RESEARCH PAPER

Journal of Biogeography

WILEY

Contemporary range position predicts the range-wide pattern of genetic diversity in balsam poplar (*Populus balsamifera* L.)

Andrew V. Gougherty[1] (iD) | Vikram E. Chhatre[2] | Stephen R. Keller[3] | Matthew C. Fitzpatrick[1]

[1]Appalachian Lab, University of Maryland Center for Environmental Science, Frostburg, MD, USA

[2]Wyoming INBRE Bioinformatics Core, University of Wyoming, Laramie, WY, USA

[3]Department of Plant Biology, University of Vermont, Burlington, VT, USA

Correspondence
Andrew V. Gougherty, Department of Botany, University of British Columbia, Vancouver, B.C., Canada.
Email: andy.gougherty@botany.ubc.ca

Abstract

Aim: Patterns of genetic diversity within species' ranges can reveal important insights into effects of past climate on species' biogeography and current population dynamics. While numerous biogeographic hypotheses have been proposed to explain patterns of genetic diversity within species' ranges, formal comparisons and rigorous statistical tests of these hypotheses remain rare. Here, we compared seven hypotheses for their abilities to describe the geographic pattern of two metrics of genetic diversity in balsam poplar (*Populus balsamifera*), a northern North American tree species.

Location: North America.

Taxon: Balsam poplar (*Populus balsamifera* L.).

Methods: We compared seven hypotheses, representing effects of past climate and current range position, for their ability to describe the geographic pattern of expected heterozygosity and per cent polymorphic loci across 85 populations of balsam poplar. We tested each hypothesis using spatial and non-spatial least-squares regression to assess the importance of spatial autocorrelation on model performance.

Results: We found that both expected heterozygosity and per cent polymorphic loci could best be explained by the current range position and genetic structure of populations within the contemporary range. Genetic diversity showed a clear gradient of being highest near the geographic and climatic range centre and lowest near range edges. Hypotheses accounting for the effects of past climate (e.g. past climatic suitability, distance from the southern edge), in contrast, had comparatively little support. Model ranks were similar among spatial and non-spatial models, but residuals of all non-spatial models were significantly autocorrelated, violating the assumption of independence in least-squares regression.

Main conclusions: Our work adds strong support for the "Central-Periphery Hypothesis" as providing a predictive framework for understanding the forces structuring genetic diversity across species' ranges, and illustrates the value of applying a robust comparative model selection framework and accounting for spatial autocorrelation when comparing biogeographic models of genetic diversity.

KEYWORDS
central/peripheral populations, genetic diversity, last glacial maximum, *Populus balsamifera*, single nucleotide polymorphisms, species distribution models

1 | INTRODUCTION

Understanding the processes shaping the macroscale pattern of genetic variation within species' ranges has been a pervasive goal of population genetics and ecology (Eckert, Samis, & Lougheed, 2008; Petit et al., 2003). Population-level genetic diversity may be shaped by historical and/or contemporary features of a species' biogeography – from past migration and shifts in population size, to recent population dynamics and response to environmental change. Because of the importance of genetic diversity to understanding the past, current and (potentially) future dynamics of species' ranges, numerous hypotheses have been proposed to explain where populations may be expected to have the highest genetic diversity. These hypotheses often fall into two broad categories: (a) *contemporary range position* – which emphasizes the relative position of populations within the contemporary geographic or climatic range and (b) *past climate effects* – which emphasize the proximity of populations to glacial refugia and effects of migration since the last glacial maximum.

(i) *Contemporary range position.* Hypotheses that emphasize the position of populations within the contemporary range often posit that genetic diversity is related to the proximity of populations to the centre of the range. One of the most commonly tested versions of this hypothesis, the central-periphery hypothesis (CPH; also known as the central-marginal hypothesis), predicts that populations near the geographic (or climatic; Lira-Noriega & Manthey, 2014) centre of the range have greater genetic diversity than populations near a range edge (Eckert et al., 2008). The CPH is based on an abundant-centre view of species' ranges, where population abundance is expected to be greatest near the geographic or climatic centre of the range due to its presumed proximity to the species' environmental optimum (Brown, 1984; although the generality of the abundant-centre model has been questioned, see Sagarin & Gaines, 2002). High population abundance near the geographic/climatic centre is expected to coincide with high effective population sizes (N_e) and elevated gene flow resulting in low genetic differentiation and high within-population diversity (Kirkpatrick & Barton, 1997). Low abundance at the range edge is expected to have opposite effects – lower gene flow, greater differentiation and lower within-population genetic diversity (Bridle & Vines, 2007; Hampe & Petit, 2005). Support for the CPH in the literature is mixed. While one review found that as many as 64.2% of studies that tested the CPH found support in favour of the hypothesis (Eckert et al., 2008), another recent review found support in fewer than 50% of studies (Pironon et al., 2017). In a test of the climatic CPH, Lira-Noriega and Manthey (2014) found that population genetic diversity of 40 species of various taxa could be better explained by distance from species' climatic niche centroid than by the geographic distance from the range centre.

(ii) *Past climate effects.* Past climate and demographic history may also shape patterns of contemporary genetic diversity within species' ranges. Most of these hypotheses arise from the concept that, following the last glacial maximum (LGM), species migrated poleward out of glacial refugia to fill their current ranges. Poleward migration is often expected to result in decreasing genetic diversity away from refugial locations, due to repeated founding events along migration routes (Excoffier, Foll, & Petit, 2009; Hewitt, 2000). The exact geographic pattern of genetic diversity that expansion from low-latitude refugia could generate, has spurred numerous hypotheses. In its simplest form, species that have undergone a strictly poleward migration following the LGM may have a latitudinal gradient in genetic diversity (Hampe & Petit, 2005). Not all species, however, migrated strictly poleward following the LGM, and some may have used refugia near (or north of) the southernmost glacial extent or in microrefugia (Anderson, Hu, Nelson, Petit, & Paige, 2006; Rull, 2009). Species may have migrated east, west or from multiple directions to fill their current ranges (e.g. Williams, Shuman, Webb, Bartlein, & Leduc, 2004), potentially obscuring a simple latitudinal gradient in genetic diversity. Refugial locations identified using distribution models have often shown elevated levels of genetic diversity (e.g. Carnaval, Hickerson, Haddad, Rodrigues, & Moritz, 2009; Yannic et al., 2013). Relatedly, landscape age (i.e. time since the landscape was last glaciated) has also been found to be a significant predictor of genetic diversity in northern plants, where recently deglaciated landscapes often harbour lower genetic diversity (Stewart et al., 2016). Past migration may also affect current population structure, and the level of mixing among historically isolated populations. In such cases, the relative affinity of current populations to ancestral groups, and the amount of mixing of ancestral groups within current populations (i.e. genetic admixture) may also be aligned with genetic diversity. Despite these past climate hypotheses being based on a similar mechanism (i.e. post-glacial migration into areas that have become newly climatically suitable), the patterns of genetic diversity predicted by each hypothesis may differ depending on the species current and past distributions (e.g. if the current range was glaciated or not during the LGM), migration direction and dispersal characteristic (e.g. propensity for long-distance dispersal).

