ISOLATION BY RESISTANCE

Brad H. McRae

National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, California 93101-5504
E-mail: mcrae@nceas.ucsb.edu

Abstract.—Despite growing interest in the effects of landscape heterogeneity on genetic structuring, few tools are available to incorporate data on landscape composition into population genetic studies. Analyses of isolation by distance have typically either assumed spatial homogeneity for convenience or applied theoretically unjustified distance metrics to compensate for heterogeneity. Here I propose the isolation-by-resistance (IBR) model as an alternative for predicting equilibrium genetic structuring in complex landscapes. The model predicts a positive relationship between genetic differentiation and the resistance distance, a distance metric that exploits precise relationships between random walk times and effective resistances in electronic networks. As a predictor of genetic differentiation, the resistance distance is both more theoretically justified and more robust to spatial heterogeneity than Euclidean or least cost path-based distance measures. Moreover, the metric can be applied with a wide range of data inputs, including coarse-scale range maps, simple maps of habitat and nonhabitat within a species' range, or complex spatial datasets with habitats and barriers of differing qualities. The IBR model thus provides a flexible and efficient tool to account for habitat heterogeneity in studies of isolation by distance, improve understanding of how landscape characteristics affect genetic structuring, and predict genetic and evolutionary consequences of landscape change.

Key words.—Gene flow, graph theory, isolation by distance, isolation by resistance, landscape connectivity, landscape genetics, resistance distance.

The emerging study of how landscape features affect microevolutionary processes (landscape genetics; Manel et al. 2003) will require tools that explicitly incorporate landscape heterogeneity into analyses of gene flow and genetic differentiation. Landscape characteristics may modify gene flow between pairs of subpopulations directly by affecting dispersal rates among them or indirectly by affecting the spatial arrangement of and dispersal rates among intervening subpopulations. Yet, few models are capable of integrating landscape data into predictions of population structure.

For example, models of isolation by distance (Wright 1943) are among the most widely applied tools in studies of genetic differentiation in natural populations. These models have provided powerful means to explain population structure (e.g., Rousset 1997, 2000; Sumner et al. 2001; Rueness et al. 2003), investigate departures from migration-drift equilibrium (Slatkin 1993; Hutchison and Templeton 1999), and address ecological questions such as whether dispersal synchronizes the dynamics of populations separated by long distances (Schwartz et al. 2002). Yet, such analyses assume homogeneous, unbounded populations, ignoring effects of range boundaries and of variation in demographic parameters within species' ranges (Maruyama 1970; Slatkin and Maruyama 1975). As Slatkin (1985) noted, most real populations are neither homogeneous nor unbounded, and when population densities or migration rates vary across space, the predictive abilities of these models may suffer.

Perhaps because of their broad utility, isolation-by-distance models have been a focus of growing efforts to incorporate landscape data into studies of natural populations. In particular, empiricists now commonly substitute least cost path (LCP) distance measures for Euclidean distances in isolation-by-distance analyses (e.g., Arnaud 2003; Coulon et al. 2004; Schweiger et al. 2004; Vignieri 2005). These measures modify geographic distances between sample pairs to account for characteristics of the intervening landscape that facilitate

or impede movement along a single, optimal pathway. Yet, despite their intuitive appeal, LCP measures cannot consider the effects of multiple dispersal pathways connecting samples and remain without theoretical foundation for most population genetic applications. To date, no theoretically justified method to incorporate landscape heterogeneity into isolation-by-distance analyses has been proposed.

To address this deficiency, I describe a new method designed to predict the effects of landscape structure on equilibrium genetic structuring in natural populations. The isolation-by-resistance (IBR) model predicts a positive relationship between genetic differentiation and the resistance distance, a graph theoretic distance metric based in circuit theory. The resistance distance provides a more appropriate predictor of equilibrium genetic differentiation than Euclidean distance because it accounts for heterogeneity in species' distributions and migration rates. It improves over LCP methods because it incorporates all possible pathways and is better supported by existing analytic theory. Finally, the resistance distance accommodates a variety of data inputs, from simple habitat or range maps to complex arrays of populations with known sizes and migration rates. Thus, the IBR model fills a gap between models that are nonspatial and those that require heavy parameterization and/or computation, and it provides a quantitative and conceptual link between the disciplines of population genetics and landscape ecology.

An Example Problem

Figure 1 shows a landscape in which suitable habitat for a species of limited dispersal ability is bounded internally and externally by inhospitable nonhabitat. The map could be based on a published range map or on a simple habitat model (e.g., forest for a forest-obligate species). The species' habitat is divided into two main blocks, *A* and *B*, connected by two corridors. Only a small subset of local populations (or individuals) has been sampled, as is common in studies of
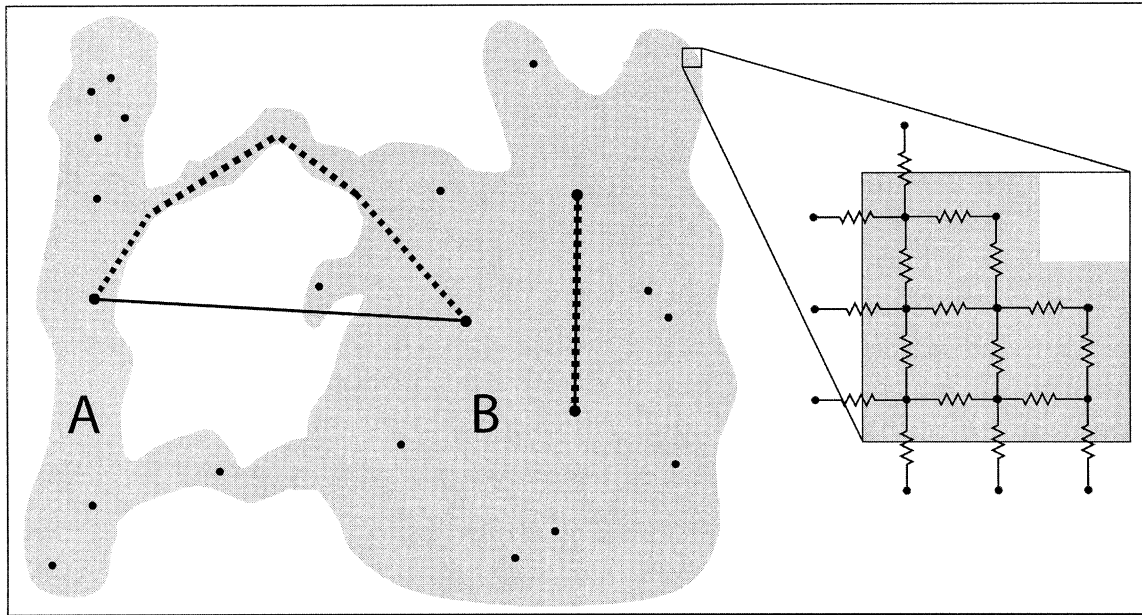
FIG. 1. Range or habitat map for a hypothetical species, with suitable habitat shown in gray. The species is assumed to be continuously distributed throughout the habitat and to have limited dispersal ability. Dots indicate locations of 21 samples from individuals or local populations. Two common distance measures between sample pairs are shown: Euclidean distances (solid lines), and cost-weighted distances from least cost path analyses (dashed lines). Inset shows discretized habitat represented as a network of nodes connected to their neighbors by resistors. Diagonal connections or connections between nonadjacent nodes could also be incorporated. Resistance distance calculations using this network would integrate all possible pathways connecting sample pairs.

natural populations (Beerli 2004; Slatkin 2005). The map is the only information available on the species' distribution, which is also very common. Questions of interest to an evolutionary biologist may include whether populations are at drift-migration equilibrium, whether events such as range expansions have occurred in the recent past, whether historical or current habitat configurations better explain patterns of genetic structuring, and how future habitat changes (e.g., elimination of one of the corridors) may affect gene flow and evolutionary potential of local populations (Templeton et al. 2001).

