Winter 2-25-2013

Landscape Genetics of Leaf-Toed Geckos in the Tropical Dry Forest of Northern Mexico

Christopher Blair

Victor H. Jimenez-Arcos
Universidad Nacional Autónoma de México

Fausto R. Mendez de la Cruz
Universidad Nacional Autónoma de México

Robert W. Murphy
University of Toronto

Recommended Citation

PLOS ONE

Landscape Genetics of Leaf-Toed Geckos in the Tropical Dry Forest of Northern Mexico

Christopher Blair[1,2]*[¤], Victor H. Jiménez Arcos[3], Fausto R. Mendez de la Cruz[3], Robert W. Murphy[1,2,4]

1 Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada, 2 Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada, 3 Laboratorio de Herpetología, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, México, 4 State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming, China

Abstract

Habitat fragmentation due to both natural and anthropogenic forces continues to threaten the evolution and maintenance of biological diversity. This is of particular concern in tropical regions that are experiencing elevated rates of habitat loss. Although less well-studied than tropical rain forests, tropical dry forests (TDF) contain an enormous diversity of species and continue to be threatened by anthropogenic activities including grazing and agriculture. However, little is known about the processes that shape genetic connectivity in species inhabiting TDF ecosystems. We adopt a landscape genetic approach to understanding functional connectivity for leaf-toed geckos (*Phyllodactylus tuberculosus*) at multiple sites near the northernmost limit of this ecosystem at Alamos, Sonora, Mexico. Traditional analyses of population genetics are combined with multivariate GIS-based landscape analyses to test hypotheses on the potential drivers of spatial genetic variation. Moderate levels of within-population diversity and substantial levels of population differentiation are revealed by F_{ST} and D_{est}. Analyses using STRUCTURE suggest the occurrence of from 2 to 9 genetic clusters depending on the model used. Landscape genetic analysis suggests that forest cover, stream connectivity, undisturbed habitat, slope, and minimum temperature of the coldest period explain more genetic variation than do simple Euclidean distances. Additional landscape genetic studies throughout TDF habitat are required to understand species-specific responses to landscape and climate change and to identify common drivers. We urge researchers interested in using multivariate distance methods to test for, and report, significant correlations among predictor matrices that can impact results, particularly when adopting least-cost path approaches. Further investigation into the use of information theoretic approaches for model selection is also warranted.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Christopher.Blair@utoronto.ca

¤ Current address: Department of Biology, Duke University, Durham, North Carolina, United States of America

Introduction

Dispersal is a fundamental process that can greatly influence ecological and demographic trajectories within and between subpopulations [1]. For example, dispersal often leads to gene flow, the transfer of genetic information from one population to another [2,3]. Maintaining adequate rates of gene flow is often beneficial because populations experiencing little gene flow are susceptible to a loss of genetic diversity due to inbreeding and drift [2,4,5]. A lack of genetic diversity may also make it difficult for populations to adapt to changing environmental conditions that may lead to local extinction [6,7].

Differences in dispersal rates can result from factors including species-specific philopatry, intra- and interspecific interactions, predation, physiological tolerances, and simple geographic (Euclidean) distance [8]. Landscape heterogeneity often plays a substantial role in the ability and/or choice of an organism to disperse or not [9–11]. Both natural and anthropogenic habitat fragmentation can detrimentally affect the connectivity and persistence of populations [12–15]. Landscape genetics seeks to explicitly quantify the influence of landscape and environmental

variables on microevolutionary processes such as gene flow and natural selection [16,17]. The approach extends traditional population genetic studies by explaining the spatial distribution of genetic variation using components of the landscape. This particularly powerful approach to studying fine-scale population structure and its application is meeting with much success [18]. By combining rapidly evolving molecular markers such as microsatellites with novel approaches to statistical analysis, landscape genetics identifies a suite of environmental variables likely to influence population genetic structure [19–21]. Identifying the landscape components facilitating or constraining gene flow can aid in delimiting areas for conservation [22], for example, by designing corridors that maximize functional connectivity [23,24]. Today, few landscape genetic studies focus on tropical areas [18], which harbour the majority of species [25]. Compared to temperate localities, relatively little is known about the processes influencing functional connectivity in species inhabiting this mega-diverse region.

Tropical deciduous or dry forests (TDF), which also occur in the Neotropics, are a major biodiversity hotspot [25]. They form a

semi-continuous belt throughout the New World from northern Mexico southwards into northern South America. Both structural and functional differences differentiate these forests from tropical rainforests [26]. The amount of annual rainfall is a primary distinction between these forests, with TDF experiencing up to eight months of arid-like conditions followed by four months of deluge. Although evidence suggests that TDF may be far more diverse than currently realized [27], habitat fragmentation due to both natural and anthropogenic factors is threatening the evolutionary potential of species inhabiting this ecosystem [28,29]. Habitat fragmentation is of particular concern in rapidly developing countries such as Mexico, where dense continuous forest is being cleared for both livestock and agriculture [28]. In Mexico, TDF reaches its northern limit near Alamos, Sonora, whereas forest density is highest in the southwestern states of Jalisco, Colima, Michoacan, and Guerrero [30]. Fragmentation of these forests is documented to have occurred for decades and continues to increase [31,32]. Unfortunately, we know relatively little about how fragmentation and other anthropogenic influences affect species and populations distributed throughout these ecosystems.

The diversity of Mexico's herpetofauna is substantial, with approximately 1,000 described species and many more awaiting formal description [27,33]. Flanking the Pacific Coast, Mexico's TDF also appears to be a centre of endemism for a variety of amphibian and reptilian taxa [34] including many species of leaf-toed geckos of the genus *Phyllodactylus* [35]. These lizards inhabit arid to semi-arid areas from southern California southwards through Middle America into northern South America and into the West Indies. Like many geckos, they are commonly found on vertical surfaces including bridges and buildings. They also appear to be common in close proximity to small streams, suggesting that riparian connectivity may be an important predictor of dispersal patterns.

Herein we use the Mexican yellow-bellied gecko, *Phyllodactylus tuberculosus*, to understand the effects of landscape configuration and anthropogenic influence on functional connectivity for a small terrestrial vertebrate presumably dependent on TDF. This species is an ideal choice to examine the relationship between landscape and genetics for several reasons. First, the geographic distribution of the gecko mirrors the distribution of TDF in Mexico. Second, abundance is relatively high when populations are isolated, providing a statistically suitable model. Third, along with others in the genus, this species is at risk of local and area-wide extirpation due to habitat fragmentation and recent introductions of non-native, all female species, such as geckos of the genus *Hemidactylus*, which appears to be displacing leaf-toed geckos (pers. obs.). Further, local people actively kill leaf-toed geckos as they are presumed to be venomous and dangerous to humans. For these reasons, leaf-toed geckos may soon pose a conservation concern and identifying landscape components that maximize genetic connectivity may be necessary for managing the persistence of populations. Our study site lay in the northernmost limit of TDF near Alamos, Sonora. We specifically test the following hypotheses: 1) high levels of genetic diversity and differentiation occur over small spatial scales; 2) anthropogenic fragmentation influences functional connectivity; 3) riparian connectivity predicts dispersal patterns, yet 4) some rivers act as dispersal barriers. We also test for an influence of slope and temperature on genetic patterns because the species is generally restricted to warm lowland habitats. Testing these hypotheses and drawing robust conclusions requires principles and practices drawn from diverse research disciplines [16]. Therefore, we use recent advances in landscape genetic techniques [18,36,37] to test our hypotheses [38] and identify which landscape variables influence functional connectivity in leaf-toed geckos. We are particularly interested in adopting multivariate approaches to model selection to assess the relative influence of multiple variables simultaneously [37].

Materials and Methods

Sampling

Our study area (Fig. 1) has a relatively high degree of forest cover compared to other locations throughout western Mexico due, in part, to the federal protection of land (Sierra de Alamos/ Rio Cuchujaqui Reserve). The landscape at lower elevations (generally under 300 m) consists of tropical dry thornscrub that gradually transitions into TDF closer to the Sierra de Alamos with increasing elevation. From 2008 to 2010, we sampled 336 leaf-toed geckos from 12 different localities (mean = 28 individuals per locality) throughout the landscape surrounding the Alamos region (Table 1). Sampling localities were chosen based on landscape characteristics to allow for testing our hypotheses. We sampled on opposite banks of two relatively large rivers or arroyos (Rio Cuchujaqui and Arroyo Tabelo) to test the hypothesis that rivers served as barriers to gene flow (in addition to conduits through opposite banks). Following Animal Use Protocols approved by the Royal Ontario Museum Animal Care Committee, tail tips were taken in the field and immediately preserved in 95% ethanol for subsequent genetic analysis. Subsequently, all individuals were released at the precise site of capture.

DNA extraction and genotyping

DNA was extracted using standard phenol-chloroform procedures. We used polymerase chain reaction (PCR) to amplify 12 polymorphic microsatellite loci developed specifically for *P. tuberculosus* [39]. Both negative and positive controls were run on each PCR plate. PCR products were visualized on an ABI 3730 automated sequencer (Applied Biosystems Inc.) at the Royal Ontario Museum. Genotyping was performed using GENEMARKER v.1.95 (SoftGenetics). We re-ran PCRs for approximately 10% of our samples to quantify any potential errors in genotyping.

Genetic diversity

We calculated diversity statistics for each site including number of alleles and allelic richness using MICROSATELLITE ANALYSER v.4.05 [40]. Observed (H_O) and expected (H_E) heterozygosites were calculated in TFPGA v.1.3 [41]. We tested for site-specific deviations from Hardy-Weinberg (exact test) and linkage equilibrium using GENEPOP v.4.0.10 [42,43]. Significance for tests was assessed using the Markov chain method using 100 batches with 1,000 iterations per batch. We used the false-discovery rate method to control for multiple comparisons [44].