Although genetic diversity within species' ranges may be shaped by both contemporary range dynamics and past climate, their relative roles are rarely formally compared. This has led some authors (e.g. Vucetich & Waite, 2003) to suggest that studies often make uncritical assumptions about the roles of contemporary and past effects on genetic diversity within species' ranges. Furthermore, many studies which formed the basis for the CPH (as discussed in Eckert et al., 2008; Pironon et al., 2017) and other hypotheses, often did not sample species' entire ranges, and rarely accounted for potential effects of spatial autocorrelation in the pattern of genetic diversity. Taken together, failure to test the effects of both past climate and contemporary range positions and account for range-wide spatial autocorrelation can leave an incomplete understanding of the drivers of genetic diversity within species' ranges.

In this study, we attempt to disentangle historical and contemporary landscape drivers of within-population genetic diversity of balsam poplar (*Populus balsamifera* L.), a northern broad-leaf tree species.

Previous work has shown a latitudinal gradient in genetic diversity within balsam poplar, consistent with range expansion from a southern refugia (Breen, Murray, & Olson, 2012; Keller, Olson, Silim, Schroeder, & Tiffin, 2010). Here, we extend this work using a larger set of sample populations to test multiple hypotheses representing both contemporary climate/range positions and past climate. Specifically, we compared seven hypotheses for their ability to describe the range-wide pattern of genetic diversity in balsam poplar. These included (a) the geographic central-periphery hypothesis (CPH), (b) the climatic CPH, (c) distance from the southern range edge, (d) past climatic refugia effects, (e) landscape age, (f) a model integrating the distance from the southern range edge with the CPH and (g) a model reflecting population structure and admixture. We also explored the effects of accounting for spatial autocorrelation on model rankings and coefficient estimates, as this helped to ensure that model assumptions are being met and coefficient estimates are unbiased.

2 | MATERIALS AND METHODS

2.1 | Study species

Balsam poplar (*Populus balsamifera* L.) is a wide-ranging boreal tree and the northernmost deciduous tree species in North America (Zasada & Phipps, 1990). Balsam poplar tends to be an early successional species found in mesic environments along floodplains, near streams and other waterways. It is relatively short-lived (rarely living longer than 200 years), fast-growing and fast to reach reproductive maturity. Being wind dispersed, balsam poplar is capable of long-distance dispersal and, like other *Populus* species, can reproduce vegetatively to form clonal stands. Like other northern tree species, balsam poplar's current range was nearly entirely glaciated during the LGM and species distribution models (SDMs) suggest the presence of refugia in the Central Rocky Mountains and possibly north of the ice sheets in Alaska (Breen et

al., 2012; Levsen, Tiffin, & Olson, 2012). Genetic studies to date suggest that if a northern refugium was present, it left no signature in the current pattern of genetic diversity, which suggests expansion from refugia in the south (Breen et al., 2012; Keller et al., 2010). Balsam poplar's large contemporary geographic range, occurrence over broad climatic gradients and post-glacial migration history make it an exemplary species to test the effects of contemporary and past drivers on genetic diversity.

2.2 | Genetic diversity

We used allele frequencies from single nucleotide polymorphisms (SNPs) to calculate two within-population metrics of genetic diversity that provide complimentary but distinct information on the genetic diversity within populations: average expected heterozygosity (H_{exp}) and per cent polymorphic loci (%P). Expected heterozygosity ($H_{exp} = 2pq$) indicates the relative evenness of allele frequencies at biallelic SNP loci, whereas %P summarizes the fraction of all variable loci (range-wide) that are polymorphic within a given focal population. Both metrics were calculated by integrating existing (Keller et al., 2010; Figure 1) and new (Chhatre et al., 2019) population genomic datasets. The Keller et al. (2010) dataset consisted of 412 single nucleotide polymorphisms (SNPs) identified by sequencing an initial discovery panel of 15 individuals (1 per population, sampled throughout the range). These SNPs were then used in targeted Sequenom genotyping assays to genotype 474 individuals across a range-wide collection of 34 populations. SNPs were randomly selected without regard to function or genomic regions under selection to quantify the background level of genetic diversity in balsam poplar. Further details on SNP genotyping can be found in Keller et al. (2010). The new SNP dataset was collected across 437 individuals from 51 additional populations using genotyping by sequencing (GBS) following the Elshire et al. (2011) protocol. Loci were filtered to remove low-quality variants (non-biallelic, minGQ <95, heterozygote excess and site missingness >20%). Full details of GBS library
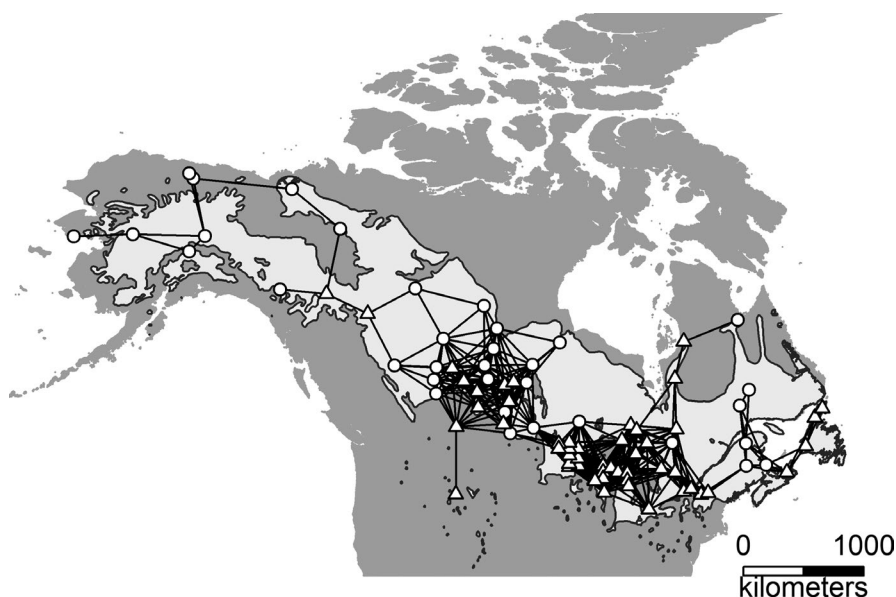
FIGURE 1 Map showing location of newly sampled *Populus balsamifera* populations (triangles; Chhatre et al., 2019), populations from Keller et al. (2010) (circles) and the neighbourhood network used in spatial analyses (black lines; see Section 2). The native North American range of *P. balsamifera* is shown in white (Little, 1971)

preparation, SNP variant calling and filtering are described in Chhatre et al. (2019).

