

# Conventional methods for enhancing connectivity in conservation planning do not always maintain gene flow

Jeffrey O. Hanson<sup>1</sup>  | Richard A. Fuller<sup>1</sup>  | Jonathan R. Rhodes<sup>2</sup> 

<sup>1</sup>School of Biological Sciences, The University of Queensland, Brisbane, Queensland, Australia

<sup>2</sup>School of Earth and Environmental Sciences, The University of Queensland, Brisbane, Queensland, Australia

## Correspondence

Jeffrey O. Hanson  
Email: jeffrey.hanson@uqconnect.edu.au

## Funding information

Australian Government Research Training Program

Handling Editor: Akira Mori

## Abstract

- 1 Protected area systems need to conserve species in places with suitable habitat that are configured to facilitate gene flow. Since genetic data require considerable resources to obtain, many proxy methods have been developed to generate plans for protected area systems (prioritizations) that facilitate gene flow without needing genetic data. However, the effectiveness of such methods—such as minimising fragmentation or enforcing contiguity among priority areas—remains largely untested.
- 2 We investigated the ability of prioritizations to maintain gene flow when they are generated using conventional methods for promoting connectivity. Using existing environmental, genetic, and occurrence datasets, we created maps of habitat suitability and resistance to gene flow for nine alpine plant species. Next, we generated multispecies prioritizations that secured 10% of the suitable habitat for each species and attempted to maintain gene flow by (a) penalizing fragmentation, (b) representing species in contiguous areas of suitable habitat, and (c) representing species in contiguous areas with minimal resistance to gene flow as modelled from genetic data.
- 3 We found that prioritizations generated using fragmentation penalties failed to represent seven of the nine species in areas that would maintain high levels of gene flow. Similarly, prioritizations that represented species in contiguous areas of suitable habitat were unable to maintain high levels of gene flow for six species—potentially because a few areas with high resistance can disrupt gene flow throughout an entire prioritization. Although prioritizations generated using genetic data successfully maintained gene flow, they also selected over three times more land than other prioritizations, suggesting that failing to account for gene flow when setting priorities may underestimate the scale of conservation action required.
- 4 *Synthesis and applications.* We found that conventional methods for enhancing connectivity in conservation planning, such as spatially clustering priority areas or providing connected sections of suitable habitat, were generally unable to maintain high levels of gene flow. Our results suggest that conservation plans could be substantially improved by directly using genetic data, although whether this is a good choice for a particular situation will also depend on the costs of obtaining these data.

**KEYWORDS**

alpine plants, connectivity, gene flow, integer programming, landscape genetics, optimization, spatial prioritization, systematic conservation planning

**1 | INTRODUCTION**

Protected areas are urgently needed to buffer species from anthropogenic threats and provide places for biodiversity processes to continue (Watson, Dudley, Segan, & Hockings, 2014). Since it is rarely feasible to conserve all remaining habitat for a species, protected area systems must encompass a cost-effective subset of places that will enable species to persist indefinitely (Margules & Pressey, 2000). Gene flow plays an important role in species' persistence (Ellstrand, 2014), by helping populations adapt to new conditions (Hoffmann & Sgrò, 2011) and maintaining genetic diversity (Schwartz & Mills, 2005). Although restoring gene flow has risks (Richardson, Brady, Wang, & Spear, 2016), there has been increasing interest in developing plans for protected area systems (termed "prioritisations") and wildlife corridors to foster gene flow (Creech, Epps, Monello, & Wehausen, 2014; Dilkina et al., 2017; Schoville et al., 2018; Wang & Önal, 2016).

Over the last decade, the field of landscape genetics has made enormous progress towards understanding and mapping the spatial patterns of gene flow (Manel & Holderegger, 2013). Resistance maps—broadly speaking, heat maps where greater values indicate areas with greater impediments to gene flow—are commonly used to describe how different areas affect gene flow for a particular species (e.g., Burkhart et al., 2017; Dudaniec et al., 2016). Today, advanced methods use genetic data, models, and optimization algorithms (e.g., Peterman, 2018) to transform maps describing the distribution of landscape features (e.g., land cover classes or environmental conditions; Dudaniec et al., 2016) into resistance maps (but see Cushman & Landguth, 2010) which, in turn, can be used to predict the effects of habitat modification on gene flow (Ruiz-Lopez et al., 2016). Although resistance maps are often used to identify wildlife corridors (e.g., Dilkina et al., 2017), they are rarely used to guide the configuration of protected area systems (Keller, Holderegger, van Strien, & Bolliger, 2015). This is, in part, because genetic data remain expensive to obtain at the scale required for planning protected area systems.

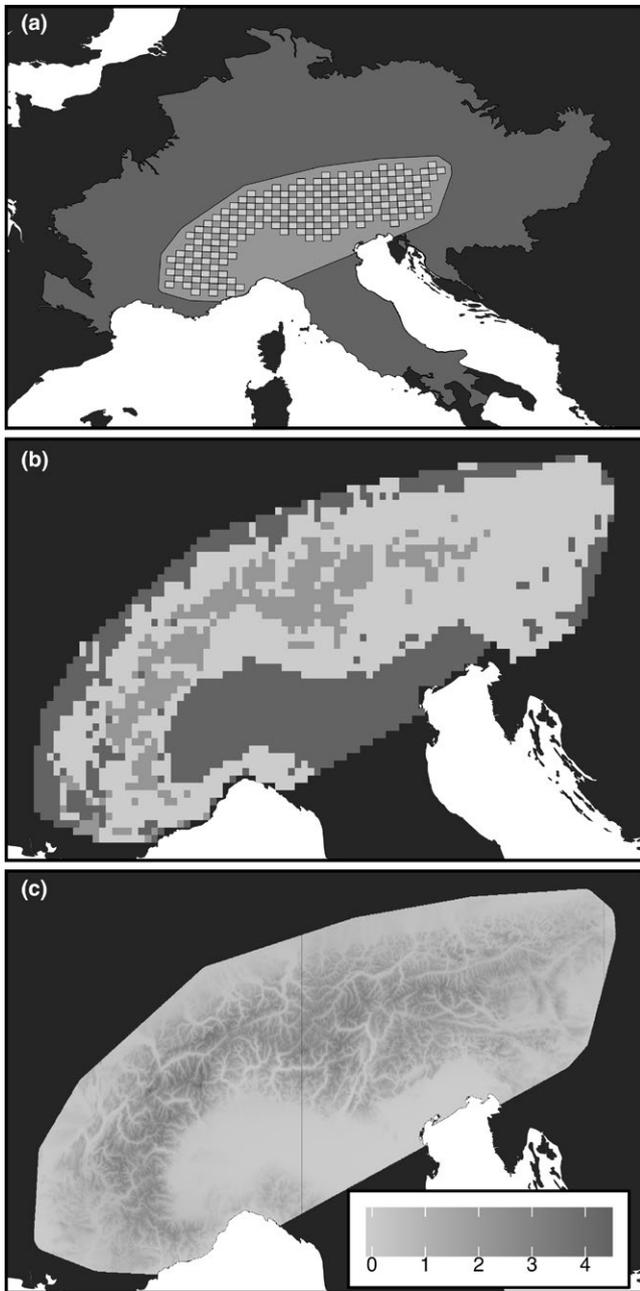
Many decision support tools for conservation planning aim to identify solutions that promote connectivity and enhance gene flow using proxies (Ball, Possingham, & Watts, 2009; Hanson, Schuster, et al., 2018; Moilanen & Wintle, 2007). Conventional approaches for increasing connectivity have involved minimizing fragmentation among priority areas (Ball et al., 2009; Nicholls & Margules, 1993). More advanced approaches—based on the idea that areas which contain more suitable habitat also facilitate greater levels of gene flow (Keeley, Beier, & Gagnon, 2016; Mateo-Sánchez et al., 2015a; Milanesi, Holderegger, Caniglia, Fabbri, & Randi, 2016; Wang, Yang, Bridgman, & Lin, 2008)—aim to represent species in connected areas