Genetic differentiation

We calculated both traditional F_{ST} and Jost's D between populations. Unlike F_{ST}, D measured the degree of allelic differentiation between populations and was particularly useful for highly polymorphic markers such as microsatellites [45]. MICROSATELLITE ANALYSER was used to calculate both global and pairwise multilocus F_{ST} [46]. Significance of comparisons was assessed using 10,000 permutations while implementing a Bonferroni correction for multiple tests. We used SMOGD v.1.2.5 [47] to calculate pairwise D_{est} based on harmonic means estimated over all loci. Unbiased estimates of all metrics were employed to account for artefacts of sample size [45,48].
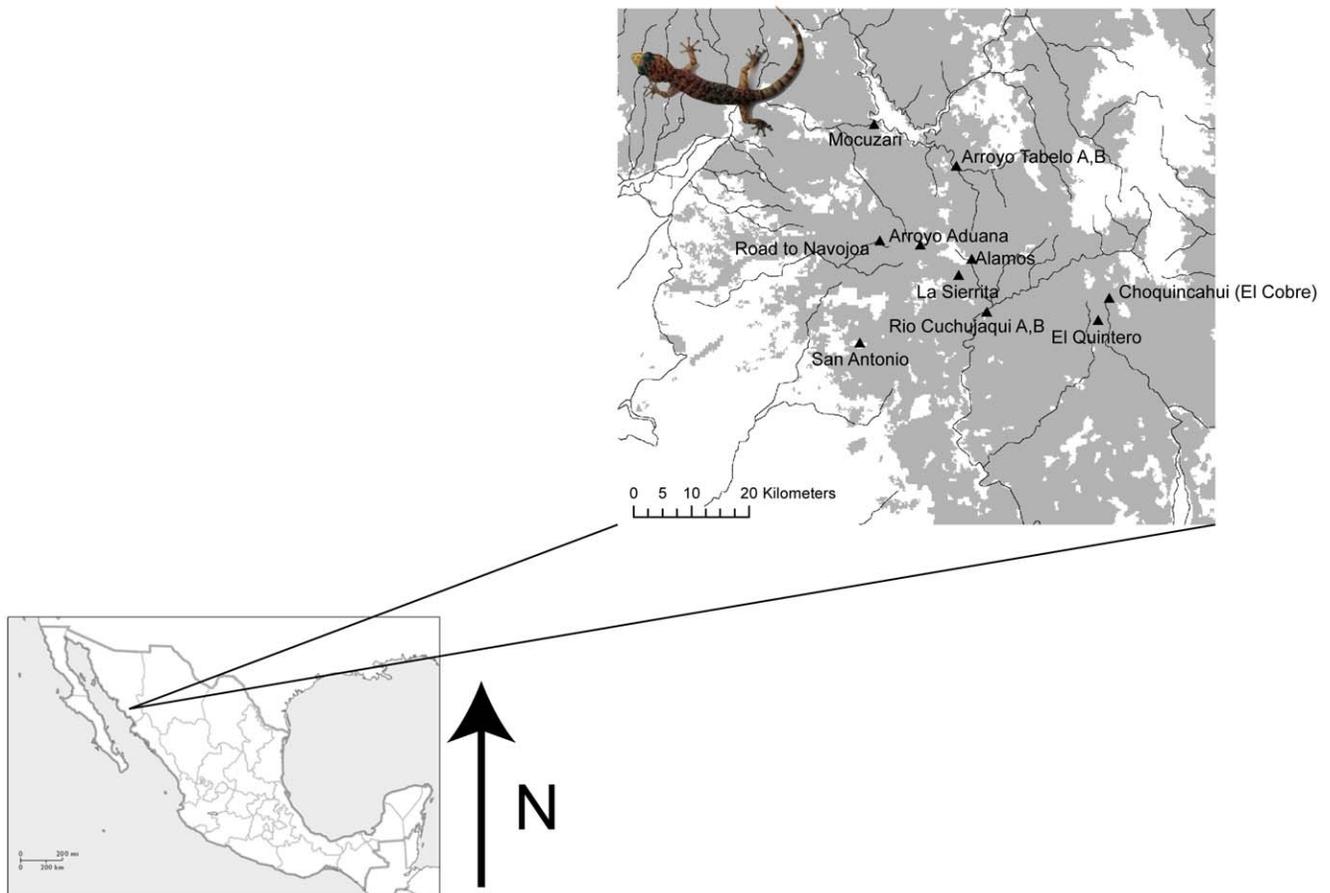
Figure 1. Sampling sites for all individuals and populations of *Phyllodactylus tuberculosus* included in this study. Darker shades of gray represent tropical dry forest. Dark lines represent rivers, streams and arroyos sampled throughout the study area.
doi:10.1371/journal.pone.0057433.g001

Table 1. Characterization of genetic diversity of *Phyllodactylus tuberculosus* at each sampling site included in this study based on data from 10 microsatellite loci.

Population	Elevation (m)	n	n(msat)	H_O	H_E	# alleles	allelic richness
Road to Navojoa	455	10	10	0.717	0.755	65	65.000
Alamos	375	36	35.2	0.762	0.773	88	64.887
Tabelo A	167	27	26.3	0.701	0.714	70	56.555
Tabelo B	199	30	29	0.721	0.734	69	54.234
Aduana	497	17	16.5	0.756	0.780	75	65.606
Rio Cuchujaqui A	358	30	28.9	0.749	0.762	86	65.451
Rio Cuchujaqui B	261	42	41.3	0.772	0.781	83	64.530
Sierrita	483	38	36.4	0.698	0.707	89	63.881
Mocuzari	124	30	29.6	0.701	0.713	70	55.220
El Quintero	361	30	29	0.690	0.703	83	61.761
Choquincahui (El Cobre)	433	31	29.2	0.709	0.721	81	59.340
San Antonio	388	15	13.6	0.724	0.752	69	62.731

n = number of individuals, n(msat) = number of individuals accounting for missing data, H_O = observed heterozygosity, H_E = expected heterozygosity; allelic richness calculated based on population with smallest sample size (n = 10).
doi:10.1371/journal.pone.0057433.t001

Population structure

We tested for the presence of genetic clusters using STRUCTURE v.2.3.3 [49,50] to infer population structure. We employed the admixture model with correlated allele frequencies while specifying a range of K-values (1–12). For simplicity, we created a batch run specifying the range of K-values to be evaluated and implemented 10 independent runs per K. Each run was composed of a burn-in of 30,000 followed by 100,000 iterations, which was sufficient to reach convergence. We evaluated the most likely number of clusters using both the ln Pr (X/K) and the ΔK method [51]. STRUCTURE HARVESTER [52] was used to visualize outputs and calculate K based on both methods. To deal with the multi-modality of utilizing multiple independent runs, we used CLUMPP v.1.1.2 [53] to permute the admixture coefficients for the runs with the chosen K-value using the "Greedy" algorithm with 1,000 random input orders. DISTRUCT v.1.1 [54] was then used to visualize the output from CLUMPP.

Geographic (spatial) information has often provided valuable insights in population genetic structure [55]. Thus, we compared our aspatial STRUCTURE results to those that incorporated information about sampling locations as prior information [56]. We introduced an additional parameter (LOCPRIOR) into the clustering analysis by specifying a different integer for each sampling location. We then ran STRUCTURE under the same conditions as the aspatial model.

Landscape genetic analysis—least-cost paths

We first calculated effective distances between populations using least-cost path modeling [19,21,57,58]. This assessed the influence of different landscape and environmental variables on population genetic structure assuming a single optimal dispersal path. These effective distances were created in a GIS environment by parameterizing different resistance surfaces that represented the hypothesized relationship between a specific habitat feature and gene flow [8]. For example, if hypothesizing that urban development versus undisturbed habitat constrained the movement of individuals, we assigned a higher cost value to cells representing urban habitat.

We tested for the relative influence of several landscape variables on genetic differentiation based on pairwise D_{est}. Landscape variables were selected based on expert knowledge of which habitat characteristics were most likely important in shaping patterns of gene flow in the species [8]. Our first model was based on isolation-by-distance (IBD) [59], which assumed genetic differentiation was a by-product of simple Euclidean distance without regard to the landscape. Next, we tested a variety of landscape genetic hypotheses that explicitly considered the intervening matrix [10]. First, we tested for the influence of land cover-type (specifically TDF vegetation) on genetic connectivity. We utilized a raster data set produced by the North American Land Change Monitoring System (NALCMS). Nineteen different land cover-types were classified at a 250 m spatial resolution. We created resistance surfaces by reclassifying the data to assign higher cost values to non-forested versus forested habitat. We tested several different cost ratios (1:2, 1:10, 1:100, 1:1000) to determine how parameterization might have influenced the results. We used the Mantel test function in the R package ECODIST [60,61] to test for both the presence of IBD and to select among the four relative cost values chosen for parameterization. Optimal values were selected based on the Mantel r correlation statistic using 10,000 randomizations.

Because *P. tuberculosus* occurred in lowland tropical environments only, our second set of analyses developed least-cost paths based on slope. These data were derived from a GTOPO30 digital elevation model (DEM) with a 1 km^2 spatial resolution produced from Natural Resources Canada and the U.S. Geological Survey. This layer consisted of seven elevation classes that were reclassified into slope to test the prediction that gene flow occurred primarily throughout lowland habitats. Our study area encompassed an elevational range between 100 to 500 m above sea level. Because slope represented continuous data and because we assumed a linear relationship between slope and gene flow [62], we simply reclassified the data into 32 classes using floating point (i.e. continuous) cell values. Higher cost values were assigned to cells with higher slope. This enabled us to test the prediction that higher slope resulted in lower levels of gene flow or higher genetic differentiation.

Because we often captured geckos adjacent to streams and arroyos, we tested if dispersal occurred primarily via stream corridors. We first obtained a polyline file representing all of Mexico's streams and tributaries from the GISDataDepot, a site that compiled multiple data layers based on ESRI's Digital Chart of the World (DCW). To represent riparian corridors, we created 1 km buffers around stream networks in the polyline file. We then converted these data into a raster file with a cell size of 100 m^2 and assigned different cost values to cells encompassing buffered streams versus those that did not. We tested the same relative cost values as our land cover analysis (1:2, 1:10, 1:100, 1:1000) and selected the best values based on Mantel correlations.

We tested for effects of minimum temperature of the coldest period of the year because these lizards are predominantly found in warm tropical lowland habitats. A significant correlation between gene flow and minimum temperatures was predicted. Temperature data were obtained from the WorldClim database at a resolution of 1 km^2. As with slope, we assumed a linear relationship between temperature and gene flow and we reclassified the data into a continuous distribution with 32 classes. We assigned higher cost values to cells representing lower minimum temperatures.