Because SNPs in the Keller et al. (2010) dataset were identified from an initial discovery panel and therefore reflect an ascertainment bias on the site frequency spectrum (Nielsen, Hubisz, & Clark, 2004), we applied a minor allele frequency (MAF) filter to the GBS SNPs to ensure H_{exp} and %P were comparable between the two datasets. Specifically, we filtered out SNPs with an MAF below 0.0333 (1/30) from the GBS dataset, which is equivalent to 1/2N diploid individuals (where $N = 15$) used for SNP discovery in Keller et al. (2010). After filtering, this left 24,087 GBS-SNPs to calculate the diversity metrics. The number of SNPs per chromosome was proportional before and after applying the MAF ($r = .966$, $p < .01$), as were both diversity metrics (H_{exp}: $r = .87$, $p < .01$, %P: $r = .95$, $p < .01$).

To ensure that the different genotyping strategies in Keller et al. (2010) and Chhatre et al. did not confound results, we conducted the analyses described below using both the combined dataset (i.e. both Keller et al. and Chhatre et al.) and the previously published Keller et al. (2010) dataset (which sampled balsam poplar's entire range although more sparsely than the combined dataset). Results were largely consistent between the combined dataset and Keller et al., hence results below refer only to the combined dataset. Model results using only data from Keller et al. (2010) are shown in Table S1. We also tested for differences in genetic diversity (Mann–Whitney test) among the two datasets for populations in the centre of the range, which belong to a single large genomic cluster. We found that H_{exp} ($p = .30$) and %P ($p = .17$) were not significantly different between the two datasets – further justifying combining the two datasets.

2.3 | Landscape variables, climate data and occurrences

We calculated eight landscape variables to use as predictors of genetic diversity (Table 1), including four variables representing balsam poplar's contemporary range and four representing effects of past climate: (a) distance from the geographic range centre, (b) distance from the range edge, (c) climatic suitability, (d) climatic distance from the climatic niche centroid, (e) climatic stability since 22 kya, (f) climatic variability since 22 kya, (g) distance from the southern range edge and (h) landscape age since the last glaciation.

We used climate data from Lorenz, Nieto-Lugilde, Blois, Fitzpatrick, and Williams (2016) to parameterize SDMs and to calculate the climatic niche centroid. This climatic dataset includes seamless and debiased climate simulations from 22 kya to the 21st century in 500-year intervals, downscaled to a resolution of 0.5°. We chose six climate variables (summer and winter mean temperature and precipitation, annual precipitation variability and average evapotranspiration ratio (actual/potential evapotranspiration)) from the Community Climate System Model to parameterize the models. Variables were chosen because of

TABLE 1 Landscape variables used in models of expected heterozygosity and per cent polymorphic loci in *Populus balsamifera*

Model	Variable description	Variable abbreviation
Geographic CPH	Distance from the geographic range edge	geoEdge
	Distance from the geographic range centre	geoCentre
Southern edge + CPH	Distance from the geographic range centre	geoCentre
	Distance from the southern range edge	southernEdge
Climatic CPH	Current climatic suitability	suitability
	Climatic distance from the climatic range centre	climDist
Past climate stability	Climatic stability since LGM	stability
	Climatic variability since LGM	stabilitySD
Landscape age	Landscape age	landAge
Distance from southern edge	Distance from the southern range edge	southernEdge
Population structure/ mixing	Admixture index	mix
	Highest average ancestry coefficient	maxCluster

their potential importance in limiting the range of balsam poplar and lack of strong correlation between variables ($|r| < .75$).

Occurrences of balsam poplar were collected from online databases (Gbif.org, 2019), the US and Canadian forest inventory programmes (Gillis, Omule, & Brierley, 2005; Woudenberg et al., 2010) and records from the literature (Soolanayakanahally, Guy, Silim, Drewes, & Schroeder, 2009). Occurrences far outside the known North American range of balsam poplar (Little, 1971) were removed. To reduce the spatial and climatic bias of the occurrence records, we thinned the points in both geographic and multidimensional climate space, similar to the approach described in Varela, Anderson, García-Valdés, and Fernández-González (2014). Briefly, first the occurrence points were thinned to 1 per 0.5° grid cell of the climate data. Next, a principal component analysis (PCA) was conducted on climate data extracted at occurrence points. The first two components of the PCA were then plotted on a grid with a resolution of 0.2 units and one occurrence was randomly selected per PCA grid cell. After removing outliers and geographic/environmental thinning, 464 occurrences remained.

2.4 | Species distribution model

We used SDMs to calculate three of the landscape predictors: current climatic suitability for balsam poplar, and climatic stability and variability since 22 kya. We used an ensemble model to predict balsam poplar's current and past distribution using the 'biomod2' package (Thuiller, Lafourcade, Engler, & Araújo, 2009) in R. Within the ensemble model, we used six algorithms including generalized linear

models, boosted regression trees, generalized additive models, flexible discriminant analysis, multiple adaptive regression splines and random forest. We used 5-fold cross-validation iterated twice to validate models, where occurrence data were split into five subsamples and models were trained with four of the subsamples (80% of the data) and tested with the remaining subsample. Model discrimination ability was tested with true skill statistic (TSS) and the final ensemble prediction was calculated as the TSS-weighted mean of all models with TSS above 0.70. Each fold of each algorithm had a TSS above 0.7 (average: 0.83, SD: 0.04), so each was included in the final ensemble. The model was then projected to each of 45 time periods between current climate and climate at 22 kya, in 500 year intervals (Lorenz et al., 2016). The temporal resolution of the Lorenz data (500-year intervals) allowed for finer assessment of past climate effects than studies limited to snapshot climate predictions for only the LGM and mid-Holocene.

Climatic stability (sensu Ortego, Gugger, & Sork, 2015; Yannic et al., 2013) was calculated as the sum of climatic suitability through time, whereas variability was calculated as the standard deviation of climatic suitability through time. These metrics provide a measure of how the climatic suitability of balsam poplar has changed over the past 22 ky. Areas that were glaciated during a given time period (based on maps by Dyke, Moore, & Robertson, 2003) were not included in the calculations. We also calculated landscape age, similar to Stewart et al. (2016), using glacial data from Dyke et al. (2003) aligned to the 45 time periods. Shapefiles of glacial extent were rasterized to match the scale, resolution and projection of the climate data. Landscape age was calculated as the time since the landscape was most recently glaciated.

2.5 | Geographic and climatic centrality

We calculated three metrics representing the position of populations in balsam poplar's current geographic and environmental ranges: distance from the geographic range edge, distance from the geographic range centre and climatic distance from the climatic niche centroid. Distance from the range edge was calculated by generating an alpha hull around occurrence records and calculating the distance between each population and the nearest edge. Alpha hulls are similar to convex hulls and are recommended as a way to decrease the bias and spatial error associated with convex hulls when estimating species range polygons (Burgman & Fox, 2003). Distance from the geographic range centre was calculated as the geographic distance from the centroid of the alpha hull, similar to that done by Lira-Noriega and Manthey (2014) and Dallas, Decker, and Hastings (2017). The alpha hull was also used to calculate population distance from the southern edge.