that contain highly suitable habitat (Krosby et al., 2015; Moilanen & Wintle, 2007; Wang & Önal, 2016). However, despite their widespread usage, the ability of these proxy-based approaches to identify prioritizations that actually maintain gene flow among protected habitats remains poorly tested.

Here, we show that prioritizations based on spatial measures of fragmentation, or habitat suitability alone, may not be able to maintain high levels of gene flow. Using occurrence and genetic data for nine plant species in the European Alps (Taberlet et al., 2012a, 2012b), we generated multispecies prioritizations that were constrained to conserve an adequate amount of habitat for each species in (a) areas that exhibited varying levels of spatial fragmentation, (b) contiguous areas that contained habitat above a specific quality, and (c) contiguous areas that facilitate specific levels of gene flow using resistance maps. Our results suggest that protected area systems will be far more likely to enhance species' evolutionary processes if they are designed using genetic data.

**2 | MATERIALS AND METHODS****2.1 | Data and study area**

We used data collected by the IntraBioDiv project (Taberlet et al., 2012a, 2012b) for nine plant species in the European Alps (*Androsace obtusifolia*, *Arabis alpina*, *Campanula barbata*, *Gypsophila repens*, *Hornungia alpina*, *Phyteuma betonicifolium*, *Phyteuma hemisphaericum*, *Rhododendron ferrugineum*, and *Saxifraga stellaris*). Briefly, project members collected genetic data for 27 species using a 20' longitude by 12' latitude grid (c. 25 × 22 km; Figure 1). They visited every second grid cell, and collected samples from approximately three individuals from each of the 27 species present in the cell (Figure 1; elevation data obtained from Jarvis, Reuter, Nelson, & Guevara, 2008). Samples were genotyped using amplified fragment length polymorphisms (AFLP; Vos et al., 1995), and used to construct matrices denoting the presence/absence of polymorphisms at different loci (mean 151.58 ± 47.8 SD markers genotyped per species; for more information see Gugerli et al., 2008). We used this dataset because it contained comparable genetic data for a diverse group of species (Meirmans, Goudet, IntraBioDiv Consortium, & Gaggiotti, 2011). Of the 27 sampled species, we used only nine species for our analysis because we excluded (a) species that primarily use wind for dispersal (Meirmans et al., 2011; Thiel-Egenter et al., 2009) as contemporary landscape genetics methods are ill-suited for asymmetric gene flow (e.g., Peterman, 2018), (b) species with insufficient occurrence data for habitat modelling (i.e., fewer than 100 occurrence records remaining after data cleaning; *Cerastium uniflorum* and *Loiseleuria procumbens*), and (c) species for which we were unable to



**FIGURE 1** Maps show (a) the area used for modelling species' distributions (dark grey), the study area (grey), and the sampling grid used by the IntraBioDiv project (light grey); (b) forest (light grey), human modified land (dark grey), and other land cover types (grey) within the study area; and (c) elevation (km) within the study area

generate resistance maps that adequately explained their genetic variation ( $R^2 \beta$  statistic less than 0.35; *Gentiana nivalis*; see below for more information). All analyses were performed using the R statistical environment (version 3.4.4; R Core Team, 2017).

We defined our study area using the grid defined by the IntraBioDiv project (buffered by 20 km; c. 26.8 million ha in size; Figure 1). We obtained data on five predictor variables to make spatial predictions of habitat suitability and landscape resistance to gene flow (Supporting Information Figure S1). These variables included

soil alkalinity (5 km resolution; Land Resources Management Unit, Institute for Environment and Sustainability, European Commission, Joint Research Centre, 2010), available growing conditions (i.e., number of days in the year with temperatures over 5°C; 2.5' resolution; Title & Bemmels, 2018), annual precipitation (BIO12), and precipitation seasonality (BIO15; 30'' resolution; <http://worldclim.org>; Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) to account for potential habitat requirements (Alvarez et al., 2009; Meineri, Skarpaas, & Vandvik, 2012). Additionally, we obtained land cover data (500 m resolution corresponding to 2015; NASA LP DAAC, 2015) and reclassified it into three categorical classes representing: forest (1–5), human modified land (12–14), and other cover classes (6–11, 15–16; Figure 1b). All data were reprojected to an equal-area coordinate system (Lambert Azimuthal Equal Area Europe; EPSG:3035), and then aggregated and resampled to a 10 × 10 km grid for subsequent analysis (using the `sp` and `RASTER` R packages; Hijmans, 2016; Pebesma & Bivand, 2005). We then used these maps for generating habitat suitability maps and resistance maps. Note that we used the same predictor maps for generating habitat suitability maps and resistance maps to ensure that differences between them did not arise due to differences in data quality.

