`landgenreport`: a new R function to simplify landscape genetic analysis using resistance surface layers

BERND GRUBER and AARON T. ADAMACK

Institute for Applied Ecology, University of Canberra, Canberra, Canberra, ACT, 2601 Australia

Abstract

We describe functions recently added to the R package POPGENREPORT that can be used to perform a landscape genetic analysis (LGA) based on landscape resistance surfaces, which aims to detect the effect of landscape features on gene flow. These functions for the first time implement a LGA in a single framework. Although the approach has been shown to be a valuable tool to study gene flow in landscapes, it has not been widely used to date, despite the type of data being widely available. In part, this is likely due to the necessity to use several software packages to perform landscape genetic analyses. To apply LGA functions, two types of data sets are required: a data set with spatially referenced and genotyped individuals, and a resistance layer representing the effect of the landscape. The function outputs three pairwise distance matrices from these data: a genetic distance matrix, a cost distance matrix and a Euclidean distance matrix. Statistical tests are performed to test whether the cost matrix contributes to the understanding of the observed population structure. A full report on the analysis and outputs in the form of plots and tables of all intermediate steps of the LGA is produced. It is possible to customize the LGA to allow for different cost path approaches and measures of genetic distances. The package is written in the R language and is available through the Comprehensive R Archive. Comprehensive tutorials and information on how to install and use the package are provided at the authors' website (www.popgenreport.org).

Keywords: circuit theory, landscape genetics, least-cost path, population genetics—empirical, resistance surface

Received 10 July 2014; revision received 26 January 2015; accepted 28 January 2015

Introduction

Landscape heterogeneity affects a species' ability to move across a landscape to meet its requirements for foraging, migration, dispersal and other movement-dependent processes (Taylor *et al.* 1993; With *et al.* 1997). Increased availability of spatial and genetic data has led to the emergence of the field landscape genetics that provides a framework to study gene flow and population structure (Manel *et al.* 2003; Storfer *et al.* 2007; Holderegger & Wagner 2008). A major aim of landscape genetics is to identify and quantify landscape features that modify the connectivity between individuals or populations within a landscape (Manel *et al.* 2003; Storfer *et al.* 2007). Commonly studied features include roads and rivers that influence connectivity (Spear & Storfer 2008; Koen *et al.* 2011) and forested areas which may enhance the connectivity of woodland species (Cushman *et al.* 2013; Etherington *et al.* 2014; Sarre *et al.* 2014). Using resistance surface models is one of the most promising approaches

for modelling the effect of landscape features on gene flow of populations in a given landscape. For a review on other landscape genetic methods (that use spatial genetic data) to study the effect of landscape characteristics on population structure, we recommend the overview provided by Manel *et al.* (2003). Resistance surfaces can be thought of as hypothesized relationships between landscape features and gene flow (Adriaensen *et al.* 2003; Wang *et al.* 2008; Spear *et al.* 2010). We describe this approach in more detail in the following section where we outline how to apply our functions, but provide a brief summary here. Instead of using Euclidean distance as a measure of the distances between individuals or subpopulations, the LGA approach calculates distances that incorporate the effects of differing permeabilities across the landscape, referred to as cost distances. Cost distances are compared to genetic-based measurements of distances between individuals or populations to evaluate the effects of landscape features on genetic connectivity (Holderegger & Wagner 2008). To control for isolation by Euclidean distance effects, cost distances are 'statistically standardized' by Euclidean distances prior to the comparison. Although the LGA approach is well

Correspondence: Bernd Gruber, Fax: +61 (2) 6201 5305;
E-mail: bernd.gruber@canberra.edu.au

established and has been used successfully to identify the effect of landscape features on gene flow (e.g. Koen *et al.* 2011; Moore *et al.* 2011; Gottelli *et al.* 2012; Munshi-South 2012; Furlan *et al.* 2013; Parks *et al.* 2013; Sarre *et al.* 2014), it is often complicated and tedious to perform because a researcher generally needs to use at least three different software programs: a geographic information system (GIS) to represent the landscape and calculate the cost distances, a program to analyse the genetic data and a statistical program to combine and analyse the outcomes of the two other programs. This process is cumbersome as the approach frequently requires repeated analyses across the three programs resulting in data being repeatedly converted and passed between programs. As a result, explorations of the potential parameter space are often limited owing to the difficulty in automating this approach.