Even in such a simple landscape, attempts to predict equilibrium patterns of genetic structuring would be compromised by spatial heterogeneity. In particular, the irregular shape of the species' habitat would weaken the predictive power of standard isolation-by-distance models. Although Euclidean distance may perform reasonably well as a predictor of genetic differentiation within each habitat block, different models would need to be applied in the two cases: a one-dimensional model (genetic differentiation increases with raw Euclidean distance) would be appropriate for block A, whereas block B would require a two-dimensional model (differentiation increases with log-transformed Euclidean distance; Rousset 1997). Log-transformed distances may also be more appropriate for the four uppermost samples in block A, where habitat is locally two-dimensional. Between habitat blocks, Euclidean distances would certainly be poor predictors of differentiation because migration is restricted to narrow corridors, resulting in greater differentiation than expected based on distance alone.

Although within-habitat distances calculated using LCP models could be substituted for Euclidean distances, they would still not adequately capture the effect of spatial heterogeneity on genetic structuring in this example. Originating in graph theory and applied in a raster GIS environment, LCP models first assign costs (also referred to as resistance or friction values) to grid cells that reflect relative movement difficulties through different habitat types. The optimal route connecting each sample pair is then plotted, and a cost-weighted distance (the cumulative cost of all cells the route runs through) is used as an index of effective distance between them (Adriaensen et al. 2003). Thus, although they may offer some predictive advantage when gene flow is genuinely constrained to single pathways (as in stream networks; e.g., Arter 1990), LCP models cannot incorporate the effects of wider habitat swaths or of independent, parallel pathways connecting samples. In the example above, cost-weighted distances would differ little from Euclidean distances within either habitat block and, in fact, would need to be log-transformed to predict genetic distances in block B. LCP methods would also not predict increased gene flow between the two blocks when two corridors are present rather than one.

Although better able to accommodate spatial heterogeneity, migration matrix models (Bodmer and Cavalli-Sforza 1968) and coalescent models (Kingman 1982) would also be of limited use in this case. These models, combined with maximum likelihood techniques, have been useful in estimating demographic parameters given genetic data from relatively small and well-sampled networks (e.g., Tufto et al. 1998; Beerli and Felsenstein 2001). But to predict genetic structuring given landscape characteristics, they would first require data on population densities and migration rates at a level of detail rarely attainable in studies of natural populations. Moreover, their computational costs increase rapidly

with the number of demes modeled, complicating their use in realistic landscapes.

What is needed, then, is a predictive model that is not only more robust to spatial heterogeneity than existing distance metrics, but also less demanding of data and computational resources than the more sophisticated analytic models. Ideally, the model should allow prediction of expected patterns of gene flow and genetic differentiation considering the extent and quality of habitat alone. It should, for example, predict increased genetic differentiation with distance within one-dimensional habitat blocks and with the logarithm of distance where habitat is two-dimensional. It should also predict greater gene flow when multiple corridors connect samples than with one and in wider swaths of habitat than in narrower swaths (Slatkin 1993).

Distance metrics based in circuit theory can fill this role, as the properties we are looking for are shared by conductive materials. Recall that two resistors connected in parallel have lower effective resistance (higher effective conductance; conductance is simply the reciprocal of resistance) to current flow than would a single resistor. Similarly, a continuous conductive material will have higher effective conductance between two points if there are multiple or wider conductive pathways between them. That is, if the map in Figure 1 represented a conductive surface, current would flow more freely in wider blocks, less freely in narrow corridors, and more freely between blocks A and B with two corridors connecting them than with one. Accordingly, effective conductance between pairs of points on the surface would behave very much like we expect gene flow to behave in the landscape. Effective resistance would behave very much like equilibrium genetic differentiation; in fact, effective resistance increases with distance in one-dimensional conductors and with its logarithm in two-dimensional conductors. These properties are also shared by networks of discrete resistors, which can be used to model electrical connectivity across continuous conductive surfaces, as in the inset of Figure 1.

THE MODEL

The conceptual basis of the IBR model lies in analogous properties of gene flow in deme networks and conductance in linear electronic circuits (Fig. 2). All else being equal, equilibrium levels of gene flow between two demes connected by migration (e.g., demes a and b) will increase if additional parallel movements of genes are allowed, either through increased direct movements of gametes or through indirect gene flow via intervening demes (deme c). Similarly, effective conductance between two nodes in a network of resistors (e.g., nodes A and B) will increase if additional connections are made, either directly through extra parallel resistors or indirectly via intervening nodes (node C).

Demes connected by migration can be represented in a graph theoretic framework as nodes connected by edges (e.g., Matsen and Wakeley 2006). In the context of this paper, edge weights are proportional to numbers of migrants exchanged between deme pairs. Numbers of migrants (and hence edge weights) can vary across the graphs, but edges are assumed to be undirected (i.e., dispersal is balanced, with numbers of
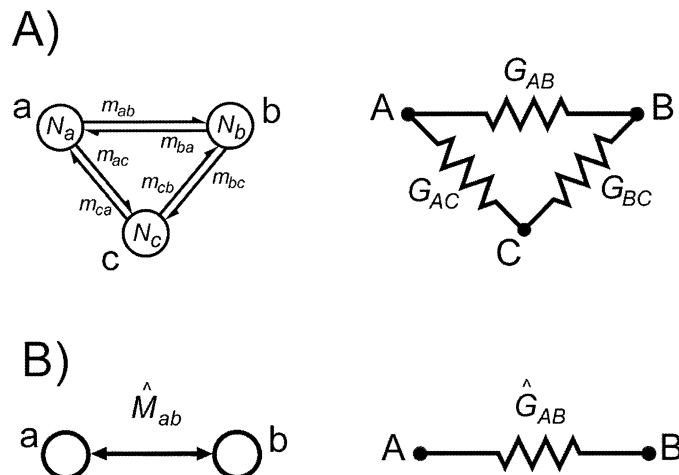

FIG. 2. (A) Three demes (open circles) connected by migration and analogous nodes (closed circles) connected by resistors. Theory discussed in this paper is limited to the balanced migration case, $N_a m_{ab} = N_b m_{ba}$. (B) Equivalent two-deme system connected by effective number of migrants $\hat{M}_{ab}$, and equivalent two-node system connected by effective conductance $\hat{G}_{ab}$. $\hat{G}_{ab}$ will approximate $\hat{M}_{ab}$ when G-values are set to corresponding Nm values.

migrants exchanged between demes being equal in both directions; McPeek and Holt 1992).

