ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

# Recurrent hybridization and gene flow shaped Norway and Siberian spruce evolutionary history over multiple glacial cycles

Qiujie Zhou[1,2] | Piyal Karunarathne[1,2,3] | Lili Andersson-Li[4] | Chen Chen[5] |
Lars Opgenoorth[6,7] | Katrin Heer[8] | Andrea Piotti[9] |
Giovanni Giuseppe Vendramin[9] | Elena Nakvasina[10] | Martin Lascoux[1,2] |
Pascal Milesi[1,2]

[1]Plant Ecology and Evolution, Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

[2]Science for Life Laboratory (SciLifeLab), Uppsala University, Uppsala, Sweden

[3]Institute of Population Genetics, Heinrich-Heine University, Düsseldorf, Universitäts Straße 1, Düsseldorf, Germany

[4]Department of Microbiology, Tumor and Cell Biology, Karolinska L2:02, Solna, Sweden

[5]Plant Pathology Group, Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland

[6]Department of Biology, Plant Ecology and Geobotany, Philipps-Universität Marburg, Marburg, Germany

[7]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

[8]Faculty of Environment and Natural Resources, Eva Mayr-Stihl Professorship for Forest Genetics, Albert-Ludwigs-Universität Freiburg, Freiburg im Breisgau, Germany

[9]Institute of Biosciences and BioResources (IBBR), National Research Council (CNR), Sesto Fiorentino, Italy

[10]Department of Forestry and Forest Management, Northern (Arctic) Federal University Named after M.V. Lomonosov, Arkhangelsk, Russian Federation

**Correspondence**
Pascal Milesi, Plant Ecology and Evolution, Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, Uppsala 75236, Sweden.
Email: pascal.milesi@scilifelab.uu.se

## Abstract

Most tree species underwent cycles of contraction and expansion during the Quaternary. These cycles led to an ancient and complex genetic structure that has since been affected by extensive gene flow and by strong local adaptation. The extent to which hybridization played a role in this multi-layered genetic structure is important to be investigated. To study the effect of hybridization on the joint population genetic structure of two dominant species of the Eurasian boreal forest, *Picea abies* and *P. obovata*, we used targeted resequencing and obtained around 480 K nuclear SNPs and 87 chloroplast SNPs in 542 individuals sampled across most of their distribution ranges. Despite extensive gene flow and a clear pattern of Isolation-by-Distance, distinct genetic clusters emerged, indicating the presence of barriers and corridors to migration. Two cryptic refugia located in the large hybrid zone between the two species played a critical role in shaping their current distributions. The two species repeatedly hybridized during the Pleistocene and the direction of introgression depended on latitude. Our study suggests that hybridization helped both species to overcome main shifts in their distribution ranges during glacial cycles and highlights the importance of considering whole species complex instead of separate entities to retrieve complex demographic histories.

## 1 | INTRODUCTION

Distribution ranges of boreal species went through cycles of contraction and expansion following glacial cycles (Fletcher et al., 2010; Nevado et al., 2018; Shuvaev et al., 2023). During the contraction phases, isolated populations could accumulate genetic divergence while the expansion phases were accompanied by secondary contact(s) and genetic admixture (Nevado et al., 2018; Petit et al., 2003; Wielstra et al., 2021; Zemlak et al., 2008). These oscillations played a major role in shaping the distribution of genetic variation within and between closely related species. Paleoecological and genetic studies have shown that the impact of glacial cycles varied extensively across species and geographical areas. In western Europe, for instance, some species could only survive in the Iberian or Italian peninsulas during glacial periods while others were able to survive as far north as Central Europe (Binney et al., 2009; Willis et al., 2000). The extent of glaciations also varied through space: while Fennoscandia was almost fully glaciated during the Last Glacial Maximum (LGM, ~26,500–19,000 years ago), Siberia was not. Siberia was a cold desert where patches of forests were able to survive (Semerikov et al., 2013 and reference therein). These repeated sequences of reproductive isolation and admixture, and the vast but non-homogeneous population movements that accompanied them at continental scales, have led to the accumulation over time of an ancient, deep-seated and complex genetic structure, as, for instance, observed in seven European tree species (Milesi et al., 2023). This structure was blurred by extensive gene flow, or reinforced by strong local adaptation, both of which are characteristics of most forest tree species (Savolainen et al., 2007). This already multi-layered structure has often been further complexified by intraspecific hybridization with closely related species.

Over the last decades, extensive genome-wide resequencing studies have highlighted the extent of hybridization and introgression between closely related species (Belokon et al., 2022; Buck et al., 2023; Chen et al., 2018; Cullingham et al., 2012; Fu et al., 2022; Keim et al., 1989; Nocchi et al., 2023; Sankararaman et al., 2014; Shuvaev et al., 2023; Thórsson et al., 2001). Secondary contacts, when previously isolated populations are re-united, have been extensively studied, in particular for their role in speciation, for instance through reinforcement of reproductive isolation (e.g. Barton & Hewitt, 1985). When not selected against, species hybridization is expected to increase genetic diversity and can allow for adaptive introgression (Chhatre et al., 2018; Jones et al., 2018; Leroy et al., 2020; Platt et al., 2019; Whitney et al., 2010), directly affect phenotypic traits (e.g. Darwin's finches beak shape, Grant & Grant, 2019) and thereby allow for niche expansion (e.g. Pfennig et al., 2016 and reference therein). In turn, ecological niche expansion

could result in an increase in census population sizes and foster populations' resilience to environmental changes. Thus, to better understand the role of hybridization in species demographic histories, it is necessary to study all hybridizing species simultaneously (Cannon & Petit, 2020). This requires extensive sampling across entire distribution ranges and examples of such large-scale study remain scarce outside model species (e.g. Bruxaux et al., 2024) and in particular in plants (Nieto Feliner et al., 2023).

*Picea abies* [L.] H. Karst (Norway spruce) and *P. obovata* Ledeb. (Siberian spruce) are keystone species of the Eurasian boreal forest, and their joint range extends from the Norwegian Coast (~10°E) in the West to the Sea of Okhotsk in the East (~140°E). Like other boreal tree species, *P. abies* and *P. obovata* experienced cycles of population expansion and contraction over multiple glacial periods. The current joint-distribution of *P. abies* and *P. obovata* shows a clear longitudinal pattern but no clear delineation between their natural range can be drawn; they occupy different but overlapping ecological niches (Karunarathne et al., 2024). Phenotypic traits also vary from one genetic background to the other (e.g. survival, growth, shape of cone-scale, Lagercrantz & Ryman, 1990; Nakvasina et al., 2019; Orlova et al., 2020; Popov, 2010) and *P. abies* can be discriminated from *P. obovata* and from their hybrid forms (Nakvasina et al., 2017, 2019; Popov, 2010, 2013). Yet, despite their great economic and ecological importance, the geographical patterns of genetic variation of the two species remain unclear, in particular in their putative contact zone in the Eastern European Plain. While *P. abies* can be separated into several well-delineated genetic clusters reflecting major geographical domains and population movements during glacial cycles (Chen et al., 2019; Tsuda et al., 2016), no clear population subdivision has been observed so far within *P. obovata* using nuclear markers, despite its much wider range (Tollefsrud et al., 2008, 2015; Tsuda et al., 2016). Current literature suggests that the last time the two species came into contact was after the LGM with a strong East to West recolonization by *P. obovata*. In the northern part of the range, this resulted in a pattern of introgression from local populations of *P. abies* into incoming *P. obovata* individuals (Chen et al., 2019; Li et al., 2022; Tsuda et al., 2016). By jointly analyzing nuclear (microsatellites) and mitochondrial markers, Tsuda et al. (2016) also detected two migration barriers, one roughly separating the northern from the southern range of *P. abies* and the other corresponding to the Ural Mountains. It was suggested that these two geographical barriers were responsible for the observed non-homogeneous longitudinal gene flow between the two species. Extensive genomic data confirmed the presence of gene flow from *P. obovata* to the northern domain of *P. abies* and suggested the existence of four main genetic ancestral clusters across the *P. abies*–*P. obovata* complex (Chen et al., 2019; Li et al., 2022). Unfortunately, a high sampling

bias towards *P. abies* and a limited sampling of individuals within the putative hybrid zone strongly limited inferences on species genetic interaction over multiple glacial cycles.

The present study investigates the role of hybridization in shaping the joint demographic history of the two spruce species over multiple glacial cycles and how it influenced their current spatial distribution. To do so, we used extensive genomic data (>480,000 SNPs) from both nuclear and chloroplast DNA from 542 individuals sampled in 55 populations. The sampling ranged from Western Europe and Fennoscandia to the Yenisei River in Siberia, with a significant sampling of the putative *P. abies–P. obovata* hybrid zone (Figure 1a). We then combined different population genetics analyses to tease apart the effect of ancient demographic events,

