



RESEARCH PAPER

Journal of
Biogeography

WILEY

Gene flow at the leading range edge: The long-term consequences of isolation in European Beech (*Fagus sylvatica* L. Kuhn)

M. Jennifer Sjölund¹ | Patricia González-Díaz¹ | Jose J. Moreno-Villena¹ |
Alistair S. Jump^{1,2}

¹Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, UK

²CREAF (Centre de Recerca Ecològica i Aplicacions Forestals, Belaterra (Barcelona), Spain

Correspondence

Alistair S. Jump, Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, UK.
Email: a.s.jump@stir.ac.uk

Present addresses

M. Jennifer Sjölund, International Foundation for Science (IFS), Stockholm, Sweden

Patricia González-Díaz, Forest Ecology and Restoration Group, Department of Life Sciences, Universidad de Alcalá, Madrid, Spain

Jose J. Moreno-Villena, Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

Funding information

Natural Environment Research Council, Grant/Award Number: NE/G002118/1

Editor: Jim Provan

Abstract

Aim: Isolation is expected to lead to negative impacts on populations due to a reduction in effective population size and gene flow, exacerbating the effects of genetic drift, which might be stronger in peripheral and fragmented populations. *Fagus sylvatica* (European beech) in southern Sweden presents a gradient of isolation towards the leading range edge of the species. We sought to determine the impact of long-term isolation on genetic diversity and population genetic structure within populations of this species.

Location: Samples were obtained from 14 sites towards the northern edge of the native range of beech in Sweden.

Taxon: Fagaceae.

Methods: Using historical sources, we obtained area- and distance-based measures of isolation. We measured genetic diversity and structure by using nuclear microsatellite marker data, and performed parentage analysis to estimate external pollen-mediated gene flow. We implemented a partial least squares regression to determine the effects of isolation on each of the genetic diversity estimators and the measures of external pollen-mediated gene flow.

Results: Long-term isolation generally had a negative impact on genetic diversity, which is exacerbated over time, further affecting progeny and suggesting that isolated populations are subject to strong genetic drift, possibly due to the combination of founder events and persistent small population sizes. Bayesian cluster analysis revealed that isolation was also acting as a barrier to gene flow in the north-eastern distribution of beech.

Main conclusions: Isolation at the leading range edge of beech in Sweden has created gradients of contemporary gene flow within the species. The long-term cumulative effects of isolation on this wind-pollinated tree species and its negative impacts on genetic diversity and gene flow, could lead to inbreeding depression and higher extinction risk where populations remain small and isolated.

KEYWORDS

colonization, establishment, forest fragmentation, genetic diversity, parentage, range margin

1 | INTRODUCTION

Peripheral populations often display decreasing density as environmental conditions depart from the optimum (Brown, 1984) such that the range edge of a species presents a matrix of increasingly naturally isolated populations the further they exist from the core of the species distribution. Considering that current range shifts are driven by contemporary climate change, peripheral populations are important for species migration, population growth and persistence (Hampe & Petit, 2005; Parmesan & Yohe, 2003; Walther et al., 2002). Highly fragmented migration fronts consisting of isolated populations are shaped by founder effects, that can lead to correlative reductions in allelic number and heterozygosity, decreasing genetic diversity and genetic bottlenecks (Comps, Gömöry, Letouzey, Thiébaud, & Petit, 2001; Eckert, Samis, & Loughheed, 2008; Nei, Maruyama, & Chakraborty, 1975; Young, Boyle, & Brown, 1996).

The consequences of isolation are larger for small populations, which are vulnerable to genetic drift (Ellstrand & Elam, 1993). Persistent reductions in gene flow can also increase the risk of inbreeding depression and genetic differentiation of outlier populations (Eckert et al., 2008; Ellstrand & Elam, 1993; Ouborg, Vergeer, & Mix, 2006), potentially compromising the adaptability of a species and the resilience of populations to environmental change (Jump & Peñuelas, 2005; Willi & Fischer, 2005). Outcrossing trees can be disproportionately sensitive to a reduction in pollen-mediated gene flow owing to their often high levels of heterozygosity that may mask deleterious recessive alleles, which if expressed can lead to a reduction in fitness (Bacles & Jump, 2011).

Widespread forest fragmentation is a consequence of an estimated loss of 13 million hectares of forests per year over the past decade (FRA, 2010). However, the effects of isolation in fragmented forest populations are difficult to measure because trees are long-lived and multiple generations are required to fully realize potential effects on population genetic variation and structure. The inconsistency in results from fragmentation studies has been described as the paradox of forest fragmentation genetics (Kramer, Ison, Ashley, & Howe, 2008). Organisms with long generation times, such as trees, may experience an effective lag to recent fragmentation events as it can take several generations for the impacts of disturbance to change population genetic structure (Aguilar, Quesada, Ashworth, Herrerias-Diego, & Lobo, 2008; Bacles & Jump, 2011; Mona, Ray, Arenas, & Excoffier, 2014). This lag is reflected in the inconsistent empirical evidence existing for the effects of fragmentation on tree populations (Kramer et al., 2008). However, natural range edges have been subject to population isolation longer than recently fragmented forests, making such populations particularly valuable for measuring the effects of isolation in long-lived species.

Some tree species have the potential to buffer the effects of genetic drift in fragmented populations through high gene flow rates via seed and pollen dispersal (Hamrick, 2004; Sork & Smouse, 2006). However, persistent isolation can lead to a reduction in gene flow. Even wind-pollinated species have shown significant effects of isolation and small population size on genetic variability (Jump

& Peñuelas, 2006; Leonardi et al., 2012 (*Fagus sylvatica*), Provan et al., 2008 (*Juniperus communis*), Aizawa et al., 2009 (*Picea jezoensis*), Liepelt et al., 2009 (*Abies alba*), Hensen et al., 2012 (*Polylepis incana*)), with effects on pollen-mediated gene flow found in Wang, Sork, Wu, & Ge, 2010 (*Pinus tabulaeformis*), Vranckx et al., 2014 (*Quercus robur*). In contrast, various studies have found no effect of isolation (Schuster & Mitton, 2000 (*Pinus flexilis*), Muir, Lowe, Fleming, & Vogl, 2004 (*Quercus petraea*), Bacles, Burczyk, Lowe, & Ennos, 2005, Bacles, Lowe, & Ennos, 2006 (*Fraxinus excelsior*), Buschbom, Yanbaev, & Degen, 2011 (*Q. robur*), Ortego, Bonal, Munoz, & Aparicio, 2014 (*Quercus ilex*)).

Given the need to better understand the impact of population isolation, we sought to determine how isolation impacts genetic diversity within populations of the long-lived, wind-pollinated tree species, *Fagus sylvatica* (European beech). We used a gradient of increasingly isolated forest patches found at the northern range edge of *F. sylvatica* in Sweden using a digitized historic map of regional beech distribution to accurately characterize the level of isolation at each study site. We measured contemporary gene flow and pollen-mediated dispersal rates to assess the effect of isolation on the genetic structure of the adult and seedling cohorts of the species. We hypothesized that in the adult cohort, long-term isolation experienced at the range-edge shape the underlying genetic structure of this species in the region, with increasingly isolated sites being highly differentiated due to founder effects and lack of gene flow between sites. Within the seedling cohort, pollen-mediated gene flow is further restricted between isolated sites leading to a further reduction in their genetic diversity. This effect should be stronger in seedlings in more isolated sites. Given the importance of population density for pollen and propagule abundance, we hypothesize that area-based methods for measuring isolation should prove more sensitive than distance-based methods when assessing genetic impacts of population isolation.

2 | MATERIALS AND METHODS

European beech is a predominantly outcrossing, wind-pollinated tree, with seed dispersed by gravity and animals. It generally lives up to 300 years (Packham, Thomas, Atkinson, & Degen, 2012), flowering at 40–60 years of age (Wagner et al., 2010). Seed dispersal has been described at distances less than 25m (Gregorius & Kownatzki, 2005). To our knowledge the pollen dispersal curve of this mast flowering wind-pollinated tree has yet to be defined but data derived from other wind-pollinated species, such as *Quercus macrocarpa*, suggest that effective pollen dispersal can reach up to hundreds of kilometres (Craft & Ashley, 2007). Beech covers a 14 million ha range in Europe (see Figure S2.1 in Appendix S2) that is predominantly climatically limited (Packham et al., 2012; Sjölund, González-Díaz, Moreno-Villena, & Jump, 2017). Locally, it is additionally influenced by anthropogenic impacts, such as the historic reduction of deciduous forests (Bradshaw & Lindbladh, 2005), human disturbance and intensive land use (Björkman, 1996; Sjölund et al., 2017). The distribution of beech in Sweden was extensively



mapped by Lindquist (1931) using aerial reconnaissance techniques data from 1927 to 1930. Lindquist charted stands of pure/beech-dominated forests, mixed stands with beech, and solitary trees. The present-day distribution of beech is substantially unaltered from Lindquist's map and demonstrates a discontinuous migration front with larger more continuous stands in the south of the country and outlying populations further north (Björkman, 1999). The map provided a comprehensive resource to derive both area- and distance-based indices of isolation, which have been shown to differ in their effectiveness as a measure of connectivity (Moilanen & Nieminen, 2002).

