Evidence for habitat fragmentation induced genetic degradation in remnant Syzygium maire (Myrtaceae) populations

Colan G Balkwill
Victoria University of Wellington
Emily Koot
Plant & Food Research
Peter Ritchie
Victoria University of Wellington
David Chagné
Plant & Food Research
Julie R Deslippe
julie.deslippe@vuw.ac.nz

Victoria University of Wellington

Research Article

Keywords: Habitat fragmentation, population genetics, genetic diversity loss, whole-genome resequencing, swamp forests, wetlands

Posted Date: January 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8283695/v1

Additional Declarations: No competing interests reported.

Abstract

Context.

Wetland restoration is critical for the recovery of biodiversity, ecosystem functions and services. In highly modified landscapes, conservation must navigate complex trade-offs among these goals. Local-scale information on the genetic structure and trajectory of species can support the design of restoration strategies that maximise diversity.

Objectives.

We investigated population genetics of *Syzygium maire* (maire tawake), a critically threatened canopy tree of swamp forests in the Greater Wellington Region of Aotearoa New Zealand.

Methods.

We sampled leaves of adult trees and seedlings in 12 sites across the region applying low-coverage whole genome resequencing to each individual. We characterised genetic structure in the adult and seedling cohorts, examining patterns of kinship and diversity using hierarchical Analysis of Molecular Variance. Further, we compared the decay in spatial autocorrelation among trees in this region to that of a relatively large and intact reference population.

Results

. Our data suggest historical homogeneity and admixture across extensive pre-fragmentation swamp forests, but fine-scale genetic structure and decay of spatial autocorrelation within a few kilometres indicates recent fragmentation now constrains gene flow. Small yet significant reductions in genetic diversity and increased inbreeding in seedlings point to early genetic erosion in remnant habitat patches. Although genetic loss is not yet severe, continued isolation risks accelerating degradation in future generations.

Conclusions.

Restoration should prioritise reconnecting fragmented populations, maximizing genetic diversity by replanting with broadly sourced seed and reestablishing habitat for avian seed-dispersers. These findings highlight how genetic information can inform restoration planning and improve the long-term resilience of habitat specialist.

Introduction

Wetlands provide numerous ecosystem services and habitat for diverse species, yet they are threatened and in decline globally. Historically, wetlands covered about 9% of Aotearoa New Zealand's land area(Dymond et al. 2021) and formed an essential taonga (treasured) life-system for Māori, the indigenous peoples of Aotearoa (Pa Ropata McGowan 2021. Since European settlement, 90%, of wetland have been lost nationally, with far higher losses across Te Ika-a-Māui (the North Island) {Dymond, 2021 #562), and more than 95% in the Greater Wellington Region (Ausseil et al. 2008, Singers et al. 2018). The most severely affected ecosystem in the region is swamp forest, occupying less than one percent of its historical range (Singers et al. 2018, Tomscha et al. 2019). Once extensive across Kāpiti Coast, the Hutt River Valley, Wainuiomata and Wairarapa (Fig. 4.1), these forests have been largely converted to farmland and urban areas, leaving small and isolated remnant swamp forests (Singers et al. 2018) (Figures S1-4).

Swamp forests are especially susceptible to the negative genetic effects of fragmentation (Vranckx et al. 2012, Schlaepfer et al. 2018). Isolated habitat patches harbour less genetic diversity and increase the risk of inbreeding and genetic drift by breaking up metapopulations and limiting dispersal (Aguilar et al. 2008, Borrell et al. 2020, Walter et al. 2020). Models suggest that fragmentation-induced genetic degradation occurs once habitat loss exceeds 70%, as reduced connectivity among remnant patches accelerates genetic drift (Pflüger et al. 2019). Much of the deforestation in Aotearoa has occurred within the last 200 years, and the genetic effects of swamp forest fragmentation are likely still to play out.

Syzygium maire, an endemic swamp forest tree, has become a central focus of wetland restoration across Aotearoa. Until 2018, it was little known and largely absent from conservation planning. The arrival of myrtle rust (*Austropuccinia psidii*), prompted an national assessment of Myrtaceae biodiversity, after which *S. maire* was classified as nationally critically threatened due to extensive habitat loss and high pathogen susceptibility (de Lange et al. 2018, McCarthy et al. 2019). Since then, *S. maire* has gained recognition as a flagship species for swamp forest health and recovery. Despite this, restoration projects may inadvertently compound risks by failing to capture the species' full genetic diversity or to enhance connectivity among remnant populations.

S. maire is particularly vulnerable to genetic degradation under fragmentation due to its highly specialised habitat requirements and reproductive biology. It is hermaphroditic and capable of self-fertilisation(Bucan 2006, Clarkson 2014). While selfing can maintain reproduction in isolation, over time it compounds inbreeding and genetic loss (Aguilar et al. 2008, Chen et al. 2017). S. maire's generalist dispersal strategy, reliant on multiple insect and birds species (Bucan 2006, Clarkson 2014), may buffer low levels of fragmentation, but effectiveness depends on disperser density, movement behaviour, and habitat continuity (Bond 1994, McRae 2006, Gómez et al. 2007, Vranckx et al. 2012). Insect typically disperse only hundreds of meters (Goverde et al. 2002, Brunet et al. 2019), while birds such as the kererū (*Hemiphaga novaeseelandiae*), a key disperser, deposit most seed within 100 m (Coates and Hamley 1999, Llorens et al. 2004) and seldom exceed 1.5 km (Wotton and Kelly 2012) (Zhang et al. 2021). Kererū and other fauna also avoid crossing non-native habitats (Coates and Hamley, Schlaepfer et al.). As most remnant swamp forest in the Greater Wellington Region now occur within urban and pastoral landscapes

(Singers et al. 2018, Tomscha et al. 2019), dispersal is likely restricted and reproduction largely confined to isolated stands, and further constrained by low disperser abundance and introduced predators (Clout and Hay 1989, Wotton 2007, Carpenter et al. 2021). Furthermore, S. maire seed is recalcitrant and does not form a natural seedbank (van der Walt 2020, van der Walt et al. 2020), and therefore once it is lost from an area there is little option for natural regeneration.

Previously, we described the national-scale genetic structure of *S. maire* (Balkwill et al. 2025), but most human-induced effects occur at finer spatial scales, within regions and individual habitat patches, especially given the species' dispersal dynamics and extreme fragmentation. Inbreeding and high kinship have been detected in isolated populations, likely reflecting founder effects and low migration among remnants (Balkwill et al. 2024). This fine-scale spatial genetic structure (FSGS) arises from the interaction of dispersal, life-history, mating system and mate density across the landscape (Vekemans and Hardy 2004). Assessing the decay of FSGS provides insight into dispersal ability and degree of reproductive isolation among remnant stands (Smouse and Peakall 1999, Rousset 2000, Vekemans and Hardy 2004). Understanding these dynamics is critical for designing restoration that maintains connectivity and genetic diversity. Although no tree species in Aotearoa has yet been assessed for FSGS, recent work suggests large areas of restorable *S. maire* habitat remain in the Greater Wellington Region (Herbert et al. 2025) and that regional genetic diversity remains comparable to other populations (Balkwill et al. 2025), indicating substantial genetic potential for recovery.

Severe habitat loss and fragmentation in the Greater Wellington Region make information on swamp forest protection particularly valuable. Given the region's historically extensive swamp forests, remaining remnants are expected to retain relatively high diversity and to provide suitable seed sources for restoration. However, limited animal-mediated dispersal combined with high levels of fragmentation will isolate remnant stands, resulting in FSGS, close kinship and genetically depauperate seed. To address these concerns, we aimed to 1) quantify standing levels of genetic diversity and kinship across the region, 2) assess historical genetic connectivity among remnant patches by determining spatial autocorrelation decay in adult trees and 3) compare diversity metrics between adult and seedling cohorts at isolated sites.

