Conserving threatened riparian ecosystems in the American West: Precipitation gradients and river networks drive genetic connectivity and diversity in a foundation riparian tree (*Populus angustifolia*)

Helen M. Bothwell1 | Samuel A. Cushman2 | Scott A. Woolbright3 | Erika I. Hersch-Green4 | Luke M. Evans5 | Thomas G. Whitham1,6 | Gerard J. Allan1,6

1Environmental Genetics & Genomics Facility, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA
2Rocky Mountain Research Station, United States Forest Service, Flagstaff, AZ, USA
3Department of Biology, University of Arkansas, Little Rock, AR, USA
4Department of Biological Sciences, Michigan Technological University, Houghton, MI, USA
5Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA
6Merriam-Powell Center for Environmental Research, Northern Arizona University, Flagstaff, AZ, USA

**Correspondence**
Helen M. Bothwell, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA.
Email: Helen.Bothwell@nau.edu

**Funding information**
NSF-IGERT Fellowships; NSF GK-12 Stem and NPS George Melendez Wright Climate Change Fellowships; NSF FIBR, Grant/Award Number: DEB-0425908; NSF MacroSystems, Grant/Award Number: DEB-1340852; NSF, Grant/Award Number: DBI-1126840

**Abstract**
Gene flow is an evolutionary process that supports genetic connectivity and contributes to the capacity of species to adapt to environmental change. Yet, for most species, little is known about the specific environmental factors that influence genetic connectivity, or their effects on genetic diversity and differentiation. We used a landscape genetic approach to understand how geography and climate influence genetic connectivity in a foundation riparian tree (*Populus angustifolia*), and their relationships with specieswide patterns of genetic diversity and differentiation. Using multivariate restricted optimization in a reciprocal causal modelling framework, we quantified the relative contributions of riparian network connectivity, terrestrial upland resistance and climate gradients on genetic connectivity. We found that (i) all riparian corridors, regardless of river order, equally facilitated connectivity, while terrestrial uplands provided 2.5× more resistance to gene flow than riparian corridors. (ii) Cumulative differences in precipitation seasonality and precipitation of the warmest quarter were the primary climatic factors driving genetic differentiation; furthermore, maximum climate resistance was 45× greater than riparian resistance. (iii) Genetic diversity was positively correlated with connectivity ($R^2 = 0.3744$, $p = .0019$), illustrating the utility of resistance models for identifying landscape conditions that can support a species’ ability to adapt to environmental change. From these results, we present a map highlighting key genetic connectivity corridors across *P. angustifolia*’s range that if disrupted could have long-term ecological and evolutionary consequences. Our findings provide recommendations for conservation and restoration management of threatened riparian ecosystems throughout the western USA and the high biodiversity they support.

**KEYWORDS**
climatic gradients, gene flow, genetic and functional connectivity, landscape genetics, landscape resistance, reciprocal causal modelling
1 | INTRODUCTION

Lack of knowledge regarding gene flow and dispersal patterns is a major “blind spot” limiting our ability to prepare for climate change impacts on natural ecosystems (Engler et al., 2009; Pauls, Nowak, Bálint, & Pfenniger, 2013). Gene flow contributes to species’ capacity to adapt to environmental change (Merriam, 1984), and disruptions to gene flow can reduce genetic diversity, hinder adaptation and increase risks of local extirpation (Allendorf & Luikart, 2006; Vranckx, Jacquemyn, Muys, & Honnay, 2012). As the human footprint has expanded sharply across the western USA (Hand, Cushman, Langduth, & Lucotch, 2014; Leu, Hanser, & Knick, 2008), an increasing number of species are being impacted by habitat loss and fragmentation (Inman et al., 2013; Kindlmann & Burel, 2008; Noss, LaRoe, & Scott, 1995). These impacts are compounded as species face added pressure to migrate and adapt in response to climate change (Mantyka-Pringle, Martin, & Rhodes, 2012). While identifying and prioritizing connectivity corridors is both a global challenge and a key mandate for US federal agencies tasked with conservation management (Langduth, Hand, Glassy, Cushman, & Savaya, 2012), little is known about the specific environmental factors that facilitate or inhibit gene flow through complex landscapes for most species (but see Castillo, Epps, Davis, & Cushman, 2014; Cushman et al., 2014; Schwartz et al., 2009; Shirk, Wallin, Cushman, Rice, & Warheit, 2010), or the effects those factors have on patterns of genetic diversity and differentiation.

Riparian ecosystems are one of the most intensely impacted habitat types worldwide (Vörösmarty et al., 2010). Although they occupy a small percentage of the landscape, riparian ecosystems support disproportionately high biodiversity (Bridgeland, Beier, Kolb, & Whitham, 2010; Johnson, Haight, & Simpson, 1977), making their conservation a priority (Davies, 2010; Gitlin et al., 2006). Damming and diversion of waterways, cattle grazing, and human encroachment have contributed to the decline of these threatened ecosystems (Rood & Mahoney, 1990) such that <3% of pre-20th-century riparian habitat remains in the American West (Noss et al., 1995). Narrowleaf cottonwood (Populus angustifolia) is a dominant tree species along many western US rivers. As a foundation species, its genetic attributes influence the diversity and structure of dependent communities (Ferrier et al., 2012; Ikeda, Bothwell, et al., 2014; Lamit et al., 2015; Whitham et al., 2012); therefore, disruption of gene flow among P. angustifolia also impacts diverse associated species. Understanding the specific environmental conditions that drive genetic connectivity, differentiation and diversity in P. angustifolia is critical for supporting resiliency of western riparian ecosystems as they face global change.

Species distributions and patterns of genetic connectivity reflect the spatially structured ecological, environmental and geographic factors that influence gene flow (Wagner & Fortin, 2005). Clinal patterns of genetic structure are common in broadly distributed tree species given their differential responses to environmental gradients (e.g., climate, photoperiod) (Alberto et al., 2013; Ingvarsson, García, Hall, Luquez, & Jansson, 2006; O’Neill, Hamann, & Wang, 2008); therefore, understanding these relationships requires statistical frameworks capable of incorporating complex spatial gradients. Classic population genetic methods that assume discretely bounded and panmictic populations can be inappropriate for such systems, where assigning distinct boundaries may introduce errors and obscure true patterns of genetic differentiation across the landscape (Balkenhol, Cushman, Storfer, & Waits, 2015; Shirk & Cushman, 2011). This has been a central motivation spurring development of the field of landscape genetics (Balkenhol et al., 2015; Manel, Schwartz, Luikart, & Taberlet, 2003) and in particular individual-based genetic approaches (Balkenhol, Waits, & Dezzani, 2009; Storfer et al., 2007).

