Received: 25 June 2021 | Revised: 9 January 2022 | Accepted: 27 January 2022

DOI: 10.1111/ddi.13490

RESEARCH ARTICLE

Diversity and Distributions WILEY

Conservation prioritization based on past cascading climatic effects on genetic diversity and population size dynamics: Insights from a temperate tree species

Yang Liu[1,2]

[1]Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, British Columbia, Canada

[2]Department of Archaeology, University of Cambridge, Cambridge, UK

Correspondence
Yang Liu, Department of Forest and Conservation Sciences, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada.
Email: y.liu@alumni.ubc.ca

Funding information
University of Cambridge

Editor: Tao Deng

Abstract

Aim: The biogeographic history affects adaptive evolution by altering the pattern of genetic variation and provides the background upon which other evolutionary processes are operating. Assessment of the consequence of such eco-evolutionary interactions is crucial to population's genetic diversity maintenance. Yet, considering how biogeographic and multiple evolutionary processes interact in complex environments to alter genetic diversity in conservation planning remains unresolved.

Location: The Pacific Northwest.

Methods: Herein I document the impact of geographic landscapes and past climatic fluctuations on the evolutionary processes that drive population divergence of *Populus trichocarpa* by assembling genomic, climatic and species occurrence data.

Results: Based on inferences of demographic history of the species and tests for niche divergence, results show that the British Columbia (BC)-North and BC-South populations are ancestral and subject to more recent bottlenecks than the Oregon population. Consistently, ecological niche modelling illustrates that compared to the Last Glacial Maximum (~21 Ka), the present niche suitability of the BC-North and -South populations is overall marginally inferior to that of the Oregon population. However, genomic analysis demonstrates that the Oregon harbours the lowest genetic diversity, possibly due to an extended bottleneck experienced by the population. Model prediction indicates that in the 2050s, the impact of climate change is minor compared to the niche suitability of the species as a whole, but climate change will likely result in considerable genetic shifts over two generations especially in the Oregon and BC-South populations, which will attenuate along a northward gradient.

Main conclusions: This study underpins the long-lasting impact of demographic histories and natural selection on genetic diversity, rather than highlighting the consequence of recent demographic events (e.g. bottleneck in the Holocene epoch; ~12 Ka) potentially cascading to eroding genetic diversity and yielding conservation urgency.

KEYWORDS
biogeography, bottleneck, climate change, genetic diversity, local adaptation, population divergence, *Populus trichocarpa*

1 | INTRODUCTION

Anthropogenic climate change is causing a wide-spread maladaptation in natural ecosystems and poses a potent threat to global biodiversity (IPBES, 2019; Shaw & Etterson, 2012; Urban, 2015). Forecasting the loss of suitable habitats caused by climate change helps inform conservation interventions and provide guidelines and tools for assisted adaptation and migration (McLachlan et al., 2007; Thomas et al., 2004; Thuiller et al., 2005). Models currently used to predict species responses to climate change often utilize a correlative niche modelling approach (Thuiller et al., 2008), which assumes static environmental tolerance over time such that organisms will track their current ecological niche in the future (Araújo & Peterson, 2012). However, this type of models could spuriously forecast less future suitable habitats because species' current distributions may not adequately represent their thermal tolerance, and future climate conditions may not be analogous to current conditions (Williams & Jackson, 2007). Decades of empirical research also contradict this assumption, as evolutionary processes are also able to counteract the impact of climate change in some species (Franks et al., 2007; Karell et al., 2011), while in others, the strong selective pressures imposed by climate change might outpace a species' capability of adapting to a changing environment (Etterson & Shaw, 2001; Jezkova & Wiens, 2016; Quintero & Wiens, 2013). In fact, the feedback between adaptive and demographic processes has large consequences for a species' response to climate change (Hoffmann & Sgrò, 2011; Urban, 2015). Hence, examination of demographic, ecological and evolutionary mechanisms influencing population divergence improves our prediction of the impacts of climate change on ecologically divergent populations.

Historically, past glaciation, most notably the Pleistocene ice ages, has profoundly shaped the genetic architecture of the flora and fauna of the Pacific Northwest and created discontinuities in the geographic distribution of many plant species (Milne, 2006; Soltis et al., 1997). Such spatial patterns of glacial persistence and post-glacial colonization led to dramatic demographic shifts of populations or species, and theoretically, large populations recover to equilibrium diversity levels more slowly (Kimura, 1983 p. 204; Nei & Graur, 1984 p. 82–89). Moreover, past demography determines the strength of genetic drift [i.e. random changes in allele frequency over time; (Caballero, 1994)] and has a profound influence on efficiency of selection (Wright, 1931) as well as genetic diversity (Hewitt, 2000) such as possible genetic erosion after a strong bottleneck. In addition, fluctuations in population size due to historical dynamics of population abundance are assessed widely as an important parameter for deriving neutral expectations (Lotterhos & Whitlock, 2014). Hence, reconstructing the demographic history of populations or species is paramount to understanding environment-driven population divergence and providing useful information for conservation planning.

Long-lived organisms, such as trees, are particularly vulnerable to maladaptation due to climate change, because the long life span limits their adaptive capacity (Aitken et al., 2008; Kuparinen et al., 2010).

BOX 1 Use of *Populus trichocarpa* as a model system for genetic and evolutionary studies

Populus trichocarpa has porous species barriers, moderate genome size (*c*. 480 Mb) and abundant genomic resources, rendering it an ideal system for molecular evolutionary studies (Cronk, 2005; Jansson & Douglas, 2007; Tuskan et al., 2006). The genomic basis of adaptive variation across a wide geographic range of *P. trichocarpa* individuals has been elucidated (Evans et al., 2014), and the adaptive and neutral evolutionary forces shaping the pattern of *P. trichocarpa* population structure comprise introgression and hybridization with the congener *P. balsamifera* (Suarez-Gonzalez et al., 2018), restricted gene flow (Geraldes et al., 2014; Xie et al., 2009), natural selection (Geraldes et al., 2014) and genetic drift (population history) (Slavov et al., 2012). Extensive gene flow is facilitated by long-distance dispersal through pollen, seeds and vegetative propagules (Galloway & Worrall, 1979) with effective distances of 2–7.6 km (DiFazio et al., 2012; Slavov et al., 2009). In addition, phenotypic plasticity has been shown to be another important factor affecting trait variability and adaptability (Liu & El-Kassaby, 2019). At the genomic level, divergent selection is evident in promoters, 5'-UTR and intronic sites, whereas intergenic regions evolve under relaxed selection (Zhou et al., 2014). Genetic divergence genes of introgression-resistant regions are found to be under positive selection (Liu & El-Kassaby, 2021). A single-gene-based mechanism determines the sexes of this dioecious species, whereby impacting species evolution (Müller et al., 2020). Latitudinal allele frequency clines are also common across the genome, hinting at large-scale patterns of adaptation in the species (Slavov et al., 2012).

Cottonwoods (*Populus* spp.) are a genus of deciduous flowering trees, in which black cottonwood (*Populus trichocarpa*) represents an important model species for the long-lived perennial plants (Box 1). Genetic analysis offers evidence that post-glacial recolonization from multiple refugia (e.g. one or more northern refugium) is a plausible scenario for the demographic history of *P. trichocarpa* (Slavov et al., 2012; Zhou et al., 2014). In addition, a glacial refugium of *Populus* trees has been located to the panhandle of Alaska or Beringia, an unglaciated region encompassing the former Bering land bridge and the land between the Lena and Mackenzie rivers (Breen et al., 2012; Brubaker et al., 2005; Levsen et al., 2012). At present, *P. trichocarpa* inhabits riparian areas in western North America from Baja California to Alaska and grows preferably west of the Cascade Mountains in northern Oregon, Washington, and southern British Columbia (DeBell, 1990). This riparian species is highly sensitive to changes in water levels (Rood et al., 2003; Rood & Mahoney, 1990) which climate change can affect by altering snowpack melt and rainfall.

Previous studies suggest that effective population size (N_e) *P. trichocarpa* has ranged between 4000 and 8000 individuals (Levsen et al., 2012; Slavov et al., 2012; Zhou et al., 2014); however, this estimate is remarkably low with respect to the vast, current census population size of *P. trichocarpa* stands and three to four times lower than those that would be obtained based on putative neutral nucleotide diversity (Gilchrist et al., 2006; Tuskan et al., 2006). Moreover, previous work reported the highest N_e in the northern population (Zhou et al., 2014), which is counterintuitive given that the present-day core growth region is in southern British Columbia and northern Oregon.

To understand the way the *P. trichocarpa* populations respond to climate change depends on the extent to which the biogeography history has shaped their genetic divergence, I use multifaceted approaches to address the question of how past climate has affected the dynamics of niche evolution, demographic history and genetic diversity. To that end, I specifically:

- decipher population structure to identify genetic clusters, followed by characterizing distinctiveness via genomic metrics (e.g. π, F_{ST} and number of private alleles) between clusters ("populations" thereafter) and assume each population with its own unique evolutionary trajectory;
- scan for recent positive or long-term balancing selection in the genome of each population and assume that populations with more adaptive genetic loci indicate their great evolutionary potential because selective sweeps could lead to shifts in the selection landscape and thus local adaptation;
- estimate the direction and magnitude of gene flow between populations, whereby inferring the demographic history for single populations and the entire species;
- perform ecological niche modelling to forecast habitat suitability under climatic scenarios and test for niche evolution (divergence or conservatism); and
- predict changes in population genetic structure in response to climate change.

By assembling the aforementioned information (Figure 1), I provide the idea that we could facilitate prioritization of target and source populations for proactive management interventions through assisted gene flow to prevent populations from genetic loss.

2 | METHODS

2.1 | Plant material and georeferenced data

The *Populus trichocarpa* provenances used in this study were collected by British Columbia Ministry of Forests, Lands and Natural Resource Operations (Xie et al., 2009), and planted in a common garden at Totem Field of the University of British Columbia. A total of 409 individuals from 139 provenances were selected. These individuals were from 29 drainages (i.e. topographic units separated by watershed barriers) extending 14° in latitude (45.6°N–59.6°N) within its geographic distribution range (31°N–62.5°N) (Appendix S1: Table A1). Georeferenced occurrence records for *P. trichocarpa* were retrieved from the literature and herbaria accessed through the Global Biodiversity Information Facility data portal at https://www.gbif.org/. A total of 886 distinct locales were gleaned and used to represent the entire range of *P. trichocarpa* (Appendix S1: Table A2). Samples were grouped into three present-day populations including British Columbia (BC)-North, BC-South and Oregon.

2.2 | Genotype data

After the raw reads were filtered into clean biallelic SNPs using a set of filtering criteria (Method S1 and Figure S1), a total of 1.6 million SNPs (Dataset I) were extracted. Population parameters (Nei's nucleotide diversity (π), Tajima's *D*, linkage disequilibrium (LD) statistics, Wall's *B* and Z_{nS} and Hudson's F_{ST}) were calculated in a 500-bp sliding window with a step size of 500-bp on PopGenome v.2.2.4 (Pfeifer et al., 2014). The diversity values were normalized by the number of total nucleotides (both variable and invariable sites) in each window. Based on the SNPs previously identified through associations with phenotypes and/or habitat climates (Table S1 and Figure S2), I aggregated 681 SNPs, termed genome-wide associated (GWA) SNPs (Dataset II), for subsequent use (Appendix S1: Table A3). To decrease false positives, all retained SNPs passed the filtering of being identifiable in two versions of the *P. trichocarpa* genome (v.2.0 and 3.0; downloaded from Ensembl Plants). Likewise, the population parameters were calculated for each population in a 1-bp sliding window with a step size of 1-bp using PopGenome v.2.2.4 (Pfeifer et al., 2014). Hardy–Weinberg equilibrium (HWE) and LD were calculated using VCFtools v.4.2 (Danecek et al., 2011). The false discovery rate (Benjamini & Hochberg, 1995) was applied to correct *P*-values from HWE and LD multiple exact tests with $\alpha = 0.05$ using the *q* value package v.2.14.0 (Bass et al., 2015) in R.