Materials and Methods

Study sites and sample collection

We studied 12 sites across the Greater Wellington Region (Figure 1, Table S1). Sites were defined as areas hosting naturally occurring *S. maire*, surrounded by less suitable habitat or human altered land (Correa Ayram et al. 2016). Study sites were a mix of remnant sites surrounded by pasture at Fensham Reserve (FHR) and Allen/Lowe's Bush (ALB), non-swamp indigenous forest at Zealandia (ZLD) and Kaitoke (KTK), urban parkland at Ngā Manu/Parkwood (NGP), Gentian Park (GEP) and Te Haukāretu Park (THP) and small isolated patches in regenerating bush in Silverlea (SLV), Sylvan Way (SYL), McLean (PMP) and Carter's Scenic Reserve (CSR). All sites existed in a complex mosaic of landcovers (Table S1,

Figures S1-4). Sites were chosen during 2020 based on iNaturalist records and consultation with local landowners and mana whenua (indigenous Māori peoples of the region). At the time, these locations were among the only recorded occurrences of *S. maire*. Sampling permits were obtained from Greater Wellington Regional, Wellington City and Upper Hutt City Councils for work in their reserves, or from landowners and through mana whenua, in this case Ngāti Kahungunu ki Wairarapa and Te Āti Awa ki Whakarongotai. For the isolation by distance analysis, we used an additional 41 samples derived from Waipoua Forest in Northland. These were collected from sites positioned closer together and were primarily interspersed with indigenous forest, providing for finer-scale assessment of naturally fragmented stands of trees.

Sample collection

Whole leaf samples were collected from adult trees and seedlings between January 2021 and May 2022. Seedlings were defined as individuals less than 5cm tall. As there is no seedbank, all seedlings represent breeding events which occurred within the previous few years. It cannot be assumed that all seedlings were from the previous season as growth rates have not been quantified and it is possible that some had germinated earlier.

Ten to twenty adult samples were collected at each study site. Where more than 20 adult trees were present, effort was made to sample individuals from across the site to minimise over representation of highly related individuals. If multiple trunks emerged from the same area of ground, leaves were only collected from one. Seedlings were collected if present and the nearest adult noted. To minimise impact, we harvested single leaves from seedlings. GPS coordinates and diameter-at-breast height were recorded for all adult samples. Samples were transported on ice, flash frozen in liquid nitrogen and stored at -80 ºC.

DNA extraction and variant calling

DNA was extracted and sequenced as in Balkwill et al. (2024) and Balkwill et al. (2025). Where specific cohorts were needed for analysis, individuals, age groups or populations were removed from the analysis using *gl.drop.pop* or *gl.drop.ind* in DartR v2.9.7 (Gruber et al. 2018) in R v4.2.1 (R Core Team 2013). Monomorphic variants were removed and locus statistics recalculated for each subset. We used a dataset filtered for missing and uncommon alleles (minor allelic frequency greater than 0.05) as the sample size was small and contained many related individuals. Missing data or under sampled alleles may therefore have a large effect on analyses. For kinship analyses a dataset without minor allele frequency (MAF) filtering was used, as recommended previously (Weir and Goudet 2017, Goudet et al. 2018).

Genetic structure

To identify genetic clusters within the data we used k-means clustering, principal component analysis and discriminant analysis of principal components. All clustering analyses were performed using the

adegenet v2.1.10 package in R (Jombart 2008). Nonparametric methods (principal component analysis and discriminant analysis of principal components) were chosen to account for potential bias in our data resulting from small numbers of potentially highly related samples. We ran principal component analysis using the *dudi.pca* function without norming for column vectors and the number of kept axes set to the number of individuals. We ran principal component analysis on all samples, as well as adult only and adult-seedling (AvS) datasets. K-means clustering was performed with the maximum number of clusters set at 20 and the number of principal components to the number of individuals in the analysis. Choice of the optimal number of clusters was made interactively based on Bayesian information criterion score. As a final clustering method, we ran discriminant analysis of principal components (Jombart et al. 2010). Prior groups were defined according to subregion (Table 4.1). The number of principal components was set to the number of individuals and all discriminant functions were retained as the number was small. For maintaining a sufficient number of principal components that discriminated between groups, we used the *optim.a.score* function, and reran discriminant analysis of principal components with the number of principal components set to that which generated the best *a*-score.

In order to test for differences in levels of population structure we used a hierarchical AMOVA (Analysis of Molecular Variance) (Meirmans 2012) with strata set as subregions with nested sampling sites. Molecular variance components for each hierarchy was estimated in the R package poppr v2.9.3 (Kamvar et al. 2014) using the function *poppr.amova* with method set to *ade4* and 99999 permutations. Statistical significance of these components was tested using *randtest* in ade4 v1.7-22 (Dray and Dufour 2007) at 99999 permutations.

Summary statistics

Summary statistics were calculated for all adult trees at the site level, except for the removal of CSR as only a single tree was sampled, All analyses were conducted in R. Population specific and overall fixation index (F_{ST}) with 2.5 and 97.5 % confidence intervals, inbreeding statistic (F_{IS}) and gene diversity was obtained using the methods of Weir and Goudet (2017) with 1000 bootstraps according to the hierfstat v0.5-11 function *betas* (Goudet 2005). Observed and unbiased heterozygosity (corrected for sample size) were obtained per site and per individual using *gl.report.heterozygosity* in DartR. Pairwise F_{ST} with confidence intervals at 2.5 and 97.5 % was calculated according to Weir and Cockerham (1984) using the *genet.dist* and *boot.ppfst* functions in hierfstat with 100 bootstraps. Pairwise private alleles, fixed alleles and allelic frequency differences for all populations were identified using the *gl.report.pa* function in DartR. Comparisons were made both for pairwise site to site comparisons, and between each site and all other sites pooled.

Spatial autocorrelation

We sampled individuals every few meters for the length and breadth of each site. However, due to the small size of remnant patches, we could not sample along the complete dispersal distance of each tree. Isolation by distance was assessed between sites and between individuals using the *gl.ibd* function in

DartR. The method was based on F_{ST} for analyses between sites. Individual assessments were run using proportion of shared alleles. Mantel tests for correlation between genetic and geographical distance was performed on 99999 permutations. Spatial autocorrelation between individuals was examined using *gl.spatial.autoCorr* in DartR. This is a nonparametric assessment developing on the methods of Smouse and Peakall (1999), Peakall et al. (2003) and Smouse et al. (2008) allowing tests for statistical significance. We used Euclidean distances for both linear geographic and genetic distances and calculated spatial autocorrelation in 10.5 bins of equal distance, representing distances increasing by approximately 10 km for each bin. Significance and confidence intervals were assessed using 1000 permutations. To obtain a better understanding of the pattern of decay in autocorrelation, we also tested these statistics for the Northland populations. Sampling of these populations was more regular and over a lesser distance, allowing decay to be calculated at smaller intervals and with less missing information. We chose to assess decay with 31.2 intervals, the maximum distance in km of any pair of individuals, to assess decay per kilometre.

Kinship

We estimated inter-individual kinship and per-individual inbreeding using the Weir and Goudet (2017) beta-estimator with the *beta.dosage* function in hierfstat (*inb = TRUE*). These measures account for small numbers of related samples (Weir and Goudet 2017, Goudet et al. 2018). Calculations were performed on the MAF0.00 dataset, as filtering for rare alleles may reduce accuracy of kinship estimators (Goudet et al. 2018).

Age group comparisons

To test for differences between cohorts, we used a hierarchical AMOVA (Meirmans 2012) with hierarchy set to sampling site and nested age groups (adult or seedling). Molecular variance components for each hierarchy were estimated in poppr using the function *poppr.amova* with method set to *ade4* and 99999 permutations. Statistical significance of these components was tested using *randtest* in ade4 at 99999 permutations.

Differences in genetic diversity between adult and seed populations were tested using the *gl.test.heterozygosity* function in DartR with 1000 replications (Gruber et al. 2018). This function calculates statistical significance through re-randomization comparison of expected heterozygosity estimates. A null hypothesis of no difference was tested at p = 0.05 and p=0.001. Pairwise private alleles, fixed alleles and allelic frequency difference for all adult-seedling pairs were identified using the *gl.report.pa* function in DartR. Individual level kinship and inbreeding statistics were calculated with the *beta.dosage* function in hierfstat (*inb = TRUE*).