2.1 | Sample collection, site selection and study sites

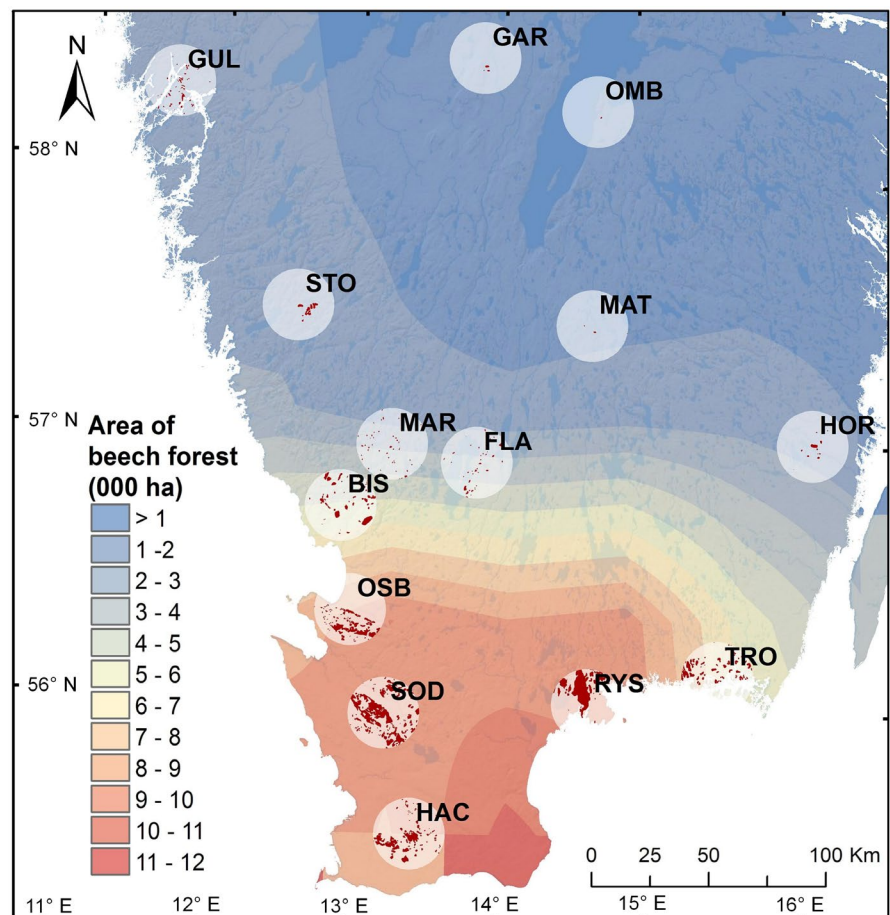
Fourteen sites were sampled based on their historic level of isolation according to Lindquist's (1931) map (Figure 1, see Table S3.1 in Appendix S3). At each plot, we sampled leaf or cambium tissue from 50 adults and 50 seedlings less than 1 year old. Seedlings were sampled towards the centre of the adult plot to improve the probability of capturing the mother tree for parentage analysis. As beech has no persistent seed bank (Packham et al., 2012), seedlings were the result of pollen-mediated gene flow during the previous year (2011), also a mast year.

2.2 | Measuring the multiple dimensions of isolation

Lindquist's (1931) map was geo-referenced in ArcMAP 10 (ESRI) allowing us to measure the approximate area of beech present in the past, closer to the time when the sampled adult trees were fertilized as seeds. We used two measures of isolation; (i) area-based methods and (ii) distance-based methods. To obtain measures of isolation based on area, circular buffer zones with a radius of 5 km, 10 km and 15 km were created around the centre of the sample plot at each site. Within each buffer zone, polygons were created for all beech stands on Lindquist's (1931) map (see Figure S2.2 in Appendix S2 for examples). Single beech trees were counted and given an arbitrary value of 78.54m², a conservative estimate of the circular crown area of a mature beech tree with a radius of 5m, which was added to the area of polygons, giving a total area of beech forest canopy within each buffer zone. This estimate was derived from unpublished field-derived estimates and was necessary to include single tree counts in area-based estimates. The area of contiguous beech forest within the immediate patch where sampling had taken place (*i.e.* the site boundary) was also measured as a proxy of population size (e.g. Jump and Peñuelas (2006)).

To obtain measures of isolation based on distance, we used two measurements commonly used to establish isolation levels (e.g. Jump and Peñuelas (2006) and Leonardi et al. (2012)), which includes the

FIGURE 1 Study sites on an interpolated map of the area of beech in 15-km buffer zones. Labelled sampling sites are surrounded by 15-km inclusive buffer boundaries, shown as light circles, with geo-referenced beech forest from Lindquist's (2002) map in dark red. Overlaid basemap (ESRI) of water bodies indicate positioning of the two largest lakes, Vänern (west) and Vättern (east), near sites GUL, GAR and OMB [Colour figure can be viewed at wileyonlinelibrary.com]



shortest distance from the centre of the sampled forest to the neighbouring forest boundary (abbreviated CB), and the shortest distance from the boundary of the sampled forest to the neighbouring forest boundary (abbreviated BB). In total, we tested four area- and two distance-based measures of isolation; their explanatory capacity was dependant on the heterogeneity in the distribution and structure of the surrounding forest patches.

2.3 | DNA isolation and microsatellite analysis

Genomic DNA was extracted from dried leaf and cambium samples, using the BIOLINE Isolate Plant Kit and the QIAGEN 96 DNeasy Plant Kit according to manufacturer's instructions. Fragment analysis was performed on an ABI 3,730 DNA Analyzer (Applied Biosystems) with scoring on GENE MARKER 2.4.0 (SoftGenetics). Amplification success ranged from 94% to 100% per population. Out of 1,400 samples, a total of 1,376 individuals were successfully genotyped at 12 polymorphic SSRs (fs1-03, fs1-15, fs3-04, fcm5, mfc7, mfs11, sfc0007-2, sfc0018, sfc0036, sfc1143, sfc1061 and sfc1063) in three multiplexes. However, analyses presented exclude fcm5 and use a total of 11 loci as 11 of 14 sites had a significant proportion of null alleles in fcm5 identified in MICRO-CHECKER (Van Oosterhout, Hutchinson, Wills, & Shipley, 2004). A total of nine loci were used for parentage analysis, as two further loci, sfc0018 and fs3-04, were removed from parentage analysis, as possible null alleles close to 10% were detected using CERVUS 3.0.6 (Kalinowski, Taper, & Marshall, 2007) and the accuracy of parentage assignment is particularly sensitive to null alleles. Gametic disequilibrium between locus pairs was tested using FSTAT 2.9.3.2 (Goudet, 1995). Significant associations were tested by random association of genotypes at pairs of loci 1,100 times, using a 5% nominal level after Bonferroni correction. The mean genotyping error rate was 0.5% for the 11 loci, and 0.6% for the 9 loci used in parentage analysis. We calculated error rate per locus as the number of erroneously assigned alleles over 80 repeated samples.

2.4 | Measuring and visualizing genetic diversity and pollen dispersal

We used the inverse distance weight methods available on the spatial analyst interpolation tool on ARCMAP 10 (ESRI) to map multilocus estimates of genetic diversity for the adults and seedlings, and the percentage of seedlings arising from external pollen-mediated gene flow obtained from parentage analysis. Rarefied allelic richness (A_R) and rarefied private allelic richness (A_p) were obtained using ADZE 1.0 (Szpiech, Jakobsson, & Rosenberg, 2008), standardized to a sample size of 47, the smallest number of successfully genotyped individuals per site. Private alleles were defined as those unique to a single site within either the adult or seedling cohort. Estimates of gene diversity (corrected for sample size) (H_S) and the inbreeding coefficient (F_{IS}) were obtained using SPAGED1 1.4b (Hardy & Vekemans, 2002). The difference between adult and seedling cohorts for each diversity estimator was tested using a Mann-Whitney U -test.

