

Scale dependent inference in landscape genetics

Samuel A. Cushman · Erin L. Landguth

Received: 26 May 2009 / Accepted: 24 February 2010 / Published online: 12 March 2010
© US Government 2010

Abstract Ecological relationships between patterns and processes are highly scale dependent. This paper reports the first formal exploration of how changing scale of research away from the scale of the processes governing gene flow affects the results of landscape genetic analysis. We used an individual-based, spatially explicit simulation model to generate patterns of genetic similarity among organisms across a complex landscape that would result given a stipulated landscape resistance model. We then evaluated how changes to the grain, extent, and thematic resolution of that landscape model affect the nature and strength of observed landscape genetic pattern–process relationships. We evaluated three attributes of scale including thematic resolution, pixel size, and focal window size. We observed large effects of changing thematic resolution of analysis from the stipulated continuously scaled resistance process to a number of categorical reclassifications. Grain and window size have smaller but statistically significant effects on landscape genetic analyses. Importantly,

power in landscape genetics increases as grain of analysis becomes finer. The analysis failed to identify the operative grain governing the process, with the general pattern of stronger apparent relationship with finer grain, even at grains finer than the governing process. The results suggest that correct specification of the thematic resolution of landscape resistance models dominates effects of grain and extent. This emphasizes the importance of evaluating a range of biologically realistic resistance hypotheses in studies to associate landscape patterns to gene flow processes.

Keywords Landscape genetics · Scale · Grain · Extent · Thematic resolution · Gradient · Pattern · Process · Gene flow · Simulation

Introduction

Scale has emerged as a central question of ecology (Levin 1992) because of the fundamental dependencies between scale and pattern and between pattern and process (Wu and Loucks 1995). Landscape ecology explicitly focuses on how relationships between patterns and process vary in heterogeneous environments in relation to spatial and temporal scale. In this paper we investigate three issues identified as key research priorities in landscape ecology (Wu and Hobbs 2002) in the context of

S. A. Cushman (✉)
US Forest Service, Rocky Mountain Research Station,
800 E Beckwith, Missoula, MT 59801, USA
e-mail: scushman@fs.fed.us

E. L. Landguth
Individualized Interdisciplinary Graduate Program,
Mathematics Building, University of Montana, 800 E
Beckwith, Missoula, MT 59801, USA
e-mail: erin.landguth@umontana.com

landscape genetic inference (Balkenhol et al. 2009). First, explicit quantification of pattern–process relationships is necessary for robust landscape ecological inferences. Second, ecological flows through landscape mosaics are foundational processes in many ecological systems. Third, reliable inferences in landscape ecology require explicit attention to the scale dependency of pattern–process relationships. Our goal in this paper is to investigate the scale dependency of inferences regarding the strength and nature of processes governing gene flow in complex landscapes.

Much research in landscape ecology has focused on investigating scale dependency of relationships between patterns and processes. For example, there is a very rich literature in landscape ecology investigating effects of different aspects of spatial scale on landscape metrics. A number of papers have investigated how classification scheme and classification accuracy affect landscape metrics calculated from categorical patch mosaics (e.g. Wickham et al. 1997; Shao et al. 2001). Several well known papers have evaluated the sensitivity and consistency of landscape metrics to variation in landscape pattern using controlled empirical sampling and neutral models (e.g. Riitters et al. 1995; Hess and Bay 1997; Hargis et al. 1998; Neel et al. 2004; Cushman et al. 2008a). Other papers have investigated how changing spatial grain (Turner et al. 1989; Wickham and Ritters 1995; Wu et al. 2000; Lausch and Herzog 2002; Wu et al. 2002; Zhao et al. 2003; Wu 2004) and landscape extent (Saura and Martinez-Millan 2001; Shen et al. 2003) affect landscape metrics. As a result of this rich body of work, there is extensive information about the scale dependency of landscape metrics in relation to grain, extent and classification scheme of categorical landscape maps (e.g. Wu et al. 2000, 2002). This body of work has revealed that landscape patterns are highly sensitive to scale of measurement and analysis (Li and Wu 2004; Wu 2004).

Less work has focused on how apparent pattern–process relationships change with scale of observation (e.g. Tischendorf 2001; Thompson and McGarigal 2002; Cushman 2006; Corry and Nassauer 2005; Cushman et al. 2008b). The drivers of an ecological process may each act at unique scales in space and time (Wiens 1989). It is critically important to correctly match the scale of each driving variable to the response process. Mismatches in scale may lead

to errors of inference and attribution (e.g. Thompson and McGarigal 2002). Measuring a response variable at incorrect scales may distort its pattern in space and time such that it appears to poorly match driving processes. Incorrect scale of measurement of driving variables may result in failure to observe a relationship between pattern and process when one exists, the observation of spurious or distorted relationship, or error in evaluation of the significance or effects size of a pattern–process relationship.

In this paper we investigate how pattern–process inferences in landscape genetic analyses of gene flow are sensitive to three attributes of scale: thematic resolution, grain, and extent. Grain refers to the resolution of the data and defines the finest resolution of patterns that can be described in the data. The extent refers to the temporal or spatial domain of the data and defines the scope of the inferences that can be drawn from the data. Thematic resolution can be considered a third attribute of ecological scale (McGarigal and Cushman 2005; Cushman et al. 2009). We define thematic resolution as how finely ecological differences are resolved in measurement. It is conceptually analogous to radiometric resolution in remote sensing, which refers to how finely a sensor measures differences in the electromagnetic spectrum. In ecology, thematic resolution refers to how finely environmental variability is measured. For example, a map of temperature measured to the nearest 1/10th degree would be a data set with fine scale thematic resolution, in comparison to a map categorizing temperature into two classes, such as above or below freezing.