Here we present a new implementation of the LGA using a single function within the R package POPGENREPORT (Adamack & Gruber 2014). The function carries out the analysis and produces an annotated report that includes maps, least-cost paths, tables and statistical analyses in a single output file (portable document format). Additionally, intermediate results can be examined, allowing the user to ensure the analysis is in line with their intentions. For example, users can check least-cost paths plots to ensure that their path results are not being affected by artefacts of the least-cost path calculations. The package comes with a full description of all functions and options and comprehensive tutorials. In line with the previously published R package POPGENREPORT (Adamack & Gruber 2014), the intention of this addition to the package is to simplify the process of analysing population genetic data. Below, we outline the LGA approach, present the various options available for this analysis and discuss potential caveats researchers should be aware of. Please note that we provide only a succinct overview on the different variants that can be used within the approach to support the decision which variant may be used in a specific case. For a comprehensive understanding of the variants, researchers need to refer to the references to the literature provided.

Outline of LGA and description of the `landgenreport` function

Assuming the necessary input data are available and the package has been installed correctly, a complete LGA can be run using a single command (Fig. 1). There are five required arguments to perform a LGA using the `landgenreport` function. The first two arguments are for the two data sets that are required for the analysis (spatial genetic data and a resistance surface layer). The three remaining arguments are options that specify
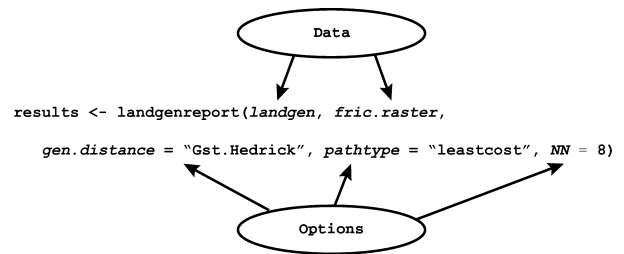

Fig. 1 Exemplary command to run a landscape genetic analysis based on landscape resistance surfaces using the `landgenreport` function.

the approach that will be used to perform the complete LGA. Each of the arguments required by the `landgenreport` function is described in the following sections. We note that there are several additional arguments for the `landgenreport` function that are optional and are described in detail in the help file for the function (`help(landgenreport)`) and the tutorials. The additional arguments are primarily used to control which outputs are produced by the function, their appearance (e.g. symbol type, size and colour), where the outputs will be physically stored, and how they will be named.

Spatial genetic data set

A spatial genetic data set (e.g. microsatellite markers collected from individuals or populations with known spatial coordinates) is required to perform a LGA.

In the `landgenreport` function, the first argument *landgen* identifies the R object which contains the spatial genetic data set (individual genotypes and spatial coordinates) formatted as a genind object in R. A genind object can be created from many standard file formats for genetic data (e.g. FSTAT, STRUCTURE, GENETIX or GENEPOP file formats) using the import functions in the `adegenet` package (Jombart 2008) or from a simple spreadsheet table using the `read.genetable` function in the `PopGenReport` package. As the function is based on the genind object format, the approach is capable of handling different ploidies and other marker types such as SNPs or allozymes. The ADEGENET package offers several functions to import SNP type data formats (such as FASTA or PLINK files) into R objects (genlight objects). A beta version of a function to convert SNP data (genlight objects) into genind objects can be obtained from the authors on request. Spatial coordinates are required for each sample within the genind object and the coordinates need to be in the same coordinate system as the resistance surface layer.