2.3 | Climate data

Climate normals for each individual in the period, 1971–2000, were obtained using ClimateNA v.5.51 (ref. Wang et al., 2016; see Supporting Information for climatic variable notation). In addition to the climate normal period, climate for the future period between 2041 and 2070 (hereafter 2050s) was also projected. I used two Representative Concentration Pathways (RCP) 4.5 and 8.5 under the two downscaled Global Climate Models (GCMs) from the Phase 5 of the Coupled Model Intercomparison Project (CMIP5; http://cmip-pcmdi.llnl.gov/): the Community Climate System Model 4 (CCSM4) and the Model for Interdisciplinary Research on Climate (MIROC). As the combined effects of the Earth's tilt and rotation generate annual sine wave variation in day length with cascading effects on temperature, rainfall and resource availability, the average day length

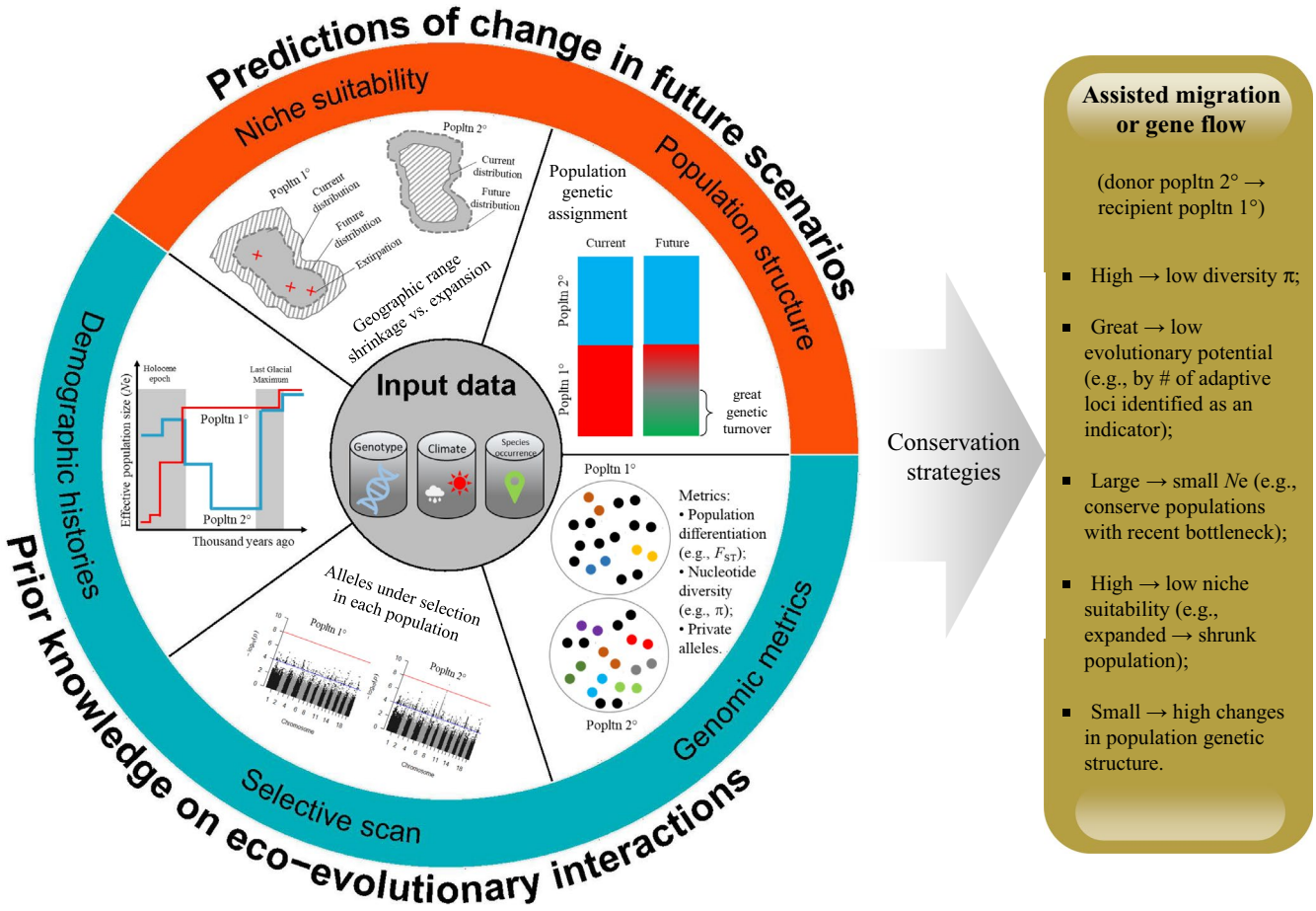

FIGURE 1 Conceptual representation of multiple genomic analyses to identify donor/recipient populations for developing conservation strategies. Illustrations are to elucidate different approaches employed to compare populations (e.g. Poplтn 1° and 2°) based on (1) prior knowledge on eco-evolutionary interactions including genomic metrics/diversity (coloured dots), signals of selection across the genome (Manhattan plot), and inferred demographic histories (changes in population sizes through time); and (2) predictions of change under future climatic scenarios using species niche suitability (shrinking or expanding geographic distribution ranges) and genetic makeup turnover (genetic structure differentiated by colour)

(lenDay; h) in June (summer solstice included) was calculated for each locale as a proxy for the photoperiodic regime using the geosphere package v.1.5.10 in R. I further extracted 19 bioclim variables from WorldClim v.2.0 (only available for the present-day conditions; last access on April 2020) and v.1.4 (past and future conditions) (ref. Hijmans et al., 2005; http://worldclim.org) at 2.5 arc-minutes resolution. These data allowed for modelling a species' full distribution range only via latitude and longitude extraction from a Spatial Polygons Data Frame or shapefile. A total of 15,602 spatial points within its natural range were extracted from a digitized map in ref. (Little, 1971), accessible from Data Basin at https://databasin.org/. As the existence of strongly correlated variables may overfit the model, I reduced collinearity ($|r| > 0.7$) among predictors by performing a pairwise Pearson's correlation analysis on the full set of bioclimatic variables and a Factor Analysis with orthogonal and oblique rotations (varimax and obvarimax solutions) for strict filtration and verification (Table S4 and Figure S9). I finally chose six representative bioclimatic variables, including Mean Annual Temperature (*MAT*; BIO1), Mean Diurnal Range (*DiurnalRange*; BIO2), Mean Temperature of Wettest Quarter (*TWettestQtr*; BIO8), Mean Annual Precipitation (*MAP*; BIO12), Precipitation Seasonality (*PSeasonality*; BIO15) and Precipitation of Warmest Quarter (*PWarmestQtr*; BIO18). In addition, I extracted 277,825 sets of scaled BioClim values (i.e. z-transformed to improve homoscedasticity) from a Raster object, which was used to define the study extent of western North America north of Mexico (150°W–115°W and 35°N–65°N).

2.4 | Population genetic assignment

To understand population genetic structure, I resorted to multiple analyses including SNP-, frequency-, distance- and likelihood (Bayesian)-based approaches. Multiple genetic differentiation metrics were calculated for populations using PEAS v1.0 (Xu et al., 2010) (Method S2). Compared to 34K genome-wide markers (Geraldes et al., 2014) previously used for genetic structure analysis, I used neutral GWA SNPs ($N = 577$ based on Dataset II; Method S3, Figures S4 and S5 for their identification details). Principle component

analysis (PCA; Method S4) based on GWA SNPs was carried out using the FactoMineR package v.2.3 (Lê et al., 2008) in R.

As the PCA does not entail the consideration of population structure, HWE and LD, I further implemented a Bayesian approach in STRUCTURE v.2.3 to infer population structure (Pritchard et al., 2000). This program iteratively ascribes individuals to *K* panmictic groups (populations) by minimizing deviations from HWE and LD within groups. I assumed that individuals followed the admixture model in structure, and allele frequencies were correlated among groups (Falush et al., 2003). For the correlated allele frequency model, the parameter λ was estimated at $K = 1$, as recommended by the software documentation for SNP data with rare alleles. According to a previous empirical estimate of the number of genetic groups using a Maximum Likelihood model (Geraldes et al., 2014), I varied *K* from 1 to 15, and for each run, I conducted five replicates with a burn-in period of 10,000 and MCMC iterations of 100,000. I applied posterior log probability (Pritchard et al., 2000), [Pr(*X*|*K*)], for each *K* and ad hoc statistic (Evanno et al., 2005), ΔK, to calculate the uppermost hierarchical level of structure (*K*). The results from five independent runs were summarized, and ln Pr(*X*|*K*) and ΔK were, respectively, plotted against *K* using the program STRUCTURE HARVESTER (Earl & vonHoldt, 2012). I used the computer program CLUMPP v1.1.2 to analyse the results from replicate analyses for optimal alignments of replicate clusters (Jakobsson & Rosenberg, 2007). The output from CLUMPP was graphically displayed by the cluster visualization program DISTRUCT v1.1 (Rosenberg, 2004). Due to significant patterns of IBD (Figure S7) and geographic barriers (Figure 2a) between populations, I also used conStruct v1.0.3 (Bradburd et al., 2018) to examine fine-scale population structure in a spatially aware context. Conceptually, conStruct is akin to STRUCTURE, except that geographic distances between sampling locales are included in the model and used for population structure inference. Due to a small number of SNPs used ($N = 577$), it was not possible to use a cross-validation to find the best-fit number of genetic clusters through conStruct. Instead, a set of spatial models was run by using different cluster numbers ($K = 2$ to 5 layers), each model with four independent MCMCs and 50,000 generations per MCMC. Convergence of independent analyses was assessed based on trace plots generated by conStruct.

Furthermore, I tested whether population structure is related to IBD or isolation-by-environment (IBE). Specifically, I tested for significant correlations between matrices of Euclidean genetic, geographic and climatic distances (Methods S5 and S6) using a Mantel test with 9999 permutations in the ade4 package v.1.7.15 (Bougeard & Dray, 2018) in R. Based on the pairwise F_{ST} and Euclidean geographic distance matrices implemented in GenAlEx 6.5 (Peakall & Smouse, 2012), I then calculated global spatial autocorrelations among three groups identified by conStruct analysis. The 95% CIs for the autocorrelation coefficients were determined using 10,000 permutations.

