Received: 21 February 2024 | Revised: 24 May 2024 | Accepted: 30 May 2024

DOI: 10.1111/mec.17430

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genomic insights into hybrid zone formation: The role of climate, landscape, and demography in the emergence of a novel hybrid lineage

Constance E. Bolte[1] | Tommy Phannareth[2] | Michelle Zavala-Paez[1] | Brianna N. Sutara[3] | Muhammed F. Can[4] | Matthew C. Fitzpatrick[5] | Jason A. Holliday[2] | Stephen R. Keller[6] | Jill A. Hamilton[1]

[1]Department of Ecosystem Science and Management, Pennsylvania State University, University Park, Pennsylvania, USA

[2]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, USA

[3]Department of Biology, Pennsylvania State University, University Park, Pennsylvania, USA

[4]Faculty of Forestry, Duzce University, Duzce, Turkey

[5]Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, Maryland, USA

[6]Department of Plant Biology, University of Vermont, Burlington, Vermont, USA

Correspondence
Jill A. Hamilton, Department of Ecosystem Science and Management, Pennsylvania State University, University Park, Pennsylvania 16801, USA.
Email: jvh6349@psu.edu

Funding information
National Institute of Food and Agriculture, Grant/Award Number: PEN04809 and VA136641; National Science Foundation, Grant/Award Number: PGR-1856450

Handling Editor: Loren Rieseberg

Abstract

Population demographic changes, alongside landscape, geographic and climate heterogeneity, can influence the timing, stability and extent of introgression where species hybridise. Thus, quantifying interactions across diverged lineages, and the relative contributions of interspecific genetic exchange and selection to divergence at the genome-wide level is needed to better understand the drivers of hybrid zone formation and maintenance. We used seven latitudinally arrayed transects to quantify the contributions of climate, geography and landscape features to broad patterns of genetic structure across the hybrid zone of *Populus trichocarpa* and *P. balsamifera* and evaluated the demographic context of hybridisation over time. We found genetic structure differed among the seven transects. While ancestry was structured by climate, landscape features influenced gene flow dynamics. Demographic models indicated a secondary contact event may have influenced contemporary hybrid zone formation with the origin of a putative hybrid lineage that inhabits regions with higher aridity than either of the ancestral groups. Phylogenetic relationships based on chloroplast genomes support the origin of this hybrid lineage inferred from demographic models based on the nuclear data. Our results point towards the importance of climate and landscape patterns in structuring the contact zones between *P. trichocarpa* and *P. balsamifera* and emphasise the value whole genome sequencing can have to advancing our understanding of how neutral processes influence divergence across space and time.

KEYWORDS
climate, forest trees, gene flow, natural hybridisation, *Populus balsamifera*, *Populus trichocarpa*, speciation

1 | INTRODUCTION

Understanding the processes influencing the formation and maintenance of species is a central goal of evolutionary biology and is crucial to management and conservation of biodiversity in a rapidly changing environment (Frankham, 2010; Hoffmann et al., 2015; Payseur & Rieseberg, 2016). As such, hybridisation and introgression have long been of interest (Dobzhansky, 1936; Heiser, 1949; Jeffrey, 1916; Stebbins, 1959), and recent advances in population-scale genome sequencing have revolutionised our understanding of the frequency and pervasiveness of hybridisation in nature (Hamilton & Miller, 2016; Janes & Hamilton, 2017; Mallet, 2005; VanWallendael et al., 2022). Multiple taxa are known to form natural hybrid zones; however, they are particularly prevalent among forest tree species (Abbott, 2017; Suarez-Gonzalez, Hefer, et al., 2018; Suarez-Gonzalez, Lexer, & Cronk, 2018; Swenson & Howard, 2004). Climatic oscillations during glacial and interglacial periods have influenced the distributions of many temperate and boreal trees, facilitating opportunities for interspecific gene flow during times of secondary contact (Hamilton et al., 2015; Hewitt, 2004; Jump & Peñuelas, 2005; Soltis et al., 2006; Swenson & Howard, 2004). However, despite widespread observations of hybridisation across forest tree species, gaps remain in our understanding of how demographic, neutral and nonneutral evolutionary processes influence the formation and maintenance of natural hybrid zones across space and time. Teasing apart these processes requires an assessment of the historical, landscape and climatic factors that underlay hybrid zone formation. Ultimately, understanding how population demographic changes, alongside landscape heterogeneity and climate adaptation contribute to the timing, stability and extent of hybridisation will be critical to managing species in a rapidly changing climate.

Population demographics, including expansion and contraction of a species' range can have substantial influence on standing genetic variation, impacting the evolutionary trajectory of populations across space and time. Hybrid zones often form when previously isolated lineages come into contact and the impacts of these processes, including genetic drift and gene flow, likely play an important role shaping the genetic structure of a contact zone (Abbott, 2017; Barton & Hewitt, 1989). Shifts between isolation and contact can coincide with changes in effective population size (N_e) influencing the direction and extent of genetic exchange, while fine-scale variance at the population or genome-level can further influence elimination or fixation of alleles. In forest trees, genetic exchange often appears to be largely asymmetrical favouring the genomic background of one species due solely or in part to differential dispersal capacity, wind-patterns and unidirectional reproductive incompatibilities (El Mujtar et al., 2017; Hamilton et al., 2013a; Lepais et al., 2009; Lexer et al., 2005, 2006). Given that an influx of genetic variants via interspecific gene flow can increase N_e and generate novel genetic recombinants upon which natural selection acts, characterising genetic structure and the impact genetic exchange may have across space and time is needed to predict evolution.

Demographic inference can be particularly useful for characterising processes underlying patterns of genetic exchange that can influence species evolutionary relationships, distributional shifts, or adaptive potential (Bacilieri et al., 1996; Petit et al., 2004). Several forest tree hybrid zones have characterised patterns of introgression, including oaks (Cannon & Petit, 2020; Eaton et al., 2015; Leroy et al., 2020; McVay et al., 2017), poplars (Chhatre et al., 2018; Christe et al., 2017; Lexer et al., 2005; Suarez-Gonzalez et al., 2016; Suarez-Gonzalez, Hefer, et al., 2018; Suarez-Gonzalez, Lexer, & Cronk, 2018) and spruce (Hamilton et al., 2013a, 2013b). However, few studies have leveraged whole genome sequencing to tease apart the temporal and spatial dynamics underlying the evolutionary history of contact zones using both biparentally and uniparentally inherited genomes. Viewing distinct evolutionary trajectories associated with biparentally and uniparentally inherited genomes through the lens of demographic change and landscape heterogeneity enables assessment of the role these evolutionary processes may play to the formation and maintenance of natural hybrid zones.

Here, we use *Populus*, an ecologically and economically important model system in forest trees that naturally hybridise with congeners where geographical ranges overlap. *Populus trichocarpa*, the first tree to have its genome fully sequenced (Tuskan et al., 2006), has played an extensive role in our understanding of tree genome biology, comparative genomics and adaptive introgression (e.g. Jansson & Douglas, 2007; Shang et al., 2020; Suarez-Gonzalez et al., 2016). Natural hybrid zones exist between *P. trichocarpa* and *P. balsamifera*, largely associated with geographically and climatically steep transitions from maritime to continental climates west and east of the Rocky Mountains. However, divergence has likely involved a combination of intrinsic and extrinsic factors, including a dynamic history of population size change throughout glacial and interglacial periods, landscape-related barriers to gene flow, and climatically-related responses to selection (Geraldes et al., 2014; Keller et al., 2010; Levsen et al., 2012; Slavov et al., 2012). We compare whole-genome sequence data for biparentally (nuclear) and uniparentally (chloroplast) inherited genomes to quantify divergence and the evolutionary relationship between species across space and time. This will provide an understanding of the role neutral and non-neutral processes play in the formation of long-lived hybrid zones (Petit et al., 2005).

Using a sampling design of latitudinally arrayed transects, we sequenced whole genomes for 576 trees to capture parental-types and hybrids across repeated zones of contact between *P. trichocarpa* and *P. balsamifera* distributed across their entire range of overlap. With these data, we observed genetic structure within and across each contact zone and asked (1) How has climate, geography and landscape-level barriers influenced genetic structure and (2) What is the history and extent of interspecific gene flow? These data provide new insights into how different evolutionary processes contribute to the formation and persistence of hybrid zones over varying spatial and temporal scales.

2 | MATERIALS AND METHODS

2.1 | Sampling of plant material

In January 2020, vegetative branch cuttings were collected from 576 poplar trees spanning repeated natural contact zones between *Populus trichocarpa* and *Populus balsamifera* (Figure 1a). These represent individual collections from seven latitudinally distributed transects (hereafter referred to as Alaska_1, Alaska_2, Cassiar, Chilcotin, Jasper, Crowsnest and Wyoming) spanning much of the species' overlapping distributions, from 40° N to 65° N and −100° W to −150° W longitude. On average the distance between samples was approximately 43 kilometres (range: <1–962 km). Vegetative branch cuttings were transported back to a greenhouse at Virginia Tech (Blacksburg, VA, USA) for propagation. Details associated with propagation and greenhouse conditions are provided in Appendix S1 of Supplementary S1. Young leaf tissue was sampled from each of the 576 propagated branch cuttings, placed immediately on dry ice, and then transferred to a −80°C freezer for long-term preservation.

2.2 | Genomic library preparation and bioinformatics

For each of the 576 samples, DNA was extracted from approximately 100 mg of leaf tissue using the Qiagen plant DNeasy kit, with one modification (details available in Supplementary S1; Appendix S2). Genomic DNA libraries were constructed at the Duke University Center for Genomic and Computational Biology, using an Illumina DNA Prep kit (Illumina Inc., San Diego, USA). Genomic libraries were sequenced on an S4 flow cell in 2 × 150 bp format on an IlluminaNovaSeq 6000 instrument with 64 samples per lane. De-indexing, QC, trimming adapter sequences, and sequence preprocessing were completed by the sequencing facility.