hybridization, and ongoing gene flow in shaping the current distribution of genetic variation. First, our study provides a precise mapping of the genetic diversity of both species at a continental scale and shows that gene flow between the two species is pervasive, with a nearly continuous transition from one genetic background to the other. Using extensive genome-wide markers, we also showed a bidirectional but asymmetrical introgression pattern between the two species. In the North, *P. obovata* alleles introgress into *P. abies* genome while *P. obovata* extends westwards. Whereas in the South *P. abies* alleles introgress into *P. obovata* genome as *P. abies* moves eastwards. Using coalescent simulations, we further demonstrated that the two species regularly came into contact across multiple ice-cycles, a process that gave rise to current genetic clusters. Hence,
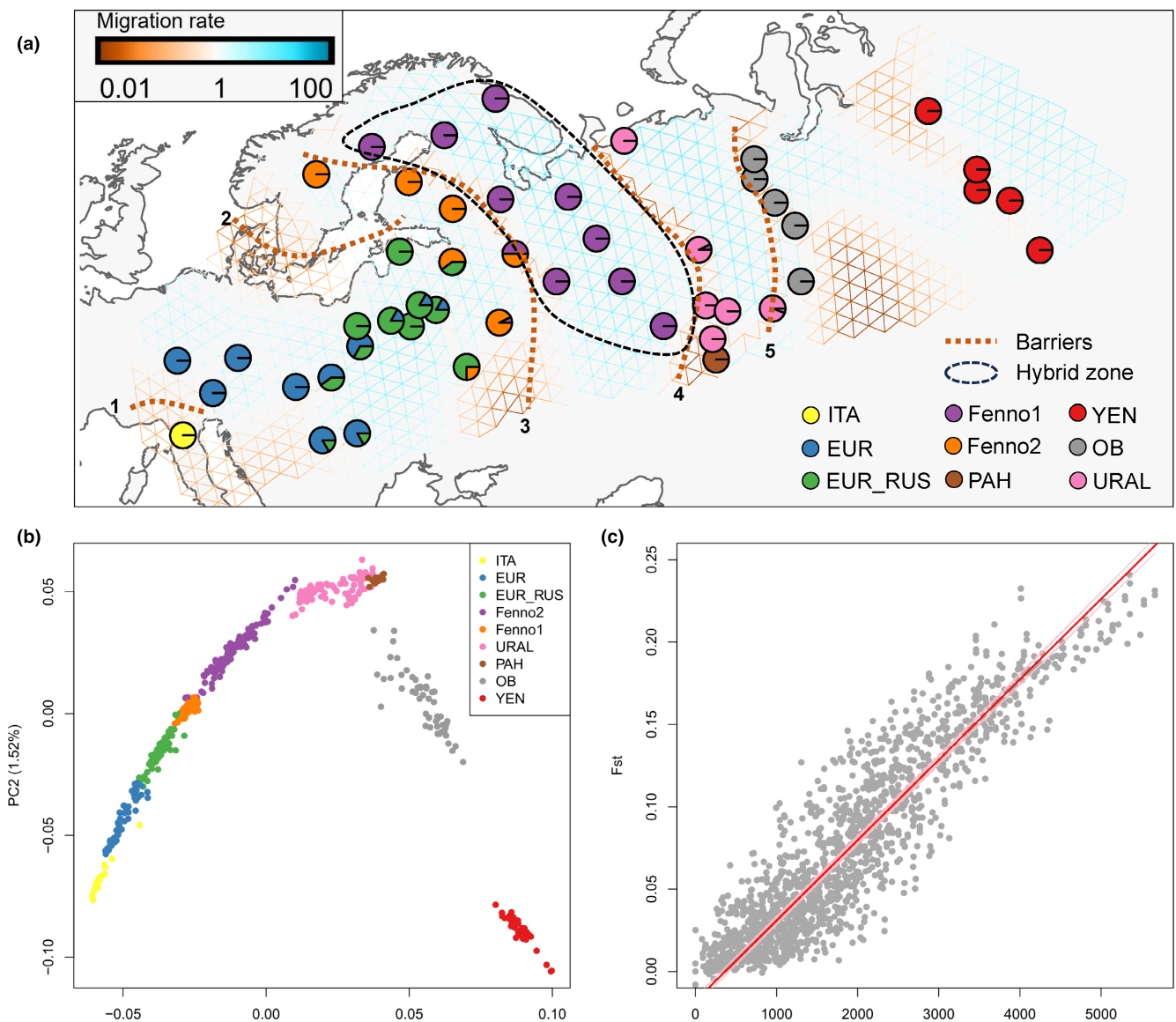


**FIGURE 1** Population structure and gene flow. (a) the pies represent the proportion of individuals in a population belonging to a given genetic cluster defined using *UMAP*. The background of the map represents effective migration surface across space estimated using *fEEMS*. The triangles represent the grids used in *fEEMS*, the colours refer to the scale on the top of the map giving the relative intensity of gene flow. Populations located inside a same deme of the grid were gathered as one single population. Main barriers identified were labelled with numbers. The core of the hybrid zone was marked with black dotted lines. Barrier 1: Alpine Mountains; Barrier 2: Baltic Sea; Barrier 5: Ural Mountains. (b) results of PCA shown as plotting PC2 against PC1. (c) Regression of population pairwise $F_{ST}$ over geodesic distances.

hybridization between the two species fostered their genetic diversity and may have helped both species overcome major climatic fluctuations over the quaternary.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling, sequencing and SNP calling

We sequenced the DNA of 542 trees from 55 populations sampled across *P. abies* and *P. obovata* natural ranges (Figure S1B, Appendices S1 and S2) using exome capture (40,018 probes, Vidalis et al., 2018). We aligned the raw reads against the reference genome of *P. abies* v1.0 (Nystedt et al., 2013) and trimmed the alignment files before SNP calling. After hard-filtering, 480,428 putatively neutral SNPs (i.e. synonymous, intronic and intergenic), pruned for high linkage disequilibrium ($r^2 < 0.5$), were kept for downstream analyses. Detailed procedures and filtering parameters are presented in Appendix S7: 'Sequencing and SNP calling'.

SNPs from cpDNA were identified following the same protocol as for nuclear SNPs but without pruning for high linkage disequilibrium. Only 484 samples were kept for the chloroplast dataset after removing samples with a too high missing rate (>30%).

### 2.2 | Population genetic structure and clustering

Population genetic structure was first investigated through principal component analysis (PCA) to assess spatial genetic patterns using *EIGENSOFT* v. 7.2.0 with default parameters (Galinsky et al., 2016). Fine-scale genetic clustering was then performed with Uniform Manifold Approximation and Projection (*UMAP*) analysis (Diaz-Papkovich et al., 2019). We estimated ancestry component coefficient using *ADMIXTURE* v. 1.3.0 (Alexander et al., 2009). We performed 10-fold cross validation with 200 bootstraps to define K, the theoretically optimal number of ancestry components (K ranges from 1 to 10). Finally, we used Bayesian inference implemented in *BayPass* v2.2 (Gautier, 2015) to estimate the empirical patterns of covariance in allele frequencies among populations and compared it with the clustering obtained with individual-based approach (*UMAP*). Detailed explanations for *UMAP* and *BayPass* are provided in Appendix S7: 'Genetic clustering'.

### 2.3 | Genetic diversity and population divergence estimation

For each population (i.e. sampling location) and each genetic cluster identified using *UMAP* approaches, we computed nucleotide diversity (Nei's π, Nei & Tajima, 1981) and Tajima's D (Tajima, 1989) using the R package '*PopGenome*' v2.7.5 (Pfeifer et al., 2014). The estimated nucleotide diversity was corrected with the length of effectively sequenced region and the proportion of filtered SNPs.

We also estimated Weir and Cockerham weighted $F_{ST}$ (Weir & Cockerham, 1984) for each population pair using *VCFtools* v0.1.17 (Danecek et al., 2011).

### 2.4 | Isolation-by-Distance and effective migration surfaces

The global pattern of Isolation-by-Distance (IBD, Morton, 2013) was first investigated by regressing population pairwise $F_{ST}$ to corresponding pairwise geodesic distances (R package '*geodist*' v0.08, Padgham, 2021); we used a Mantel test (Mantel, 1967) with 10,000 permutations to test the significance of the correlation between the two distance matrices (R, '*vegan*' v2.6, Dixon, 2003). We then identified corridors and barriers to gene flow using the *fEEMS* v1.0.1 (Marcus et al., 2021) software, that estimates effective migration surfaces across space. We adopted a grid size of $100\,km^2$ that was the best compromise between computational burden and model misspecification. All 55 populations were assigned to 48 different grids. Individuals from different sampling locations within the same grid that were genetically close were grouped, leaving a minor risk of bias in migration surface estimation with adjacent grids.

### 2.5 | Chloroplast haplotype group analysis

SNPs located in the chloroplast genome were used for unrooted maximum likelihood (ML) phylogenetic inference using *IQtree* v2.03 (Nguyen et al., 2015), see Appendix S7: 'Model comparison' – *IQtree*.

### 2.6 | Inference of ancient migration events between the main clusters

We used *TreeMix* v1.13 (Pickrell & Pritchard, 2012) to infer the pattern of population splits and major historical migration events among the main genetic entities. A maximum likelihood phylogenetic tree of the main genetic clusters identified using *UMAP* was first built by *TreeMix*. The support of the resulting phylogeny was evaluated by bootstrapping blocks of 500 SNPs. The homogenous genetic cluster formed by populations located along the Yenisei River was used as the outgroup to root the tree and up to 10 edges ('migration events') were added to the tree (see Appendix S7: 'Model comparison' – *TreeMix*).

### 2.7 | Demographic history inference of the main clusters

In order to test alternative demographic models (see Appendix S7: 'Model comparison' – *FastSimCoal2*), we used the coalescent-based composite likelihood method implemented in *FastSimCoal2*

(Excoffier et al., 2021), considering a generation time of 25 years and a mutation rate of $2.75 \times 10^{-8}$ per site per generation (Ann et al., 2007; Hanlon et al., 2019; Nystedt et al., 2013).

We used the output of the *TreeMix* analysis to define the demographic models to be tested. The models were built with increasing complexity and migration events were added progressively in the order suggested by the *TreeMix* analysis. Five scenarios with different numbers of edges added (ranging from 0 to 4) were initially tested. We then tested additional models, notably considering different origins for the clusters with ambiguous phylogenetic positions in *TreeMix* (see Section 3: '*Biased ancient migration from P. obovata into P. abies*'). Migration was only allowed between adjacent clusters to avoid over-parametrization (the various models are presented in Appendices S3 and S4; also see Appendix S7: 'Site frequency spectra (SFS)' and 'Model comparison' – *FastSimCoal2*). We also explored simpler models based on the four clusters defined using *ADMIXTURE* (see also Appendix S7: '*FastSimCoal2* simulation with four clusters', Table S1 and Figure S5).

