

Circuit theory predicts gene flow in plant and animal populations

Brad H. McRae*[†] and Paul Beier[‡]

*National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, CA 93101; and [‡]School of Forestry and Merriam–Powell Center for Environmental Research, Northern Arizona University, Flagstaff, AZ 86011

Edited by John C. Avise, University of California, Irvine, CA, and approved September 18, 2007 (received for review July 12, 2007)

Maintaining connectivity for broad-scale ecological processes like dispersal and gene flow is essential for conserving endangered species in fragmented landscapes. However, determining which habitats should be set aside to promote connectivity has been difficult because existing models cannot incorporate effects of multiple pathways linking populations. Here, we test an ecological connectivity model that overcomes this obstacle by borrowing from electrical circuit theory. The model vastly improves gene flow predictions because it simultaneously integrates all possible pathways connecting populations. When applied to data from threatened mammal and tree species, the model consistently outperformed conventional gene flow models, revealing that barriers were less important in structuring populations than previously thought. Circuit theory now provides the best-justified method to bridge landscape and genetic data, and holds much promise in ecology, evolution, and conservation planning.

Gulo gulo | isolation by resistance | landscape connectivity | *Swietenia macrophylla* | landscape genetics

Preserving and restoring connectivity for broad-scale ecological processes, such as dispersal and gene flow, has become a major conservation priority (1, 2). Conservation organizations are investing considerable resources—and asking governments to do the same—to set aside land to promote connectivity (3). A major impediment to this goal is the difficulty in predicting how different land use, climate change, or reserve design scenarios will affect connectivity, and conservation planning decisions are often made without quantifying benefits for the ecological processes they are meant to conserve. If scarce conservation dollars are to be spent effectively, conservation biologists need clear, efficient, and reliable tools relating landscape composition and pattern to important ecological processes (4).

Gene flow is a critical ecological process with conservation benefits ranging from promoting the persistence of small populations to spreading adaptive traits in changing environments (5–8). Because of these important ecological and evolutionary roles, a new and rapidly growing field—landscape genetics—is primarily dedicated to understanding and predicting how landscape characteristics affect gene flow (5). By combining genetic, computational, and spatial analytic tools unavailable a decade ago, the interdisciplinary field has yielded insights relevant not only to conservation (e.g., refs. 8 and 9), but to fields such as evolution (10), infectious disease ecology (11), and population ecology (12).

Yet progress in all of these fields has been hampered by a lack of models capable of predicting gene flow from landscape structure. As a result, typical landscape genetic analyses simply detect genetic discontinuities and propose ad hoc explanations based on coincident landscape features, rather than testing *a priori* predictions of how such features are expected to influence genetic structure. This hinders hypothesis testing and leaves conservation planners without validated metrics of how landscape change will affect genetic connectivity. Although a growing number of studies are incorporating landscape data into genetic predictions and conservation plans using least-cost path

(LCP) modeling (9, 13–16), gene flow among real populations is not restricted to single, optimal pathways, as LCP models assume; rather, it occurs over multiple pathways and often involves indirect allele movements spanning many generations (17). Thus, to support conservation planning and to enable hypothesis testing in landscape genetics, theoretically justified models of how landscape features facilitate or impede gene flow over multiple, direct and indirect pathways are needed.

Here, we use electrical circuit theory to overcome these limitations and test its performance using data from threatened plant and animal species. The isolation-by-resistance (IBR) model improves over previous gene flow models by using the resistance distance, a graph-theoretic distance metric based in circuit theory (18), to simultaneously consider all possible pathways connecting population pairs (19). The analogy between electrical and genetic connectivity is simple: as multiple or wider conductors connecting two electrical nodes allow greater current flow than would a single, narrow conductor, multiple or wider habitat swaths connecting populations allow greater gene flow. The model is further grounded in rigorous connections between effective resistances in electrical networks and gene coalescence times in population networks (19), giving it a stronger theoretical foundation than either conventional isolation-by-distance (IBD) models (17, 20) or popular LCP models. The IBR model has been shown to reliably predict the effects of landscape heterogeneity on gene flow and genetic structure in artificial population networks (19) but has thus far not been tested with real landscape or genetic data.