Finally, I confirmed the patterns identified in population structure by analysing the level and partitioning of genetic variation within and among populations using locus-by-locus analysis of molecular variance (AMOVA) (Excoffier et al., 1992) performed on Arlequin v.3.5.2 (Excoffier et al., 2005). Statistical significance was determined by comparison against 1,000 permutations. Three

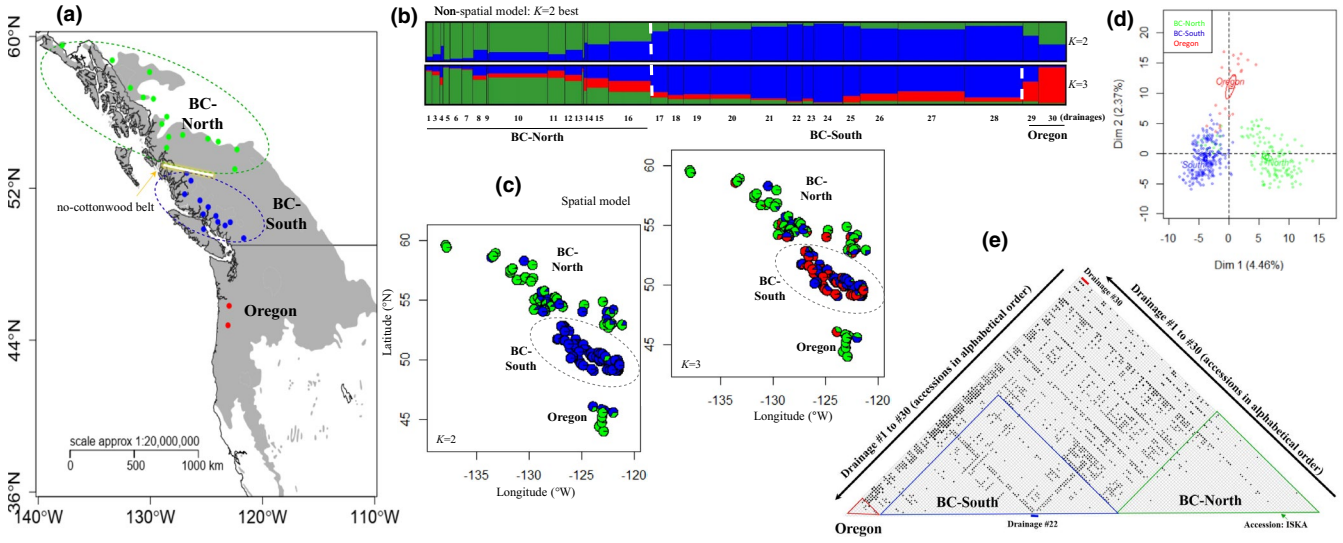

FIGURE 2 Population structure revealed by GWA SNPs. (a) Three present-day populations of *P. trichocarpa*. The grey-shaded background represents the distribution range of the species. The no-cottonwood belt divides the species distribution into the BC-North and -South populations and filled circles mark drainages used as a unit of habitat specificity. 409 individuals used in this study are from 29 drainages including 139 provenances. The same colour coding is used for the three population throughout this study. (b, c) Admixture proportions for drainages using Bayesian modelling and spatial conStruct model, respectively. (d) Individual factor map via PCA using GWA SNPs. Eigenvectors for the first two components are visualized on the graph and eigenvalues (inertia) are shown in the parenthesis. (e) Pairwise multi-locus F_{ST} matrix for provenances. I only display significant pairs ($\alpha = 0.05$) with F_{ST} value above 0.1 (N.B. negative F_{ST} is biologically uninterpretable and removed) and the range of F_{ST} values displayed in the matrix is bounded between 0.1 and 0.37. Three triangles frame between-provenance F_{ST} within populations

different hierarchical structures (i.e. groups classified by genetics, geography and climate) and differentiation indices (F_{ST}, F_{IS} and F_{IT}) were used to test differentiation strength between groups.

2.5 | Genome-wide scan for selective sweeps

Scans for selective sweeps were carried out first using integrated haplotype scores (*iHS*) implemented in the rehh package v.2.0.2 (Gautier et al., 2017) in R. The rationale for this method is that increase in the frequency of a beneficial allele occurs too quickly for recombination to homogenize its genetic background. Hence, a positively selected allele is embedded in a less common haplotype that extends further than what haplotypes at neutral loci do. This analysis was performed independently for each of the three populations. SNPs under significant strong selection were identified by selecting $|iHS| \geq 5.0$ (*c*. top 0.05%) for each population. The ancestral state of alleles at each focal SNP was extracted from raw VCF files (i.e. REF alleles). I used default settings for all analyses, and the threshold of missing data for haplotypes and SNPs was set to 90%. Moreover, I detected putative regions under long-term balancing selection, conducive to the maintenance of standing genetic variation, by estimating normalized β statistics implemented in BetaScan (Siewert & Voight, 2017). This measure identifies the clusters of variants with an excess number of intermediate frequency polymorphisms (Siewert & Voight, 2017). I used a window size of 1 kb to calculate the unfolded version of β (allelic state as described above) for each core SNPs in each of the three populations. SNPs with extreme β scores in the top 0.05% were deemed significant.

2.6 | Historical and contemporary gene flow

To investigate gene flow scenarios between population groups, I employed MIGRATE-N v3.6.11 (Beerli, 2006; Beerli & Felsenstein, 2001; Beerli & Palczewski, 2010) and BayesAss v3.0.4 (Beerli & Felsenstein, 2001), which estimate historical and contemporary migration rates, respectively. I used a Bayesian method to estimate historical gene flow parameters: mutation-scaled effective population size, θ ($\theta = 4N_e\mu$, where N_e is effective population size and μ is the mutation rate per generation) and mutation-scaled migration rates, *M* ($M = m/\mu$, where *m* is the migration rate per generation) for multiple populations in a coalescent framework (Beerli, 2009). Relative mutation rates were estimated from the data on the basis of the number of alleles for each locus as implemented on MIGRATE-N (mutation = DATA). Before choosing priors for θ and *M*, I experimented with a full model with starting values estimated from F_{ST} (Beerli & Felsenstein, 1999) and refined in subsequent trial runs using θ and *M* obtained in the first run. Search parameters were shown as follows: (1) uniform priors for both θ and *M* ranging from 0 to 1000 with an interval of 100; (2) three long chains with an increment of 100 and sampling of 100,000; and (3) a burn-in of 10,000, heating = ADAPTIVE: 1 {1.0, 1.5, 3.0, 1.0×10^6}. The results of the final long chain search were averaged over five independent runs. I observed Markov chain Monte Carlo (MCMC) chain convergence for the final output using Tracer v. 1.7.1 (http://beast.community/tracer). Recent migration rate was used as a measure for an indirect estimate of gene flow among population groups over the last several generations. The program calculates unidirectional estimates of *m* for each group pair (Fraser et al., 2007), and this method is based on the model in which individuals are exchanged between groups over generations (Goossens et al., 2005). I run the program with the common-line flag "BA3 –v –i10000000 –b1000000 –n1000" on BayesAss (Beerli & Felsenstein, 2001). As the data exceeded the program's single load size, I divided the data into two equal sets (i.e. 341 or 340 loci per load), and the averaged output was used for the final results.

To test scenarios of gene flow among provenances of the three populations, I carried out a multivariate provenance (subpopulation) graphing approach via PopGraph v.1.4 (Dyer & Nason, 2004) using gstudio (Dyer, 2009). Based on genetic covariance structures among provenances (i.e. MATRIXFILE in Bayenv2.0; Methods S3), PopGraph creates a framework of connections among provenances, and then nodes and edges are mapped onto a real space. Social network node-specific parameters including closeness, degree, betweenness and eigenvector centrality were estimated from PopGraph. I interpreted lower values of closeness and higher values of degree, betweeness and eigenvector centrality as indicative of higher gene flow (Method S7).

2.7 | Demographic history inference

The single population demographic histories were estimated using 200 bootstrap iterations with stairway plot package v.2.0 beta2 (Liu & Fu, 2015) based on observed unfolded SFS from each population. Considering that the exclusion of singletons would skew the entire frequency spectrum potentially impacting population genetic processes inference, I included singletons when estimating the SFS. I used default 2/3 of the data for training, and four numbers of sequences (nseq), including (nseq)-2)/4, (nseq-2)/2, 3*(nseq-2)/4 and nseq-2 as the number of random breakpoints. According to the stairway plot pipeline, the best number of random breakpoints was chosen based on the training data. To calibrate the results, I used a generation time of 15 years (Ingvarsson, 2008) and two different mutation rates, $\mu = 1.75 \times 10^{-8}$ and 2.5×10^{-9} (Method S8).

To elucidate the population split history, I used the joint 2D-SFS (Method S9) based on the composite likelihood (CL) demographic modelling approach described by Excoffier et al. (2013) and implemented in fastsimcoal v.2.6. I assumed two models with alternative scenarios of lineage divergence or formation: Model one assumes migrants from BC-South (source) to BC-North (sink) due to admixture or introgression events, whereas Model two assumes the same events with a reverse direction of migration. In both models, population sizes were allowed to vary in historical events, and thus the bottleneck may occur given new population size inferior to the previous one in simulations. The historical gene flow was set according

to the result shown in Figure S13. To enable the estimation of other parameters (Excoffier et al., 2013), the parameter for ancestral population size (N_e) was anchored to be 3048, estimated by the formula, $N_e = N_{ANC} = \theta_\pi \times (4 \times \mu \times L)^{-1}$, where the genetic diversity (θ_π) was set as 0.0032 (Levsen et al., 2012), the neutral mutation rate (μ) was set as 1.75×10^{-8} (Method S8), and the generation time (L) was assumed to be 15 years (Ingvarsson, 2008). The recombination rate per generation (c) was set as 9.37×10^{-8} given $c = \rho \times (4N_e)^{-1}$, where ρ is the scaled recombination rate (Zhou et al., 2014). Monomorphic sites and singletons were ignored in the likelihood computation (N.B. only SNPs that are singletons across all samples are ignored), and SFS entries with support from less than six SNPs were collapsed into a single entry, as suggested by Excoffier et al. (2013). To maximize the likelihood of the model, I considered 200,000 coalescent simulations for the CL calculation, 10–40 expectation-conditional maximization (ECM) cycles, and a convergence criterion of 0.001. I performed 50 independent runs with randomized starting points to retain the global maximum likelihood model. Two alternative models were compared, using Akaike information criterion (AIC) and Akaike's weight of evidence following Excoffier et al. (2013). The 95% CIs for each estimate of the best model were computed from 100 parametric bootstrapping runs from simulated SFS of the point estimates. The simulated SFS was compared with the observed SFS to evaluate the fit of the best demographic model.

In contrast with N_e inferred from best-supported demographic models, N_e was also calculated for each population on NeEstimator v2.1 (Do et al., 2014) using GWA SNPs. This program employs an LD method to calculate N_e estimates. I assumed a random mating model and calculated separate estimates after screening out alleles with frequencies below P_{Crit} values of 0.05, 0.02 and 0.01, respectively. Note that these two N_e estimation approaches may not be comparable, as SFS data and short-range LD are very sensitive to long-term N_e, whereas long-range LD tends to be influenced by very recent changes in population size (e.g. bottleneck and recent expansion) (Rogers, 2014; Tenesa et al., 2007).