Here, we employ a spatially explicit, individual-based modelling framework to better understand the fine-scale relationships between environmental and genetic gradients in narrowleaf cottonwood. We utilize the concept of landscape resistance, which quantifies the relative permeability of landscape features to gene flow. This approach assesses the ability of organisms or propagules to travel through their environment, with landscape features either hindering or facilitating gene flow depending on the cost or relative resistance they impose on movement (Spear, Balkenhol, Fortin, McRae, & Scribner, 2010). Given hypothesized cost–movement relationships, landscape resistance models (Cushman, McKelvey, Hayden, & Schwartz, 2003; McRae, 2006), using least-cost path (Schwartz et al., 2009) or circuit theory (McRae, Dickson, Keitt, & Shah, 2008) approaches, can be used to quantify the relative influence of environmental variables on connectivity and to identify connectivity corridors. Furthermore, landscape resistance models provide a means for testing classic population genetic hypotheses linking genetic connectivity to genetic diversity.

Our study contributes empirical tests of two key hypotheses central to population genetics. First, the central–marginal (Eckert, Samis, & Lougheed, 2008) or abundant centre hypothesis (Sagarin & Gaines, 2002) predicts that large effective population sizes ($N_e$) and high gene flow maintain high genetic diversity in the core of species’ ranges. Conversely, small $N_e$ lower gene flow and higher rates of genetic drift are characteristic of peripheral populations, thereby predicting lower genetic diversity at range margins (Vucetich & Waite, 2003). Yet, phylogeographic history also contributes to contemporary spatial genetic patterns. A second hypothesis, the leading edge–rear edge model (Hampe & Petit, 2005; Hewitt, 2000), suggests that opposing dynamics at different range margins can lead to departures from central–marginal hypothesis predictions (Eckert et al., 2008; Sagarin, Gaines, & Gaylord, 2006). Range expansion along leading edges often results in low genetic diversity as growth of a limited number of colonizing individuals follows rare, long-distance dispersal events (Hewitt, 2000). In contrast, old, stable, rear edges can harbour rich diversity and be valuable repositories for conservation (Hampe & Petit, 2005). Given these hypotheses and knowledge of a species’ phylogeographic history, we would predict a positive relationship between genetic diversity and connectivity as a function of landscape resistance to gene flow. Although several simulation studies have demonstrated support for connectivity–diversity relationships (Macdonald, Cushman, & Macdonald, in review; Wasserman,
Cushman, Littell, Shirk, & Landguth, 2013; Wasserman, Cushman, Shirk, Landguth, & Littell, 2012), no studies to date have conducted an empirical test in wild populations.

In the context of these hypotheses, we tested three overarching predictions. First, we predicted that riparian connectivity is a primary driver of genetic connectivity in *P. angustifolia*, and more specifically, we considered how different stream orders within the river network continuum may differentially affect gene flow. We anticipated that 1st-order streams (e.g., narrow headwaters with shallow channel depth) would be more restrictive to gene flow than higher-order streams (i.e., two 1st-order streams merge to form a 2nd-order stream) that typically support larger populations. In the closely related species, *P. fremontii*, Cushman et al. (2014) found that gene flow was facilitated by mid-sized to large rivers and constrained by lower-order streams and terrestrial uplands. Given that *P. angustifolia* occupies higher elevation portions of drainages wherever the two species co-occur, we expected that *P. angustifolia*’s more disjunct distribution would be associated with greater genetic differentiation than found in *P. fremontii* (Cushman et al., 2014; Evans, Allan, & Whitham, 2012; Evans et al., 2015). In northern latitudes, however, *P. angustifolia* is more commonly found at lower elevations along higher-order rivers. Therefore, we predicted that (i) gene flow in *P. angustifolia* would be facilitated by a broader range of river orders, and headwater streams would provide less resistance to *P. angustifolia* than *P. fremontii*; (b) gene flow would be restricted by terrestrial uplands; and (c) genetic connectivity would be greater among northern versus southern populations.

Second, we predicted that *P. angustifolia* would be genetically differentiated along climate gradients. In addition to neutral processes, selection can influence neutral genetic patterns, particularly if it occurs during establishment (Orsini, Vanoverbeke, Swillen, Meegeay, & Meester, 2013; Sexton, Hangartner, & Hoffmann, 2014). In cottonwood, seedling establishment is closely tied to flood regimes (Baker, 1990): temperature and precipitation gradients may contribute to differentiation through their influence on timing of spring flood events and cueing of reproductive phenology (e.g., flowering, seed release; Perry, Andersen, Reynolds, Nelson, & Shafroth, 2012; Rood et al., 2005; Yamamoto & Sota, 2009). Asynchrony in seedling recruitment has been documented across adjacent river drainages and even among different reaches of the same river (Coble & Kolb, 2013), suggesting that variation in temperature and precipitation may be influencing genetic differentiation at very fine spatial scales.

Third, we predicted that genetic diversity and connectivity would be positively correlated, supporting central-marginal hypothesis predictions (Eckert et al., 2008). Previous work by Evans et al. (2015) determined that southern populations have contracted and northern populations have expanded since the last glacial maxima (LGM). Given *P. angustifolia*’s phylogeographic history, we refined central-marginal hypothesis predictions such that we expected high connectivity in the north and lower connectivity in the south. Initial postglacial expansion likely resulted in decreased genetic diversity along the expanding front; however, we anticipated that ~18,000 years of population persistence have likely contributed to substantial present-day genetic connectivity and diversity in the north. Genetic diversity in southern *P. angustifolia* depends upon whether this region represents a stable or trailing rear edge (Hampe & Petit, 2005), and the degree to which population contraction has impacted genetic connectivity.

## METHODS

### 2.1 Study species

Narrowleaf cottonwood (*Populus angustifolia*) is a foundation tree species that occurs along riparian corridors throughout mountainous regions of western North America from northeastern Sonora, Mexico, to southern Alberta, Canada (Little, 1976). Riparian corridors encompass stream channels and adjacent alluvial habitats, and terrestrial-aquatic feedback loops drive both cottonwood persistence and overall riparian ecosystem integrity (Compson et al., 2016; Helfield & Naiman, 2001). Riparian corridors influence pollen dispersal by channelling wind flow, while both wind and water contribute to seed dispersal. Long-distance dispersal and long lifespan have the potential to contribute to extensive connectivity in this species. Typical of pioneer species, its seeds contain minimal endosperm provisioning (Braatne, Rood, & Heilman, 1996). Successful germination requires that seed release coincides with disturbance in the form of flood-scorched virgin habitat, followed by gradual floodwater recession (Rood et al., 2005). Flooding, ice, high wind and mammalian browsing can also disperse broken twigs, contributing to clonal propagation where branch fragments root and establish downstream (Braatne et al., 1996; Rood, Goater, Mahoney, Pearce, & Smith, 2007). Both seed production and vegetative propagation vary greatly among cottonwood species and individual genotypes (Schweitzer, Martinsen, & Whitham, 2002), further contributing to demographic variation among watersheds.