Resistance surface layer(s)

The landscape component of the analysis is included as a resistance surface layer. The resistance surface layer is a raster image of the landscape with each raster cell representing the difficulty for an individual or population to traverse it. Higher resistance values represent species specific factors that limit the ability of an individual to traverse a cell. The ability of an individual or population to traverse a cell could be a function of the mortality risk, the energetic costs, the food availability or any of a number of other factors associated with traversing a cell and the interactions between each of the factors (resistance surfaces) used. A number of approaches exist for inferring the magnitude of the resistance values for a resistance surface layer. The most commonly used approaches are using expert opinion or testing a wide range of resistance values. In some cases, data from other sources (e.g. telemetry or movement studies) are used. For a recent review of resistance value estimation, see Zeller *et al.* (2012).

The argument *fric*`.raster` identifies the R object that contains the resistance surface layer. A resistance surface layer can be loaded into R using the functions in the `raster` package (Hijmans 2014). These functions are capable of loading a wide variety of raster file types including images (BMP, JPEG or TIFF), and GIS produced raster files such as GEOTIFFS, ESRI (ASC) or IMAGINE formats (IMG). As resistance surface layers can be interpreted as simple images, it is in principle possible to simply sketch a resistance surface layer in any image creation program (e.g. MICROSOFT PAINT, GIMP) and use the colour codes as resistance values. This approach may be useful for performing preliminary testing of hypotheses and to get familiar with the package. However, we do not recommend taking this approach as actual maps based on spatial data sets are often readily available from providers such as Google Maps or government agencies, eliminating the need of using a specialist to create a resistance surface layer. We provide an example in our tutorial on how to use maps provided by Google to create resistance surface layers within R, although we note that the preferred option is to use GIS software such as ARCGIS (commercial), GRASS (open source) or QGIS (open source) to create resistance surface layers as they allow for accurate and fine-tuned manipulations of spatial landscape features.

To facilitate the process of testing several resistance surface layers, our function is capable of processing several resistance surfaces at once using a raster stack. This can simplify the task of testing the effects of different resistance values in a given landscape, as a raster stack could contain several variants of the same landscape using different resistance values for a particular landscape feature. For example, one could test the effect of roads on population structure using increasing values of resistance for this feature. A cost matrix would then be calculated for each variant of the resistance surface layer provided and then tested for its effectiveness in explaining the observed population structure.

Genetic distances

Once the genetic data set has been imported, measures of pairwise genetic distance between each individual or population can be calculated. Pairwise genetic distances between samples can be based on measures of individual relatedness such as number of shared alleles or variants thereof (Smouse & Peakall 1999; Kosman & Leonard 2005) or on measures of subpopulation relatedness such as F_{ST}, G_{ST} and D (Jost 2008; Meirmans & Hedrick 2011). The third argument in the `landgenreport` function (`gen.distance`) specifies the type of genetic distance index to be used for calculating pairwise genetic distances between samples. Depending on the type of distance, the analysis will use individuals if individual pairwise genetic distances (options: '`propShared`', '`smouse`', '`Kosman`') are specified and subpopulations if Jost's D, Hedrick's G'_{ST}, Nei's G_{ST} (options: '`D`', '`Gst.Hedrick`', '`Gst.Nei`') are specified. There has been an ongoing discussion in the literature about which kind of genetic distance should be used under which circumstances. The first decision that has to be made is whether an individual-based distance or a subpopulation distance should be used. This obviously depends on the spatial distribution of the sampled locations and the hypothesis that the researcher had in mind. If there are clearly identified and preferable evenly sampled—in terms of sample size—subpopulations, a subpopulation-based distance might be the preferred option. Alternatively if individuals were sampled fairly evenly across the study area, an individual-based distance might be preferable. In a simulation study on the time lag to identify barriers, Blair *et al.* (2012) found that individual-based distances reacted 'faster' to a newly placed barrier, so the time to detect an effect of a barrier was short compared to an F_{ST}-based distance (G'_{ST}). As the simulation was based on fairly evenly distributed individuals, it is not evident whether the finding holds in a more clustered sampling design. The different versions of individual distance indices (proportion of shared alleles, Smouse and Kosman) basically reflect how much weight is given to individuals that share some alleles at a locus. Based on our experience, all three indices are often very highly correlated and most likely lead to the same conclusion in a LGA. There has been an