Results

Study sites and sampling

Sites were visited from early summer to late autumn and flowering or fruiting trees were observed at the majority of sites. Seedlings were noted at only a few sites, primarily those with boggy ground and closed canopy cover. Myrtle rust was not present at the time but has since been reported in the region.

Sites were small, fragmented fenced off reserves experiencing no understory browsing from farm animals with most receiving active invasive predator control. All sites represent small reserves, apart from KTK and PMP which intersect the broader Tararua Forest Park. All stands comprised fewer than 50 trees, barring ALB, which had fewer than 100 (Table S1). Less than 10 trees were observed at CSR, PMP, THP and KTK. The largest populations of intact swamp forest were ALB, FHR, NGP and GEP. No larger populations were known at the time. Plant cover at ALB, FHR and GEP consisted of waterlogged pukatea (*Laurelia novae-zelandiae*) swamp forest surrounded by kahikatea (*Dacrycarpus dacrydioides*). NGP was primarily dry ephemeral mixed swamp forest, waterlogged swamp forest and open water vegetation shared with kiekie (*Freycinetia banksia*). Smaller populations were also sampled at SVL (waterlogged, regenerating swamp forest in rural context), ZLD (dry stream fan in regenerating urban reserve), SYL (suburban gully in regenerating bush), THP (waterlogged bush remnant near roadside in urban setting) and PMP (boggy, flax with emergent trees). Only a single individual was sampled at CSR, an open water wetland site with flax and emergent trees. Seedlings were collected at GEP, ALB, FHR and THP as they could not be located at other sites. All patches were isolated and more than 1.5 km from the next nearest patch apart from GEP-THP which were ~1.4 km. The next closest pair was NGP-SLV at 3.5 km.

Variant calling and datasets

Variants were successfully detected for 186 individuals across the Greater Wellington Region (Table S2). Of these, 111 were adults of breeding age and 74 were seedlings (Table S1). We further filtered this dataset for adult-seedling (AvS) pairs derived from ALB, FHR, GEP and THP for a total of 54 adults and 74 seedlings. At ALB and FHR, sample sizes were roughly equal for age cohorts, while at GEP we obtained twice as many seedlings. We were only able to sample two seedlings at THP. At MAF > 0.05 the dataset yielded 55,709 SNPs across all 186 individuals, while that for just adults was based on 55,708 SNPs and AvS on 53,633 SNPs (Table S2). For kinship analysis, the dataset filtered for a MAF of 0.00 and no missing data yielded 171,172 SNPs for adult analyses and 156,672 for AvS.

Genetic structure

Weak clustering by site was observed in the principal component analysis (Figure S5). For the dataset containing only adults, axis one explained 6.1% of the variation, while axis two explained 5.1% and axis three 3.6%. Axis one explained variance separating subregions, a rough Kāpiti Coast-Hutt Valley cluster and a Wairarapa group. ZLD and PMP formed tight groups segregated from the rest. K-means clustering suggested that five clusters existed in the dataset. However, when examined more closely, clustering did not seem to fully align to sites or subregions, with some individuals grouped with others from sites across the region (e.g., Wairarapa samples with the Kāpiti Coast). For discriminant analysis of principal components, LD1 (48.8%) primarily explained the separation of ZLD from the rest, and LD2 (35.8%) and

LD3 (15.5%) explained segregation of broader subregions (Figure 2). In general, clusters were tight, barring some FHR and THP individuals and CR01 approaching the NGP/SLV cluster.

The AMOVA suggested that most genetic variation of S. maire in the Greater Wellington Region was contained within sites, followed by among sites and finally among subregions (Table 1).

Table 1: Hierarchical AMOVA for adult S. maire trees in the Greater Wellington Region. Percentage covariance, phi statistics and significance of results at 99999 permutations are shown. Variance was tested for sampling sites nested within subregions.

Comparison	% covariance	Phi	P-value
Variation among subregions	5.74	0.06	0.00952
Variation among sites within a subregion	18.72	0.20	0.00001
Variation within sites	75.53	0.24	0.00001

Summary statistics

We observed medium fixation index (0.18) and inbreeding (0.1) across all of the Greater Wellington Region, (Table 2). At the site level, NGP showed the lowest fixation index (0.04) followed by ALB (0.09) and FHR (0.07), while the highest fixation indices occurred at PMP (0.256) and KTK (0.44). Confidence intervals for all fixation index values fell within 0.01 or less of the estimate. Trends in allelic frequency difference values were similar, although differences were less extreme.

Heterozygosity values were comparable (uH_e from 0.22 to 0.25) for all populations except KTK (0.15) and PMP (0.11) (Table 2, Figure S6). Multi-locus beta-diversities were also lowest in PMP and KTK, and near identical to heterozygosity values, with medians at 0, suggesting that most loci were homozygous across these trees (Table 2, Figure S6). Standard deviations were all high for heterozygosity and variation in beta-diversity and individual H_o was high within sites (Figure S6, Table S3). Inbreeding values within sites were lowest at NGP (0.01), THP (0.04), FHR and SLV (0.06) and ALB (0.07), and highest at KTK (0.37 and PMP (0.57) (Table 2, Figure S7). Variation in individual inbreeding statistics per site was lowest in NGP and PMP.

Pairwise fixation indices among sampling sites ranged between 0.04 (NGP-SVL) and 0.52 (KTK-PMP) (Table S3). All the highest pairwise differences (>0.20) were between KTK or PMP and other sites. The lowest values (<0.10) were between SYL and NGP, FHR and ALB. Confidence intervals ranged within less than 0.01 of the estimated pairwise fixation index.

Table 2: Population-specific and overall summary statistics for adult S. maire trees sampled across the Greater Wellington Region.

Site	$F_{ST\ 2.5\%\ CI}$	F_{ST}	$F_{ST\ 97.5\%\ CI}$	F_{IS}	uH_e	β-gene diversity	AFD
ALB	0.09	0.09	0.09	0.07	0.24	0.24	0.09
FHR	0.06	0.07	0.07	0.06	0.24	0.24	0.08
GEP	0.11	0.12	0.12	0.12	0.23	0.23	0.10
KTK	0.43	0.44	0.45	0.37	0.15	0.15	0.19
NGP	0.04	0.04	0.05	0.01	0.25	0.25	0.10
PMP	0.56	0.57	0.57	0.42	0.11	0.11	0.19
SLV	0.13	0.14	0.14	0.06	0.22	0.22	0.12
SYL	0.13	0.13	0.13	0.17	0.23	0.23	0.10
THP	0.10	0.10	0.11	0.04	0.23	0.23	0.12
ZLD	0.14	0.15	0.15	0.13	0.22	0.22	0.12
Overall	0.18	0.18	0.19	0.11		0.26	

Inbreeding values are the average for individual F_{IS}. Unbiased expected heterozygosities (uH_e) are corrected according to sample size. Allelic frequency differences (AFD) were calculated relative to all other samples.

Spatial autocorrelation

We did not detect significant isolation by distance across the dataset using Mantel tests between subregions (r = 0.71, p = 0.08) or between sites (r = -0.08, p = 0.59) (Table 3, Table S5). Significant isolation by distance at the individual level (r = 0.50, p = $1e^{-5}$), however, was detected for both Greater Wellington (Table 3, Figure S8A) and Northland (Table 3, Figure S8B).

Spatial autocorrelation declined rapidly from the first distance class (10 km, r = 0.07) to the second (20 km, r = 0.01) in the Greater Wellington Region (Figure 3A, Table S6), with both values exceeding the upper limit expected under the null hypothesis of no spatial autocorrelation. All other distance classes were negatively correlated and exceeded the lower bound of the null hypothesis. Correlation was, however, weak at all levels. Significant negative correlation was observed at 19 and 20 km. All other distance classes were not different from zero. In the Northland samples, autocorrelation decreased rapidly from one (r = 0.029) to two (r = 0.006) and became negative by three kilometres (Figure 3B, Table S7). Significant positive correlation was detected only in the first distance class, although the strength of this association was weak.