Parentage analysis was performed within each site to quantify the proportion of seedlings arising from trees existing outside the

sampled plot using maximum-likelihood based methods in CERVUS (Kalinowski et al., 2007). We have previously tested three methods for parentage analysis: (a) parent pair analysis, (b) paternity analysis and (c) counts of foreign alleles, to obtain an estimate of pollen flow and present the most sensitive method, parent pair analysis (for further details see Sjölund, 2014). A combined exclusion probability of >99.99% for parent pairs was obtained for nine loci. We followed a method used by Buschbom et al. (2011) to estimate external pollen-mediated gene flow. We used a two-step procedure within each site to assign parentage. In the first step, seedlings were primarily assigned a maternal tree based on LOD score at 90% confidence level (CI). LOD scores are obtained from the natural log of the overall likelihood ratio, which is calculated for candidate parents by dividing the likelihood that they are the true parent by the likelihood that they are not the true parent, with larger ratios indicating that the candidate parent is likely to be the true parent. Individuals with identified parents (single or parent pair) at 90% CI were assumed to have a maternal tree present, whereas seedlings with no identified parents at 90% CI were considered as not having a maternal tree present in the adult cohort and therefore excluded from further analysis. For the purpose of maternal tree assignment, seedlings were assumed to be primarily dispersed by gravity (Wagner et al., 2010), which is likely as the seedlings originated from a mast year, known to satiate primary predators, including animals involved in seed dispersal. In the second step, we estimated seedlings originating from external pollen dispersal from parent pair analysis (Figure 2). To avoid potential bias generated by comparing populations of different sizes when estimating pollen immigration, we standardized seedling sample size to 15 per population (i.e. the lowest number of maternal trees assigned for any given site using subsamples by random selection).

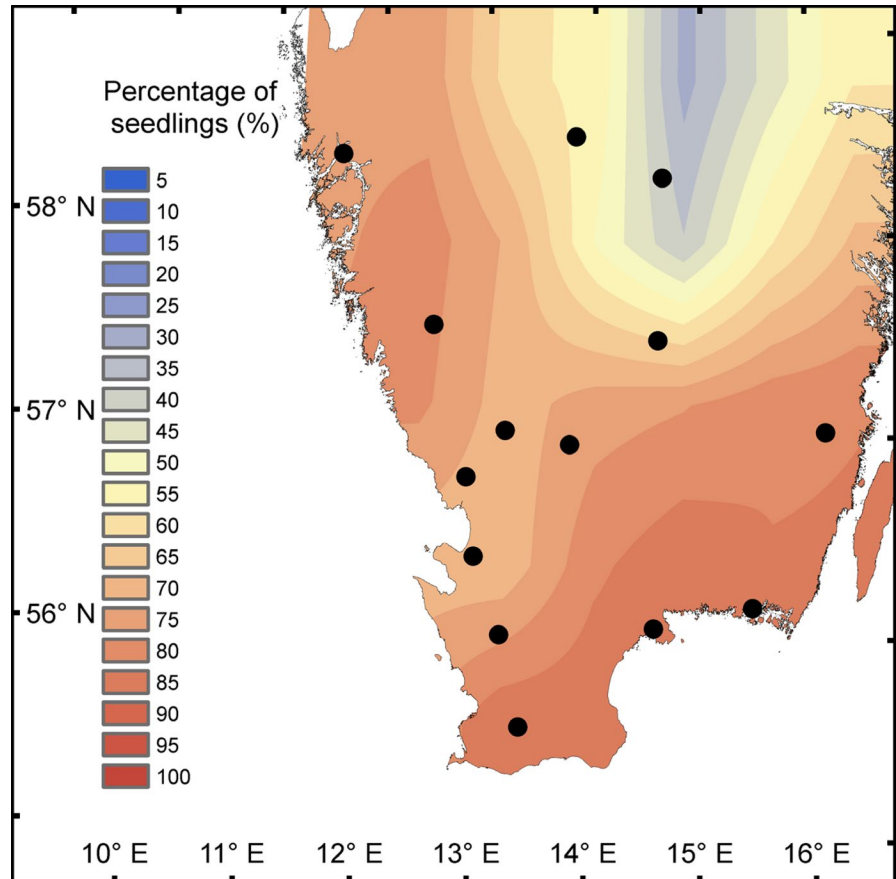
Critical LOD scores and Delta (the difference in LOD scores) at 90% confidence between the most likely candidate parent and the second most likely candidate parent were obtained by simulating 100,000 offspring according to the following settings: all adult trees were included as candidate parents and represented 100% of potential parents, the proportion of loci typed varied from 0.998 to 1, the proportion of loci mistyped was set to 1%, as the minimum recommended by the program.

2.5 | Modelling the effects of isolation on genetic diversity and pollen-mediated gene flow

We performed a partial least squares regression (PLSR) using the R package PLSDEPOT 0.1.17 to determine the effects of isolation on each of the genetic diversity estimators and the measures of external pollen-mediated gene flow. PLSR is relatively robust to small sample sizes compared to multiple regression and has the added benefit of allowing easy visualization of data consisting of correlated predictor variables (Carrascal, Galva, & Gordo, 2009). PLSR deals with the lack of independence among predictor variables by grouping them into one or more orthogonal, linear gradients of covariation while maximizing the explained variance in the response variable (Palomino & Carrascal, 2007).



FIGURE 2 Percentage of seedlings resulting from external pollen dispersal. Interpolated maps of parentage analysis using nine loci depicting the percentage of seedlings (%) with no local parent pair at 95% CI (No LPP 95% CI) [Colour figure can be viewed at wileyonlinelibrary.com]



We tested three variables describing genetic diversity in adults and seedlings: allelic richness (A_R); private allelic richness (A_P) and gene diversity (H_S). We did not perform tests on F_{IS} as patterns are confounded by the significance of the F_{IS} values. For measures of pollen-mediated gene flow, we tested the number of seedlings with no local parent pairs (LPP) at 95% CI, that is, seedlings that could not be assigned to two local adult trees. The predictor variables used were the six measures of isolation derived from Lindquist's (1931) map, comprising the four area-based (5-km buffer, 10-km buffer, 15-km buffer, site boundary) and two distance-based (CB, BB) measurements. Latitude and longitude were also included as predictors to account for geographical variation. Predictor variables were $\log(x + 1)$ transformed to standardize the scale of predictors. In models with two significant components, the secondary components were found to be redundant, revealing similar trends as the main component and are therefore not presented. We present weights of significant predictor variables, which indicate the trend and the importance of the relationship with the component.

We used the program BOTTLENECK 1.2.02 (Cornuet & Luikart, 1996) to test for recent bottlenecks in the adult and seedling cohort using 11 loci. Populations were tested under the Two-phase model (TPM), allowing 95% single-step mutations and 5% multi-step mutations with a variance of 12 among multi-step mutations under 1,000 simulation iterations, as recommended for microsatellites by Piry & Luikart (1999). Significance tests for $H_S > H_{eq}$ (where, H_S is gene diversity and H_{eq} is the heterozygosity expected under

mutation-drift equilibrium) were performed using Wilcoxon's test implemented in the program. A recent reduction in population effective size that causes bottlenecks, can be detected as a reduction in allelic richness and gene diversity H_S , where allelic richness is reduced faster than H_S leading to a larger heterozygosity than expected under mutation-drift equilibrium.

2.6 | Identifying regional population structure at the leading edge

To characterize regional genepools and identify potential barriers to dispersal, individual-based assignment methods were performed on the adult cohort using GENELAND 4.0.4 (Guillot, Estoup, Mortier, & Cosson, 2005) a spatially explicit Bayesian clustering model. Geneland results were validated using STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000), as recommended by Guillot, Lebloit, Coulon, and A.C., F. (2009). We followed standard approaches when implementing both programmes, as detailed in Appendix S1.

3 | RESULTS

3.1 | Isolation indices

The level of isolation obtained by area-based measurements ranged from 32.94 ha to 8,396.37 ha of surrounding forest. Distance measures of isolation ranged from 0.065 km to 50.429 km (see Table S3.1

in Appendix S3). Mapping the total area of beech within a 15-km buffer zone around each site revealed a north-easterly trend of increasing isolation (Figure 1), reflecting the reduction of population density visible in Linquist's (1931) map.