We used Mahalanobis distance as a metric of population distance from the climatic niche centroid. Mahalanobis distance is a measure of the multivariate distance between climate extracted at each population location and the average climate of all balsam poplar

occurrences. Mahalanobis distances account for correlation among variables by scaling the distances by the covariance between climate variables. The covariance matrix and average climate were based on the climate at the climatically thinned balsam poplar locations. Climate variables were the same as those used for the distribution models.

2.6 | Population structure and admixture

In addition to effects of geographic/climatic centrality and past climate, we also tested for effects of population structure and admixture among genetic clusters on genetic diversity. For the previously published dataset, we used admixture proportions from Keller et al. (2010). For the new MAF-filtered GBS dataset, we estimated admixture proportions using ADMIXTURE 1.30 (Alexander, Novembre, & Lange, 2009) and chose $K = 3$ for consistency with Keller et al. (2010). The three inferred genetic clusters in the GBS data were spatially congruent with the previous dataset – both showed a distinctive eastern cluster, a large central cluster and a northern cluster that tended to be well-mixed with the central cluster (Figure S1). Using these admixture proportions, we calculated a population-level index of admixture similar to that done by Ortego et al. (2015). To do so, first we averaged admixture proportions across individuals within populations and calculated the standard deviation of the average proportions. Next, we rescaled this value between 0 and 1, such that the index was 1.0 when populations were evenly mixed among clusters and 0 when populations were entirely affiliated with a single cluster. For each population, we also determined which of the three clusters had the highest average admixture proportion.

2.7 | Models and statistical analyses

We assessed the local and global spatial pattern of H_{exp} and %P using Moran's Index (I), a measure of spatial autocorrelation, where −1.0 indicates perfect dispersion and 1.0 indicates perfect clustering. Correlograms of Moran's I for H_{exp} and %P were estimated in 100 km increments. Significance was determined for both the correlograms and global statistic by comparing the observed statistic to 999 random permutations.

Using the eight landscape variables, we compared statistical support for models representing seven hypotheses listed in Table 1. For each hypothesis, we created spatial and non-spatial models to assess the effect of accounting for spatial autocorrelation on model performance. To account for spatial autocorrelation, we used conditional autoregressive (CAR) models, which integrated a weighted estimate of the response variable (here, the metrics of genetic diversity) at neighbouring locations, in addition to the explanatory variables, in parameterizing the model (Lichstein, Simons, Shriner, & Franzreb, 2002). Neighbourhoods were defined as all populations within 600 km of one another (Figure 1). This distance was chosen as it ensured each

population had at least one neighbour and was the approximate maximum distance of continuous significant positive spatial autocorrelation (see Section 3). A complementary set of non-spatial ordinary least squares (OLS) regression models were fit and compared with the CAR models. Models were compared using Nagelkerke R^2, Akaike information criterion (AIC) and Akaike weights, which are recommended as a way to compare AIC scores across models (Wagenmakers & Farrell, 2004). Each explanatory variable was scaled to a mean of 0 and a standard deviation of 1, to facilitate the comparison of coefficient estimates (Schielzeth, 2010). All modelling and statistical analyses were performed in R (R Core Team, 2019).

3 | RESULTS

3.1 | Spatial pattern and autocorrelation

Both metrics of genetic diversity (H_{exp} and %P) were highest near the geographic centre of the range and declined towards the latitudinal

(southern and northern) and longitudinal (eastern and western) range edges. Both H_{exp} and %P showed moderate (I = 0.28 and 0.45, respectively) but significant spatial autocorrelation ($p < .05$) among the 85 populations, indicating that adjacent populations tended to have more similar levels of genetic diversity than distant neighbours. Correlograms of Moran's I revealed that H_{exp} and %P were significantly positively autocorrelated up to ~600 km and at ~2,000 km, and significantly negatively correlated around 1,000 km and between 3,000 and 4,000 km (Figure 2b, d). Genetic diversity metrics were significantly correlated among the 85 populations (Pearson's r = .88, $p < .001$).

3.2 | Spatial models of genetic diversity

The top performing models for H_{exp} and %P were the models of population structure, and the geographic CPH, respectively, and had by far the highest Akaike weights and highest Nagelkerke R^2's (Tables 2 and 3). Inspection of the coefficients revealed that both diversity
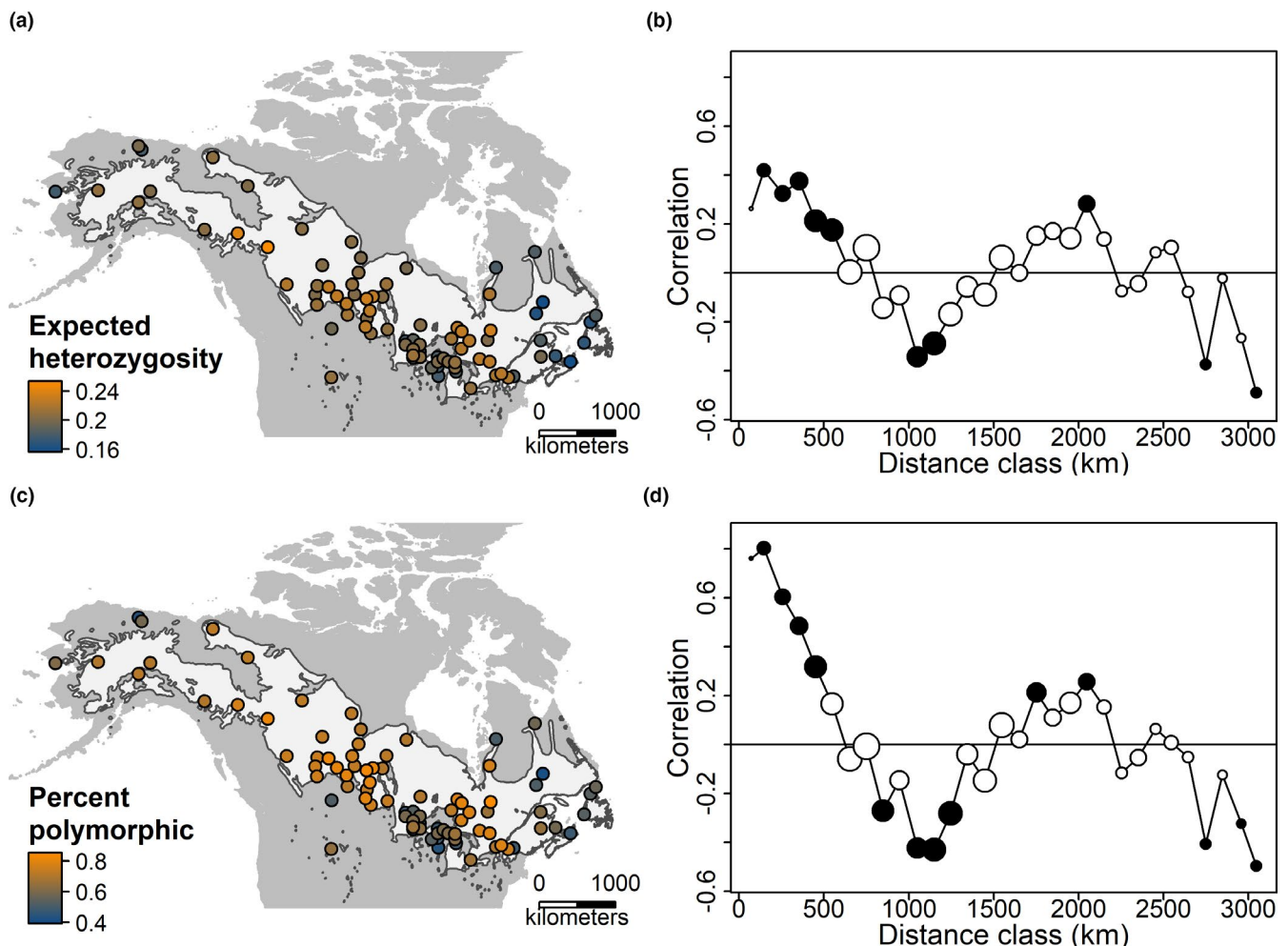
FIGURE 2 Maps and correlograms of (a, b) expected heterozygosity and (c, d) per cent polymorphic loci among 85 *Populus balsamifera* populations. Circle size in the correlograms is proportional to the number of records within each distance class and filled circles indicate significant autocorrelation at particular distance classes (two sided, $p > .975$ or $p < .025$). The native North American range of *P. balsamifera* is shown in white (Little, 1971), in (a) and (c) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Summary statistics for conditional autoregressive models for a range-wide sample of expected heterozygosity in *Populus balsamifera*, ranked by relative support