### 2.2 Genetic collections and analyses

We collected leaf material from 696 trees at 40 different sampling locations spanning the full range of geographic and environmental niche space occupied by *P. angustifolia* within the USA (Figure 1, Table S1). We extracted genomic DNA from leaf material following the methods in Cushman et al. (2014) and standardized DNA to 15 ng/µl. From a preliminary screening of 17 microsatellite loci developed under the *Populus trichocarpa* Genome Project (Tuskan et al., 2004, 2006; http://www.ornl.gov/sci/ipgc/ssr_resource.htm), we selected 12 loci that demonstrated substantial polymorphism, successful amplification across the species’ range, and scoring consistency (Table S2). Loci were amplified via touchdown PCR in 10 µl reactions (see Appendix S1 for PCR protocol). Fragment analysis of PCR products was performed on an ABI 3730xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) with Genescan LI500 internal size standard (ABI); allele fragment sizes were scored with GeneMarker v2.2.0 (Softgenetics LLC, State College, PA, USA). Data quality control (genotyping error rate, null allele testing) is available in Appendix S2.
2.3 Genetic diversity and differentiation

We assessed genetic diversity and differentiation using GENALEX 6.5 (Peakall & Smouse, 2006, 2012). Prior to conducting analyses, we tested for and removed clonal replicates. Mean unbiased expected heterozygosity ($uH_e$) was calculated for each locus and sampling location (Tables S1 and S2); $uH_e$ provides greater accuracy for diversity comparisons including small and uneven sample sizes (Nei, 1978). For comparison, we also calculated individual and mean site observed heterozygosity ($H_o$), mean expected heterozygosity ($H_e$) and mean rarefied allelic richness ($A_r$; HP-RARE 1.0; Kalinowski, 2005). Much of the rear edge of *P. angustifolia*’s distribution is characterized by small, relatively isolated stands. Small sample sizes ($n < 10$) included in our study represent all genets present at a given site; diversity measures reported for these sites are not estimates but rather a comprehensive assessment of genetic variation at selected loci. We also conducted an analysis of molecular variance (AMOVA) to estimate among site differentiation ($F_{ST}$; Wright, 1965), to assess how genetic variation is partitioned within and among individuals and among sampling locations, and to facilitate cross-species comparisons.

2.4 Genetic distance

To test how landscape resistance (e.g., relative permeability of landscape variables to gene flow) influences genetic connectivity, we first calculated pairwise genetic distances among all individuals using a principal components analysis (PCA)-based method (Shirk et al., 2010; R code available in Code S1). First, we converted microsatellite data into a matrix of individual allele frequencies (*import2genind* function, ADEGENET package; Jombart, 2008; Jombart & Ahmed, 2011; R 3.1.2, R Development Core Team 2014). We then derived eigenvectors from the allele frequency data and generated a pairwise genetic distance matrix based on distance among individuals along

**FIGURE 1** *Populus angustifolia* collection locations. We sampled 696 trees from 40 different collection locations spanning the full range of narrowleaf cottonwood within the western USA.
the first eigenvector (Euclidean distance function in ECODIST; Goslee & Urban, 2007). For comparison, we also calculated interindividual AMOVA genetic distance (Dyer, Westfall, Sork, & Smouse, 2004; Excoffier, Smouse, & Quattro, 1992; GSTUDIO R package, Dyer, 2014). Individualistic approaches based on allele frequency distributions have been shown to better reflect contemporary influences on genetic connectivity, compared to population- and heterozygosity-based metrics such as $F_{ST}$ (Murphy, Evans, Cushman, & Storfer, 2008; Murphy, Evans, & Storfer, 2010).

### 2.5 Reciprocal causal modelling and multivariate restricted optimization

To test the prediction that river network connectivity, terrestrial upland resistance, and climate gradients jointly control genetic connectivity in *P. angustifolia*, we employed multivariate restricted optimization (Shirk et al., 2010) in a reciprocal causal modelling (RCM) framework (Cushman, Wasserman, Landguth, & Shirk, 2013). The Shirk et al. (2010) approach systematically varies values for each resistance parameter across a range of values of the other predictor variables, facilitating identification of an optimal model across a very large parameter space. RCM then enables comparison of a large number of alternative hypotheses within a single modelling framework. Figure 2 illustrates an example of RCM workflow and model selection.

First, we parameterized landscape resistance models in geographic information system (GIS) raster layers by assigning values to each grid cell based on hypothesized relationships between environmental variables and cost of movement (Cushman et al., 2006; Spear et al., 2010; see Fig. S1 for example). Raster layers were processed in ArcGIS or ARCINFO WORKSTATION 10.2.2 (Environmental Systems Research Incorporated (ESRI), Redlands, CA, USA, 2011); all restricted optimization in ARCGIS or ARCINFO WORKSTATION 10.2.2 (Environmental Systems Research Incorporated (ESRI), Redlands, CA, USA, 2011); all mental variables and cost of movement (Cushman et al., 2006; Spear et al., 2010; see Fig. S1 for example). Raster layers were processed in ARCGIS or ARCINFO WORKSTATION 10.2.2 (Environmental Systems Research Incorporated (ESRI), Redlands, CA, USA, 2011); all

We then used RCM to identify which landscape resistance model best describes spatial genetic patterns in cottonwood. For each set of isolation by resistance (IBR) hypotheses, we calculated partial Mantel correlation coefficients ($r$; Legendre & Legendre, 1998; Mantel, 1967: Figure 2a) between genetic distance (GD) and each landscape resistance hypothesis ($H_{p1}$ to $H_{pn}$) while partialling out the effect of all other hypotheses ($H_{p1}$ to $H_{pn}$) following the general format:

$$GD \sim H_{p1} \ldots H_{pt} | H_{p1} \ldots H_{pt}$$

We also compared each set of IBR hypotheses with isolation by distance (IBD; Wright, 1943). If the null IBD model receives greater support, this suggests that an IBR hypothesis performs no better than simple Euclidean distance for explaining genetic differentiation. Partial Mantel tests were calculated using ECODIST; computationally intensive, high-value resistance surface correlations required additional $z^*$ software (Bonnet & Van de Peer, 2002).

Best-supported models were then selected by comparing relative support, or the difference between reciprocal partial Mantel tests

**Figure 2** Reciprocal causal modelling workflow and model selection. (a) First, a matrix of partial Mantel correlation coefficients is generated for all hypothesis pairs. This measures the correlation between genetic distance (GD) and a focal hypothesis (H1) while partialling out the effect of a second hypothesis (H2). (b) Next, relative support is calculated as the difference between reciprocal tests. If the first hypothesis is supported, $A$ will be positive and $B$ will be zero or negative, and vice versa. (c) For each set of hypotheses, the full matrix of reciprocal causal modelling (RCM) results is presented as a coloured block diagram in which each cell is colour-coded according to its relative support (e.g., eqns. A and B). In this example, $A = -0.70491$, and hence is coded navy; $B = 0.70491$, so is coded bright yellow. H13 is fully supported independent of all other hypotheses as indicated by its column containing all positive values and its row containing all negative values. In cases where more than one column exhibited all positive values, we selected the best-supported hypothesis based on a second criterion, column average – row average ($C - R$).
If H1 is supported independently of H2, the result will be positive; conversely, when H2 is supported independently of H1, the difference will be zero or negative. RCM matrices comparing relative support among all hypotheses are presented as coloured block diagrams with each cell colour-coded by relative support (Figure 2c). A hypothesis is fully supported independently of all others if its column contains all positive values, and its row contains all negative values (Ruiz-Gonzalez, Cushman, Madeira, Randi, & Gomez-Molinero, 2015; Yang, Cushman, Yang, Yang, & Bao, 2013).