2.8 | Ecological niche modelling and niche divergence

To test how climate has driven species adaptation, I performed climatic suitability prediction using ecological niche modelling (ENM) for past, current and future time periods with MaxEnt v.3.4.1 (Phillips et al., 2006). I used 524 georeferenced points as species presence locales and six representative climatic variables (see Section 2.3) in the construction of environmental GIS data layers. The program DIVA-GIS v.7.5 (Hijmans et al., 2005) was used to create mask files of the population's background ranges from which to generate the random sample points. Of the 886 locales compiled, there were 362 pruned due to being outside of the boundary of the focus region I set or having no corresponding climate data point extractable from the environmental layers. Model performance was assessed using the area under the receiving operator characteristics curve (AUC). Predicted distribution probability (on a logistic scale) in each grid cell is an index of suitability or suitability scores, ranging from 0 to 1 (unsuitable to highly suitable [blue to red portrayed on the graph]). I produced a current distribution using default settings except only allowing hinge features, a 20-fold cross-validation, and a regularization multiplier of 2.5 to fit more general models (Elith et al., 2010). To obtain a simulated current distribution, interpolations of observed data during the 1971–2000 period were used. To obtain the potential distribution at the Last Glacial Maximum (LGM; ~21K years ago), I used paleoclimatic layers simulated under CCSM4 and MIROC. Likewise, I used the same GCMs under both low and high greenhouse gas (CO_2 accumulation) scenarios, RCP 4.5 and 8.5, to obtain potential distributions in the 2050s, respectively.

Further, I used the locales with suitability scores above 0.1 in the present-day scenario to investigate the patterns of ecological niche evolution across this species' full distribution range. Based on empirical data, I defined the three populations by individual's latitude, that is, ≥53°N as BC-North, ≤45°N as Oregon, and in-between these two transects as BC-South. To assess levels of climate niche divergence between populations, I first calculated the niche overlap statistic Schoener's D (Schoener, 1968) between populations using the ENMTools package v.1.0 (Warren et al., 2008) in R. I then performed a Bioclim background randomization test in ENMTools to account for potential biases towards niche divergence introduced through latitude-associated environmental variation in the present-day range of each population pair. The two resulting null distributions obtained through this test correspond to the background level of niche divergence for each population pair. They were generated by comparing the niche model of a focal population to 1,000 models generated from randomly drawn points from a background range of the other population. Such pairwise comparisons for the three populations were conducted with 200 replicates. An observed value of Schoener's D much smaller than expected after accounting for background differences indicates a greater level of similarity than expected by chance (i.e. niche divergence), whereas a value much larger than expected reflects niche conservatism (Warren et al., 2008).

2.9 | Population genetic structure in future climatic scenarios

To predict the dynamics of genetic clusters for GWA SNPs under selection by nine climate drivers in the context of climate change, I conducted spatial simulations using POPS v.1.2 (Jay et al., 2015), which implements Bayesian clustering based on genetic, geographic and climatic variables. By using 104 adaptive GWA SNPs (Method S3) and nine climate variables, I first simulated genetic clusters under the current climate scenario (1971–2000) using an MCMC estimation algorithm implemented on POPS v.1.2 (Jay et al., 2015). The MCMC runtime was set to 100,000 sweeps and the burn-in period to 10,000 sweeps, and the degree of the spatial trend surface was set to 1. These settings were launched using models with admixture. To define the number of clusters, K, I run the simulations five times

for each *K* ranging between 2 and 15. A subset of runs that minimize the deviance information criterion (DIC) scores (Spiegelhalter et al., 2002), namely, the lowest DIC values, were selected. Finally, I forecasted the genetic clusters for two future scenarios (RCP 4.5 and 8.5) under two GCMs: CCSM4 and MIROC. To average multiple runs with the same number of clusters, I used the computer program CLUMPP to analyse the results from replicate analyses for optimal alignments of replicate clusters (Jakobsson & Rosenberg, 2007). The output from CLUMPP was graphically portrayed using R. The grid from an ascii raster file created in *P. trichocarpa* distribution modelling under the current scenario was used as background to display estimated and predicted coefficients of this analysis on the map.

3 | RESULTS

3.1 | Population structure revealed by associative genomic loci in a spatially aware context

According to a previous genetic clustering based on genome-wide SNP variants of *Populus trichocarpa* (Geraldes et al., 2014), the species mainly consists of three present-day populations: populations in the geographic north and south of British Columbia (BC-North and -South, hereafter) and the Oregon population (Figure 2a), in which BC-North and -South are geographically adjacent but isolated by a no-cottonwood belt (Figure 2a). For comparison, this study used GWA SNPs to decipher population genetic structure. Given a total of 681 GWA SNPs compiled from the literature (Table S1), there were 194 SNPs that showed significant deviations from the expectations of HWE and LD (Appendix S1: Tables A4 and A5). The best model for hierarchical population structure supported a significant disparity between BC-North and -South and a partial genetic similarity between BC-South and Oregon (Figure 2b), as evaluated by the posterior log probability and *ad hoc* statistic ΔK (Figure S6a). While STRUCTURE analysis indicates a most likely *K* of 2 (Figure 2b), the spatial analysis by conStruct with a geographic layer incorporated showed a similar result for $K = 2$ (Figure 2c). However, given $K = 3$ or 4, the spatial model deciphered subpopulation structures of BC-South (Figure 2c and Figure S6b). The Oregon as an independent cluster became evident when $K = 5$ in the spatial model (Figure S6b), in contrast to $K = 3$ in the non-spatial model (Figure 2b). The PCA result showed a clear separation of BC-North and -South along the first axis, whereas axis 2 separated BC-North and -South from part of Oregon (i.e. drainage #30; individuals of drainage #29 overlapped with the BC-South) (Figure 2d). Pairwise F_{ST} values for provenances were small within populations except for drainage #22 of BC-South (Figure 2e).

3.2 | Population genetic diversity, differentiation, LD and signal of selection

Patterns of Nei's nucleotide diversity (π) were different between populations (Figure 3a), and BC-North harboured significantly greater nucleotide diversity (average $\pi = 2.71e^{-3}$), compared with BC-South and Oregon ($2.66e^{-3}$ and $2.59e^{-3}$, respectively; all adjusted *p*-values <.0001 based on Wilcoxon test) (Figure 3a). Two LD statistics, Wall's *B* and Z_{nS}, were consistently highest in Oregon (average $B = 0.188$; $Z_{nS} = 0.057$), followed by BC-South (0.170; 0.054) and BC-North (0.145; 0.056) (all $p < .0001$; Figure 3a). Pairwise genetic differentiation showed that Oregon (Hudson's $F_{ST} = 0.033$) differed from BC-North and -South (0.025 for both; $p < .0001$) (Figure 3a). Moreover, BC-North and -South had more private SNPs ($n = 1{,}609{,}193$ and 1,607,971, respectively) than Oregon (1,457,717). The entire genome of BC-North and -South had significantly higher Tajima's *D* values (1.68 and 1.96, respectively), compared to Oregon (0.90; *p*-values <.0001) (Figure 3a). After a stringent filtering for missing data (≤10%), 1.43, 1.45 and 1.33 million callable SNPs were used in the BC-North, -South and Oregon populations, respectively, in EHH-based statistics for selection scans. Based on *iHS* above 2 and 5, similar fractions of loci (27.03% and 0.055%) were significant and under strong selection across the populations, respectively (Figure 3b). More loci were under strong selection in BC-South ($n = 1072$) than BC-North and Oregon (664 and 602, respectively), with 73 loci identified across the three populations. Given the same proportion of loci (top 0.05%) deemed under balancing selection for each population using BetaScan, I found that 35.86% of these loci were under balancing selection across populations and 25%–35% loci under balancing selection exclusively in single populations (Figure 3c).

3.3 | Contribution of IBD and IBE to population genetic divergence

Genetic differentiation between BC-South and Oregon was highest (Wright's $F_{ST} = 0.16$) (Table S3). Molecular variance by AMOVA (Table S4) revealed that most of genetic variation was distributed within individuals (86.04%), and individuals were genetically more divergent among populations ($F_{IT} = 0.14$) than within population ($F_{IS} = 0.06$). The spatial autocorrelation analysis showed significantly positive genetic correlations for populations within 200 km (Figure S7). There was a moderately high correlation between climate and geography ($r = .627$, $p < .0001$) (Figure S8a). The correlation between genetic and climatic distances increased ($r = .49$, $p < .0001$) when the three populations were pooled for analysis (Figure S8b). Note that drainages #29 and #30 of Oregon formed two distinct clouds of points, and the correlation was weak based on either drainage (average $r = .23$, $p < .0001$) (Figure S8b). Nonetheless, some climatic variables (e.g. mean annual precipitation, *MAP*) were weakly-to-moderately correlated with niche climate and geography ($r = .13$–$.49$, $p < .0001$) (Figure S9). In addition, average day length (photoperiod) was highly correlated with geography as expected, showing a latitudinal pattern (spearman's $\rho = 0.955$, $p < .0001$; Figure S2), but the correlation between day length and niche climate was much weaker ($\rho = 0.629$, $p < .0001$; Figure S2). There were intermediate-to-weak correlations of *MAP* and mean annual temperature (*MAT*) with geography ($\rho = 0.594$–0.132, all $p < .0001$; Figure S9).