Finally, we utilized a multivariate resistance surface representing the combined effects of anthropogenic land-use (anthropogenic model). Data were obtained from the Wildlife Conservation Society (WCS) and the Center for International Earth Science Information Network (CIESIN). These data represented the combined effects of population density, built-up areas, roads, railroads, navigable rivers, coastlines, land-use, and nighttime lights (The Last of the Wild, Version Two 2005). The data were categorized based on the Human Influence Index at a spatial resolution of 817 m^2. Cell values ranged between 0 and 64, with 0 representing no anthropogenic influence and 64 representing maximum influence. To create least-cost paths, we reclassified the data into a continuous distribution with 32 classes and assigned cost values ranging from 0 to 31, with 31 representing the highest cost to gene flow for cells with the highest anthropogenic influence. Like slope and temperature, we assumed a linear relationship between the degree of disturbance and gene flow.

For all least-cost path-analyses, we used the LANDSCAPE GENETICS TOOLBOX 1.2.3 [63] implemented in ARCMAP 10 to calculate effective distances between sampling localities. This calculated both the cumulative cost-distance and the length of the least-cost path between any two sampling points. Because both distances could have been sensitive to relative cost values [64], we tested several different relative values for categorical variables as described above. For all least-cost path-analyses, we used the cumulative cost-distance because this metric minimized the degree of multicollinearity in our predictors.

Landscape genetic analysis—circuit theory

We also modelled patterns of gene flow using circuit theory [65–67]. This so-called isolation-by-resistance (IBR) approach has been shown to be powerful in modeling functional connectivity in both simulated and empirical data sets [65,66]. We calculated resistance distances between populations using CIRCUITSCAPE 3.5.7 [68]. Each calculation used focal points in pairwise mode and an eight-neighbors connection scheme. Due to memory issues with the original 100 m stream data, we aggregated cells in this raster to a resolution of 200 m to obtain reasonable computing times. Resistance distances based on all other variables were calculated using the original resolution of the data layer (e.g. 250 m for the land cover). All calculations were based on values of per-cell resistance.

Statistical analysis

We used multiple regression analyses on distance matrices (MRM; [69,70]) in ECODIST to evaluate landscape-genetic relationships. Although a potentially powerful method for landscape-genetic inference [37], few studies have incorporated MRM analyses (e.g. [71]). Similar to the commonly used partial Mantel test [72], MRM was developed to test for significant relationships between a dependent distance matrix (e.g. linearized D_{est}) and a number of indicator matrices and identify the contribution of each explanatory variable to the overall fit of the model [69]. Further, MRM modeled polynomial and nonlinear relationships [70]. Each distance matrix was unfolded into vectors representing pairwise distances. The response vector (i.e. linearized D_{est}) was then regressed against each indicator vector (i.e. least-cost or resistance distances) and the significance of the model was assessed by permuting the objects of the response vector.

MRM models using all six explanatory variables were not created for the least-cost analyses due to a relatively high degree of collinearity among the predictors, which could have resulted in coefficients with large variances and lead to erroneous conclusions regarding the direction and magnitude of slope. Some authors have suggested calculating Variance Inflation Factors (VIF) for each predictor in a model to ascertain if collinearity might be a problem in parameter estimation [73]. Like previous studies, we used VIF values>10 as evidence for substantial multicollinearity [71]. Thus, for least-cost paths we selected candidate models (see below) based on both our hypotheses of interests and to minimize the potential error in estimated regression coefficients. Because multicollinearity was minimal with our resistance distances calculated from CIRCUITSCAPE, we included all variables in the model selection procedure. We predicted a negative relationship between genetic differentiation and several of our landscape features including stream networks and the degree of undisturbed habitat. Because multiple regression models account for the effects of all included predictors, we anticipated that the regression coefficients may change depending upon the other variables included in the model (e.g. Euclidean distance). Univariate MRM models all resulted in positive coefficients due to spatial autocorrelation. Thus, particular attention was placed on regression coefficients in highly supported models containing Euclidean distance as a predictor.

We utilized information theoretic criteria to select among candidate models hypothesized to be important predictors of spatial genetic variation in our system [74]. Specifically, we calculated second-order AIC values (AICc) for competing candidate models based on either least-cost or resistance distances. Candidate models were selected based on a priori hypotheses regarding which combination of variables best explained patterns of genetic structure. In all cases we tested fewer than 20 candidate

models [75]. The best model minimized the amount of information lost as represented by the combination of variables with the lowest AICc value [76]. Different combinations of variables were compared to the null model of IBD to determine if the incorporation of landscape variables explained more of the variation in D_{est}. We followed previous recommendations in assessing the relative importance of models [75]. We also used MuMIn [77] to calculate AICc weights for each model and we estimated the 95% confidence set of candidate models [74].

Because it remained unknown how information theoretic criteria performed when evaluating models based on pairwise distances, we compared our MRM results with a linear mixed modeling approach [78]. This approach was based on a maximum likelihood population effects (MLPE) model that explicitly accounts for non-independence of values in regressions on distance matrices [79]. We created linear mixed models using the R package LME4 [80] defining populations as the random effect and each predictor matrix as a fixed effect. Parameter estimation was performed using restricted maximum likelihood (REML). All predictor matrices were centered around their mean prior to analysis. Statistical significance of both fixed and random effects were calculated using the R package MIXMOD [81]. Finally, to select among competing models, we calculated R^2_β [82] for each model based on the Kenward-Roger F and degrees of freedom [83] calculated using the R package PBKRTEST [84]. MLPE models were calculated for the top set of candidate models as determined from the MRM analysis to compare relative performance.

Results

Genetic diversity

All microsatellite data were deposited in Dryad (Provisional DOI: doi:10.5061/dryad.tj1k5). The genotyping and scoring of microsatellite alleles had an error rate of less than 1%, and, thus, high reproducibility. After controlling for false discovery rates, some loci showed significant deviations from Hardy-Weinberg expectations within collecting sites. For example, locus G2_96 showed heterozygote deficits at six of the 12 sites, locus P7 at eight of 12, and locus G2_59 at four of 12 (Table S1). However, only three alleles were present at locus G2_59 and, thus, there was a high probability that random chance resulted in significance. Because loci G2_96 and P7 showed a significant heterozygote deficit at multiple sites, we ran preliminary analyses with and without these loci to see how results changed. Although results did not differ substantially, we adopted a conservative approach and chose to report results from subsequent analyses excluding these two loci.

After controlling for false discovery rates, a few loci showed signs of linkage disequilibrium. For example, locus G2_22 showed linkage to loci G2_96, P2, P7, P12, and P19. Locus G2_85 showed signs of linkage with locus P7, and locus P2 with P15. However, linkage occurred at only two of our 12 sites (Arroyo Tabelo B and Mocuzari) suggesting that our loci were, in fact, independent, unlinked markers.

In general, within-site diversity was moderate as shown by both expected heterozygosity and allelic richness (Table 1). Expected heterozygosity ranged from 0.7025 to 0.7812 and allelic richness from 54.2341 to 65.6058. Allelic diversity within loci over all populations ranged from three alleles at locus G2_59 to 26 alleles at locus G2_37 (mean number of alleles per locus = 13.1). In general, diversity estimates were fairly similar among sites. A highly significant positive relationship occurred between elevation and genetic diversity that generally corresponded to habitat-type (tropical thornscrub versus TDF; $R^2 = 0.5836; P = 0.004$; Fig. 2).
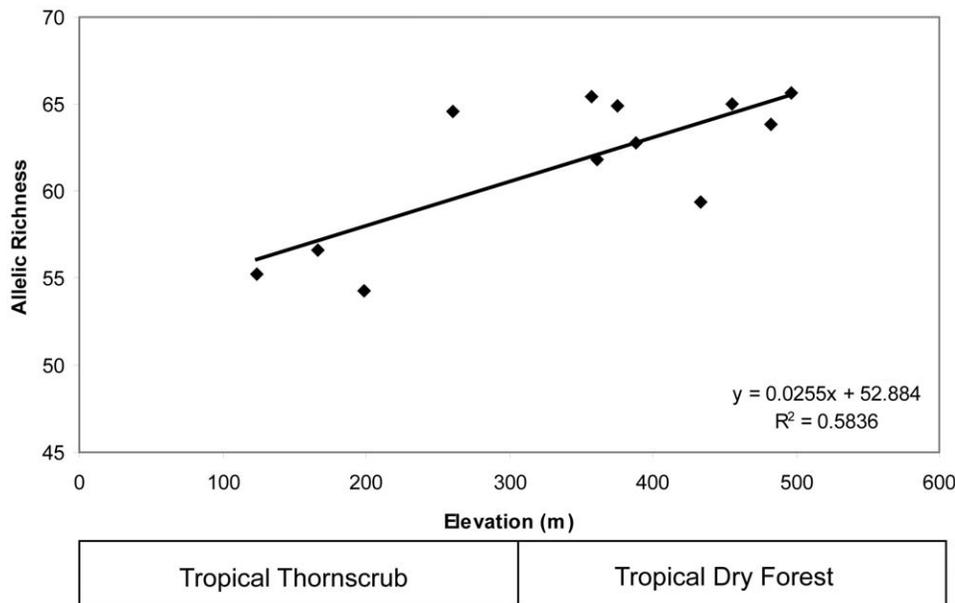
Figure 2. Relationship between elevation and genetic diversity (allelic richness) for all populations of *Phyllodactylus tuberculosus* included in this study.

Genetic differentiation

Moderate levels of global genetic differentiation were resolved based on both F_{ST} ($F_{ST} = 0.0869; P = 0.0001$) and D ($D_{est} = 0.248$). Pairwise measures of differentiation revealed moderate to high levels of population divergence (Table 2). Pairwise F_{ST} values ranged from zero between Aduana and the road to Navojoa (RN) to 0.201 between La Sierrita and El Quintero. Further, the majority of F_{ST} values were significant based on 10,000 random permutations of alleles after a Bonferroni correction ($P<0.00075$). Values of D_{est} ranged from moderate to high, showing similar relative values between populations. El Quintero and Choquinca-hui were the most divergent from the remaining populations based on both statistics. We found a high correlation between F_{ST} and D_{est} (Pearson r = 0.940). Because of this correlation, all subsequent landscape genetic analyses were performed with D_{est} only. We detected a significant positive correlation between Euclidean distance and D_{est} (Mantel r = 0.4897, P<0.001).

Population structure

An optimal *K*-value based on ΔK suggested a $K = 3$ ($\Delta K = 138.332$; Fig. 3a; Fig. S1b). A *K*-value of 5 was obtained using the ln Pr($X|K$) method (Fig. 3b; Fig. S1a). Individuals on opposite banks of Rio Cuchujaqui and Arroyo Tabelo did not form distinct clusters. However, individual-based Mantel tests found a highly significant barrier effect for both Rio Cuchujaqui (Mantel r = 0.1148; P <0.0001) and Arroyo Tabelo (Mantel r = 0.1126; P = 0.0006).