Some controversy has been raised over the use of Mantel tests in landscape genetics, particularly associated with their tendency towards high type I error rates (Castellano & Balletto, 2002; Guillot & Rousset, 2013). This issue is common to individual Mantel tests and especially prevalent when data exhibit autocorrelation (Amos et al., 2012; Meirmans, 2012). However, directly competing many alternative IBR and IBD models against each other greatly reduces the risk of type I errors (Cushman et al., 2013). Simulation modelling by Cushman and Landguth (2010) and Cushman et al. (2013) validated the specific use of partial Mantel testing within a simultaneous RCM framework to identify robust relative support among alternative hypotheses. Additionally, Castillo et al. (2014) and Shirk et al. (2010) used complementary simulation approaches to demonstrate that multivariate restricted optimization with Mantel tests can effectively identify the correct drivers of gene flow among a large number of highly correlated alternative models. Finally, Zeller et al. (2016) compared model selection criteria and found that performance evaluation via relative support in an RCM framework (Castillo et al., 2014; Cushman et al., 2013) improved correct identification of the true resistance model when compared with original causal modelling selection criteria (Cushman et al., 2006); model selection via relative support correctly discriminated among alternative resistance hypotheses comparing different landscape predictor variables >98% of the time. While critiques have shed light on important misuses of Mantel tests in spatial ecology (Legendre, Fortin, & Borcard, 2015), ample testing has demonstrated that use of partial Mantel tests within the narrow application of discriminating among alternative landscape resistance models via relative support is valid.

2.6 | Riparian network and terrestrial upland resistance hypotheses

To test the prediction that terrestrial uplands (i.e., all nonriparian corridor habitat) and smaller, lower-order rivers provide greater resistance to gene flow than larger, higher-order rivers, we used a four-step procedure. We obtained high-resolution Strahler stream order data from the National Hydrology Dataset Plus version 2 (NHDPLUSV2: http://www.horizon-systems.com/NHDPplus/NHDPlusV2_data.php; McKay et al., 2012; Strahler, 1957). In Step 1, we tested nine IBR hypotheses of the influence of riparian network connectivity on genetic connectivity. Hypothesis 1 proposes that all river orders facilitate gene flow and terrestrial uplands provide high resistance (river resistance = 1, upland resistance = 20). Hypothesis 2 states that only river orders 2–9 facilitate gene flow, while first-order rivers and uplands provide high resistance and so on with increasing resistance assigned to lower-order rivers with each successive hypothesis (Table 1).

Step 2 used RCM on the best-supported hypothesis from Step 1 to evaluate 21 hypotheses of relative resistance among river orders.

### Table 1

Experiment 1 used RCM to test the relative resistance of terrestrial uplands and river orders on gene flow in *Populus angustifolia*. We compared nine hypotheses of isolation by river network resistance (IBR; Hyp1-9) and isolation by distance (IBD). Terrestrial uplands were assigned high resistance; Hyp1-9 tested a gradient of low (1) to high (20) relative resistance among river orders. A river resistance value of 1 is equal to simple Euclidean cost distance; crossing a single raster cell confers the lowest potential cost in both cases. A separate GIS raster surface was generated for each hypothesis.

<table>
<thead>
<tr>
<th>River order</th>
<th>Resistance models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hyp1</td>
</tr>
<tr>
<td>Terrestrial Uplands</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

> BOTHWELL ET AL.
with terrestrial uplands held constant. We tested three suites of hypotheses that varied the rate of increase in relative resistance of river orders with respect to each other. Low resistance hypotheses tested slow rates of increase in resistance from higher- to lower-order rivers. Similarly, medium and high resistance hypotheses tested increasingly rapid rates of resistance to gene flow with decreasing river order (Table 2, Fig. S1).

Step 3 maintained the optimized relative resistance of river orders established in Step 2 and varied hypothesized values for upland resistance. Model h1b suggests that terrestrial uplands provide 1.5× greater resistance to gene flow than river networks, whereas h1m suggests that uplands provide 25× more resistance (Table 3).

Step 4 sought stable convergence of both relative resistance of river orders with respect to each other and river networks relative to upland resistance through iterative repetition of Steps 2 and 3 (Table 4).

2.7 Climate hypotheses

To better understand how climate contributes to patterns of genetic distance, we first identified relevant bioclimatic variables using a species distribution model (SDM; Maxent 3.3.3k; Phillips, Anderson, & Schapire, 2006). WorldClim bioclimatic variables (1961–1990 averages; Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) were obtained from the Global Climate Model Downscaled Data Portal (www.ccafs-climate.org/data/). After removing highly correlated variables (r > .8), we derived jackknife estimates of per cent contribution and permutation importance of each variable to the model building process. Variables with per cent contributions <5% were excluded (Ikeda, Grady, Shuster, & Whitham, 2014). We retained five bioclimatic variables for landscape genetic modelling: mean diurnal temperature range (BIO2), temperature seasonality (BIO4), mean temperature of the warmest quarter (BIO10), precipitation seasonality (BIO15) and precipitation of the warmest quarter (BIO18).

We compared three overarching climate hypotheses to determine which model best describes observed patterns of genetic distance in *P. angustifolia*: (i) isolation by environment (IBE) according to regional climate clusters, (ii) pairwise IBE and (iii) isolation by landscape resistance (IBR). Isolation by environment (Wang & Summers, 2010) can arise when nonlocal migrants fail to establish or when hybrid offspring exhibit reduced fitness in a new environment (Wang & Bradburd, 2014). IBE models describe a spatial genetic pattern in which genetic and environmental distances are positively correlated, independent of geographic distance or landscape matrix (Wang & Bradburd, 2014); selection acts as an environmental filter at occurrence locations (island-based model). In contrast, IBR models (McRae, 2006) describe how gene flow has been shaped by the movement of organisms or propagules. IBR models consider how the landscape matrix has influenced genetic distance by quantifying the probability of dispersal between individuals given resistance of the intervening landscape. Following, we explore whether genetic distance in cottonwood is better described by models of environmental filtering (IBE) or functional connectivity (IBR) (Mateo-Sánchez et al., 2015).

First, we modelled genetic distance as a function of multivariate climate cluster membership, using an iterative k-means classification method (IsoCluster; ARCGIS 2012) to identify regions of similar climate across *P. angustifolia*’s range. We tested five regional IBE hypotheses (k = 3–7).