Subsequent bioinformatic analysis was performed on Virginia Tech's Advanced Research Computing System using Burrows-Wheeler Aligner to map reads to the *P. trichocarpa* reference genome (v4.0), and the resulting SAM files were converted to BAM format with SAMtools (Li & Durbin, 2010). The Genome Analysis Toolkit (v3.7) Haplotype Caller algorithm was then used to generate individual gVCF files, which were merged in a single VCF file with the GATKGenotype GVCFs function. The raw VCF had ~82 million variant sites, for which missing data were on average 0.7% (SD: 0.1%–1.3%) for 575 of the 576 samples. One sample (GPR-14) failed and was removed. Candidate variants were filtered to remove singletons and those with >10% missingness, those with poor map quality (MQ < 40), elevated strand bias (FS > 40, SOR > 3), differential map quality between reference and alternative alleles (MQRankSum < −12.5), positional bias between reference and alternate alleles (ReadPosRankSum < −8), or low quality-by-depth (QD < 2). Finally, we removed INDELS and SNPs with >2 alternate alleles, which resulted in a final dataset comprised of 29,663,130 bi-allelic SNPs.

2.3 | Analysis of population structure

To account for potential sampling of non-focal poplar species and remove those samples, we merged the genetic data for our 575 samples with genetic data available for 48 samples of *P. angustifolia*, which overlaps in portions of the southern distribution of our sampling region (Chhatre et al., 2018). Methods in Supplementary S1 (Appendix S3) detail the analyses used. While we found no need to remove samples due to admixture with *P. angustifolia* (Q < 0.01 across all 575 samples), we removed 29 of the 575 samples with unknown ancestry (Q > 0.01; Supplementary S1; Figures S1 and S2), leaving 546 samples for subsequent genetic analyses.

To assess the genetic structure of the contact zones between *P. balsamifera* and *P. trichocarpa*, we first thinned variant sites in high linkage disequilibrium using *Plink* (v1.9) based on 10,000 bp windows sliding in increments of 1000 bp along each chromosome, flagging site pairs with an $r^2 > 0.1$. One site for each pair exceeding the r^2 threshold, along with singletons (--mac 2) and those with missing data (--max-missing 1.0), were removed using *vcftools*, version 0.1.16. The resulting 299,335 low LD variant sites were used to assess heterozygosity, identity-by-descent (*Plink* v1.9) and visualise the scale and spatial extent of hybridisation.

Genomic structure was assessed using a principal component analysis (PCA) with the 'prcomp' function of the *stats* package in R, version 4.2.1. We then performed an ADMIXTURE analysis (Alexander et al., 2009) using cluster assignments ranging from $K = 2$ to $K = 6$. The cluster assignment (K) with the lowest cross validation error (k-folds = 10) was determined to best fit the data.

2.4 | Assessing evolutionary relationships from nuclear data

To evaluate the evolutionary relationships between genetic clusters, we used the three genetic clusters identified from ADMIXTURE ($K = 3$; see Results) to determine the topology with the greatest support. We identified parental-type samples for each genetic cluster, hereafter referred to as *P. balsamifera*, *coastal P. trichocarpa*, and *interior P. trichocarpa*, based on K = 3 using an ancestry coefficient threshold of Q > 0.90. Of the 546 samples included in the analysis, 185 parental-types were identified (112 *P. balsamifera*, 51 *coastal P. trichocarpa* and 22 *interior P. trichocarpa*). Using parental-types, we evaluated three different topologies based on the proposed divergence history between *interior P. trichocarpa* and the other two parental-type genetic lineages (Table 1). The first topology evaluated expectations based on a recent divergence from *coastal P. trichocarpa*, while the second tested for a recent divergence from *P. balsamifera* and the third evaluated whether *interior P. trichocarpa* was a hybrid lineage formed from an admixture event between *coastal P. trichocarpa* and *P. balsamifera*.

The evolutionary relationship of the three genetic clusters was evaluated using Diffusion Approximation for Demographic Inference (*∂α∂i* v.2.1.1; Gutenkunst et al., 2009). We used the

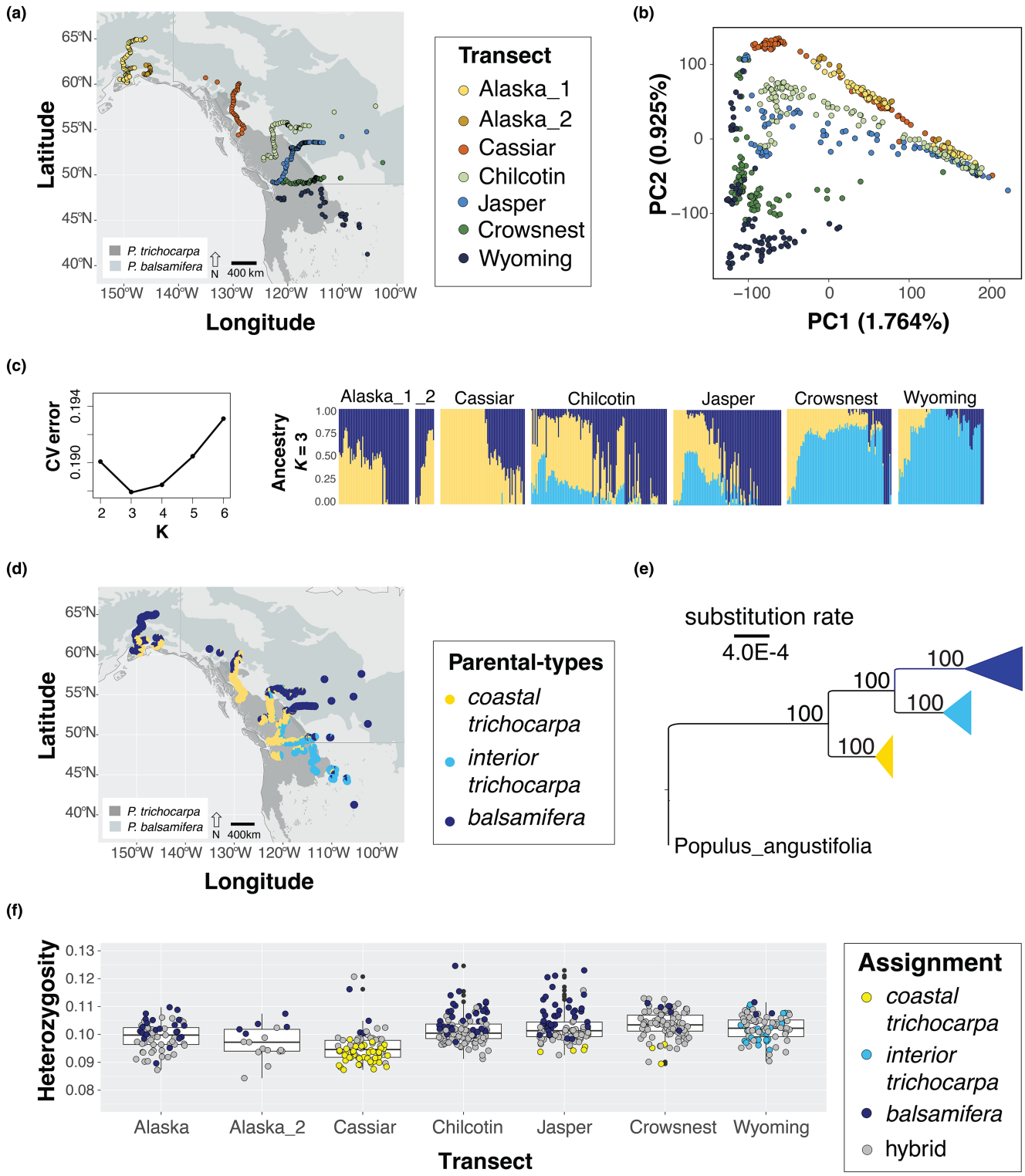
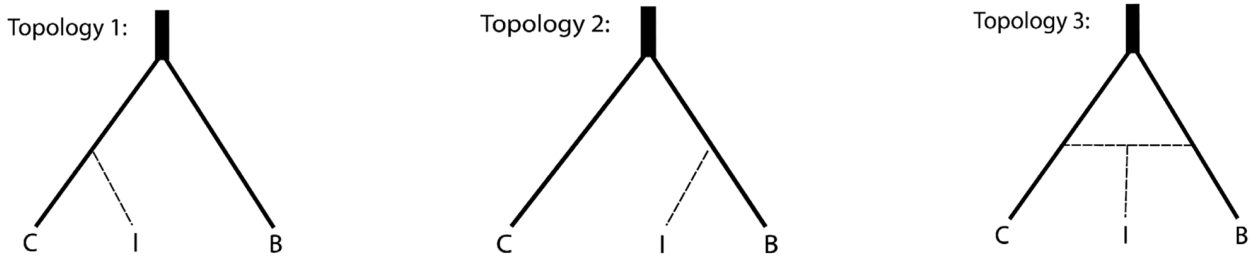


FIGURE 1 Spatial patterns of genetic structure and ancestry across 546 trees sampled from (a) seven transects of the *Populus balsamifera* and *P. trichocarpa* hybrid zone, labelled by transect; (b) transect-specific genomic structure based on principal component analysis for PC1 and PC2; (c) results of admixture analysis for each sample at $K = 3$, the best supported cluster assignment on the basis on lowest cross validation (CV) error; (d) geographical distribution of $K = 3$ ancestry proportions illustrated as pie charts; (e) ML phylogeny of chloroplast evolutionary relationships among parental-type samples from $K = 3$ admixture results ($Q > 0.90$ for each cluster). Bootstrap confidence is provided on each branch. (f) Boxplots of the distribution of observed heterozygosity for each sample is provided with each sample colour-coded based on its respective $K = 3$ ancestral assignment (i.e. parental-type or hybrid).

TABLE 1 Demographic inference AIC comparisons across three-population models.

Model[a]	*k*	AIC	Delta AIC
(a)			
Topology 1: SI	6	68,945.17	4942.96
Topology 2: SI	6	75,772.18	11,769.97
Topology 3: SI	6	64,002.21	**0.00**
(b)			
Topology 3: T2_SC	12	34,935.00	136.26
Topology 3: T2_IM	12	43,391.55	8592.80
Topology 3: T3_IM_SC	14	**34,798.75**	**0.00**
Topology 3: T4_SC_SC	14	35,426.09	627.34



Note: Model abbreviations are defined in the footnote. The lower the AIC and delta AIC the better the model fit to the data. The first three rows (a) are results from the topology test to identify the evolutionary relationship of *interior P. trichocarpa* (I) to *P. balsamifera* (B) and *coastal P. trichocarpa* (C). The following four rows (b) are results of four models with gene flow parameters tested in various scenarios under the best fit topology (Topology 3) of admixed origin. Bolded models in each of the two inference routines signify best fit with delta AIC = 0.00.