3.2 | Estimates of genetic diversity and external pollen-mediated gene flow

The 11 loci used in this study were found to be in gametic equilibrium. The maximum number of alleles for all samples had a multilocus average of 13.36, ranging from 4 to 28 alleles per locus. Rarefied allelic richness (A_R) in adults ranged from 5.06 to 7.17, and 4.50 to 6.17 in seedlings, with adults having a significantly higher level of allelic richness compared to seedlings ($U(12) = 147, Z = 2.25, p < .05$) (Table 1, Figure 3). Adult rarefied private allelic richness (A_p) ranged from 0.02 ± 0.02 to 0.39 ± 0.22 , while in seedlings A_p ranged from 0.00 to 0.53 ± 0.36 . There were no significant differences between levels of private allelic richness in adults and seedlings ($U(12) = 87, Z = -0.51, p = .63$), although the highest levels of A_p were consistently found in the most southerly site (HAC) in both adults and seedlings (Figure 3). Gene diversity (H_S) estimates ranged from 0.601 to 0.703 in adults, and 0.534 to 0.688 in seedlings, with no significant differences found between adults and seedlings ($U(12) = 126, Z = 1.28, p = .21$). No evidence of homozygote excess was found in either cohort. However, significantly negative inbreeding coefficients (F_{IS}) indicated a heterozygote excess in one site for the adult cohort and in six sites for the seedling cohort (Table 1, Figure 3), with significantly lower values of F_{IS} in seedlings ($U(12) = 154, Z = 2.57, p < .01$). Parent pair analysis identified a range of 13.3% to 86.7% of seedlings with no local parent pairs at 95% CI.

3.3 | PLSR models of genetic diversity and external pollen-mediated gene flow

Significant PLSR components were found for rarefied allelic richness (A_R) and gene diversity (H_S) in adults and seedlings, and for measures of external pollen-mediated gene flow (Table 2). No significant relationships between components and response variables were found for private allelic richness (A_p) in adults or seedlings. It should be noted that significant correlations between buffer-based isolation measures and latitude were found (Pearson's r and significances, 5 km: $r = -.713, p < .01$; 10 km: $r = -.732, p < .01$; 15 km: $r = -.733, p < .01$; Site boundary: $r = -.520, p = .06$; CB: $r = .373, p = .19$; BB: $r = .424, p = .130$). No correlations were found between isolation indices and longitude ($p > .40$ for all indices).

The number of seedlings with no local parent pairs at 95% CI increased in sites with lower distances between forests and higher areas of surrounding beech, following a north to south gradient ($R^2 = 58.2\%, p < .01$). This pattern is visible in Figure 2. The contribution of significant isolation variables and geographical variables were similar, with the distance-based variable, CB, explaining 10.2% ($p < .05$) of the variation within the response, the area of beech within the 5-km buffer zone explaining 10.1% ($p < .001$), and latitude

explaining 6.5% ($p < .001$). Overall, isolation indices and latitude explained more variation in pollen-mediated gene flow, with an R^2 ranging from 5.7% to 10.2%, compared to the adult and seedling measured of genetic diversity (Table 2). Measures that were taken closest to the sampling area, for example, 5-km buffer and CB, explained the largest amount of variation 10.1% and 10.2%, respectively, reflecting the sensitivity of measures of pollen-mediated gene flow to local measures of isolation.

There was no evidence for recent genetic bottlenecks in any of the 14 sites in the adult or seedling cohort (see Table S3.2 in Appendix S3). For genetic diversity estimators, the strongest relationship between predictor variables and response was found in seedling A_R ($R^2 = 58.3\%, p < .05$), which was significantly negatively related to isolation. The remaining genetic diversity response variables in order of the original variance explained were seedling H_S ($R^2 = 33.8\%, p < .05$), seedling A_R ($R^2 = 30.1\%, p < .05$) and adult H_S ($R^2 = 26.2\%, p < .05$).

An increase in A_R in adults was associated with southern sites, with a high area of surrounding beech forest (Table 2, Figure 3). Latitude explained the largest variation in adult A_R ($R^2 = 7.3\%, p < .001$), with site boundary explaining the most variation out of the area-based measurements ($R^2 = 5.8\%, p < .001$). There was an unexpected significant, positive relationship of increasing distance and increasing adult A_R . However, this explained $<0.1\%$ of R^2 ($p < .05$) and therefore was not considered biologically relevant. Seedling A_R revealed a similar trend to that found in adults, although isolation was the primary driver of variation in the response, instead of latitude. Increased A_R in seedlings was associated with a high area of surrounding forests and low distances between forests in southern latitudes. The area-based measure of beech in the 15-km buffer zone was the primary contributor to variation in the response ($R^2 = 8.3\%, p < .001$) with a comparable amount of variance explained by the distance measure, boundary to boundary (BB) ($R^2 = 8.2\%, p < .001$). Latitude only explained 0.4% of the total variation in the response ($p < .001$).

Adult H_S was primarily related to increased area of surrounding beech, specifically associated with the 15-km buffer zone ($R^2 = 14.9\%, p < .001$). As with adult allelic richness, southern sites also displayed higher levels of H_S . However, in the maps of genetic diversity (Figure 3), the trend for adult H_S was not as clear as that displayed by adult A_R . Seedling H_S was the only response significantly influenced by longitude, with H_S increasing on a west to east gradient, explaining the largest amount of variation in the response ($R^2 = 14.7\%, p < .001$) (Table 2, Figure 3). The relationship with isolation was contradictory to that found in the other genetic diversity response variables, as higher levels of seedling H_S were associated with a decrease in surrounding beech area (10 km: $R^2 = 4.3\%, p < .01$) and lower distances between forests (CB: $R^2 = 2.7\%, p < .05$).

When considering all models for both genetic diversity and external pollen-mediated gene flow, area-based measurements significantly contributed to the explanation of the response for all presented response variables, whereas distance-based measures failed to explain significant variation in the response in one model, adult H_S . Concerning area-based measures, significant contributions

TABLE 1 Multilocus estimates of genetic diversity and differentiation

Site code	Adults (N)	Seedlings (N)	A_R		A_P		H_S		F_{IS}	
			Adult	Seedling	Adult	Seedling	Adult	Seedling	Adult	Seedling
SOD	49	50	6.38 ± 0.58	6.11 ± 0.48	0.08 ± 0.04	0.20 ± 0.13	0.620	0.611	0.016	-0.054*
RYS	49	50	7.07 ± 0.68	6.16 ± 0.63	0.27 ± 0.14	0.12 ± 0.06	0.678	0.673	0.018	-0.016
HAC	50	49	6.86 ± 0.73	6.11 ± 0.55	0.39 ± 0.22	0.53 ± 0.36	0.681	0.666	-0.029	-0.083***
OSB	50	50	6.30 ± 0.62	5.77 ± 0.57	0.13 ± 0.05	0.14 ± 0.08	0.667	0.659	-0.028	-0.032
TRO	50	47	5.85 ± 0.45	4.98 ± 0.36	0.03 ± 0.02	0.00 ± 0.00	0.679	0.643	0.002	-0.088***
BIS	49	50	6.52 ± 0.61	5.82 ± 0.48	0.05 ± 0.03	0.12 ± 0.06	0.646	0.636	-0.007	-0.003
FLA	50	50	5.89 ± 0.61	5.76 ± 0.43	0.02 ± 0.02	0.19 ± 0.12	0.677	0.683	-0.010	-0.044
GUL	49	48	6.07 ± 0.62	5.34 ± 0.56	0.12 ± 0.06	0.15 ± 0.09	0.667	0.603	0.027	-0.053*
STO	47	47	5.61 ± 0.46	5.16 ± 0.46	0.20 ± 0.10	0.13 ± 0.06	0.601	0.534	-0.011	-0.043
MAR	49	50	6.53 ± 0.69	6.17 ± 0.67	0.21 ± 0.10	0.40 ± 0.12	0.699	0.687	0.000	-0.043
HOR	50	49	5.74 ± 0.57	5.55 ± 0.60	0.19 ± 0.12	0.36 ± 0.20	0.639	0.648	-0.047	-0.006
GAR	47	50	5.06 ± 0.38	4.50 ± 0.34	0.13 ± 0.09	0.18 ± 0.12	0.644	0.627	-0.104***	-0.061*
MAT	49	48	5.27 ± 0.57	4.91 ± 0.57	0.11 ± 0.08	0.08 ± 0.08	0.685	0.645	-0.011	-0.001
OMB	50	50	7.17 ± 0.66	5.76 ± 0.48	0.30 ± 0.14	0.11 ± 0.05	0.703	0.688	0.002	-0.066**
		Mean ± SE	6.17 ± 0.17*	5.58 ± 0.14*	0.16 ± 0.03	0.19 ± 0.04	0.663 ± 0.008	0.648 ± 0.011	-0.013 ± 0.009***	-0.042 ± 0.008***

Note: Terms for genetic diversity estimators are as follows: A_R , allelic richness (Petit, El Mousadik, & Pons, 1998); A_P , private allelic richness (Szpiech et al., 2008); H_S , gene diversity corrected for sample size (Nei, 1978); F_{IS} , inbreeding coefficient (Weir & Cockerham, 1984). The minimum number of gene copies (k) used for rarefaction analysis of A_R and A_P is 94 with standard errors of the rarefaction procedure provided. p -values for F_{IS} are obtained after 10,000 permutations of gene copies within adult or seedling individuals of each site. p -values for mean genetic diversity estimates indicate significant differences between adults and seedlings. Significant two-sided p -values are indicated as * $p < .05$, ** $p < .01$ and *** $p < .001$. The sites are ordered in terms of total area of beech within the 15 km inclusive buffer zone. Multilocus estimates of genetic diversity were attained from 11 loci.