Second, we tested 31 pairwise IBE hypotheses in which we modelled genetic distance as a function of pairwise Euclidean climate distance among all individuals (extracted from standardized climate

<table>
<thead>
<tr>
<th>River order</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>h1</td>
<td>h2</td>
<td>h3</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>25</td>
<td>34</td>
</tr>
<tr>
<td>Terrestrial Uplands</td>
<td>150</td>
<td>150</td>
<td>150</td>
</tr>
</tbody>
</table>

*TABLE 2* Step 2 evaluated 21 hypotheses of relative resistance among river orders with terrestrial uplands held constant. Step 1 indicated that all river orders 1–9 facilitate gene flow, and terrestrial uplands provide comparatively high resistance. Low hypotheses tested a slow rate of increase in resistance from higher-order to lower-order rivers. Similarly, medium and high slope hypotheses tested more rapid rates of increase in resistance to gene flow with decreasing river order, whereas h1m suggests that uplands provide 25× more resistance (Table 3).
Hypotheses included genetic distance as a function of each of the five climate variables singly (a) and all combinations of summed climate distance (b), for example:

\[ GD \sim BIO2 \]  
\[ GD \sim BIO2 + BIO4 + BIO18. \]

Third, we tested an IBR model in which genetic distance was a function of cumulative differences in climate traversing the landscape between individuals. Whereas IBE models consider the influence of environmental filtering at occurrence locations, IBR models describe how the landscape matrix between individuals influences genetic connectivity. To understand how *P. angustifolia* gene flow tracks climate variables across the landscape, we first tested 13 hypotheses in which we varied the shape of the response function for each climate variable. Following the method of Shirk et al. (2010), we varied the response shape exponent \( x \) from 0.1 to 7 to generate strongly convex, to linear, to concave response functions (Figure 3a). After identifying optimal climate slope relationships for each variable, we tested 31 hypotheses modelling genetic distance as a function of cumulative climate resistance distance among individuals. As with the pairwise IBE model above, hypotheses explored genetic distance as a function of each climate variable singly, as well as all combinations of summed climate resistance distances.

Finally we competed the best-supported model from each of the three overarching climate hypotheses against each other to identify the optimal relationship between genetic distance and climate (Cushman et al., 2014).

### Table 3

<table>
<thead>
<tr>
<th>River order</th>
<th>Resistance models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h1a</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Terrestrial Uplands</td>
<td>10</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>River Order</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h1</td>
<td>h2</td>
<td>h3</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Terrestrial Uplands</td>
<td>125</td>
<td>125</td>
<td>125</td>
</tr>
</tbody>
</table>

Rasters using ECODIST. Hypotheses included genetic distance as a function of each of the five climate variables singly (a) and all combinations of summed climate distance (b), for example:

\[ GD \sim BIO2 \]  
\[ GD \sim BIO2 + BIO4 + BIO18. \]
2.8 | Optimizing river, upland and climate resistance

After identifying the best-supported river network, terrestrial upland and climate resistance models, we applied multivariate restricted optimization to reach a stable solution of optimal climate resistance parameters relative to river and upland resistance. We tested 133 models in which we systematically varied maximum resistance ($R_{\text{max}}$; Shirk et al., 2010) of our best-supported climate hypothesis from 1.0125 to 77.5, while holding river network and upland parameter values constant (rivers = 1, terrestrial uplands = 2.5).

2.9 | Genetic connectivity corridors and connectivity–diversity relationships

To visualize synoptic patterns and identify genetic connectivity corridors, we modelled cumulative resistant kernel densities (Compton, 2014; Wasserman et al., 2012). Although our sampling was stratified to include locations that appeared to span very remote to highly connected sites, our models revealed an over-representation of cottonwood in regions of moderate-to-high genetic connectivity. To correct for this sampling bias, connectivity–diversity relationships were assessed on a reduced data set representing a random, stratified sample of connectivity values. Connectivity values were stratified by steps of 10 (e.g., 0–10, 10–20), with five random sites selected from each strata. If <5 sites were available within a given range, all sites were used for that strata, resulting in 23 sites and 190 individuals included in regression models.

3 | RESULTS

3.1 | Genetic analyses

We detected moderate-to-high genetic variation, with allelic richness varying from 4 to 25 and $u_{H}$ ranging from 0.035 to 0.813 across loci (Table S2). Across sampling locations, $u_{H}$ varied from 0.219 to 0.679 ($k = 0.409$, Table S1). Pairwise differentiation among sampling locations ranged from 0.000 to 0.607 (Table S3). Global genetic differentiation was also moderate to high ($F_{\text{ST}} = 0.28$, $p = .001$), with 72% of genetic variation occurring within individuals, 28% distributed among sampling locations, and negligible variation among individuals. Our estimate of global genetic differentiation ($F_{\text{ST}} = 0.28$) is similar to previously reported estimates for this species (Evans et al. 2012; Evans et al., 2015; $F_{\text{ST}} = 0.27$ and 0.21, respectively), and within the range of multispecies means established for long-lived, outcrossing and wind/water dispersed species (Nybom & Bartish, 2000). Broader sampling and inclusion of more disjunct, isolated sites likely contributed to our slightly higher estimate. Comparing differentiation across cottonwood species, $F_{\text{ST}}$ for _P. angustifolia_ was greater than that reported for _P. fremontii_ (Cushman et al., 2014; $F_{\text{ST}} = 0.22$), as predicted given the species’ contrasting ecological niches. Clonality in _P. angustifolia_, in contrast with little to no asexual regeneration in _P. fremontii_, may also be contributing to higher $F_{\text{ST}}$ detected for this species. Accordingly, these estimates are consistent with multispecies, meta-analysis findings comparing outcrossing ($F_{\text{ST}} = 0.22$) and mixed ($F_{\text{ST}} = 0.26$) breeding systems (Nybom, 2004).
PCA-based genetic distance and interindividual AMOVA genetic distance were strongly correlated ($r = .77$, $p = .01$). In nearly all cases, both measures provided qualitatively the same results. Only with the RCM comparing isolation by climate resistance models did the two genetic distance measures generate different results (see Discussion below). For brevity, we only present AMOVA results.

3.2 | Riparian network and terrestrial upland resistance

In Step 1 of optimizing riparian resistance, Hypothesis 1 was best-supported, indicating that all river orders support *Populus angustifolia* gene flow (Table 1, Fig. S2). The IBD model performed slightly better than the best river network hypothesis; however, we retained Hypothesis 1 as we expected it would outcompete IBD when fully optimized in Step 4. In Step 2, the first hypothesis (h1) was again best-supported, indicating that all river orders equally facilitate genetic connectivity (Table 2, Fig. S3). In Step 3, model h1d was best-supported; terrestrial uplands provide $2.5\times$ more resistance to gene flow than riparian corridors (Table 3, Fig. S4). In the final step seeking stable convergence of both relative resistance of river orders with respect to each other and river networks relative to upland resistance, the best-supported model (h1) confirmed that all river orders provide equally low resistance to gene flow, and terrestrial uplands are $2.5\times$ more resistant than river networks (Table 4, Fig. S5). Furthermore, this optimized resistance model was fully supported relative to IBD.