In stark contrast to the rich understanding of the scale dependency of landscape pattern analysis with landscape metrics, virtually nothing is known about the sensitivity of landscape genetic analyses to variation in the spatial scale of measurement and analysis (Balkenhol et al. 2009; Segelbacher et al. 2010). Landscape genetics is a synthetic field combining spatial population genetics with landscape ecology (Manel et al. 2003; Storfer et al. 2007). Landscape genetics explicitly quantifies the effects of landscape composition, configuration, and matrix quality on spatial patterns in neutral and adaptive genetic variation and underlying microevolutionary processes (Holderegger and Wagner 2008; Balkenhol et al. 2009). Recent landscape genetic approaches largely focus on describing and mapping populations

(e.g., Pritchard et al. 2000; Dupanloup et al. 2001; Francois et al. 2006) and on identifying factors that influence rates and patterns of gene flow within and between populations (e.g., Coulon et al. 2004; Cushman et al. 2006; McRae and Beier 2007; Schwartz et al. 2009).

The predominant analytical approach to associate landscape patterns with gene flow processes uses pair-wise calculation of cost distances, using least cost paths (e.g. Cushman et al. 2006) or multi-path circuit approaches (McRae 2006). These pair-wise cost distances among many individuals across a landscape resistance model are then correlated with pair-wise genetic distances among the same group of individuals with methods such as Mantel (1967) and partial Mantel (Smouse et al. 1986) tests. This framework allows direct association of landscape pattern (expressed in cost distances among individuals) and population genetic process (expressed in terms of pair-wise genetic differentiation among spatially referenced genotypes; e.g. Cushman and Landguth in press).

There has been no formal evaluation of how changing grain, extent or thematic resolution of the underlying landscape resistance model affects ability to correctly and reliably identify the driving causal process (Balkenhol et al. 2009; Segelbacher et al. 2010). Landscape genetic analyses have largely been naïve with respect to the effects of grain, extent, and thematic resolution on observed pattern–process relationships. However, organisms exploit their environments over temporal and spatial scales and time periods that are species-specific and may further depend on life-stage, sex, or season. Balkenhol et al. (2009) argue that scale is a central unexplored question in landscape genetics and that each study should carefully evaluate the appropriate spatial and temporal scale for analysis.

This paper is the first formal evaluation of scale in landscape genetic analyses. Nearly all past landscape genetic analyses of relationships between spatial genetic substructure and landscape features have utilized simplistic binary representations of barrier and matrix, or habitat and non-habitat (Balkenhol et al. 2009), and to date no published papers in landscape genetics have evaluated effects of grain, extent and thematic resolution on detected relationships between spatial genetic structure and landscape features. In this paper we explore how variation in

grain, extent, and thematic resolution affect identification of the correct ecological processes driving gene flow across complex landscapes.

Simulation modeling approaches are particularly valuable to explore scale dependency of landscape genetic pattern–process relationships as they allow explicit control over both scale and the implemented process (Balkenhol et al. 2009; Epperson et al. in press). For an investigation of the effects of scale on landscape genetics it would be desirable to stipulate the operative scales of the pattern process relationship, so that they would be known without error and thus provide a basis for evaluating how departures from these “correct” scales affect observed relationships. Simulation modeling is the approach to accomplish this. In this paper we use an individual-based, spatially explicit simulation program to generate spatial patterns of genetic similarity among organisms across a complex landscape that would result given a stipulated landscape resistance model (Landguth and Cushman 2010). We then evaluate how changes to the grain, extent, and thematic resolution of that landscape resistance model affect the nature and strength of observed landscape genetic pattern–process relationships.

Objectives

The overall objective for this analysis is to evaluate how mismatches between the scale of research and the scale of operative process affect the apparent relationships between gene flow and landscape structure. The specific goal is to evaluate the sensitivity of the Mantel’s test for the relationships between genetic distance and least-cost distance to changes in thematic resolution, focal window width (extent), and pixel size (grain). Within this overall objective we evaluate three hypotheses. (1) We hypothesize that changes in thematic resolution away from the simulated model will result in large reduction in the strength of the apparent relationship between genetic distance and landscape structure. (2) We hypothesize that changes in research grain away from the simulated model likewise will reduce the apparent strength of the relationship. Specifically, we expect that changes in pixel size will have a substantially larger effect on the strength of apparent relationship than holding pixel size constant and

changing the width of a focal window. (3) Following Cushman et al. (2006) and Cushman and Landguth (in press), we expect that the effects of changing research scale will be statistically stronger when analyzed based on partial Mantel tests removing the affect of Euclidean distance than with simple Mantel tests between genetic distance and cost distance alone.

Methods

Simulation

We use the CDPOP (Landguth and Cushman 2010) spatially explicit, landscape genetics simulation program to generate spatially referenced genotypes for a population of 1,248 individual organisms that would result if a stipulated landscape resistance model determined spatial patterns of breeding and dispersal. CDPOP models genetic exchange for a given resistance surface and $n - (x, y)$ located individuals as functions of individual-based movement through mating and dispersal, vital dynamics, and mutation. The program represents landscape structure flexibly as resistance surfaces whose value represents the step-wise cost of crossing each location. Mating and dispersal are modeled as probabilistic functions of cumulative cost across these resistance surfaces. It provides a framework for simulating the emergence of spatial genetic structure in populations resulting from specified landscape resistance processes governing organism movement behavior.

In this analysis we simulated genetic exchange with k -alleles mutation rate of 0.0005 in a population of 1,248 individuals under an isolation by landscape resistance model (Fig. 1). The landscape resistance surface used was identified by Cushman et al. (2006) as the most supported out of 110 alternative models for black bear (*Urusus americanus*) gene flow in northern Idaho, USA. We calculated cost distance matrices as the cumulative cost associated with traversing the least cost path from each individual's location to every other individual's location using ArcGIS COSTDISTANCE function (ESRI 1999–2008). We populated the landscape resistance surface by placing a 1.6 km spaced grid over the landscape. We placed simulated organism on each pixel on this grid with a landscape resistance value of <6 (e.g.