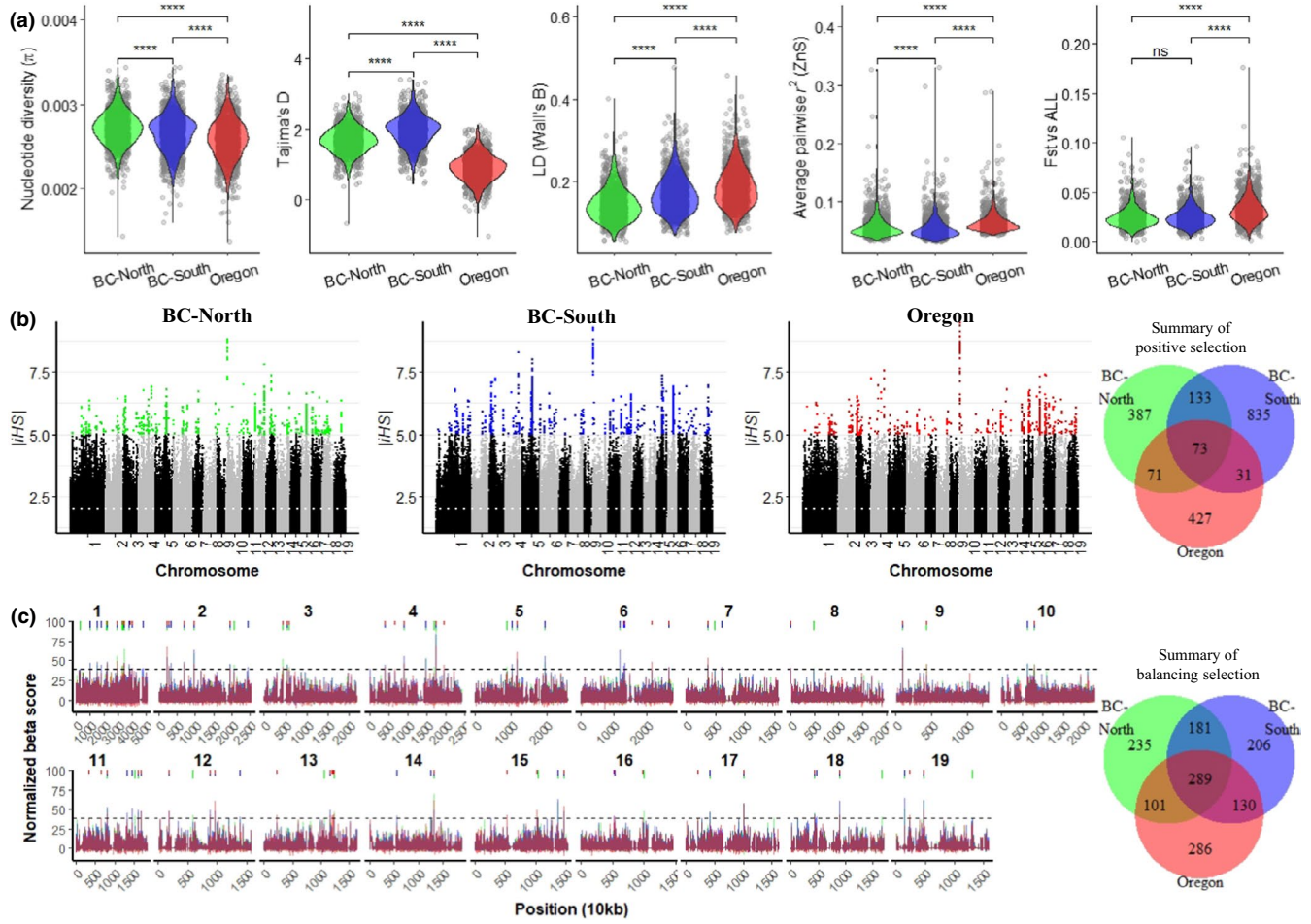

FIGURE 3 Genome-wide scan for genetic diversity, LD, positive and balancing selection in each population. (a) A 500-bp window size with a 500-bp step size was used to define the slide window for scanning across the 19 chromosomes using the filtered sequentially concatenated 1.6 M SNPs. The π values were weighted by using the total genome length (i.e. both variable and invariable sites). Wilcoxon signed-rank tests were performed between each population pairs. All *p*-values were adjusted using the sequential Bonferroni method (****: $p < .0001$; n.s.: $p > .05$). F_{ST} vs. ALL denotes fixation index (Hudson's F_{ST}), contrasting one population against all the other individuals. (b) Genome-wide scans for recent positive selection using the standardized integrated haplotype score |*iHS*| in each of the three populations. Dots show |*iHS*| values for each SNP within the populations BC-North, -South and Oregon, respectively. The most significant SNPs (|*iHS*| > 5) are highlighted in green, blue and red for the three populations, respectively, and summarized by a Venn diagram. (c) Genome-wide scans for long-term balancing selection using BetaScan. The horizontal dashed line marks the average cut-off of the β statistics across the three populations ($\beta = 38.66$). Signals of balancing selection across the chromosomes are delineated at the top of the panels with colour bars corresponding to the three populations. The Venn diagrams summarize shared and unique selected SNPs in each population

3.4 | Ecological niche model and niche evolution reveal niche suitability and divergence over time

Six of the original 19 bioclimatic variables were selected for ecological niche modelling (ENM) based on collinearity among the variables (Table S2 and Figure S3). The ENM accuracy was relatively high (area under the curve [AUC] = 0.893; SD = 0.026). The present-day distribution model was mostly concordant with its extant distribution (Figure S10), but under-predicted areas of occurrence in British Columbia, Canada (Figure S10). The climatic variables that contributed most to the model predictions, measured as the percentage of total model gain contributed by a variable, were *MAP* (40.2%) followed by *MAT* (38.7%) (Figure S11), with higher spatial heterogeneity in *MAP* than in *MAT* over the distribution range (Figure S9). Paleodistribution models indicated a history of relatively stable range in *P. trichocarpa* (Figure S10). By illustrating changes in niche suitability based on the predictions of *P. trichocarpa* past (Last Glacial Maximum, LGM) and present distributions, I found no significantly different changes in niche suitability between BC-North and Oregon based on the two models (Figure 4a). The inferred LGM and present distributions indicate that most suitable habitats had been in northern Oregon, Washington and southern British Columbia, but also a much larger northern extent of the distribution during the LGM (e.g. southern Alaska), relative to the present (Figure S10). I additionally

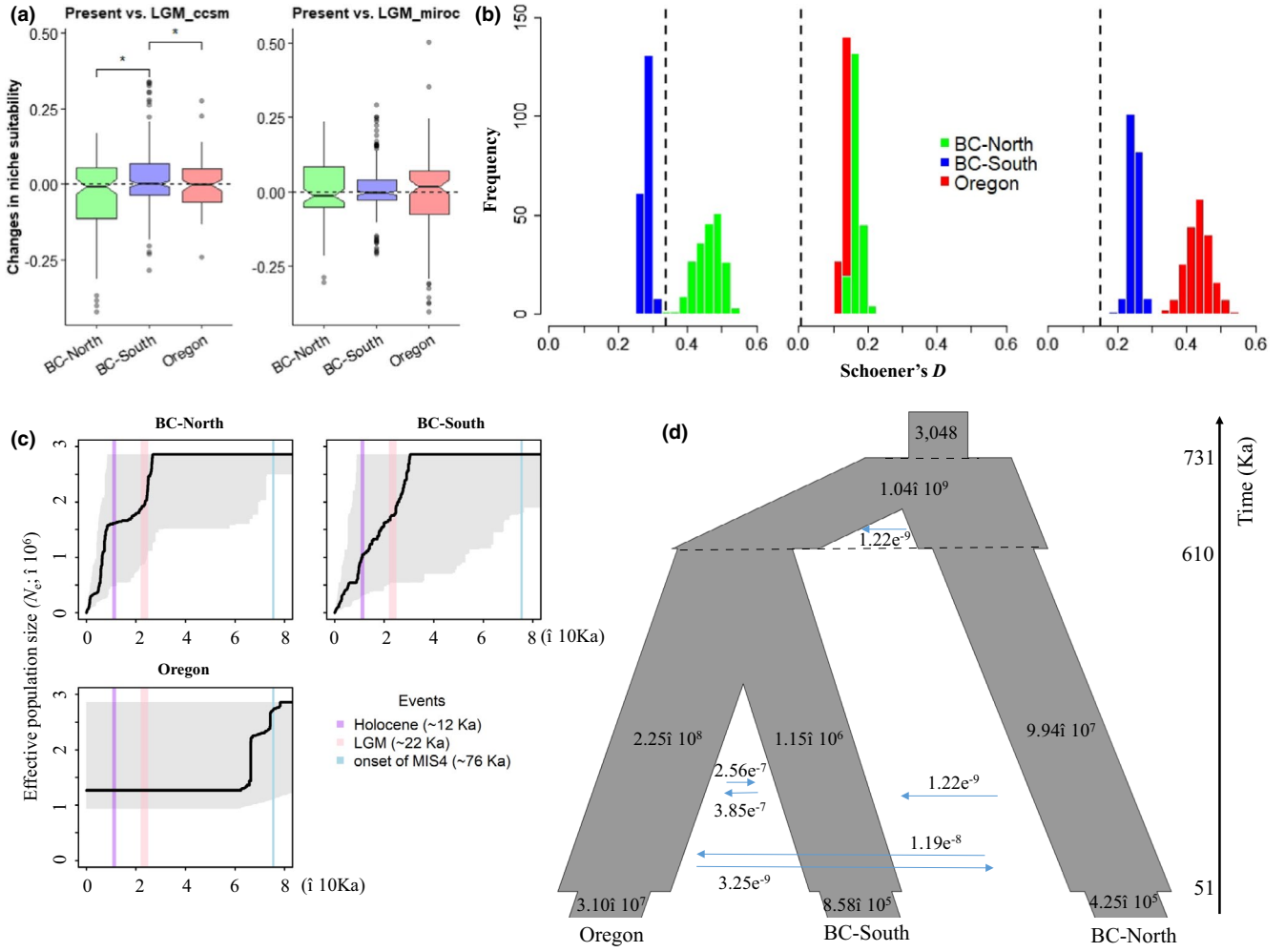

FIGURE 4 Ecological niche model, niche evolution and demographic inference. (a) Changes in niche suitability for 524 *P. trichocarpa* locales between present-day (1971–2000) and Last Glacial Maximum (LGM) based on two models, CCSM4 and MIROC. Hindcasted suitability maps can be viewed in Figure S10. An asterisk * for significance at $p < .05$ based on Wilcoxon test. Non-significant pairs are not displayed. (b) Assessment of niche divergence and conservatism based on pairwise niche divergence tests (Schoener's *D*). Histograms for null distributions indicate the background levels of niche divergence, and dashed lines indicate the observed value of Schoener's *D* (niche overlap) for each pair compared. An overlap value larger than the null distributions indicates niche conservatism, whereas a value significantly smaller than the distributions indicates niche divergence. All *p*-values from background tests are smaller than 0.01. (c) Estimated single population demographic histories. The bold line represents the median estimate from 200 bootstrap replicates with 95% CIs shown via light grey shading. MIS4 denotes Marine Isotope Stage 4. Mutation rate was set as 1.75×10^{-8} per site per generation, and generation time was assumed to be 15 years in the model. Other demographic trajectories modelled with 7× lower mutation rate can be viewed in Figure S15. (d) Schematic illustration of the best-fitting demographic model for the *P. trichocarpa* colonization history. The width of the boxes represents the relative effective population size and arrows represent directional gene flow between populations. Maximum-likelihood parameter estimates for the model are given in the graph, including split times (Ka), population sizes and migration rates (95% CIs given in Table S6). The model assumes that BC-North and -South split from a common ancestral and Oregon then separated from BC-South. The three populations underwent recent genetic bottlenecks. Fit of models was assessed and given in Figure S16 and Table S6

note that British Columbia was almost entirely covered by continental ice sheets at the LGM (Figure S12), and Alaska, therefore, provided a possible refugium for the BC-North population during the LGM. With respect to testing for niche divergence, background randomization tests using all the six bioclimatic variables revealed statistically significant niche divergence for two of the three comparisons (Figure 4b). Niche divergence was asymmetrical between BC-North and -South, with the BC-North being conserved relative to the background of BC-South (Figure 4b).

3.5 | Direction of historical and contemporary gene flow between populations

Historical gene flow was low and unidirectional from BC-North to –South, but there was a strong and bidirectional gene flow between BC-South and Oregon (Figure S13). Contemporary gene flow within populations was higher than between populations (Figure S13), and this was also supported by high genetic connectivity among provenances within populations (Figure S14a). However, high

between-provenance connectivity was not homogenous, and some provenances were dominant gene flow donors or recipients based on low "closeness" and high "degree," "betweenness" and "eigenvector centrality" in PopGraph (Table S5). Moreover, the between-provenance correlations were overall moderate within BC-North, -South and Oregon (average Spearman's r = .63, .56, and .35 with average p = .002, .001, and .456, respectively) (Figure S14b).