3.3 | Climate resistance

Testing five hypotheses of genetic distance as a function of climate cluster membership, we found that genetic distance best reflected a model based on five climate clusters. However, ISO5 performed poorly compared to IBD, with IBD $C = R$ indicating $7.8\times$ better model performance than ISO5 (Fig. S6).

Second, comparing 31 pairwise IBE models, we determined that model d4.18 was best-supported (Fig. S7). Temperature seasonality and precipitation of the warmest quarter jointly explained the most variation in genetic distance by pairwise climate distance.

Third, modelling genetic distance as a function of IBR, we found that linear and strongly concave response functions best described the relationships between genetic distance and bioclimatic variables. Genetic distance was $–$linearly related to mean temperature and precipitation of the warmest quarter (BIO10$^{0.7}$, BIO18$^{0.7}$); strongly concave relationships best described the remaining bioclimatic variables (BIO2$^2$, BIO4$^2$, BIO15$^2$; Figure 3b; Figs S8–S12). Using the optimal response function model for each bioclimatic variable, we then tested 31 models of isolation by cumulative climate resistance distance. The best-supported model using PCA-based genetic distance included BIO2, BIO4, BIO15 and BIO18, while findings based on AMOVA genetic distance indicated the best model included only BIO15 and BIO18 (Fig. S13). This was the only instance where findings differed between genetic distance measures. We chose the most parsimonious result, selecting model r15.18 to compete in the final climate model optimization step.

Comparing the best of each of the three overarching climate models against each other and with IBD, IBR was the only fully supported model, and no other models were supported independently of it (Fig. S14).

3.4 | Optimization of river, upland and climate resistance

We used RCM to compare 133 models that systematically varied maximum resistance ($R_{\text{max}}$) of the best-supported climate resistance hypothesis (including the additive effects of precipitation seasonality and precipitation of the warmest quarter) such that climate was $1.0125–77.5\times$ as influential as river network and $0.40–31\times$ as influential as terrestrial upland resistance. In these models, riparian resistance is equal to Euclidean cost distance; crossing a single raster cell confers a cost of one, whereas crossing a single cell of terrestrial upland costs as much as crossing 2.5 riparian corridor cells. The best-supported model (CR7040) indicated that maximum climate resistance was $45\times$ greater than $R_{\text{max}}$ of river networks and $18\times$ greater than terrestrial upland resistance in driving spatial patterns of genetic distance in *P. angustifolia* (Figure 4). In other words, it is $45\times$ more costly to disperse against a maximally different climate gradient than to disperse across a single riparian raster cell with no change in climate. Similarly, dispersing across a single cell of terrestrial upland habitat with maximally different climate costs as much as crossing 18 upland cells with the same climate as the focal sample. Riparian corridors confer very low resistance, followed by terrestrial uplands; the cost to disperse against strong climate gradients far outweighs both of these. We also confirmed there was no independent relationship with IBD after accounting for the influence of riparian network, upland and climate resistance. As predicted, environmental distance provided a better explanation for genetic distance than simple Euclidean distance.

3.5 | Genetic connectivity corridors and connectivity–diversity relationships

Mapping dispersal corridors on the landscape revealed high levels of connectivity within the northern and southern portions of *P. angustifolia*’s range, but few connections between these regions (Figure 5). A few key dispersal corridors traversing western Utah link northern and southern populations; however, the central and eastern portions of the species’ range are characterized by low expected density of dispersing individuals. Connectivity corridors were derived using a maximum threshold of 577,191 cumulative cost distance units (Mantel $r = .124$, $p = .01$), based on Mantel correlogram peak autocorrelation indicating significant genetic connectivity within this threshold (Table S4, Fig. S15a). While the majority of dispersal likely occurs within a few kilometres, the selected threshold accommodates the very long tail of the dispersal distance frequency distribution known from previous *Populus* research (Slavov et al., 2009; Fig. S15b).
4.1 Riparian corridors as ecological drivers of genetic connectivity

Supporting our prediction, riparian networks emerged as a primary driver of genetic connectivity, while terrestrial uplands and climate gradients were much stronger contributors to genetic distance. Compared to adjacent semiarid landscapes, riparian corridors confer very low resistance to gene flow; therefore, maintenance of river network connectivity is essential for supporting genetic connectivity. Comparing connectivity models across species, as predicted, gene flow in *P. angustifolia* was facilitated by a broader range of river orders (all orders conferring equal resistance) compared to *P. fremontii*. Whereas medium-to-large, higher-order rivers provided low resistance to both species, smaller, lower-order rivers conferred high resistance to *P. fremontii* (Cushman et al., 2014). Similarly, terrestrial uplands were 2.5× more resistant to gene flow than river networks for *P. angustifolia*, while uplands and headwater streams provided 12× greater resistance to *P. fremontii* (Cushman et al., 2014). These findings illustrate the species-specific nature of landscape resistance and suggest the potential for niche separation along gradients in resistance to differential reproductive strategies.

Fluvial geomorphological processes (e.g., hillside erosion, sedimentation) vary along river network continuums, with predictable states of dynamic equilibrium leading to characteristic land forms and associated plant communities (Hupp & Osterkamp, 1996). These processes present different opportunities and constraints on germination and establishment, and different modes of reproductive ecology may be associated with different riparian zones. In addition to differences in climate niche, we hypothesize that *P. angustifolia*'s more generalist reproductive strategy may be contributing to its broader distribution along all river orders. While both species reproduce sexually, *P. angustifolia* also exhibits prolific clonal reproduction (Eckenwalder, 1996). Common garden (Schweitzer et al., 2002) and wild site studies (Gom & Rood, 1999) comparing clonality among *Populus* species have found that section Tacamahaca (*P. angustifolia*, *P. balsamifera*) produces $2N$ the number of ramets as species in section Aigeiros (*P. fremontii*, *P. deltoides*). This greater clonality is likely adaptive to the higher elevation, lower-order reaches where *Tacamahaca* ssp. typically occur (Gom & Rood, 1999). Narrow, headwater streams characterized by steep gradients, decreased sinuosity, low frequency of predictable flooding and minimal sediment available for nursery formation (Hupp & Osterkamp, 1996; Samuelson & Rood, 2004) confer low opportunity for seedling establishment and greater selection pressure for clonality. In contrast, successful seedling establishment in *P. fremontii* requires predictable spring flooding coupled with low gradient channels, broad floodplains and ample
FIGURE 5 Cumulative resistant kernel density map illustrating genetic connectivity corridors. Gene flow in *P. angustifolia* is facilitated by riparian network connectivity, and differentiation is driven by terrestrial upland resistance and cumulative differences in climate. Warm colours indicate high expected density of dispersing individuals or propagules; cool colours indicate low connectivity.
alluvial substrate for nursery formation (Rood et al., 2005). Future research is needed to elucidate how specific geomorphological characteristics may be constraining different reproductive strategies along the river network continuum, thereby contributing to niche differentiation among *Populus* species.