On such graphs, the resistance distance between any pair of nodes is defined as the effective resistance between the nodes when each graph edge is replaced by a resistor, the conductance of which is set equal to the corresponding edge weight (Klein and Randic 1993). A key property of the resistance distance is that adding additional pathways between nodes always reduces the resistance distance separating them; the metric also takes advantage of well-known connections between effective resistances and random walks on graphs (Doyle and Snell 1984). Although it has been used in other fields (e.g., in the analysis of social, biochemical, and neural networks; Balaban and Klein 2002; Vast et al. 2005; Yen et al. 2005), the resistance distance has yet to be applied to network analyses in ecology or evolution.

For deme networks or continuous populations that can be represented using connected, undirected graphs, the IBR model predicts increased equilibrium genetic differentiation between samples with increased resistance distances between nodes. Applying the model first involves replacing networks of demes or habitat cells connected by migration with networks of nodes connected by resistors, as in Figures 1 and 2. Conductances assigned to resistors reflect numbers of migrants exchanged between corresponding demes or cells (or relative numbers of migrants, depending on the level of detail of hypothesized migration patterns). Effective conductances between node pairs are then used as proxies for effective migration between corresponding demes or grid cells, whereas resistance distances are used as indices of genetic differentiation.

Calculating the Resistance Distance

Nodal analysis applies Kirchhoff's (1845) laws in matrix form to efficiently solve for effective resistances across pairs

of nodes within circuits (Dorf and Svoboda 2003). Consider an array of n nodes connected in any configuration by resistors, with the conductance of the resistor connecting each pair of nodes i and j denoted by G_{ij}. Define **G** to be an $n \times n$ conductance (or weighted Laplacian) matrix with each non-diagonal element at row i and column j set equal to $-G_{ij}$ (or zero if there is no direct connection between the nodes), and each diagonal element (i, i) set equal to the sum of the conductances of all resistors connected to node i. For the circuit shown in Figure 2,

$$\mathbf{G} = \begin{pmatrix} G_{AB} + G_{AC} & -G_{AB} & -G_{AC} \\ -G_{AB} & G_{AB} + G_{BC} & -G_{BC} \\ -G_{AC} & -G_{BC} & G_{AC} + G_{BC} \end{pmatrix}. \quad (1)$$

Note that this matrix is easily generated from a standard migration matrix. To calculate the resistance distance between any pair of nodes x and y ($x < y$ for convenience), a reduced conductance matrix $\mathbf{G}_y$ is defined, which is equal to **G** except that row y and column y have been removed. Next, define an $n - 1$ column vector **i**, which is all zeros except for the xth element, which is set to one. $\mathbf{G}_y$ and **i** are then substituted into the equation $\mathbf{G}_y\mathbf{v} = \mathbf{i}$, which is solved for the vector **v**. The xth element of **v** will equal the resistance distance between nodes x and y, $\hat{R}_{xy}$, while its reciprocal will equal the effective conductance between the two nodes, $\hat{G}_{xy}$. Because **v** can be obtained using matrix techniques that do not require inverting the reduced conductance matrix, very large networks can be solved using this method. Other algorithms (e.g., Babic et al. 2002; Vast et al. 2005) allow computation of all resistance distances with a single matrix inversion and may be more efficient for moderate-sized networks with large numbers of pairwise resistances to calculate.

Equivalence of Analytic Theories in Isotropic Deme Arrays

The theoretical basis of the IBR model lies in combining previously established relationships between resistance distances and random walk times with those between pairwise F_{ST}-values and coalescence times. In this section, I demonstrate a basic equivalence between theories describing effective resistance in circuits and genetic differentiation in isotropic deme networks (i.e., networks in which migration patterns are identical among all demes; Strobeck 1987). The nonisotropic case will then be addressed in the following section.

In the equations below, I depart somewhat from previous practice in defining the effective number of migrants measured between any two demes x and y, $\hat{M}_{xy}$, to be the number of migrants that would produce the same level of genetic differentiation in a two-deme (two-island) system. This differs slightly from Slatkin's (1993) $\hat{M}$, which assumes an infinite island model. For the two-deme case, Crow and Aoki's (1984) equation for equilibrium pairwise F_{ST} in terms of deme sizes and migration rates can be rewritten in linearized form (Rousset 1997) as:

$$\frac{F_{ST}}{1 - F_{ST}} \approx \frac{1}{16Nm}, \quad (2)$$

where N is the size of each deme and m is the proportion of genes in each deme that are derived from immigrants from

the other deme. From this equation, the number of migrants between two demes that would be estimated from their pairwise F_{ST} value is

$$\hat{M}_{xy} \approx \frac{1}{16}\left(\frac{1}{F_{ST}} - 1\right), \quad (3)$$

which is one-quarter the value inferred assuming an infinite island model (Slatkin 1993).

Slatkin (1991, 1993) showed that for the weak mutation case, pairwise F_{ST}-values between two demes can be calculated from coalescence times of pairs of genes sampled within and between the demes. Slatkin's equation relating these quantities can be rewritten in terms of linearized F_{ST} as:

$$\frac{F_{ST}}{1 - F_{ST}} \approx \frac{\bar{t}_1 - \bar{t}_0}{2\bar{t}_0}, \quad (4)$$

where $\bar{t}_0$ is the average expected coalescence time of two different genes sampled from the same deme, and $\bar{t}_1$ is the expected coalescence time of two genes sampled from different demes. The second quantity is the sum of two parts: $\bar{t}'_1$, the expected time before two alleles are first found in the same deme and the time until they coalesce thereafter (Slatkin 1991).

For demes connected by isotropic migration, the time until two alleles are first in the same deme will be the same as in the case of one allele moving at twice the rate with the other's position fixed, provided m is small (Matsen and Wakeley 2006). Thus, the time until the two alleles are in the same deme is half the time for an allele to move from one deme to the other, or one-quarter the commute time, that is, the expected time for the allele to make a round trip from one deme to the other and back again. Chandra et al. (1997) showed that commute times on weighted graphs are precisely related to effective resistances in analogous circuits when conductances are proportional to movement probabilities. When conductances between adjacent pairs of nodes are set to the number of migrants, $N_x m_{xy}$ between corresponding demes (so that the sum of all conductances for node x, including self loops to reflect the probability that an allele does not migrate, is N_x), Chandra et al.'s (1997, theorem 2.2) expression for commute time simplifies to:

$$\bar{t}_c = \hat{R}_{xy} \sum_{i=1}^{d} N_i, \quad (5)$$

where $\bar{t}_c$ is the commute time between demes x and y and $\hat{R}_{xy}$ is the effective resistance (i.e., the resistance distance) between corresponding nodes. If all deme sizes are equal and d is the number of demes in the system, then $\bar{t}_c = \hat{R}_{xy}Nd$, and $\bar{t}'_1 = \hat{R}_{xy}Nd/4$, where N is the deme size. The expected coalescence times are then $\bar{t}_0 = 2Nd$ (Slatkin 1991), and $\bar{t}_1 = \hat{R}_{xy}Nd/4 + 2Nd$. Substituting these values into equation (4), the relationship between equilibrium pairwise genetic distance and the resistance distance will be:

$$F_{ST}/(1 - F_{ST}) \approx \hat{R}_{xy}/16. \quad (6)$$

Combined with equation (3), this result also implies:

$$\hat{M}_{xy} \approx \hat{G}_{xy}, \quad (7)$$

where $\hat{G}_{xy}$ is the effective conductance between nodes x and y.
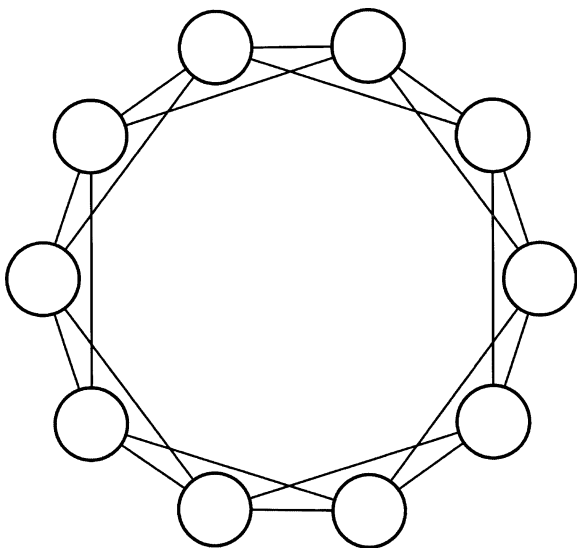
FIG. 3. An isotropic migration pattern with long-distance dispersal (adapted from Matsen and Wakeley 2006). The figure could represent, for example, breeding habitat for fish along the shoreline of a lake or for amphibians in ponds separated by dry land. Resistance distances are precisely related to F_{ST}-values in isotropic cases.

Equations (6) and (7) establish practical links between circuit and gene flow theories and form the basis for using the resistance distance as an index of genetic differentiation. Note that the same correspondence can be obtained without reference to commute times. For example, $\hat{G}$ calculated using Jeng's (2000) equation for effective resistance on a torus is equivalent to the equation for $\hat{M}$ on a torus derived by Slatkin (1993), adjusted for the two-island case. Equivalence for demes arrayed on a circle is easily demonstrated using elementary circuit theory. The equivalence between the two theories holds for any isotropic migration scheme, including symmetric island models and symmetric arrays with long-distance migration (Fig. 3).

A convenient property of the model is that multiplying conductances of all resistors in a circuit by a scalar will result in all effective conductances being multiplied by the same scalar and all resistance distances being multiplied by its inverse. So, if node-to-node conductances reflect relative rather than absolute numbers of migrants, $\hat{G}_{xy}$ and $\hat{R}_{xy}$ will still be proportional to $\hat{M}_{xy}$ and linearized F_{ST}, respectively.

Model Performance in Nonisotropic Cases

The above results demonstrate that the IBR model is consistent with existing analytic theory for isotropic migration patterns; however, the main purpose of the model is to predict genetic structuring when migration patterns differ from this ideal. In such cases, resistance distances still precisely capture commute times as long as balanced migration is assumed, but commute times only approximate the expected time until two alleles are first found in the same deme. So, although equation (5) still holds, we can no longer rely on equations (6) and (7) to justify the model. Instead, the predictive power of resistance distances can be evaluated by comparing them to pairwise F_{ST}-values expected under a range of nonisotropic migration scenarios.

The standard: F_{ST} calculated using Markov chains

In small networks with known deme sizes and migration rates, exact pairwise F_{ST}-values can be obtained by using standard Markov chain approaches to calculate within- and between-deme coalescence times. A full description of the Markov chain algorithm used here is available in Appendix 1 (available online only at http://dx.doi.org/10.1554/05-321.1.s1). An example application of the algorithm to solve the three-deme system in Figure 2 is given in Appendix 2 (available online only at http://dx.doi.org/10.1554/05-321.1.s2), along with a formal derivation of the nodal analysis equations for the analogous circuit.

Aside from its data requirements, the main limitation of the Markov chain method is computational: for an array of d demes, the algorithm requires specifying $d(d + 1)/2$ states. As a result, equation (A6) in Appendix 1 can quickly become difficult to solve (in my experience, this becomes limiting at about 100 demes on a standard desktop computer). By contrast, solving for resistance distances requires only operations involving matrices of dimension $d \times d$. The smaller dimensions and greater sparseness of conductance matrices allow effective resistances to be calculated on much larger arrays ($\sim$100,000 demes on the same computer). Still, IBR predictions can be checked against those of the Markov chain algorithm in deme arrays small enough to be accommodated by the algorithm.

Correlations between distance measures and linearized F_{ST}

To test the IBR model in nonideal cases, I regressed exact $F_{ST}/(1 - F_{ST})$ values calculated using the Markov chain algorithm against resistance distances calculated using nodal analysis for each of 10 example deme networks. Note that correlations will be the same whether we use the resistance distance as an index of differentiation or estimate $F_{ST}/(1 - F_{ST})$ using equation (6). I compare IBR model predictions with those obtained using LCP and isolation-by-distance models, and use R^2-values to assess model fit because this is how model predictions will most likely be evaluated with empirical data. Higher R^2-values will also mean greater power to detect significant relationships using Mantel tests.

Table 1 summarizes correlations between Markov chain solutions and resistance distances, Euclidean distances, and cost-weighted (LCP) distances for the 10 networks. The regressions show the resistance distance to be a reliable predictor of equilibrium genetic differentiation across a wide range of conditions, consistently outperforming the other distance measures. The results also illustrate several important properties of the model.

First, performance of the resistance distance improves with increasing population sizes and/or migration rates in all of the networks. This is due to the assumption that, within a network, no demes are more isolated from the network as a whole than others. When numbers of migrants are few (typically fewer than one migrant per generation), demes in corners of arrays or with lower migration rates than other demes will have shorter than average within-deme coalescence times (i.e., they will be more inbred than other demes). Pairs of demes that are separated by large distances and more inbred than other demes will exhibit long between-deme coalescence

TABLE 1. Isolation-by-resistance (IBR) model performance in finite networks. Table entries are R^2-values for regressions of $F_{ST}/(1 - F_{ST})$ as calculated by the Markov chain algorithm in 10 deme networks against Euclidean distance (IBD), cost-weighted distance (LCP), and resistance distance (IBR). $R \times C$ deme arrays are finite networks of R rows by C columns of demes (total of RC demes) connected by migration rates that are equal across the networks. The remaining arrays correspond to the deme maps shown in Figure 4 (case I). Following Rousset (1997), Euclidean distances are $\log_{10}$ transformed except in the 1×8, 1×64, and 4×16 arrays. Cost-weighted distances are calculated along the least cost path connecting deme pairs, with per step costs equal to the reciprocal of migration rates.

Deme array	$N = 2$			$N = 100$			$N = 10,000$		
	IBD	LCP	IBR	IBD	LCP	IBR	IBD	LCP	IBR
4×4, $m = 0.1$	0.86	0.79	0.99	0.90	0.82	1.0	0.90	0.82	1.0
4×4, $m = 0.001$	0.79	0.73	0.96	0.84	0.78	0.99	0.90	0.82	1.0
8×8, $m = 0.001$	0.67	0.69	0.93	0.75	0.76	0.98	0.83	0.81	1.0
1×16, $m = 0.1$	0.92	0.92	0.92	1.0	1.0	1.0	1.0	1.0	1.0
1×64, $m = 0.1$	0.83	0.83	0.83	1.0	1.0	1.0	1.0	1.0	1.0
4×16, $m = 0.1$	0.95	0.92	0.98	0.99	0.96	1.0	0.99	0.96	1.0
Figure 4A	0.27	0.89	0.93	0.33	0.93	1.0	0.34	0.92	1.0
Figure 4B	0.33	0.71	0.89	0.37	0.79	0.97	0.42	0.85	1.0
Figure 4C	0.14	0.85	0.90	0.11	0.93	0.99	0.10	0.94	1.0
Figure 4D	0.40	0.54	0.81	0.45	0.57	0.99	0.40	0.53	1.0

times relative to within-deme coalescence times, inflating F_{ST} values at large distances relative to IBR predictions.