Table 3: Isolation by distance of *S. maire* trees across the Greater Wellington Region using the Mantel test statistic.

Comparison	r (Mantel statistic)	p-value	Upper quantiles of permutations			
			90%	95%	97.5%	99%
Among subregions	0.71	0.08	0.59	0.60	0.66	0.70
Among sites	-0.08	0.59	0.31	0.39	0.45	0.49
Among individuals	0.50	1E-05	0.03	0.04	0.05	0.06

Isolation by distance was tested at three levels: between subregions (Wairarapa, Kāpiti, Hutt Valley and Wellington), between sites and between all individuals across Greater Wellington. Geographic distances were based on Euclidian distance in meters, while genetic distance was F_{ST} for subregions and sites and proportion shared alleles for indivduals. Significance was tested for 99999 permutations.

Allelic analyses

In pairwise comparisons between populations, we observed the lowest allelic frequency differencesfor ALB-FHR (0.07), the largest populations (Table S8), while the highest values (>0.2) were obtained for comparisons of small, inbred populations. For comparisons between each sampling site and the rest of the pooled dataset, the lowest allelic frequency differences were ascribed to FHR, ALB, NGP and GEP (Table 2, Table S9). The largest were for CSR and PMP. No unique alleles were observed for any one site. However, differences in allelic frequency were found when comparing site to site pairwise (Table S9).

Kinship

Across the Greater Wellington Region, we find evidence for both strongly inbred and outcrossed individuals, with individual kinship ranging from -0.15 to 0.55. Varying kinship was often observed between individuals within sites (Figure 4). Relatedness blocks between sites existed for Kāpiti Coast and Hutt Valley populations. Within sites, PMP and KTK were the most related, with high inbreeding and kinship among all individuals in PMP. By contrast, ALB, FHR and NGP had the lowest within site inbreeding and kinship values. Outliers were present at all sites.

Age group comparisons

Seed for all sites appeared to represent the complete variation of adults in principal component graph space, suggesting they captured a substantial proportion of total adult diversity (Figure 5). For the principal component analysis of only the sites where seed was sampled, axis one explained 10.6% of the variance, axis two 5.5% and axis 3, 3.0%. GEP and Wairarapa sites (ALB, FHR) were differentiated along axis one, with THP in the centre. Axis two primarily explained variation within GEP. Clustering was far tighter in the Wairarapa samples, with GEP samples representing a far looser association.

Differences in heterozygosity between sites were small, ranging from -0.005 (FHR_A - ALB_S) to 0.110 (FHR_A - THP_S) (Table S10). Excluding THP, the maximum deviation was 0.013 (FHR_A to GEP_S). The difference in heterozygosity between ALB cohorts was 0.001, which was statistically significant (Table

4). Differences between FHR and GHP cohorts respectively were 0.004 and 0.001, both at p-values < 0.001. Interestingly, non-significant differences were observed for ALB_A-FHR_S and ALB_S-FHR_S comparisons (Table S10). uH_e ranged from 0.224 to 0.242 for all sample sets apart from THP_S at 0.166 (Table S11).

Table 4: Comparison of genetic diversity statistics for adult and seedling *S. maire* trees from single populations.

Site	F_{IS} Adult	F_{IS} Seed	uH_e Adult	uH_e Seed	H_e diff	p-value H_e diff	AFD
ALB	0.0338	0.0554	0.2362	0.2341	0.0009	0.022	0.07
FHR	0.0101	0.0155	0.2420	0.2371	0.0043	<0.001	0.072
GEP	0.1226	0.1418	0.2292	0.2243	0.0013	<0.001	0.048
THP	0.0685	0.2992	0.2330	0.1659	0.0892	<0.001	0.154

Inbreeding values are the average for individual F_{IS}. Unbiased expected heterozygosities (uH_e) are corrected according to sample size. Differences in expected heterozygosity for adult-seedling pairs (H_e diff) was calculated with 1000 replications for significance testing. Allelic frequency differences (AFD) were calculated relative to all other samples.

Pairwise allelic frequency difference between adults and seedlings ranged from 0.048 for GEP_{AvS} to 0.203 for FHR_{AvS}-THP_{AvS} (Table 4, Table S10). Excluding THP, the maximum allelic frequency difference was GEP_A-FHR_S (0.143). Allelic frequency difference for adult seedling pairs were 0.070, 0.072 and 0.154 for ALB, FHR and THP respectively. For the GEP pair, seedlings had 4,308 alleles (7.8% of the SNP dataset) not represented in adults, while adults had 974 not present in seedlings (Table S10). FHR and ALB adults lacked 2,435, and 2,563 seedling alleles. The seedling-adult relationships showed the least number of pairwise private alleles. Except in the case of ALB, seedlings showed higher allelic frequency difference than adults at the same site (Table S8). At the regional level, adult and seedlings from the same site have similar allelic frequency difference and number of missing alleles compared to all other pooled populations (Table S9).

For the adult-seedling dataset, pairwise individual kinship was generally highest within sampling sites and seedlings were most often most closely related to adults within sites (Figure S9). There were, however, exceptions, with some individuals showing higher kinship with individuals at other sites compared to their own. For instance, two adults in THP displayed similar kinship to adults and seedlings in GHP. There was also a range of individual inbreeding values and inbreeding was consistently higher for seedling cohorts compared to adults (Table 4, Figure 6). Seedlings in GEP appeared most related to each other and GEP had higher within-site kinship compared to ALB and FHR.

AMOVA suggested that most variation existed withing age groups (80.39%) followed by between sites (18.48%) (Table 5). Differences between age groups explained a very small proportion of the variation and was marginally significant with a p-value of 0.03. When THP cohorts were removed from the analysis, between age group variance was no longer significant (0.07).

Table 5: Hierarchical AMOVA for adult versus seedling trees.

Comparison	% covariance	Phi	P-value
Including THP			
Variation between site	18.48	0.19	0.00001
Variation between age group within site	1.14	0.01	0.02979
Variation within age group	80.39	0.19	0.00001
Excluding THP			
Variation between site	18.16	0.19	0.00001
Variation between age group within site	0.86	0.01	0.06759
Variation within age group	80.99	0.18	0.00001

Variance was tested for age group nested within sample sites. Percentage covariance, phi statistics and significance of results at 99999 permutations are presented.

Discussion

We assessed local-scale genetic structure, isolation by distance and fragmentation effects on the genetic diversity of remnant *S. maire* stands across the Greater Wellington Region. We identified evidence of weak structure and admixture, as well as varying heterozygosity, inbreeding and kinship within sites, likely relating to isolation by distance and natural fragmentation in micro-habitats. Significant and rapid decay of spatial autocorrelation over just a few kilometres indicates that remnant patches will likely experience genetic erosion due to habitat fragmentation. Small but significant differences in heterozygosity and higher inbreeding in seedlings suggest remnants are experiencing effects of fragmentation. Restoration and conservation goals would best be served by capturing diversity within the region and improving population connectivity to better emulate the natural dispersal state of the species, especially through expanding existing swamp forest and restoring swamp forest habitat within a few kilometres of remnant patches.

Across the region, there is evidence of weak genetic structure. Principal component and discriminant analysis of principal component analyses suggest broad clusters corresponding to subregions (Wairarapa, Hutt Valley, Wellington and Kāpiti Coast), and medium to high fixation index values further suggest population differentiation. However, AMOVA results and deviations from subregion-based

clustering in k-means and pairwise F_{ST} analyses indicate that this structure is weak. The absence of private alleles at any one site relative to others suggests that observed structure reflects differences in allele frequencies rather than unique genetic variants. Small sample sizes and exclusion of rare alleles may also have limited detection of site specific variation. Apart from PMP and KTK, diversity measures are similar across subregions, again, implying that structure arises from frequency shifts or allelic combinations, rather than distinct genetic content. Because adaptation may depend on specific combinations of alleles, some of this structure could reflect selective processes. Nevertheless, adaptive variation is unlikely to be highly localised as historically extensive swamp forests would have promoted outbreeding and admixture with neutral processes, isolation by distance or patch-level selection shaping site-level allelic frequencies. Homogenization of allelic content through admixture, followed by selection favouring locally advantageous genotypes, could maintain overall genetic diversity, a process consistent with balancing selection (Slotte et al. 2010, Delph and Kelly 2014, Kutschera et al. 2020).