A second set of analyses utilized a model that explicitly incorporated prior information for sampling localities to aid in clustering. These results differed from the clustering results that did not utilize sampling localities. For example, whereas the ΔK method suggested a *K* of 2 ($\Delta K = 80.915$; Fig. 4a; Fig. S2b), the plot of *K* versus ln Pr($X|K$) reached a slight peak at 9 (Fig. 4b; Fig. S2a). For $K = 9$, some structure was resolved across opposite banks of Rio Cuchujaqui and Tabelo (Fig. 4b). The plot of *K* versus ln Pr($X|K$) showed that likelihood values began to stabilize at about $K = 3$, which was a value similar to that chosen using the ΔK

method in the aspatial analysis. At $K = 3$, cluster memberships and admixture coefficients were very similar to those of the aspatial analysis.

Landscape genetics—least-cost paths

For both forest cover and streams, Mantel *r* values were highest with a cost ratio of 1:2 (respectively Mantel r = 0.4344, P = 0.002; r = 0.6090, P<0.001). The MRM analysis based on least-cost path distances suggested that the incorporation of landscape variables explained significantly more variance in genetic differentiation than a simple IBD model (Table 3; Table S2). Whereas Euclidean distance was able to explain approximately 44% of the variation in D_{est} values, the incorporation of landscape variables increased this value to approximately 62%. The statistically best supported model (using AICc weights) was based on a combination of Euclidean distance, stream connectivity, and the degree of anthropogenic disturbance ($R^2 = 0.621$; $w_i = 0.856$). Our least-cost paths based on temperature and slope explained slightly more variation in D_{est} than Euclidean distance (0.548 and 0.539 versus 0.44, respectively; Table S1). Least-cost paths based on forest connectivity received relatively little support with similar weights to the Euclidean distance model. However, these two distances were highly correlated (Mantel r = 0.995; P<0.001) leading to very large VIF values. After accounting for Euclidean distance in models, regression coefficients for stream connectivity, forest, and undisturbed habitat (anthropogenic model) were negative and, thus, associated with a lower D_{est} and higher gene flow.

Similar results were obtained from the MLPE models with a model containing Euclidean distance, anthropogenic disturbance, and stream connectivity receiving the highest support ($R^2_\beta = 0.517$; Table 3). All fixed effects for each model were statistically significant as was the population (random) effect. In general, the sign of regression coefficients was identical between MRM and MLPE models. However, for Model C stream had a negative coefficient for MRM and a positive coefficient in MLPE. The relative support for top models was the same for both MRM and MLPE.

Table 2. Pairwise genetic differentiation between populations estimated from 10 microsatellite loci.

	Road to Navojoa	Alamos	Tabelo A	Tabelo B	Aduana	Cuch A	Cuch B	La Sierrita	Mocuzari	El Quintero	Choquincahui	San Antonio
Rd to Navojoa	--	0.106	0.181	0.174	0.000	0.119	0.075	0.048	0.094	0.325	0.311	0.113
Alamos	0.032	--	0.197	0.166	0.072	0.086	0.095	0.139	0.151	0.217	0.309	0.032
Tabelo A	**0.080**	**0.080**	--	0.050	0.125	0.196	0.223	0.257	0.133	0.260	0.309	0.248
Tabelo B	**0.065**	**0.059**	0.026	--	0.143	0.224	0.232	0.252	0.100	0.293	0.377	0.197
Aduana	−0.003	0.021	**0.058**	**0.059**	--	0.074	0.084	0.069	0.123	0.207	0.269	0.083
Cuch A	**0.039**	**0.028**	**0.082**	**0.076**	**0.024**	--	0.059	0.131	0.150	0.170	0.195	0.134
Cuch B	**0.033**	**0.029**	**0.087**	**0.075**	**0.031**	**0.021**	--	0.137	0.197	0.236	0.319	0.101
La Sierrita	**0.042**	**0.056**	**0.154**	**0.123**	**0.047**	**0.078**	**0.069**	--	0.176	0.345	0.378	0.155
Mocuzari	**0.045**	**0.063**	**0.068**	**0.049**	**0.056**	**0.055**	**0.070**	**0.097**	--	0.267	0.318	0.172
El Quintero	**0.149**	**0.118**	**0.147**	**0.151**	**0.128**	**0.096**	**0.111**	**0.201**	**0.150**	--	0.009	0.217
Choquincahui	**0.140**	**0.122**	**0.142**	**0.150**	**0.125**	**0.092**	**0.114**	**0.199**	**0.142**	0.008	--	0.336
San Antonio	0.040	0.019	**0.099**	**0.078**	0.027	**0.040**	**0.042**	**0.069**	**0.083**	**0.135**	**0.144**	--

Values above diagonal represent D_{est} and values below diagonal F_{ST}. Bold values of F_{ST} indicate significance (P<0.00075) after Bonferroni correction.
doi:10.1371/journal.pone.0057433.t002

Landscape genetics—circuit theory

Results of MRM based on resistance distances were similar to those based on least-cost path distances (Table 4). However, slope appeared to be an important variable influencing gene flow under a circuit-theoretic approach. The best supported model based on AICc weights included Euclidean distance, slope, and stream connectivity ($w_i = 0.585$), with stream being the only variable with a negative coefficient. The next most supported model, which included all six variables, received considerably less support ($\Delta AICc = 2.37$; $w_i = 0.167$), although well within the estimated confidence set. Visual examination of cumulative current based on a composite map of Euclidean distance, stream networks, and slope was highly congruent with the genetic clusters inferred from the STRUCTURE analysis (Fig. 5). In most cases AICc weights were higher for multivariate models incorporating landscape variables versus a model of simple IBD (Table S3). VIF values were less than 10 for all models suggesting that multicollinearity among predictors was not likely to be a problem.

Relative support for top candidate models differed between MRM and MLPE. Unlike MRM, MLPE favored the model containing all predictors ($R_\beta^2 = 0.673$) versus a model with Euclidean distance, slope, and stream connectivity ($R^2_\beta = 0.615$; Table 4). However, in the former many of the fixed effects were not significant. In general, direction of slope was similar for both MRM and MLPE models. Incongruence was only detected in models with non-significant predictors. The population (random) effect was highly significant in all models.

Discussion

This study highlights the power of adopting a landscape genetics approach to understanding functional connectivity for tropical organisms. We show how the incorporation of landscape heterogeneity into our models can enhance our understanding of processes shaping the genetic structure of populations. Rates of deforestation and habitat fragmentation are increasing exponentially throughout these regions. Coupled with climate change, we still know relatively little about how organisms will respond to these continuing threats [32,85–87]. Based on our results we cannot reject the hypothesis that landscape composition is an important predictor of spatial genetic variation in this system. This suggests that if we are to fully comprehend how natural and anthropogenic habitat alteration influences functional connectivity in tropical organisms, additional landscape genetic studies are required to infer species-specific responses to continued habitat change.

Genetic diversity and population structure

Our results detect moderate to high levels of genetic diversity within populations of P. tuberculosus. Diversity is not substantially higher in the population near Alamos versus the other areas, as might be predicted given the tendency for many gecko species to aggregate near human settlements. Further, a statistically significant relationship occurs between elevation and allelic richness. These results corroborate field observations that these lizards are predominantly encountered in TDF habitat. For example, elevation of our sites ranges from 100 m to 500 m. This range in elevation spans two distinct habitat types: tropical thornscrub and TDF [27]. An abrupt change in tropical vegetation occurs at approximately 400 m as well as an apparent change in abundance of geckos; more individuals are encountered per unit of time at El Quintero (361 m) and Choquincahui (433 m) than the other sites.

Significant population structure occurs based on both pairwise differentiation statistics and results from STRUCTURE. Most pairwise F_{st} values are high and statistically significant, with the localities El Quintero and Choquincahui being most divergent from all other populations. Values of D_{est} yield similar results. Thus, leaf-toed geckos appear to exhibit substantial population differentiation over relatively fine spatial scales.