[a]Models; SI (strict isolation), IM (isolation-with-migration), SC (secondary contact), T# (number of time intervals considered).

unfolded joint site frequency spectrum (jSFS) of derived allele counts following polarisation of genotypes using the approach of Luqman et al. (2023). Details pertaining to data sources and program implementation are provided in Supplementary S1, Appendix S4 and Table S1. We assigned allelic states as either ancestral or derived based on the consensus ancestral state: 6495 sites (2.260%) had the reference allele assigned to the derived state and 275,751 sites (95.959%) had the reference allele assigned to the ancestral state. The remaining 1.781% of sites had undefined ancestral alleles. To increase computational efficiency, we used the derived allele frequencies for a random sample of 10 diploid individuals per parental-type cluster to build an unfolded three-dimensional jSFS using the function 'GenerateFs' in the dadi-cli program (Huang et al., 2023). We performed 10 replicate runs of each model in *∂α∂i* with a 40 × 50 × 60 grid space and the nonlinear Broyden–Fletcher–Goldfarb–Shannon (BFGS) optimisation routine. Akaike information criterion (AIC; Akaike, 1987) was employed for model selection, with the best replicate run (highest log composite likelihood) for each model used to calculate ΔAIC ($AIC_{model\ i} - AIC_{best\ model}$) scores (Burnham & Anderson, 2002).

To estimate the timing of divergence and quantify the extent and directionality of gene flow between genetic clusters, we tested four hypotheses, or models, of connectivity and divergence based on the best supported topology (Table 1). The first model assumed a scenario of recurrent gene flow, while the other three models assumed scenarios of isolation that varied based on the timing and frequency of secondary contact (Supplementary S1: Figure S3). We performed 10 replicate runs of each model in *∂α∂i* with a 40 × 50 × 60 grid space and the nonlinear BFGS optimisation routine. AIC (Akaike, 1987) was employed for model selection, and Fisher Information Matrix-based uncertainty analysis was then conducted for the best supported model to obtain upper and lower 95% confidence intervals (CIs) for the inferred parameter estimates. To unscale parameter estimates and their 95% CIs we used a mutation rate of 1.33×10^{-10} substitutions/site/year rate for *Populus trichocarpa* (Hofmeister et al., 2020), a generation time of 15 years (Hofmeister et al., 2020; Ingvarsson, 2008), and a calculated effective genome length (Supplementary S3). To calculate effective genome length, we multiplied the proportion of variant sites used in the modelling (count of filtered sites in each pairwise inference with no missing data and singletons removed/count of variant sites prior to filtering) by the number of base pairs from the *P. trichocarpa* reference genome considered in variant calling (389,204,664 bp).

2.5 | Assessing evolutionary relationships from chloroplast data

To gain insight into the evolutionary relationships among the three genetic clusters identified from nuclear data, we used the 185 parental-types to examine differentiation across uni-parentally inherited chloroplast genomes. Chloroplast (cp) genomes were assembled using NOVOPlasty (Dierckxsens et al., 2017). The published

Nisqually-1 *P. trichocarpa* chloroplast genome (NC_009143.1) was used as the reference sequence for genome assembly. The ribulose-bisphosphate carboxylase (*rbc*L) gene was manually extracted and used as a seed, serving as an initial reference point for the chloroplast genome assembly. Assembled chloroplast genomes were annotated using Geneious v7.1.4 (Kearse et al., 2012). Following annotation, inverted repeats and coding regions were manually checked for each annotated individual. In total, 184 out of the 185 parental-types were successfully assembled, with one individual (genotype 831) excluded due to incomplete assembly.

A maximum likelihood (ML) phylogenetic tree was built to assess the relationships among the chloroplast genomes for each parental type. The chloroplast genomes from two *P. angustifolia* samples (NC_037413 and MW376761) were downloaded from NCBI and selected as an outgroup. Chloroplast genomes for all 184 samples and outgroups were aligned using MAFFT V.7.453 (Katoh & Standley, 2013) to perform a ML phylogeny in IQ-tree 1.6.11 (Trifinopoulos et al., 2016). The best fit evolutionary model for the sequence alignment was calculated by ModelFinder (Kalyaanamoorthy et al., 2017), with the selected model being the transversion model with empirical base frequencies and FreeRate model (TVM + F + R2). A ML tree was generated following 1000 bootstrap replications and a consensus phylogenetic tree was created based on all parental-type samples using the FigTree software v1.4.13.

2.6 | Associating genetic structure with climate and geography

To quantify the multivariate relationships between genetic structure, climate and geographical variation across the hybrid zone, a redundancy analysis (RDA) was conducted within the *vegan* package, version 2.6–2 (Oksanen et al., 2020) in R version 4.2.1. Climate data at 30-second resolution averaged across the years 1961–1990 for 25 variables associated with geographical origin were extracted using *ClimateNA* (version 5.2; Wang et al., 2016) for each sampled tree (Supplementary S2) and then tested for collinearity using variance inflation factor (VIF) analysis implemented in the *usdm* R package (version 1.1–18; Naimi et al., 2014). Six climatic variables with a VIF less than 10 were retained for subsequent analysis, including continentality (TD), mean annual temperature (MAT), mean annual precipitation (MAP), climate moisture deficit (CMD), relative humidity (RH) and precipitation as snow (PAS). Geographical variables included for subsequent analysis were latitude, longitude and elevation. We examined correlations between geographical and climate variables using the 'cor' function and Spearman method available in base R. Pairwise correlation coefficients were less than or equal to $|r^2| \leq 0.70$ with one exception. Latitude and MAT was $r^2 = -0.82$ (Figure S4).

LD-pruned nuclear genotypic data (described above; 299,335 variants) was converted to minor allele counts per sample using vcftools for all 546 samples (version 0.1.16, output format --012). A permutation-based analysis of variance (ANOVA) procedure with 999 permutations (Legendre & Legendre, 2012) was used to test the overall statistical significance of the RDA model ($\alpha = 0.05$) as well as the significance of each RDA axis separately. Variance partitioning was then performed using the 'varpart' function of the *vegan* package to quantify the influence of predictor variables and their confounding effects to observed genetic variation. This analysis was repeated without hybrid samples ($0.10 < Q < 0.90$) to compare the multivariate relationships among the 185 parental-type samples. The association between genetic variation and environment was also performed for the 2828 chloroplast SNPs.

To validate variance partitioning and RDA, the relationship between genetic distances and climate distance (i.e. isolation-by-environment) and genetic distance with geographical distance (i.e. isolation-by-distance) was evaluated using partial Mantel tests within the *vegan* package. We calculated pairwise Euclidean distances for genetic data (scores for the first and second axes of the genetic PCA), geographical data (longitude and latitude), and climate data (scores for the first and second axes of the climate PCA). The climate PCA for the 546 samples (Supplementary S1: Figure S5) was informed by the six climate variables included in the RDA described above (i.e. no genetic constraints). Statistical significance of each partial Mantel test was determined from comparisons of the empirical data to 999 random permutations of the data.

2.7 | Visualising barriers to gene flow

To evaluate genetic connectivity across the contact zones between *P. trichocarpa* and *P. balsamifera*, we used estimated effective migration surfaces (EEMS; Petkova et al., 2016) and observed correspondence between landscape features, ancestry and patterns of genetic connectivity. A genetic dissimilarity matrix was generated using the 'bed2diffs script' included within the EEMS program informed by the 299,335 genome-wide nuclear SNPs across 546 individuals. Two additional geographical files were created using geographical coordinates; including the location data of each individual (.coord extension) and a large general habitat polygon (.outer extension) which was manually obtained using Google Maps API v3 Tool. The geographical coordinates for each sample alongside the parameterised grid space (500 demes with each grid cell ~21,000 km^2) assigned individuals to one of 49 demes, defined as pseudo populations of close geographical proximity to a vertex of the defined grid space (Supplementary S1: Figure S6). We then estimated the effective migration rate using a stepping-stone model. Model parameters included a burn-in of 25,000,000 with an MCMC length of 50,000,000 and thinning to sample 1 in every 100,000 iterations. The results of three replicate runs were averaged to visualise effective diversity (q) and effective migration rate (m) surfaces (in $\log_{10}$ scale) across the contact zone using the R package *rEEMSplots* v 0.0.1 (Petkova et al., 2016).

3 | RESULTS

3.1 | Genomic structure of hybrid zone identifies three distinct genetic clusters

Using all 546 individuals, the nuclear genomic PCA ascribed genetic structure along PC1 (1.764% variance explained) to species-specific differences between *P. trichocarpa* (far left of PC1) and *P. balsamifera* (far right of PC1) and PC2 (0.925% variance explained) to genetic structure within *P. trichocarpa* (Figure 1b). Admixture analysis best supported the presence of three genetic clusters ($K=3$), with admixed individuals present across all transects (Figure 1c). Two of the three genetic groups clustered geographically within the range of *P. trichocarpa* (Little Jr., 1971) – this included a *coastal P. trichocarpa* cluster which had a northern, coastal distribution, and an *interior P. trichocarpa* cluster, which had a southern, interior distribution (Figure 1d). The third genetic cluster was *P. balsamifera* in the $K=3$ admixture analysis. Pairwise estimates of genome-wide genetic differentiation (F_{ST}) were similar across the three genetic clusters, ranging from 0.053 for *coastal – interior P. trichocarpa*, 0.054 for *interior P. trichocarpa – P. balsamifera* and 0.056 for *coastal P. trichocarpa – P. balsamifera* comparisons. The cross-validation error for the cluster assignment of $K=4$ was comparable with $K=3$ (Figure 1c). At $K=4$, an additional genetic cluster differentiated Alaska and Cassiar genotypes in the north from the $K=3$ *coastal P. trichocarpa* cluster (Supplementary S1: Figure S7).