Larger, historically continuous swamp forest sites (ALB, FHR, NGP) showed greater genetic diversity and lower differentiation (fixation indices, inbreeding statistics and allelic frequency differences) than smaller, isolated patches (KTK and PMP). Even low migration rates can homogenise variation across a species' range, and extensive swamp forest would have provided abundant mating opportunities. However, insect pollinators tend to remain, within forest patches rather than move between them (Didham et al. 1996, Goverde et al. 2002), and birds are unlikely to move far if food resources are sufficient (Wotton and Kelly 2012). Chance long-distance dispersal into small or marginal habitats is therefore likely followed by inbreeding and selfing. This pattern suggests that large, admixed populations maintain higher genetic diversity, whereas dispersal into ephemeral or isolated habitats produces founder effects and selfing. Thus, population context should be carefully considered when selecting sites for conservation and seed sourcing.

Stands that were once part of extensive swamp forest also show much lower kinship, suggesting that local substructure results from the dispersal dynamics of this animal dispersed tree. Spatial genetic structure was detected at the individual level for both the Northland and Wellington populations, with Mantel tests indicating moderate to strong isolation by distance. Spatial autocorrelation declined sharply within 1–10 km, beyond which it became non-significant or negative, consistent with the maximum estimated seed dispersal distance of kererū (Wotton and Kelly 2012). No isolation by distance was evident within individual patches, either reflecting limited sampling or or genuinely weak structure at this scale. The rapid decay and weak autocorrelation suggest that historic populations in large swamps experienced sufficient gene flow to maintain admixture. In wind-pollinated beech and oak, spatial autocorrelation decays within as little as 60 meters due to pollen influx from outside stands (Nakanishi et al. 2009, Sandurska et al. 2024). Insect-mediated pollen transfer rarely exceeds a few hundred meters (White et al. 2002, Hardesty et al. 2006, Hoebee et al. 2007, Byrne et al. 2008, Dick et al. 2008). For insect-pollinated, bird-dispersed species like *Copaifera langsdorffii*, pollination occurs within 150 meters (Sebbenn et al. 2011) and most seedlings are clustered within 100 meters of their mother tree. In the context of these studies, admixture at the kilometre scale in large *S. maire* populations is plausible but

fragmentation likely leads to dispersal failure once remnants are widely separated and disperser density is low.

In *S. maire* seedlings from isolated stands, we observed significant loss of heterozygosity, increased inbreeding and small but significant shifts in allele frequencies. Despite this, seedlings retained most alleles present in their parents, and allele frequencies remain comparable to their parents and other outbred populations. Similar patterns have been observed for temperate, wind dispersed (Guillardin et al. 2024, Liesebach et al. 2024) and tropical, animal dispersed species with mixed mating system (Sebbenn et al. 2011, Waqar et al. 2021). Molecular variance analysis revealed significant age cohort differences only when THP was included, indicating possible bias from small sample size, a likely reality for many *S. maire* populations with limited recruitment. These results indicate that the region's genetic diversity remains broadly represented across populations. Moreover, because alleles present at any single site are also found elsewhere, there is still time to capture and propagate this diversity through well-designed restoration. However, fragmentation effects are expected to intensify if patches are kept isolated, and passive restoration alone will not sustain genetic diversity. Active, wide-ranging seed collection is the best strategy. Sources at NGP, FHR and ALB were the most diverse and should be prioritised for conservation and seed sourcing, though relatedness within stands cautions against relying on only a few populations.

Given the species' natural dispersal dynamics, the current spatial arrangement is unlikely to maintain sufficient levels of migration or viable meta-population sizes. A rule-of-thumb for the species may be that any patch greater than 2 km from another should be considered isolated. Restoration should prioritise sites within 2 km of existing patches. Where this is not possible, artificial migration with mixed provenancing would serve to maximise retained diversity. *Ex situ* collections should also be established to insure against mortality due to myrtle rust incursion, drought or other challenges. Although seed banking is not an option, nursery propagation is well described (Bettoni et al. 2022). Furthermore, increasing native forest cover as disperser habitat, even in non-swamp areas, would likely enhance connectivity among *S. maire* remnants. Conservation and restoration planning should also ensure that disperser populations remain sufficiently dense to maintain gene flow, by expanding suitable habitat in the surrounding matrix and reducing predation pressure. In some cases, however, it may be beneficial to retain a degree of isolation to buffer against the spread of myrtle rust.

The limited number of sampling sites, especially in adult-seedling comparisons, and non-exhaustive sampling design is acknowledged as a limitation of this study. Future studies would benefit from exhaustive sampling to achieve larger sample sizes. Since data collection, additional *S. maire* populations have been discovered, creating opportunities for expanded analysis. With a reference genome available (Balkwill et al. 2024), and this preliminary information, ultra-low coverage whole genome resequencing could provide cost-effective means to genotyping many more individuals (Martin et al. 2021). Collecting seed from known maternal trees would increase power to detect crosspatch pollination events and to describe pollination-dependant dispersal dynamics. Finally, integrating these

data into landscape-scale dispersal models that incorporate standing genetic diversity would provide a valuable basis for evidence-based restoration planning in the region.

In conclusion, restoration in the Greater Wellington Region should prioritise reestablishing habitat patches that enhance connectivity among existing *S. maire* remnants. Where this is not feasible, or where sites are located more than 2 km from the nearest patch, planting should incorporate seed sourced multiple populations to maintain genetic diversity. Local authorities should also focus on conserving the larger, more outbred populations at ALB, FHR and NGP. These recommendations are likely to be relevant to *S. maire* populations throughout Aotearoa, as well as other habitat-specialist species world-wide. However, restoration strategies should always account for species-specific dispersal mechanisms and the surrounding landscape context to determine acceptable levels of habitat fragmentation and ensure long-term population resilience.

Declarations

Acknowledgements

We thank Peter Bellingham (Manaaki-Whenua Landcare Research) for support with the Ngā Rakau Taketake project, Prof. Bruce Weir for discussions about kinship analysis and Karen van der Walt for insight into the ecology and locations of *S. maire*. Computational analyses were performed with the support of the New Zealand eScience Infrastructure (NeSI, https://www.nesi.org.nz/).

Author contributions

CB performed sampling and data analysis with assistance from EMK, JRD, and DC. DC, JRD and PR aided in conceiving and developing the study. CB wrote the manuscript with input from all authors.

Funding

Laboratory and operational costs and sequencing was supported by a Biological Heritage New Zealand Ngā Rakau Taketake project grant (https://bioheritage.nz/about-us/nga-rakau-taketake/, grant number E3784 – 3339), Wellington City Council Our Natural Capital Research Grant (grant numbers E3784 – 3339 and 2021-44-017) allocated to JRD and the New Zealand Ministry for Business Innovation and Employment (MBIE) Strategic Science Investment Fund (SSIF) platform ‘Genomics Aotearoa’. Other costs were supported by Tararua Tramping Club and Wellington Botanical Society grants allocated to CB. DC and PR are supported by the MBIE platform Genomics Aotearoa (https://www.genomics-aotearoa.org.nz/). CB was supported by a Te Herenga Waka Victoria University of Wellington Doctoral Scholarship and Bioprotection Aotearoa.

Data availability

Raw fastq files and filtered vcf are available from the Aotearoa Genomic Data Repository (Te Aika et al. 2023) at https://data.agdr.org.nz/ (https://doi.org/10.57748/1q9v-pr20) with permission access to be

consented from mana whenua where it was requested to do so.

Conflicts of Interest

The authors do not have any conflicts of or competing interests to disclose.