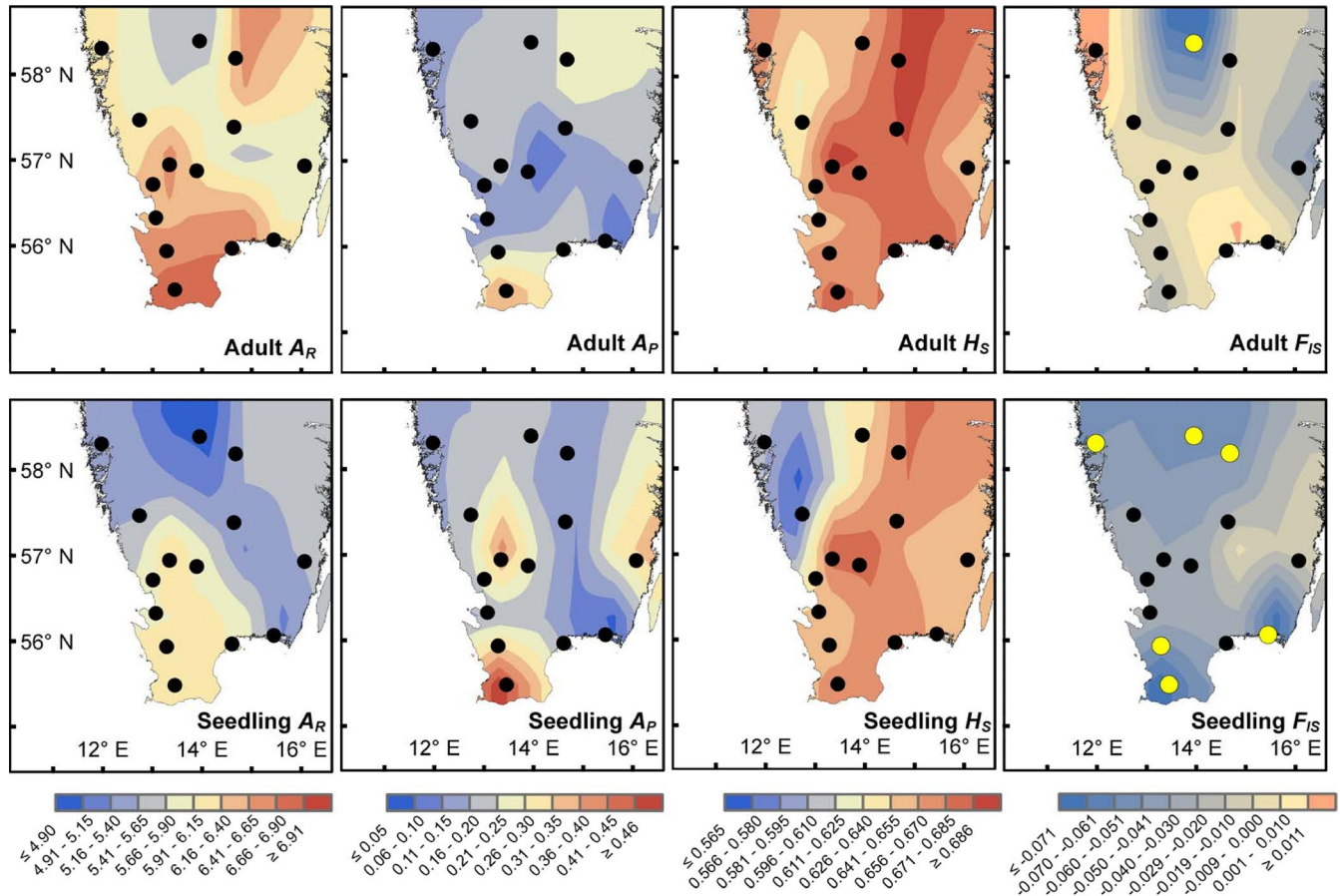


FIGURE 3 Distribution of genetic diversity in adults and seedlings. Multilocus genetic diversity estimators using 11 loci are abbreviated as follows: rarefied allelic richness (A_R); rarefied private allelic richness (A_P); gene diversity corrected for sample size (H_S) and the inbreeding coefficient (F_{IS}). p -values for F_{IS} are obtained after 10,000 permutations of gene copies within adult or seedling individuals of each site. Significant two-sided p -values less than .05 are indicated by yellow circle [Colour figure can be viewed at wileyonlinelibrary.com]

were made by the addition of each buffer zone in all response variables except for H_S , which in adults only related to the 15-km buffer, and in seedlings, the 10-km buffer. The boundary-based distance measure, BB, which was significant in most models, did not significantly explain variation in adult or seedling H_S .

3.4 | Regional genetic structure at the leading range edge

Three clusters were identified using individual-based assignment methods in 8 of 10 runs using the uncorrelated model, with consistent results in 9 of 10 runs with the subsequent correlated model. The spatially explicit models in GENELAND presented the highest average posterior probabilities for the clusters that extended over three regions: (a) the west; (b) the south-east and (c) the north-east (Figure 4; see Figure S2.3 in Appendix S2 for maps of cluster posterior probabilities), reflecting the north-easterly gradient of increasing isolation (Figure 1). Further substructuring was found in the south-eastern and north-eastern cluster 2 and 3, when analysed separately, with each of the four populations clustered individually with relatively low admixture levels within each (Figures 4 and 5; see Figure S2.4 in

Appendix S2 for maps of cluster posterior probabilities). No further substructuring was found in the western cluster 1. All inferred clusters using 13 sites were found to be significantly differentiated and in gametic equilibrium, except for site OMB, which showed significant disequilibrium at one pair of loci (data not shown).

Analysis with STRUCTURE was in agreement with GENELAND and indicated a presence of three clusters in the data (see Figure S2.5 in Appendix S2 for the log probability of the data and ΔK). The levels of admixture were higher according to STRUCTURE as opposed to GENELAND, with a trend of decreasing admixture as isolation (i.e. area of beech forest in the 15-km buffer zone) increased (Figure 5).

4 | DISCUSSION

4.1 | Founder events and isolation shape genetic diversity

We found evidence of reduced allelic richness (A_R) in both adults and seedlings in isolated sites, with latitude also significantly explaining a large proportion of variation in adults (Table 2). Adult gene diversity (H_S) revealed a similar trend to that found in allelic

**TABLE 2** Significant PLSR models with predictor weights and their contributions to R^2

	Adult A_R	Adult H_S	Seedling A_R	Seedling H_S	No LPP 95% CI
Response R^2	30.1*	26.2*	58.3*	33.8*	58.2**
R^2 contributions					
5 km	4.5		6.9		10.1
10 km	0.4		2.3	4.3	8.7
15 km	4.6	14.9	8.3		7.9
Site boundary	5.8		5.7		5.7
CB				2.7	10.2
BB	<0.1		8.2		9.0
Latitude	7.3	7.0	0.4		6.5
Longitude				14.7	
Predictor weights					
5 km	0.386***		0.450***		0.417***
10 km	0.113*		0.262**	-0.358**	0.387***
15 km	0.393**	0.755**	0.495***		0.367***
Site boundary	0.439***		0.410***		0.313**
CB				0.282*	-0.419*
BB	0.004*		-0.162***		-0.394***
Latitude	-0.492***	-0.518**	-0.511***		-0.334***
Longitude				0.660***	

Note: Only response variables which had significant relationships to components were included in the table. All predictor variables include four area-based measures (m^2), the 5-km, 10-km and 15-km buffer zones, and site boundary; two distance-based measures (m), the centre to boundary (CB), and boundary to boundary (BB); and two geographical measures, latitude and longitude. Predictor weights and their contribution to R^2 are given for those significantly related to the component. Significant p -values are indicated as, * $p < .05$, ** $p < .01$ and *** $p < .001$.

richness. The observed pattern in genetic diversity also resembled the *southern richness and northern purity* paradigm, coined by Hewitt (1999), and was based on the observed reduction in population size through founder events at the leading edge, resulting in a loss of alleles through genetic drift (Excoffier, Foll, & Petit, 2009; Lande, 1988; Nei et al., 1975). Low migration between isolated populations contributes to this trend (Vucetich & Waite, 2003). As rare alleles are at risk of disappearing first, a reduction in population size can affect A_R disproportionately more than H_S (Comps et al., 2001; Jump & Peñuelas, 2006; Piry & Luikart, 1999). This pattern was reflected in both adults and seedlings, which displayed a higher amount of variation explained by predictor variables for A_R compared to H_S . Additionally, the most southerly site, HAC displayed the highest level of rarefied private allelic richness (A_p) in adults and seedlings (Figure 3), although PLSR models for A_p were not significant.