Of the 546 samples included in the analysis, 185 parental-types associated with $K=3$ genetic clusters were identified based on $Q>0.90$ (112 individuals identified as *P. balsamifera*, 51 individuals identified as *coastal P. trichocarpa* and 22 individuals identified as *interior P. trichocarpa). Coastal P. trichocarpa* and *P. balsamifera* genetic ancestry was observed in Alaska and Cassiar transects, although the Alaska transect did not appear to have any *coastal P. trichocarpa* parental-types. All three genetic clusters were observed in Chilcotin, Jasper, Crowsnest and Wyoming transects (Figure 1c). Shared ancestry associated with *P. balsamifera* along the Crowsnest transect was restricted to the eastern edge of the *P. trichocarpa* distribution. The Wyoming transect was dominated by *interior P. trichocarpa* parental-type ancestry ($Q>0.90$ at $K=3$; Supplementary S2), whereas individuals sampled near the western edge of this transect exhibited mixed ancestry between the coastal and interior genetic clusters of *P. trichocarpa*. Consistent with the Crowsnest transect, *P. balsamifera* ancestry was limited to the far eastern distribution of *P. trichocarpa*, with *P. balsamifera* parental-types restricted to fragmented populations geographically distant from the main distribution of either species.

On average, identity-by-descent (PI_HAT) was 0.17 across pairwise comparisons (range 0.00–0.78 across the 148,785 pairwise assessments; Supplementary S1: Figure S8). Only 11 pairwise comparisons (0.007%) had a PI_HAT greater than 0.50 (Supplementary S1: Table S2). Given the low degree of kinship observed, we used all 546 genotypes to assess the scale and spatial extent of hybridisation. Observed heterozygosity ranged from 0.084 to 0.125 with an average of $H_o=0.101$ across all samples (Supplementary S2). The distribution of genetic variation between parental-types and associated hybrids were compared across transects using the $K=3$ determination of parental-type and hybrid individuals (Figure 1f). *Coastal P. trichocarpa* parental-types exhibited reduced heterozygosity ($H_o=0.093$, range = 0.087–0.104; Supplementary S1: Figure S9) relative to other parental-types and were largely confined to three transects (Cassiar, Jasper, and Crowsnest; Figure 1f). In contrast, parental-types of *P. balsamifera* were present across all seven contact zones and on average, exhibited greater heterozygosity ($H_o=0.106$, range = 0.089–0.125; Supplementary S1: Figure S9). *Interior P. trichocarpa* parental-types and hybrids exhibited intermediate values of heterozygosity ($H_o=0.101$ and $H_o=0.100$), but *interior P. trichocarpa* was solely observed within the Wyoming contact zone (Figure 1f) while hybrids of varying ancestral backgrounds were observed across the sampled range (Figure 1).

3.2 | Nuclear and chloroplast genomes identify three genetic lineages

Given the similar estimates of genome-wide nuclear differentiation (F_{ST}), we performed a topology test to better understand the relationships among the three genetic clusters. The best fit topology as determined by AIC (Table 1a) supported *interior P. trichocarpa* as an ancient lineage of admixed origin. The second-best fit topology placed *interior P. trichocarpa* as a split from *coastal P. trichocarpa*, but delta AIC was 4943 units higher.

To complement assessment of nuclear topologies, evolutionary relationships were also assessed using ML analysis of chloroplast genomes. The alignment of the 184 individuals and outgroup genomes produced a single matrix with 164,183 nucleotide sites, including 1020 polymorphic sites. Three distinct clades were detected using the ML tree, including *coastal P. trichocarpa*, *interior P. trichocarpa* and *P. balsamifera* respectively (Figure 1e). The *coastal P. trichocarpa* clade appeared ancestral to both *interior P. trichocarpa* and *P. balsamifera*, with all three exhibiting robust bootstrap support. Lineage assignments were congruent between nuclear and chloroplast genomes, except for individual 757, which clustered with *interior P. trichocarpa* in the nuclear admixture analysis but in the chloroplast ML tree belonged to the *P. balsamifera* clade.

3.3 | Climate strongly associates with genetic structure

Climate and geography explained 3.79% (r^2) of the genetic variance across the 546 sampled trees and 299,335 SNP dataset. Most of the explainable genetic variance (PVE) was accounted for by the first RDA axis (38.26%, Figure 2a) with TD and MAT having the highest predictor loadings (eigenvectors). Individuals identified as *P.*

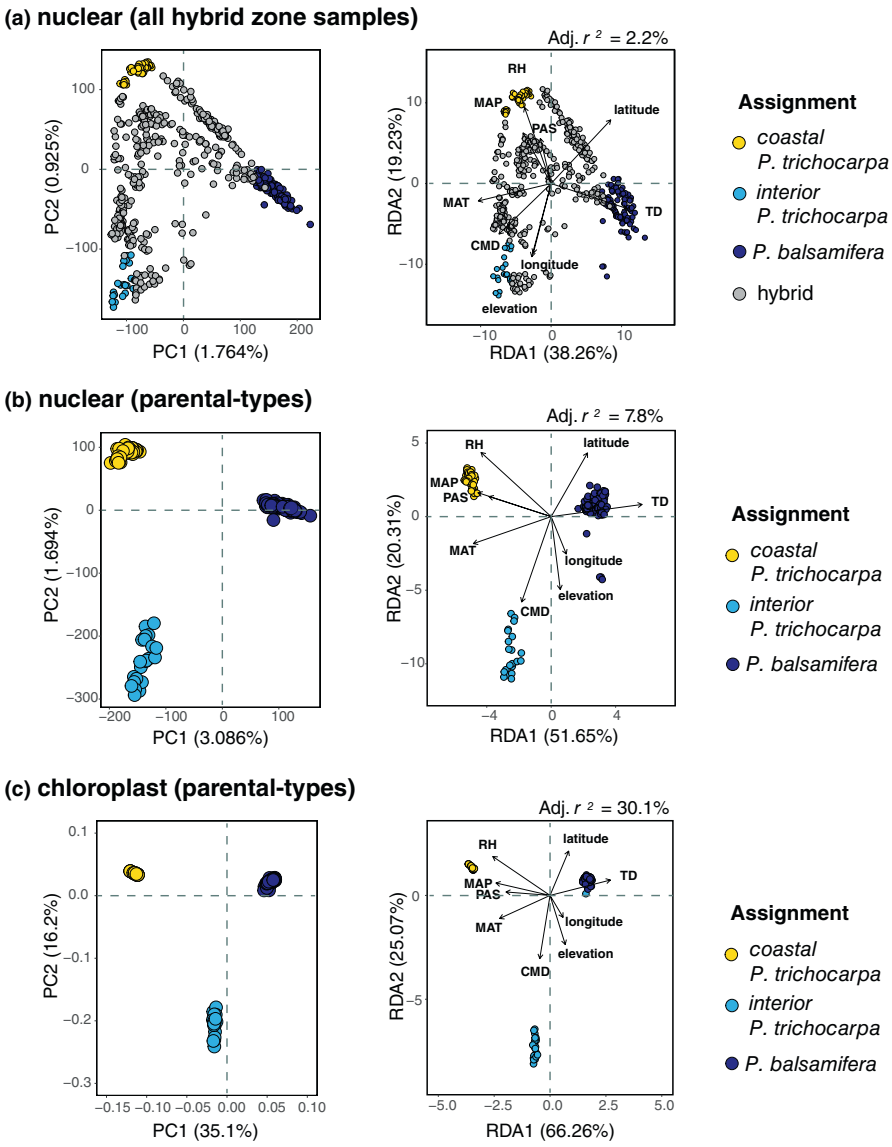


FIGURE 2 Genetic structure (PCA) and redundancy analysis (RDA) based on associations between six climate, three geographical variables for (a) genome-wide nuclear data of all 546 samples analysed, versus (b) genome-wide nuclear and (c) chloroplast differentiation across parental-types. In each RDA plot, the direction and length of each arrow corresponds with the strength of prediction (i.e. loading) along each RDA axis. Climate and geographical variables include MAT (mean annual temperature), TD (continentality), PAS (precipitation as snow), RH (relative humidity), MAP (mean annual precipitation), CMD (climate moisture deficit), latitude, longitude and elevation. Patterns between chloroplast and nuclear DNA are congruent with the exception of one sample (ID = 757_S202_L004) which has *interior P. trichocarpa* nuclear genetic background with *P. balsamifera* chloroplast identity.

balsamifera were associated with higher TD and lower MAT on average. The second RDA axis captured 19.23% (PVE) of the explainable genetic variance, with the highest predictor loadings on the second RDA axis being RH, elevation and longitude. These variables were associated with differentiation within *P. trichocarpa* across the seven transects. Associations between climate, geography and genetic variation were not observed along RDA2 for *P. balsamifera* (Figure 2). Combining loadings for RDA1 and RDA2, we found latitude, TD and CMD had the strongest associations with genetic variation within and among sampled transects.

Because geographical and climatic predictor variables are often correlated, we performed variance partitioning to tease apart the respective contributions of climate, geography, and their interactions to genomic variance. We observed a large, confounded effect (58.69%) between the climatic and geographical variables chosen for analysis. Removing confounded effects through conditioning revealed climate was a stronger predictor of genetic variance. Climate, conditioned on geography, accounted for 27.28% of explainable genetic variance, whereas geography, conditioned on climate, accounted for just 14.03% of explainable genetic variance (Supplementary.S1: Table S3).

Using the same suite of analyses to investigate associations between environment and ancestry for the 185 parental-types revealed higher contributions of climate and geography to explainable genetic structure. From side-by-side comparisons of nuclear and chloroplast genetic structure in the PCA, we observed similar clustering (Figure 2). Considering the differentiation along PC1 axis, *coastal P. trichocarpa* and *interior P. trichocarpa* appear less diverged based on nuclear data when compared to chloroplast. Similar clustering and environmental associations were also observed in the RDA (Figure 2) and aligned with the associations presented for the 546 samples (Figure 2a). However, with hybrids removed from the analysis, the explanatory variance increased (adj. $r^2 = 7.8\%$; PVE = 51.65%). The explanatory variance based on

chloroplast structure was the highest observed (adj. $r^2 = 30.1\%$; PVE = 66.26%). Variance partitioning was also conducted with the hybrids excluded, and with the confounding effects of geography removed, climate explained approximately 40–42% of the genetic variance in both the chloroplast and nuclear data. Conversely, after conditioning on climate, geography explained only 9.46% of the nuclear genetic variance and 2.87% of the chloroplast genetic variance (Supplementary S1: Table S3).

Partial Mantel tests provided further support that genetic differentiation is driven by climate more than geography. Evaluations of correlation across distance metrics for all 546 samples indicated greater genetic isolation-by-environment (IBE = 0.37) than isolation-by-distance (IBD = 0.16). However, both IBE and IBD were statistically significant (p-value = 0.001). With hybrids excluded, IBE increased by nearly 50% (IBE = 0.70), and IBD was no longer significant (*p*-value = 1; IBD = −0.10). Evaluations of IBD and IBE using the genetic distance based on chloroplast indicated environment was the major contributing factor to the parental-type chloroplast genome ancestry (IBE = 0.66 vs. IBD = −0.13; Supplementary S1: Table S4).