ongoing debate in the literature about the usefulness and difference between the current three main differentiation metrics [G_{ST} (= option `Gst.Nei`), G'_{ST} (= option `Gst.Hedrick`) and D (= option `D`)] (Jost 2008; Meirmans & Hedrick 2011). An in-depth discussion of this issue is beyond this manuscript, but a nice summary of the ongoing debate has been recently published by Verity & Nichols (2014). In their publication, they study the appropriateness of the three indices to detect differentiation as defined by Wright's fixation index F_{ST}, which he defined 'as the correlation between uniting gametes that make up a diploid individual' (Verity & Nichols 2014). They show that most of the current debate stems from the fact that there are two different versions of F_{ST} that can follow from the above definition: a mutation-dependent and an mutation-independent version of F_{ST}. They conclude that it is necessary to define beforehand which type of differentiation is studied and which kind of model is assumed for the markers and population under study. For some simple cases, they give specific recommendations; for example, under the Wright's simple island model with no mutation, G_{ST} is a good estimator of differentiation, whereas if the mutation rate is assumed to be high, G_{ST} needs to be supplemented by G'_{ST} or D.

Cost distance methods

There are two main methods currently being used to calculate pairwise cost distances between locations on a resistance surface layer. Finding the route of maximum efficiency is the most commonly used method for calculating the least-cost distance between pairs of locations. The least-cost distance is the route that offers the shortest cost-weighted accumulated distance between pairs of locations (sum of per-pixel resistance values). The second most commonly used method is based on electrical resistance theory and has until now only been implemented in the program Circuitscape (McRae 2006; McRae *et al.* 2008). The basic idea is that in Circuitscape, pairwise distances are calculated treating one location as an electrical source and the other one as an electrical ground. Landscape cells between these two locations are assumed to be connected by electrical resistors using the resistance values of the cost surface. Simulating an electrical current, the effective resistance distance between a pair of locations is then calculated. In contrast to the least-cost distance, the Circuitscape-based distance integrates the cost distance over all possible pathways and therefore takes the effect of alternative routes between locations into account (McRae *et al.* 2008). Although there has not been a rigorous comparison

between the two methods, they are currently being treated as complementary methods, in the sense that least-cost path identifies the single most cost-effective path between two locations, whereas Circuitscape takes into account the cost of alternative routes and the width of potential paths (McRae *et al.* 2008; Spear *et al.* 2010). A third method that thus far has not been explored is the randomized shortest-path distance (Saerens *et al.* 2009). Through the use of an additional parameter, a degree of randomness can be introduced when the cost path is calculated. If no randomness is specified, this method is identical to the least-cost approach. Theoretically, this method should result in cost distances that are intermediate between the two former methods. All three approaches are implemented in the `landgenreport` function, which in turn relies on functions provided by the GDISTANCE package (van Etten 2014) and further details on their calculations can be found in the help sections of this package. There are a number of approaches published in the literature summarized under the name of 'least-cost corridor'. Here, the identified least-cost paths between all locations are used to infer dispersal corridors using additional GIS tools such as kernel and moving windows methods to integrate over all least-cost paths (for an overview on these approaches, see Schwartz *et al.* 2009). These approaches are currently not implemented in our package, but it can be used for the first step, calculating the least-cost paths, which in turn can then be further analysed using additional GIS tools.

Number of neighbouring cells

For all three cost distance methods, it is necessary to specify the connectivity rule between neighbouring raster cells (NN = 4 or NN = 8 neighbouring cells). The most commonly used option is to use eight neighbouring cells in order to reduce directional bias introduced by the orthogonal raster geometry and to allow path directions in 45-degree intervals (Adriaensen *et al.* 2003). The downside of the eight neighbouring cells rule option is that this may lead to the exploitation of 'cracks' in the resistance surface layer (Rothley 2005). 'Cracks' in turn may lead to erroneous identification of shortcuts across truly expensive barriers and therefore lead to vastly distorted cost distances. This can be important, particularly if linear structures are the main features in the landscape being studied as the use of NN = 8 is prone to creating these artefacts. We therefore highly recommend inspecting visually the actual least-cost paths and if necessary, recreating the resistance surface layer by filling in the 'cracks' (Rothley 2005) or alternatively, rerunning the analysis using the four neighbouring cells rule if mainly

linear structures are being examined. However, note that the four neighbouring cells rule has a stronger directional bias in the least-cost paths. An example of a least-cost path plot for two different cost surfaces and comparing the NN = 4 and NN = 8 options can be seen in Fig. 2.