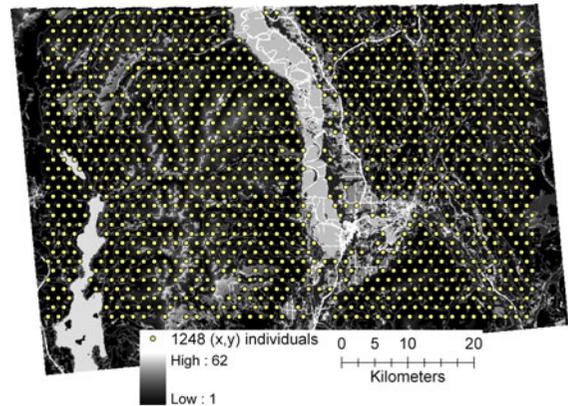


Fig. 1 Resistance map and locations of 1,248 individuals used in the CDPOP simulation of gene flow as a function of landscape resistance

Landguth and Cushman 2010). This was done to place individuals in habitat that was relatively suitable for the species given this landscape resistance hypothesis. We initialized genotypes for these 1,248 individuals by randomly assigning allelic states across 30 loci each with 30 alleles.

We simulated change in individual genotypes across this grid of 1,248 individuals following random assortment in Mendelian inheritance. We ran all simulations for 1,000 non-overlapping generations. We simulated mating with male replacement and without female replacement at each generation. Each mate pair was specified to have a number of offspring following a Poisson process with mean of 4, to guarantee a positive lambda value to ensure that all locations were filled at each time step and to avoid immigrants. We maintained a constant population of 1,248 by discarding the remaining offspring once all the 1,248 grid locations were occupied. This is equivalent to forcing emigration out of the study area once all available home ranges are occupied (Landguth and Cushman 2010). We used an inverse square probability function to select mates and offspring dispersal destinations as functions of cost distance, with a maximum movement threshold set at 10,000 cost units. This is reasonable given that spatial processes on a two-dimensional surface usually are governed by an inverse square relationship, and probability of moving to a particular location on a plane is inversely related to the square of the cost distance from the origin to the destination. We ran 10 Monte Carlo

replicate runs and wrote out individual genotypes for all 1,248 individuals every 10 years.

Scale of simulated resistance model

We used the CDPOP simulation described above to generate spatially referenced genotypes based on an isolation by landscape resistance process (Fig. 2). The resistance map used in the simulation is continuously scaled floating point grid, with resistance values ranging from a minimum of one to a maximum of approximately 62 (Cushman et al. 2006). The pixel size is 90 m. For this simulation, we applied a focal mean moving window with 270 m radius. Thus the stipulated process that was simulated is isolation by landscape resistance on a continuously scaled resistance map consisting of a 270 m focal mean of the resistance values and a 90 m pixel size, which reflects the resistance process governing gene flow in American black bears in Northern Idaho, USA (Cushman et al. 2006).

Alternative scaled resistance models

Our analysis is based on a comparing the strength of pattern–process relationships between gene flow and landscape pattern for the simulated resistance map with the strength of apparent pattern–process relationships when the scale of that resistance map was

altered. We evaluated a combination of alternative landscape resistance hypotheses consisting of a factorial of thematic resolution, focal window width, and pixel (Fig. 2). Thematic resolution was varied across five levels: (1) the native continuously scaled resistance, and equal interval classification into (2) 17 classes, (3) nine classes, (4) five classes, and (5) two classes. We assigned resistance values for each of these classes as the mean of the pixels in the continuously scaled map that were binned to form each output resistance class. Focal window scale was varied across five levels: (1) native 90 m pixel, (2) 270 m radius focal window, (3) 450 m radius focal window, (4) 630 m radius focal window, and (5) 810 m radius focal window. For the continuously scaled map the focal mean was applied; for each of the reclassified categorical resistance maps described above the focal majority was applied. Pixel size was varied across five levels: (1) native 90 m pixel, (2) 270 m radius pixel, (3) 450 m radius pixel, (4) 630 m radius pixel, and (5) 810 m radius pixel.

We produced two-factorials of 25 alternative resistance models by changing the three scale parameters for the resistance model (Fig 1). The two factorials are: (1) combination of the five levels of thematic resolution with the five levels of focal window width, and (2) the five levels of thematic resolution with the five levels of pixel size. The purpose of these factorials was to evaluate the

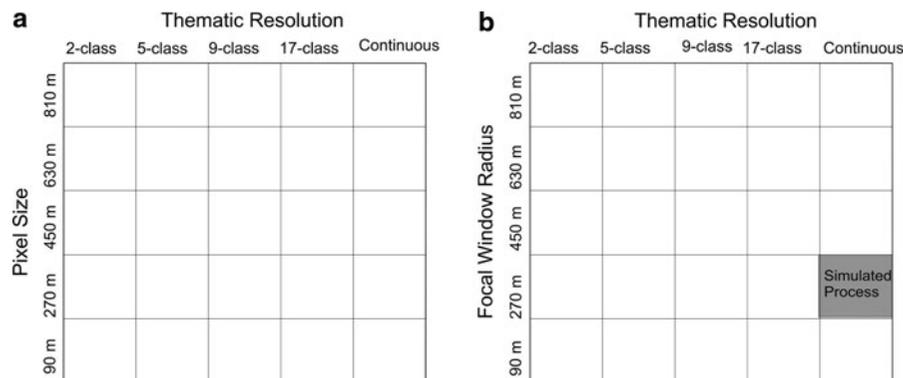


Fig. 2 Schematic outlining the factorial modeling experiment. The modeling experiment involves two combinations of thematic resolution and grain: **a** thematic resolution, with five levels (continuous floating point, 17, 8, 4, and 2 class categorical maps), grain, with five levels (90, 270, 540, 630, 810 m) pixel diameter; **b** thematic resolution, with five levels (continuous floating point, 17, 9, 4, and 2 class categorical maps), grain, with five levels (90, 270, 540, 630, 810 m) focal

mean with 90 m pixel diameter. We simulated genetic structure resulting from one stipulated resistance process reflecting the shaded cell in **(b)**. We conducted 10 replicate runs of the simulation model for this stipulated process and conducted Mantel and partial Mantel tests between the matrix of inter-individual genetic distances produced for the simulated resistance process and each of the alternative hypotheses represented by the other cells in **(a)** and **(b)**

interaction between thematic resolution, grain and extent.