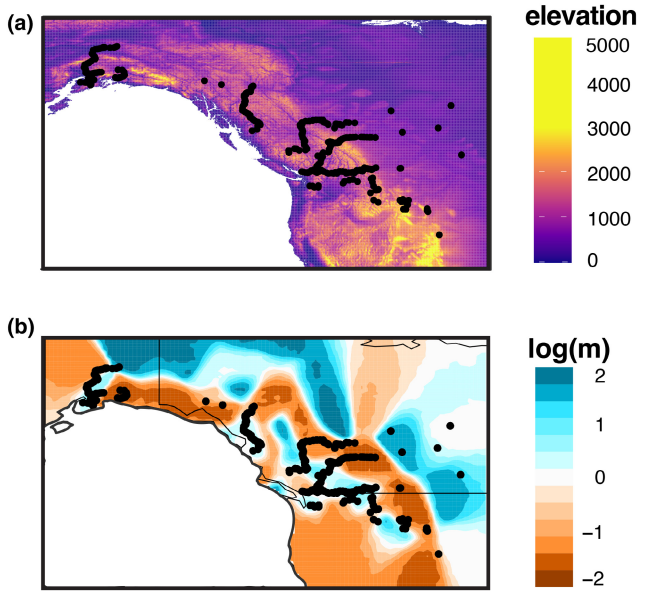


FIGURE 3 Estimated effective migration surfaces (EEMS) analysis of the relationship between genetic distance and geographical distance for 546 sampled trees across seven transects of the *P. trichocarpa* and *P. balsamifera* hybrid zone to visualise potential barriers to gene flow. Panel (a) is a raster map of elevational differences (in metres) across the studied region and panel (b) illustrates putative barriers to gene flow within and among sampled transects. The black dots represent the geographical location of each of the 546 samples. Regions with negative migration rate predictions (brown colours on the log 10 scale) indicate lower genetic connectivity (gene flow) than expected. Regions with positive migration rate predictions (blue colours on the log 10 scale) indicate higher genetic connectivity than expected.

3.4 | Gene flow dynamics shaped by physical barriers

Using departures from IBD to indicate barriers to gene flow, EEMS analysis identified the continental divide of the Rocky Mountains as a large barrier between *P. trichocarpa* and *P. balsamifera*. Additional barriers to gene flow were identified between transects associated with elevational gradients (Figure 3). For example, *interior P. trichocarpa* within the most southern transect was observed in a region of high local connectivity but was isolated from neighbouring individuals identified as *coastal P. trichocarpa* and *P. balsamifera* in other analyses. The EEMS results indicated both Alaska transects were more isolated from the other northern transect, Cassiar, than expected based on IBD. In contrast, individuals spanning *coastal P. trichocarpa* parental-type and hybrids across the Chilcotin, Jasper and Crowsnest transects were more similar than expected under IBD, suggesting substantial intraspecific migration across latitudes within the southern range of this hybrid zone. This pattern was not observed for *P. balsamifera*, as only individuals located east of the Rocky Mountains exhibited evidence of greater effective migration than expected.

3.5 | Demographic inference reveals periods of interspecific vicariance, gene flow and the emergence of a hybrid lineage

Of the four demographic models of connectivity and divergence evaluated (Table 1b), the model with the greatest support inferred an initial divergence between *coastal P. trichocarpa* – *P. balsamifera* approximately 2.12 million years ago (95% CI: 2.11–2.13 million years ago). Following initial divergence, limited gene flow (*mA* ~ 1.90e-04) continued until approximately 860,000 years ago when gene flow ceased, during which time the two lineages were isolated (Figure 4). Following approximately 96,000 years of isolation, the two lineages experienced secondary contact and 764,000 years ago formed the *interior P. trichocarpa* lineage via admixture. During the initial formation of *interior P. trichocarpa* the proportions of each progenitor lineage were inferred to be approximately 34.8% *coastal P. trichocarpa* (parameter *f*; Table 2) and 65.2% *P. balsamifera* (1 − *f*). Following secondary contact, asymmetrical gene flow was inferred until present, with the highest rates of gene flow observed between *coastal P. trichocarpa* and *interior P. trichocarpa* (*m13* ~ 5.08e-05 and *m31* ~ 4.58e-05; Table 2), with a slightly greater rate of genetic exchange from *interior* into *coastal* (*m13*). Gene flow appeared to be strongly asymmetrical, with more movement from *P. balsamifera* into *coastal* and *interior P. trichocarpa* (*m12* ~ 3.64e-05, *m32* ~ 3.13e-05, respectively) and less *coastal* or *interior P. trichocarpa* movement into *P. balsamifera* (*m21* ~ 1.30e-05 and *m23* ~ 9.08e-06; Table 2). The highest effective population size inferred was for *P. balsamifera* (N_2 ~ 100,000) followed by *interior P. trichocarpa* (N_3 ~ 53,800) and *coastal P. trichocarpa* (N_1 ~ 36,500). The residuals between the model and data fit were small and evenly distributed, suggesting good

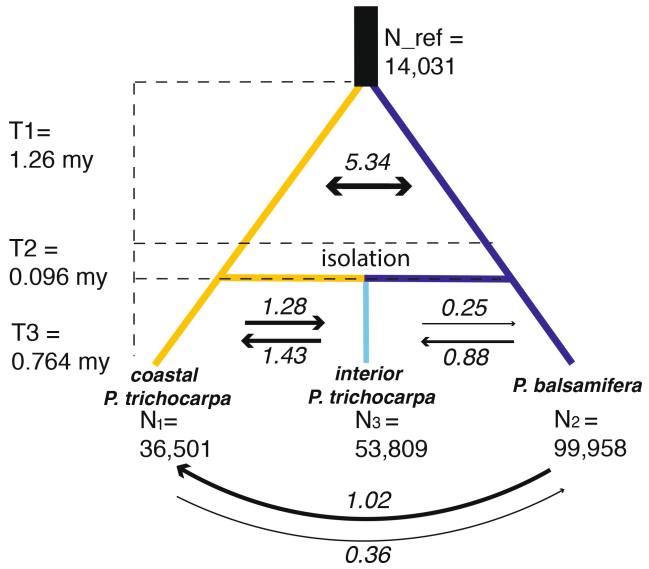


FIGURE 4 The best fit model of divergence among *coastal P. trichocarpa* (in yellow), *P. balsamifera* (in dark blue) and *interior P. trichocarpa* (in light blue). N_ref is the effective population size estimated for the ancestral group prior to *coastal P. trichocarpa* and *P. balsamifera* divergence. N_1 – N_3 report contemporary effective population sizes for each of the three diverged lineages. Migration rates (M) in number of migrants per generation are reported contextually above or below arrows. Line and arrow weights correspond with M- the larger the value of M, the thicker the arrowed line. Parameter estimates and confidence intervals for the three-population model are provided in Table 2. Note that T1, T2 and T3 are time intervals in millions of years (my).

model fit (Supplementary S1: Figure S10). In addition, CIs for each parameter were tightly bound to the estimations inferred (Table 2).

4 | DISCUSSION

This study provides valuable insights into how complex spatial and temporal dynamics influence hybridisation, highlighting the role neutral processes can play in shaping the genetic structure of hybrid zones. A dynamic history between glacial and interglacial cycles suggests that secondary contact over time has had a substantial influence on hybrid zone structure. We identified three distinct genetic clusters geographically distributed across our seven repeated zones of species overlap. Differentiation within and across nuclear and chloroplast genomes supported the presence of three genetic lineages, with tests of topology suggesting one genetic lineage reflects an ancient admixture event between *coastal P. trichocarpa* and *P. balsamifera* leading to the origin of a new hybrid lineage, *interior P. trichocarpa*. While physical geographical barriers have played a significant role influencing historical and contemporary asymmetrical genetic exchange, our results highlight the important role climate gradients alongside geography have played influencing hybrid zone structure. Together, this suggests that genetic patterns observed could be sensitive to climatic shifts, which has implications for understanding the future of these contact zones in a changing climate.

4.1 | Demographic shifts associate with shifts in historical climate

According to our best fit 3D demographic model, initial divergence between *P. balsamifera and coastal P. trichocarpa* (~2.12 million years ago) aligns with a shift from the Pliocene to the Pleistocene Climate Era (Tan et al., 2018). This initial divergence was inferred to occur amid recurrent gene flow for approximately 1.25 million years, followed by a period of isolation lasting ~96,000 years. Interestingly, this period of isolation not only aligns with the Mid-Pleistocene Transition, when glacial cycles became longer (from ~40,000 to ~100,000 years long) and shifts to interglacial climates became abrupt and extreme, but it also coincides in timing with a glacial maximum (~800,000 years ago; Chalk et al., 2017; Ellis & Palmer, 2016; Willeit et al., 2019). Secondary contact between *P. balsamifera* and *coastal P. trichocarpa*, along with the origin of *interior P. trichocarpa*, was estimated to have occurred ~764,000 years ago and coincides with a transition to interglacial climate (Ellis & Palmer, 2016). This suggests opportunities for genetic exchange were regained following expansion from isolated refugia.

While climate has continued to oscillate since the inferred secondary contact event, which might be expected to reduce opportunities for gene flow given the different climatic affinities of each lineage, our demographic model suggested continued genetic exchange. One explanation may be that species distributions have moved in tandem with climate such that populations were proximal enough to maintain gene flow, resulting in a stable hybrid zone. Indeed, the long-distance dispersal of seed and pollen in *Populus* and other forest trees blurs definitions for how geographically isolated species need to be to disrupt gene flow. An alternative explanation may be that periods of strict isolation did occur during extreme climatic shifts but were not sustained long enough for genomic patterns of divergence associated with allopatry to develop, or patterns were broken down following subsequent contact. This latter possibility may characterise hybridising species with porous genomes whose species boundaries vary temporally with repeated cycles of reinforcement and homogenisation (e.g. Mix-Isolation-Mix hypothesis; He et al., 2019) or due to changes in the strength and direction of selection over historically varying climatic optima (e.g. Shang et al., 2020). Among the divergence histories inferred for forest tree species, there is growing evidence for divergence with recurrent gene flow despite abrupt shifts in climate (e.g. Bolte et al., 2022; He et al., 2019; Lexer et al., 2006; Menon et al., 2018).