### 4.2 Relative importance of climate versus riparian corridors on genetic differentiation

While riparian corridors support genetic connectivity, climate gradients are a much stronger influence on genetic differentiation. Climate is a primary factor structuring plant species distributions across broad spatial scales (Araújo & Peterson, 2012; Woodward, 1987), and a wealth of studies have documented adaptive divergence along environmental clines in forest trees, even with substantial gene flow (Aitken, Yeaman, Holliday, Wang, & Curtis-McLane, 2008; Evans et al., 2014; Grady et al., 2011; McKown et al., 2014; Savolainen, Pyhájärvi, & Knürr, 2007; Slavov & Zhelev, 2010). We found that maximum climate resistance was 45% more influential in driving patterns of genetic distance compared to riparian corridors. Furthermore, variation in precipitation emerged as a much stronger predictor than temperature, similar to recent meta-analysis findings (Siepielski et al., 2017). Studies investigating differentiation among populations of riparian species often focus on temperature variation among provenances (Grady et al., 2011, 2013), with precipitation conventionally considered less of a limiting factor given groundwater availability. Notably, we found the opposite was true for *P. angustifolia*. Riparian species can be particularly sensitive to precipitation and vapour pressure deficit during drought years when increased water diversion increases depth to groundwater, leading to greater canopy dieback (Horton, Kolb, & Hart, 2001) and decreased growth (Coble & Kolb, 2012). With increasing drought and aridity (Dai, 2013), genetic variation related to water use efficiency and drought tolerance will likely play an increasingly important role in buffering riparian tree stress and ultimately structuring adaptive variants across the landscape.

*Populus angustifolia* encounters significant variation in both quantity and timing of precipitation across its range. Genetic distance exhibited a strongly concave relationship with cumulative differences in precipitation seasonality (Figure 3b), consistent with low levels of resistance within regions of similar precipitation seasonality and sharp increases across regions. In Nevada’s Great Basin region, western Utah and southern Idaho, ~70% of annual precipitation falls during winter months; in contrast, New Mexico, eastern Colorado, western Montana and southern Alberta receive the majority of annual moisture during summer months (ClimaTemps 2015; Lowe, 1964; Trimble, 1999). The Sonoran Desert and Mogollon Rim regions of southern and central Arizona lie at the confluence of these major precipitation regimes and receive equal contributions from summer and winter precipitation. In addition to seasonality, the quantity of summer precipitation also varied substantially across *P. angustifolia*’s range, from 6 cm in Nevada up to nearly 4 x that (23 cm) in Arizona. Genetic distance increased linearly with cumulative differences in summer precipitation (Figure 3b). Variation in quantity and timing of precipitation are strong predictors of phylogeographic structure in *P. angustifolia*. Our research supports a growing number of studies finding that environmental constraints may be just as common as neutral processes in driving genomewide neutral differentiation patterns; this may be particularly true when selection acts on establishment (Orsini et al., 2013; Sexton et al., 2014).

In cottonwood, climate is closely linked to sexual regeneration and establishment through its influence on timing of flood events. Successful recruitment requires that flowering phenology and seed release coincide with the availability of flood-scoured virgin habitat, coupled with gradual floodwater recession to match seedling root growth (Rood et al., 2005). Variation in timing of seasonal precipitation may exert strong selection pressure against nonlocal migrants if the narrow window of seed viability fails to correspond with local flood events, thereby influencing genomewide patterns of neutral phylogeographic structure.

Similar to our findings, several other recent studies also found seasonal precipitation to be the primary climatic factor influencing genetic differentiation, including the Inner Mongolia shrub species *Reaumuria soongorica* (Yang, Cushman, Song, Yang, & Zhang, 2015) and those of the *Caragana* genus complex (Yang et al., 2013), as well as *P. fremontii* in the southwestern USA (Cushman et al., 2014). These findings point to an emerging trend that cumulative differences in seasonal precipitation are a primary ecological driver constraining genetic connectivity from shrubs to tree species and across habitats as varied as deserts of central Asia to riparian forests of the southwestern USA.

#### 4.3 Isolation by environment versus landscape resistance

To better understand how climate has influenced patterns of genetic distance in cottonwood, we compared IBE (both regional climate
clustering and pairwise climate distance), IBR (landscape resistance), and IBD (geographic distance) models. Simultaneous comparison of alternative hypotheses within a single RCM framework is essential to avoid spurious selection of highly correlated alternative hypotheses (Castillo et al., 2014; Cushman & Landguth, 2010; Cushman et al., 2013). Indeed, while individual Mantel tests demonstrated support for IBD and IBE, neither of these models was independently supported after accounting for IBR (Fig. S14). While clustering regions by similar climate is a common strategy used to inform seed sourcing and restoration (Doherty, Butterfield, & Wood, 2017), environmental similarity does not necessarily equate with genetic similarity nor shared evolutionary history. Restoration stock from similar climates may exhibit phenotypic similarities due to convergent evolution (homoplasy); however, this approach risks confining unique ancestral lineages into a single gene pool. Modelling environmental and genetic variation as continuous gradients (Cushman, Gutzweiler, Evans, & McCarigal, 2010) enables identification of the factors that both promote and constrain genetic connectivity, thereby facilitating management of not only evolutionary endpoints (current genotypes), but also maintenance of the evolutionary processes (e.g., gene flow) that have shaped phylogeographic patterns and will support evolutionary potential into the future.

4.4 Genetic connectivity–diversity relationships

An additional goal of our study was to test predictions of the central–marginal hypothesis. Several recent simulation studies have demonstrated support for connectivity–diversity relationships, finding strong, positive correlations between connectivity, as a function of cumulative resistant kernel values, and both genetic diversity (Macdonald et al., in review; Wasserman et al., 2012, 2013) and population density (Puyraudeau, Cushman, Davidson, & Madappa, 2017). Our study is the first to our knowledge to empirically demonstrate that genetic diversity increases as a function of connectivity (Figure 6, Fig. S16). While our data exhibit positive trends, supporting the central–marginal hypothesis, our study is constrained by small sample size in low connectivity regions. To assess robustness, we reanalysed relationships removing the single low connectivity–diversity sample. While both correlative strength and significance decreased for all site-based diversity metrics, relationships remained marginally significant. Interestingly, strength and significance increased for individual-based $H_o$. Given that we sampled some of the most remote margins of $P. \text{angustifolia}$’s distribution and still detected substantial diversity and connectivity, this leads us to believe that low connectivity–diversity sites are rare on the landscape. Thus, we consider the "outlier" site to be an anomaly reflecting $P. \text{angustifolia}$’s natural distribution. Located in the Santa Catalina Mountains near Tucson, AZ, the outlier site (ML) is extremely remote, separated both geographically and ecologically from other cottonwood. This sky island habitat sits near the species’ southernmost distribution limit, surrounded on all sides by the Sonoran Desert. Had we been able to discover more sites representing the lower distribution, we expect the strong, positive correlations would have been maintained.