## 2.2 | Habitat suitability maps

We fitted environmental niche models to generate the habitat suitability maps for each species. To increase prediction accuracy inside the study area, we used a much larger extent than the study area to train the models (c. 108.9 million ha in size; Figure 1; defined by merging the study area with overlapping and neighbouring World Wildlife Fund ecoregions; Olson et al., 2001). Although the IntraBioDiv dataset contains locality data, we did not use this data for generating habitat maps so that the maps would be of comparable quality to those commonly used in prioritizations, and more independent of the genetic data. Thus, we obtained 53,671 occurrence records for the nine study species from the Global Biodiversity Information Facility (using the `RGBIF` R package; Chamberlain, 2017).

We cleaned the occurrence records prior to model fitting. Records were removed if they were duplicates, low in precision (i.e., coordinates were integers), not collected through direct human observations of wild individuals (e.g., data for captive specimens were omitted), entered with invalid or suspicious coordinates, or tagged with an incorrect country. This resulted in a dataset containing on average 525 records per species  $\pm 348.93$  SD (Supporting Information Table S1). Additionally, to control for sampling bias during model fitting, we extracted the total number of plant records in the Global Biodiversity Information Facility database over a 50 × 50 km grid, excluding records associated with the previously discussed issues, and used them to generate a bias map (Supporting Information Figure S2).

We fitted `MAXENT` models (version 3.4.0; Phillips, Dudík, & Schapire, 2017) for each species using their presence records and 10,000 background records generated from the bias map. The models were trained using 10-fold cross-validation and fitted using

the five predictor variables. They were tuned using various feature class combinations (“L”, “LQ”, “H”, “LQH”, “LQHP”, and “LQHPT”; see Phillips et al., 2017 for explanation) and regularization penalties (0.1, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, and 4; using the ENMEVAL R package; Muscarella et al., 2014). To avoid over-fitting, the best model for each species was identified by calculating the average area under the receiver operating characteristic curve (AUC) statistic using the data held out from each fold. In other words, we selected models that had the greatest ability to predict presence and pseudo-absence data not used to train them. To assess the importance of different predictor variables for each species’ habitat suitability map, the percent contribution of each variable was calculated for the best model (Supporting Information Table S2). The continuous suitability maps (Supporting Information Figure S3) were then thresholded by maximizing the sum of the sensitivity and specificity statistics (Supporting Information Figure S4; Liu, Berry, Dawson, & Pearson, 2005). The thresholded habitat maps adequately described the species’ geographic distributions (mean 0.88 AUC; ranging from 0.76 to 0.94; Supporting Information Table S1).

## 2.3 | Resistance maps

We generated resistance maps for each species using the genetic and location data collected by the IntraBioDiv project (Supporting Information Figure S5). Since the dataset contained replicate samples per sampling grid cell, we grouped together samples in the same cell and calculated pairwise Nei’s (1978) distances between them to obtain measures of genetic dissimilarity for each species that described how different the individuals in each cell were compared to other cells (using the POPPR R package; Kamvar, Tabima, & Grünwald, 2014). Although such metrics can be affected by biases (Whitlock & McCauley, 1999), they are routinely used to model gene flow (e.g., Mateo-Sánchez et al., 2015b; Ruiz-Lopez et al., 2016). After calculating the genetic distances, they were used to generate resistance maps.

Briefly, the resistance maps were generated using genetic algorithms (via the RESISTANCEGA R package; Peterman, 2018). These algorithms were run using a population size of 250 and terminated when 50 consecutive iterations showed no improvement or until 500,000 iterations had elapsed. In a given iteration, the algorithm (a) generated coefficient values; (b) applied the coefficient values to rescale the predictor variable maps using Ricker functions; (c) summed the rescaled predictor maps together to form a candidate resistance map (if multiple predictor maps were input); (d) calculated resistance distances between each pair of cells using the candidate resistance map (using the GDISTANCE R package; van Etten, 2017); (e) fitted a maximum likelihood population effects model (MLPE; Clarke, Rothery, & Raybould, 2002) correlating the genetic distances between cells to their resistance distances; and (f) used the model’s log-likelihood statistic to evaluate the coefficients and the candidate resistance map. Resistance distances were calculated using commute distances because they are computationally efficient and can accommodate multiple pathways (similar to Circuitscape; van Etten, 2017).

Since it was not computationally feasible to use cross-validation to limit over-fitting, we used genetic algorithms to generate five preliminary single-surface resistance maps for each species to identify potential variables for generating higher quality multisurface resistance maps (Supporting Information Table S3; similar to Burkhart et al., 2017). Specifically, the MLPE models were refitted using restricted maximum likelihood,  $R^2\beta$  statistics were calculated (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008; using the r2GLMM R package; Jaeger, 2016), and variables were identified which adequately explained genetic variation for each species ( $R^2\beta \geq 0.35$ ). If multiple variables were identified for a given species, then the variables were inputted to the genetic algorithms to generate a multisurface resistance map for that species. The multisurface resistance maps and best performing single-surface maps for species without multisurface resistance maps were then collated for subsequent analysis (hereafter referred to as resistance maps). To validate the resistance maps, we calculated  $R^2\beta$  statistics and confirmed that they explained an adequate proportion of the sampled genetic variation for each species (mean  $0.63 R^2\beta \pm 0.13 SD$ ; Supporting Information Table S4). We also verified that these statistics are robust to uncertainty using bootstrap analyses (10,000 replicates with 80% of the data; Supporting Information Table S4; following Dudaniec et al., 2016). As mentioned earlier, the analysis originally included one additional species (*G. nivalis*) but it was omitted because its resistance maps had poor predictive abilities ( $R^2\beta \leq 0.35$ ; following Keller et al., 2015). To assess the importance of different variables, the percent contribution of each variable was calculated for each species’ resistance map (Supporting Information Table S5).

## 2.4 | Prioritizations

We generated 2,649 planning units inside the study area using a  $10 \times 10$  km grid (i.e., the same grid used for the habitat and resistance maps). Next, we overlaid the planning units with the habitat and resistance maps, and associated each planning unit with the following data for each species: a continuous value for resistance to gene flow, a continuous value for habitat suitability, and a binary value indicating the presence or absence of suitable habitat. Using the planning units and their corresponding data, we then generated nine multispecies prioritizations to examine the performance of different approaches for maintaining gene flow. These prioritizations all shared the same primary objective—minimize the number of selected planning units in the solution—and contained constraints to secure 10% of each species’ distribution of suitable habitat (using the binary habitat suitability data). They differed in their approach to promote gene flow.