Varying mutation rate and generation time can impact the unscaling of demographic parameter estimates into meaningful time and population size units. A generation time of 15 years was chosen based on previous *Populus* studies (e.g. Hofmeister et al., 2020; Ingvarsson, 2008) and are within the predicted range for both species (i.e., ~10 years old at first reproduction; Burns & Honkala, 1990).

TABLE 2 Parameter estimates and 95% confidence intervals (CI) associated with the best fit three-population demographic model (Topology 3 – T3_IM_SC_IM). Calculations used to unscale parameter estimates (using theta, mutation rate per generation and genome length) into meaningful units are provided in Supplementary S3.

Parameter	∂α∂i scaled	∂α∂i unscaled	95% CI (lower)	95% CI (upper)
N_ref (ancestral, theta)	4938	14,031	13,650	14,413
N_1 (*coastal trichocarpa*)	2.60	36,501	35,851	37,150
N_2 (*balsamifera*)	7.12	99,958	96,125	103,792
N_3 (*interior trichocarpa*)	3.83	53,809	51,106	56,512
T1 (ancient gene flow)	2.99	1,259,951	1,254,996	1,264,907
T2 (no gene flow)	0.23	96,645	93,272	100,018
T3 (admixture)	1.82	764,327	763,623	765,031
$M_A \rightarrow mA$[a]	5.34	1.90E-04	1.58E-04	2.22E-04
$M_{12} \rightarrow m12$[a]	1.02	3.64E-05	3.56E-05	3.71E-05
$M_{21} \rightarrow m21$[a]	0.36	1.30E-05	1.15E-05	1.45E-05
$M_{23} \rightarrow m23$[a]	0.25	9.08E-06	7.48E-06	1.07E-05
$M_{32} \rightarrow m32$[a]	0.88	3.13E-05	2.89E-05	3.37E-05
$M_{13} \rightarrow m13$[a]	1.43	5.08E-05	5.01E-05	5.14E-05
$M_{31} \rightarrow m31$[a]	1.28	4.58E-05	4.28E-05	4.88E-05
f	0.348	34.8%	0.346	0.350

Note: Each value for T (time) is an interval. Interpretation of each migration (gene flow) parameter is provided in the footnote of this table. The parameter *f* is an estimate of the proportion of *coastal P. trichocarpa* in the formation of *interior P. trichocarpa* at the start of T3. The proportion of *P. balsamifera* is thus 1 − *f*.

[a]M is a scaled estimate of the number of migrants per generation. *m* is an estimate of gene flow rate. After M (or *m*) the notation of A is ancestral migration, 12 is migrations from *balsamifera* into *coastal trichocarpa*, 21 is migration from *coastal trichocarpa* into *balsamifera*, 23 is migration from *interior trichocarpa* into *balsamifera*, 32 is migration from *balsamifera* into *interior trichocarpa*, 13 is migrations from *interior trichocarpa* into *coastal trichocarpa* and 31 is migration from *coastal trichocarpa* into *interior trichocarpa*.

Compared to the effects of generation time, varying the mutation rate is more impactful to unscaled parameter estimates. However, we consider the mutation rate inferred by Hofmeister et al. (2020) for *P. trichocarpa* to be reliable given its similarity to other forest tree genera (e.g. Quercus – Schmid-Siegert et al., 2017; Pinaceae – De La Torre et al., 2017).

4.2 | Gene flow asymmetries in the context of future climate change

Gene flow was nearly three times higher from *P. balsamifera* into *coastal P. trichocarpa* than from *coastal P. trichocarpa* into *P. balsamifera*. This observation is corroborated by previous estimates of genetic exchange in this hybrid zone by Suarez-Gonzalez, Hefer, et al. (2018) and Suarez-Gonzalez, Lexer, & Cronk (2018), who noted asymmetrical introgression from *P. balsamifera* into *P. trichocarpa*. This asymmetry may suggest *P. balsamifera* has a demographic advantage within the ecotonal transition (e.g. dispersal down an elevational gradient or with river flow from inland to coast). On the other hand, lower rates of introgression between *P. balsamifera* and *interior P. trichocarpa* may be indicative of stronger geographical isolation, a lack of hybrids geographically positioned between these genetic clusters to facilitate backcrossing, or even more likely, asynchronous reproductive phenology. According to descriptions within (Haeussler et al., 1990) flowering times for interior populations of *P. trichocarpa* can be nearly 1 month later than those reported for populations of *P. trichocarpa* (coastal) and *P. balsamifera* in British Columbia, Canada.

Gene flow between *coastal P. trichocarpa* and *interior P. trichocarpa* were greatest relative to all inferred between-lineage contemporary estimates (Figure 4). Indeed, *interior P. trichocarpa* is genetically more similar to *coastal P. trichocarpa* than *P. balsamifera* based on patterns of nuclear differentiation. However, despite this, the demographic model that received the greatest support suggested that in the formation of *interior P. trichocarpa* (~764,000 years ago) only 35% of the hybrid genome reflected ancestry from *coastal P. trichocarpa*. Taken together, these results may indicate that higher rates of contemporary gene flow between *coastal* and *interior P. trichocarpa* homogenised population differences following the initial divergence of the *interior P. trichocarpa* lineage. Regardless, the influx of new genetic recombinants through introgression is evident across all three parental-types and may impact adaptation to a changing climate (Aitken et al., 2008; Barton & Bengtsson, 1986; Martinsen et al., 2001).

4.3 | Climate and landscape heterogeneity contribute to genetic structure

Landscape-induced restrictions to gene flow have contributed to patterns of differentiation across the hybrid zone. EEMS suggest

gene flow has likely been suppressed by elevational barriers through the Rocky Mountains of northwestern North America. West of the Rocky Mountains, we observed variation in effective migration rates within the *P. trichocarpa* distribution described by (Little Jr., 1971). For example, effective migration was greater within the northern transects of Alaska and Cassiar but not between these two transects. Similarly, there appeared to be substantial gene flow between hybrids within the 'mid-latitudinal' transects (Chilcotin, Jasper and Crowsnest), but migration was restricted along the coastal regions of these three transects where parental-types of *coastal P. trichocarpa* were identified. Based on expectations of IBD, this suggests gene flow between *P. trichocarpa* populations may be lower than expected. Within this region, landscape barriers to gene flow likely contribute to strong population structure (Geraldes et al., 2014; Slavov et al., 2012) and lower levels of heterozygosity for *coastal P. trichocarpa* parental-types. Estimates of heterozygosity for *coastal P. trichocarpa* (H_o = 0.093) and *interior P. trichocarpa* (H_o = 0.101) are in line with previous reports by Slavov et al. (2010) for coastal and interior populations. Interestingly, greater heterozygosity in *P. balsamifera* (H_o = 0.106) suggest *P. balsamifera* populations may have maintained substantial genetic variation following post glacial re-colonisation (Keller et al., 2010). This may result from fewer landscape barriers to gene flow east of the Rocky Mountains. In contrast, *P. trichocarpa* populations are within the Cascade, Rocky, and Canadian Coast Mountain terrains, where dynamic expansions and contractions act as strong selective sieves to reduce genetic variation (Hamilton & Aitken, 2013).

RDA, variance partitioning and Mantel tests for isolation-by-environment revealed genetic structure and climate are significantly correlated across the hybrid zone. The Rocky Mountains present environmental transitions from drier and colder conditions on the eastern side (favouring *P. balsamifera*) to more maritime climates on the western side of the Rockies (favouring *coastal P. trichocarpa*). The distributions and climate associations we observed between *P. balsamifera*, *coastal P. trichocarpa*, and *interior P. trichocarpa* (Figure 2a) are comparable to other hybrid zones within the region, including *Picea glauca*, *P. sitchensis*, and *P. engelmannii* hybrid zones (A. De La Torre et al., 2015; Hamilton et al., 2013b, 2015). In both systems, maritime to continental climate transitions were associated with genomic ancestry. Because the frequency of hybrid genotypes is largely associated with ecotonal transitions, this suggests these hybrid zones likely reflect a model of bounded hybrid superiority. However, quantifying fitness in reciprocal transplant experiments is needed to test the bounded hybrid superiority model.

The admixed lineage, *interior P. trichocarpa*, appears to be bounded to a drier and warmer climate than either of its progenitors. The genetic structure of *interior P. trichocarpa* is associated with high CMD. If selection maintains genetic structure among the lineages observed with our study, we may find trait divergence, especially those underlying water-use, across genetic lineages. Interestingly, *Populus* hybrid lineages of the Qinghai-Tibet Plateau also inhabit landscapes with higher aridity relative to their progenitors. However, survival and growth measurements within common garden experiments suggested phenotypic plasticity may have contributed to adaptive divergence and persistence in environments beyond the parental species' range (Shi et al., 2024). A similar experimental design could be employed for *coastal P. trichocarpa*, *interior P. trichocarpa*, and *P. balsamifera* to test if the admixed lineage (*interior P. trichocarpa*) exhibits transgressive traits relative to parents while also evaluating the role of plasticity to success in novel environments. In the absence of trait data, we can leverage genomic data to differentiate between, and functionally assess, genomic regions associated with restricted or excess introgression that may underlay adaptive divergence.

4.4 | Nuclear and chloroplast genomes provide complementary insights into hybrid zone dynamics

The estimates observed herein suggest 2.11–2.13 million years ago *P. trichocarpa* and *P. balsamifera* diverged and are congruent with observations for this species-pair from (Sanderson et al., 2023). The differentiation we observed in the chloroplast not only aligns with patterns of nuclear differentiation among the three assigned parental-types but also supports the observations and interpretations of previous chloroplast assessments of divergence in *Populus*. Huang et al. (2014) ascribed the deep chloroplast divergence of *P. trichocarpa* from *P. balsamifera* to chloroplast capture of an ancient, now extinct, lineage. While the genome-wide patterns of nuclear differentiation (e.g. PCA, F_{ST}) suggests *coastal P. trichocarpa* and *interior P. trichocarpa* are more closely related, our phylogenetic inference from the chloroplast placed *coastal P. trichocarpa* outside the clade of *interior P. trichocarpa* and *P. balsamifera*. Only one individual out of 185 did not have a matching chloroplast – nuclear parental-type assignment. Considering our demographic modelling and the fact that our focal taxa exhibit dioecy with maternal inheritance of the chloroplast genome, the origin of *interior P. trichocarpa* likely resulted from admixture between *P. balsamifera* seed parents and *coastal P. trichocarpa* pollen parents. Given the inferred length of time that has passed since the admixed origin (~764,000 years), the *interior P. trichocarpa* chloroplast genome has likely evolved following divergence. Co-evolution and co-adaptation between the chloroplast and nuclear genomes may have contributed to chloroplast divergence of *interior P. trichocarpa* from the original seed parent (Greiner & Bock, 2013) and would be an interesting future avenue for research.