Results

To test the IBR model, we used landscape and genetic data from two threatened species. Both are vulnerable to extinction because of habitat loss and overexploitation (21), and conservation efforts rely on our ability to identify important landscape connections and barriers. The plant data set consisted of eight Central American populations of big-leaf mahogany (*Swietenia macrophylla*), an insect-pollinated tropical tree, with 287 individuals typed at seven microsatellite loci by Novick *et al.* (22) (Fig. 1A). The animal data set consisted of 12 western north American populations of the wolverine (*Gulo gulo*), a large circumpolar mustelid, with 461 animals typed at 12 microsatellite loci by Kyle and Strobeck (23) (Fig. 1B). These previous studies had used conventional IBD analyses to draw conclusions regarding the evolutionary history and future conservation needs of the species. However, such analyses assume habitat to be unbounded and were unable to incorporate effects of spatial heterogeneity, such as range limits, into predictions of genetic differentiation.

Author contributions: B.H.M. and P.B. designed research; B.H.M. performed research; B.H.M. contributed new reagents/analytic tools; B.H.M. analyzed data; and B.H.M. and P.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[†]To whom correspondence should be addressed. E-mail: mcrae@nceas.ucsb.edu.

© 2007 by The National Academy of Sciences of the USA

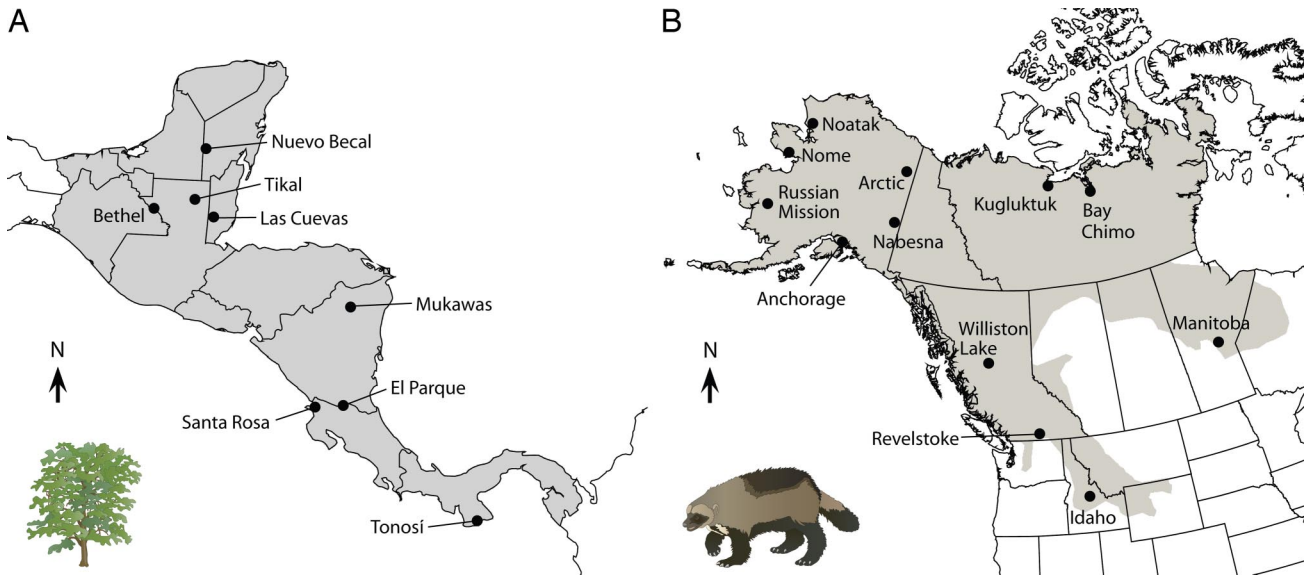


Fig. 1. Sample locations and modeled ranges. (A) Eight big-leaf mahogany populations (22), and modeled range following the Central American landmass from the Isthmus of Tehuantepec south through Panama (shading); extension of modeled range further north or south did not alter results. Sample sizes ranged from 16 to 55 individuals per population. (B) Twelve wolverine populations (23), and contiguous wolverine range in the western United States and Canada. Sample sizes ranged from 14 to 67 individuals per population. All results were consistent using alternative map projections, raster cell sizes, and genetic distance measures.

We reanalyzed each data set using IBR to determine whether incorporating known spatial heterogeneity could improve predictions of genetic differentiation among the sampled populations. Although IBR can accommodate multiple habitat types and landscape features, each with different effects on gene flow (19), for this initial empirical test we used simple, readily obtainable spatial data representing coarse-scale limits to each species' range (Fig. 1). For comparison we repeated analyses with LCP models using the same spatial data. We then evaluated the performance of IBR, LCP, and IBD models as predictors of observed differentiation between population pairs using Mantel tests (24) and causal modeling (25).