Second, while the resistance distance scales with Euclidean distance and with its logarithm in purely one- and two-dimensional arrays, respectively, it improves over both in intermediate cases. For example, R^2-values for resistance distances are the same as those for Euclidean distances in the 1×16 and 1×64 cases, but are higher in the 4×16 array.
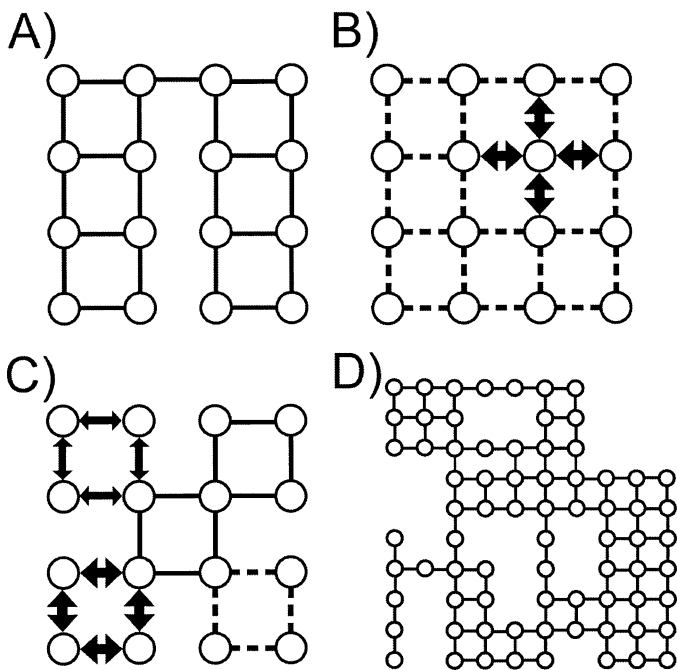

FIG. 4. Four deme maps used in comparisons of isolation-by-resistance predictions with Markov chain predictions. For case I, deme sizes are equal across the networks and are given in Table 1; large arrows correspond to $m = 0.1$, small arrows $m = 0.05$, solid lines $m = 0.01$, and dotted lines $m = 0.001$. For cases II and III, large arrows correspond to $Nm = 10$, small arrows $Nm = 5$, solid lines $Nm = 2$, and dotted lines $Nm = 1$. For case II, deme sizes in the upper half of each network are $N = 100$, and those in the lower half are $N = 1000$. For case III, deme sizes in the upper-left quadrant of each networks are $N = 1000$, upper-right $N = 500$, lower-left $N = 200$, lower-right $N = 100$.

The latter array represents a situation in which habitat is locally two-dimensional, but becomes linear at larger distances. Here the resistance distance obviates choosing between raw and log-transformed Euclidean distances, and it is superior to both.

Third, the advantages of the resistance distance are particularly evident in arrays with irregular shape and/or variation in migration rates. The array in Figure 4A includes a partial barrier between demes in the left and right portions of the array, making Euclidean distance an inappropriate predictor of genetic differentiation when measured across the barrier. The irregular shape is automatically factored into the resistance distance, however, resulting in much better predictions of patterns of genetic differentiation (Table 1). Cost-weighted distances do not perform as well, but are still an improvement over Euclidean distances because they skirt the barrier and measure within-habitat distance. Figure 4B shows a network with extreme (100-fold) variation in migration rates, and Figure 4C combines such variation with an irregular array shape. Cost-weighted distances help to compensate for differences in migration rates in these two arrays but still underperform resistance distances, which simultaneously account for heterogeneity in migration rates and multiple pathways between demes. The remaining array is analogous to the continuous case shown in Figure 1: demes of equal size and connected by equal migration rates but arrayed in an irregular network. Figure 5 compares scatterplots of genetic distance versus Euclidean distance, cost-weighted distance, and resistance distance for the array. Cost-weighted distances perform particularly poorly here, because multiple pathways connecting demes are especially important.

A fourth important result is that resistance distances appear reasonably robust to variation in deme sizes within arrays, as long as migration remains balanced. Table 2 shows correlations for cases II and III described in the caption for Figure 4, in which deme sizes varied from 100 to 1000. IBR predictions depend on the number of migrants exchanged between deme pairs—rather than on individual deme sizes and migration rates per se—and this simplification provides a good approximation for equilibrium genetic structuring in these cases as well.
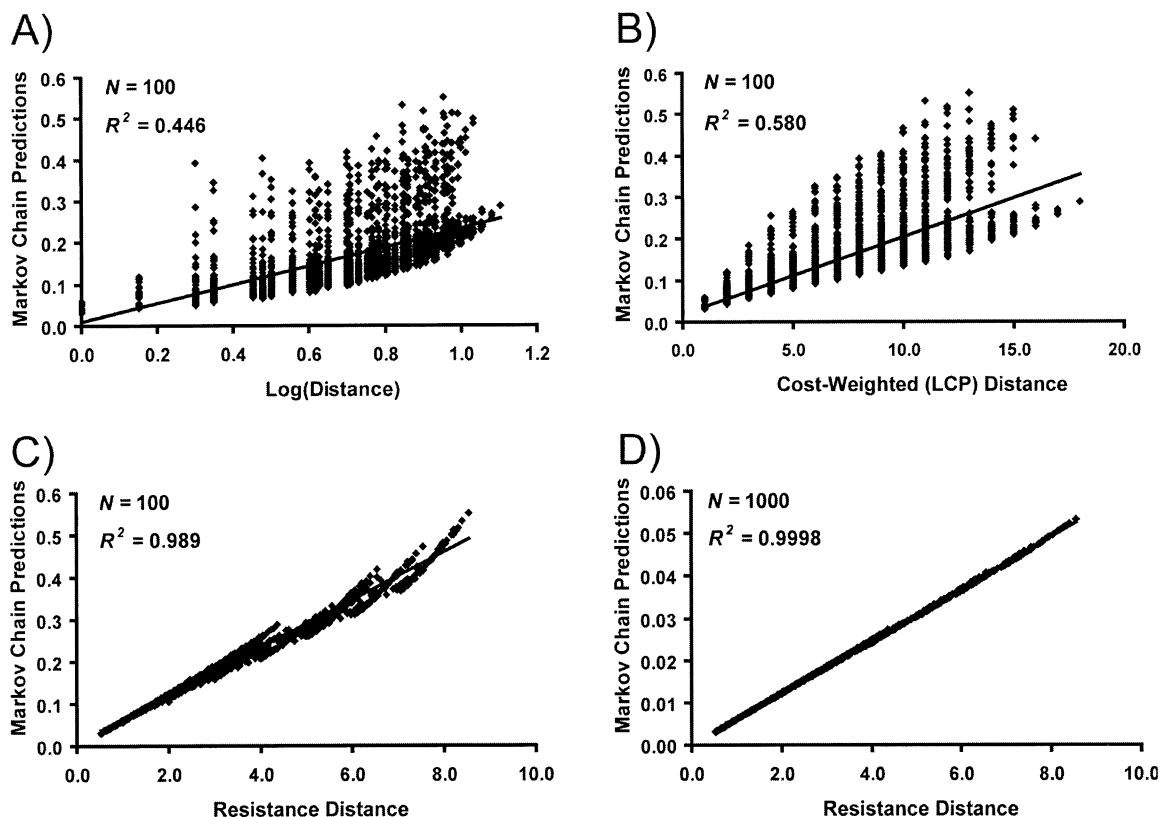
FIG. 5. Relationship of $F_{ST}/(1 - F_{ST})$ values predicted by Markov chain model for the deme array shown in Figure 4D, $N = 100$ case to (A) $\log_{10}$(distance), (B) cost-weighted (LCP) distance, and (C) resistance distance. (D) $F_{ST}/(1 - F_{ST})$ values for the same deme array, but with $N = 1000$ plotted against resistance distance. The resistance distance improves as a predictor of genetic differentiation with increased deme sizes and migration rates; this is not always the case for Euclidean or cost-weighted distances.