Considerable recent debate exists as to the best F_{ST}-like analogue for assessing genetic differentiation between populations [48,88–90]. The traditional F_{ST} metric (G_{ST} for multiple loci and alleles) is highly sensitive to within-population heterozygosity, making it difficult to compare values between studies and markers [45,91]. Further, because F_{ST} is dependent on the number of alleles and heterozygosity, two populations can have low F_{ST} values despite sharing no alleles. For example, studies have shown that in two hypothetical populations with a total of 16 alleles, F_{ST} is constrained to less than 0.1 even when the two populations share no alleles [90]. New metrics such as G'_{ST} [91] and D [45] circumvent some of the more common issues found with F_{ST} [92]. Although these new metrics are not without criticism [93], a
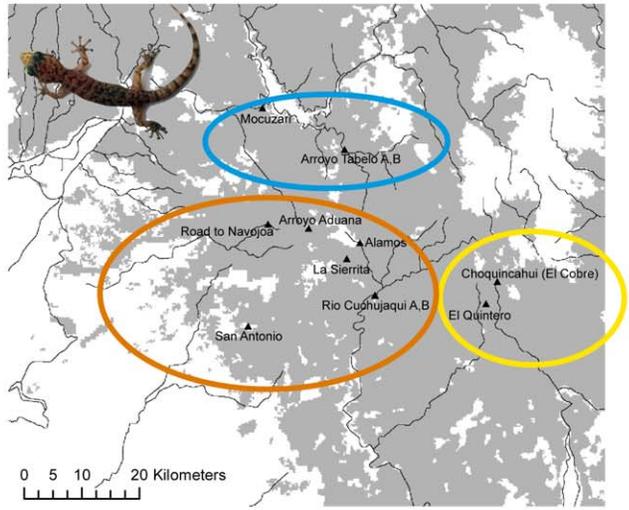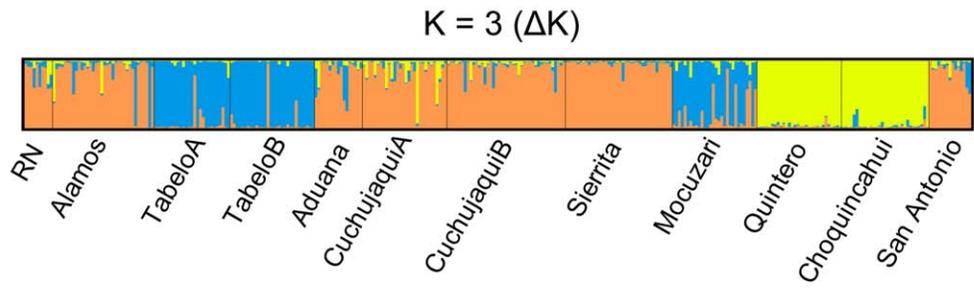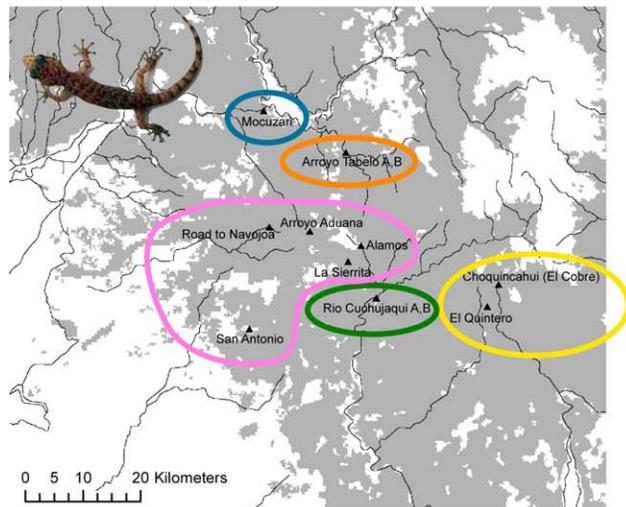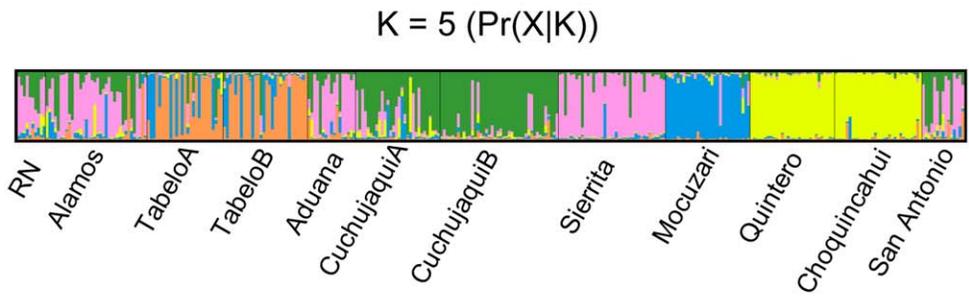
Figure 3. STRUCTURE results excluding prior information on sampling locality for *Phyllodactylus tuberculosus* included in this study. a) Results for $K = 3$ selected using the ΔK method. b) Results for $K = 5$ selected using ln Pr($X|K$). Each vertical bar represents a single individual with different colors representing the proportion of an individual's genome originating from that specific cluster. Names below plots represent population codes referred to in the text. Colored circles represent the most likely ancestry.
doi:10.1371/journal.pone.0057433.g003

comparison of multiple statistics can maximize information [92]. Whereas D is a useful metric for landscape-genetic inference, spatially explicit computer simulations may provide a greater understanding of how this measure compares to other commonly used differentiation statistics.

Our STRUCTURE analyses reveal cryptic population structure. We infer values of K both with and without incorporating prior information on sampling localities. Incorporating locality data as a prior changes our inference of K. The ΔK method suggests $K = 2$, whereas the plot of K versus ln Pr($X|K$) suggests $K = 9$. Populations at El Quintero and Choquincahui always group together into a single cluster and the populations at Mocuzari and Tabelo fall into separate clusters from the remaining populations in several STRUCTURE analyses. STRUCTURE plots generally corroborate the levels of differentiation based on D_{est} and F_{ST}.

We never detect structure on opposite banks of the Río Cuchujaqui or Arroyo Tabelo based on aspatial Bayesian clustering (i.e. individuals on opposite banks cluster as one group with similar admixture coefficients). However, spatial clustering at $K = 9$ reveals different ancestries on opposite banks. Further, Mantel tests suggest a significant barrier effect for these features. Although Bayesian clustering methods, and STRUCTURE in particular, can be a powerful tool for inferring recent linear barriers to gene flow [94,95], their high Type-1 error rates, relatively low power, and difficulty in interpretation require caution when using these methods to test barrier hypotheses. Of interest is that we were able to detect a barrier effect even though on several occasions we witnessed geckos dispersing across the underside of bridges. This suggests that some individuals are crossing streams and suggests that Mantel tests may be the most sensitive to detect a relatively weak barrier effect.

New, sophisticated algorithms achieve difficult genetic clustering. Although early algorithms are entirely aspatial in nature [49], recent applications incorporate prior locality information into the analysis [96–99]. STRUCTURE incorporates geographic information by assigning different codes to different populations [56]. Although relatively underutilized, this method is appealing in cases like ours where geographic information is available for populations and not individuals. Different Bayesian clustering programs often obtain different results [94,95] and this necessitates additional empirical and simulation studies to test the power of the spatial approach in STRUCTURE in comparison to the fully spatial models implemented in other software packages.

To date, few studies examine the population genetic structure of other lineages of geckos. For example, a recent study on two species of sympatric gecko species report significant differences in genetic diversity and structure [100]. Their results suggest different dispersal abilities in sympatric species, with one species exhibiting a maximum dispersal distance of only 500 m. The results of our analysis of *P. tuberculosus* also suggest that many gecko species may have limited dispersal abilities and rely on landscape characteristics to facilitate dispersal.

Landscape genetics: least-cost paths

To date, few studies investigate landscape-genetic relationships in Neotropical vertebrates [18,101]. Our results identify several landscape variables important in shaping the genetic connectivity of leaf-toed geckos. Landscape variables explain significantly more

variation in genetic differentiation than IBD. For example, our qualitative observations of relative abundance in different habitats suggest that forest fragmentation will have detrimental effects on functional connectivity. Several landscape genetic studies also report a negative relationship between forest fragmentation and genetic connectivity in small vertebrates [62,102], but most of these focus on temperate systems. After controlling for Euclidean distance, our least-cost path results show a negative relationship between forest connectivity and genetic differentiation, suggesting that gene flow is higher through forest patches (Table S2). However, we view these results with caution for several reasons. First, models including forest are less well-supported than alternative candidate models. Second, VIF values are exceptionally high for these models and this may be causing large variances in regression coefficients. Third, our study area contains a relatively large amount of undisturbed forest compared to localities in southern Mexico. Thus, Euclidean distances and least-cost path distances based on forest are nearly identical. Additional landscape genetic studies are necessary in areas experiencing rapid loss of TDF in order to understand the effect of forest patch dynamics on functional connectivity.

Our least-cost path analysis suggests that anthropogenic disturbance is influencing functional connectivity in geckos. Because we parameterize the multivariate anthropogenic resistance surface by assigning higher costs to disturbed areas, the significantly negative regression coefficient (after controlling for Euclidean distance) suggests that undisturbed habitat is associated with a lower D_{est} or higher gene flow. These results are concordant with other studies that show a negative relationship between anthropogenic disturbance and rates of gene flow [103–105]. This concordance suggests that although geckos are frequently encountered in close proximity to human settlements, these areas have a detrimental impact on genetic connectivity.

Although geckos are common on abandoned houses in TDF habitat, individuals of *P. tuberculosus* are only present in and around houses in absence of introduced geckos of the genus *Hemidactylus* (pers. obs.). Very few leaf-toed geckos can be found syntopically with *Hemidactylus* and on one occasion we witnessed the head of a *P. tuberculosus* in the jaws of *H. frenatus*. Thus, in heavily anthropogenically influenced areas, it appears that introduced *Hemidactylus* are directly competing with native *Phyllodactylus*. To exacerbate this issue, people often kill leaf-toed geckos on site as they believe the darker colour of these geckos indicates that they are venomous. Conservation efforts should focus on educating local people on differences between native and non-native flora and fauna to aid in the maintenance and protection of the native species. This is especially important for species commonly found close to human settlements.