The process of calculating the least-cost path between pairs of sampling locations is repeated for each of the $n \times (n - 1)/2$ possible pairs, resulting in a cost distance matrix.

Tests for detecting the effect of cost distances on population structure

The final step in the LGA is to analyse the relationship between the two distance matrices (the genetic distance and cost distance matrix). A simple approach would be to plot the two distance matrices against each other and perform a linear regression on them. This approach is not recommended due to the nonindependence of the elements within the distance matrices (Legendre *et al.* 1994). Additionally, the effects of isolation by distance should be tested to ensure that simple pairwise Euclidean distances are not better at explaining the observed population structure than the cost matrix. The current practice is to either perform a series of partial Mantel tests (Cushman *et al.* 2006, 2013; Wasserman *et al.* 2010) or, more recently, to perform multiple matrix regression on distance matrices (MMRR) (Wang 2013). The approach suggested by Cushman *et al.* (2006) and refined

by Wasserman *et al.* (2010) uses a series of partial Mantel tests (also called causal modelling). The basic idea is that there are (at least) two competing hypothesis (cost matrices) about the effect of the landscape on population structure, for example one could be a cost matrix representing Euclidean distances ($c1$) and the other one a cost matrix based on the least-cost path algorithm ($c2$). The two competing matrices are tested against each other. If a partial Mantel test of the genetic distances against $c2$ partialling out $c1$ is significant, but a partial Mantel test of the genetic distances against $c1$ partialling out $c2$ is not, then this is evidence that the cost matrix $c2$ shows an association with the genetic distances. The approach can be iterated across all pairs of competing cost matrices. If neither of the two tests follow the described pattern, one would conclude there is no correlation between either cost matrix and the genetic distance matrix.

Wang (2013) implemented the MMRR approach, which was initially suggested by Legendre *et al.* (1994). This approach is similar to Mantel and partial Mantel tests and according to Wang (2013) 'provides easily interpretable output in the form of a multiple regression equation'. The main advantage of his regression approach is that the output is very similar to well-known ANOVA tables, providing a significance test for each cost matrix. Moreover, the MMRR analyses not only asks whether the two distance matrices are significantly correlated, but also 'quantifies how genetic distances respond to changes in geographic and environmental distances (regression coefficients) and calculates the overall fit of the model' (Wang 2013).

MMRR and causal modelling have both been shown to be capable of identifying the cost matrix layer(s) that best explain(s) a population's spatial genetic structure using simulations (Landguth *et al.* 2010; Spear *et al.* 2010) and real-world examples (Spear & Storfer 2008; Wang *et al.* 2009; Spear *et al.* 2010; Gottelli *et al.* 2012). However, both approaches have difficulties in discriminating between competing cost matrices if they are highly correlated and there is currently no preferred solution for this problem. This difficulty in discriminating between highly correlated competing cost matrices is particularly problematic when comparing the same surfaces using different resistance values as these cost matrices are most likely highly correlated. In this case, when using the causal modelling approach, the partial Mantel tests tend to have increased type I error rates, identifying a relationship between the cost and genetic distance matrices even though there truly is none (Cushman *et al.* 2013). Similarly, for the MMRR approach, Wang (2013) recommends exercising caution if the explanatory cost matrices are strongly correlated.