References

1. Aguilar, R., M. Quesada, L. Ashworth, Y. Herrerias-Diego, and J. Lobo. 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. Molecular ecology 17:5177–5188.
2. Ausseil, A.-G., P. Gerbeaux, W. L. Chadderton, T. Stephens, D. Brown, and J. Leathwick. 2008. Wetland ecosystems of national importance for biodiversity: criteria, methods and candidate list of nationally important inland wetlands. Landcare Research Contract Report LC0708/158.
3. Balkwill, C. G., J. R. Deslippe, P. Horton, C. David, C. Wu, E. Koot, P. Ritchie, W. Blissett, and D. Chagné. 2024. De-novo assembly of a reference genome for a critically threatened Aotearoa New Zealand tree species, *Syzygium maire* (Myrtaceae). Tree Genetics & Genomes 20:1–20.
4. Balkwill, C. G., E. Koot, P. Ritchie, D. Chagné, and J. R. Deslippe. 2025. Adaptive Potential of Syzygium maire, a Critically Threatened Habitat Specialist Tree Species in Aotearoa New Zealand. Evolutionary applications 18:e70161.
5. Bettoni, J. C., K. van der Walt, J. A. Souza, A. McLachlan, and J. Nadarajan. 2022. Sexual and asexual propagation of *Syzygium maire*, a critically endangered Myrtaceae species of New Zealand. New Zealand Journal of Botany:1–18.
6. Bond, W. 1994. Do mutualisms matter? Assessing the impact of pollinator and disperser disruption on plant extinction. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 344:83–90.
7. Borrell, J. S., J. Zohren, R. A. Nichols, and R. J. Buggs. 2020. Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. Evolutionary applications 13:161–175.
8. Brunet, J., Y. Zhao, and M. K. Clayton. 2019. Linking the foraging behavior of three bee species to pollen dispersal and gene flow. PLoS One 14:e0212561.
9. Bucan, A. 2006. Karori Sanctuary Trust newsletter “Nature’s Corner”.
10. Byrne, M., C. P. Elliott, C. J. Yates, and D. J. Coates. 2008. Maintenance of high pollen dispersal in *Eucalyptus wandoo*, a dominant tree of the fragmented agricultural region in Western Australia. Conservation genetics 9:97–105.
11. Carpenter, J. K., S. Walker, A. Monks, J. Innes, R. N. Binny, and A.-K. V. Schlesselmann. 2021. Factors limiting kererū (*Hemiphaga novaeseelandiae*) populations across New Zealand. New Zealand Journal of Ecology 45:1–15.
12. Chen, J., S. Glémin, and M. Lascoux. 2017. Genetic diversity and the efficacy of purifying selection across plant and animal species. Molecular biology and evolution 34:1417–1428.

13. Clarkson, W. 2014. Waiwaka: Taranaki's forgotten swamp tree. Page 44. WM Clarkson, New Plymouth, New Zealand.
14. Clout, M. N., and J. Hay. 1989. The importance of birds as browsers, pollinators and seed dispersers in New Zealand forests. New Zealand Journal of Ecology:27–33.
15. Coates, D. J., and V. L. Hamley. 1999. Genetic divergence and the mating system in the endangered and geographically restricted species, *Lambertia orbifolia* Gardner (Proteaceae). Heredity 83:418–427.
16. Correa Ayram, C. A., M. E. Mendoza, A. Etter, and D. R. P. Salicrup. 2016. Habitat connectivity in biodiversity conservation: A review of recent studies and applications. Progress in Physical Geography 40:7–37.
17. de Lange, P. J., J. R. Rolfe, J. W. Barkla, S. P. Courtney, B. SM, F. KA, B. I, S. I, and H.-W. R. 2018. Conservation status of New Zealand indigenous vascular plants. Department of Conservation, Wellington.
18. Delph, L. F., and J. K. Kelly. 2014. On the importance of balancing selection in plants. New Phytologist 201:45–56.
19. Dick, C. W., O. J. Hardy, F. A. Jones, and R. J. Petit. 2008. Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. Tropical Plant Biology 1:20–33.
20. Didham, R. K., J. Ghazoul, N. E. Stork, and A. J. Davis. 1996. Insects in fragmented forests: a functional approach. Trends in Ecology & Evolution 11:255–260.
21. Dray, S., and A.-B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of statistical software 22:1–20.
22. Dymond, J. R., M. Sabetizade, P. F. Newsome, G. R. Harmsworth, and A.-G. Ausseil. 2021. Revised extent of wetlands in New Zealand. New Zealand Journal of Ecology 45:1–8.
23. Gómez, J. M., J. Bosch, F. Perfectti, J. Fernández, and M. Abdelaziz. 2007. Pollinator diversity affects plant reproduction and recruitment: the tradeoffs of generalization. Oecologia 153:597–605.
24. Goudet, J. 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184–186.
25. Goudet, J., T. Kay, and B. S. Weir. 2018. How to estimate kinship. Molecular ecology 27:4121–4135.
26. Goverde, M., K. Schweizer, B. Baur, and A. Erhardt. 2002. Small-scale habitat fragmentation effects on pollinator behaviour: experimental evidence from the bumblebee *Bombus veteranus* on calcareous grasslands. Biological Conservation 104:293–299.
27. Gruber, B., P. J. Unmack, O. F. Berry, and A. Georges. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular Ecology Resources 18:691–699.
28. Guillardin, L., J. MacKay, G. Kerr, and E. Glover. 2024. Low effective population size but high heterozygosity revealed by SNP analyses in adult and juvenile *Thuja plicata* in UK Woodlands. bioRxiv:2024.2003. 2018.585477.

29. Hardesty, B. D., S. P. Hubbell, and E. Bermingham. 2006. Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. Ecology Letters 9:516–525.
30. Herbert, S. M., S. A. Tomscha, H. R. Lai, R. Benavidez, C. G. Balkwill, P. R. Ruston, B. Jackson, and J. R. Deslippe. 2025. Identifying potentially suitable and accessible refugia to mitigate impacts of an emerging disease on a rare tree. Conservation biology:e70088.
31. Hoebee, S., U. Arnold, C. Düggelin, F. Gugerli, S. Brodbeck, P. Rotach, and R. Holderegger. 2007. Mating patterns and contemporary gene flow by pollen in a large continuous and a small isolated population of the scattered forest tree *Sorbus torminalis*. Heredity 99:47–55.
32. Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.
33. Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC genetics 11:1–15.
34. Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281.
35. Kutschera, V. E., J. W. Poelstra, F. Botero-Castro, N. Dussex, N. J. Gemmell, G. R. Hunt, M. G. Ritchie, C. Rutz, R. A. W. Wiberg, and J. B. Wolf. 2020. Purifying selection in corvids is less efficient on islands. Molecular biology and evolution 37:469–474.
36. Liesebach, H., P. Eusemann, A. M. Höltken, U. Tröber, O. Kuchma, M. Karopka, F. Becker, R. Kätzel, and B. Fussi. 2024. Effective population size of adult and offspring cohorts as a genetic monitoring tool in two stand-forming and wind-pollinated tree species: *Fagus sylvatica* L. and *Picea abies* (L.) Karst. Conservation genetics:1–15.
37. Llorens, T., D. Ayre, and R. Whelan. 2004. Evidence for ancient genetic subdivision among recently fragmented populations of the endangered shrub *Grevillea caleyi* (Proteaceae). Heredity 92:519–526.
38. Martin, A. R., E. G. Atkinson, S. B. Chapman, A. Stevenson, R. E. Stroud, T. Abebe, D. Akena, M. Alemayehu, F. K. Ashaba, and L. Atwoli. 2021. Low-coverage sequencing cost-effectively detects known and novel variation in underrepresented populations. The American Journal of Human Genetics 108:656–668.
39. McCarthy, J. K., S. J. Richardson, J. A. Cooper, P. J. Bellingham, and S. K. Wiser. 2019. Species distribution models of the native New Zealand Myrtaceae. Page 93. Ministry for Primary Industries, Wellington.
40. McRae, B. H. 2006. Isolation by resistance. Evolution 60:1551–1561.
41. Meirmans, P. G. 2012. AMOVA-based clustering of population genetic data. Journal of Heredity 103:744–750.
42. Nakanishi, A., N. Tomaru, H. Yoshimaru, T. Manabe, and S. Yamamoto. 2009. Effects of seed-and pollen-mediated gene dispersal on genetic structure among *Quercus salicina* saplings. Heredity 102:182–189.