3.6 | Demographic history shows more recent bottlenecks in BC-North and -South than in Oregon

The stairway plot revealed that bottlenecks in the three populations corresponded roughly with known events of climate upheaval (Figure 4c). For Oregon, an early N_e drop was coincident with the onset of Marine Isotope Stage 4 (~76 Ka) (Figure 4c). Following this climatic event, Oregon has maintained a stable N_e (Figure 4c). A later bottleneck occurred in BC-North and -South after the LGM (Figure 4c). That the BC-North and -South were greatly impacted during the LGM agrees with the fact that the southern margin of the LGM ice sheet roughly coincided with a transect line separating BC-South and Oregon (Figure S12). A recent bottleneck occurred in BC-North during the Holocene (~12 Ka) that corresponded with the onset of anthropogenic environmental changes (Figure 4c). During this time period, N_e in BC-South continued declining (Figure 4c). The same analysis based on a sevenfold lower mutation rate also revealed bottleneck and gradual N_e decline over time in BC-North and -South, respectively, and stable N_e in Oregon over the past 60 Ka (comparison with a shorter time period due to a smaller mutation rate used; Figure S15).

Further, I studied the population split times using a simulation-based approach implemented in fastsimcoal2. The model and parameter search ranges were selected based on the results in previous sections (see Methods for setting details; full estimates in Table S6). Two alternative models with different unidirectional gene flow between BC-North and -South were tested. The best-fitting model favoured unidirectional gene flow from BC-North to -South (Table S6 and Figure S16). The model also supported the proposed genetic bottlenecks that the species might have undergone (Figure 4d), highlighting more recent bottlenecks in BC-North and -South than in Oregon (Figure 4c). The model also indicates that the north-to-south gene flow into the Oregon was higher than that in the opposite direction (Figure 4d), as shown by one-magnitude higher migration rate for BC-South to Oregon than the reverse direction and 1.4 times higher for BC-North to Oregon than the opposite (Table S6). Under the best-supported model, the current N_e of Oregon was largest, in line with the results from the single population demographic models (Figure 4c; but see Table S7 based on an LD approach using GWA SNPs with a general interpretation in Section 2.7): ~31 million with a 95% CI of 29–52 million from a large ancestral population (~225 million, 95% CI = 98–217 million) due to a bottleneck before separating with BC-South ~51 Ka (95% CI =67–206 Ka) (Figure 4d and Table S6). BC-North and -South diverged ~610 Ka (95% CI =606–617 Ka) from a large ancestral population (~1.04 million, 95% CI = 0.9–1.05 million) (Figure 4d and Table S6). The simulated SFS from the model showed that both BC-North and -South have an excess of low frequency alleles and a deficit of mid-to-high frequency alleles, consistent with the decline of range-wide empirical SFS (Figure S17). Such an SFS distribution pattern in the two populations fits with the inference of bottlenecks that they have likely undergone. The Oregon exhibited an exponential decay (Figure S17a–c) indicating population stationarity, concurring with the outcome of its single population demographic model (Figure 4c).

3.7 | Projections of change in niche suitability and genetic structure in response to global warming

Compared to the present, four future distribution models consistently showed no dramatic range shifts over the next 30 years, that is, two generations given 15 years to reproductive maturity in *Populus* ($\Delta \approx \pm 0.05$ or $\pm 5\%$ averaged across the four models; Figure 5a and Figure S11). However, due to the high spatial landscape heterogeneity, findings showed that across populations, niches along the coastal area will likely be less suitable, while the interior has a propensity for increased suitability (Figure 5b). In predicted changes in genetic makeup due to climate change, eight genetic clusters were used based on the deviance information criterion (DIC) scores (Table S8). Compared with the present-day (1971–2000), the CCSM4 model under both low and high greenhouse gas emission scenarios (RCP 4.5 and 8.5) showed that the main genetic cluster in the BC-South population (cluster 1) substantially shrank and would be superseded by the Oregon main genetic cluster in the 2050s (cluster 4) (Figure 5c). By contrast, the MIROC_ESM model predicts that the current major population-level genetic clusters (clusters 1, 2 and 4) would be replaced by current rare clusters (e.g. clusters 6 and 8) particularly for Oregon and BC-South under RCP 8.5 in the 2050s (Figure 5c and Figure S18).

4 | DISCUSSION

The three *P. trichocarpa* populations have disparate nucleotide diversity π (Figure 3a) and experienced different positive and balancing selection at loci across the genome (Figure 3b,c), indicating the discrepancy in adaptive forces that maintains genetic variation in these populations. In quest of exploring conservation strategies for populations, I used multiple genomic approaches to investigate how demography and selective processes interact to constrain genetic diversity and how past and future climate scenarios have shaped or would alter species niche suitability and population structure (Figure 1). Significance of this study attaches importance to the use of comprehensive genomic analyses to identify the populations greatly impacted by climate change at the genetic level and then to advocate for implementing practices to prevent population's genetic risks (e.g. loss of diversity) based on that information.

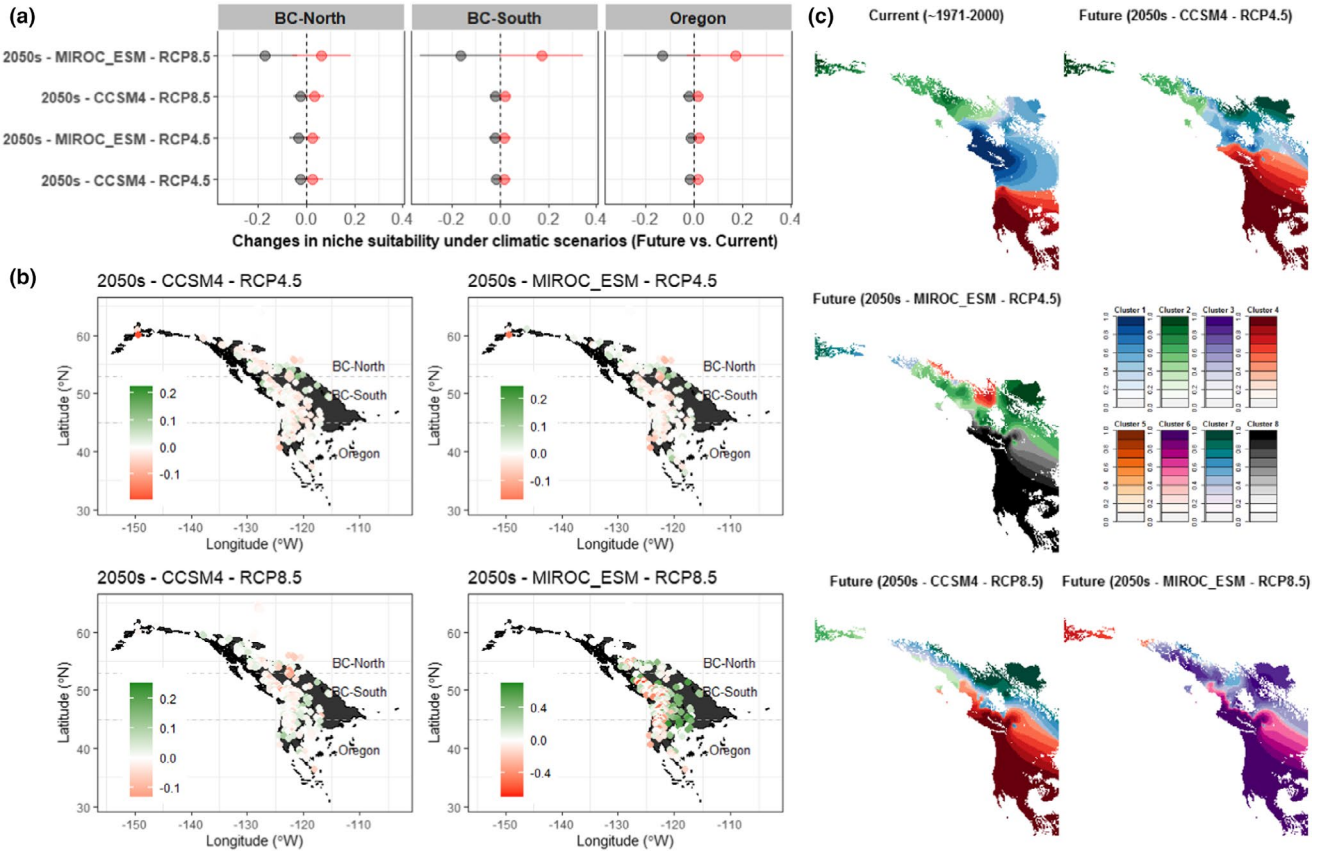

FIGURE 5 Forecasted changes in niche suitability and genetic clusters due to global warming. (a, b) Changes in niche suitability for 524 *P. trichocarpa* locales between present-day (1971–2000) and a future period (2050s) based on two models, CCSM4 and MIROC. The future climate predictions considered low and high greenhouse gas scenarios, RCP4.5 and 8.5. Increase (red) and decrease (black) in suitability are respectively graphed in (a) with error bars. The grey-shaded background in (b) is the distribution range of the species. The dashed horizontal lines in (b) mark the boundary for BC-North, Oregon and BC-South, defined as north of 53°N, south of 45°N and between the two transects, respectively. Future suitability maps can be viewed in Figure S10. (c) Spatial distribution of genetic clusters conditioned to climatic variables under low and high greenhouse gas scenarios. Each colour set in the legend represents a genetic cluster. Only grids with suitability above 0.1 projected for the present-day scenario (1971–2000) in the ENM were used for clustering and at each point, the cluster for which the coefficient is maximal was plotted

4.1 | Gene flow and environmental drivers of population divergence

Environmental heterogeneity influences gene flow through geographic barriers that impede dispersal or by creating contrasting selective pressures which elicit local adaptation and select against gene flow between locally adapted forms (Rundle & Nosil, 2005; Sobel et al., 2010). Such influence varies from one system to another, depending on the magnitude and spatial configuration of landscape heterogeneity. In this study, both the best-fitting demographic models (Table S6 and Figure S16) and historical gene flow estimation (Figure S13) favoured unidirectional gene flow from BC-North to -South, in accordance with the species southward migration after deglaciation from a possible glacial refugium in Alaska, a far northern territory of the BC-North population. The gene flow direction is attuned to the post-glacial recolonization route from an Alaska glacial refugium followed by secondary contact in the species. As a result of repeated founder events and allele surfing along the colonization front (de Lafontaine et al., 2013), this process creates genetic barriers, that is, isolation by colonization (IBC). While there is a natural barrier between BC-North and -South (Figure 2a), evidence of niche divergence (Figure 4b) and climate (Figure S4a) indicates a starkly heterogeneous landscape between BC-North vs. BC-South. This indicates that gradients of IBD, IBE and IBC covary and jointly affect gene flow between BC-North and -South. Findings also indicate a pronounced IBD within populations (Figure S7). However, there is a climatic similarity between BC-South and Oregon (Figure S4a), indicating that IBE is not the main cause of their genetic separation. Moreover, IBE was weak within population, as evidenced by significant divergence between adjacent drainages #29 and #30 (Figure 2d and Figure S8b). This indicates a limited contribution of niche climate adaptation or geographic separation to population structure within *P. trichocarpa* populations. Furthermore, 11 frost and drought-related climate variables were identified as potential selective drivers (Figure S5a), in which the primary three included

mean annual temperature (*MAT*), mean warmest month temperature (*MWMT*) and frost-free period (*FFP*) (Figure S5b). However, these climatic stressors were not from single climate clusters (Figure S4b,c), suggesting that specific climatic variables rather than niche climate similarity (or an aspect of climate) are the true drivers of population genetic divergence.