Model	Coefficient	Estimate	p-value	AIC	AIC weight	Nagelkerke R^2	Moran's I (residuals)	Moran's I p-value (residuals)
Population structure/ mixing	Intercept	0.208	<.01	−456.40	0.84	.39	0.00	.37
	mix	0.028	.06					
	maxCluster-Eastern	−0.034	<.01					
	maxCluster-Northern	−0.007	.27					
Geographic CPH	Intercept	0.205	<.01	−451.81	0.08	.35	0.07	.03
	geoEdge	0.004	.11					
	geoCentre	−0.009	<.01					
Climatic CPH	Intercept	0.208	<.01	−450.73	0.05	.34	0.05	.06
	suitability	0.006	<.01					
	climDist	−0.007	<.01					
Southern edge + CPH	Intercept	0.204	<.01	−449.63	0.03	.33	0.09	.01
	geoCentre	−0.011	<.01					
	southernEdge	−0.001	.55					
Landscape age	Intercept	0.201	<.01	−431.12	0.00	.15	0.10	.01
	landAge	−0.001	.72					
Distance from southern edge	Intercept	0.201	<.01	−431.01	0.00	.14	0.10	.01
	southernEdge	2.9E−04	.91					
Past climate stability	Intercept	0.200	<.01	−430.64	0.00	.16	0.11	.01
	stability	0.002	.35					
	stabilitySD	−0.002	.33					

metrics were greatest near the range centre and lowest near the range edge (Figure 3). Furthermore, we found that climatic distance from the climatic niche centroid was correlated with distance from the range centre ($r = .30$, $p < .01$) and negatively correlated with distance from the range edge ($r = −.51$, $p < .01$) – indicating that populations near the range centre tended to be near the niche centroid, whereas populations near a range edge tended to be more distant (Figure S2).

In general, the spatial models incorporating past climate had less support than models incorporating current climate/range positions. For both H_{exp} and %P, past climate CAR models (i.e. landscape age, distance from the southern edge, past climatic stability) consistently had lower Akaike weights (i.e. all models had weights near zero) and tended to have lower Nagelkerke R^2's compared with models representing contemporary range positions. Coefficient estimates for all variables in the past climate models, other than the intercepts, were not significant ($p > .05$).

3.3 | Non-spatial models of genetic diversity

The rank (based on Akaike weights) of non-spatial OLS models for both H_{exp} and %P was similar to that of the spatial models. The top model for both H_{exp} and %P was the population structure model and the geographic CPH respectively (Tables S2 and S3). Coefficient

estimates for these models again showed higher diversity near range centre and lower near the range edges. The model integrating distance from the southern range edge and the CPH had the second highest support for %P (Akaike weight = 0.15), but like the CAR models, only the coefficient for the distance from the geographic centre was significant. Similar to the CAR models, OLS models representing past climate effects (i.e. landscape age, distance from the southern edge, past climate stability) had low support, with Akaike weights near zero.

3.4 | Residual spatial autocorrelation

Residuals of most models, spatial and non-spatial, were spatially autocorrelated, as quantified by Moran's *I*. Of the CAR models, only four models (population structure models for H_{exp} and %P, and geographic and climatic CPH for %P) had uncorrelated residuals ($p > .05$; Tables 2 and 3). Residuals of CAR models representing contemporary climate/range positions tended to have lower autocorrelation than past climate models, but all spatial models had relatively low autocorrelation (all <0.11). In contrast, autocorrelation in OLS models was often nearly as strong as the actual diversity metrics being modelled. This suggests that the assumption of independent residuals is being violated in nearly all the non-spatial models.

TABLE 3 Summary statistics for conditional autoregressive models for a range-wide sample of per cent polymorphic loci in *Populus balsamifera*, ranked by relative support