The first three prioritizations were generated using additional penalties to punish fragmented solutions. To understand how fragmentation influences performance, they were generated using different penalty factors (low: 0, medium:  $4 \times 10^{-11}$ , and high:  $2 \times 10^{-4}$ ). These penalties were formulated using the mathematical definitions used by the Marxan decision support tool to exemplify conventional approaches (Ball et al., 2009; Beyer, Dujardin, Watts, & Possingham, 2016).

The next three prioritizations were generated using additional constraints to represent each species in a spatially contiguous area

containing habitat with continuous suitability values above a certain threshold (see Supporting Information Appendix S1 for mathematical formulation that builds on Beyer et al., 2016; Cerdeira, Pinto, Cabeza, & Gaston, 2010; Önal & Briers, 2006). To explore how different thresholds influenced performance, we generated the prioritizations using different thresholds (low: 0th, medium: 75th, and high: 100th percentiles; Supporting Information Figure S6). Since different species had different patterns of fragmentation in their habitat suitability maps, these thresholds were expressed along a continuum where the 0th percentile corresponded to all of the suitable habitat for each species and the 100th percentile corresponded to the highest possible threshold of the continuous suitability data which resulted in a contiguous area containing 10% of the suitable habitat for each species (defined using the binary habitat data).

The final three prioritizations were generated using additional constraints to represent each species in a spatially contiguous area containing resistance values below a specified threshold (constraints detailed in Appendix S1). For convenience, we expressed the resistance values as conductance values (i.e., the reciprocal value). Thus, planning units with higher conductance values have a greater capacity to facilitate gene flow. Similar to the scaling procedure described for the continuous habitat data, the conductance thresholds were expressed along a continuum where the 0th percentile corresponded to all of the conductance data inside areas with suitable habitat and the 100th percentile corresponded to the highest possible threshold for the continuous conductance data which resulted in a contiguous area containing 10% of the suitable habitat for each species (defined using the binary suitability data). Given this scaling procedure, the thresholds were set as the 60th (low), 80th (medium), and 100th (high) percentiles of the continuous conductance data for each species (see Supporting Information Figure S7 for maps).

All prioritizations were solved to within 10% of optimality using GUROBI (version 8.0.1; Gurobi Optimization Inc., 2018) and the PRIORITIZER R package (Hanson et al., 2018b). Thus, the level of fragmentation in species' distributions did not affect solution quality. After generating the prioritizations, we assessed their ability to promote gene flow by calculating the maximum amount of contiguous habitat for each species given various conductance thresholds (using the IGRAPH R package; Csardi & Nepusz, 2006).

### 3 | RESULTS

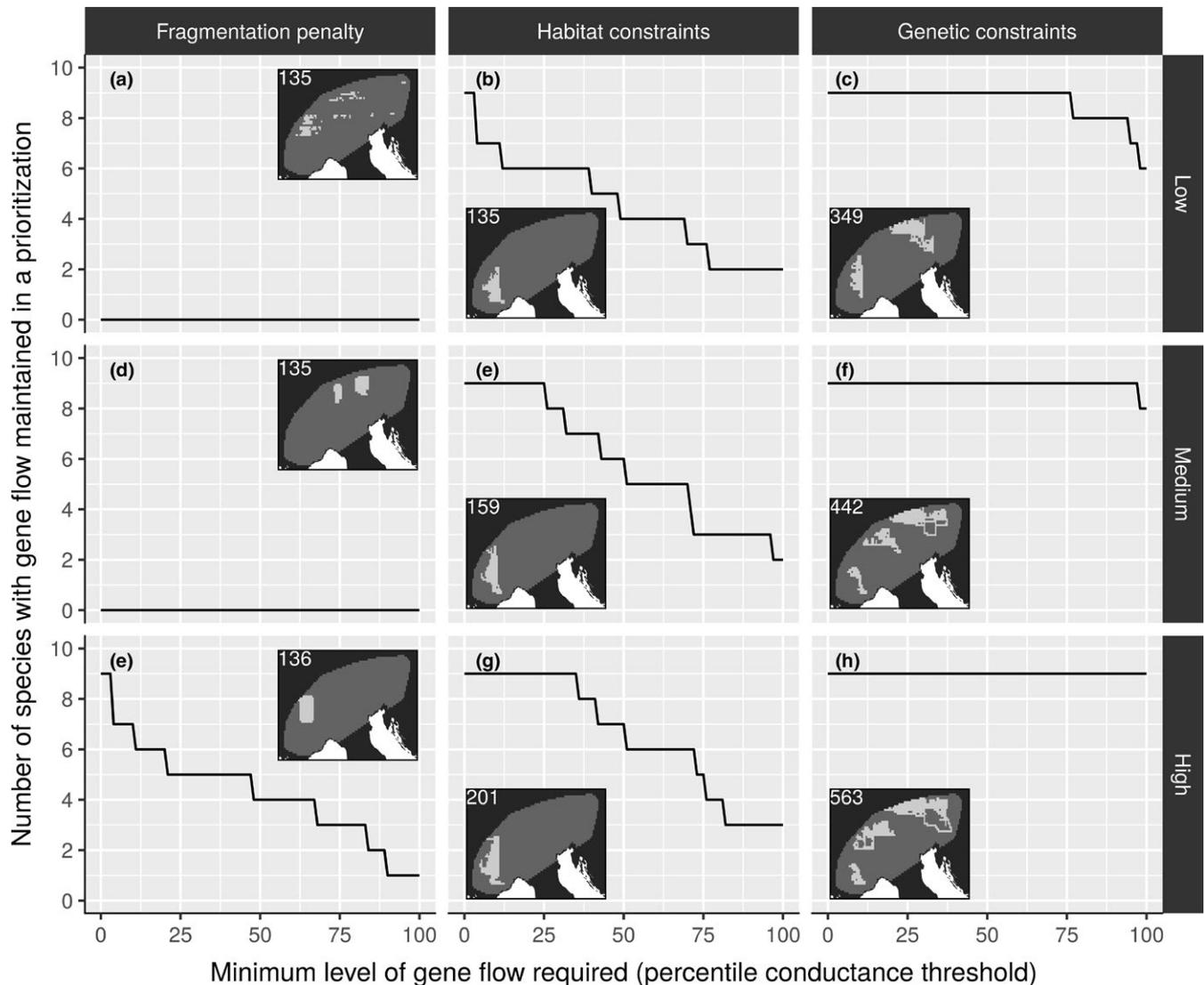
We found that none of the prioritizations generated using fragmentation penalties selected planning units in a configuration that would maintain high levels of gene flow for most species (i.e., minimum level of gene flow required greater than the 85th percentile; Figure 2a,d,g). The prioritization generated without any fragmentation penalties was extremely fragmented (Figure 2a), and so it was not able to represent any species in a contiguous area that contained 10% of its habitat in the study area. Similarly, the prioritization generated with medium fragmentation penalties contained two