Model Performance. Circuit theory consistently outperformed IBD models for both species. Model fit improved 37–73% for mahogany (Fig. 2) and 188% for wolverines (Fig. 3). By contrast, LCP models yielded more modest improvements or none at all (Figs. 2C and 3B). Causal modeling also strongly supported the IBR hypothesis. For both species, a strong positive relationship between IBR predictions and genetic distance remained after controlling for all competing models (mahogany, $r > 0.82$, $P < 0.0005$; wolverines, $r > 0.79$, $P < 0.0001$). Conversely, partial correlations between all other models and genetic distances were either nonsignificant or negative ($r < 0$) after controlling for IBR.

Barrier Effects. Incorporating range shape using circuit theory not only improved gene flow predictions, but also illuminated the effects of other important factors driving population structure. Both of the original studies had found residual variation unexplained by distance and had attributed that variation to migration barriers or other factors driving differentiation within each species' range. For example, Novick *et al.* (22) noted that the southernmost mahogany population [Tonosí (Fig. 1A)] appeared to be more differentiated from other populations than would be expected from distance alone, and speculated that the increased differentiation may result from geographic barriers and/or historical events. We confirmed that the population was more differentiated than expected from either its distance [$r = 0.82$, $P = 0.0002$ (Fig. 2A)] or log-transformed distance [$r = 0.87$,

$P = 0.0001$ (Fig. 2B)] from the other populations. But evidence for increased differentiation was considerably weakened after controlling for IBR [$r = 0.40$, $P = 0.025$ (Fig. 2D)], and the additional differentiation, as measured by partial regression coefficients, was 51–55% smaller.

Correcting for range shape also diminished the inferred importance of barriers among the wolverine populations. As noted by Kyle and Strobeck (23), the Idaho population (Fig. 1B) was more differentiated from others than expected under IBD [$r = 0.92$, $P = 0.0001$ (Fig. 3A)]. The authors concluded that this likely resulted from low effective population sizes and restricted gene flow in the southern portion of the species' range, possibly due to barriers or population fragmentation (23). We found that Idaho was still more differentiated than expected from IBR predictions [$r = 0.87$, $P = 0.0001$ (Fig. 3C)], but the increased differentiation, as measured by partial regression coefficients, was 23% less than under IBD.

Robustness to Alternative Model Parameterizations. To examine the effects of different modeling choices on model performance, we repeated analyses with alternative genetic distance measures, map projections, node connection schemes, and extents and resolutions of habitat maps. The improved performance of IBR was consistent across all alternative model parameterizations, as shown in Tables 1 and 2. The use of different cell sizes and genetic distance measures had only modest effects on model fit, and the IBR model consistently outperformed IBD and LCP models in each case. Resistance distances calculated using alternative map projections, with diagonal connections between nodes, or with the modeled mahogany range extended to the north or south were highly correlated with original calculations ($r > 0.99$) and had no effect on correlations with any of the genetic distance measures.

Discussion

Three main results from our analyses have relevance for landscape genetics and for conservation. First, circuit theory consistently outperformed competing models for both the plant and animal data sets. Combined with its stronger theoretical foun-

Table 1. Results of mahogany analyses using different spatial resolutions and genetic distance metrics

Spatial distance metric	Genetic distance metric	
	$F_{ST} / (1 - F_{ST})$	$R_{ST} / (1 - R_{ST})$
Resistance distance (5-km cell size)	0.85	0.77
Resistance distance (50-km cell size)	0.83	0.73
Geographic distance	0.62	0.46
Log ₁₀ (geographic distance)	0.49	0.41
Least-cost distance	0.59	0.44

Table entries are R^2 values. The resistance distance was superior to geographic and least-cost distances in all cases.

distances differed little from Euclidean distances and offered no improvement over IBD predictions.

Improvement in model fit was greatest for the wolverine data (Fig. 3), where a more irregular range shape had larger effects on patterns of genetic differentiation. Northern populations are well connected by broad swaths of habitat, but southern populations such as Revelstoke, Manitoba, and especially Idaho (Fig. 1B) are in peninsular portions of the range, where we would expect genetic differentiation to increase more rapidly with distance. Here, LCP predictions improved over IBD by measuring within-habitat distances (Fig. 3B), but IBR also considered habitat width and availability of additional pathways (Fig. 3C), greatly improving connectivity estimates.