Genetic and cost distance matrix creation

We used the ArcGIS COSTDISTANCE function (ESRI 1999–2008) to calculate the least cost distance for traversing from each individual's location to every other individual's location on each of the 50 alternative resistance hypotheses comprising the two factorials described above. We concatenated these cost distances into $n \times n$ cost-distance matrices recording the least cost distance among all pairs of individual locations for all 50 alternative resistance hypotheses. These cost distance matrices are the independent variable sets for the Mantel analyses. The dependent variable set was the proportion of shared alleles (Dps; Bowcock et al. 1994) genetic distance among all individuals at time 1,000 in each of the 10 CDPOP runs.

Mantel analyses

We used Mantel and partial Mantel tests as implemented in the ECODIST package in *r* (R Development Team 2009) to test the correlation between genetic distances and least cost distances. We computed two sets of Mantel tests, corresponding to simple Mantel tests between genetic and landscape cost distance, and partial Mantel tests between genetic and landscape cost distance, controlling for the relationship between genetic distance and Euclidean distance among individuals. We computed these full and partial Mantel tests between the 10 replicates of the year 500 genetic distance matrix produced from the CDPOP simulation on the stipulated resistance model and the landscape cost-distance matrix for each of the 50 alternative scaled resistance hypotheses.

Analysis of variance

We used two-way analysis of variance to test for significant differences between levels of thematic resolution, focal window size and pixel size in terms of their mean Mantel r value. As Mantel r values are truncated at 0 and 1, we used the arcsin square-root transformation to correct the distribution prior to analysis of variance. In these models we treated thematic resolution, window size, and pixel size as

fixed effects, with CDPOP replicate runs as the random effects error term. We conducted 2 two-way ANOVAs, one for each of the two factorials described above.

Orthogonal contrasts

We used orthogonal contrasts to evaluate the significance of differences between the stipulated landscape resistance model simulated in CDPOP and each alternative level of thematic resolution, focal window, and pixel size. This enabled us to evaluate the significance of differences between the simulated process scale and the particular levels of the alternative scale hypotheses.

Results

We plotted the mean values of the Mantel r statistic for each of the two factorials for both simple and partial Mantel tests (Fig. 3). These surfaces clearly show the dominant effect of thematic resolution, with very large relative drop in Mantel r between the continuously scaled resistance map and any of the categorically scaled alternative models. This suggests that landscape genetic analysis using least-cost paths and Mantel testing is highly sensitive to correct specification of the thematic content and resolution of the resistance model. In contrast, the surfaces also show a weaker effect of grain that is only apparent in the categorical levels of thematic resolution. This distinct difference in apparent effects size and relation to analysis grain between the continuously scaled and categorically scaled hypotheses is notable.

The peak of support across categorical grain models was uniformly at a 90 m pixel and window size, despite the fact that the simulated process was a 90 m cell size within a 270 m window. This suggests that incorrectly fine grain of analysis may not decrease power, but may sometimes increase power in cost-distance based Mantel analysis of landscape genetic pattern–process relationships. For the categorically scaled resistance maps there was only modest (relative to the magnitude of the thematic resolution effect) decrease in Mantel r values as window and pixel size increase to almost an order of magnitude larger than the simulated pattern–process relationship, while Mantel r showed no consistent response across levels of window size or grain in the

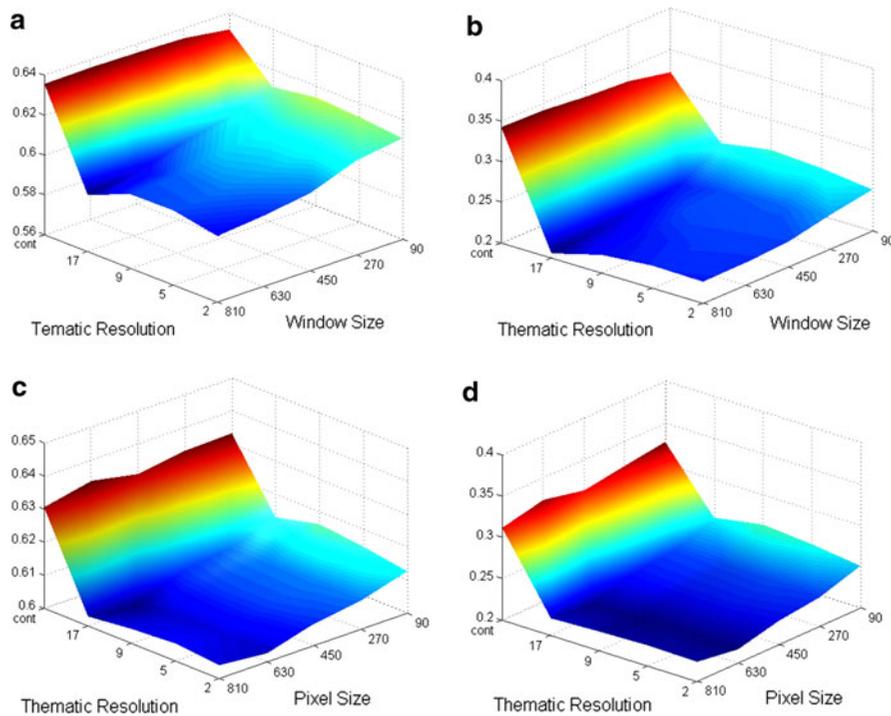


Fig. 3 Surface plots of mean Mantel r values across 10 CDPOP Monte Carlo replicates for the two factorials (thematic resolution \times window size; thematic resolution \times pixel size) for simple Mantel and partial Mantel testing of correlation between genetic distance and landscape cost distance. The Mantel r is between the genetic structure produced with CDPOP on the stipulated true process of isolation by landscape resistance, where resistance is a continuously scaled property, at a 90 m cell size, and within a 270 m radius focal window and the alternatively scaled resistance hypothesis associated