The minimum temperature of the coldest period influences gene flow, which explains slightly more of the variance in D_{est} than Euclidean distance (0.548 vs. 0.444). TDF is a seasonal forest with approximately eight months of warm, wet conditions and four months of dry, cooler conditions [26]. These geckos commonly occur in hot tropical lowland environments, and less gene flow will occur in localities that experience lower temperatures. Our MRM and MLPE results suggest that these geckos are avoiding areas experiencing colder temperatures. Seasonality and climate are important variables shaping connectivity for other species [106–
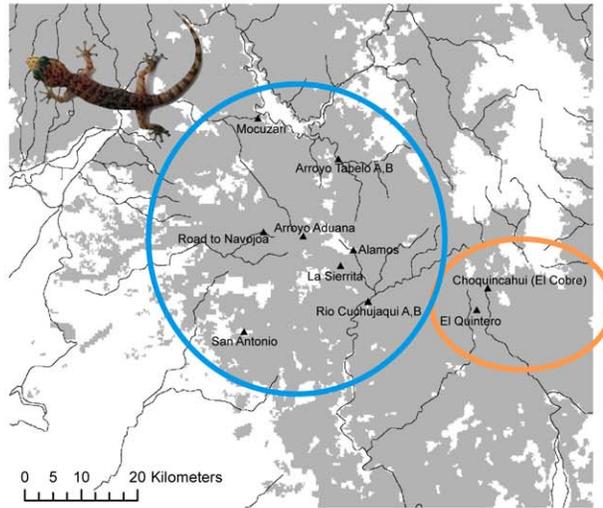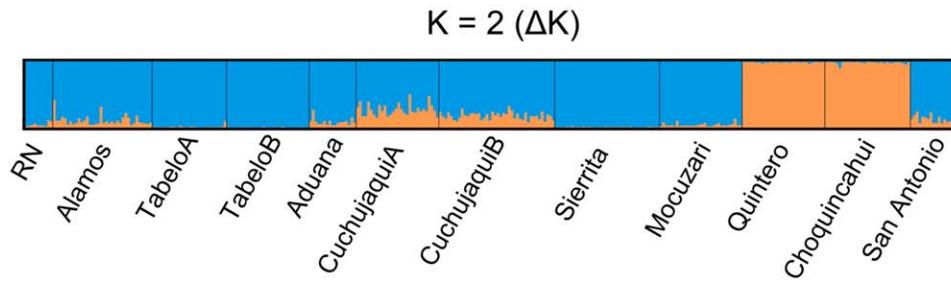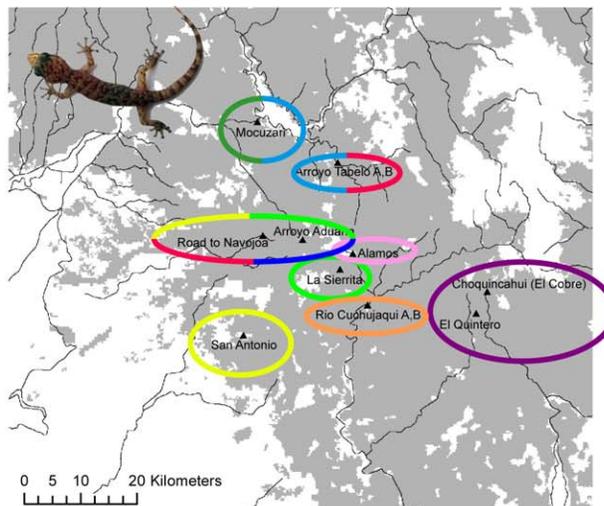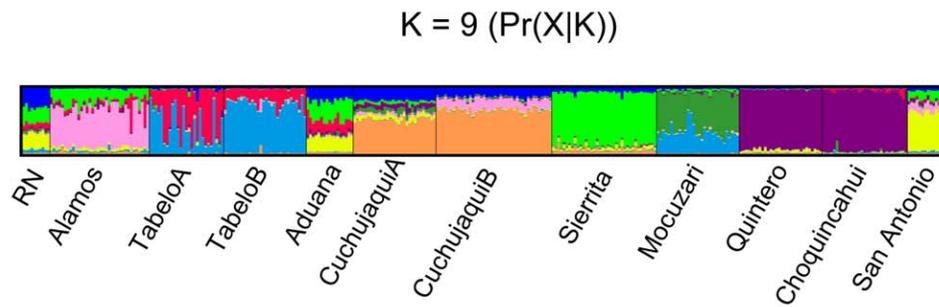
Figure 4. STRUCTURE results including prior information on sampling locality for *Phyllodactylus tuberculosus* included in this study. a) Results for $K = 2$ selected using the ΔK method. b) Results for $K = 9$ selected using ln Pr($X|K$). Each vertical bar represents a single individual with different colors representing the proportion of an individual's genome originating from that specific cluster. Names below plots represent population codes referred to in the text. Colored circles represent the most likely ancestry.
doi:10.1371/journal.pone.0057433.g004

109]. Future studies will identify if this is a common trend in temperature-dependent species and if/how patterns will change with the continual threat of global warming.

Many geckos are from near bridges and rocky outcroppings adjacent to streams and rivers. Thus, we test the hypothesis that riparian networks are an important component shaping patterns of gene flow. Our results suggest that riparian connectivity is an important predictor for patterns of dispersal, as this variable always occurs in our top models. However, these features may also serve as a genetic barrier. Thus, it appears as if geckos may disperse along riparian networks, but seldom cross them. Previous studies conflict as to the role riparian networks play in shaping functional connectivity in small terrestrial vertebrates. For example, streams facilitate gene flow among populations of blotched tiger salamanders (*Ambystoma tigrinum melanosticum*) [19] as they do in Rocky Mountain tailed frogs, (*Ascaphus montanus*) [110], and Pacific jumping mice, (*Zapus trinotatus*) [20]. Conversely, gene flow occurs terrestrially in coastal tailed frogs (*Ascaphus truei*) and does not follow riparian corridors [62]. Combined, these results illustrate the utility of a GIS-based landscape genetic approach to understanding the influence of stream networks on genetic connectivity of small terrestrial vertebrates and reaffirm the necessity for examining species-specific processes [18]. Nevertheless, evidence suggests that riparian corridors should be given conservation priority in many cases.

Landscape genetics: circuit theory

Our results based on resistance-distances derived from a circuit theoretic approach are similar to those based on least-cost paths. However, slope becomes a more important predictor of genetic variation in the former models, where populations separated by higher slopes experience lower rates of gene flow (Table 4). These results are similar to numerous other studies that show a direct relationship between topological relief, elevation, and slope on rates and patterns of gene flow in terrestrial vertebrates [19,111–113]. Riparian connectivity and anthropogenic disturbance also

appear as important components based on resistance-distances. Similar to the least-cost models, we find relatively weak evidence for an influence of forest structure on genetic connectivity. However, when forest is included in a model with Euclidean distance its coefficient is negative, suggesting that intact forest may facilitate gene flow.

Circuit theoretic approaches complement least-cost path modeling and often explain more of the variance in genetic differentiation than more traditional methods [66,67]. For example, slope becomes an important predictor of genetic differentiation in our resistance models. This makes sense intuitively and biologically because geckos are unlikely to disperse along one narrow strip of optimal slope. Thus, the application of circuit theoretic approaches is particularly attractive when a single optimal dispersal route is unlikely. The relative utility of least-cost versus resistance distances will likely depend on the scale of the study and the specific landscape feature in question. However, testing for congruence with both approaches will result in more robust conclusions regarding the influence of specific landscape features. Our model with the largest AICc weight (0.626) and R^2_β (0.673) based on resistance distances explains similar variation in D_{est} as our least-cost models. However, unlike the least-cost analysis, we can combine all resistance distances into our models due to the lack of collinearity among the predictors. Our results corroborate previous findings and suggest that a combination of circuit theoretic and least-cost models provides a powerful tool for investigating functional connectivity in dynamic landscapes.

Statistics and landscape genetics

Landscape genetics is still a relatively new discipline [16] and a large number of recent studies focus on testing the power of various analytical techniques for understanding the influence of landscape variables on microevolutionary processes [37,38,94,95,114,115]. Although the Mantel and partial Mantel tests continue to be the most widely used methods to link landscape and genetic data [18], recent research suggests that

Table 3. Multiple regression on distance matrices (MRM) and maximum likelihood population effects (MLPE) results showing the relationship between pairwise genetic distance (linearized D_{est}) and least-cost path cost distances incorporating landscape heterogeneity.

Model	Variables	β	P	Model R^2	P	VIF	Model AICc	ΔAICc	Akaike Weight (w_i)	β (MLPE)	P (MLPE)	R^2_β	P_{RE}
A	Euclidean	2.75E-05	0.0014	0.621	0.0001	**20.26**	−118.13	0.00	0.856	1.89E-05	<2E-16	0.517	0.0200
	Anthropogenic	−1.72E-06	0.0077			4.20				−1.12E-06	<2E-16		
	Stream	−6.77E-06	0.1152			**13.64**				−4.11E-06	0.0090		
B	Euclidean	1.62E-05	0.0002	0.573	0.0001	3.86	−112.60	5.53	0.054	1.16E-05	<2E-16	0.467	0.0100
	Anthropogenic	−1.46E-06	0.0426			3.86				−8.42E-07	<2E-16		
C	Temperature	1.99E-06	0.0052	0.566	0.0001	5.39	−111.53	6.60	0.032	1.09E-06	<2E-16	0.454	0.0020
	Stream	−2.65E-06	0.3308			5.39				2.01E-07	0.0300		
D	Temperature	1.44E-06	0.0001	0.548	0.0001		−111.06	7.07	0.025	1.14E-06	<2E-16	0.446	0.0050

Candidate models tested were based on a priori hypotheses and to minimize collinearity among predictors. For clarity, only models with relatively high support based on ΔAICc and w_i are shown (i.e. confidence set of candidate models; [74]). Optimal cost values used to parameterize resistance surfaces prior to calculating each least-cost path were selected based on Mantel r correlation coefficients. VIF = Variance Inflation Factor. P_{RE} represents P-value for population effect.
doi:10.1371/journal.pone.0057433.t003

Table 4. Multiple regression on distance matrices (MRM) and maximum likelihood population effects (MLPE) results showing the relationship between pairwise genetic distance (linearized D_{est}) and resistance distances incorporating landscape heterogeneity.

Model	Variables	β	P	Model R²	P	VIF	Model AICc	ΔAICc	Akaike Weight (w_i)	β (MLPE)	P (MLPE)	R²β	P_RE
A	Euclidean	7.81E-06	0.0005	0.626	0.0001	1.65	−118.94	0	0.545	4.95E-06	<2E-16	0.615	3.00E-04
	Slope	9.00E-02	0.0070			1.77				1.19E-01	<2E-16		
	Stream	−1.95E-01	0.0734			2.11				−1.28E-01	0.0010		
B	Euclidean	8.47E-06	0.0006	0.654	0.0003	2.18	−116.57	2.37	0.167	4.75E-06	<2E-16	0.673	9.00E-05
	Anthropogenic	−2.03E-02	0.4354			3.31				7.61E-03	0.1200		
	Stream	−2.02E-01	0.1782			3.33				−8.56E-02	0.0700		
	Forest	8.64E-02	0.7342			4.62				−1.04E-01	0.9700		
	Slope	4.74E-02	0.4369			6.48				1.47E-01	0.0400		
	Temperature	2.75E-02	0.5779			8.30				−1.19E-02	0.3200		
C	Euclidean	7.90E-06	0.0006	0.626	0.0002	1.92	−116.56	2.38	0.166	5.20E-06	<2E-16	0.624	2.00E-04
	Forest	−1.83E-02	0.9152			2.52				−5.99E-02	0.3000		
	Stream	−1.87E-01	0.1851			3.21				−8.95E-02	0.1000		
	Slope	8.92E-02	0.0111			1.88				1.17E-01	<2E-16		
D	Euclidean	8.32E-06	0.0003	0.606	0.0001	2.06	−115.53	3.41	0.099	5.11E-06	<2E-16	0.568	1.00E-05
	Slope	5.05E-02	0.0670			1.47				8.60E-02	<2E-16		
	Anthropogenic	−2.53E-02	0.1106			1.54				−9.21E-03	2.00E-04		

Candidate models tested were based on a priori hypotheses. For clarity, only models with relatively high support based on ΔAICc and w_i are shown (i.e. confidence set of candidate models; [74]). Optimal cost values used to parameterize resistance surfaces prior to calculating resistance distances were selected based on Mantel r correlation coefficients. VIF = Variance Inflation Factor. P_{RE} represents P-value for population effect.

these methods suffer from low power and high Type-1 errors [37,116]. Recognizing these limitations, recent studies have concluded that partial Mantel tests implemented in a causal modeling framework are a powerful tool [38,114]. However, landscape-genetic relationships are often multivariate and are best represented in models that simultaneously consider multiple landscape and environmental variables [8]. Thus, a MRM approach serves as a powerful method to understanding the complex suite of factors important in shaping the spatial distribution of genetic variation [37]. Our study highlights the value of applying MRM analyses to both least-cost and resistance distances. Surprisingly, few studies use this analytical method [71].