Finally, the resistance distance has distinct advantages when long-distance dispersal is important (e.g., in Fig. 3). If individuals in Figure 3 were five times more likely to disperse one step than two, correlations between genetic distance and both log-transformed Euclidean and cost-weighted distances would be high ($R^2 = 0.98$ and 0.93, respectively). If individuals were five times more likely to disperse two steps than one, the correlation would remain high for cost-weighted distance ($R^2 = 0.88$) but would drop substantially for Euclidean distance ($R^2 = 0.047$). When background migration is allowed such that $m = 0.05$ between adjacent demes and $m = 0.01$ between all nonadjacent demes, the predictive power of Euclidean distance ($R^2 = 0.95$) again exceeds that of cost-weighted distance ($R^2 = 0.68$). Clearly, traditional distance measures are inconsistent predictors of genetic differentiation here, and they could lead to spurious conclusions about actual migration patterns or degree of departure from genetic equilibrium (for similar examples see Matsen and Wakeley 2006). In contrast, the resistance distance and linearized F_{ST} are precisely related in these examples because of the isotropic migration pattern ($R^2 = 1.0$ for each case). Even when migration is not isotropic, long-distance dispersal improves IBR predictions (all else being equal), because within-deme coalescence times across networks become more similar.

When Absolute Demographic Parameters Are Unavailable

In each of the above cases, the resistance distance is used to predict genetic differentiation when deme sizes and migration rates are known. However, the metric is equally suited to the vast majority of cases when actual demographic parameters are unknown but relative values can be hypothesized. For example, consider a network of demes where we know the topology of the network (i.e., which pairs of demes exchange migrants and which do not). Either constant numbers of migrants between neighboring demes across the network are assumed, or relative rates are hypothesized for different pairs of demes (perhaps based on distance separating them or on direct observations of movement behavior in different habitat types). Conductance values that reflect the rel-

TABLE 2. Isolation-by-resistance (IBR) model performance when deme sizes vary across networks. Table entries are R^2-values for log-transformed distance (IBD), cost-weighted distance (LCP), and resistance distance (IBR) for cases II and III described in Figure 4.

Deme array	Case II			Case III		
	IBD	LCP	IBR	IBD	LCP	IBR
Figure 4A	0.34	0.93	1.0	0.45	0.89	1.0
Figure 4B	0.51	0.86	1.0	0.51	0.85	1.0
Figure 4C	0.38	0.92	1.0	0.40	0.92	1.0
Figure 4D	0.41	0.53	1.0	0.44	0.57	0.99

ative magnitude of migration can be applied to the network, and resistance distances can be calculated as indices of relative genetic differentiation between sampling locations. Had we applied relative values to the arrays considered above (e.g., $G = 1$ for all conductances in Fig. 4D, or conductances proportional to numbers of migrants exchanged in Fig. 4C), correlations and scatterplots would have been unchanged save for the x-axis labels.

Applying the Model to Continuous Landscapes

Although resistance distance calculations operate on discrete networks, the continuous case may be approximated using discretized habitat maps. This approach is not without precedent; Slatkin (1985) noted that results from unbounded continuum and discrete models are similar, a result supported for the finite case by the simulations of Slatkin and Barton (1989). Likewise, engineers commonly use discrete resistor networks as models of continuous conductive materials.

Approximating the continuous case involves discretizing maps of habitat and replacing habitat cells with nodes. Pairs of nodes representing cells that exchange migrants are then connected by resistors, the conductance values of which reflect relative numbers of migrants exchanged, thus forming a network. These conductance values are then entered into the same **G** matrix used to solve discrete deme networks. In the case shown in Figure 1 (binary habitat/nonhabitat, only local dispersal allowed), nodes representing habitat cells would be connected to neighboring habitat cells with resistors, while nonhabitat cells would be dropped, as shown in the inset. When diverse habitat types are represented in the same landscape, conductance values between habitat cells of differing type are set to reflect relative numbers of migrants exchanged between those habitats. Long-distance dispersal can be incorporated by adding resistors between nonadjacent nodes, as in Figure 3. Often the simplifying assumption of stepping-stone migration will be preferable for parsimony, but long-distance dispersal can have important consequences for genetic structuring (Slatkin 1985; Rousset 1997; Matsen and Wakeley 2006), and considering such movements will be essential when habitat blocks are disjunct.

DISCUSSION

Empiricists have most commonly accounted for spatial heterogeneity in studies of isolation by distance by adjusting Euclidean distances to reflect landscape features that increase dispersal or act as barriers (e.g., King 1987; Arter 1990; Rowe et al. 2000; Vos et al. 2001; Lugon-Moulin and Hausser 2002; Arnaud 2003; Coulon et al. 2004; Schweiger et al. 2004; Vignieri 2005) or testing for discrete barrier effects using partial Mantel tests (Smouse et al. 1986). While these approaches are useful, they may mask the more subtle effects of limited and irregular habitat extent or spatial variation in demographic parameters. It is not difficult to imagine situations in which such phenomena (e.g., constrictions in a species' habitat along an isthmus between two larger, contiguous blocks of habitat) would produce structuring that could be misinterpreted as due to a barrier or departure from genetic equilibrium. Whenever possible, studies of natural populations should consider these more subtle effects of spatial

heterogeneity rather than assuming habitat to be homogeneous and unbounded or simply divided by discrete barriers.

The IBR framework offers an intuitive and flexible approach to more thoroughly integrate spatial heterogeneity into studies of natural populations. Like migration matrix and coalescent models, the IBR model can predict effective migration rates and pairwise F_{ST}-values when population sizes and migration rates are known. The model's most promising contribution, however, is its ability to generate indices of gene flow and genetic differentiation using large spatial datasets that provide only relative values of local demographic parameters. Rarely are actual deme sizes, population densities, or migration rates known for natural populations, but relative suitabilities of different habitat types and their permeabilities to movement can be readily generated using ecological data, published literature, or expert opinion (e.g., Boone and Hunter 1996; Vos et al. 2001; Clevenger et al. 2002; Arnaud 2003; Coulon et al. 2004).

Model limitations include assumptions of balanced, gametic dispersal, genetic equilibrium, and no mutation or selection. Model accuracy is also reduced when some demes are more inbred than others. This reduction in accuracy tends to zero with increasing numbers of migrants (Table 1).