reserves (Figure 2d) and was not able to represent any species in a contiguous area that met its habitat target. Next, the prioritization generated with high fragmentation penalties contained a single reserve (Figure 2g). As a consequence, this prioritization represented all species in a contiguous section of land that contained 10% of their suitable habitat. It selected planning units in a configuration that would maintain moderate levels of gene flow for four species (minimum level of gene flow required greater than the 50th percentile). These results provide evidence that fragmentation penalties can help prioritizations maintain gene flow for some species. However, they also suggest that simply clustering priority areas is not likely to facilitate high levels of gene flow because spatially clustered areas may not necessarily contain habitat that can facilitate high levels of gene flow.

None of the prioritizations generated to represent species in contiguous areas of suitable habitat were able to maintain moderate levels of gene flow for all nine of the studied species (minimum level of gene flow required greater than 50th percentile; Figure 2b,e,h). The prioritization generated to represent species in contiguous areas containing marginally suitable habitat (Figure 2b) was able to maintain moderate levels of gene flow for four species, and high levels of gene flow for two species (minimum level of gene flow required greater than 85th percentile). Similarly, the prioritizations generated to represent species in contiguous areas containing moderately suitable habitat (Figure 2e) and extremely suitable habitat (Figure 2h) were only able to maintain high levels of gene flow for three species. These results show that, for the species studied here, representing populations in contiguous areas of suitable habitat is not an effective strategy for maintaining high levels of gene flow—regardless of the suitability of the habitat located in these contiguous areas.

We detected significant negative correlations between the spatial patterns of habitat suitability and resistance to gene flow for all nine of the studied species (one-sided Kendall's rank correlation tests; Supporting Information Table S6). Thus, areas that contained larger amounts of suitable habitat also tended to offer a smaller impediment to gene flow. This result was expected because we selected these species because they exhibit dispersal mechanisms which predispose them towards showing this relationship. Since the prioritizations generated to represent species in contiguous areas of suitable habitat (Figures 2b,e,h) were generally unable to maintain high levels of gene flow (minimum level of gene flow required greater than 85th percentile), these results suggest that the correlation between habitat suitability and resistance to gene flow—despite being statistically significant—may not necessarily be strong enough to operationalize for conservation planning given the strict contiguity requirements for gene flow.

We confirmed that generating prioritizations to represent species in contiguous areas that offer low resistance to gene flow actually resulted in prioritizations that were able to maintain gene flow (Figure 2c,f,i). All three of the prioritizations generated to represent species in areas with prespecified levels of resistance to gene flow were able to maintain moderate levels of gene flow for all nine of the studied species (minimum level of gene flow required greater than



**FIGURE 2** The ability of different prioritizations to maintain gene flow. Each panel corresponds to a different prioritization, and shows the number of species that are represented in a contiguous area that maintains gene flow and contains 10% of their habitat as a function of the minimum level of gene flow required. Columns show prioritizations generated using different methods, and rows show prioritizations generated using different parameters. Map insets show the planning units (dark grey), prioritizations (light grey), and number of selected planning units

50th percentile). Additionally, the prioritizations generated to represent species in areas that contained medium and high conductance to gene flow (Figures 2f,i) were able to maintain high levels of gene flow for all nine species (minimum level of gene flow required greater than 85th percentile). Furthermore, all of the prioritizations generated using gene flow contiguity constraints contained far more selected planning units than the prioritizations generated using other approaches, suggesting that failing to account for gene flow during reserve selection may underestimate the scale of conservation actions required.

#### 4 | DISCUSSION

We aimed to understand whether commonly used approaches for enhancing connectivity in prioritizations can facilitate gene flow. By

evaluating the performance of prioritizations generated using fragmentation penalties and habitat contiguity constraints according to measurements derived from genetic data, we found that these approaches can produce prioritizations that do not maintain high levels of gene flow. This occurred despite the fact that areas which contained more suitable habitat tended to offer lower resistance to gene flow. Our results add to the growing body of evidence that integrating genetic data into conservation can improve decision-making (Creech et al., 2014; Hendry et al., 2010; Schoville et al., 2018; Taberlet et al., 2012a). Fortunately, the rapidly decreasing costs for obtaining genomic data using Next Generation Sequencing mean that this will become less challenging.

We found that prioritizations needed to be guided using genetic data to maintain high levels of gene flow for all nine of the study species. These results mean that commonly used proxy approaches for

increasing genetic connectivity, such as penalizing spatial fragmentation (Ball et al., 2009; Nicholls & Margules, 1993) or representing species in contiguous areas of suitable habitat (Krosby et al., 2015; Saura & Pascual-Hortal, 2007) may not be as effective as previously thought. Although setting higher representation targets (e.g., 70%) and spatially clustering priority areas might yield prioritizations that maintain higher levels of gene flow among a subset of the priority areas (see Supporting Information Appendix S2 and Figure S8), this approach may not be reliable or efficient given the large number of selected planning units required. Since earlier studies have identified effective surrogates for genetic variation (Hanson, Rhodes, Riginos, & Fuller, 2017; Ponce-Reyes, Clegg, Carvalho, McDonald-Madden, & Possingham, 2014), it would seem that conservation planners may be able to use surrogates for some applications of genetic data but not others. One explanation could be that small changes to a prioritization have a much larger influence on its ability to maintain gene flow compared to its ability to represent genetic variation. For instance, consider a prioritization that contains priority areas arranged in a chain. Here, the removal of a single priority area from the middle of the chain would approximately halve the maximum amount of habitat represented in a contiguous set of priority areas, but may only slightly reduce the level of represented genetic diversity if the remaining priority areas still capture multiple genetic types. Thus, conservation planners may need to account for gene flow using genetic data, especially where protected areas are being designated to help populations adapt to changing conditions (Hoffmann & Sgrò, 2011).