## 3 | RESULTS

### 3.1 | Continuous change in genomic background from one species to the other

First, we measured genetic diversity across all populations from 480,428 putatively neutral nuclear SNPs. Nei's nucleotide diversity, $\pi$, ranges from 0.00524 to 0.00687 and shows a longitudinal gradient, increasing from West (*P. abies*) to East (*P. obovata*, Spearman's $rho = 0.5$, $S = 13,413$, $p < 0.001$, Appendix S2). Tajima's D statistics were negative in all populations (−0.91 to −0.031, median: −0.397, Appendix S2). Considering that Tajima's D is expected to roughly follow a beta distribution of mean 0 and variance 1 (Tajima, 1989), the populations do not significantly depart from the standard coalescent model even if the global negative trend might suggest recent population expansion, as expected after the LGM. We then investigated genetic structure with a PCA (Figure 1b and Figure S2). The first two principal components (PCs) together explain ~5.1% of the total genetic variation and clustering reflects geography (Figure S2E,F). No clear delineation between the two species can be drawn.

As some population structure is still captured by other PCs (e.g. PC3 and 4, see Figure S2C,D), we performed dimensionality reduction incorporating the first five PCs using *UMAP* (Diaz-Papkovich et al., 2019). Nine genetic clusters can be distinguished. Among them, two include only one or two populations (Figure 1a, Figure S2A and Appendices S1 and S2). The first one (ITA) includes two populations located at different altitudes (ITAL and ITAH) in the Apennine region in Italy (Figure 1, Figure S1B). The second cluster (PAH) is from a population located on an elevated plateau in Southern Urals (Mount Iremel). These two clusters are isolated from the rest of the range of *P. abies* and *P. obovata*. The seven other clusters grouped individuals according to their geographical distribution (Figure 1a). In *P. obovata*,

the main population structure runs along the longitude where three main clusters can be discriminated. The Yenisei (YEN) cluster only contains individuals from populations sampled along the Yenisei River. Populations located along the Ob River, East of the Ural Mountains range, form a cluster of their own (OB) and those located West of the Ural Mountains range gather into a third cluster (URAL, Figure 1a). It is worth noting that one isolated population from the northern distribution of *P. obovata*, Indigo, belonged to the URAL cluster. In contrast with *P. obovata*, the main population structure runs along latitude in *P. abies*. The EUR cluster gathers populations from the Alpine and Carpathian Mountain ranges, and the EUR_RUS cluster groups populations from the Russia-Baltic region. The last two clusters, Fenno1 and Fenno2, cover most of the hybrid zone that extends from Fennoscandia up to the Ural Mountains and split it into a Northern-Eastern (Fenno1) and a Southern-Western (Fenno2) cluster (Figure 1a,b). A population-based clustering approach (Günther & Coop, 2013) confirmed the clustering based on individual data (Figure S1).

### 3.2 | Extensive gene flow explains the continuous change in genomic background

Population structure analyses suggested the occurrence of extensive ongoing gene flow between *P. abies* and *P. obovata* genetic backgrounds. We first investigated the pattern of Isolation-by-Distance (IBD) by regressing population pairwise $F_{ST}$ over the geodesic distance between each population in a pair. The divergence among populations was globally low with the highest $F_{ST}$ (~0.25) measured between Yenisei River (*P. obovata*) and Italian populations (*P. abies*). The regression revealed a striking IBD pattern across the whole range of both species ($r^2 = 0.82$, Figure 1c, Mantel's correlation coefficient $r$: 0.92, $p < 0.001$) and no gap in the distribution was observed. In order to capture local deviations in gene flow (i.e. barriers and corridors), we estimated the geographical distribution of effective migration rates using *fEEMS* (Marcus et al., 2021). As expected, areas with reduced migration rates (hereafter 'barriers') delineate the main genetic clusters identified with *UMAP* while the migration rate was higher than the average within clusters (Figure 1a). Within *P. obovata*, the Ural Mountains clearly constitute a barrier to gene flow separating URAL and OB genetic clusters (barrier 5, Figure 1a). However, in *P. abies*, no clear topological barrier can be found to explain the separation of EUR_RUS from EUR. The Baltic Sea was identified as a barrier to gene flow between *P. abies* clusters and Fenno2 likely due to recent recolonization history after LGM (barrier 2, Figure 1a). Importantly, we found two main barriers to gene flow delineating a corridor between Southern Urals and Northern Fennoscandia that do not correspond to any clear topological variation but delineate the hybrid zone. The first barrier (barrier 3) separates Fenno2 from Fenno1 in the West and the second (barrier 4) separates Fenno1 from URAL in the East (Figure 1a). These two barriers are probably best explained by a combination of ancient demographic events and natural selection. For instance, barrier 3 matches the transition in climatic groups defined by Karunarathne et al. (2024).

## 3.3 | Four ancestral populations contributed to current genetic variation

To investigate how ancient demographic history shaped the current distribution of genetic variation, we used *ADMIXTURE* (Alexander et al., 2009) to first estimate ancestry components and better understand the origin of the current population structure. Based on 10-fold cross-validation (CV), *K*=3 or 4 best explained the data (Figure 2, Figure S3). In both cases, *ADMIXTURE* revealed a continuous genetic composition change from 'pure' EUR (*P. abies*) to 'pure' YEN (*P. obovata*)

genetic background, which is consistent with the result of the PCA analysis. However, in contrast with what would be expected if *K*=2, the admixed populations do not directly share EUR and YEN genetic components. Instead, the analysis revealed a two-fold admixture pattern, with populations sharing either EUR (EUR_RUS, Fenno1 and Fenno2) or YEN (PAH, URAL and OB) genetic component with a cluster of third origin (Figure 2a). The main ancestry component found for these clusters for *K*=3 (green in Figure 2a, top panel) is split into two ancestral components when *K*=4 (see AC_FEN, green, and AC_UR, olive, in Figure 2a, bottom panel).
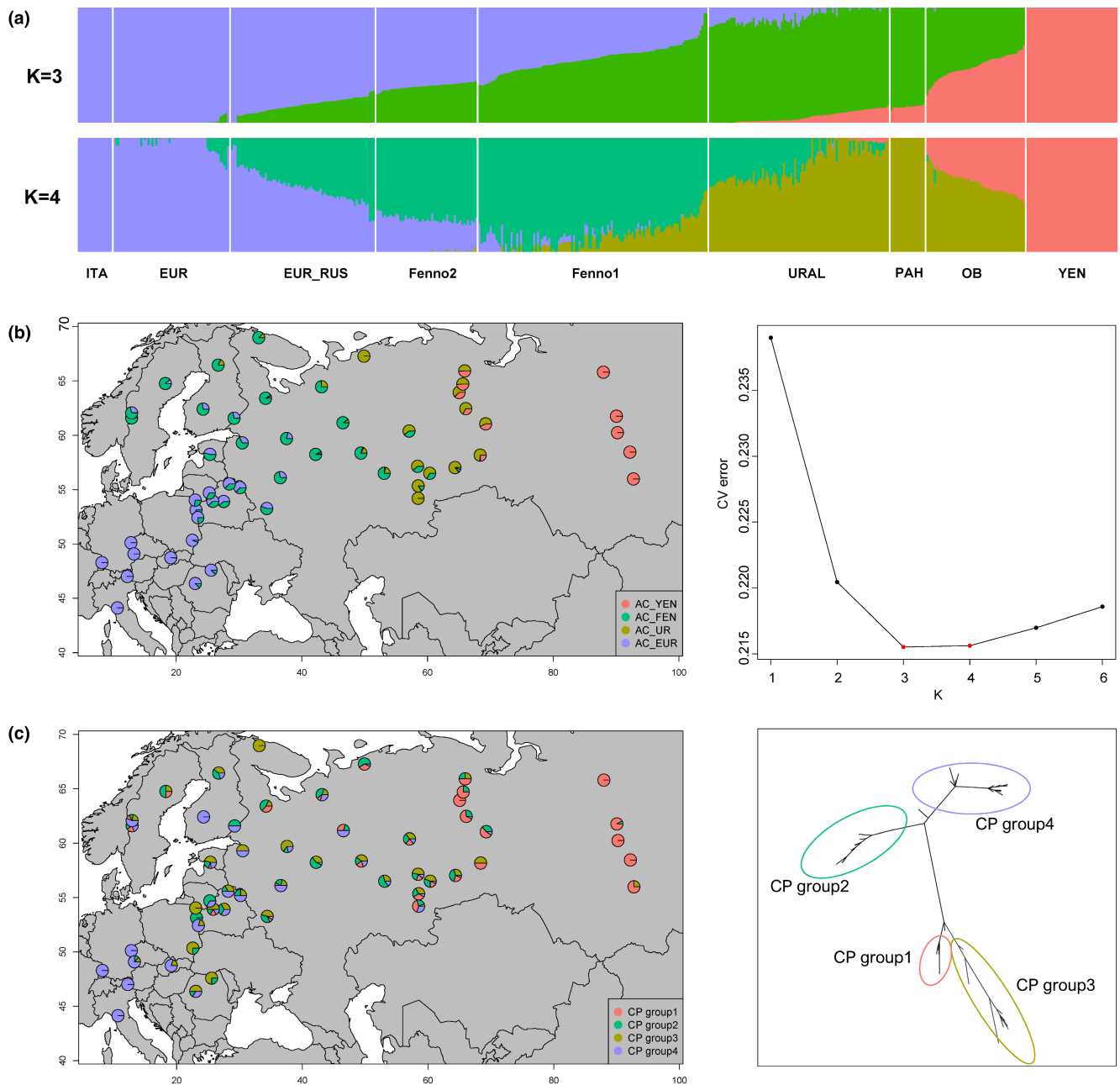


**FIGURE 2** Distribution of ancestry components and chloroplast haplotypes. (a) *ADMIXTURE* plot showing the ancestry components (nuclear DNA) distribution for *K*=3 (upper panel) and *K*=4 (bottom panel). (b) Geographical distribution of ancestry components estimated with *ADMIXTURE* analysis (*K*=4, left panel) and scree plot of cross-validation errors for different *K* values (right panel). (c) Geographical distribution of the four main chloroplast DNA haplotype groups (left) and maximum likelihood tree showing their phylogenetic relationships (right panel).