Both methods are implemented in the POPGENREPORT package, but again no rigorous comparisons between
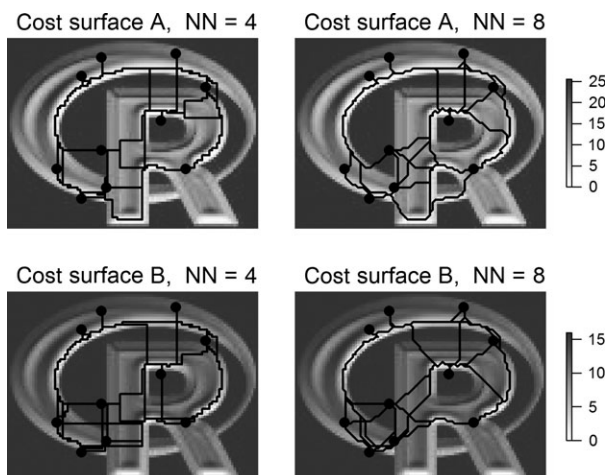

Fig. 2 An example plot of a least-cost path analysis (`path-type = least-cost`) for two different cost surfaces (A and B) and two different neighbouring rules (NN= 4, NN= 8). As can be seen, darker areas (= high resistance values) are avoided to traverse the landscape. This is more prominent in cost surface A (top row), where the resistance values on average are higher compared to cost surface B (bottom row). Please note not all pairwise paths can be seen as they overlap.

methods have been published as of yet. From our experience, both methods perform similarly in detecting the effect of landscape features.

Output and customization of a standard LGA

The results of a LGA using the `landgenreport` function (e.g. all least-cost path plots and pairwise distance matrices) are stored in the specified 'result' object and can be examined further or be used in downstream analyses. Additionally, all plots and tables created as a part of the analysis are stored as individual files (tables as CSV files, plots as PNG, PDF and SVG files) in an output folder. If a LaTeX environment has been installed, a complete report detailing all of the least-cost paths and the intermediate and overall results is provided in a single PDF file (see Supporting Information for an example). To facilitate the installation of the package, we have assembled 'POPGENPACK' for the Windows operating system, which provides a portable version of R that includes all required R packages and a running LaTeX environment. POPGENPACK can be run from a USB stick or from any folder on a computer without the need to install the programs and packages.

It is also possible to perform a customized analysis using this package. Instead of running the complete analysis in a single step, each of the steps of the analysis can be completed separately using the functions: `genleastcost`, `wassermann` and `lgrMMRR`, which are documented in the R package. A stepwise analysis allows the user to specify other measures of genetic distance measures or to use only a subset of the pairwise distances within the matrices in order to avoid spatial autocorrelation (e.g. Etherington *et al.* 2014). Examples of how these analyses can be performed are provided in the tutorials that are provided in the Supporting Information or can be accessed from within R by typing: browseVignettes('PopGenReport'). The latest version of the package can be obtained from the R Cran repositories or via the package website: www.popgenreport.org.

Acknowledgements

References

Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, **5**, 384–387.

Adriaensen F, Chardon JP, De Blust G *et al.* (2003) The application of "least-cost" modelling as a functional landscape model. *Landscape and Urban Planning*, **64**, 233–247.

Blair C, Weigel DE, Balazik M *et al.* (2012) A simulation-based evaluation of methods for inferring linear barriers to gene flow. *Molecular Ecology Resources*, **12**, 822–833.

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

Cushman SA, Wasserman T, Landguth E, Shirk A (2013) Re-evaluating causal modeling with Mantel tests in landscape genetics. *Diversity*, **5**, 51–72.

Etherington TR, Perry GLW, Cowan PE, Clout MN (2014) Quantifying the direct transfer costs of common brushtail possum dispersal using least-cost modelling: a combined cost-surface and accumulated-cost dispersal kernel approach. *PLoS One*, **9**, e88293.

van Etten J (2014) gdistance: distances and routes on geographical grids. R package version 1.1.5. http://cran.r-project.org/package=gdistance.

Furlan EM, Griffiths J, Gust N *et al.* (2013) Dispersal patterns and population structuring among platypuses, Ornithorhynchus anatinus, throughout south-eastern Australia. *Conservation Genetics*, **14**, 837–853.

Gottelli D, Sillero-Zubiri C, Marino J, Funk SM, Wang J (2012) Genetic structure and patterns of gene flow among populations of the endangered Ethiopian wolf (T Katzner, J Johnson, Eds,). *Animal Conservation*, **16**, 234–247.