These results raise the question: How effective is genetic connectivity as a metric for predicting genetic diversity? Our data suggest the strength of this relationship may largely depend on a species’ ecology and evenness of available sampling sites spanning the full distribution of the connectivity-diversity spectrum. With $P. \text{angustifolia}$, its particular life-history characteristics (e.g., long lifespan, long-distance dispersal) appear to be maintaining moderate-to-high genetic connectivity throughout most of its range. $Populus$ genets may persist for hundreds to thousands of years through asexual regeneration (Mitton & Grant, 1996; Mock, Rowe, Hooten, DeWoody, & Hipkins, 2008); thus, high levels of heterozygosity may be maintained through clonal propagation, and even rare long-distance dispersal events may be enough to preserve the "appearance" of substantial connectivity. For example, small relict stands in Nevada’s remote sky island mountains showed surprisingly high genetic connectivity and diversity; however, this likely reflects past conditions when $P. \text{angustifolia}$ was more abundant in this region. While these factors may help buffer cottonwood from environmental change, these same life-history characteristics are also likely contributing to the weaker connectivity–diversity relationships detected in our study, as compared to the simulation studies of obligate-outcrossing species noted above.

4.5 Conservation management recommendations

Our model informs conservation efforts in two key ways. First, maintenance of low resistance riparian corridors is essential for preserving cottonwood genetic connectivity by supporting the neutral processes (e.g., gene flow) that maintain evolutionary potential (i.e., genetic variability that natural selection can act upon) (Graudal et al., 2014). Second, precipitation gradients are a primary factor driving genetic differentiation. Because precipitation gradients have likely selected for locally adapted phenotypes, managers should avoid transferring restoration stock across regions with strongly differentiated precipitation patterns. It is important to note that genetic distance does not necessarily reflect contemporary gene flow. Particularly with long-lived clones, genetic patterns may reflect past conditions when trees may have been more abundant and connected, thereby giving the appearance of high connectivity despite much reduced present-day gene flow. While connectivity and differentiation provide important measures to guide conservation actions, they should be part of a comprehensive management plan including other critical targets like effective population size.

Comparing connectivity models across cottonwood species suggests important differences among river orders in their capacity to support sexual regeneration and hence adaptive responses to environmental change. Higher-order rivers supporting larger populations should be less vulnerable than headwater reaches where adaptive responses may be less efficient due to smaller $N_e$ and fewer opportunities for sexual regeneration. Following warming and drying since the LGM, southern populations of cottonwood have declined, contracting to lower-order rivers; conversely, northern populations have expanded into broad valleys along higher-order rivers (Evans et al.,
2015). Surprisingly, we found no evidence of greater connectivity among northern (χ² = 40.285) compared to southern (χ² = 40.583) populations (p = .9493; two-sample, two-tailed homoscedastic t test). Although genetic diversity and connectivity in *P. angustifolia* appear to be fairly resilient to reductions in population size, increasing climate stress poses the greatest risk to southern populations and small, isolated stands at range margins (e.g., Nevada sky islands). Hydrological models suggest many permanently flowing western rivers will become intermittent with climate change (Reynolds, Shafroth, & Poff, 2015); genetic connectivity along these rivers may be particularly at risk.

Given intraspecific variation, several macroscale conservation recommendations emerged from our model. High connectivity within the core of southern and northern portions of *P. angustifolia*’s range will likely maintain substantial genetic diversity; however, the Rocky Mountains and southern Colorado Plateau aridlands pose major barriers to E-W and N-S dispersal, respectively. Maintaining dispersal corridors connecting northwestern Arizona and western Utah is critical as this region represents one of the only natural bridges between southern and northern populations (Figure 5). East of the Rockies, connectivity corridors between Colorado and both northern New Mexico and southern Wyoming are largely lacking; therefore, cottonwood along Colorado’s front range is particularly vulnerable to the impacts of isolation. As climatic stress and habitat fragmentation continue, Colorado populations may benefit from assisted gene flow (Aitken & Whitlock, 2013). High-resolution GIS raster layers of our landscape resistance and genetic connectivity models are available at http://www.sega.nau.edu/data to support local-scale conservation and restoration planning.

As climate change alters precipitation regimes, local maladaptation is predicted to increase, accompanied by reduced productivity and potential local extirpation as trees become decoupled from their climate niches (Ikeda et al., 2017). Conservation and restoration efforts should be aimed at preserving connectivity along riparian corridors to support gene flow and the maintenance of genetic variation within major precipitation regimes. Common garden experiments are needed to elucidate relationships between precipitation gradients and adaptive phenotypic variation (e.g., drought tolerance, productivity). Future research investigating contact zones, where strong gradients across precipitation ecoregions are likely to generate “evolutionary hotspots” of genetic diversity and divergence, may be particularly informative for understanding adaptive divergence among regions (Vandergast et al., 2013; Whitney, Randell, & Riesenberg, 2010). Studies of landscape genetic connectivity provide valuable data to support management of ecological variants and the evolutionary processes that maintain the capacity for species’ resiliency in the face of rapid global change (Moritz, 2002).

**ACKNOWLEDGEMENTS**

This research was supported by NSF-IGERT Fellowships (HM Bothwell, LM Evans), NSF GK-12 Stem and NPS George Melendez Wright Climate Change Fellowships (HM Bothwell), NSF FIBR grant DEB-0425908 (TG Whitham, GJ Allan), NSF MacroSystems grant DEB-1340852 (GJ Allan, SA Cushman, TG Whitham) and NSF DBI-1126840 (TG Whitham, SA Cushman) for establishing the Southwest Experimental Garden Array. We are grateful to the Cottonwood Ecology & Community Genetics Laboratory group, S DiFazio, and five anonymous reviewers for valuable feedback. We also thank R Durben, C Sanfiorenzo-Barnhard and C Van-Camp for assistance with field collections and Bandelier National Monument, Great Basin National Park, and Grand Teton National Park for collection permits and the opportunity to contribute research to continued conservation of these protected areas.

**DATA ACCESSIBILITY**

Genotype data (sample IDs, geographic coordinates, microsatellite allele calls), genetic distance matrices and GIS raster layers for the landscape resistance and genetic connectivity models are available from the Southwest Experimental Garden Array data archive (http://www.sega.nau.edu/data).

**AUTHOR CONTRIBUTIONS**

H.M.B. and S.A.C. designed the study; H.M.B., S.A.W., E.I.H.-G. and L.M.E. collected field data and performed laboratory analyses; H.M.B. conducted data analyses; G.J.A., T.G.W. and S.A.C. provided valuable discussions and debates on ecological theory; and H.M.B. led writing with all authors contributing to the writing and revision process.

**ORCID**

Helen M. Bothwell http://orcid.org/0000-0003-0916-8355

Scott A. Woolbright http://orcid.org/0000-0002-7886-1009

**REFERENCES**