with each combination of thematic resolution and grain. **a** Factorial of thematic resolution \times window size for simple Mantel tests, **b** factorial of thematic resolution \times window size for partial Mantel tests, removing the relationship between genetic distance and Euclidean distance, **c** factorial of thematic resolution \times window size for simple Mantel tests) and **d** factorial of thematic resolution \times pixel size for partial Mantel tests, removing the relationship between genetic distance and Euclidean distance

continuously scaled resistance models. This suggests that landscape genetic methods based on least cost distances may be relatively insensitive to grain of analysis. Interestingly, the surfaces produced by varying window size and pixel size are extremely similar. The high degree of comparability between these suggests that focal extent (window size) and grain (pixel size) may be relatively interchangeable in landscape genetic analysis of pattern–process relationships using cost distances and pair-wise distance-based regression such as Mantel tests.

Analysis of variance

We computed two-way factorial analysis of variance for two combinations of input variables. First, we ran

the analysis of variance for the simple Mantel correlation between genetic structure and alternative resistance hypotheses (Land and Land_G). Second, we computed the same analysis but with the partial Mantel correlation between genetic structure and alternative resistance hypotheses, removing the effect of Euclidean distance (Land_D and Land_DG). In all four factorials there are highly significant differences in Mantel correlation between genetic and cost distance across levels of thematic resolution (Table 1). In contrast, significant grain main effects only appear in the two factorials corresponding to analysis of partial Mantel r values (Land_D and Land_DG). In all four factorials, partialling out Euclidean distance improved ability to detect differences between the correct resistance model and the incorrectly scaled alternative models (Table 1). In the

case of grain main effects, this enabled detection of significant differences, while in the case of thematic resolution main effects, it enhanced the strength of effect as measured by ANOVA *P*-value.

Orthogonal contrasts enable evaluation of differences among levels of the main effects. We conducted two sets of orthogonal contrasts. First, we computed contrasts between the correct model and each incorrect level of each main effect for each of the four factorial ANOVA (Table 2). Second, as we had observed an apparently non-linear interaction between thematic resolution and grain in Fig. 3, we re-ran the orthogonal main-effect contrasts for grain, limiting thematic resolution to the four classified map versions and omitting the continuously scaled resistance map (Table 3).

The first set of contrasts highlighted the dominant effect of changes in thematic resolution on differences between Mantel correlation between the correct and incorrectly scaled models (Table 2). In all four factorials, continuously scaled resistance maps had highly significantly greater Mantel correlation with genetic data than did the alternative hypotheses based on classification of the resistance maps. In the first set

Table 1 Analysis of variance *P*-values for main effects and interaction testing for difference in mean Mantel *r* between levels of grain and thematic resolution, for each of four series of Mantel tests

	Land	Land_D	Land_G	Land_DG
Grain	0.2319	0.0003715	0.74	5.68E-06
Thematic	8.28E-08	2.20E-16	0.00024	2.20E-16
G*T	0.9999	0.1426	1	0.9408

The dependent variable is Mantel correlation between simulated genetic data and alternative resistance hypotheses. Ten replicates of genetic data were produced with CDPOP simulation of gene flow across a continuously scaled map of resistance with 90 m pixel size, and 270 m radius focal mean. The alternative resistance hypotheses are represented as two-way factorials of thematic resolution and grain. In both factorials, thematic resolution has five levels, corresponding to the continuously scaled map, and equal interval reclassifications into four categorizations of resistance levels with 17, 9, 5, and 2 classes, respectively. Grain was varied in two ways. In the first factorial (Land and Land_D) we varied the radius of a focal moving window calculating either focal mean (for the continuously scaled map) or focal majority (for the classified maps). We calculated this over five window sizes corresponding to radii of 90, 270, 450, 630, and 810 m. In the second factorial (Land_G and Land_DG) we resampled the input maps to pixel sizes of 90, 270, 450, 630 and 810 m

Table 2 Orthogonal contrasts to compare Mantel *r* values for the correct resistance model simulated in CDPOP with the levels of the main effects for each of the incorrectly scaled alternative hypotheses

	Land	Land_D	Land_G	Land_DG
Grain B	0.504	0.123	0.553	0.0775
Grain C	0.566	0.759	0.793	0.5156
Grain D	0.296	0.598	0.549	0.1636
Grain E	0.194	0.469	0.563	0.232
Thematic B	3.98E-08	2.00E-16	7.77E-05	2.00E-16
Thematic C	2.12E-06	2.00E-16	1.96E-04	2.00E-16
Thematic D	1.48E-06	2.00E-16	2.01E-04	2.00E-16
Thematic E	2.25E-07	2.00E-16	9.43E-05	2.00E-16

Thematic B—contrast between continuously scaled map (simulated truth) and 17-class map, across levels of grain; Thematic C—contrast between continuously scaled map and 9-class map, across levels of grain; Thematic D—contrast between continuously scaled map and 5-class map, across levels of grain; Thematic E—contrast between continuously scaled map and 2-class map, across levels of grain. For Land and Land_D: Grain B—contrast between 90 m cell, 270 m window (simulated truth) and 90 m cell, 90 m window, across levels of thematic; Grain C—contrast between 90 m cell, 270 m window and 90 m cell, 450 m window, across levels of thematic; Grain D—contrast between 90 m cell, 270 m window and 90 m cell, 630 m window, across levels of thematic; Grain E—contrast between 90 m cell, 270 m window and 90 m cell, 810 m window, across levels of thematic. For Land_G and Land_DG: Grain B—contrast between 90 m cell, 270 m window (simulated truth) and 90 m cell size, with no window; Grain C—contrast between 90 m cell, 270 m window and 450 m cell size; Grain D—contrast between 90 m cell, 270 m window and 630 m cell size; Grain E—contrast between 90 m cell, 810 m window

of contrasts we observed no significant differences between models with the correct grain (90 m cell size and 270 m focal window) and the incorrect focal window size (for Land and Land_D) or pixel size (for Land_G and Land_DG). Partialling our Euclidean distance did not result in significance of these contrasts, although it did substantially reduce *P*-values toward a significant level.