In conifers, chloroplasts are haploid, paternally inherited (i.e. through pollen) but dispersed with both pollen and seeds. Investigating the distribution of genetic variation of chloroplast DNA (cpDNA) thus provides additional information on pollen and seed dispersal (Scotti et al., 2008; Tollefsrud et al., 2015). Four well-supported chlorotype groups were identified using a maximum likelihood (ML) phylogeny built on 87 chloroplast SNPs and their geographical distribution globally matches the distribution of the ancestry components inferred from nuclear markers (*ADMIXTURE* analysis, *K*=4) but with a larger geographical spread (Figure 2c). Group 1 and group 4, that are predominantly found along the Yenisei River and in Western Europe, respectively, are also found in populations located within the hybrid zone while the two corresponding main ancestry components estimated from nDNA, AC_EUR and AC_YEN, are not into contact in the hybrid zone. The other two groups (CP groups 2 and 3) are widely distributed across the hybrid zone with no clear geographical pattern even though they likely originate from the two different species (Figure 2c). Finally, most of the populations from the hybrid zone display chlorotypes belonging to at least three different groups. This analysis supports the existence of cross-species long-range pollen flow and shows that the chlorotypes originating from *P. obovata* (CP group 1) introgress further West in the North and those originating from *P. abies* (CP group 4) introgress further East in the South.

The distribution of nuclear ancestries and chlorotype groups together suggest that four instead of two main ancestral populations contributed to the contemporary genetic variation across the joint distribution of the two species.

## 3.4 | Biased ancient migration from *P. obovata* into *P. abies*

The previous analyses revealed a gradual transition from the genetic background of one species to the other that extends from Northern Fennoscandia up to Southern Urals, suggesting ancient admixture events between *P. abies* and *P. obovata*. Hence, we used *TreeMix* (Pickrell & Pritchard, 2012) to quantify the number, intensity and direction of main migration events between the two species. The best model explained 99.8% of the total variance and was obtained by adding four edges (i.e. main migration events) to the initial tree (Figure 3a and Figure S4). The resulting phylogenetic tree first separates YEN and OB from EUR, Fenno1, Fenno2, and EUR_RUS clusters with URAL branching out from the branch connecting these two groups. EUR_RUS cluster splits early from the common ancestor of EUR, Fenno1 and Fenno2 clusters. The model suggests that the first migration event occurred from *P. obovata* to what is today, Fenno2 (arrow 1 in Figure 3a). More importantly, we detected two concomitant migration events occurring from URAL towards OB (*P. obovata*) and Fenno1 (*P. abies*) (arrows 2 and 3, respectively, in Figure 3a). URAL is thus an ancestral cluster that influenced both *P. obovata* and *P. abies* genetic diversity, shedding a new light on the patterns observed with *ADMIXTURE* and *fEEMS* analyses. The

most recent migration event was from EUR towards EUR_RUS (arrow 4, Figure 3a). This explains the close relationship between the two clusters observed with PCA and *ADMIXTURE*, despite the long divergence time suggested by *TreeMix*. All migration events were confirmed with *f3*-tests as suggested by (Chen et al., 2018) (Appendix S5). However, one needs to be cautious when interpreting the reticulate evolutionary history revealed by *TreeMix*. When the migration weights are close to 0.5, the branching and the source population for migration could be flipped, both being theoretically indistinguishable from one another, see for instance arrows 3 and 4 in Figure 3a.

## 3.5 | Demographic history inference

The results of *TreeMix* provide new insights into the complex evolutionary dynamics of the two species but did not allow for a detailed inspection of their demographic histories. It is not clear whether URAL split from YEN or EUR. Similarly, the origin of EUR_RUS cluster remains unclear due to the roughly equal contribution of the two source populations. To solve these issues, we explored different demographic scenarios with different sets of parameters and topologies with *FastSimCoal2* (Excoffier et al., 2021). Among the 36 models we explored, one model stands out and recovered the observed site frequency spectra well (CLR: 747,198, Appendix S3, S4 and S6). The best model and the confidence intervals for the estimates are presented in Figure 3b and Table 1.

The use of a coalescent-based approach confirmed that the URAL cluster is ancient ($T_{UR1}$: ~820 kya) and diverged from the YEN cluster after the split between the ancestors of *P. abies* and *P. obovata* ($T_{anc}$: ~2.5 mya). It also emphasizes the importance of admixture events in the history of the two species and confirms the hybrid origin of some clusters. Fenno2 resulted from the admixture between URAL and EUR ($T_{Fenno2}$: ~230 kya) with a higher contribution from the *P. abies* genetic background (relative contribution, 60%) while Fenno1 originated slightly later from the admixture between URAL and Fenno2 (~187 kya), with a major contribution from Fenno2 (~90%). At roughly the same time, the OB cluster formed from the admixture between YEN (~90%) and URAL (~10%). The concomitance of these two events suggests that they occurred during the same interglacial period. Finally, the EUR_RUS cluster resulted recently from the admixture between EUR (~50%) and Fenno2 (~50%) clusters, likely as a result of the post-LGM population expansion ($T_{EUR\_RUS}$: ~3150 years ago).

In summary, *TreeMix* and *FastSimCoal2* analyses highlight three key inferences of the evolutionary history of *P. abies* and *P. obovata*. First, most of the clusters originated much before the LGM. Second, the main clusters except for the two 'pure' ones (EUR and YEN) were the results of admixture between more ancient genetic entities. Third, recurring hybridization between the two species played a crucial role in shaping the current distribution of genetic variation, with a central role played by the URAL cluster that contributed to several admixture events involving both species.
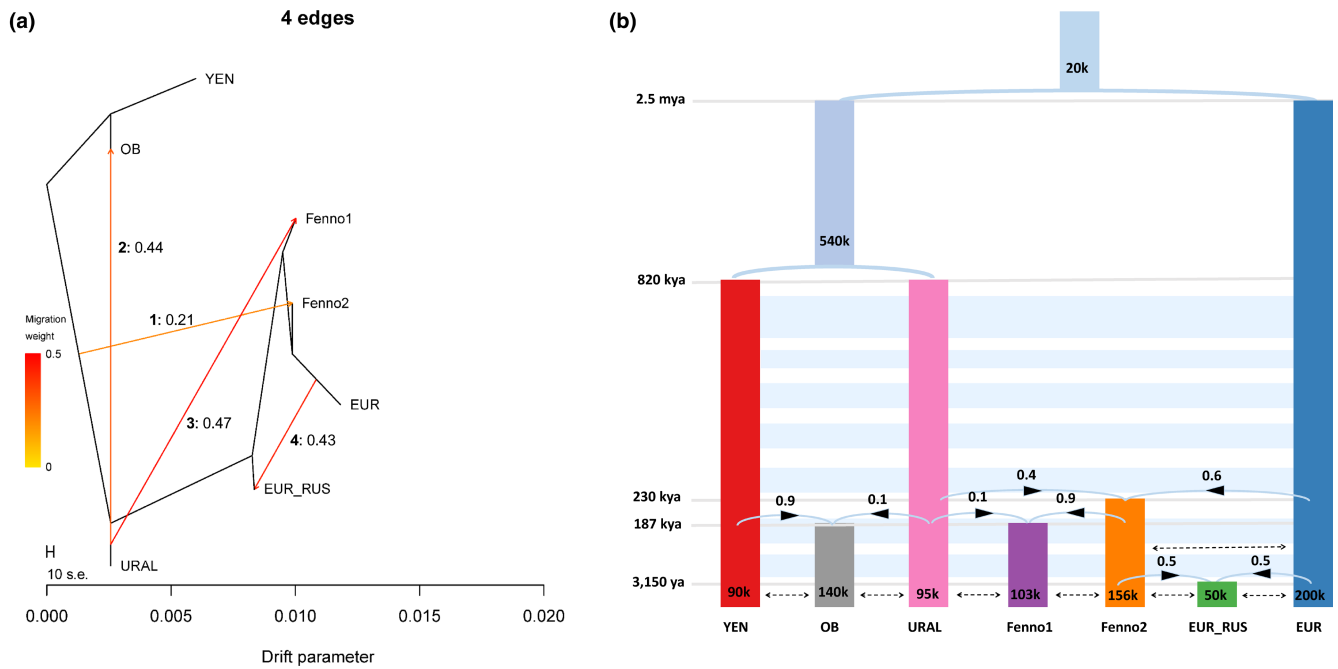
**FIGURE 3** Demography history inference. (a) The tree with the highest support in *TreeMix* analysis is represented. The main migration events are shown with arrows (labelled from 1 to 4). (b) *FastSimCoal2* analysis, the best model with four admixture events is represented. Time estimates (number on the left) were obtained from the number of generations considering a mean generation time of 25 years. Horizontal blue bars indicate roughly the glacial periods for the last 800,000 years. For both plots, the numbers above/next to the arrows indicate the relative contribution of the source populations to the sink populations.

## 4 | DISCUSSION

In the present study, we investigated the joint demographic history of two dominant spruce species in Eurasian boreal forests. Our results highlight three main features of this demographic history. First, most pivotal events occurred before the Last Glacial Maximum, indicating that the current population structure is not solely a result of post-LGM events. Secondly, bidirectional gene flow/admixture played an extremely important role in shaping the geographical distribution of the current diversity. Thirdly, the divergence of the main genetic ancestries predated LGM and persisted across several glacial–interglacial cycles despite substantial gene flow occurring within the common habitat of these two species. This remarkable resilience to gene flow highlights the enduring impact of historical events on the genetic makeup of these species. Below we discuss the implications of these features on the inferences that can be drawn on the demographic history of the two species.

