N, Lill A, Radford JQ *et al.* (2012). Fine-scale effects of habitat loss and fragmentation despite large-scale gene flow for some regionally declining woodland bird species. *Landscape Ecol* **27**: 813–827.
- Homer C, Huang CQ, Yang LM, Wylie B, Coan M (2004). Development of a 2001 national land-cover database for the United States. *Photogramm Eng Remote Sensing* **70**: 829–840.
- Hurston H, Voith L, Bonanno J, Foufopoulos J, Pafilis P, Valakos E *et al.* (2009). Effects of fragmentation on genetic diversity in island populations of the Aegean wall lizard *Podarcis erhardii* (Lacertidae, Reptilia). *Mol Phylogenet Evol* **52**: 395–405.
- Hurvich CM, Tsai CL (1989). Regression and time-series model selection in small samples. *Biometrika* **76**: 297–307.
- Hutchison DW, Templeton AR (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**: 1898–1914.
- Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.
- Jenkins DG, Carey M, Czerniewska J, Fletcher J, Hether T, Jones A *et al.* (2010). A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? *Ecography* **33**: 315–320.
- Jordan MA, Snell HL (2008). Historical fragmentation of islands and genetic drift in populations of Galapagos lava lizards (*Microlophus albemarlensis* complex). *Mol Ecol* **17**: 1224–1237.
- Kalinowski ST (2006). HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Mol Ecol Notes* **5**: 187–189.
- Kamler JF, Pennock DS (2004). Microhabitat selection of *Peromyscus leucopus* and *P. maniculatus* in mid-successional vegetation. *Trans Kans Acad Sci* **107**: 89–92.
- Krohne DT, Dubbs BA, Baccus R (1984). An analysis of dispersal in an unmanipulated population of *Peromyscus leucopus*. *Am Midl Nat* **112**: 146–156.
- Krohne DT, Hoch GA (1999). Demography of *Peromyscus leucopus* on habitat patches: the role of dispersal. *Can J Zool* **77**: 1247–1353.
- Kyle CJ, Strobeck C (2001). Genetic structure of North American wolverine (*Gulo gulo*) populations. *Mol Ecol* **10**: 337–347.
- Mantel N (1967). Detection of disease clustering and a generalized regression approach. *Cancer Res* **27**: 209–220.
- Mossman CA, Waser PM (2001). Effects of habitat fragmentation on population genetic structure in the white-footed mouse (*Peromyscus leucopus*). *Can J Zool* **79**: 285–295.
- Munshi-South J, Kharchenko K (2010). Rapid, pervasive genetic differentiation of urban white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Mol Ecol* **19**: 4242–4254.
- Munshi-South J (2012). Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Mol Ecol* **21**: 1360–1378.
- Myers P, Lundrigan BL, Hoffman SMG, Haraminac AP, Seto SH (2009). Climate-induced changes in the small mammal communities of the northern Great Lakes Region. *Glob Chang Biol* **15**: 1434–1454.
- Narum SR (2006). Beyond Bonferroni: less conservative analyses for conservation genetics. *Conserv Genet* **7**: 783–787.
- Neigel JE, Ball RM, Avise JC (1991). Estimation of single generation migration distances from geographic-variation in animal mitochondrial-DNA. *Evolution* **45**: 423–432.
- Nupp TE, Swihart RK (2000). Landscape-level correlates of small-mammal assemblages in forest fragments of farmland. *J Mammal* **81**: 512–526.
- Parren SG, Capen DE (1985). Local distribution and coexistence of 2 species of *Peromyscus* in Vermont. *J Mammal* **66**: 36–44.
- Pielou EC (1992). *After the Ice Age: The Return of Life to Glaciated North America*. The University of Chicago Press: Chicago.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna: Austria. Available from <http://www.R-project.org>.
- Rangel T, Diniz-Filho JAF, Bini LM (2006). Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Glob Ecol Biogeogr* **15**: 321–327.
- Rasmussen DI (1964). Blood group polymorphism and inbreeding in natural populations of the deer mouse *Peromyscus maniculatus*. *Evolution* **18**: 219–229.
- Rizkalla CE, Swihart RK (2007). Explaining movement decisions of forest rodents in fragmented landscapes. *Biol Cons* **140**: 339–348.
- Robinson SJ, Samuel MD, Lopez DL, Shelton P (2012). The walk is never random: subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. *Mol Ecol* **21**: 4190–4205.
- Rogic A, Tessier N, Legendre P, Lapointe F-J, Millien V (2013). Genetic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Quebec. *Ecol Evol* **3**: 2075–2088.
- Rosenberg MS (2001). *PASSaGE: Pattern Analysis, Spatial Statistics, and Geographical Exegesis. Version 1*. Department of Biology, Arizona State University: Tempe, AZ.
- Rousset F (2008). GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resources* **8**: 103–106.
- Rowe KC, Heske EJ, Paige KN (2006). Comparative phylogeography of eastern chipmunks and white-footed mice in relation to the individualistic nature of species. *Mol Ecol* **15**: 4003–4020.
- Selkoe KA, Watson JR, White C, Ben Horin T, Iacchi M, Mitarai S *et al.* (2010). Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Mol Ecol* **19**: 3708–3726.
- Short Bull RA, Cushman SA, Mace R, Chilton T, Kendall KC, Landguth EL *et al.* (2011). Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Mol Ecol* **20**: 1092–1107.
- Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 1463–1463.
- Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006). Comparative phylogeography of unglaciated eastern North America. *Mol Ecol* **15**: 4261–4293.
- Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010). Landscape genetics: where are we now? *Mol Ecol* **19**: 3496–3514.
- Van Oosterhout C, Hutchinson WF, Willis DPM, Shipley P (2004). Micro-checker: software for identifying and correcting null alleles. *Mol Ecol Notes* **4**: 725–727.
- Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- Wolff JO (1996). Coexistence of white-footed mice and deer mice may be mediated by fluctuating environmental conditions. *Oecologia* **108**: 529–533.
- Zachos FE, Hacklaender K (2011). Genetics and conservation of large mammals in Europe: a themed issue of *Mammal Review*. *Mamm Rev* **41**: 85–86.
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009). *Mixed Effects Models and Extensions in Ecology with R*. Springer: New York.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)