The second set of contrasts revealed a significant grain effect for levels of thematic resolution other than continuously scaled maps (Table 3). Partialling out Euclidean distance effects greatly improved ability to detect differences difference between the correct and incorrectly scaled hypotheses. With distance not partialled out (Land and Land_G) there were no significant contrasts between the correct grain model and the alternative, incorrectly scaled

Table 3 Orthogonal contrasts for grain main effects across only the categorically scaled levels of thematic

	Land	Land_D	Land_G	Land_DG
Grain B	0.3525	0.00018	0.5	0.00113
Grain C	0.46	0.3825	0.799	0.32496
Grain D	0.17887	0.1105	0.448	0.00399
Grain E	0.0999	0.3686	0.537	0.0463

This was done because of an apparently non-linear interaction between thematic resolution and grain in Fig. 3. For Land and Land_D: Grain B—contrast between 90 m cell, 270 m window (simulated truth) and 90 m cell, 90 m window, across categorical levels of thematic; Grain C—contrast between 90 m cell, 270 m window and 90 m cell, 450 m window, across categorical levels of thematic; Grain D—contrast between 90 m cell, 270 m window and 90 m cell, 630 m window, across categorical levels of thematic; Grain E—contrast between 90 m cell, 270 m window and 90 m cell, 810 m window, across categorical levels of thematic. For Land_G and Land_DG: Grain B—contrast between 90 m cell, 270 m window (simulated truth) and 90 m cell size, with no window across categorical levels of thematic; Grain C—contrast between 90 m cell, 270 m window and 450 m cell size across categorical levels of thematic; Grain D—contrast between 90 m cell, 270 m window and 630 m cell size across categorical levels of thematic; Grain E—contrast between 90 m cell, 270 m window and 810 m window across categorical levels of thematic

models. With distance partialled out, however, we observe significant differences between the correct grain model (90 m cell, 270 m window) and Grain B for Land_D and Land_DG. The Grain B model had significantly higher Mantel r values than the simulated truth (90 m cell and 270 m window), indicating that incorrectly fine grain can increase power to detect an effect. No other contrasts were significant for grain in the Land_D model. However, in the Land_DG model, both contrast Grains D and E were significant. This indicates that the Mantel correlation between genetic structure and the simulated truth (90 m cell and 270 m window) is significantly higher than incorrectly scaled grain models with 630 and 810 m cell size.

Discussion

The goal of this analysis was to evaluate how mismatches between the scale of research and the scale of operative process affect the apparent relationships between gene flow and landscape structure.

We specifically sought to evaluate the sensitivity of Mantel tests between genetic distance and least-cost distance to changes thematic resolution, focal window width, and pixel size, with focus on three hypotheses: (1) We hypothesized that changes in thematic resolution away from the simulated scale will result in large reduction in the strength of the apparent relationship between genetic distance and landscape structure. (2) We hypothesized that changes in research grain away from the simulated model likewise will reduce the apparent strength of the relationship. Specifically, we expect that changes in pixel size will have a substantially larger effect on the strength of apparent relationship than holding pixel size constant and changing the width of a focal window. (3) Following Cushman et al. (2006) and Cushman and Landguth (in press), we expect that the effects of changing research scale will be statistically stronger when analyzed based on partial Mantel tests removing the affect of Euclidean distance than with simple Mantel tests between genetic distance and cost distance alone.

Hypothesis 1 Changes in thematic resolution will reduce the strength of the apparent relationship between genetic distance and landscape structure.

Thematic resolution refers to the how precisely differences in the quantitative value of a variable are measured or represented. In our case, it refers to how finely differences in resistance are represented among pixels. We analyzed six levels of thematic resolution. The level with the highest thematic resolution was the continuously scaled map. This map represents resistance as a floating point, continuous variable ranging between 1 and 62 (Cushman et al. 2006). This resolution was stipulated as the true process and simulated in CDPOP. The other levels of thematic resolution each represent a binning of this continuously scaled variable into equal-interval classifications with 32, 17, 9, 5 and 2 classes. Our hypothesis represented an expectation that the strength of Mantel tests would decrease from the stipulated truth of the continuously scaled resistance map to each of these reclassified lower-thematic resolution maps. Specifically, we expected that the strength of the apparent relationship would weaken monotonically as the number of classes was reduced. Our results confirm this expectation. Across all levels of pixel and window size, we observed a dramatic difference

between the Mantel r values for the continuously scaled resistance map and the reclassified resistance maps, with Mantel r values decreasing monotonically with reduction in thematic resolution (Fig. 3). Importantly, the major change was between the continuously scaled map and the 32 class classified resistance map, with relatively little further reduction in apparent effects size among the different classifications (Fig. 3).

The predominant effect of a change from a continuously scaled to a categorically scaled resistance map and the lesser effects of changes among categorical maps of different number of classes has several potentially important implications. First, it suggests that expression of continuous spatial processes as categorical maps may substantially distort the nature and strength of landscape genetic relationships. The analyses presented here demonstrate that analyses of the relationships between landscape structure and gene flow are highly sensitive to the thematic resolution at which landscape structure is represented. Specifically, the analyses show a relatively dramatic decrease in the apparent strength of relationship between the continuous and all categorical representations of the resistance map. Importantly, this difference was nearly as strong between the continuously scaled map and the 17 class map as between the continuously scaled map and the two class map. This suggests that when gene flow is a function of gradients of landscape pattern (McGarigal and Cushman 2005) landscape genetic analyses based on categorical landscape resistance maps may dramatically underestimate the strength of the true pattern–process relationship.