Even in low migration cases, IBR predictions can greatly improve over commonly applied isolation-by-distance models, which share the limitations listed above and are further limited by assumptions of spatial homogeneity. Improvements over Euclidean distance are evident even in homogeneous arrays of regular shape (e.g., square arrays, Table 1).

Furthermore, the resistance distance provides a unified framework for scaling of distance. In elongated rectangular arrays, Rousset (1997) suggested using the two-dimensional model (i.e., regressing log-transformed distance versus $F_{ST}/[1 - F_{ST}]$) when habitat is locally two-dimensional and using the one-dimensional model at larger distances. The IBR model makes this choice unnecessary by subsuming both cases within a single model. Many empirical studies do not clearly specify whether habitat is assumed to be one- or two-dimensional, but the IBR model makes assumptions of habitat configuration explicit.

Decisions of how to apply the IBR model should involve critical thought about the spatial and temporal scales of the processes being modeled, recalling that patterns of genetic structuring result from the cumulative effects of drift and migration over long time periods. In some cases, it may be most appropriate to correct only for the presence of persistent features, such as oceans or mountain ranges. The approach will be particularly forgiving with simple binary range maps as long as grid cell sizes capture the salient features of the species' range (e.g., the two corridors in Fig. 1). The model can accommodate more complex formulations when clear hypotheses about landscape structure and population connectivity exist; however, including fine-scaled movement behaviors, particularly in response to ephemeral landscape features, may not always be necessary or justified.

Evaluating Model Fit

Visual inspection of scatterplots of resistance distances and genetic distances can be used along with R^2-values to in-

vestigate how well the model fits the data. A poor model fit may indicate that model assumptions have been violated or an inappropriate spatial model has been applied.

Chief among assumptions of IBR and isolation-by-distance models is that populations are in genetic equilibrium. Differentiation is assumed to be due to recurrent processes of gene flow and genetic drift, yet equation (4) depends on coalescence times over which significant landscape or population changes may have occurred. For example, maximum expected coalescence times for the arrays shown in Table 1 for the $N = 2$ case range from 108 generations (4×4, $m = 0.1$) to 6422 generations (Fig. 4D); when $N = 10,000$, maximum coalescence times exceed 10^6 generations (Fig. 4D). This discouraging result is at least partially mitigated by the fact that F_{ST} values are known to approach equilibrium quickly following demographic perturbations (Crow and Aoki 1984; Slatkin 1985). In addition, the simulations of Slatkin (1993) showed that patterns of isolation by distance can become evident long before coalescence times are reached. Still, model assumptions will hold best at relatively small geographic scales, where equilibrium is approached more rapidly and where mutation has less of an effect on differentiation (Slatkin 1985, 1993; Rousset 1997). Effects of historical events and of differences in selective pressures may also be less evident at short distances (Rousset 1997).

Even when populations are far from equilibrium, IBR analyses can be used to test hypotheses of how different factors contribute to genetic structuring. For example, the scatterplot diagnostics described by Hutchison and Templeton (1999) can be applied in conjunction with IBR predictions to investigate the interplay of migration, drift, and population history at different geographic scales. Indeed, applying such diagnostics without first correcting for spatial heterogeneity could easily lead to mistaken conclusions about the relative influence of these factors. In addition, multiple IBR runs can be used to investigate whether historical or current habitat configurations better explain genetic structuring, or to compare competing hypotheses about past landscapes and population distributions. Finally, it may often be of interest to know toward what pattern of genetic structuring the landscape is driving populations, regardless of whether equilibrium has been reached.

A second assumption of IBR and isolation-by-distance models is that all demes are equally outbred. Lack of fit can be due to differing levels of isolation of individual demes from the network as a whole, resulting in some demes being more inbred than others. Again, visual inspection of scatterplots can help to determine sources of residual variation. Pairwise genetic distances between samples that are highly isolated from the network as a whole (e.g., from populations with low numbers of migrants or at range edges) will tend to be higher than predicted, while those including more outbred samples will be lower than predicted. For example, the point most above the regression line in Figure 5C reflects distances between the demes in the upper-left and lower-left corners of the array in Figure 4D. Points furthest below the line tend to involve at least one well-connected deme from the array's interior.

Most importantly, poor model fit may also result from applying an incorrect spatial model. Well-connected habitat may have been assumed where cryptic barriers exist, important dispersal pathways may have been ignored, or long-distance dispersal may not have been considered. Alternative models may be investigated by calculating resistance distances using multiple candidate habitat maps or assigning different conductances to different habitat types. In this way, competing hypotheses regarding habitat configuration or dispersal behavior may be investigated.

Isolation-by-Resistance Versus Other Equilibrium Models

Choices of appropriate predictive models of genetic structuring will depend on the level of detail known about species' distributions and migration patterns. As with spatial connectivity metrics (Calabrese and Fagan 2004), these choices will involve trade-offs between data requirements and realism. At one extreme, we may have samples from different locations in a species' range, with no information about its habitat extent or range limits. At the other, we may have samples from a discrete array of demes, for which we know actual deme sizes and migration rates. Traditional isolation-by-distance models would be most appropriate for the first case, whereas the accuracy and generality of migration matrix and coalescent models make them most appropriate for the second.

Between these two ideals, there is a wide spectrum of situations for which the IBR approach can strike a useful balance between simpler and more complicated models. Unlike the alternatives, resistance distances can be calculated using a range of data inputs, including (in order of increasing detail) spatially homogeneous arrays with unknown demographic parameters, coarse-scale range maps, maps of habitat distribution within a species' range, maps of habitat types of differing qualities, and maps of discrete demes with known deme sizes and migration rates. The method's most promising use will likely be in the intermediate cases, when only range maps or models of habitat suitability (including potential barriers) are available, as this is an unfilled niche among gene-flow models.

Moreover, as described above, the IBR model has substantial computational advantages over Markov chain and coalescent models, greatly increasing the number of demes that can be analyzed. The ability to accommodate large networks is particularly relevant to continuously distributed populations, which could easily require solving discretized arrays with tens of thousands of grid cells.

Software

Software to calculate resistance distances in both discrete stepping-stone arrays and in continuous GIS habitat maps is available at http://www.nceas.ucsb.edu/~mcrae/software. The software accommodates habitats of differing qualities and allows connections between nonadjacent grid cells based on cell values and user-input distance functions.

Conclusions

Linear circuits can provide simple and intuitive models of gene flow and genetic differentiation in heterogeneous landscapes. Effective conductance is a readily calculable analog

to effective migration, and the resistance distance predicts relative levels of equilibrium genetic differentiation. IBR analyses can improve over standard isolation-by-distance approaches by incorporating hypotheses of how habitat heterogeneity constrains gene flow without the data requirements or computational complexities of migration matrix or coalescent models. After correcting for effects of habitat configuration, empiricists can then more reliably attribute residual variation to other factors such as cryptic barriers or historical events.

The IBR model should also allow prediction of genetic and evolutionary consequences of landscape change, particularly if first calibrated with contemporary landscape and genetic data. In this capacity, the model could be used to predict how conservation efforts or land-use scenarios may enhance or disrupt gene flow and species' capacities for adaptation, speciation, and macroevolutionary change (Templeton et al. 2001). The framework introduced here thus provides theoretical and practical links between the fields of evolutionary biology, landscape ecology, and conservation biology and should further efforts to synthesize these disciplines.