### 4.1 | ITA, PAH and indigo are likely 'relic' genetic clusters/populations

Out of the nine genetic clusters identified using *UMAP*, the ITA and PAH clusters included only few individuals from isolated populations from the Apennines (Northern Italy) and from the Southern Urals, respectively. ITAH and ITAL were collected from the southern limits of the natural range of *P. abies* and are geographically isolated from the other populations of the species, as confirmed by *fEEMS* analysis

(Barrier 1, Figure 1a and Figure S1). The low nucleotide diversity of these populations further supports a higher level of genetic drift than in other populations. Similarly, higher Tajima's D estimates than in surrounding populations indicate that their demographic histories also differed (Appendix S2). The autochthonous origin of ITA was first hypothesized by Chiarugi (1936a, 1936b), and was supported by phytological, morphological and genetic data (Borghetti et al., 1988; Bucci & Vendramin, 2000; Ferrarini, 1977; Giannini et al., 1991, 1994; Magini et al., 1980; Vendramin et al., 2000). The most recent palynological records also support the local persistence of *P. abies* at low altitudes in the northern Apennines during the LGM (Magri et al., 2015; Morales-Molino et al., 2021; Ravazzi, 2002; Vescovi et al., 2010).

Individuals from PAH, a population located at high altitude, clustered separately from all nearby populations (*UMAP*), including PAL, a nearby population of lower altitude. In contrast, PAH is genetically closer than expected, given the geographical distance, to Indigo, a population located at the northern limit of the *P. obovata* range (Figure 1, Figure S1). Both PAH and Indigo share phenotypic traits characteristic of adaption to harsh habitats with extreme temperatures and strong winds; typically, small trees with a 'bushy' architecture (V. Semerikov and M. Lascoux, personal observation), as also encountered at high altitude in Norwegian and Swedish mountain ranges (for a picture, see Nota et al., 2022). The observed pattern could therefore be the combined result of convergent local adaptation to extreme environments and long-term survival through glacial periods as small and isolated refugia. Sediment coring analysis confirmed the presence of spruce pollen during LGM in Indigo (Väliranta et al., 2011). It is also

**TABLE 1** Demographic parameters estimated for the model using *FastSimCoal2*.

| Parameters | Point estimation | 2.50% | 97.50% |
|---|---|---|---|
| $N_{EUR}$ | 200,301 | 198,875 | 268,481 |
| $N_{OB}$ | 139,799 | 123,490 | 169,277 |
| $N_{YEN}$ | 90,266 | 63,134 | 148,423 |
| $N_{URAL}$ | 95,476 | 91,911 | 142,727 |
| $N_{Fenno1}$ | 103,151 | 97,447 | 112,885 |
| $N_{Fenno2}$ | 156,520 | 152,277 | 202,673 |
| $N_{EUR\_RUS}$ | 50,108 | 50,386 | 123,478 |
| $M_{YEN-OB}$ | 0.94 | 0.81 | 0.98 |
| $M_{EUR-Fenno2}$ | 0.60 | 0.56 | 0.69 |
| $M_{URAL-Fenno1}$ | 0.08 | 0.02 | 0.18 |
| $M_{EUR-EUR\_RUS}$ | 0.46 | 0.46 | 0.49 |
| $T_{Fenno2}$ | 230,650 | 214,191 | 281,793 |
| $T_{URAL}$ | 818,900 | 571,852 | 1,088,353 |
| $T_{Fenno1}$ | 188,650 | 181,676 | 206,930 |
| $T_{OB}$ | 187,225 | 105,660 | 244,877 |
| $T_{anc}$ | 2,439,200 | 2,300,048 | 2,790,453 |
| $T_{EUR\_RUS}$ | 3150 | 3411 | 9683 |

*Note*: $M$, migration proportion from source cluster to sink cluster for admixture events; $N$, effective population size; $T$, time for divergence or admixture events; $T_{anc}$, coalescent time for all clusters.

**TABLE 2** Genetic differentiation among the putative ancestry populations inferred by *ADMIXTURE*.

| | AC_YEN | AC_EUR | AC_UR |
|---|---|---|---|
| AC_EUR | 0.173 | | |
| AC_UR | 0.087 | 0.139 | |
| AC_FEN | 0.124 | 0.062 | 0.064 |

well known that high mountains served as refugia for *P. abies* populations during glacial periods (Nota et al., 2022; Parducci et al., 2012; Tollefsrud et al., 2008, 2015) and the same probably applies to the PAH population. In Fennoscandia, these high-altitude relics did not have a significant contribution to the re-colonization after the LGM (Nota et al., 2022) and the same seems to be true in PAH that genetically differs from nearby PAL. PAH and Indigo in *P. obovata*, as well as ITA in *P. abies* appear to be relics of past distributions.

## 4.2 | Two ancient cryptic refugia bridge *P. abies* and *P. obovata* genetic backgrounds

In *P. abies*, our clustering corroborates the primary population structure identified in earlier investigations (Chen et al., 2019; Li et al., 2022; Milesi et al., 2023). In particular, we confirmed that the genetic contribution of *P. abies* expands as far East as the Southern Ural Mountains (Krutovskii & Bergmann, 1995; Tsuda et al., 2016). Much less was known about *P. obovata* genetic structure, and our study revealed the existence of at least three main genetic clusters: from East to West, a first cluster comprising populations sampled along the Yenisei River (YEN), a second one grouping populations along the Ob River (OB), and a third one west of the Ural Mountains (URAL) (Figure 1, Figure S1).

Recent literature showed a pattern of introgression of *P. abies* populations by *P. obovata* genetic background, in particular in the Northern range of *P. abies* (Chen et al., 2019; Li et al., 2022; Tollefsrud et al., 2015; Tsuda et al., 2016). The analyses of mitochondrial DNA and SSR markers by Tsuda et al. (2016) suggested that cryptic *P. abies* populations might have acted as stepping stones for *P. obovata* westwards move. The presence of those 'stepping stones' would help explain the asymmetric patterns of introgression observed for mtDNA and nDNA (see Figures 1a and 3b in Tsuda et al., 2016) and would also be consistent with Currat et al. (2008) introgression model. In line with these results, the admixture analysis suggested that K=3 or K=4 ancestry components were at the origin of today's population structure (Figure 2b). It implies that at least one (K=3), or more probably two (K=4), ancestry components contributed to the genetic variation we observe today. Additionally, we detected four main chloroplast haplotypes. Like nDNA, cpDNA is dispersed through both pollen and seeds (Petit et al., 2005) but it has been shown to introgress more readily than nDNA (Soltis et al., 1992 and reference therein). Their phylogenetic relationships and spatial distribution suggest that these haplotypes arose from the same ancestral populations as the ones identified using nuclear polymorphisms. This further supports our hypothesis of the existence of two cryptic refugia, one with a more *P. abies* genetic background (AC_FEN and CP group2, Figure 2) and one with a more *P. obovata* background (AC_UR and CP group3, Figure 2).

By comparing the distribution of nuclear and chloroplast genetic diversity from our study with the distribution of mitochondrial genetic diversity from Tsuda et al. (2016), we can infer that past populations in the hybrid zone were likely founded by westward colonization from the *P. obovata* domain, the mitochondrial haplotype being retained into the two refugia in the hybrid zone. Individuals from one refugium then received a massive gene flow from *P. abies* through pollen dispersal and the local chloroplast genome was replaced by the one from *P. abies* (chloroplast capture, Tsitrone et al., 2003; Schulte et al., 2021; Yang et al., 2021) while the other refugium was more influenced by *P. obovata*. This hypothesis would also explain the inconsistency of nuclear and chlorotype divergence between the two main ancestral components/groups in the hybrid zone: CP group 2 and CP group 3 originate from different genetic backgrounds (Figure 2c), while the $F_{ST}$ between the two nuclear ancestral components (AC_FEN and AC_UR) estimated with *ADMIXTURE* was among the lowest (Table 2), suggesting possible different origins of nuclear and chloroplast sequences for one of the two ancestral populations in the hybrid zone. Alternatively, the observed cytonuclear discordance can also be explained by incomplete lineage sorting instead of chloroplast capture. However, it is difficult to distinguish the two scenarios (e.g. Liu et al., 2020; Morales-Briones et al., 2018; Rose et al., 2021), especially with extensive ongoing gene flow and no proper outgroup, which it is the case in this study.

Both, a tree-based approach (*TreeMix*) and coalescent simulations (*FastSimCoal2*) supported recurrent admixture events between the main genetic entities that gave rise to the main genetic clusters. Even if a precise timing of these admixture events is out of reach given the uncertainties about mutation rate and generation time in spruce species, it is worth noting that they all date back before the LGM (Figure 3b and Table 1). Our best estimates showed that the main admixture events between *P. abies* and *P. obovata* genetic background occurred during the Pleistocene in agreement with what was already described (Chen et al., 2019; Tsuda et al., 2016) and likely during inter-glacial recolonization. Note that a direct comparison of the estimates between these studies is not possible as the demographic models and the gene pool considered differ across studies. Two genetic clusters, URAL and Fenno2, were revealed to be ancient genetic entities that survived multiple glacial cycles and play substantial parts in shaping the current distribution of genetic variation. The two cryptic refugia likely inhabited the ancestral populations of each of the two clusters during glacial periods, promoting genetic divergence and giving birth to the two ancestries (AC_UR and AC_FEN, Figure 2b).

The location of these refugia remains unclear but our analyses point towards Western Russia-Eastern Fennoscandia and the Southern Ural Mountains. Tollefsrud et al. (2008) showed that early Holocene pollen records were enriched in the Alps and Carpathian Mountains, which are known refugia for *P. abies*, but also in a large area in Western Russia. The latter could thus be the location of a refugium during glacial periods as also suggested by Giesecke and Bennett (2004) and Latałowa and van der Knaap (2006). Locating the other refugium could be more challenging as the classical concept of refugia may not apply due to the lack of complete glaciation of the region. According to palaeobiological records, Siberia was a dry desert interspersed with pockets of forested areas (Semerikov et al., 2013 and reference therein). Still, Karunarathne et al. (2024) suggested that the southern part of the Siberian Plain was a suitable habitat for *P. obovata* during glacial periods and the presence of spruce during the LGM in the Baikal region is attested by fossil pollen (Kobe et al., 2022). Using ecological niche modelling and backward climate projection, Karunarathne et al. (2024) also suggested that the distribution ranges of *P. abies* and *P. obovata* overlapped during interglacial periods in their Northern range (~145 kya). This could, for instance, explain both the admixture of the two species at the origin of the Fenno1 cluster (~180 kya, Figure 3b) and the fact that *P. obovata* contribution expands further West in the Northern range of *P. abies*.