Cushman and Landguth (in press) show that an individual-based, gradient perspective for landscape genetic analysis provides high power to correctly attribute causality in landscape genetic analysis and to correctly reject incorrect alternative explanations. The analyses presented here confirm that Mantel testing of relationships between gene flow and landscape structure are very powerful, but depend fundamentally on correct representation of landscape structure in terms of how it actually influences the spatial processes that drive gene flow. The simulation approach used in this analysis allowed us to stipulate a true process and assess the degree to which the strength of apparent relationships changed as thematic resolution departs from that of the simulated process. Incorrect

expression of a continuous resistance process as a categorical map resulted in large reduction in the apparent strength of the relationship, with the dominant effect appearing even when resistance was classified in a large number of bins, suggesting a general failure of categorical maps to adequately represent continuous processes in landscape genetics.

Hypothesis 2 Changes in research grain will reduce the apparent strength of the relationship.

While there is extensive knowledge of how changing grain of analysis affects landscape structure (Turner et al. 1989; Wickham and Ritters 1995; Wu 2004), there is relatively little known about how grain affects pattern–process relationships. Thompson and McGarigal (2002) showed that changes in research grain in terms of the pixel size and focal landscape area evaluated in landscape-level species habitat relationships can dramatically alter the apparent nature and strength of species–environment relationships. Specifically, they observed several apparent reversals in the nature of the relationship between species occurrence and landscape attributes across a range of pixel sizes and focal landscape areas. However, no previous work has been completed to evaluate the effects of changes in research grain on the apparent nature and strength of relationships between landscape structure and gene flow.

Our hypothesis was that changes in grain from the stipulated process simulated in CDPOP would reduce the apparent strength of the relationship between gene flow and landscape structure. Specifically, we proposed that changes in pixel size would have a larger effect than changes in focal window size holding pixel size constant. Our results confirm the existence of significant effects of grain on landscape genetic analyses, but in somewhat different ways than we expected. First, there was no clear difference between changing pixel size and window size in the strength and nature of observed landscape genetic relationships (Fig. 3). This suggests that resampling to larger pixel size and conducting moving window analysis at constant pixel size produce similar results and may be used interchangeably in some landscape genetic analyses. This is important given the immensely higher computational time of focal window analysis than for resampling to larger pixel size. Second, our results show that affects of thematic resolution are highly dominant compared to those of changes in

grain, such that grain effects were only statistically significant when removing the dominant difference due to thematic resolution (Table 3). Also, grain analyses consistently identified the finest tested grain (90 m pixel and window size) as the most highly supported, even though the stipulated grain size simulated in CDPOP was a 90 m pixel size within a 270 m radius focal window. This suggests that landscape genetic analyses of the relationship between gene flow and landscape pattern may gain power by increasing the spatial grain of analysis even when the governing spatial process acts at a coarser grain. This has potentially important implications. First, it suggests that power in landscape genetics is related to grain of analysis as well as sample size. This implies that researchers should be explicit in describing the grain of their data and consider how it may affect power to detect an effect. Second, it suggests that scaling analysis in landscape genetics may not be effective at determining the operative grain at which the governing process acts. In contrast to our results for thematic resolution, which strongly implicated the correct thematic scale, the grain analysis failed to identify the operative grain, with the general pattern of stronger apparent relationship with finer grain, even at grains finer than the governing process.

Hypothesis 3 Effects of research scale will be statistically stronger when analyzed using partial Mantel tests.

Previous work has shown that partial Mantel tests provide more coherent and stronger prediction of landscape genetic processes than simple Mantel tests (Cushman et al. 2006; Cushman and Landguth in press). Specifically, Cushman and Landguth (in press) showed that simple Mantel tests produce highly equivocal support for alternative explanations due to the very high possibility of spurious correlations leading to the error of affirming the consequent, while partial Mantel tests were nearly perfect in their ability to correctly identify the driving causal model and reject incorrect alternative models. In this analysis we evaluated how partial Mantel tests compare to simple Mantel tests in identifying the correct spatial and thematic scale of landscape genetic processes. Consistent with our hypothesis, partial Mantel tests produced stronger differences between the correct scale and incorrect scales of landscape data, both for thematic resolution and

grain. The partial Mantel test in this case statistically removes the linear relationship between genetic structure and Euclidean distance, and tests the residual relationship between genetic structure and landscape pattern. The results suggest that partial Mantel tests removing distance effects provide improved power to detect scale effects in landscape genetic analysis relative to simple Mantel tests.

Conclusion

Simulation modeling allows the researcher to stipulate a process and thus provide a basis for evaluating how departures from the “correct” process and scale affects observed relationships. In this paper we used a spatially explicit simulation program to generate spatial patterns of genetic similarity among individuals across a complex landscape that would result given a stipulated landscape resistance model, and then evaluated how changes to the grain, extent, and thematic resolution of that landscape model affect the nature and strength of observed landscape genetic pattern–process relationships.

Ecological relationships between patterns and processes are highly scale dependent. Previous work has explored scale dependence in a number of areas of ecological research. However, this is the first formal exploration of how changing scale of research away from the scale of the processes governing gene flow affects the results of landscape genetic analysis. We evaluated three attributes of scale including thematic resolution, pixel size, and focal window size. We observed large effects of thematic resolution from a correct continuously scaled resistance process to a number of categorical reclassifications. We argue this suggests that categorical maps fail to adequately represent continuous processes in landscape genetics. Analyses of effect of grain and window size on landscape genetic results showed smaller but statistically significant effects.