Habitat suitability was significantly negatively correlated with landscape resistance to gene flow for all nine of the studied plant species (Supporting Information Table S6). Although these relationships were not strong enough to operationalize for conservation planning, the fact that these relationships exist at least partially validates some of the key assumptions used to enhance connectivity in prioritizations (e.g., Krosby et al., 2015; Moilanen & Wintle, 2007) and supports previous studies (Keeley et al., 2016; Milanese et al., 2016; Wang et al., 2008). Furthermore, since these species were selected because they exhibit traits that make associations between habitat suitability and gene flow more likely, this result means that genetic data may be especially useful when conserving species whose gene flow is driven by other factors. For instance, genetic data may be informative for plant species which have different habitat preferences to their pollinators (e.g., Arroyo, Primack, & Armesto, 1982) or rely on wind for dispersal. Perhaps one reason for why habitat suitability was a poor surrogate for landscape resistance to gene flow—despite being significantly correlated—is that they are each driven by different landscape features.

The landscape features which had the largest contribution to habitat suitability generally made little contribution to resistance to gene flow (Supporting Information Tables S2 and S5). For all nine species, habitat suitability was largely influenced by available growing conditions and annual precipitation. On the other hand, landscape resistance to gene flow was mostly determined by soil alkalinity for three species, and different climatic variables for the

remaining six species. These results corroborate previous work which found that climatic conditions can influence habitat suitability and gene flow in alpine plant species (Hirao & Kudo, 2004; Meineri et al., 2012). Additionally, soil substrate conditions have played a large role in shaping the recolonization patterns of these species following the Last Glacial Maximum and so it is not surprising that these conditions may still be affecting their evolution today (Alvarez et al., 2009).

This study is associated with several limitations. Firstly, we used larger scaled planning units than commonly used in conservation planning (e.g., Moilanen & Wintle, 2007). Although we could have interpolated the data to a finer resolution, this would have introduced artificial spatial autocorrelation and biased our analysis. Secondly, we used nine alpine plant species, and although this constitutes more species than most landscape genetics studies (Manel & Holderegger, 2013), further work is needed to verify these findings in other taxa and biomes. Thirdly, the prioritization approaches for representing species in contiguous areas required thresholding continuous data. Although this means that practitioners will need to determine appropriate thresholds, ideally linked to persistence (Keller et al., 2015), such thresholds could provide a transparent and effective method for identifying solutions that exhibit a minimum level of performance similar to representation targets (Carwardine, Klein, Wilson, Pressey, & Possingham, 2009). Fourthly, although future advances in exact algorithm solvers may resolve this issue, the prioritization approaches used here for representing species in contiguous areas will currently not scale to hundreds of species or hundreds of thousands of planning units. Fifthly, future studies could examine the cost effectiveness of genetic data using value of information analyses. Sixthly, since genetic data can fail to reflect contemporary dispersal (Manel & Holderegger, 2013), conservation planners should also aim to incorporate movement data. Finally, to successfully employ the methods outlined here, conservation planners will require high-quality resistance maps based on robust sampling schemes, carefully selected genetic markers, and careful evaluation of modelling assumptions (Richardson et al., 2016; Whitlock & McCauley, 1999).

Current approaches for promoting connectivity in prioritizations often involve spatially clustering priority areas or representing species in contiguous areas containing suitable habitat. Here, we show that such approaches may not necessarily secure species in areas that facilitate high levels of gene flow. Although further studies are needed to verify our findings in other systems, in the meantime, we urge conservation planners to carefully consider using genetic data to maintain gene flow in prioritizations—especially where the costs for land acquisition eclipse those for obtaining genetic data.

## ACKNOWLEDGEMENTS

J.O.H. was supported by an Australian Government Research Training Program (RTP) Scholarship. R.A.F. had an Australian Research Council Future Fellowship. This work was supported by a research allocation from the National eResearch Collaboration Tools and Resources (NeCTAR) project. We are grateful to members of

the IntraBioDiv project for collecting the genetic data and making it freely available.

## AUTHORS' CONTRIBUTIONS

J.O.H., J.R.R., and R.A.F. conceived the study and designed the methodology; J.O.H. analysed the data; J.O.H. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## DATA ACCESSIBILITY

Code and data, with exception to the "Map of Soil pH in Europe" dataset (Land Resources Management Unit, Institute for Environment and Sustainability, European Commission, Joint Research Centre, 2010), available via the Zenodo Digital Repository <https://doi.org/10.5281/zenodo.1255387> (Hanson, Fuller, Rhodes, 2018a). The "Map of Soil pH in Europe" is a third party dataset and can be obtained from <http://esdac.jrc.ec.europa.eu>.