4.2 | Relationship among genetic diversity, selection and demography

Genetic diversity is a key basis for species survival and ecosystem functioning (Toro & Caballero, 2005). Greater genetic diversity prompts higher evolutionary potential and a greater capability of responding to climate change (Frankham et al., 2002). Conserving genetic diversity therefore becomes crucial and high-priority strategies for biodiversity conservation (Frankham et al., 2002). This study provides evidence that eco-evolutionary processes interact to affect genetic diversity. First, the Tajima's *D* pattern (Figure 3a) indicates that BC-North and -South have deviations from mutation-drift equilibrium and both populations might have undergone a demographic contraction, resulting in the loss of low frequency alleles and thus yielding positive Tajima's *D* values (Figure 3a). However, based on this alone, we cannot exclude the possibility of within-population genetic structure making the frequency spectrum shift towards alleles of intermediate frequencies by pooling locally common alleles in subpopulations. Further, the SFS distribution was graphically used to reveal whether a population is non-bottlenecked at mutation-drift equilibrium or not (Nei et al., 1976). Deviations from an L-shape would indicate a population bottleneck (Luikart et al., 1998). BC-North and -South did not exhibit L-shaped SFS with slight distortions (Figure S17a,b), indicating that both populations have experienced bottlenecks. In addition, drainage #22 of BC-South showing high F_{ST} even within populations (Figure 2e) indicates its bottleneck. Analyses of the demographic history of the species supported that two ancestral populations, BC-North and -South, have been subjected to more recent bottlenecks than the Oregon population (Figure 4c,d).

Neutral theory predicts that genetic diversity and population size are positively correlated, according to $\pi = 4N_e\mu$ where μ is the per site per generation mutation rate. Nonetheless, selection could suppress diversity due to greater impact of linked selection in large populations (Corbett-Detig et al., 2015) possibly driven by higher rates of adaptive substitutions in large populations (Ohta, 1992). A confluence of these viewpoints suggests that selection and demography jointly shape variation in N_e and comprise two main factors affecting genetic diversity (Ellegren & Galtier, 2016). A reduction in N_e, *a priori*, reduces the efficiency of selection, although increased homozygosity could make selection more efficient by exposing recessive mutations. As opposed to the neutral theory, the impact of selection may lead to a negative relationship between π and N_e or between estimated heterozygosity (H_e) and N_e (Table S7). Indeed, based on demographic models, the Oregon population had the highest N_e (Figures 3d and 4c) but the lowest π across the genome (Figure 3a). Moreover, an early N_e drop in the Oregon coincided with the onset of Marine Isotope Stage 4 (~76 Ka; Figure 4c), a period with a rapid increase in global ice volume (Landais et al., 2004). This climatic event occurred 55–64 Ka earlier than the LGM (~21 Ka) or Holocene (~12 Ka), which coincided with a dramatic N_e drop in BC-South and -North, respectively (Figure 4c). Furthermore, demographic history determines the strength of genetic drift, that is, the inverse of N_e represents the magnitude of genetic drift and elevated genetic drift results in increased fixation or loss of alleles and therefore low genetic diversity (Wright, 1931). The lowest diversity in the Oregon population might be explained by a possibly extended bottleneck in the Oregon, in contrast with a more recent N_e decline in BC-North and -South (Figure 4c); therefore, BC-North and -South are likely not in equilibrium and thus retain high π.

High genetic diversity prompts more standing genetic variation for selection to act on and thus increases the probability that a population contains genotypes that are favourable under new environmental conditions (Hughes et al., 2008; Reed & Frankham, 2003). Such genotypes can be selected for and are more likely to persist following disturbances, allowing great adaptability in future climatic scenarios. Consistent with the level of genetic diversity in BC-North and -South vs. Oregon, BC-South contained the highest number of loci under positive selection, whereas the Oregon had the least number of adaptive loci (Figure 3b). This indicates that the BC-South population has the greatest evolutionary potential. Altogether, the BC-South could be a good donor population for assisted gene flow towards the Oregon population to mitigate its genetic risk. Nonetheless, as contemporary N_e dictates potential genetic variation available for a robust evolutionary response to a changing environment (Charlesworth, 2009), large populations are thought to be less vulnerable to extinction from demographic and environmental stochasticity. The large N_e in the Oregon (Figure 4d) indicates that this population would not confront an imminent extinction risk.

4.3 | Impact of climate change on species niche suitability and population structure

Forecasts of future niche suitability (Figure 5b) indicate that the impending climate change along the coast may outstrip the species generation time, and the possibility and time of occurrence will depend on future emissions and warming scenarios. Given the same climatic scenarios, projections of future genetic clusters suggest that the impact of climate change on population genetic structure is most dramatic in the south (i.e. Oregon and BC-South), at the lagging end of the species distribution, and the impact of future climate change wanes along a northward cline across the species' distribution range (Figure 5c). As the adaptive genetic markers used for this analysis are associated with multiple traits, the projected south-to-north pattern of change in genetic clusters has twofold implications: (1) global warming has greater impacts on warm locations (i.e. southern or low

latitude); and (2) the northward genetic change and bidirectional gene flow may hasten the amalgamation of BC-South and Oregon in the future.

Collectively, the key finding of the study supports the impact of demographic dynamics on genetic diversity, and the consequences of this change can be used to infer the vulnerability and persistence of populations to respond under future climatic scenarios. Given that the Oregon population may experience a prolonged bottleneck and climate change would substantially affect its genetic structure, the Oregon may confront a relatively high risk of demographic collapses as climate changes. This evaluation points to conserving the Oregon relative to the other two populations. Broadly, the study underpins that (1) evidence of recent demographic events alone does not suffice to place a population in conservation priority; and (2) genomic diversity estimate, evolutionary potential, future niche suitability and genetic makeup are equally important to knowing whether the population has undergone a genetic diversity loss or even erosion due to recent demographic processes.

ACKNOWLEDGEMENTS

This work was supported by the University of Cambridge open access funding and funds of the Genome Canada Large-Scale Applied Research Program (POPCAN project 168BIO) to leading PIs of the project team (C. J. Douglas, Y. A. El-Kassaby, R. D. Guy, S. D. Mansfield, and Q. C. B. Cronk). I appreciate insightful suggestions from P. K. Ingvarsson (SLU) and Y. A. El-Kassaby (UBC) on the previous versions of the manuscript. I thank R. D. Guy and F. Unda (UBC) for providing background information of the experimental details of the project and Z. Guo (Sun Yat-sen Univ.) for general feedback on this work. I am equally grateful for Compute Canada (Cedar system) to undertake part of computational simulations. Finally, I extend my gratitude to previous colleagues who contributed to the project on which this study is based, and to the Editor and two anonymous reviewers for their helpful suggestions.

CONFLICT OF INTEREST

The author declares no competing interests.

DATA AVAILABILITY STATEMENT

Whole-genome shotgun sequencing data and alignment information have been deposited on SRA under the accession, PRJNA276056.

ORCID

Yang Liu https://orcid.org/0000-0002-3479-9223

REFERENCES

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T. L., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evolutionary Applications*, *1*, 95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, *93*, 1527–1539. https://doi.org/10.1890/11-1930.1

Bass, A. J., Dabney, A., & Robinson, D. (2015). *qvalue: Q-value estimation for false discovery rate control*. R package version 2.10.0. http://github.com/jdstorey/qvalue

Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, *22*, 341–345. https://doi.org/10.1093/bioinformatics/bti803

Beerli, P. (2009). *How to use MIGRATE or why are Marko chain Monte Carlo programs difficult to use?*. Cambridge University Press.

Beerli, P., & Felsenstein, J. (1999). Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, *152*, 763–773. https://doi.org/10.1093/genetics/152.2.763

Beerli, P., & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 4563–4568. https://doi.org/10.1073/pnas.081068098

Beerli, P., & Palczewski, M. (2010). Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*, *185*, 313–326. https://doi.org/10.1534/genetics.109.112532

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bougeard, S., & Dray, S. (2018). Supervised multiblock analysis in R with the *ade4* package. *Journal of Statistical Software*, *86*, 1–17.

Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring continuous and discrete population genetic structure across space. *Genetics*, *210*, 33–52. https://doi.org/10.1534/genetics.118.301333

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. *Journal of Biogeography*, *39*, 918–928. https://doi.org/10.1111/j.1365-2699.2011.02657.x

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*, 833–848. https://doi.org/10.1111/j.1365-2699.2004.01203.x

Caballero, A. (1994). Developments in the prediction of effective population size. *Heredity*, *73*, 657–679. https://doi.org/10.1038/hdy.1994.174

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*, 195–205. https://doi.org/10.1038/nrg2526

Corbett-Detig, R. B., Hartl, D. L., & Sackton, T. B. (2015). Natural selection constrains neutral diversity across a wide range of species. *PLoS Biology*, *13*, e1002112. https://doi.org/10.1371/journal.pbio.1002112

Cronk, Q. C. B. (2005). Plant eco-devo: The potential of poplar as a model organism. *New Phytologist*, *166*, 39–48. https://doi.org/10.1111/j.1469-8137.2005.01369.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

de Lafontaine, G., Ducousso, A., Lefèvre, S., Magnanou, E., & Petit, R. J. (2013). Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology*, *22*, 4397–4412. https://doi.org/10.1111/mec.12403

DeBell, D. S. (1990). *Populus trichocarpa Torr. & Gray, black cottonwood*. US Department of Agriculture, Forest Service, Agriculture Handbook.

DiFazio, S. P., Leonardi, S., Slavov, G. T., Garman, S. L., Adams, W. T., & Strauss, S. H. (2012). Gene flow and simulation of transgene dispersal from hybrid poplar plantations. *New Phytologist*, *193*, 903–915. https://doi.org/10.1111/j.1469-8137.2011.04012.x

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*, 209–214.

Dyer, R. J. (2009). GeneticStudio: A suite of programs for spatial analysis of genetic-marker data. *Molecular Ecology Resources*, *9*, 110–113. https://doi.org/10.1111/j.1755-0998.2008.02384.x

Dyer, R. J., & Nason, J. D. (2004). Population Graphs: The graph theoretic shape of genetic structure. *Molecular Ecology*, *13*, 1713–1727. https://doi.org/10.1111/j.1365-294X.2004.02177.x

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*, 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*, 422–433. https://doi.org/10.1038/nrg.2016.58

Etterson, J. R., & Shaw, R. G. (2001). Constraint to adaptive evolution in response to global warming. *Science*, *294*, 151–154. https://doi.org/10.1126/science.1063656

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., Brunner, A. M., Schackwitz, W., Gunter, L., Chen, J.-G., Tuskan, G. A., & DiFazio, S. P. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nature Genetics*, *46*, 1089–1096. https://doi.org/10.1038/ng.3075

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*, e1003905. https://doi.org/10.1371/journal.pgen.1003905

Excoffier, L., Estoup, A., & Cornuet, J. M. (2005). Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, *169*, 1727–1738. https://doi.org/10.1534/genetics.104.036236

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479–491. https://doi.org/10.1093/genetics/131.2.479

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2002). *Introduction to conservation genetics*. Cambridge University Press.