References

- Balkenhol N, Gugerli F, Cushman SA, Waits LP, Coulon A, Arntzen JW, Holderegger R, Wagner HH, Arens P, Campagne P, Dale VH, Niecieza AG, Smulders MJM, Tedesco E, Wang H, Wasserman TN (2009) Identifying

- future research needs in landscape genetics: where to from here? *Landscape Ecol* 24:455–463
- Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic micorsatellites. *Nature* 368:455–457
- Corry RC, Nassauer JI (2005) Limitations of using landscape pattern indices to evaluate the ecological consequences of alternative plans and designs. *Landscape Urban Plan* 72:265–280
- Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Mol Ecol* 13:2841–2850. doi: [10.1111/j.1365-294X.2004.02253.x](https://doi.org/10.1111/j.1365-294X.2004.02253.x)
- Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: a review and prospectus. *Biol Conserv* 128:231–240. doi: [10.1016/j.biocon.2005.09.031](https://doi.org/10.1016/j.biocon.2005.09.031)
- Cushman SA, Landguth EL (in press) Spurious correlations and inference in landscape genetics. *Mol Ecol*
- Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene-flow in complex landscapes: testing multiple models with causal modeling. *Am Nat* 168:486–499
- Cushman SA, McGarigal K, Neel M (2008a) Parsimony in landscape metrics: strength, universality, and consistency. *Ecol Indic* 8:691–703
- Cushman SA, McKelvey K, Flather C, McGarigal K (2008b) Testing the use of forest communities to evaluate biological diversity. *Frontiers Ecol Environ* 6:13–17
- Cushman SA, McGarigal K, Gutzwiller K, Evans J (2009) The gradient paradigm: a conceptual and analytical framework for landscape ecology, chap 5. In: Cushman SA, Huettman F (eds) *Spatial complexity, informatics and wildlife conservation*. Springer, Tokyo
- Dupanloup I, Schneider S, Excoffier L (2001) A simulated annealing approach to define genetic structure of populations. *Mol Ecol* 58:2021–2036
- Epperson BK, McRae B, Scribner K, Cushman SA, Rosenberg MS, Fortin M-J, James PMA, Murphy M, Manel S, Legendre L, Dale MRT (in press) Utility of computer simulations in landscape genetics. *Mol Ecol*
- ESRI (1999–2008) ArcGIS: Release 9.3. Environmental System Research Institute, Redlands, CA
- Francois O, Ancelet S, Guillot G (2006) Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* 174:805–816
- Hargis CD, Bissonette JA, David JL (1998) The behavior of landscape metrics commonly used in the study of habitat fragmentation. *Landscape Ecol* 13:167–186
- Hess G, Bay JM (1997) Generating confidence intervals for composition-based landscape indexes. *Landscape Ecol* 12:309–320
- Holderegger R, Wagner HH (2008) Landscape genetics. *Bio-science* 58:199–207
- Landguth EL, Cushman SA (2010) CDPOP: an individual-based, cost-distance spatial population genetics model. *Mol Ecol Resour* 10:156–161
- Lausch A, Herzog F (2002) Applicability of landscape metrics for the monitoring of landscape change: issues of scale, resolution and interpretability. *Ecol Indic* 2:3–15
- Levin SA (1992) The problem of pattern and scale in ecology. *Ecology* 73:1943–1967
- Li H, Wu J (2004) Use and misuse of landscape indices. *Landscape Ecol* 19:389–399
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* 18:189–197
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220
- McGarigal K, Cushman SA (2005) The gradient concept of landscape structure. In: Wiens J, Moss M (eds) *Issues and perspectives in landscape ecology*. Cambridge University Press, Cambridge, pp 112–119
- McRae BH (2006) Isolation by resistance. *Evol Int J Org Evol* 60:1551–1561
- McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci USA* 104:19885–19890. doi: [10.1073/pnas.0706568104](https://doi.org/10.1073/pnas.0706568104)
- Neel MC, Cushman SA, McGarigal K (2004) Behavior and stability of landscape metrics across controlled gradients of landscape structure. *Landscape Ecol* 19:435–455
- Pritchard JK, Stephens M, Peter D (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- R Development Core Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>
- Riitters KH, O'Neill RV, Hunsaker CT, Wickham JD, Yankee DH, Timmins SP, Jones KB, Jackson BL (1995) A factor analysis of landscape pattern and structure metrics. *Landscape Ecol* 10:23–39
- Saura S, Martinez-Millan J (2001) Sensitivity of landscape pattern metrics to map spatial extent. *Photogramm Eng Remote Sens* 67:1027–1036
- Schwartz MK, Copeland JP, Anderson NJ, Squires JR, Inman RM, McKelvey KS, Pilgrim KL, Waits LP, Cushman SA (2009) Wolverine gene flow across a narrow climatic niche. *Ecology* 90:3222–3232
- Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy OJ, Holderegger R, Manel S (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conserv Genet*. doi: [10.1007/s10592-009-0044-5](https://doi.org/10.1007/s10592-009-0044-5)
- Shao G, Liu D, Zhao G (2001) Relationships of image classification accuracy and variation of landscape statistics. *Can J Remote Sens* 27:33–43
- Shen W, Wu J, Ren H (2003) Effects of changing spatial extent on landscape pattern analysis. *Acta Ecol Sin* 23:2219–2231
- Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* 35:627–632
- Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the “landscape” in landscape genetics. *Heredity* 98:128–142
- Thompson CM, McGarigal K (2002) The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). *Landscape Ecol* 17: 569–586

- Tischendorf L (2001) Can landscape indices predict ecological processes consistently? *Landscape Ecol* 16:235–254
- Turner MG, O'Neill RV, Gardner RH, Milne BT (1989) Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecol* 3:153–162
- Wickham JD, Ritters KH (1995) Sensitivity of landscape metrics to pixel size. *Int J Remote Sens* 16:3585–3594
- Wickham JD, O'Neill RV, Riitters KH, Wade T, Jones KB (1997) Sensitivity of selected landscape pattern metrics to land-cover misclassification and differences in land-cover composition. *Photogram Eng Remote Sens* 63:397–402
- Wiens JA (1989) Spatial scaling in ecology. *Funct Ecol* 3:385–397
- Wu J (2004) Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecol* 19:125–138
- Wu J, Hobbs R (2002) Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. *Landscape Ecol* 17:355–365
- Wu J, Loucks OL (1995) From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology. *Q Rev Biol* 70:439–466
- Wu J, Jelinski DE, Luck M, Tueller PT (2000) Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Geogr Inf Syst* 6:6–19
- Wu J, Shen W, Sun W, Tueller PT (2002) Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecol* 17:761–782
- Zhao W, Fu B, Chen L (2003) The effects of grain change on landscape indices. *Quat Sci* 23:326–333