Although powerful, MRM approaches have limitations that need to be addressed [69,70]. Collinearity often occurs among independent variables because they are in the form of distances. Often, this will manifest itself in least-cost analyses and its severity depends on landscape structure and the chosen distance metric [37]. Multicollinearity will not affect predictions of the variance of the dependent variable, but it may have a consequence on individual regression coefficients because of higher standard errors. Regression coefficients for predictors can change drastically depending on what other predictors are included in the model [69]. Researchers implementing MRM should examine the influence of multicollinearity on model results and how regression coefficients change with different models. Rigorous model-selection criteria, such as information theoretic or stepwise regression methods, can identify the best combination of explanatory variables.

Although multicollinearity occurs in our least-cost path data, coefficients change little with different explanatory models (Table 3; Table S2). Thus, we are confident in our conclusions regarding landscape-genetic relationships in this system. In cases where coefficients change drastically between models, VIF values will identify the degree of correlation among the predictors. We

recommend that future landscape genetic studies using methods such as MRM and Mantel tests report VIF for each model examined. We also encourage the exploration of ridge regression techniques for landscape-genetic inference.

It remains unclear how the pairwise nature of distance matrices can influence model-selection using information theoretric metrics such as AIC, which generally assume independent observations. Other recent studies acknowledge this potential issue and propose alternatives through the use of Delaunay triangulation [102] or linear mixed models [78]. Unfortunately, the former drastically simplifies landscape heterogeneity while model selection in linear mixed models brings another component of statistical uncertainty. Further, both the sign and magnitude of coefficients may differ depending on the method of analysis used (Table 3,4). The relative performance of these methods for landscape-genetic inference requires evaluation using spatially explicit simulations before strict recommendations can be made.

Conservation implications

Habitat fragmentation and extirpation continue to threaten tropical ecosystems throughout the globe [86,117]. Fragmentation of TDF in Mexico is of particular concern as these forests form the predominant vegetation-type and are known to be a biodiversity hotspot [27,30,31]. A time-series analysis of Mexican TDF reported that by 1990 only 27% of intact forest remained due to unabated anthropogenic conversion for agriculture and pasture-land [28]. The TDF near Alamos constitutes one of the most undistributed tracts of continuous forest in Mexico due, in part, to federally protected reserves [27]. However, even in areas of relatively high forest cover, slight anthropogenic disturbance may have detrimental impacts to functional connectivity. Our study highlights the need for additional landscape genetic studies focusing on TDF ecosystems to better understand how habitat fragmentation and climatic change will influence ecological and
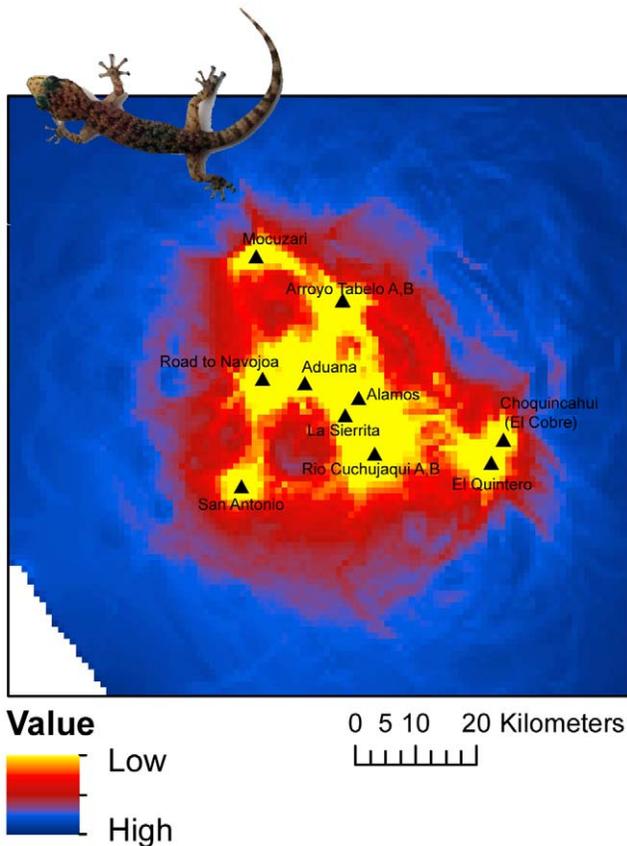
Figure 5. Results from CIRCUITSCAPE **depicting cumulative resistance between populations of** *Phyllodactylus tuberculosus* **based on the combined effect of Euclidean distance, slope and riparian connectivity.** 'Low' in figure legend represents low levels of resistance whereas 'High' represents high levels of resistance.
doi:10.1371/journal.pone.0057433.g005

evolutionary processes. To this end, researchers should focus on developing geospatial data sets at finer spatial resolutions. This will allow a far more comprehensive examination of the effect of landscape-level processes on the spatial distribution of genetic variation. Analyses incorporating high-resolution landscape layers, highly polymorphic genetic markers, and sophisticated analytical techniques will allow the design of movement corridors to maximize functional connectivity for species inhabiting this threatened ecosystem.

Supporting Information

Figure S1   a) STRUCTURE **results illustrating changes in ln Pr(X|K) under the aspatial model.** b) STRUCTURE results

based on the second order rate of change (ΔK method) under the aspatial model. For each K, 10 independent simulations were performed.
(TIF)

Figure S2   a) STRUCTURE **results illustrating changes in ln Pr(X|K) under the spatial model.** b) STRUCTURE results based on the second order rate of change (ΔK method) under the spatial model. For each K, 10 independent simulations were performed.
(TIF)

Table S1   Genetic diversity statistics per locus and population for *Phyllodactylus tuberculosus* **sampled throughout the Alamos, Sonora region.**
(DOCX)

Table S2   Multiple regression on distance matrices (MRM) results showing the relationship between pairwise genetic distance (linearized D_{est}**) and least-cost path cost distances incorporating landscape heterogeneity.** Candidate models tested were based on a priori hypotheses and to minimize collinearity among predictors. Optimal cost values used to parameterize resistance surfaces prior to calculating each least-cost path were selected based on Mantel r correlation coefficients. VIF = Variance Inflation Factor.
(DOCX)

Table S3   Multiple regression on distance matrices (MRM) results showing the relationship between pairwise genetic distance (linearized D_{est}**) and resistance distances incorporating landscape heterogeneity.** Candidate models tested were based on a priori hypotheses. Optimal cost values used to parameterize resistance surfaces prior to calculating resistance distances were selected based on Mantel r correlation coefficients. VIF = Variance Inflation Factor.
(DOCX)

Author Contributions

Conceived and designed the experiments: CB VHJ FRM RWM. Performed the experiments: CB VHJ. Analyzed the data: CB. Contributed reagents/materials/analysis tools: RWM. Wrote the paper: CB RWM.

References

1. Clobert J, Danchin E, Dhondt AA, Nichols JD (2001) Dispersal. Oxford University Press, New York. Pp. 452.
2. Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236: 787–792.
3. Bohonak AJ (1999) Dispersal, gene flow, and population structure. Q Rev Biol 74: 21–45.
4. Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. New York, Cambridge University Press, 617 p.
5. Vilá C, Sundqvist AK, Flagstad O, Seddon J, Björnerfeldt S, et al. (2002) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. Proc R Soc Lond B 270: 91–97.
6. Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am Nat 142: 911–927.
7. Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC (2011) Broken barriers: human-induced changes to gene flow and introgression in animals. BioEssays 33: 508–518.
8. Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Mol Ecol 19: 3576–3591.
9. With KA, Garder RH, Turner MG (1997) Landscape connectivity and population distributions in heterogeneous environments. Oikos 78: 151–169.
10. Ricketts TH (2001) The matrix matters: effective isolation in fragmented landscapes. Am Nat 158: 87–99.

11. Bowne DR, Bowers MA (2004) Interpatch movements in spatially structured populations: a literature review. Land Ecol 19: 1–20.

12. Fahrig L (1997) Relative effects of habitat loss and fragmentation on population extinction. J Wildl Manage 61: 603–610.

13. Fahrig L (1998) When does fragmentation of breeding habitat affect population survival? Ecol Model 105: 273–292.

14. Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34: 487–515.

15. Fahrig L (2007) Non-optimal animal movement in human-altered landscapes. Funct Ecol 21: 1003–1015.

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

17. Holderegger R, Wagner HH (2008) Landscape genetics. BioScience 58: 199–207.

18. Storfer A, Murphy MA, Spear SP, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? Mol Ecol 19: 3496–3514.

19. Spear SF, Peterson CR, Matocq M, Storfer A (2005) Landscape genetics of the blotched tiger salamander (Ambystoma tigrinum melanostictum). Mol Ecol 14: 2553–2564.

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

21. Wang IJ, Savage WK, Shaffer BH (2009) Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (Ambystoma californiense). Mol Ecol 18: 1365–1374.

22. Neel MC (2008) Patch connectivity and genetic diversity conservation in the federally endangered and narrowly endemic plant species Astragalus albens (Fabaceae). Biol Conserv 141: 938–955.

23. Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. J Appl Ecol 44: 714–724.

24. Braunisch V, Segelbacher G, Hirzel AH (2010) Modelling functional landscape connectivity from genetic population structure: a new spatially explicit approach. Mol Ecol 19: 3664–3678.

25. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 409: 853–858.

26. Murphy PG, Lugo AE (1986) Ecology of tropical dry forest. Ann Rev Ecol Syst 17: 67–88.

27. Robichaux RH, Yetman DA (2000) The tropical deciduous forest of Alamos. Biodiversity of a threatened ecosystem in Mexico.The University of Arizona Press.

28. Trejo I, Dirzo R (2000) Deforestation of seasonally dry tropical forest: a national and local analysis in Mexico. Biol Conserv 94: 133–142.

29. Sigala-Rodríguez JJ, Greene HW (2009) Landscape change and conservation priorities: Mexican herpetofaunal perspectives at local and regional scales. Rev Mex Biodivers 80: 231–240.