4.5 | Conclusions

Here, we provide insight into the evolutionary processes shaping divergence and hybridisation across space and time in a *Populus* system of western North America. Demographic models support the origin of a hybrid lineage following secondary contact of ancestral lineages and a history of asymmetrical gene flow between focal taxa. Our results highlight the influence of climate and geography to genetic structure and divergence, as well as the influence of landscape-level

attributes to gene flow dynamics. Overall, our study sheds light on the role demographic processes can play within a long-lived hybrid zone across space and time and highlight the need for fine-scale assessment of regions of the genome important to introgression and adaptation for species management in a changing climate.

AUTHOR CONTRIBUTIONS

MCF, JH, SRK and JAH conceived the study and collected the samples. MC extracted DNA. TP processed genomic data. CEB analysed genomic data, conducted demographic inference and performed genetic-environment associations. MZ and BS assembled chloroplast genomes and performed chloroplast analyses. SRK, JH and JAH advised analyses. CEB and JAH led the writing of the manuscript. SRK, JH and MCF provided valuable feedback.

ACKNOWLEDGEMENTS

We thank Joe Braasch, Lionel Di Santo, Sonia DeYoung, Sara Klopf, Jack Woods, Ben Woods and Raju Soolanayakanahally for assisting with field sample collection. We also thank Kyle Peer, Clay Sawyers and Deborah Bird from the Virginia Tech Reynold's Homestead Forestry Research Station for their assistance with plant propagation. We thank Ryan Gutenkunst and Andrew Eckert for advice and communications related to demographic inference, and Christian Rellstab and the anonymous reviewer for their advice during manuscript revisions. We acknowledge the valuable questions from and conversations with our team members during project calls and meetings: Cigdem Kansu, Susanne Lachmuth, Baxter Worthing and Alayna Mead.

FUNDING INFORMATION

National Science Foundation grant PGR-1856450, USDA National Institute of Food and Agriculture project and Hatch Appropriations (PEN04809 and Accession 7003639), NIFA project VA-136641, Schatz Center for Tree Molecular Genetics.

CONFLICT OF INTEREST STATEMENT

All authors claim no competing interests.

DATA AVAILABILITY STATEMENT

Sequence reads for the 546 samples analysed in this manuscript have been archived on NCBI (PRJNA996882). Chloroplast genomes are archived under BankIt2793276 (Accession numbers PP274173–PP274682). Scripts for variance partitioning, RDA, and demographic inference can be found at https://doi.org/10.5281/zenodo.11490191.

ORCID

Constance E. Bolte https://orcid.org/0000-0003-0438-7102
Tommy Phannareth https://orcid.org/0009-0007-1042-6098
Michelle Zavala-Paez https://orcid.org/0009-0009-7939-5351
Brianna N. Sutara https://orcid.org/0009-0006-9690-1576
Muhammed F. Can https://orcid.org/0009-0000-3649-0863
Matthew C. Fitzpatrick https://orcid.org/0000-0003-1911-8407
Jason A. Holliday https://orcid.org/0000-0002-2662-8790
Stephen R. Keller https://orcid.org/0000-0001-8887-9213
Jill A. Hamilton https://orcid.org/0000-0002-3041-0508

REFERENCES

Abbott, R. J. (2017). Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics and Evolution*, *55*(4), 238–258. https://doi.org/10.1111/JSE.12267

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

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–332.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Bacilieri, R., Ducousso, A., & Kremer, A. (1996). Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. *Annales des Sciences Forestières*, *53*(1), 79–91. https://doi.org/10.1051/forest:19960106

Barton, N., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, *57*(3), 357–376. https://doi.org/10.1038/hdy.1986.135

Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, *341*(6242), 497–503. https://doi.org/10.1038/341497a0

Bolte, C. E., Faske, T. M., Friedline, C. J., & Eckert, A. J. (2022). Divergence amid recurring gene flow: Complex demographic histories for two north American pine species (Pinus pungens and P. Rigida) fit growing expectations among forest trees. *Tree Genetics & Genomes*, *18*(5), 35. https://doi.org/10.1007/s11295-022-01565-8

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer.

Burns, R., & Honkala, B. (1990). *Silvics of North America* (Vol. *2*). United States Department of Agriculture.

Cannon, C. H., & Petit, R. J. (2020). The oak syngameon: More than the sum of its parts. *New Phytologist*, *226*(4), 978–983. https://doi.org/10.1111/nph.16091

Chalk, T. B., Hain, M. P., Foster, G. L., Rohling, E. J., Sexton, P. F., Badger, M. P. S., Cherry, S. G., Hasenfratz, A. P., Haug, G. H., Jaccard, S. L., Martínez-García, A., Pälike, H., Pancost, R. D., & Wilson, P. A. (2017). Causes of ice age intensification across the mid-Pleistocene transition. *Proceedings of the National Academy of Sciences*, *114*(50), 13114–13119. https://doi.org/10.1073/pnas.1702143114

Chhatre, V. E., Evans, L. M., DiFazio, S. P., & Keller, S. R. (2018). Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of Populus. *Molecular Ecology*, *27*(23), 4820–4838. https://doi.org/10.1111/mec.14820

Christe, C., Stölting, K. N., Paris, M., Fraïsse, C., Bierne, N., & Lexer, C. (2017). Adaptive evolution and segregating load contribute to the genomic landscape of divergence in two tree species connected by episodic gene flow. *Molecular Ecology*, *26*(1), 59–76. https://doi.org/10.1111/mec.13765

De La Torre, A., Ingvarsson, P. K., & Aitken, S. N. (2015). Genetic architecture and genomic patterns of gene flow between hybridizing species of picea. *Heredity*, *115*(2), 153–164. https://doi.org/10.1038/hdy.2015.19

De La Torre, A. R., Li, Z., Van De Peer, Y., & Ingvarsson, P. K. (2017). Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution*, *34*(6), 1363–1377. https://doi.org/10.1093/molbev/msx069

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: *De novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *gkw955*, e18. https://doi.org/10.1093/nar/gkw955

Dobzhansky, T. H. (1936). Studies on hybrid sterility. 11. localization of sterility factors i n drosophila pseudoobscura hybrids. *Genetics*, *21*(113), 113–135.

Eaton, D. A. R., Hipp, A. L., González-Rodríguez, A., & Cavender-Bares, J. (2015). Historical introgression among THE American live oaks and THE comparative nature of tests for introgression: Introgression in the american live oaks. *Evolution*, *69*(10), 2587–2601. https://doi.org/10.1111/evo.12758

El Mujtar, V., Sola, G., Aparicio, A., & Gallo, L. (2017). Pattern of natural introgression in a Nothofagus hybrid zone from south American temperate forests. *Tree Genetics & Genomes*, *13*(2), 49. https://doi.org/10.1007/s11295-017-1132-1

Ellis, R., & Palmer, M. (2016). Modulation of ice ages via precession and dust-albedo feedbacks. *Geoscience Frontiers*, *7*(6), 891–909. https://doi.org/10.1016/j.gsf.2016.04.004

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, *143*(9), 1919–1927. https://doi.org/10.1016/j.biocon.2010.05.011

Geraldes, A., Farzaneh, N., Grassa, C. J., Mckown, A. D., Guy, R. D., Mansfield, S. D., Douglas, C. J., & 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*(11), 3260–3280. https://doi.org/10.1111/evo.12497

Greiner, S., & Bock, R. (2013). Tuning a ménage à trois: Co-evolution and co-adaptation of nuclear and organellar genomes in plants. *BioEssays*, *35*(4), 354–365. https://doi.org/10.1002/bies.201200137

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*(10), e1000695. https://doi.org/10.1371/journal.pgen.1000695

Haeussler, S., Coates, D., & Mather, J. (1990). *Autecology of common plants in British Columbia: A literature review*. British Columbia Ministry of Forests, Research Branch.

Hamilton, J. A., & Aitken, S. N. (2013). Genetic and morphological structure of a spruce hybrid (*Picea sitchensis* × *P. Glauca*) zone along a climatic gradient. *American Journal of Botany*, *100*(8), 1651–1662. https://doi.org/10.3732/ajb.1200654

Hamilton, J. A., De la Torre, A. R., & Aitken, S. N. (2015). Fine-scale environmental variation contributes to introgression in a three-species spruce hybrid complex. *Tree Genetics & Genomes*, *11*(1), 1–14. https://doi.org/10.1007/s11295-014-0817-y

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013a). Differential introgression reveals candidate genes for selection across a spruce (*Picea sitchensis* × *P. glauca*) hybrid zone. *New Phytologist*, *197*(3), 927–938. https://doi.org/10.1111/nph.12055

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013b). Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). *Molecular Ecology*, *22*(3), 827–841. https://doi.org/10.1111/mec.12007

Hamilton, J. A., & Miller, J. M. (2016). Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), 33–41. https://doi.org/10.1111/cobi.12574

He, Z., Li, X., Yang, M., Wang, X., Zhong, C., Duke, N. C., Wu, C. I., & Shi, S. (2019). Speciation with gene flow via cycles of isolation and migration: Insights from multiple mangrove taxa. *National Science Review*, *6*(2), 275–288. https://doi.org/10.1093/nsr/nwy078

Heiser, C. (1949). Natural hybridization with particular reference to introgression. *The Botanical Review*, *15*, 645–687.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the quaternary. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *359*(1442), 183–195. https://doi.org/10.1098/rstb.2003.1388

Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., Jordan, R., Oakeshott, J., Weeks, A., Joseph, L., Lockhart, P., Borevitz, J., & Sgrò, C. (2015). A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses*, *2*(1), 1–24. https://doi.org/10.1186/s40665-014-0009-x