## 4.3 | Ecological factors could explain the maintenance of distinct genetic clusters despite extensive gene flow

Current genomic variation shows a continuous spatial distribution between *P. obovata* and *P. abies*. The low differentiation between populations geographically distant and belonging to different genetic clusters having diverged long ago highlights the maintenance of large effective population sizes through time and extensive gene flow sustained by long-distance pollen dispersal (Figures 1c, 2b). Despite such a high gene flow, well-delineated genetic clusters persist at the nuclear level, probably because of the recurrent sheltering by well-defined glacial refugia. While distinct barriers to gene flow explain the distribution of current genetic diversity (e.g. mountain ranges, the Baltic Sea), there is no direct evidence of a topographic cause for the barrier within the hybrid zone (barrier 2 and 3, Figure 1a). This barrier could thus be the result of the interplay of the ancient demographic events described above with ecological factors and local adaptation. A whole body of literature discriminates *P. abies* from *P. obovata* and their hybrids based on the shape of the cone scale. Notably, the distribution range of these groups aligns well with the distribution of the main genetic clusters we identified within the hybrid zone (e.g. Popov, 2010, 2013; Pravdin, 1975). Two types of hybrids can be further discriminated using morphological traits, hybrids with properties of *P. abies* or with properties of *P. obovata* (e.g. Nakvasina et al., 2019; Orlova et al., 2020) mirroring the distribution of the ancestry components within the hybrid zone. Nakvasina et al. (2017) showed that the various hybrid forms display different phenotypic plasticity and that both hybrid forms had higher survival rates than the non-hybrids. The provenance test was conducted in the Arkhangelsk Region, the highest survival rate being achieved by the local genotype, namely hybrid with '*P. obovata*' properties. Some of the samples from the provenance tests were also included in this study (samples from population R1, R2, R5, R8, R20, R24, R27, R29, R35 and R38, see Figure S1B). In line with these studies, Karunarathne et al. (2024) showed that the hybridization between the two species enlarged both species' ecological niches, with hybrids occupying a specific ecological niche. Li et al. (2022) have further evidenced the role of local adaptation in the maintenance of the contact zone between Fenno1 and Fenno2 in Sweden, and that despite extensive gene flow. These studies thus further support the role of natural selection in maintaining hybrid zones between locally adapted ecotypes. It would certainly be interesting to extend the approach developed in Li et al. (2022) to other contact zones across the ranges of *P. abies* and *P. obovata*.

## 4.4 | Hybridization enhances genetic diversity but does not affect species boundaries

Taxonomic status of *P. abies* and *P. obovata* has been recurrently discussed and is likely to remain controversial (see Lockwood et al., 2013 and references therein). Our study shed a new light on the intensity of the gene flow between *P. abies* and *P. obovata* across what can be perceived as one of the largest hybrid zones but does not challenge their species status. Instead, we propose to consider *P. abies* and *P. obovata* more as a syngameon (Grant, 1981) as the deep-seated genetic structure between the two species also matches already characterized phenotypic differences (e.g. Nakvasina

et al., 2017; Popov, 2010, 2013; Pravdin, 1975) and ecological niche preferences (Karunarathne et al., 2024). Also, fixation index values between non-admixed populations of *P. abies* and *P. obovata* (from $F_{ST}=0.18$ to 0.25) are the highest reported so far for closely related spruce species (e.g. compared to *P. asperata* and *P. crassifolia*, Feng et al., 2023; *P. glauca* and *P. engelmannii*, De La Torre et al., 2014; *P. meyeri*, *P. koraiensis* and *P. mongolica*, Liu et al., 2024). In striking contrast, for instance, Scots pine (*P. sylvestris*) displays a much lower genetic differentiation across an even broader distribution range (Bruxaux et al., 2024).

Our study is in line with studies highlighting the extent and role of hybridization in forest trees species' evolutionary dynamics (e.g. Leal et al., 2024; Budde et al., 2023; Wang et al., 2019; Pfeilsticker et al., 2022; Larson et al., 2021, respectively for birch, beech, poplar, eucalypts or rainforest trees species). When not selected against, hybridization has been shown to enhance genetic diversity and foster species response to main climatic changes (e.g. Buck et al., 2023 in pines, Cannon & Petit, 2020 in oaks, Karunarathne et al., 2024 in spruce), contribute to local adaptation through adaptive introgression (e.g. Chhatre et al., 2018; Jones et al., 2018; Leroy et al., 2020; Platt et al., 2019; Whitney et al., 2010) or facilitate tree invasion dynamics (Gaskin, 2017 and reference therein) and should therefore be considered for forest conservation and management plan and trees breeding programme (Janes & Hamilton, 2017).

## 5 | CONCLUSION

Thanks to the development of Next Generation Sequencing technologies, it is now affordable to conduct population level studies at continental scale with extensive genomic markers even for non-model species. This offers new opportunities for apprehending the evolutionary importance of species hybridization, in particular for species with long range gamete dispersal. Our study highlights the importance of considering whole species complexes, instead of separate entities, to retrieve complex demographic histories. It showed that the recurrent hybridization between the two-spruce species that occurred during multiple glacial–interglacial cycles likely helped both species to overcome drastic contraction of their ranges and potentially fostering recolonization using refugia of hybrid nature as stepping stones. The fine mapping of genetic diversity changes at continental scale also allowed us to describe an extensive hybrid zone with an unprecedented resolution. The hybrid zone extends from Northern-Fennoscandia to Southern Urals where no topological barriers are found, strongly suggesting a main role of natural selection in its maintenance.

## AUTHOR CONTRIBUTIONS

PM and ML conceived and designed the research. QZ, PK, LA and CC performed experiments and analysed data. LO, KH, AP, GGV, EN and ML collected samples. QZ, ML and PM wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

## CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

## DATA AVAILABILITY STATEMENT

The raw-reads are available in SRA (https://www.ncbi.nlm.nih.gov/sra) under bioproject numbers PRJNA511374 and PRJNA1007582.

## ORCID

*Qiujie Zhou* https://orcid.org/0000-0001-7351-2371
*Chen Chen* https://orcid.org/0009-0008-0539-7517
*Lars Opgenoorth* https://orcid.org/0000-0003-0737-047X
*Katrin Heer* https://orcid.org/0000-0002-1036-599X
*Martin Lascoux* https://orcid.org/0000-0003-1699-9042
*Pascal Milesi* https://orcid.org/0000-0001-8580-4291

## REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

Ann, W., Syring, J., Gernandt, D. S., Liston, A., & Cronn, R. (2007). Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for Pinus. *Molecular Biology and Evolution*, *24*, 90–101.

Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, *16*, 113–148.

Belokon, M. M., Belokon, Y. S., Petrova, E. A., Vasilyeva, G. V., Efimova, A. P., Zakharov, E. S., Goroshkevich, S. N., & Politov, D. V. (2022). Ancient and modern hybridization of Siberian stone pine and dwarf Siberian pine in the south of Yakutia. *Russian Journal of Genetics*, *58*, 1287–1296.

Binney, H. A., Willis, K. J., Edwards, M. E., Bhagwat, S. A., Anderson, P. M., Andreev, A. A., Blaauw, M., Damblon, F., Haesaerts, P., Kienast, F., Kremenetski, K. V., Krivonogov, S. K., Lozhkin, A. V., MacDonald, G. M., Novenko, E. Y., Oksanen, P., Sapelko, T. V., Väliranta, M., & Vazhenina, L. (2009). The distribution of late-quaternary woody taxa in northern Eurasia: Evidence from a new macrofossil database. *Quaternary Science Reviews*, *28*, 2445–2464.

Borghetti, M., Giannini, R., & Menozzi, P. (1988). Geographic variation in cones of Norway spruce (*Picea abies karst*). *Silvae Genetica*, *37*, 178–184.

Bruxaux, J., Zhao, W., Hall, D., Curtu, A. L., Androsiuk, P., Drouzas, A. D., Gailing, O., Konrad, H., Sullivan, A. R., Semerikov, V., & Wang, X. R. (2024). Scots pine – panmixia and the elusive signal of genetic adaptation. *New Phytologist*, *243*(3), 1231–1246.

Bucci, G., & Vendramin, G. G. (2000). Delineation of genetic zones in the European Norway spruce natural range: Preliminary evidence. *Molecular Ecology*, *9*, 923–934.

Buck, R., Ortega-Del Vecchyo, D., Gehring, C., Michelson, R., Flores-Rentería, D., Klein, B., Whipple, A. V., & Flores-Rentería, L. (2023). Sequential hybridization may have facilitated ecological transitions in the southwestern pinyon pine syngameon. *The New Phytologist*, *237*(6), 2435–2449.

Budde, K. B., Hötzel, S., Müller, M., Samsonidze, N., Papageorgiou, A. C., & Gailing, O. (2023). Bidirectional gene flow between *Fagus sylvatica* L. and F. *Orientalis* Lipsky despite strong genetic divergence. *Forest Ecology and Management*, *537*, 120947.

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

Chen, J., Li, L., Milesi, P., Jansson, G., Berlin, M., Karlsson, B., Aleksic, J., Vendramin, G. G., & Lascoux, M. (2019). Genomic data provide new insights on the demographic history and the extent of recent material transfers in Norway spruce. *Evolutionary Applications*, *12*, 1539–1551.

Chen, N., Cai, Y., Chen, Q., Li, R., Wang, K., Huang, Y., Hu, S., Huang, S., Zhang, H., & Zheng, Z. (2018). Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nature Communications*, *9*, 1–13.

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*, 4820–4838.

Chiarugi, A. (1936a). Cicli forestali postglaciali nell'Appenino Etrusco attraverso l'analisi pollinica di torbe e depositi lacustri presso l'Alpe delle Tre Potenze e il Monte Rondinaio. *Nuovo Giornale Botanico Italiano*, *XLIII*, 6–61.