LITERATURE CITED

Adriaensen, F., J. P. Chardon, G. De Blust, E. Swinnen, S. Villalba, H. Gulinck, and E. Matthysen. 2003. The application of ''least-cost'' modeling as a functional landscape model. Landscape Urban Plan. 64:233–247.

Arnaud, J. F. 2003. Metapopulation genetic structure and migration pathways in the land snail *Helix aspersa*: influence of landscape heterogeneity. Landscape Ecol. 18:333–346.

Arter, H. E. 1990. Spatial relationship and gene flow paths between populations of the alpine snail *Arianta arbustorum* (Pulmonata: Helicidae). Evolution 44:966–980.

Babic, D., D. J. Klein, I. Lukovits, S. Nikolic, and N. Trinajstic. 2002. Resistance-distance matrix: a computational algorithm and its application. Int. J. Quantum Chem. 90:166–176.

Balaban, A. T., and D. J. Klein. 2002. Co-authorship, rational Erdös numbers, and resistance distances in graphs. Scientometrics 55:59–70.

Beerli, P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol. Ecol. 13:827–836.

Beerli, P., and J. Felsenstein. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using the coalescent approach. Proc. Natl. Acad. Sci. USA 98:4563–4568.

Bodmer, W. F., and L. L. Cavalli-Sforza. 1968. A migration matrix model for the study of random genetic drift. Genetics 59:565–592.

Boone, R. B., and M. L. Hunter. 1996. Using diffusion models to simulate the effects of land use on grizzly bear dispersal in the Rocky Mountains. Landscape Ecol. 11:51–64.

Calabrese, J. M., and W. F. Fagan. 2004. A comparison-shopper's guide to connectivity metrics. Front. Ecol. Environ 2:529–536.

Chandra, A. K., P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari. 1997. The electrical resistance of a graph captures its commute and cover times. Comput. Compl. 6:312–340.

Clevenger, A. P., J. Wierzchowski, B. Chruszcz, and K. Gunson. 2002. GIS-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. Conserv. Biol. 16:503–514.

Coulon, A., J. F. Cosson, J. M. Angibault, B. Cargnelutti, M. Galan, N. Morellet, E. Petit, S. Aulagnier, and A. J. M. Hewison. 2004. Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. Mol. Ecol. 13:2841–2850.

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proc. Natl. Acad. Sci. USA 81:6073–6077.

Dorf, R. C., and J. A. Svoboda. 2003. Introduction to electric circuits. 6th ed. John Wiley and Sons, New York.

Doyle, P. G., and J. L. Snell. 1984. Random walks and electrical networks. Mathematical Association of America, Washington, DC.

Hutchison, D., and A. R. Templeton. 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.

Jeng, M. 2000. Random walks and effective resistances on toroidal and cylindrical grids. Am. J. Phys. 68:37–40.

King, P. S. 1987. Macro- and microgeographic structure of a spatially subdivided beetle species in nature. Evolution 41:401–416.

Kingman, J. F. C. 1982. On the genealogy of large populations. Pp. 27–43 *in* J. Gani and E. J. Hannan, eds. Essays in statistical science. Applied Probability Trust, London.

Kirchhoff, G. 1845. Über den Durchgang eines elektrischen Stromees durch eine Ebene, insbesondere durch eine kreisförmige. Ann Physik Chemie 64:497–514.

Klein, D. J., and M. Randic. 1993. Resistance distance. J. Math. Chem. 12:81–85.

Lugon-Moulin, N., and J. Hausser. 2002. Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of the common shrew (*Sorex araneus*). Mol. Ecol. 11:785–794.

Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. 18:189–197.

Maruyama, T. 1970. Effective number of alleles in a subdivided population. Theor. Popul. Biol. 1:273–306.

Matsen, F. A., and J. W. Wakeley. 2006. Convergence to the island-model coalescent process in populations with restricted migration. Genetics 172:701–708.

McPeek, M. A., and R. D. Holt. 1992. The evolution of dispersal in spatially and temporally varying environments. Am. Nat. 140:1010–1027.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219–1228.

———. 2000. Genetic differentiation between individuals. J. Evol. Biol. 13:58–62.

Rowe, G., T. J. Beebee, and T. Burke. 2000. A microsatellite analysis of natterjack toad, *Bufo calamita*, metapopulations. Oikos 88:641–651.

Rueness, E. K., P. E. Jorde, L. Hellborg, N. C. Stenseth, H. Ellegren, and K. S. Jacobsen. 2003. Cryptic population structure in a large, mobile mammalian predator: the Scandinavian lynx. Mol. Ecol. 12:2623–2633.

Schwartz, M. K., L. S. Mills, K. S. McKelvey, L. F. Ruggiero, and F. W. Allendorf. 2002. DNA reveals high dispersal synchronizing the population dynamics of Canada lynx. Nature 415:520–522.

Schweiger, O., M. Frenzel, and W. Durka. 2004. Spatial genetic structure in a metapopulation of the land snail *Cepaea nemoralis* (Gastropoda: Helicidae). Mol. Ecol. 13:3645–3655.

Slatkin, M. 1985. Gene flow in natural populations. Annu. Rev. Ecol. Syst. 16:393–430.

———. 1991. Inbreeding coefficients and coalescence times. Genet. Res. 58:167–175.

———. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264–279.

———. 2005. Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. Mol. Ecol. 14:67–73.

Slatkin, M., and N. H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368.

Slatkin, M., and T. Maruyama. 1975. The influence of gene flow on genetic distance. Am. Nat. 109:597–601.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst. Zool. 35:627–632.

Strobeck, C. 1987. Average number of nucleotide differences in a sample subpopulation: a test for population subdivision. Genetics 117:149–153.

Sumner, J., F. Rousset, A. Estoup, and C. Moritz. 2001. 'Neighbourhood' size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandiae*) using individual genetic and demographic methods. Mol. Ecol. 10:1917–1927.

Templeton, A. R., R. J. Robertson, J. Brisson, and J. Strasburg. 2001. Disrupting evolutionary processes: the effect of habitat fragmentation on collared lizards in the Missouri Ozarks. Proc. Natl. Acad. Sci. USA 98:5426–5432.

Tufto, J., A. F. Raybould, K. Hindar, and S. Engen. 1998. Analysis of genetic structure and dispersal patterns in a population of sea beet. Genetics 149:1975–1985.

Vast, S., P. Dupont, and Y. Deville. 2005. Automatic extraction of relevant nodes in biochemical networks. 7e Conférence Francophone sur l'Apprentissage Automatique, May 31–June 3, 2005, Nice, France.

Vignieri, S. N. 2005. Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). Mol. Ecol. 14:1925–1937.

Vos, C. C., A. G. Antonisse-De Jong, P. W. Goedhart, and M. J. Smulders. 2001. Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). Heredity 86:598–608.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138.

Yen, L., D. Vanvyve, F. Wouters, F. Fouss, M. Verleysen, and M. Saerens. 2005. Clustering using a random-walk based distance measure. Pp. 317–324 *in* Proceedings of the 13th European symposium on artificial neural networks. D-side Publications, Evere, Belgium.

Corresponding Editor: B. Nürnberger