43. Pa Ropata McGowan, R. 2021. Te reo o te repo kei konei tonu au.*in* Y. Taura, C. v. Schravendijk-Goodman, and B. Clarkson, editors. Manaaki Whenua – Landcare Research.
44. Peakall, R., M. Ruibal, and D. B. Lindenmayer. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. Evolution 57:1182–1195.
45. Pflüger, F. J., J. Signer, and N. Balkenhol. 2019. Habitat loss causes non-linear genetic erosion in specialist species. Global ecology and conservation 17:e00507.
46. R Core Team, R. 2013. R: A language and environment for statistical computing. R foundation for statistical computing Vienna, Austria.
47. Rousset, F. 2000. Genetic differentiation between individuals. J Evol Biol 13:58–62.
48. Sandurska, E., B. Ulaszewski, K. Meyza, E. Sztupecka, and J. Burczyk. 2024. Factors determining fine-scale spatial genetic structure within coexisting populations of common beech (*Fagus sylvatica* L.), pedunculate oak (*Quercus robur* L.), and sessile oak (*Q. petraea* (Matt.) Liebl.). Annals of Forest Science 81:3.
49. Schlaepfer, D. R., B. Braschler, H. P. Rusterholz, and B. Baur. 2018. Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. Ecosphere 9:e02488.
50. Sebbenn, A., A. Carvalho, M. Freitas, S. Moraes, A. P. d. S. d. C. Gaino, J. Da Silva, C. Jolivet, and M. Moraes. 2011. Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. Heredity 106:134–145.
51. Singers, N., P. Crisp, and O. Spearpoint. 2018. Forest ecosystems of the Wellington Region. Greater Wellington Regional Council.
52. Slotte, T., J. P. Foxe, K. M. Hazzouri, and S. I. Wright. 2010. Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. Molecular biology and evolution 27:1813–1821.
53. Smouse, P. E., and R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82:561–573.
54. Smouse, P. E., R. Peakall, and E. Gonzales. 2008. A heterogeneity test for fine-scale genetic structure. Molecular ecology 17:3389–3400.
55. Te Aika, B., L. Liggins, C. Rye, E. O. Perkins, J. Huh, R. Brauning, T. Godfery, and M. A. Black. 2023. Aotearoa genomic data repository: an āhuru mōwai for taonga species sequencing data. Molecular Ecology Resources.
56. Tomscha, S., J. Deslippe, M. de Róiste, S. Hartley, and B. Jackson. 2019. Uncovering the ecosystem service legacies of wetland loss using high-resolution models. Ecosphere 10:e02888.
57. van der Walt, K. 2020. Investigating cryopreservation options for'Syzygium maire', a threatened endemic New Zealand tree. Australasian Plant Conservation: Journal of the Australian Network for Plant Conservation 29:22–25.

58. van der Walt, K., P. Kemp, S. Sofkova-Bobcheva, D. J. Burritt, and J. Nadarajan. 2020. Seed development, germination, and storage behaviour of *Syzygium maire* (Myrtaceae), a threatened endemic New Zealand tree. New Zealand Journal of Botany:1–19.
59. Vekemans, X., and O. J. Hardy. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular ecology 13:921–935.
60. Vranckx, G., H. Jacquemyn, B. Muys, and O. Honnay. 2012. Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. Conservation biology 26:228–237.
61. Walter, G. M., S. Catara, J. R. Bridle, and A. Cristaudo. 2020. Population variation in early development can determine ecological resilience in response to environmental change. New Phytologist:1.
62. Waqar, Z., R. C. S. Moraes, M. Benchimol, J. C. Morante-Filho, E. Mariano-Neto, and F. A. Gaiotto. 2021. Gene flow and genetic structure reveal reduced diversity between generations of a tropical tree, *Manilkara multifida* Penn., in atlantic forest fragments. Genes 12:2025.
63. Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution:1358–1370.
64. Weir, B. S., and J. Goudet. 2017. A unified characterization of population structure and relatedness. Genetics 206:2085–2103.
65. White, G., D. Boshier, and W. Powell. 2002. Increased pollen flow counteracts fragmentation in a tropical dry forest: an example from *Swietenia humilis* Zuccarini. Proceedings of the National Academy of Sciences 99:2038–2042.
66. Wotton, D. M. 2007. Consequences of dispersal failure: kereru and large seeds in New Zealand.
67. Wotton, D. M., and D. Kelly. 2012. Do larger frugivores move seeds further? Body size, seed dispersal distance, and a case study of a large, sedentary pigeon. Journal of Biogeography 39:1973–1983.
68. Zhang, J., J. L. Pannell, B. S. Case, G. Hinchliffe, M. C. Stanley, and H. L. Buckley. 2021. Interactions between landscape structure and bird mobility traits affect the connectivity of agroecosystem networks. Ecological Indicators 129:107962.

Figures

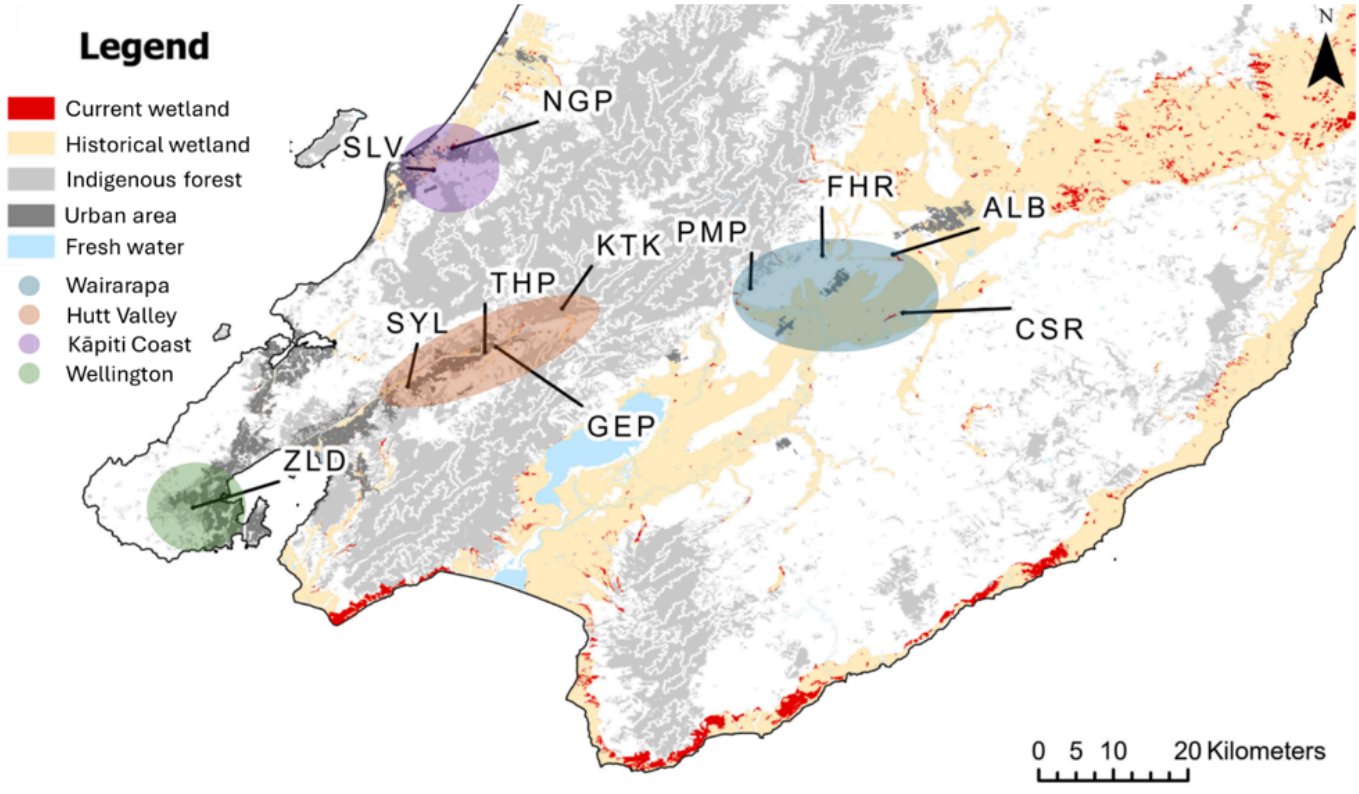

Figure 1

Locations of sampling sites used for this study. Red depicts extant wetland habitat, while light orange depicts historical wetland. Light grey relief represents surviving native forest. Urban areas are coloured dark grey. Information was derived from the New Zealand Landcover Database version 5 (LRIS), geographical borders from NZ Coastlines and Islands Polygons (LINZ) and Greater Wellington Regional Council.