Chiarugi, A. (1936b). L'indigenato della "Picea excelsa" nell'Appennino Etrusco. *Nuovo Giornale Botanico Italiano*, *43*, 131–166.

Cullingham, C. I., James, P. M. A., Cooke, J. E. K., & Coltman, D. W. (2012). Characterizing the physical and genetic structure of the lodgepole pine × jack pine hybrid zone: Mosaic structure and differential introgression. *Evolutionary Applications*, *5*, 879–891.

Currat, M., Ruedi, M., Petit, R. J., & Excoffier, L. (2008). The hidden side of invasions: Massive introgression by local genes. *Evolution*, *62*, 1908–1920.

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.

De La Torre, A. R., Roberts, D. R., & Aitken, S. N. (2014). Genome-wide admixture and ecological niche modelling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Molecular Ecology*, *23*(8), 2046–2059.

Diaz-Papkovich, A., Anderson-Trocmé, L., Ben-Eghan, C., & Gravel, S. (2019). UMAP reveals cryptic population structure and phenotype heterogeneity in large genomic cohorts. *PLoS Genetics*, *15*, e1008432.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*, 927–930.

Excoffier, L., Marchi, N., Marques, D. A., Matthey-Doret, R., Gouy, A., & Sousa, V. C. (2021). Fastsimcoal2: Demographic inference under complex evolutionary scenarios. *Bioinformatics*, *37*, 4882–4885.

Feng, S., Wan, W., Li, Y., Wang, D., Ren, G., Ma, T., & Ru, D. (2023). Transcriptome-based analyses of adaptive divergence between two closely related spruce species on the Qinghai–Tibet plateau and adjacent regions. *Molecular Ecology*, *32*(2), 476–491.

Ferrarini, E. (1977). Cenosi a Picea abies relitte sull'Appennino. *Annali Accademici Delle Scienze Forestali*, *2*, 185–237.

Fletcher, W. J., Sánchez Goñi, M. F., Allen, J. R. M., Cheddadi, R., Combourieu-Nebout, N., Huntley, B., Lawson, I., Londeix, L., Magri, D., Margari, V., Müller, U. C., Naughton, F., Novenko, E., Roucoux, K., & Tzedakis, P. C. (2010). Millennial-scale variability during the last glacial in vegetation records from Europe. *Quaternary Science Reviews*, *29*, 2839–2864.

Fu, R., Zhu, Y., Liu, Y., Feng, Y., Lu, R. S., Li, Y., Li, P., Kremer, A., Lascoux, M., & Chen, J. (2022). Genome-wide analyses of introgression between two sympatric Asian oak species. *Nature Ecology & Evolution*, *6*, 924–935.

Galinsky, K. J., Bhatia, G., Loh, P. R., Georgiev, S., Mukherjee, S., Patterson, N. J., & Price, A. L. (2016). Fast principal-component analysis reveals convergent evolution of ADH1B in Europe and East Asia. *The American Journal of Human Genetics*, *2016*, 456–472. https://doi.org/10.1016/j.ajhg.2015.12.022

Gaskin, J. F. (2017). The role of hybridization in facilitating tree invasion. *AoB Plants*, *9*(1), plw079.

Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, *201*, 1555–1579.

Giannini, R., Morgante, M., & Vendramin, G. G. (1991). Allozyme variation in Italian population of *Picea abies* (L.) karst. *Silvae Genetica*, *40*, 160–166.

Giannini, R., Petri, L., Rossi, P., & Vendramin, G. G. (1994). Differenziazione e variabilità genetica nel genere *Picea* mediante analisi dei polimorfismi isoenzimatici. *Annali Accademici Delle Scienze Forestali*, *XLIII*, 309–331.

Giesecke, T., & Bennett, K. D. (2004). The Holocene spread of *Picea abies* (L.) karst. In Fennoscandia and adjacent areas. *Journal of Biogeography*, *31*, 1523–1548.

Grant, P. R., & Grant, B. R. (2019). Hybridization increases population variation during adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 23216–23224.

Grant, V. (1981). *Plant speciation*. Columbia University Press.

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*, 205–220.

Hanlon, V. C., Otto, S. P., & Aitken, S. N. (2019). Somatic mutations substantially increase the per-generation mutation rate in the conifer Picea sitchensis. *Evolution Letters*, *3*(4), 348–358.

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.

Jones, M. R., Scott Mills, L., Alves, P. C., Callahan, C. M., Alves, J. M., Lafferty, D. J. R., Jiggins, F. M., Jensen, J. D., Melo-Ferreira, J., & Good, J. M. (2018). Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science*, *360*, 1355–1358.

Karunarathne, P., Zhou, Q., Lascoux, M., & Milesi, P. (2024). Hybridization mediated range expansion and climate change resilience in two conifers. *Global Change Biology*, *30*(4), e17262.

Keim, P., Paige, K. N., Whitham, T. G., & Lark, K. G. (1989). Genetic analysis of an interspecific hybrid swarm of Populus: Occurrence of unidirectional introgression. *Genetics*, *123*, 557–565.

Kobe, F., Leipe, C., Shchetnikov, A. A., Hoelzmann, P., Gliwa, J., Olschewski, P., Goslar, T., Wagner, M., Bezrukova, E. V., & Tarasov, P. E. (2022). Not herbs and forbs alone: Pollen-based evidence for the presence of boreal trees and shrubs in cis-Baikal (eastern Siberia) derived from the last glacial maximum sediment of Lake Ochaul. *Journal of Quaternary Science*, *37*, 868–883.

Krutovskii, K. V., & Bergmann, F. (1995). Introgressive hybridization and phylogenetic relationships between Norway, Picea abies (L.) karst., and siberian, p. obovata ledeb., spruce species studied by isozyme loci. *Heredity*, *74*, 464–480.

Lagercrantz, U., & Ryman, N. (1990). Genetic structure of Norway spruce (*Picea abies*): Concordance of morphological and allozymic variation. *Evolution*, *44*, 38–53.

Larson, D. A., Vargas, O. M., Vicentini, A., & Dick, C. W. (2021). Admixture may be extensive among hyperdominant Amazon rainforest tree species. *New Phytologist*, *232*(6), 2520–2534.

Latałowa, M., & van der Knaap, W. O. (2006). Late quaternary expansion of Norway spruce *Picea abies* (L.) *karst*. In Europe according to pollen data. *Quaternary Science Reviews*, 25, 2780–2805.

Leal, J. L., Milesi, P., Hodková, E., Zhou, Q., James, J., Eklund, D. M., & Lascoux, M. (2024). Complex polyploids: Origins, genomic composition, and role of Introgressed alleles. *Systematic Biology*, 12, 392–418.

Leroy, T., Louvet, J. M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J. M., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *The New Phytologist*, 226, 1171–1182.

Li, L., Milesi, P., Tiret, M., Chen, J., Sendrowski, J., Baison, J., Chen, Z., Zhou, L., Karlsson, B., & Berlin, M. (2022). Teasing apart the joint effect of demography and natural selection in the birth of a contact zone. *The New Phytologist*, 236, 1976–1987.

Liu, B. B., Campbell, C. S., Hong, D. Y., & Wen, J. (2020). Phylogenetic relationships and chloroplast capture in the Amelanchier-Malacomeles-Peraphyllum clade (Maleae, Rosaceae): Evidence from chloroplast genome and nuclear ribosomal DNA data using genome skimming. *Molecular Phylogenetics and Evolution*, 147, 106784.

Liu, Y., Xiao, W., Wang, F., Wang, Y., Dong, Y., Nie, W., & Jia, Z. (2024). Adaptive divergence, historical population dynamics, and simulation of suitable distributions for *Picea meyeri* and *P. Mongolica* at the whole-genome level. *BMC Plant Biology*, 24(1), 479.

Lockwood, J. D., Aleksić, J. M., Zou, J., Wang, J., Liu, J., & Renner, S. S. (2013). A new phylogeny for the genus Picea from plastid, mitochondrial, and nuclear sequences. *Molecular Phylogenetics and Evolution*, 69(3), 717–727.

Magini, E., Pelizzo, A., Proietti Placidi, A. M., & Tonarelli, F. (1980). La Picea dell'Alpe delle Tre Potenze. Areale-Caratteristiche-Posizione sistematica. *Annali Accademici Delle Scienze Forestali*, XXIX, 107–210.

Magri, D., Agrillo, E., Di Rita, F., Furlanetto, G., Pini, R., Ravazzi, C., & Spada, F. (2015). Holocene dynamics of tree taxa populations in Italy. *Review of Palaeobotany and Palynology*, 218, 267–284.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27, 209–220.

Marcus, J., Ha, W., Barber, R. F., & Novembre, J. (2021). Fast and flexible estimation of effective migration surfaces. *eLife*, 2021, 61927. https://doi.org/10.7554/eLife.61927

Milesi, P., Kastally, C., Dauphin, B., Cervantes, S., Bagnoli, F., Budde, K. B., Cavers, S., Ojeda, D. I., Fady, B., & Faivre-Rampant, P. (2023). Synchronous effective population size changes and genetic stability of forest trees through glacial cycles. *bioRxiv*. https://doi.org/10.1101/2023.01.05.522822

Morales-Briones, D. F., Liston, A., & Tank, D. C. (2018). Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the neotropical genus Lachemilla (Rosaceae). *The New Phytologist*, 218(4), 1668–1684.

Morales-Molino, C., Steffen, M., Samartin, S., van Leeuwen, J. F. N., Hürlimann, D., Vescovi, E., & Tinner, W. (2021). Long-term responses of Mediterranean mountain forests to climate change, fire and human activities in the northern Apennines (Italy). *Ecosystems*, 24, 1361–1377.

Morton, N. E. (2013). Isolation by distance. In *Brenner's encyclopedia of genetics* (2nd ed., p. 139). Academic Press.