Hijmans RJ (2014) raster: raster: Geographic data analysis and modeling. R package version 2.2-31. http://cran.r-project.org/package=raster.

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

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Koen EL, Bowman J, Garroway CJ, Mills SC, Wilson PJ (2011) Landscape resistance and American marten gene flow. *Landscape Ecology*, **27**, 29–43.

Kosman E, Leonard KJ (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology*, **14**, 415–424.

Landguth EL, Cushman SA, Schwartz MK *et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, **19**, 4179–4191.

Legendre P, Lapointe F-J, Casgrain P (1994) Modeling brain evolution from behavior: a permutational regression approach. *Evolution*, **48**, 1487–1499.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

McRae BH (2006) Isolation by resistance. *Evolution*, **60**, 1551–1561.

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, **89**, 2712–2724.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Moore JA, Tallmon DA, Nielsen J, Pyare S (2011) Effects of the landscape on boreal toad gene flow: does the pattern-process relationship hold true across distinct landscapes at the northern range margin? *Molecular Ecology*, **20**, 4858–4869.

Munshi-South J (2012) Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leuc-*

opus) populations in New York City. *Molecular Ecology*, **21**, 1360–1378.

Parks SA, McKelvey KS, Schwartz MK (2013) Effects of weighting schemes on the identification of wildlife corridors generated with least-cost methods. *Conservation Biology: The Journal of the Society for Conservation Biology*, **27**, 145–154.

Rothley K (2005) Finding and filling the "Cracks" in resistance surfaces for Least-cost Modeling. *Ecology and Society*, **10**, 4.

Saerens M, Achbany Y, Fouss F, Yen L (2009) Randomized shortest-path problems: two related models. *Neural Computation*, **21**, 2363–2404.

Sarre SD, Aitken N, Adamack AT *et al.* (2014) Creating new evolutionary pathways through bioinvasion: the population genetics of brushtail possums in New Zealand. *Molecular Ecology*, **23**, 3419–3433.

Schwartz MK, Copeland JP, Anderson NJ *et al.* (2009) Wolverine gene flow across a narrow climatic niche. *Ecology*, **90**, 3222–3232.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, **82**, 561–573.

Spear SF, Storfer A (2008) Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology*, **17**, 4642–4656.

Spear SF, Balkenhol N, Fortin M-J, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, **19**, 3576–3591.

Storfer A, Murphy MA, Evans JS *et al.* (2007) Putting the "landscape" in landscape genetics. *Heredity*, **98**, 128–142.

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. *Oikos*, **68**, 571–573.

Verity R, Nichols RA (2014) What is genetic differentiation, and how should we measure it—G_{ST}, D, neither or both? *Molecular Ecology.*, **23**, 4216–4225.

Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**, 3403–3411.

Wang Y-H, Yang K-C, Bridgman CL, Lin L-K (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, **23**, 989–1000.

Wang IJ, Savage WK, Bradley Shaffer H (2009) Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). *Molecular Ecology*, **18**, 1365–1374.

Wasserman TN, Cushman SA, Schwartz MK, Wallin DO (2010) Spatial scaling and multi-model inference in landscape genetics: *Martes americana* in northern Idaho. *Landscape Ecology*, **25**, 1601–1612.

With KA, Gardner RH, Turner MG, With A (1997) Landscape connectivity and population distributions in heterogeneous environments. *Oikos*, **78**, 151–169.

Zeller KA, McGarigal K, Whiteley AR (2012) Estimating landscape resistance to movement: a review. *Landscape Ecology*, **27**, 777–797.

Data Accessibility

All data are publicly available. Data files are part of the `PopGenReport` package on CRAN (http://cran.r-project.org/web/packages/PopGenReport).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 An example of a report produced via the `landgenreport` function.

Appendix S2 Tutorial on how to install and use the package POP-GENREPORT.

Appendix S3 Tutorial on how to do a landscape genetic analysis based on landscape resistance surfaces using the `landgenreport` function.