Explicitly incorporating the effects of range shape can alter conclusions regarding the roles that barriers and/or historical events have had in generating observed patterns of genetic structure. For mahogany, doing so called into question whether the southernmost population was more differentiated than would be expected in the absence of barriers. Although habitat gaps may currently separate northern and southern mahogany populations (26, 27), they apparently have not yet resulted in substantial divergence (Fig. 2D). Conversely, although correcting for range shape diminished the inferred importance of barriers among the wolverine populations, doing so also increased confidence in their effects. We found that much of Idaho's isolation can be explained by the shape of the wolverine's range, particularly its peninsular southern limit where habitat becomes more constricted (Fig. 3C). Still, because we have accounted for range shape, the substantial remaining variance can now be more confidently attributed to the factors cited by Kyle and Strobeck (23), namely barriers to dispersal or lower population densities in the south.

Importantly, the improved performance of the IBR model was evident when using even coarse-scale, binary habitat data. Moreover, model performance was relatively insensitive to choice of cell size (Tables 1 and 2). These results indicate that process-based connectivity models can still be informative in the many cases for which only poor-quality spatial data are available.

When spatial and genetic data permit, the IBR model can also incorporate more subtle effects of spatial heterogeneity within species' ranges. Model formulations can include fine-scaled variation in habitat quality, landscape resistances derived from detailed movement models, and/or connections between non-adjacent grid cells to accommodate long-distance dispersal (19). Such refinements will necessitate careful choice of appropriate scales of analysis. Even in complex landscapes, however, preliminary simulations indicate that as long as raster grids capture relevant landscape elements (e.g., corridors and barriers), resistance distances measured using different cell sizes are highly correlated (B.H.M., unpublished data).

Of course, genetic structuring is affected not only by spatial heterogeneity, but also by population history, and many factors other than recurrent processes of drift and migration likely affect

Table 2. Results of wolverine analyses using different spatial resolutions and genetic distance metrics

Spatial distance metric	Genetic distance metric		
	$F_{ST} / (1 - F_{ST})$	D_{LR}	D_S
Resistance distance (5-km cell size)	0.68	0.70	0.75
Resistance distance (50-km cell size)	0.71	0.73	0.77
Log ₁₀ (geographic distance)	0.24	0.24	0.28
Least-cost distance	0.37	0.37	0.43

Table entries are R^2 values. The resistance distance was superior to geographic and least-cost distances in all cases.

structuring in these (and nearly all) natural populations. Indeed, both species are undoubtedly undergoing rapid demographic changes due to anthropogenic pressure. However, if equilibrium processes are to be invoked at all (as they are in IBD analyses), circuit theory better predicts their effects than do traditional approaches.

Still, the utility of circuit theory is not limited to equilibrium situations. As with IBD analyses, scatterplots of predicted vs. observed genetic differentiation can be examined to test for equilibrium and evaluate the historical influences of gene flow and genetic drift on population structure (28). IBR can improve upon this method because it accounts for spatial heterogeneity in null model predictions, rather than assuming habitat to be unbounded or simply divided by discrete barriers. Applying these diagnostics to our IBR plots (Figs. 2D and 3C) indicates no evidence for the nonequilibrium effects detectable by the method (28).

We also see potential for directly modeling nonequilibrium processes by drawing on well established methods for analyzing dynamic circuit behavior. The methods incorporate circuit elements such as capacitors and inductors that store and release energy over time and resist instantaneous changes in currents and voltages. Because they can predict transient behavior resulting from initial conditions, perturbations, or changes in network configurations over time, these methods may be useful in predicting lasting effects of historical events in complex population networks. Recent work relating dynamic circuit behavior to random walks (29) is particularly encouraging. Just as the IBR model is based on links between static circuit analyses and random walk theory (19), these findings may pave the way for modeling nonequilibrium processes as well.

More immediately, circuit theory should be of broad use in other areas of connectivity research relevant to ecology, evolution, and conservation. Such applications appear overdue. For example, it has long been known that patterns of current flow across electrical networks have precise relationships with movement patterns and fates of random walkers in analogous graphs (30); this means that circuit theory can be applied to predict not only emergent patterns of gene flow over large distances, but also patterns of individual movement, mortality, and dispersal success at finer scales. As with gene flow predictions, such analyses can better assess the importance of multiple pathways for dispersal than popular LCP models, while well established connections with random walk theory give them a stronger theoretical foundation. We hope that these new models will be used to identify and map critical linkages connecting threatened populations in landscapes under development pressure. Coupled with the present work, such efforts should make increasingly clear how spatial heterogeneity affects connectivity for important ecological and evolutionary processes.