Franks, S. J., Sim, S., & Weis, A. E. (2007). Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 1278–1282. https://doi.org/10.1073/pnas.0608379104

Fraser, D. J., Hansen, M. M., Østergaard, S., Tessier, N., Legault, M., & Bernatchez, L. (2007). Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology*, *16*, 3866–3889. https://doi.org/10.1111/j.1365-294X.2007.03453.x

Galloway, G., & Worrall, J. (1979). Cladoptosis: A reproductive strategy in black cottonwood? *Canadian Journal of Forest Research*, *9*, 122–125. https://doi.org/10.1139/x79-022

Gautier, M., Klassmann, A., & Vitalis, R. (2017). REHH2.0: A reimplementation of the R package REHH to detect positive selection from haplotype structure. *Molecular Ecology Resources*, *17*, 78–90.

Geraldes, A., Farzaneh, N., Grassa, C. J., McKown, A. D., Guy, R. D., Mansfield, S. D., & Cronk, Q. C. B. (2014). Landscape genomics of *Populus trichocarpa*: The role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. *Evolution*, *68*, 3260–3280.

Gilchrist, E. J., Haughn, G. W., Ying, C. C., Otto, S. P., Zhuang, J., Cheung, D., Hamberger, B., Aboutorabi, F., Kalynyak, T., Johnson, L., Bohlmann, J., Ellis, B. E., Douglas, C. J., & Cronk, Q. C. B. (2006). Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Molecular Ecology*, *15*, 1367–1378. https://doi.org/10.1111/j.1365-294X.2006.02885.x

Goossens, B., Chikhi, L., Jalil, M. F., Ancrenaz, M., Lackman-ancrenaz, I., Mohamed, M., Andau, P., & Bruford, M. W. (2005). Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia. *Molecular Ecology*, *14*, 441–456. https://doi.org/10.1111/j.1365-294X.2004.02421.x

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913. https://doi.org/10.1038/35016000

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978. https://doi.org/10.1002/joc.1276

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*, 479–485. https://doi.org/10.1038/nature09670

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*, 609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x

Ingvarsson, P. K. (2008). Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*. *Genetics*, *180*, 329–340.

IPBES (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. IPBES. https://ipbes.net/

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jansson, S., & Douglas, C. J. (2007). *Populus*: A model system for plant biology. *Annual Review of Plant Biology*, *58*, 435–458.

Jay, F., François, O., Durand, E. Y., & Blum, M. G. B. (2015). POPS: A software for prediction of population genetic structure using latent regression models. *Journal of Statistical Software*, *68*, 1–19.

Jezkova, T., & Wiens, J. J. (2016). Rates of change in climatic niches in plant and animal populations are much slower than projected climate change. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20162104. https://doi.org/10.1098/rspb.2016.2104

Karell, P., Ahola, K., Karstinen, T., Valkama, J., & Brommer, J. E. (2011). Climate change drives microevolution in a wild bird. *Nature Communications*, *2*, 208. https://doi.org/10.1038/ncomms1213

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kuparinen, A., Savolainen, O., & Schurr, F. M. (2010). Increased mortality can promote evolutionary adaptation of forest trees to climate change. *Forest Ecology and Management*, *259*, 1003–1008. https://doi.org/10.1016/j.foreco.2009.12.006

Landais, A., Barnola, J. M., Masson-Delmotte, V., Jouzel, J., Chappellaz, J., Caillon, N., Huber, C., Leuenberger, M., & Johnsen, S. J. (2004). A continuous record of temperature evolution over a sequence of Dansgaard-Oeschger events during Marine Isotopic Stage 4 (76 to 62 kyr BP). *Geophysical Research Letters*, *31*, L22211. https://doi.org/10.1029/2004GL021193

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*, 1–18.

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

Little, E. L. (1971). *Atlas of United States trees*. US Department of Agriculture, Forest Service.

Liu, X., & Fu, Y. X. (2015). Exploring population size changes using SNP frequency spectra. *Nature Genetics*, *47*, 555–559. https://doi.org/10.1038/ng.3254

Liu, Y., & El-Kassaby, Y. A. (2019). Phenotypic plasticity of natural *Populus trichocarpa* populations in response to temporally environmental change in a common garden. *BMC Evolutionary Biology*, *19*, 231. https://doi.org/10.1186/s12862-019-1553-6

Liu, Y., & El-Kassaby, Y. A. (2021). Transcriptome-wide analysis of introgression-resistant regions reveals genetic divergence genes under positive selection in *Populus trichocarpa*. *Heredity*, *126*, 442–462. https://doi.org/10.1038/s41437-020-00388-4

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, *23*, 2178–2192.

Luikart, G., Allendorf, F. W., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, *89*, 238–247. https://doi.org/10.1093/jhered/89.3.238

McLachlan, J. S., Hellmann, J. J., & Schwartz, M. W. (2007). A framework for debate of assisted migration in an era of climate change. *Conservation Biology*, *21*, 297–302. https://doi.org/10.1111/j.1523-1739.2007.00676.x

Milne, R. I. (2006). Northern Hemisphere plant disjunctions: A window on tertiary land bridges and climate change? *Annals of Botany*, *98*, 465–472. https://doi.org/10.1093/aob/mcl148

Müller, N. A., Kersten, B., Leite Montalvão, A. P., Mähler, N., Bernhardsson, C., Bräutigam, K., Carracedo Lorenzo, Z., Hoenicka, H., Kumar, V., Mader, M., Pakull, B., Robinson, K. M., Sabatti, M., Vettori, C., Ingvarsson, P. K., Cronk, Q., Street, N. R., & Fladung, M. (2020). A single gene underlies the dynamic evolution of poplar sex determination. *Nature Plants*, *6*, 630–637. https://doi.org/10.1038/s41477-020-0672-9

Nei, M., Chakraborty, R., & Fuerst, P. A. (1976). Infinite allele model with varying mutation rate. *Proceedings of the National Academy of Sciences of the United States of America*, *73*, 4164–4168. https://doi.org/10.1073/pnas.73.11.4164

Nei, M., & Graur, D. (1984). Extent of protein polymosphism and the neutral mutation theory. *Evolutionary Biology*, *17*, 73–118.

Ohta, T. (1992). The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, *23*, 263–286. https://doi.org/10.1146/annurev.es.23.110192.001403

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics*, *28*, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*, 1929–1936. https://doi.org/10.1093/molbev/msu136

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. https://doi.org/10.1093/genetics/155.2.945

Quintero, I., & Wiens, J. J. (2013). Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecology Letters*, *16*, 1516.

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, *17*, 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Rogers, A. R. (2014). How population growth affects linkage disequilibrium. *Genetics*, *197*, 1329–1341. https://doi.org/10.1534/genetics.114.166454

Rood, S. B., Braatne, J. H., & Hughes, F. M. (2003). Ecophysiology of riparian cottonwoods: Stream flow dependency, water relations and restoration. *Tree Physiology*, *23*, 1113–1124. https://doi.org/10.1093/treephys/23.16.1113

Rood, S. B., & Mahoney, J. M. (1990). Collapse of riparian poplar forests downstream from dams in western prairies: Probable causes and prospects for mitigation. *Environmental Management*, *14*, 451–464. https://doi.org/10.1007/BF02394134

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Schoener, T. W. (1968). Anolis lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, *49*, 704–726.

Shaw, R. G., & Etterson, J. R. (2012). Rapid climate change and the rate of adaptation: Insight from experimental quantitative genetics. *New Phytologist*, *195*, 752–765. https://doi.org/10.1111/j.1469-8137.2012.04230.x

Siewert, K. M., & Voight, B. F. (2017). Detecting long-term balancing selection using allele frequency correlation. *Molecular Biology and Evolution*, *34*, 2996–3005. https://doi.org/10.1093/molbev/msx209

Slavov, G. T., DiFazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., & Tuskan, G. A. (2012). Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist*, *196*, 713–725.

Slavov, G. T., Leonardi, S., Burczyk, J., Adams, W. T., Strauss, S. H., & Difazio, S. P. (2009). Extensive pollen flow in two ecologically contrasting populations of *Populus trichocarpa*. *Molecular Ecology*, *18*, 357–373.

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, *64*, 295–315. https://doi.org/10.1111/j.1558-5646.2009.00877.x

Soltis, D. E., Gitzendanner, M. A., Strenge, D. D., & Soltis, P. S. (1997). Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, *206*, 353–373. https://doi.org/10.1007/BF00987957

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*, 583–639. https://doi.org/10.1111/1467-9868.00353

Suarez-Gonzalez, A., Hefer, C. A., Lexer, C., Douglas, C. J., & Cronk, Q. C. B. (2018). Introgression from *Populus balsamifera* underlies adaptively significant variation and range boundaries in *P. trichocarpa*. *New Phytologist*, *217*, 416–427.

Tenesa, A., Navarro, P., Hayes, B. J., Duffy, D. L., Clarke, G. M., Goddard, M. E., & Visscher, P. M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome Research*, *17*, 520–526. https://doi.org/10.1101/gr.6023607

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., de Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Townsend Peterson, A., Phillips, O. L., & Williams, S. E. (2004). Extinction risk from climate change. *Nature*, *427*, 145–148. https://doi.org/10.1038/nature02121

Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G. F., Paterson, J., Schurr, F. M., Sykes, M.

T., & Zimmermann, N. E. (2008). Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*, 137–152. https://doi.org/10.1016/j.ppees.2007.09.004

Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T., & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8245–8250. https://doi.org/10.1073/pnas.0409902102

Toro, M. A., & Caballero, A. (2005). Characterization and conservation of genetic diversity in subdivided populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 1367–1378. https://doi.org/10.1098/rstb.2005.1680

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., & Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, *313*, 1596–1604.

Urban, M. C. (2015). Accelerating extinction risk from climate change. *Science*, *348*, 571–573. https://doi.org/10.1126/science.aaa4984

Wang, T. L., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One*, *11*, e0156720. https://doi.org/10.1371/journal.pone.0156720

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, *62*, 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x

Williams, J. W., & Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, *5*, 475–482. https://doi.org/10.1890/070037

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*, 97–159. https://doi.org/10.1093/genetics/16.2.97

Xie, C.-Y., Ying, C. C., Yanchuk, A. D., & Holowachuk, D. L. (2009). Ecotypic mode of regional differentiation caused by restricted gene migration: A case in black cottonwood (*Populus trichocarpa*) along the Pacific Northwest coast. *Canadian Journal of Forest Research*, *39*, 519–525. https://doi.org/10.1139/X08-190

Xu, S. H., Gupta, S., & Jin, L. (2010). PEAS V1.0: A package for elementary analysis of SNP data. *Molecular Ecology Resources*, *10*, 1085–1088. https://doi.org/10.1111/j.1755-0998.2010.02862.x

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*, 2486–2499.

BIOSKETCH

Yang Liu is a Research Associate at the University of Cambridge. His research seeks to reveal a role for demographic, ecological and genetic mechanisms in plant adaptive evolution, and how adaptive and non-adaptive evolutionary processes have shaped biodiversification. His ongoing work is part of the "Crops, Pollinators and People" project, in an effort to demonstrate the expansion and evolution of buckwheat across Eurasia in relation to honeybee populations and human management in prehistory.

Author contribution: Y.L. conceived this study, performed data analyses, and wrote the manuscript. Two advisors, Pär K. Ingvarsson (SLU) and Yousry A. El-Kassaby (UBC), provided quite a few insightful ideas and suggestions on this work.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Liu, Y. (2022). Conservation prioritization based on past cascading climatic effects on genetic diversity and population size dynamics: Insights from a temperate tree species. *Diversity and Distributions*, 00, 1–17. https://doi.org/10.1111/ddi.13490