Model	Coefficient	Estimate	p-value	AIC	AIC weight	Nagelkerke R^2	Moran's I (residuals)	Moran's I p-value (residuals)
Geographic CPH	Intercept	0.694	<.01	−161.44	0.77	.43	0.04	.12
	geoEdge	0.037	.01					
	geoCentre	−0.040	<.01					
Climatic CPH	Intercept	0.705	<.01	−158.53	0.18	.41	0.02	.19
	suitability	0.038	<.01					
	climDist	−0.034	<.01					
Southern edge + CPH	Intercept	0.680	<.01	−154.76	0.03	.38	0.06	.04
	geoCentre	−0.058	<.01					
	southernEdge	−0.006	.64					
Population structure/ mixing	Intercept	0.716	<.01	−154.38	0.02	.39	0.00	.35
	mix	0.088	.32					
	maxCluster-Eastern	−0.168	<.01					
	maxCluster-Northern	−0.056	.15					
Landscape age	Intercept	0.667	<.01	−140.95	0.00	.26	0.08	.02
	landAge	−0.016	.16					
Distance from southern edge	Intercept	0.664	<.01	−139.04	0.00	.24	0.08	.02
	southernEdge	0.003	.85					
Past climate stability	Intercept	0.655	<.01	−139.73	0.00	.26	0.09	.01
	stability	0.003	.79					
	stabilitySD	−0.021	.10					
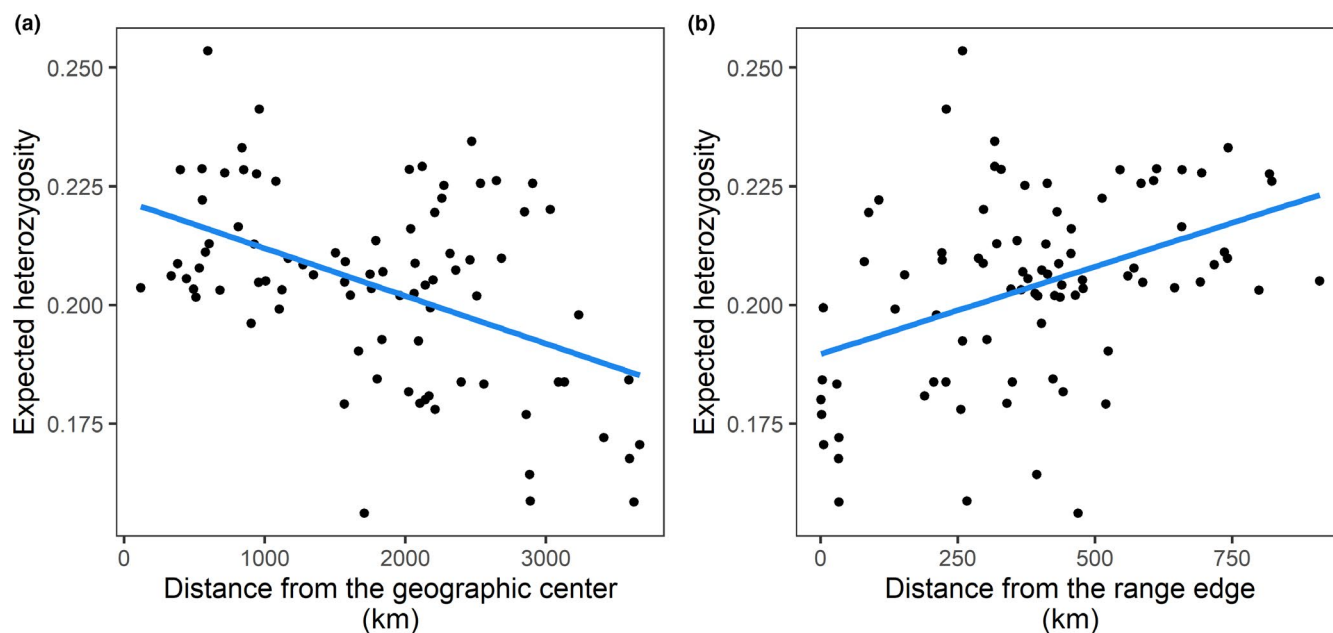
FIGURE 3 Relationship between *Populus balsamifera* expected heterozygosity and (a) distance from the geographic range centre and (b) distance from the range edge. Least-squares regression lines are shown in blue [Colour figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

Quantifying patterns of genetic diversity within species' ranges can reveal important insights into species' biogeography, effects of past climate and where populations may be best positioned to adapt to future climates. However, because these patterns can arise from multiple historic and contemporary processes, it is important to evaluate the comparative strength of different hypotheses. In this study, we compared multiple biogeographic hypotheses established in the literature and tested them in a model selection framework using the range-wide patterns of genetic diversity in balsam poplar. We found the greatest support for hypotheses representing populations' position in the contemporary geographic and climatic range, in particular the centre-periphery hypothesis (CPH) and population structure models, whereas we found comparatively little support for hypotheses that included variables for historical climate or distance from potential refugia. Our work illustrates the value of applying statistical model selection among multiple competing biogeographic hypotheses, representing both current and past climate, to better understand the landscape-scale predictors of genetic diversity across species' ranges.

4.1 | Geographic pattern of diversity

Like studies of other tree species (e.g. Walter & Epperson, 2005), we found that genetic diversity in balsam poplar populations was spatially autocorrelated over large distances (here, hundreds of kilometres). Interestingly, the strength of spatial autocorrelation did not simply decay with increased distance, but rather oscillated between (significant) positive and negative autocorrelation over thousands of kilometres, indicating that the autocorrelation spanned multiple spatial scales. This pattern seems to be the result of multiple, discontinuous hotspots of diversity in balsam poplar's sampled range, in particular near the centre of the range in Saskatchewan and in populations north of the Great Lakes region. Both the diversity hotspots and positive autocorrelation among nearby populations are likely the result of nearby populations undergoing similar processes (such as gene flow and drift) as well as shared common ancestry. Indeed, individuals from these two regions have been shown to belong to a large genetic cluster (as identified by admixture analyses; Keller et al., 2010; Figure S1) that coalescent models have shown has a large effective population size (N_e) and is the source of asymmetric migration from the centre towards the periphery of the range (Keller et al., 2010).

4.2 | CPH and the abundant-centre model

Diversity in balsam poplar was highest in the centre of the range and tended to decline towards the range edge, consistent with the CPH. The CPH is presumed to be driven by population abundance – where high abundance in the centre of the range promotes high gene flow and N_e, while low abundance at the range edge results

in isolation, reduced gene flow and ultimately low genetic diversity. Because we were unable to directly test if abundance peaked in the centre of balsam poplar's range and declined towards the edge, it remains unclear if population abundance is the ultimate driver of genetic diversity within balsam poplar's range. There is reason to suspect, however, that abundance in balsam poplar's range does not have a monotonic decline towards the edge. First, multiple studies have shown that the pattern expected by the abundant centre model is rarely observed (Dallas et al., 2017; Sagarin & Gaines, 2002). Dallas et al. (2017), for instance, showed that most North American tree species that the authors assessed (~97%) did not peak in abundance in the centre of their ranges, but rather species more often had higher abundance near the range edge than the range centre. Furthermore, balsam poplar, specifically, shows considerable variability in abundance along its range edges. Recent analyses by the US Forest Service (Prasad, Iverson, Peters, & Matthews, 2014), for instance, show relatively high balsam poplar abundance at the range edge in the upper Midwest (e.g. northern Minnesota), lower abundance near the Great Lakes and north-eastern US and very low abundance in the Rocky Mountains. This high spatial variability in abundance suggests that, like other tree species, proximity to a range edge may not be the sole driver of abundance in balsam poplar's range. Hence, other hypotheses may be necessary to explain the high diversity in the centre of balsam poplar's range irrespective of abundance.

Other processes could plausibly result in the patterns expected by the CPH. For instance, if migration following the LGM occurred mainly from the centre of the (current) range towards the range edges, rather than strictly poleward, diversity could be highest in the centre of the range and lowest near the edges. Keller et al. (2010) suggested balsam poplar had refugia in the Rocky Mountains during the LGM south of the centre of the current range, from which it expanded eastward and northward following glacial retreat. Bottlenecks and founding events along the migratory paths from the centre of the current range towards the edges likely left a gradient in genetic diversity often documented in tree species that have undergone long-distance migration (Hewitt, 2000; Petit, Bialozyt, Brewer, Cheddadi, & Comps, 2001). This hypothesized migration history may also explain the population structure observed in balsam poplar, where populations at the periphery of the range tended to belong to genetic clusters (i.e. northern and eastern clusters) that had lower diversity than the cluster at the centre of the range (Tables 2 and 3; Keller et al., 2010) and likely explains the relatively strong support for the population structure model. Although we cannot conclusively determine whether range/climatic marginality or population structure is the ultimate driver of genetic diversity (as populations structure is correlated with the distance from the range centre), both clearly indicate high diversity in the centre of the range (which coincides with a genetic cluster with high N_e) and lower diversity in the range edges (coincident with two other genomic clusters), which would be consistent with range expansion from the centre of the current range to the periphery.