The interacting effects of latitude and isolation on A_R in adults suggest that marginal, isolated populations are subject to strong genetic drift, possibly due to the combination of founder events and persistent small population sizes that lead to the loss of alleles over time. The relatively weaker effect of isolation on A_R in adults, compared to latitude, may also be influenced by an outlier, site OMB, which displayed relatively high levels of allelic richness (Figure 3). Although OMB is not the most northerly site, it is the most isolated in terms of surrounding area and the distance to the nearest neighbouring forest (see Table S3.1 in Appendix S3), influencing the trend

found with isolation indices. Seedling A_R was strongly affected by isolation, with very little contribution of latitude. This stronger negative impact of isolation on seedlings compared to adults (Table 1, Figure 3) reflects temporal effects of genetic drift on small populations, which intensify over time (Aguilar et al., 2008; Ellstrand & Elam, 1993). This pattern is in agreement with the observed adult to seedling relative reduction of A_R in site OMB, compared to remaining sites (Figure 3). As the life span of beech is between 150 and 300 years, the oldest populations, having established approximately at 3,000 BP (Bradshaw & Lindbladh, 2005), have experienced at least 10 to 20 generations. Even without considering overlapping generations, this is in agreement with the values obtained from a simulation study on habitat fragmentation by Mona et al. (2014) who recommend at least 10 generations, ideally more than 100 generations needed for isolation to affect genetic diversity at the leading range edge. Overall, these patterns are indicative of a relatively strong effect of isolation on genetic diversity in these range-edge populations.

4.2 | Isolation shapes regional genetic structure

Isolation appears to be acting as a barrier to gene flow at the range edge in Sweden, indicated by the reduction in external pollen-mediated gene flow with increased isolation (Figure 2). Although we found isolation-by-distance in the adult cohort, it appears that paired sites

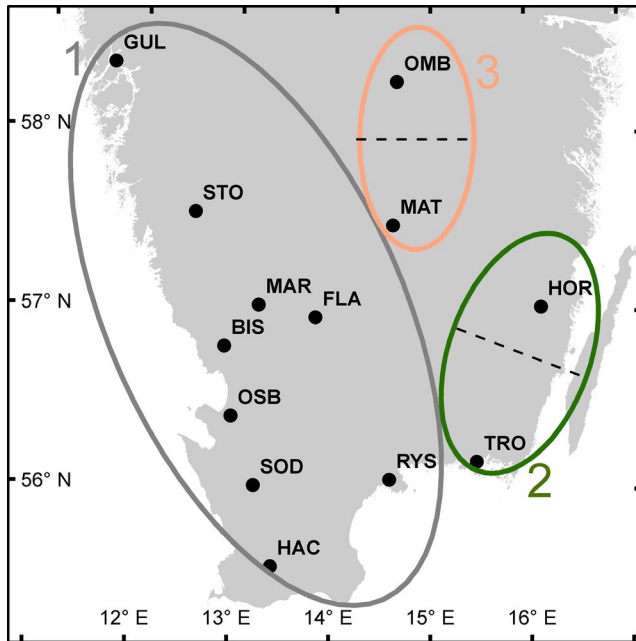


FIGURE 4 Inference of genetic clusters in the adult cohort over 13 sites. Sites are displayed as small black circles, with large open ellipses indicating the grouping of the three inferred population clusters: the western cluster 1 (grey); the south-eastern cluster 2 (green) and the north-eastern cluster 3 (orange). The dotted lines indicate further substructuring found in further analysis of the subset of sites in cluster 2 and 3. A total of 11 loci were used in analysis [Colour figure can be viewed at wileyonlinelibrary.com]

for each geographical distance displayed a greater genetic distance between them, if sites originated from different, instead of the same cluster, suggesting the existence of barriers to gene flow (Fontaine et al., 2007; McRae, Beier, Dewald, Huynh, & Keim, 2005; Rosenberg et al., 2005) (see Figure S2.6 in Appendix S2 for plot of geographical and genetic distance within and between clusters). In terms of geographical barriers, the southern area of Sweden sampled in this study is quite flat with the highest peak, Tomtabacken, being 361m in elevation. The sampling area does enclose Vättern, Sweden's second largest lake, which may have been a barrier to gene flow between GAR and OMB northerly sites that lie on either side of the lake. Clustering results using STRUCTURE revealed a trend of increased admixture in south-westerly sites with increased homogeneity among individuals between north-eastern clusters (Figure 5). A similar trend has been found in leading-edge populations of *Acer campestre* in Poland, which also display further genetic structuring and less admixture with latitude (Chybicki, Waldon-rudziołek, & Meyza,

2014). The large western cluster 1 identified by GENELAND in the adult cohort (Figure 4), contains sites displaying a range of isolation levels. Although our markers are unable to confirm phylogeographical patterns, the pattern is in agreement with palynological evidence on the initial colonization of beech in Sweden around 3,000 BP (Bradshaw & Lindbladh, 2005). Further population substructure observed for clusters 2 and 3 (Figures 4 and 5) is likely to have arisen from genetic drift in marginal isolated populations (Excoffier et al., 2009).

4.3 | Pollen dispersal and its relation to forest patch density

The directional relationships observed for pollen dispersal with isolation and latitude were consistent between models, although the amount of variation explained varied. As our study plots exist within forest fragments, measures of external-mediated pollen dispersal reflect the density of the pollen cloud produced by surrounding forest as opposed to strict long-distance pollen flow. As pollen production is related to the number of reproductive trees, it is likely that pollen production is higher in continuous populations compared to fragmented isolated populations. Therefore, the probability of fertilization by pollen grains from external trees decreases along with isolation. This pattern is in agreement with previous studies that have found an increase in pollen dispersal with tree density (Vranckx et al., 2014; Wang et al., 2010). Although the isolation indices used in our study do not incorporate density in terms of tree numbers, the buffer measures do incorporate the density of forest fragments. The reduction in allelic richness between adult and seedling cohorts suggests that the pollen donor diversity is not large enough to safeguard against the effects of genetic drift under small population sizes (Table 1, Figure 3). As seedlings were the result of pollen dispersal during a mast year, we would expect a stronger negative impact of isolation on pollen dispersal during non-mast years when flowering and pollination success are significantly lower (Hilton & Packham, 1997; Linquist, 1931; Nilsson & Wästljung, 1987).

4.4 | Effective measurements of isolation

A combination of isolation variables was effective at explaining variation in genetic diversity and pollen dispersal. Area-based measurements using buffer zones present a standardized measure for the surrounding area of forest. We found a general additive effect of increasing buffer zone size, implying a sensitivity to buffer size and also the importance of local and regional isolation levels on

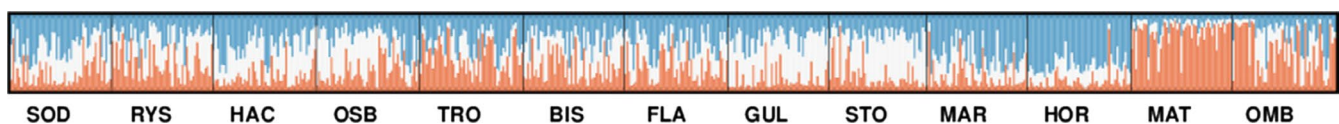


FIGURE 5 Individual assignment to each cluster using STRUCTURE. The Q-matrix presents the average assignment probabilities over 10 consecutive runs for $K = 3$. Site codes are indicated below, and are ordered by increasing isolation from left to right obtained from estimates for the area of beech (ha) in the 15-km inclusive buffer zone. A total of 11 loci were used in analysis [Colour figure can be viewed at wileyonlinelibrary.com]



genetic diversity and pollen dispersal of beech. In a meta-analysis by Moilanen and Nieminen (2002), buffer-based measurements were found to be superior to distance-based measurements when quantifying isolation. Our findings emphasize the utility of such area-based measurements to provide a relatively easy and quick way of measuring isolation that, in our study, outperforms commonly used site boundary and distance measures.