## ORCID

Jeffrey O. Hanson  <https://orcid.org/0000-0002-4716-6134>

Richard A. Fuller  <http://orcid.org/0000-0001-9468-9678>

Jonathan R. Rhodes  <http://orcid.org/0000-0001-6746-7412>

## REFERENCES

- Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönswetter, P., ... IntraBioDiv Consortium. (2009). History or ecology? Substrate type as a major driver of spatial genetic structure in alpine plants. *Ecology Letters*, 12, 632–640. <https://doi.org/10.1111/j.1461-0248.2009.01312.x>
- Arroyo, M. T. K., Primack, R., & Armesto, J. (1982). Community studies in pollination ecology in the high temperate Andes of central Chile. I. Pollination mechanisms and altitudinal variation. *American Journal of Botany*, 69, 82–97. <https://doi.org/10.1002/j.1537-2197.1982.tb13237.x>
- Ball, I., Possingham, H. P., & Watts, M. E. (2009). Marxan and relatives: Software for spatial conservation prioritisation. In A. Moilanen, K. A. Wilson, & H. P. Possingham (Eds.), *Spatial conservation prioritisation: Quantitative methods and computational tools* (pp. 185–189). Oxford, UK: Oxford University Press.
- Beyer, H. L., Dujardin, Y., Watts, M. E., & Possingham, H. P. (2016). Solving conservation planning problems with integer linear programming. *Ecological Modelling*, 328, 14–22. <https://doi.org/10.1016/j.ecolmodel.2016.02.005>
- Burkhardt, J. J., Peterman, W. E., Brocato, E. R., Romine, K. M., Willis, M. M. S., Ousterhout, B. H., ... Eggert, L. S. (2017). The influence of breeding phenology on the genetic structure of four pond-breeding salamanders. *Ecology and Evolution*, 7, 4670–4681. <https://doi.org/10.1002/ece3.3060>
- Carwardine, J., Klein, C. J., Wilson, K. A., Pressey, R. L., & Possingham, H. P. (2009). Hitting the target and missing the point: Target-based conservation planning in context. *Conservation Letters*, 2, 4–11. <https://doi.org/10.1111/j.1755-263X.2008.00042.x>
- Cerdeira, J. O., Pinto, L. S., Cabeza, M., & Gaston, K. J. (2010). Species specific connectivity in reserve-network design using graphs. *Biological Conservation*, 143, 408–415. <https://doi.org/10.1016/j.biocon.2009.11.005>
- Chamberlain, S. (2017). *rgbif: Interface to the global biodiversity information facility API*. R package version 0.9.7. Retrieved from <https://CRAN.R-project.org/package=rgbif>
- Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, 7, 361. <https://doi.org/10.1198/108571102320>
- Creech, T. G., Epps, C. W., Monello, R. J., & Wehausen, J. D. (2014). Using network theory to prioritize management in a desert bighorn sheep metapopulation. *Landscape Ecology*, 29, 605–619. <https://doi.org/10.1007/s10980-014-0016-0>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695, 1–9.
- Cushman, S. A., & Landguth, E. L. (2010). Spurious correlations and inference in landscape genetics. *Molecular Ecology*, 19, 3592–3602. <https://doi.org/10.1111/j.1365-294X.2010.04656.x>
- Dilkina, B., Houtman, R., Gomes, C. P., Montgomery, C. A., McKelvey, K. S., Kendall, K., ... Schwartz, M. K. (2017). Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. *Conservation Biology*, 31, 192–202. <https://doi.org/10.1111/cobi.12814>
- Dudaniec, R. Y., Worthington Wilmer, J., Hanson, J. O., Warren, M., Bell, S., & Rhodes, J. R. (2016). Dealing with uncertainty in landscape genetic resistance models: A case of three co-occurring marsupials. *Molecular Ecology*, 25, 470–486. <https://doi.org/10.1111/mec.13482>
- Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An  $R^2$  statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, 27, 6137–6157. <https://doi.org/10.1002/sim.3429>
- Ellstrand, N. C. (2014). Is gene flow the most important evolutionary force in plants? *American Journal of Botany*, 101, 737–753. <https://doi.org/10.3732/ajb.1400024>
- Gugerli, F., Englisch, T., Niklfeld, H., Tribsch, A., Mirek, Z., Ronikier, M., ... IntraBioDiv Consortium. (2008). Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation—A project synopsis. *Perspectives in Plant Ecology, Evolution and Systematics*, 10, 259–281. <https://doi.org/10.1016/j.ppees.2008.07.001>
- Gurobi Optimization Inc. (2018). Gurobi optimizer reference manual. Retrieved from <http://www.gurobi.com>
- Hanson, J. O., Fuller, R. A., & Rhodes, J. R. (2018a). Code and data from: Conventional methods for enhancing connectivity in conservation planning do not always maintain gene flow. *Zenodo Digital Repository*, <https://doi.org/10.5281/zenodo.1255387>
- Hanson, J. O., Rhodes, J. R., Riginos, C., & Fuller, R. A. (2017). Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 12755–12760. <https://doi.org/10.1073/pnas.1711009114>
- Hanson, J. O., Schuster, R., Morrell, N., Strimas-Mackey, M., Watts, M. E., Arcese, P., ... Possingham, H. P. (2018b). *prioritizr: Systematic conservation prioritization in R*. R package version 4.0.0.9. <https://github.com/prioritizr/prioritizr>
- Hendry, A. P., Lohmann, L. G., Conti, E., Cracraft, J., Crandall, K. A., Faith, D. P., ... Donoghue, M. J. (2010). Evolutionary biology in biodiversity science, conservation, and policy: A call to action. *Evolution*, 64, 1517–1528. <https://doi.org/10.1111/j.1558-5646.2010.00947.x>
- Hijmans, R. J. (2016). *raster: Geographic data analysis and modeling*. R package version 2.5-8. <https://CRAN.R-project.org/package=raster>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978. <https://doi.org/10.1002/joc.1276>