The possibility that post-glacial migration left a pattern of genetic diversity similar to that expected by the CPH would be unsurprising given recent work showing that migration following the LGM in North American trees rarely left a latitudinal gradient in genetic diversity, as has often been found for European species (Lumibao, Hoban, & McLachlan, 2017). The lack of distinctive migration barriers in North America (Soltis, Morris, McLachlan, Manos, & Soltis, 2006), combined with a large ice-free area north of the southern ice margin (Brubaker, Anderson, Edwards, & Lozhkin, 2005), may have allowed North American species to fill their current ranges from multiple directions, precluding a monotonic trend in genetic diversity within the range. This is apparent in other *Populus* species, such as *P. trichocarpa* a sister species of balsam poplar, which has been shown to have low diversity in the centre of its range (Zhou, Bawa, & Holliday, 2014), possibly reflective of refugia north and south of the current range. In contrast, *P. tremuloides* has been shown to have the lowest diversity in the south-eastern portion of its range and a peak near the centre of its latitudinal range (Callahan et al., 2013). The similarity to *P. tremuloides* is particularly interesting, as balsam poplar and quaking aspen share similar current ranges and have co-occurred in the past (evidenced from the North American pollen database) – possibly suggestive that they have undergone similar migratory histories and could be under similar forces shaping their genetic diversity.

4.3 | Spatial autocorrelation

Our comparison between the CAR and OLS models revealed the importance of accounting for spatial autocorrelation when assessing the drivers of genetic diversity at large spatial scales. Although the model ranks differed only slightly among the spatial and non-spatial models, the residuals of the non-spatial models were comparatively strongly autocorrelated, indicating lack of independence and violation of model assumptions. In fact, for multiple OLS models, residuals were nearly as strongly spatially autocorrelated as the actual diversity metric being modelled. While the spatial models did not completely account for the autocorrelation (many of these models also had significant residuals autocorrelation), residuals of all spatial models were substantially less autocorrelated than the actual diversity metrics.

Although spatial effects are infrequently accounted for when assessing landscape drivers of genetic variability, when model residuals are strongly autocorrelated, spatial models should be used to ensure coefficient estimates are unbiased and OLS assumptions are not violated. Failing to account for spatial relationships of genetic variability can affect the sign and magnitude of model coefficients and the associated inference (Dormann et al., 2007). Despite spatial models requiring additional steps to be fit (e.g. defining a spatial neighbourhood), future studies should account for spatial non-independence, or at least test for (and report) the presence of spatial autocorrelation in model residuals to ensure this assumption is not being violated.

5 | CONCLUSIONS

Our analyses indicated that genetic diversity in balsam poplar reflects distances from the geographic range centre and edges, consistent with high N_e in the range centre and lower N_e towards the range edges. In general, effects of past climate were not well supported, suggesting that the main demographic centre of the species has migrated into mid-latitudes during range expansion following the LGM and maintained high diversity there, whereas edge populations show low diversity due to low N_e and/or reduced connectivity. Furthermore, our results point to the benefit of comparing multiple competing hypotheses when assessing the pattern of genetic diversity across species ranges, as well as the advantage of considering spatial effects to ensure assumptions are not violated and results are not biased.

DATA AVAILABILITY STATEMENT
Climate data and occurrences are publicly available at the cited sources. Metrics of genetic diversity and landscape variables used to fit models are available in the supplemental material.

ORCID
Andrew V. Gougherty 🆔 https://orcid.org/0000-0002-3905-8539

REFERENCES
Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Anderson, L. L., Hu, F. S., Nelson, D. M., Petit, R. J., & Paige, K. N. (2006). Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(33), 12447–12450. https://doi.org/10.1073/pnas.0605310103

Breen, A. L., Murray, D. F., & Olson, M. S. (2012). Genetic consequences of glacial survival: The late Quaternary history of balsam poplar (*Populus balsamifera* L.) in North America: Genetic consequences of glacial survival. *Journal of Biogeography*, *39*(5), 918–928. https://doi.org/10.1111/j.1365-2699.2011.02657.x

Bridle, J. R., & Vines, T. H. (2007). Limits to evolution at range margins: When and why does adaptation fail? *Trends in Ecology & Evolution*, *22*(3), 140–147. https://doi.org/10.1016/j.tree.2006.11.002

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*(2), 255–279. https://doi.org/10.1086/284267

Brubaker, L. B., Anderson, P. M., Edwards, M. E., & Lozhkin, A. V. (2005). Beringia as a glacial refugium for boreal trees and shrubs: New perspectives from mapped pollen data. *Journal of Biogeography*, *32*(5), 833–848. https://doi.org/10.1111/j.1365-2699.2004.01203.x

Burgman, M. A., & Fox, J. C. (2003). Bias in species range estimates from minimum convex polygons: Implications for conservation and options for improved planning. *Animal Conservation Forum*, 6(1), 19–28. https://doi.org/10.1017/S1367943003003044

Callahan, C. M., Rowe, C. A., Ryel, R. J., Shaw, J. D., Madritch, M. D., & Mock, K. E. (2013). Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). *Journal of Biogeography*, 40(9), 1780–1791. https://doi.org/10.1111/jbi.12115

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, 323(5915), 785–789. https://doi.org/10.1126/science.1166955

Chhatre, V. E., Fetter, K. C., Gougherty, A. V., Fitzpatrick, M. C., Soolanayakanahally, R. Y., Zalensy, R. S., & Keller, S. R. (2019). Climatic niche predicts the landscape structure of locally adaptive standing genetic variation. *BioRxiv*, 817411, https://doi.org/10.1101/817411

Dallas, T., Decker, R. R., & Hastings, A. (2017). Species are not most abundant in the centre of their geographic range or climatic niche. *Ecology Letters*, 20(12), 1526–1533. https://doi.org/10.1111/ele.12860

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., … Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, 30(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Dyke, A. S., Moore, A., & Robertson, L. (2003). Deglaciation of North America. Thirty-two maps at 1:70000000 scale with accompanying digital chronological database and one poster (in two sheets) with full map series. Geol. Surv. Can. Tech. Rep. Open File 1574.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: The central–marginal hypothesis and beyond. *Molecular Ecology*, 17(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Gbif.org. (2019). GBIF occurrence download.https://doi.org/10.15468/dl.0wxmwg

Gillis, M. D., Omule, A. Y., & Brierley, T. (2005). Monitoring Canada's forests: The national forest inventory. *The Forestry Chronicle*, 81(2), 214–221. https://doi.org/10.5558/tfc81214-2

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters: Rear edges and climate change. *Ecology Letters*, 8(5), 461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913.