30. Becerra JX (2005) Timing the origin and expansion of the Mexican tropical dry forest. Proc Natl Acad Sci USA 31: 10919–10923.

31. Ceballos G, García A (1995) Conserving neotropical biodiversity: the role of dry forests in western Mexico. Conserv Biol 9: 1349–1356.

32. Burgos A, Maass MJ (2004) Vegetation change associated with land-use in tropical dry forest areas of Western Mexico. Agric Ecosyst Environ 104: 475–481.

33. Wilson LD, Townsend JH, Johnson JD (2010) Conservation of Mesoamerican Amphibians and Reptiles. Eagle Mountain Publishing, LC, Eagle Mountain, Utah, 812 pp.

34. García A (2006) Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. Biol Conserv 130: 25–46.

35. Dixon JR (1964) The systematics and distribution of lizards of the genus Phyllodactylus in North and Central America. New Mexico State University Research Center, University Park, New Mexico, Scientific Bulletin 64-1: 1–139.

36. Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, et al. (2007) Putting the 'landscape' in landscape genetics. Heredity 98: 128–142.

37. Balkenhol N, Waits LP, Dezzani RJ (2009) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32: 818–830.

38. Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. Am Nat 168: 486–499.

39. Blair C, Jiménez Arco VH, Méndez de la Cruz FR, Murphy RW (2012a) Using next-generation DNA sequencing for rapid microsatellite discovery in Mexican leaf-toed geckos (Phyllodactylus tuberculosus). Conserv Genet Resour 4: 807–810.

40. Dieringer D, Schlötterer C (2003) (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Resour 3: 167–169.

41. Miller MP (1997) Tools for Population Genetic Analyses (TFPGA). A Windows program for the analysis of allozyme and molecular population genetic data. Distributed by author.

42. Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86: 248–249.

43. Rousset F (2008) GENEPOP '007: a complete reimplementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8: 103–106.

44. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57: 289–300.

45. Jost L (2008) G_{ST} and its relatives do not measure differentiation. Mol Ecol 17: 4015–4026.

46. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

47. Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. Mol Ecol Resour 10: 556–557.

48. Jost L (2009) D vs. G_{ST}: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). Mol Ecol 18: 2088–2091.

49. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

50. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587.

51. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611–2620.

52. Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361.

53. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

54. Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4: 137–138.

55. Guillot G, Leblois R, Coulon A, Frantz A (2009) Statistical methods in spatial genetics. Mol Ecol 18: 4734–4756.

56. Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with this assistance of sample group information. Mol Ecol Resour 9: 1322–1332.

57. Michels E, Cottenie K, Neys L, De Gelas K, Coppin P, et al. (2001) Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modelling of the effective geographical distance. Mol Ecol 10: 1929–1938.

58. Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, et al. (2003) The application of 'least-cost' modelling as a functional landscape model. Landsc Urban Plan 64: 233–247.

59. Wright S (1943) Isolation by distance. Genetics 28: 139–156.

60. R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

61. Goslee S, Urban D (2008) Dissimilarity-Based Functions for Ecological Analysis. CRAN repository. Available from http://cran.r-project.org/web/packages/ecodist/.

62. Spear SF, Storfer A (2008) Landscape genetic structure of coastal tailed frogs (Ascaphus truei) in protected vs. managed forests. Mol Ecol 17: 4642–4656.

63. Etherington T (2011) Python based GIS tools for landscape genetics: visualising genetic relatedness and measuring landscape connectivity. Methods Ecol Evol 2: 52–55.

64. Rayfield B, Fortin MJ, Fall A (2010) The sensitivity of least-cost habitat graphs to relative cost surface values. Land Ecol 25: 519–532.

65. McRae BH (2006) Isolation by resistance. Evolution 60: 1551–1561.

66. McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. Proc Natl Acad Sci USA 104: 19885–19890.

67. McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology and conservation. Ecology 10: 2712–2724.

68. Shah VB, McRae BH (2008) CIRCUITSCAPE: a tool for landscape ecology. In: Varoquaux G,Vaught T, Millman J(eds). Proceedings of the 7th Python in Science Conference (SciPy 2008), pp. 62–66.

69. Legendre P, Lapointe FJ, Casgrain P (1994) Modeling brain evolution from behaviour: a permutational regression approach. Evolution 48: 1487–1499.

70. Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecol 188: 117–131.

71. Dyer RJ, Nason JD, Garrick RC (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. Mol Ecol 19: 3746–3759.

72. Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst Zool 35: 627–632.

73. Kutner M, Nachsheim C, Netter J (2004) Applied Linear Regression Models. McGraw-Hill, New York, NY.

74. Johnson JB, Omland KS (2004) Model selection in ecology and evolution. Trends Ecol Evol 19: 101–108.

75. Burnham KP, Anderson DR (2002) Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed.Springer-Verlag.

76. Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22: 79–86.

77. Barton K (2009) MuMIn: Multi-Model Inference. R package version 0.12.2/r18, URL http://R-Forge.R-project.org/projects/mumin/.

78. van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. Mol Ecol 21: 4010–4023.

79. Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. J Agric Biol Envir S 7: 361–372.

80. Bates D, Maechler M, Bolker B (2012) lme4: linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4.

81. Kuznetsova A, Brockhoff PB (2012) MixMod: Analysis of Mixed Models. R package version 1.0. http://CRAN.R-project.org/package=MixMod.

82. Edwards LJ, Muller KE, Wolfinger RD, Qaqish BF, Schabenberger (2008) An R^2 statistic for fixed effects in the linear mixed model. Stat Med 27: 6137–6157.

83. Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53: 983–997.

84. Halekoh U, Højsgaard S (2012) pbkrtest: parametric bootstrap and Kenward Roger based methods for mixed model comparison. R package version 0.3-2. http://CRAN.R-project.org/package=pbkrtest.

85. Cascante A, Quesada M, Lobo JJ, Fuchs EA (2002) Effects of dry tropical forest fragmentation on the reproductive success and genetic structure of the tree Samanea saman. Conserv Biol 16: 137–147.

86. Miles L, Newton AC, DeFries RS, Ravilious C, May I, et al. (2006) A global overview of the conservation status of tropical dry forests. J Biogeogr 33: 491–505.

87. Becerra JX, Venable LD (2008) Sources and sinks of diversification and conservation priorities for the Mexican tropical dry forest PLoS One 3: e3436.

88. Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation G_{ST}, D_{EST} and G'_{ST}: how wrong have we been? Mol Ecol 18: 2080–2083.

89. Ryman N, Leimar O (2009) G_{ST} is still a useful measure of genetic differentiation—a comment on Jost's D. Mol Ecol 18: 2084–2087.

90. Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of population differentiation based on G_{ST} and D: forget G_{ST} but not all of statistics!. Mol Ecol 19: 3845–3852.

91. Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59: 1633–1638.

92. Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Mol Ecol Resour 11: 5–18.

93. Whitlock MC (2011) G'_{ST} and D do not replace F_{ST}. Mol Ecol 20: 1083–1091.

94. Safner T, Miller MP, McRae BH, Fortin MJ, Manel S (2011) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. Int J Mol Sci 12: 865–889.

95. Blair C, Weigel DE, Balazik M, Keeley AT, Walker FM, et al. (2012b) A simulation-based evaluation of methods for inferring linear barriers to gene flow. Mol Ecol Resour 12: 822–833.

96. Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. Genetics 170: 1261–1280.

97. Guillot G, Mortier F, Estoup A (2005b) Geneland: a program for landscape genetics. Mol Ecol Notes 5: 712–715.

98. Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7: 747–756.

99. Durand E, Jay F, Gaggiotti OE, François O (2009) Spatial inference of admixture proportions and secondary contact zones. Mol Biol Evol 26: 1963–1973.

100. Hoehn M, Sarre SD, Henle K (2007) The tales of two geckos: does dispersal prevent extinction in recently fragmented populations? Mol Ecol 16: 3299–3312.

101. Sork VL, Waits LP (2010) Contributions of landscape genetics—approaches, insights, and future potential. Mol Ecol 19: 3489–3495.

102. Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. Mol Ecol 19: 3650–3663.

103. Wofford JEB, Gresswell RE, Banks MA (2005) Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. Ecol Appl 15: 628–637.

104. Leclerc E, Mailhot Y, Mingelbier M, Bernatchez L (2008) The landscape genetics of yellow perch (Perca flaviscens) in a large fluvial ecosystem. Mol Ecol 17: 1702–1717.

105. Raeymaekers JAM, Maes GE, Geldof S, Hontis I, Nackaerts K, et al. (2008) Modeling genetic connectivity in sticklebacks as a guideline for river restoration. Evol Appl 1: 475–488.

106. Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. Mol Ecol 13: 2481–2490.

107. Hirao AS, Kudo G (2004) Landscape genetics of alpine snowbed plants: comparisons along geographic and snowmelt gradients. Heredity 93: 290–298.

108. Hirao AS, Kudo G (2008) The effect of segregation of flowering time on fine-scale spatial genetic structure in an alpine-snowbed herb Primula cuneifolia. Heredity 100: 424–430.

109. Banks SC, Piggott MP, Williamson JE, Bove U, Holbrook NJ, et al. (2007) Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecology 88: 3055–3064.

110. Spear SF, Storfer A (2010) Anthropogenic and natural disturbance lead to differing patterns of gene flow in the Rocky Mountain tailed frog, Ascaphus montanus. Biol Conserv 143: 778–786.

111. Funk WC, Blouin MS, Corn PS, Maxell BA, Pilliod DS, et al. (2005) Population structure of Columbia spotted frogs (Rana luteiventris) is strongly affected by the landscape. Mol Ecol 14: 483–496.

112. Pérez-Espona S, Perez-Barberia FJ, McLeod JE, Jiggins CD, Gordon IJ, et al. (2008) Landscape features affect gene flow of Scottish Highland red deer (Cervus elaphus). Mol Ecol 17: 981–996.

113. Murphy M, Dezzani R, Pilliod DS, Storfer A (2010) Landscape genetics of high mountain frog metapopulations. Mol Ecol 19: 3634–3649.

114. Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. Mol Ecol 19: 3592–3602.

115. Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy M, et al. (2010) Quantifying the lag time to detect barriers in landscape genetics. Mol Ecol 19: 4179–4191.

116. Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol Ecol Resour 10: 831–844.

117. Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. Science 277: 494–499.