5 | CONCLUSION

The gradient of isolation at the leading range edge of beech in Sweden has shaped contemporary gene flow in adult and seedling cohorts of the species. Isolation has a negative impact on allelic richness, which is exacerbated over time, further affecting progeny. In adult populations, gene diversity follows a similar trend to allelic richness. North-eastern populations appear to be mainly shaped by barriers to gene flow imposed by isolation, with the possibility of some barriers from geographical features in the north. This study highlights the long-term cumulative effects of isolation on a wind-pollinated tree species and its negative impacts on genetic diversity and gene flow, which can lead to inbreeding depression and higher extinction risk and could shed light on the potential effects fragmented populations may experience in the future. Given that much European forest now persists in highly human-modified landscapes, opportunities for migration and forest expansion in response to the warming climate are strongly limited. Our study highlights the need to consider climate smart forest management strategies to aid range shifts in species and reduce the negative impacts on genetic diversity experienced by isolated populations at the leading edge of a species distribution. Further research into the genetic condition of marginal populations can help us understand the consequences for the future expansion and persistence of forests under contemporary climate-driven range shifts.

ACKNOWLEDGEMENTS

We thank landowners and Skoggstyrelsen for their support. We also thank E. Herridge, G. Flint, J. Brunet, J. McArthur, L. Carrascal, P. Ruiz-Benito, T. Bimson and Ö. Fritz for their help and advice on the project. We also thank the anonymous referees whose comments have led to significant improvements of the paper. This work was funded by the Natural Environment Research Council as part of the ERA-Net BiodivERsA Project 'European Beech Forests for the Future' (BEFOFU) [Grant NE/G002118/1]. PGD was supported by the grants FUNDIVER, CGL2015-69186-C2-2-R and REMEDINAL, TE-CM S2018/EMT-4338

AUTHOR CONTRIBUTIONS

M. J. S. and A. S. J. designed the research. M. J. S. conducted field-based work. M. J. S., J. J. M.V. and P. G. D. conducted lab-based work. M. J. S. and P. G. D. conducted data analysis. A. S. J. supervised the

research project. M. J. S., A. S. J. and P. G. D. wrote the manuscript with the contribution of J. J. M.V.

DATA AVAILABILITY STATEMENT

All microsatellite data, geographical coordinates for adult and seedling specimens are available from DRYAD <https://doi.org/10.5061/dryad.8q51152> (Sjölund, González-Díaz, Moreno-Villena, & Jump, 2019).

Title: Data from: Gene flow at the leading range edge: the long-term consequences of isolation in European Beech (*Fagus sylvatica* L. Kuhn)

DOI: doi:10.5061/dryad.8q51152

Journal: Journal of Biogeography

Journal manuscript number: JBI-18-0316

ORCID

M. Jennifer Sjölund  <https://orcid.org/0000-0001-9599-3516>

Patricia González-Díaz  <https://orcid.org/0000-0003-3664-2913>

Alistair S. Jump  <https://orcid.org/0000-0002-2167-6451>

REFERENCES

- Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, *17*, 5177–5188. <https://doi.org/10.1111/j.1365-294X.2008.03971.x>
- Aizawa, M., Yoshimaru, H., Saito, H., Katsuki, T., Kawahara, T., Kitamura, K., ... Kaji, M. (2009). Range-wide genetic structure in a north-east Asian spruce (*Picea jezoensis*) determined using nuclear microsatellite markers. *Journal of Biogeography*, *36*, 996–1007. <https://doi.org/10.1111/j.1365-2699.2008.02074.x>
- Bacles, C. F. E., Burczyk, J., Lowe, A. J., & Ennos, R. A. (2005). Historical and Contemporary Mating Patterns in Remnant Populations of the Forest Tree *Fraxinus excelsior* L. *Evolution*, *59*, 979–990. <https://doi.org/10.1111/j.0014-3820.2005.tb01037.x>
- Bacles, C. F. E., & Jump, A. S. (2011). Taking a tree's perspective on forest fragmentation genetics. *Trends in Plant Science*, *16*, 13–18. <https://doi.org/10.1016/j.tplants.2010.10.002>
- Bacles, C. F. E., Lowe, A. J., & Ennos, R. A. (2006). Effective seed dispersal across a fragmented landscape. *Science*, *311*, 628. <https://doi.org/10.1126/science.1121543>
- Björkman, L. (1996). Long-term population dynamics of *Fagus sylvatica* at the northern limits of its distribution in southern Sweden: A palaeoecological study. *The Holocene*, *6*, 225–234. <https://doi.org/10.1177/095968369600600208>
- Björkman, L. (1999). The establishment of *Fagus sylvatica* at the stand-scale in southern Sweden. *The Holocene*, *9*, 237–245. <https://doi.org/10.1191/095968399668494320>
- Bradshaw, R. H. W., & Lindbladh, M. (2005). Regional spread and stand-scales establishment of *Fagus sylvatica* and *Picea abies* in Scandinavia. *Ecology*, *86*, 1679–1686. <https://doi.org/10.1890/03-0785>
- Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*, 255–279. <https://doi.org/10.1086/284267>
- Buschbom, J., Yanbaev, Y., & Degen, B. (2011). Efficient long-distance gene flow into an isolated relict oak stand. *Journal of Heredity*, *102*, 464–472. <https://doi.org/10.1093/jhered/esr023>