Keller, S. R., Olson, M. S., Silim, S., Schroeder, W., & Tiffin, P. (2010). Genomic diversity, population structure, and migration following rapid range expansion in the balsam poplar. *Populus Balsamifera*. *Molecular Ecology*, 19(6), 1212–1226. https://doi.org/10.1111/j.1365-294X.2010.04546.x

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *The American Naturalist*, 150(1), 1–23. https://doi.org/10.1086/286054

Levsen, N. D., Tiffin, P., & Olson, M. S. (2012). Pleistocene speciation in the genus *Populus* (Salicaceae). *Systematic Biology*, 61(3), 401–412. https://doi.org/10.1093/sysbio/syr120

Lichstein, J. W., Simons, T. R., Shriner, S. A., & Franzreb, K. E. (2002). Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, 72(3), 445–463. https://doi.org/10.1890/0012-9615(2002)072[0445:SAAAMI]2.0.CO;2

Lira-Noriega, A., & Manthey, J. D. (2014). Relationship of genetic diversity and niche centrality: A survey and analysis. *Evolution*, 68(4), 1082–1093. https://doi.org/10.1111/evo.12343

Little, E. L. (1971). *Atlas of United States trees*. Washington, DC: U.S. Dept. of Agriculture, Forest Service.

Lorenz, D. J., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C., & Williams, J. W. (2016). Downscaled and debiased climate simulations for North America from 21,000 years ago to 2100AD. *Scientific Data*, 3, https://doi.org/10.1038/sdata.2016.48

Lumibao, C. Y., Hoban, S. M., & McLachlan, J. (2017). Ice ages leave genetic diversity 'hotspots' in Europe but not in Eastern North America. *Ecology Letters*, 20(11), 1459–1468. https://doi.org/10.1111/ele.12853

Nielsen, R., Hubisz, M. J., & Clark, A. G. (2004). Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. *Genetics*, 168(4), 2373–2382. https://doi.org/10.1534/genetics.104.031039

Ortego, J., Gugger, P. F., & Sork, V. L. (2015). Climatically stable landscapes predict patterns of genetic structure and admixture in the Californian canyon live oak. *Journal of Biogeography*, 42(2), 328–338. https://doi.org/10.1111/jbi.12419

Petit, R. J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., … Vendramin, G. G. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science*, 300(5625), 1563–1565. https://doi.org/10.1126/science.1083264

Petit, R. J., Bialozyt, R., Brewer, S., Cheddadi, R., & Comps, B. (2001). From spatial patterns of genetic diversity to postglacial migration processes in forest trees. *Special Publication-British Ecological Society*, 14, 295–318.

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017). Geographic variation in genetic and demographic performance: New insights from an old biogeographical paradigm. *Biological Reviews*, 92(4), 1877–1909. https://doi.org/10.1111/brv.12313

Prasad, A. M., Iverson, L. R., Peters, M. P., & Matthews, S. N. (2014). *Climate change atlas*. Delaware, OH: Northern Research Station, U.S. Forest Services. Retrieved from https://www.fs.fed.us/nrs/atlas/

R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Rull, V. (2009). Microrefugia. *Journal of Biogeography*, 36(3), 481–484. https://doi.org/10.1111/j.1365-2699.2008.02023.x

Sagarin, R. D., & Gaines, S. D. (2002). The 'abundant centre' distribution: To what extent is it a biogeographical rule? *Ecology Letters*, 5(1), 137–147. https://doi.org/10.1046/j.1461-0248.2002.00297.x

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America: Phylogeography of unglaciated eastern North America. *Molecular Ecology*, 15(14), 4261–4293. https://doi.org/10.1111/j.1365-294X.2006.03061.x

Soolanayakanahally, R. Y., Guy, R. D., Silim, S. N., Drewes, E. C., & Schroeder, W. R. (2009). Enhanced assimilation rate and water use efficiency with latitude through increased photosynthetic capacity and internal conductance in balsam poplar (*Populus balsamifera* L.). *Plant, Cell & Environment*, 32(12), 1821–1832. https://doi.org/10.1111/j.1365-3040.2009.02042.x

Stewart, L., Alsos, I. G., Bay, C., Breen, A. L., Brochmann, C., Boulanger-Lapointe, N., … Pellissier, L. (2016). The regional species richness and genetic diversity of Arctic vegetation reflect both past glaciations and current climate: Glaciation effect on arctic vegetation

diversity. *Global Ecology and Biogeography*, *25*(4), 430–442. https://doi.org/10.1111/geb.12424

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – A platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Varela, S., Anderson, R. P., García-Valdés, R., & Fernández-González, F. (2014). Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, *37*, 1084–1091. https://doi.org/10.1111/j.1600-0587.2013.00441.x

Vucetich, J. A., & Waite, T. A. (2003). Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics*, *4*(5), 639–645.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192–196. https://doi.org/10.3758/BF03206482

Walter, R., & Epperson, B. K. (2005). Geographic pattern of genetic diversity in *Pinus resinosa*: Contact zone between descendants of glacial refugia. *American Journal of Botany*, *92*(1), 92–100. https://doi.org/10.3732/ajb.92.1.92

Williams, J. W., Shuman, B. N., Webb, T., Bartlein, P. J., & Leduc, P. L. (2004). Late-Quaternary vegetation dynamics in North America: Scaling from taxa to biomes. *Ecological Monographs*, *74*(2), 309–334. https://doi.org/10.1890/02-4045

Woudenberg, S. W., Conkling, B. L., O'Connell, B. M., LaPoint, E. B., Turner, J. A., & Waddell, K. L. (2010). *The forest inventory and analysis database: Database description and users manual version 4.0 for Phase 2*. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 336 pp. https://doi.org/10.2737/RMRS-GTR-245

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., … Côté, S. D. (2013). Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change*, *4*(2), 132–137. https://doi.org/10.1038/nclimate2074

Zasada, J. C., & Phipps, H. M. (1990). *Populus balsamifera* L. *Silvics of North America*, *2*, 518–529.

Zhou, L., Bawa, R., & Holliday, J. A. (2014). Exome resequencing reveals signatures of demographic and adaptive processes across the genome and range of black cottonwood (*Populus trichocarpa*). *Molecular Ecology*, *23*(10), 2486–2499. https://doi.org/10.1111/mec.12752

BIOSKETCH

Andrew Gougherty recently obtained his PhD from the University of Maryland Center for Environmental Science. His thesis assessed various ways of using genetic information to understand species' responses to climate change. He is currently a postdoctoral researcher at the University of British Columbia.

Author contributions: AG, MF and SK conceived the ideas; VC and SK collected and processed the genetic data; AG analysed the data and led the writing, with contributions from all authors.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gougherty AV, Chhatre VE, Keller SR, Fitzpatrick MC. Contemporary range position predicts the range-wide pattern of genetic diversity in balsam poplar (*Populus balsamifera* L.). *J Biogeogr*. 2020;47:1246–1257. https://doi.org/10.1111/jbi.13811