- Hirao, A. S., & Kudo, G. (2004). Landscape genetics of alpine-snowbed plants: Comparisons along geographic and snowmelt gradients. *Heredity*, 93, 290–298. <https://doi.org/10.1038/sj.hdy.6800503>
- Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, 470, 479–485. <https://doi.org/10.1038/nature09670>
- Jaeger, B. (2016). r2glmm: Computes R squared for mixed (multilevel) models. R package version 0.1.1. Retrieved from <https://CRAN.R-project.org/package=r2glmm>
- Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). *Hole-filled SRTM for the Globe*. Version 4. Retrieved from <http://srtm.csi.cgiar.org>
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>
- Keeley, A. T. H., Beier, P., & Gagnon, J. W. (2016). Estimating landscape resistance from habitat suitability: Effects of data source and nonlinearities. *Landscape Ecology*, 31, 2151–2162. <https://doi.org/10.1007/s10980-016-0387-5>
- Keller, D., Holderegger, R., van Strien, M. J., & Bolliger, J. (2015). How to make landscape genetics beneficial for conservation management? *Conservation Genetics*, 16, 503–512. <https://doi.org/10.1007/s10592-014-0684-y>
- Krosby, M., Breckheimer, I., John Pierce, D., Singleton, P. H., Hall, S. A., Halupka, K. C., ... Schuett-Hames, J. P. (2015). Focal species and landscape “naturalness” corridor models offer complementary approaches for connectivity conservation planning. *Landscape Ecology*, 30, 2121–2132. <https://doi.org/10.1007/s10980-015-0235-z>
- Land Resources Management Unit, Institute for Environment and Sustainability, European Commission, Joint Research Centre. (2010). Map of Soil pH in Europe. Retrieved from <https://esdac.jrc.ec.europa.eu/content/soil-ph-europe>
- Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385–393. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>
- Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution*, 28, 614–621. <https://doi.org/10.1016/j.tree.2013.05.012>
- Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, 405, 243–253. <https://doi.org/10.1038/35012251>
- Mateo-Sánchez, M. C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A., & Saura, S. (2015a). A comparative framework to infer landscape effects on population genetic structure: Are habitat suitability models effective in explaining gene flow? *Landscape Ecology*, 30, 1405–1420. <https://doi.org/10.1007/s10980-015-0194-4>
- Mateo-Sánchez, M. C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A., & Saura, S. (2015b). Estimating effective landscape distances and movement corridors: Comparison of habitat and genetic data. *Ecosphere*, 6, 1–16. <https://doi.org/10.1890/ES14-00387.1>
- Meineri, E., Skarpaas, O., & Vandvik, V. (2012). Modeling alpine plant distributions at the landscape scale: Do biotic interactions matter? *Ecological Modelling*, 231, 1–10. <https://doi.org/10.1016/j.ecolmodel.2012.01.021>
- Meirmans, P. G., Goudet, J., IntraBioDiv Consortium, & Gaggiotti, O. E. (2011). Ecology and life history affect different aspects of the population structure of 27 high-alpine plants. *Molecular Ecology*, 20, 3144–3155. <https://doi.org/10.1111/j.1365-294X.2011.05164.x>
- Milanesi, P., Holderegger, R., Caniglia, R., Fabbri, E., & Randi, E. (2016). Different habitat suitability models yield different least-cost path distances for landscape genetic analysis. *Basic and Applied Ecology*, 17, 61–71. <https://doi.org/10.1016/j.baae.2015.08.008>
- Moilanen, A., & Wintle, B. A. (2007). The boundary-quality penalty: A quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21, 355–364. <https://doi.org/10.1111/j.1523-1739.2006.00625.x>
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MaxEnt ecological niche models. *Methods in Ecology and Evolution*, 5, 1198–1205. <https://doi.org/10.1111/2041-210X.12261>
- NASA LP DAAC. (2015). MCD12Q1 land cover type yearly L3 global 500 m SIN Grid. Sioux Falls, South Dakota: NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center. Version 051. Retrieved from <https://lpdaac.usgs.gov>
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583–590.
- Nicholls, A. O., & Margules, C. R. (1993). An upgraded reserve selection algorithm. *Biological Conservation*, 64, 165–169. [https://doi.org/10.1016/0006-3207\(93\)90654-J](https://doi.org/10.1016/0006-3207(93)90654-J)
- Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., ... Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, 51, 933–938. [https://doi.org/10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)
- Önal, H., & Briers, R. A. (2006). Optimal selection of a connected reserve network. *Operations Research*, 54, 379–388. <https://doi.org/10.1287/opre.1060.0272>
- Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5, 9–13.
- Peterman, W. E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, 9, 1638–1647. <https://doi.org/10.1111/2041-210X.12984>
- Phillips, S. J., Dudík, M., & Schapire, R. E. (2017). Maxent software for modeling species niches and distributions. Version 3.4.0. Retrieved from [http://biodiversityinformatics.amnh.org/open\\_source/maxent](http://biodiversityinformatics.amnh.org/open_source/maxent)
- Ponce-Reyes, R., Clegg, S. M., Carvalho, S. B., McDonald-Madden, E., & Possingham, H. P. (2014). Geographical surrogates of genetic variation for selecting island populations for conservation. *Diversity and Distributions*, 20, 640–651. <https://doi.org/10.1111/ddi.12195>
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org>
- Richardson, J. L., Brady, S. P., Wang, I. J., & Spear, S. F. (2016). Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology*, 25, 849–863. <https://doi.org/10.1111/mec.13527>
- Ruiz-Lopez, M. J., Barelli, C., Rovero, F., Hodges, K., Roos, C., Peterman, W. E., & Ting, N. (2016). A novel landscape genetic approach demonstrates the effects of human disturbance on the Udzungwa red colobus monkey (*Procolobus gordonorum*). *Heredity*, 116, 167–176. <https://doi.org/10.1038/hdy.2015.82>
- Saura, S., & Pascual-Hortal, L. (2007). A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. *Landscape and Urban Planning*, 83, 91–103. <https://doi.org/10.1016/j.landurbplan.2007.03.005>
- Schoville, S. D., Dalongeville, A., Viennois, G., Gugerli, F., Taberlet, P., Lequette, B., ... Manel, S. (2018). Preserving genetic connectivity in the European Alps protected area network. *Biological Conservation*, 218, 99–109. <https://doi.org/10.1016/j.biocon.2017.12.017>
- Schwartz, M. K., & Mills, L. S. (2005). Gene flow after inbreeding leads to higher survival in deer mice. *Biological Conservation*, 123, 413–420. <https://doi.org/10.1016/j.biocon.2004.11.016>
- Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., ... IntraBioDiv Consortium. (2012a). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters*, 15, 1439–1448. <https://doi.org/10.1111/ele.12004>

- Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., ... IntraBioDiv Consortium. (2012b). Data from: Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.s4q6s>
- Thiel-Egenter, C., Gugerli, F., Alvarez, N., Brodbeck, S., Cieślak, E., & Colli, L., ... IntraBioDiv Consortium. (2009). Effects of species traits on the genetic diversity of high-mountain plants: A multi-species study across the Alps and the Carpathians. *Global Ecology and Biogeography*, *18*, 78–87. <https://doi.org/10.1111/j.1466-8238.2008.00421.x>
- Title, P. O., & Bemmels, J. B. (2018). ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, *41*, 291–307. <https://doi.org/10.1111/ecog.02880>
- van Etten, J. (2017). R package gdistance: Distances and routes on geographical grids. *Journal of Statistical Software*, *76*, 1–21. <http://doi.org/10.18637/jss.v076.i13>
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., ... Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, *23*, 4407–4414. <https://doi.org/10.1093/nar/23.21.4407>
- Wang, Y., & Önal, H. (2016). Optimal design of compact and connected nature reserves for multiple species. *Conservation Biology*, *30*, 413–424. <https://doi.org/10.1111/cobi.12629>
- Wang, Y.-H., Yang, K.-C., Bridgman, C. L., & Lin, L.-K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, *23*, 989–1000. <https://doi.org/10.1007/s10980-008-9262-3>
- Watson, J. E. M., Dudley, N., Segan, D. B., & Hockings, M. (2014). The performance and potential of protected areas. *Nature*, *515*, 67–73. <https://doi.org/10.1038/nature13947>
- Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration:  $F_{ST} \approx 1/(4Nm + 1)$ . *Heredity*, *82*, 117–125. <https://doi.org/10.1038/sj.hdy.6884960>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Hanson JO, Fuller RA, Rhodes JR. Conventional methods for enhancing connectivity in conservation planning do not always maintain gene flow. *J Appl Ecol*. 2019;56:913–922. <https://doi.org/10.1111/1365-2664.13315>