Nakvasina, E., Demina, N., Prozherina, N., & Demidova, N. (2019). Assessment of phenotypic plasticity of spruce species *Picea abies* (L.) karst. and P. *Obovata* (Ledeb.) on provenances tests in European north of Russia. *Central European Forestry Journal*, 65, 121–128.

Nakvasina, E. N., Volkov, A. G., & Prozherina, N. A. (2017). Provenance experiment with spruce (*Picea abies* (L.) karst. and *Picea obovata* (Ledeb.)) in the north of Russia (Arkhangelsk region). *Folia Forestalia Polonica, Series A – Scimago*, 59, 219–230.

Nei, M., & Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics*, 97, 145–163.

Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *The New Phytologist*, 219, 779–793.

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274.

Nieto Feliner, G., Criado Ruiz, D., Álvarez, I., & Villa-Machío, I. (2023). The puzzle of plant hybridisation: A high propensity to hybridise but few hybrid zones reported. *Heredity*, 131, 307–315. https://doi.org/10.1038/s41437-023-00654-1

Nocchi, G., Wang, J., Yang, L., Ding, J., Gao, Y., Buggs, R. J., & Wang, N. (2023). Genomic signals of local adaptation and hybridization in Asian white birch. *Molecular Ecology*, 32(3), 595–612.

Nota, K., Klaminder, J., Milesi, P., Bindler, R., Nobile, A., van Steijn, T., Bertilsson, S., Svensson, B., Hirota, S. K., & Matsuo, A. (2022). Norway spruce postglacial recolonization of Fennoscandia. *Nature Communications*, 13, 2531.

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hällman, J., Keech, O., Klasson, L., ... Jansson, S. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature*, 497, 579–584.

Orlova, L., Gussarova, G., Glazkova, E., Egorov, A., Potokin, A., & Ivanov, S. (2020). Systematics and distribution of spruce species in the north-west of Russia. *Dendrobiology*, 84, 12–29.

Padgham, M. (2021). geodist: Fast, Dependency-Free Geodesic Distance Calculations. *R Packag. version 0.0.7*.

Parducci, L., Jørgensen, T., Tollefsrud, M. M., Elverland, E., Alm, T., Fontana, S. L., Bennett, K. D., Haile, J., Matetovici, I., & Suyama, Y. (2012). Glacial survival of boreal trees in northern Scandinavia. *Science*, 335, 1083–1086.

Petit, R. J., Aguinagalde, I., de Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., & Mohanty, A. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science*, 300(5625), 1563–1565.

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Vendramin, G. G. (2005). Invited review: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, 14(3), 689–701.

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.

Pfeilsticker, T. R., Jones, R. C., Steane, D. A., Harrison, P. A., Vaillancourt, R. E., & Potts, B. M. (2022). Expansion of the rare Eucalyptus risdonii under climate change through hybridization with a closely related species despite hybrid inferiority. *Annals of Botany*, 129(1), 1–14.

Pfennig, K. S., Kelly, A. L., & Pierce, A. A. (2016). Hybridization as a facilitator of species range expansion. *Proceedings of the Royal Society B: Biological Sciences*, 283, 1648.

Pickrell, J., & Pritchard, J. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *Nature Preceding*, 2, 2012. https://doi.org/10.1038/npre.2012.6956.1

Platt, R. N., McDew-White, M., Le Clec'H, W., Chevalier, F. D., Allan, F., Emery, A. M., Garba, A., Hamidou, A. A., Ame, S. M., & Webster, J. P. (2019). Ancient hybridization and adaptive introgression of an Invadolysin gene in schistosome parasites. *Molecular Biology and Evolution*, 36, 2127–2142.

Popov, P. P. (2010). Form structure and geographic differentiation of spruce populations in northwestern Russia. *Russian Journal of Ecology*, 41, 378–385.

Popov, P. P. (2013). Structure and differentiation of spruce populations in the Komi Republic. *Russian Journal of Ecology*, 44, 193–198.

Pravdin, L. F. (1975). *El' evropeiskaya i el' sibirskaya v SSSR (Norway spruce and Siberian spruce in the Soviet Union)*. Nauka.

Ravazzi, C. (2002). Late quaternary history of spruce in southern Europe. *Review of Palaeobotany and Palynology*, *120*, 131–177.

Rose, J. P., Toledo, C. A., Lemmon, E. M., Lemmon, A. R., & Sytsma, K. J. (2021). Out of sight, out of mind: Widespread nuclear and plastid-nuclear discordance in the flowering plant genus Polemonium (Polemoniaceae) suggests widespread historical gene flow despite limited nuclear signal. *Systematic Biology*, *70*(1), 162–180.

Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., Patterson, N., & Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*, *507*, 354–357.

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 595–619.

Schulte, L., Bernhardt, N., Stoof-Leichsenring, K., Zimmermann, H. H., Pestryakova, L. A., Epp, L. S., & Herzschuh, U. (2021). Hybridization capture of larch (*Larix mill.*) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Molecular Ecology Resources*, *21*, 801–815.

Scotti, I., Gugerli, F., Pastorelli, R., Sebastiani, F., & Vendramin, G. G. (2008). Maternally and paternally inherited molecular markers elucidate population patterns and inferred dispersal processes on a small scale within a subalpine stand of Norway spruce (*Picea abies* [L.] *karst.*). *Forest Ecology and Management*, *255*, 3806–3812.

Semerikov, V. L., Semerikova, S. A., Polezhaeva, M. A., Kosintsev, P. A., & Lascoux, M. (2013). Southern montane populations did not contribute to the recolonization of west Siberian plain by Siberian larch (*Larix sibirica*): A range-wide analysis of cytoplasmic markers. *Molecular Ecology*, *22*, 4958–4971.

Shuvaev, D. N., Semerikov, V. L., Kuznetsova, G. V., & Putintseva, Y. A. (2023). Late quaternary history of Siberian stone pine as revealed by genetic and paleoecological data. *Tree Genetics & Genomes*, *19*, 1–15.

Soltis, D. E., Soltis, P. S., & Milligan, B. G. (1992). Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. In *Molecular Systematics of Plants* (pp. 117–150). Springer.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.

Thórsson, Æ. T., Salmela, E., & Anamthawat-Jónsson, K. (2001). Morphological, cytogenetic, and molecular evidence for introgressive hybridization in birch. *The Journal of Heredity*, *92*, 404–408.

Tollefsrud, M. M., Kissling, R., Gugerli, F., Johnsen, Ø., Skrøppa, T., Cheddadi, R., Van Der Knaap, W. O., Latałowa, M., Terhürne-Berson, R., & Litt, T. (2008). Genetic consequences of glacial survival and postglacial colonization in Norway spruce: Combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology*, *17*, 4134–4150.

Tollefsrud, M. M., Latałowa, M., van der Knaap, W. O., Brochmann, C., & Sperisen, C. (2015). Late quaternary history of north Eurasian Norway spruce (*Picea abies*) and Siberian spruce (*Picea obovata*) inferred from macrofossils, pollen and cytoplasmic DNA variation. *Journal of Biogeography*, *42*, 1431–1442.

Tsitrone, A., Kirkpatrick, M., & Levin, D. A. (2003). A model for chloroplast capture. *Evolution*, *57*, 1776–1782.

Tsuda, Y., Chen, J., Stocks, M., Källman, T., Sønstebø, J. H., Parducci, L., Semerikov, V., Sperisen, C., Politov, D., Ronkainen, T., Väliranta, M., Vendramin, G. G., Tollefsrud, M. M., & Lascoux, M. (2016). The extent and meaning of hybridization and introgression between Siberian spruce (*Picea obovata*) and Norway spruce (*Picea abies*):

Cryptic refugia as stepping stones to the west? *Molecular Ecology*, *25*, 2773–2789.

Väliranta, M., Kaakinen, A., Kuhry, P., Kultti, S., Salonen, J. S., & Seppä, H. (2011). Scattered late-glacial and early Holocene tree populations as dispersal nuclei for forest development in north-eastern European Russia. *Journal of Biogeography*, *38*, 922–932.

Vendramin, G. G., Anzidei, M., Madaghiele, A., Sperisen, C., & Bucci, G. (2000). Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). *Genome*, *43*, 68–78.

Vescovi, E., Ammann, B., Ravazzi, C., & Tinner, W. (2010). A new late-glacial and Holocene record of vegetation and fire history from Lago del Greppo, northern Apennines, Italy. *Vegetation History and Archaeobotany*, *19*, 219–233.

Vidalis, A., Scofield, D., Neves, L., Bernhardsson, C., García-Gil, M. R., & Ingvarsson, P. (2018). Design and evaluation of a large sequence-capture probe set and associated SNPs for diploid and haploid samples of Norway spruce (*Picea abies*). *bioRxiv*. https://doi.org/10.1101/291716

Wang, D., Wang, Z., Kang, X., & Zhang, J. (2019). Genetic analysis of admixture and hybrid patterns of *Populus hopeiensis* and *P. tomentosa*. *Scientific Reports*, *9*(1), 4821.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Whitney, K. D., Randell, R. A., & Rieseberg, L. H. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower Helianthus annuus. *The New Phytologist*, *187*, 230–239.

Wielstra, B., Salvi, D., & Canestrelli, D. (2021). Genetic divergence across glacial refugia despite interglacial gene flow in a crested newt. *Evolutionary Biology*, *48*, 17–26.

Willis, K. J., Rudner, E., & Sümegi, P. (2000). The full-glacial forests of central and southeastern Europe. *Quaternary Research*, *53*, 203–213.

Yang, Y. Y., Qu, X. J., Zhang, R., Stull, G. W., & Yi, T. S. (2021). Plastid phylogenomic analyses of Fagales reveal signatures of conflict and ancient chloroplast capture. *Molecular Phylogenetics and Evolution*, *163*, 107232.

Zemlak, T. S., Habit, E. M., Walde, S. J., Battini, M. A., Adams, E. D. M., & Ruzzante, D. E. (2008). Across the southern Andes on fin: Glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of Galaxias platei in Patagonia. *Molecular Ecology*, *17*, 5049–5061.

## SUPPORTING INFORMATION

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