Methods

We chose plant and animal data sets from the literature according to the following criteria. We required that there be published

We thank M. Slatkin, P. Flikkema, E. Damschen, J. Novembre, R. Fredrickson, P. Keim, C. Kyle, R. Novick, L. DeWald, T. Theimer, J. Gregg, N. Schumaker, and the Sisk and Waits laboratory groups for discussions. We also thank the Integration and Application Network for supplying wolverine and mahogany symbols. This work was supported in part by graduate fellowships from the Environmental Protection Agency

Science to Achieve Results program and the Merriam–Powell Center for Environmental Research. B.H.M. was supported as a Postdoctoral Associate at the National Center for Ecological Analysis and Synthesis, a center funded by National Science Foundation Grant DEB-0553768, the University of California at Santa Barbara, and the State of California.

1. Rosenberg DK, Noon BR, Meslow EC (1998) *Bioscience* 47:677–687.
2. Crooks KR, Sanjayan M, eds (2006) *Connectivity Conservation* (Cambridge Univ Press, Cambridge, UK).
3. Kareiva P (2006) in *Connectivity Conservation: Maintaining Connections for Nature*, eds Crooks K, Sanjayan MA (Cambridge Univ Press, Cambridge, UK), pp 293–295.
4. Calabrese JM, Fagan WF (2004) *Front Ecol Environ* 2:529–536.
5. Manel S, Schwartz MK, Luikart G, Taberlet P (2003) *Trends Ecol Evol* 18:189–197.
6. Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) *Nature* 392:491–494.
7. Slatkin M (1987) *Science* 236:787–792.
8. Templeton AR, Robertson RJ, Brisson J, Strasburg J (2001) *Proc Natl Acad Sci USA* 98:5426–5432.
9. Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) *J Appl Ecol* 44:714–724.
10. Petren K, Grant PR, Grant BR, Keller LF (2005) *Mol Ecol* 14:2943–2957.
11. Real LA, Henderson JC, Biek R, Snaman J, Jack TL, Childs JE, Stahl E, Waller L, Tinline R, Nadin-Davis S (2005) *Proc Natl Acad Sci USA* 102:12107–12111.
12. Rueness EK, Stenseth NC, O'Donoghue M, Boutin S, Ellegren H, Jakobsen KS (2003) *Nature* 425:69–71.
13. Adriaansen F, Chardon JP, Blust GD, Swinnen E, Villalba S, Gulink H, Matthysen E (2003) *Landsc Urban Plann* 64:233–247.
14. Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004) *Mol Ecol* 13:2841–2850.
15. Vignieri SN (2005) *Mol Ecol* 14:1925–1937.
16. Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) *Am Nat* 168:486–499.
17. Slatkin M (1993) *Evolution (Lawrence, Kans)* 47:264–279.
18. Klein DJ, Randic M (1993) *J Math Chem* 12:81–85.
19. McRae BH (2006) *Evolution (Lawrence, Kans)* 60:1551–1561.
20. Rousset F (1997) *Genetics* 145:1219–1228.
21. World Conservation Union (2006) *IUCN Red List of Threatened Species* (World Conservation Union, Gland, Switzerland).
22. Novick RR, Dick CW, Lemes MR, Navarro C, Caccone A, Bermingham E (2003) *Mol Ecol* 12:2885–2893.
23. Kyle JC, Strobeck C (2001) *Mol Ecol* 10:337–347.
24. Mantel N (1967) *Cancer Res* 27:209–220.
25. Legendre P, Legendre L (1998) *Numerical Ecology* (Elsevier, Amsterdam).
26. Lamb FB (1966) *Mahogany of Tropical America: Its Ecology and Management* (Univ of Michigan Press, Ann Arbor, MI).
27. Calvo J (2000) *Evaluation of Mahogany (Swietenia macrophylla King) in Mesoamerica* (PROARCA/CAPAS, Tropical Science Center, Costa Rica).
28. Hutchison DW, Templeton AR (1999) *Evolution (Lawrence, Kans)* 53:1898–1914.
29. Qian H, Nassif SR, Sapatnekar SS (2005) *IEEE Trans Comput Aided Des Integr Circuits Syst* 24:1204–1224.
30. Doyle PG, Snell JL (1984) *Random Walks and Electric Networks* (Math Assoc of Am, Washington, DC).
31. Wilson DE (1982) in *Wild Mammals of North America*, eds Chapman JA, Feldhamer GA (Johns Hopkins Univ Press, Baltimore), pp 644–652.
32. Pennington TD, Styles BT, Taylor DAH (1981) *Flora Neotrop* 28:1–449.
33. Ray N (2005) *Mol Ecol Notes* 5:177–180.
34. Smouse PE, Long JC, Sokal RR (1986) *Syst Zool* 35:627–632.
35. Legendre P, Vaudor A (1991) *The R Package: Multidimensional Analysis, Spatial Analysis* (Département de Sciences Biologiques, Université de Montréal, Montréal).