Hofmeister, B. T., Denkena, J., Colomé-Tatché, M., Shahryary, Y., Hazarika, R., Grimwood, J., Mamidi, S., Jenkins, J., Grabowski, P. P., Sreedasyam, A., Shu, S., Barry, K., Lail, K., Adam, C., Lipzen, A., Sorek, R., Kudrna, D., Talag, J., Wing, R., … Schmitz, R. J. (2020). A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial *Populus trichocarpa*. *Genome Biology*, *21*(1), 259. https://doi.org/10.1186/s13059-020-02162-5

Huang, D. I., Hefer, C. A., Kolosova, N., Douglas, C. J., & Cronk, Q. C. B. (2014). Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytologist*, *204*(3), 693–703. https://doi.org/10.1111/nph.12956

Huang, X., Struck, T. J., Davey, S. W., & Gutenkunst, R. N. (2023). dadi-cli: Automated and distributed population genetic model inference from allele frequency spectra [preprint]. *bioRxiv*. https://doi.org/10.1101/2023.06.15.545182

Ingvarsson, P. K. (2008). Multilocus patterns of nucleotide polymorphism and the demographic history of Populus tremula. *Genetics*, *180*(1), 329–340. https://doi.org/10.1534/genetics.108.090431

Janes, J. K., & Hamilton, J. A. (2017). Mixing it up: The role of hybridization in forest management and conservation under climate change. *Forests*, *8*(7), 237. https://doi.org/10.3390/f8070237

Jansson, S., & Douglas, C. J. (2007). Populus: A model system for plant biology. *Annual Review of Plant Biology*, *58*(1), 435–458. https://doi.org/10.1146/annurev.arplant.58.032806.103956

Jeffrey, E. C. (1916). Hybridism and the rate of evolution in angiosperms. *The American Naturalist*, *50*, 129–143.

Jump, A. S., & Peñuelas, J. (2005). Running to stand still: Adaptation and the response of plants to rapid climate change. *Ecology Letters*, *8*(9), 1010–1020. https://doi.org/10.1111/j.1461-0248.2005.00796.x

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649.

Keller, S. R., Olson, M. S., Salim, S., William, S. A., & Peter, T. (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

Legendre, P., & Legendre, L. (2012). *Numerical ecology*. Elsevier.

Lepais, O., Petit, R. J., Guichoux, E., Lavabre, J. E., Alberto, F., Kremer, A., & Gerber, S. (2009). Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, *18*(10), 2228–2242. https://doi.org/10.1111/j.1365-294X.2009.04137.x

Leroy, T., Louvet, J., Lalanne, C., Le Provost, G., Labadie, K., Aury, J., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks.

New Phytologist, *226*(4), 1171–1182. https://doi.org/10.1111/nph.16095

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

Lexer, C., Fay, M. F., Joseph, J. A., Nica, M. S., & Heinze, B. (2005). Barrier to gene flow between two ecologically divergent Populus species, *P. alba* (white poplar) and *P. tremula* (European aspen): The role of ecology and life history in gene introgression. *Molecular Ecology*, *14*(4), 1045–1057. https://doi.org/10.1111/j.1365-294X.2005.02469.x

Lexer, C., Kremer, A., & Petit, R. J. (2006). Shared alleles in sympatric oaks: Recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology*, *15*(7), 2007–2012. https://doi.org/10.1111/j.1365-294X.2006.02896.x

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics*, *26*(5), 589–595. https://doi.org/10.1093/bioinformatics/btp698

Little, E. L., Jr. (1971). *Atlas of United States trees. Vol. 1. Conifers and important hardwoods*. USDA Misc. Publ. 1146 (200 maps).

Luqman, H., Wegmann, D., Fior, S., & Widmer, A. (2023). Climate-induced range shifts drive adaptive response via spatio-temporal sieving of alleles. *Nature Communications*, *14*(1), 1080. https://doi.org/10.1038/s41467-023-36631-9

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, *20*(5), 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Martinsen, G. D., Whitham, T. G., Turek, R. J., & Keim, P. (2001). Hybrid populations selectively filter gene introgression between species. *Evolution*, *55*(7), 1325–1335.

McVay, J. D., Hipp, A. L., & Manos, P. S. (2017). A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1854), 20170300. https://doi.org/10.1098/rspb.2017.0300

Menon, M., Bagley, J. C., Friedline, C. J., Whipple, A. V., Schoettle, A. W., Leal-Sàenz, A., Wehenkel, C., Molina-Freaner, F., Flores-Rentería, L., Gonzalez-Elizondo, M. S., Sniezko, R. A., Cushman, S. A., Waring, K. M., & Eckert, A. J. (2018). The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). *Molecular Ecology*, *27*(5), 1245–1260. https://doi.org/10.1111/MEC.14505

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, *37*(2), 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., & Legendre, P. (2020). Package "vegan"—Community ecology package version 2.5-6. Cran. Ism. Ac. Jp.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, *25*(11), 2337–2360. https://doi.org/10.1111/mec.13557

Petit, R. J., Bodénès, C., Ducousso, A., Roussel, G., & Kremer, A. (2004). Hybridization as a mechanism of invasion in oaks. *New Phytologist*, *161*(1), 151–164. https://doi.org/10.1046/j.1469-8137.2003.00944.x

Petit, J. R., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Ven-dramin, G. G. (2005). Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, *14*, 689–701.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. https://doi.org/10.1038/ng.3464

Sanderson, B. J., Ghambir, D., Feng, G., Hu, N., Cronk, Q. C., Percy, M., Molina Freaner, F., Johnson, M. G., Smart, L. B., Yin, T., Ma, T., DiFazio, S. P., & Liu, J. (2023). Phylogenomics reveals patterns of ancient hybridization and differential diversification that contribute to phylogenetic conflict in willows, poplars, and close relatives. *Systematic Biology*, *72*, 1220–1232. https://doi.org/10.1093/sysbio/syad042

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., Cattaneo, P., Schütz, F., Farinelli, L., Pagni, M., Schneider, M., Voumard, J., Jaboyedoff, M., Fankhauser, C., Hardtke, C. S., Keller, L., Pannell, J. R., Reymond, A., Robinson-Rechavi, M., … Reymond, P. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nature Plants*, *3*(12), 926–929. https://doi.org/10.1038/s41477-017-0066-9

Shang, H., Hess, J., Pickup, M., Field, D. L., Ingvarsson, P. K., Liu, J., & Lexer, C. (2020). Evolution of strong reproductive isolation in plants: Broad-scale patterns and lessons from a perennial model group: Evolution of strong barriers in plants. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *375*(1806), 20190544. https://doi.org/10.1098/rstb.2019.0544

Shi, Y.-J., Huang, J.-L., Mi, J.-X., Li, J., Meng, F.-Y., Zhong, Y., He, F., Tian, F.-F., Zhang, F., Chen, L.-H., Yang, H.-B., Hu, H.-L., & Wan, X.-Q. (2024). A model of hybrid speciation process drawn from three new poplar species originating from distant hybridization between sections. *Molecular Phylogenetics and Evolution*, *190*, 107966. https://doi.org/10.1016/j.ympev.2023.107966

Slavov, G. T., Difazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., Lipphardt, M. F., Pennacchio, C. P., Hellsten, U., Pennacchio, L. A., Gunter, L. E., Ranjan, P., Vining, K., Pomraning, K. R., Wilhelm, L. J., Pellegrini, M., Mockler, T. C., Freitag, M., Geraldes, A., … Tuskan, G. A. (2012). Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist*, *196*(3), 713–725. https://doi.org/10.1111/j.1469-8137.2012.04258.x

Slavov, G. T., Leonardi, S., Adams, W. T., Strauss, S. H., & DiFazio, S. P. (2010). Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity*, *105*(4), 348–357. https://doi.org/10.1038/hdy.2010.73

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

Stebbins, G. L. (1959). The role of hybridization in evolution. *Proceedings of the American Philosophical Society*, *103*, 231–251.

Suarez-Gonzalez, A., Hefer, C. A., Christe, C., Corea, O., Lexer, C., Cronk, Q. C. B., & Douglas, C. J. (2016). Genomic and functional approaches reveal a case of adaptive introgression from *Populus balsamifera* (balsam poplar) in *P. Trichocarpa* (black cottonwood). *Molecular Ecology*, *25*(11), 2427–2442. https://doi.org/10.1111/mec.13539

Suarez-Gonzalez, A., Hefer, C. A., Lexer, C., Cronk, Q. C. B., & Douglas, C. J. (2018). Scale and direction of adaptive introgression between black cottonwood (*Populus trichocarpa*) and balsam poplar (*P. Balsamifera*). *Molecular Ecology*, *27*(7), 1667–1680. https://doi.org/10.1111/mec.14561

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, *14*(3), 20170688. https://doi.org/10.1098/rsbl.2017.0688

Swenson, N. G., & Howard, D. J. (2004). Do suture zones exist? *Evolution*, *58*(11), 2391–2397. https://doi.org/10.1111/j.0014-3820.2004.tb00869.x

Tan, N., Ladant, J.-B., Ramstein, G., Dumas, C., Bachem, P., & Jansen, E. (2018). Dynamic Greenland ice sheet driven by pCO2 variations across the Pliocene Pleistocene transition. *Nature Communications*, *9*(1), 4755. https://doi.org/10.1038/s41467-018-07206-w

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, *44*(W1), W232–W235. https://doi.org/10.1093/nar/gkw256

Tuskan, G. A., DiFazio, S., Jansson, S., & Bohlmann, J. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, *313*, 1596–1605.

VanWallendael, A., Lowry, D. B., & Hamilton, J. A. (2022). One hundred years into the study of ecotypes, new advances are being made through large-scale field experiments in perennial plant systems. *Current Opinion in Plant Biology*, *66*, 102152. https://doi.org/10.1016/j.pbi.2021.102152

Wang, T., 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*(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Willeit, M., Ganopolski, A., Calov, V., & Brovkin, V. (2019). Mid-Pleistocene transition in glacial cycles explained by declining CO2 and regolith removal. *Science Advances*, *5*(4), 1–8. https://doi.org/10.1126/sciadv.aav7337

SUPPORTING INFORMATION

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

How to cite this article: Bolte, C. E., Phannareth, T., Zavala-Paez, M., Sutara, B. N., Can, M. F., Fitzpatrick, M. C., Holliday, J. A., Keller, S. R., & Hamilton, J. A. (2024). Genomic insights into hybrid zone formation: The role of climate, landscape, and demography in the emergence of a novel hybrid lineage. *Molecular Ecology*, *00*, e17430. https://doi.org/10.1111/mec.17430