- Carrascal, L. M., Galva, I., & Gordo, O. (2009). Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos*, *118*, 681–690. <https://doi.org/10.1111/j.1600-0706.2008.16881.x>
- Chybicki, I. J., Waldon-rudzioek, B., & Meyza, K. (2014). Population at the edge: Increased divergence but not inbreeding towards northern range limit in *Acer campestre*. *Tree Genetics & Genomes*, *10*, 1739–1753. <https://doi.org/10.1007/s11295-014-0793-2>
- Comps, B., Gömöry, D., Letouzey, J., Thiébaud, B., & Petit, R. J. (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European Beech. *Genetics Society of America*, *157*, 389–397.
- Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, *144*, 2001–2014.
- Craft, K. J., & Ashley, M. V. (2007). Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. *Forest Ecology and Management*, *239*, 13–20. <https://doi.org/10.1016/j.foreco.2006.11.005>
- Eckert, C. G., Samis, K. E., & Loughheed, S. C. (2008). Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology*, *17*, 1170–1188. <https://doi.org/10.1111/j.1365-294X.2007.03659.x>
- Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics*, *24*, 217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>
- Excoffier, L., Foll, M., & Petit, J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 481–501. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173414>
- Fontaine, M. C., Baird, S. J. E., Piry, S., Ray, N., Tolley, K. A., Duke, S., ... Michaux, J. R. (2007). Rise of oceanographic barriers in continuous populations of a cetacean: The genetic structure of harbour porpoises in Old World waters. *BMC Biology*, *5*, 1–16. <https://doi.org/10.1186/1741-7007-5-30>
- FRA (2010). *Global forest resources assessment 2010*. Rome, Italy: FAO.
- Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, *86*, 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Gregorius, H. R., & Kownatzki, D. (2005). Spatiogenetic characteristics of beech stands with different degrees of autochthony. *BMC Ecology*, *5*, 8. <https://doi.org/10.1186/1472-6785-5-8>
- Guillot, G., Estoup, A., Mortier, F., & Cosson, J. F. (2005). A spatial statistical model for landscape genetics. *Genetics Society of America*, *170*, 1261–1280. <https://doi.org/10.1534/genetics.104.033803>
- Guillot, G., Lebloit, R., Coulon, A., & Frant, A. C. (2009). Statistical methods in spatial genetics. *Molecular Ecology*, *18*, 4734–4756. <https://doi.org/10.1111/j.1365-294X.2009.04410.x>
- Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, *8*, 461–467. <https://doi.org/10.1111/j.1461-0248.2005.00739.x>
- Hamrick, J. L. (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management*, *197*, 323–335. <https://doi.org/10.1016/j.foreco.2004.05.023>
- Hardy, O. J., & Vekemans, X. (2002). Spagedi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*, 618–620. <https://doi.org/10.1046/j.1471-8286.2002.00305.x>
- Hensen, I., Cierjacks, A., Hirsch, H., Kessler, M., Romoleroux, K., Renison, D., & Wesche, K. (2012). Historic and recent fragmentation coupled with altitude affect the genetic population structure of one of the world's highest tropical tree line species. *Global Ecology and Biogeography*, *21*, 455–464. <https://doi.org/10.1111/j.1466-8238.2011.00691.x>
- Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*, 87–112. <https://doi.org/10.1111/j.1095-8312.1999.tb01160.x>
- Hilton, G. M., & Packham, J. R. (1997). A sixteen-year record of regional and temporal variation in the fruiting of beech (*Fagus sylvatica* L.) in England (1980–1995). *Forestry*, *70*, 7–16. <https://doi.org/10.1093/forestry/70.1.7>
- Jump, A. S., & Penuelas, J. (2005). Running to stand still: Adaptation and the response of plants to rapid climate change. *Ecology Letters*, *8*, 1010–1020. <https://doi.org/10.1111/j.1461-0248.2005.00796.x>
- Jump, A. S., & Peñuelas, J. (2006). Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 8096–8100. <https://doi.org/10.1073/pnas.0510127103>
- Kalinowski, S., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*, 1099–1106. <https://doi.org/10.1111/j.1365-294X.2007.03089.x>
- Kramer, A. T., Ison, J. L., Ashley, M. V., & Howe, H. F. (2008). The paradox of forest fragmentation genetics. *Conservation Biology*, *22*, 878–885. <https://doi.org/10.1111/j.1523-1739.2008.00944.x>
- Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*, 1455–1460. <https://doi.org/10.1126/science.3420403>
- Leonardi, S., Piovani, P., Scalfi, M., Piotti, A., Giannini, R., & Menozzi, P. (2012). Effect of habitat fragmentation on the genetic diversity and structure of peripheral populations of beech in central Italy. *Journal of Heredity*, *103*, 408–417. <https://doi.org/10.1093/jhered/ess004>
- Liepert, S., Cheddadi, R., de Beaulieu, J.-L., Fady, B., Gömöry, D., Hussendörfer, E., ... Ziegenhagen, B. (2009). Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) — A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, *153*, 139–149. <https://doi.org/10.1016/j.revpalbo.2008.07.007>
- Lindquist, B. (1931). *Den Skandinaviska bokskogens biologi*. Stockholm, Sweden: Centraltryckeriet.
- McRae, B. H., Beier, P., Dewald, L. E., Huynh, L. Y., & Keim, P. (2005). Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. *Molecular Ecology*, *14*, 1965–1977. <https://doi.org/10.1111/j.1365-294x.2005.02571.x>
- Moilanen, A., & Nieminen, M. (2002). Simple connectivity measures in spatial ecology. *Ecological Society of America*, *83*, 1131–1145. [https://doi.org/10.1890/0012-9658\(2002\)083\[1131:SCMISE\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[1131:SCMISE]2.0.CO;2)
- Mona, S., Ray, N., Arenas, M., & Excoffier, L. (2014). Genetic consequences of habitat fragmentation during a range expansion. *Heredity*, *112*, 291–299. <https://doi.org/10.1038/hdy.2013.105>
- Muir, G., Lowe, A. J., Fleming, C. C., & Vogl, C. (2004). High nuclear genetic diversity, high levels of outcrossing and low differentiation among remnant populations of *Quercus petraea* at the margin of its range in Ireland. *Annals of Botany*, *93*, 691–697. <https://doi.org/10.1093/aob/mch096>
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*, 583–590.
- Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, *29*, 1–10. <https://doi.org/10.1111/j.1558-5646.1975.tb00807.x>
- Nilsson, S. G., & Wästljung, U. (1987). Seed predation and cross-pollination in mast-seeding Beech (*Fagus sylvatica*) patches. *Ecological Society of America*, *68*, 260–265. <https://doi.org/10.2307/1939256>
- Ortego, J., Bonal, R., Munoz, A., & Aparicio, J. M. (2014). Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology*, *7*, 384–395. <https://doi.org/10.1093/jpe/rtt049>
- Ouborg, N. J., Vergeer, P., & Mix, C. (2006). The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology*, *94*, 1233–1248. <https://doi.org/10.1111/j.1365-2745.2006.01167.x>



- Packham, J. R., Thomas, P. A., Atkinson, M. D., & Degen, T. (2012). Biological flora of the British Isles: *Fagus sylvatica*. *Journal of Ecology*, 100, 1557–1608. <https://doi.org/10.1111/j.1365-2745.2012.02017.x>
- Palomino, D., & Carrascal, L. M. (2007). Habitat associations of a raptor community in a mosaic landscape of Central Spain under urban development. *Landscape and Urban Planning*, 83, 268–274. <https://doi.org/10.1016/j.landurbplan.2007.04.011>
- Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421, 37–42. <https://doi.org/10.1038/nature01286>
- Petit, R. J., El Mousadik, A., & Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12, 844–855. <https://doi.org/10.1111/j.1523-1739.1998.96489.x>
- Piry, S., & Luikart, G. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90, 502–503. <https://doi.org/10.1093/jhered/90.4.502>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Provan, J., Beatty, G., Hunter, A., McDonald, R., McLaughlin, E., Preston, S., & Wilson, S. (2008). Restricted gene flow in fragmented populations of a wind-pollinated tree. *Conservation Genetics*, 9, 1521–1532. <https://doi.org/10.1007/s10592-007-9484-y>
- Rosenberg, N. A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J. K., & Feldman, M. W. (2005). Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genetics*, 1, e70. <https://doi.org/10.1371/journal.pgen.0010070>
- Schuster, W. S. F., & Mitton, J. B. (2000). Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity*, 84, 348–361.
- Sjölund, M. J. (2014). *Interactions between natural and anthropogenic impacts on the genetic diversity and population structure of European Beech Forest*. Stirling, Scotland: University of Stirling. PhD Thesis.
- Sjölund, M. J., González-Díaz, P., Moreno-Villena, J. J., & Jump, A. S. (2017). Understanding the legacy of widespread population translocations on the post-glacial genetic structure of the European beech, *Fagus sylvatica* L. *Journal of Biogeography*, 44, 2475–2487. <https://doi.org/10.1111/jbi.13053>
- Sjölund, M. J., González-Díaz, P., Moreno-Villena, J. J., & Jump, A. S. (2019). Data from: Gene flow at the leading range edge - the long-term consequences of isolation in European Beech (*Fagus sylvatica* L. Kuhn). *Dryad Digital Repository*. <https://doi.org/10.5061/dryad.8q51152>
- Sork, V. L., & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, 21, 821–836. <https://doi.org/10.1007/s10980-005-5415-9>
- Szpiech, Z. A., Jakobsson, M., & Rosenberg, N. A. (2008). ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, 24, 2498–2504. <https://doi.org/10.1093/bioinformatics/btn478>
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). Micro-Checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Vranckx, G., Mergeay, J., Cox, K., Muys, B., Jacquemyn, H., & Honnay, O. (2014). Tree density and population size affect pollen flow and mating patterns in small fragmented forest stands of pedunculate oak (*Quercus robur* L.). *Forest Ecology and Management*, 328, 254–261. <https://doi.org/10.1016/j.foreco.2014.05.044>
- Vucetich, J. A., & Waite, T. A. (2003). Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics*, 4, 639–645. <https://doi.org/10.1023/A:1025671831349>
- Wagner, S., Collet, C., Madsen, P., Nakashizuka, T., Nyland, R. D., & Sagheb-talebi, K. (2010). Beech regeneration research: From ecological to silvicultural aspects. *Forest Ecology and Management*, 259, 2172–2182. <https://doi.org/10.1016/j.foreco.2010.02.029>
- Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., ... Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, 416, 389–395. <https://doi.org/10.1038/416389a>
- Wang, H., Sork, V. L., Wu, J., & Ge, J. (2010). Effect of patch size and isolation on mating patterns and seed production in an urban population of Chinese pine (*Pinus tabulaeformis* Carr.). *Forest Ecology and Management*, 260, 965–974. <https://doi.org/10.1016/j.foreco.2010.06.014>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.
- Willi, Y., & Fischer, M. (2005). Genetic rescue in interconnected populations of small and large size of the self-incompatible *Ranunculus reptans*. *Heredity*, 95, 437–443. <https://doi.org/10.1038/sj.hdy.6800732>
- Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, 11, 413–418. [https://doi.org/10.1016/0169-5347\(96\)10045-8](https://doi.org/10.1016/0169-5347(96)10045-8)

BIOSKETCH

Alistair Jump's research team focuses on understanding biogeographical impacts of past and present environmental changes from population genetics to demography and remote sensing and how they interact with human interventions.

SUPPORTING INFORMATION

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

How to cite this article: Sjölund MJ, González-Díaz P, Moreno-Villena JJ, Jump AS. Gene flow at the leading range edge: The long-term consequences of isolation in European Beech (*Fagus sylvatica* L. Kuhn). *J Biogeogr.* 2019;46:2787–2799. <https://doi.org/10.1111/jbi.13701>