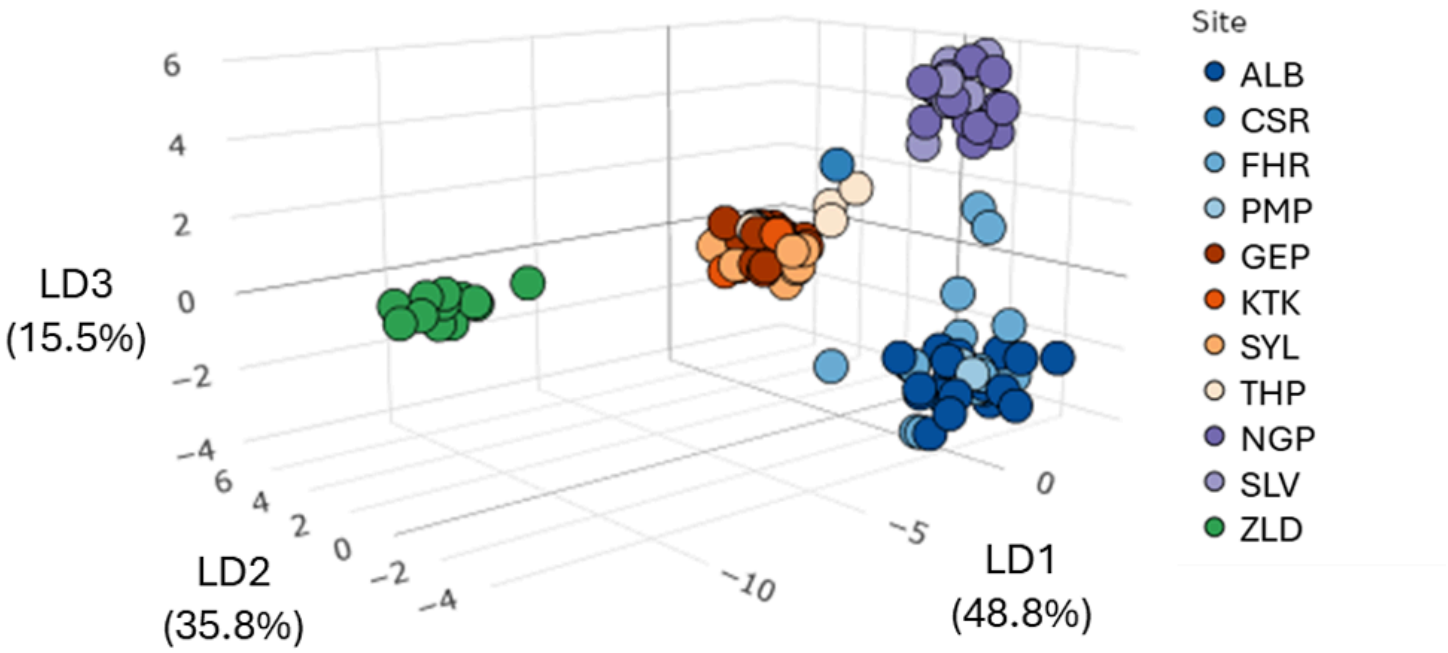

Figure 2

Three-dimensional representation of discriminant analysis of principal components for 111 reproductive age *S. maire* trees sampled in the Greater Wellington Region. Analyses were based on 55,708 SNPs at a minor allele frequency of 0.05, filtered for linkage disequilibrium and without missing information. Individual points represent single trees coloured according to sampling sites. Points are coloured according to region, with tones according to sample site.

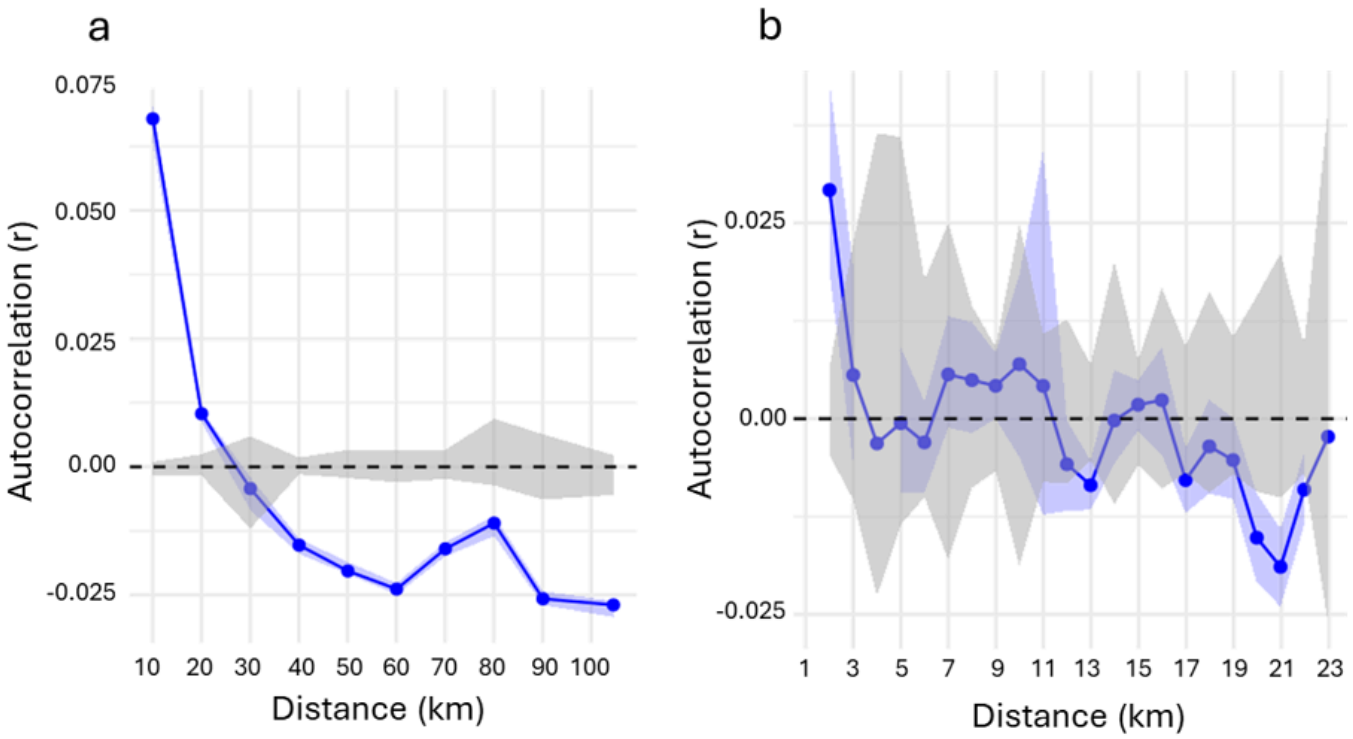

Figure 3

Decay in spatial autocorrelation among *S. maire*trees in Greater Wellington (A) and Northland (B) populations. The blue points, line and error ribbon depict estimated autocorrelation coefficient (r) and its bootstrapped upper and lower limits. The grey ribbon shows the calculated upper and lower limits for the null hypothesis of no spatial autocorrelation. Autocorrelation is significant if its estimate exceeds the limits of the null hypothesis.

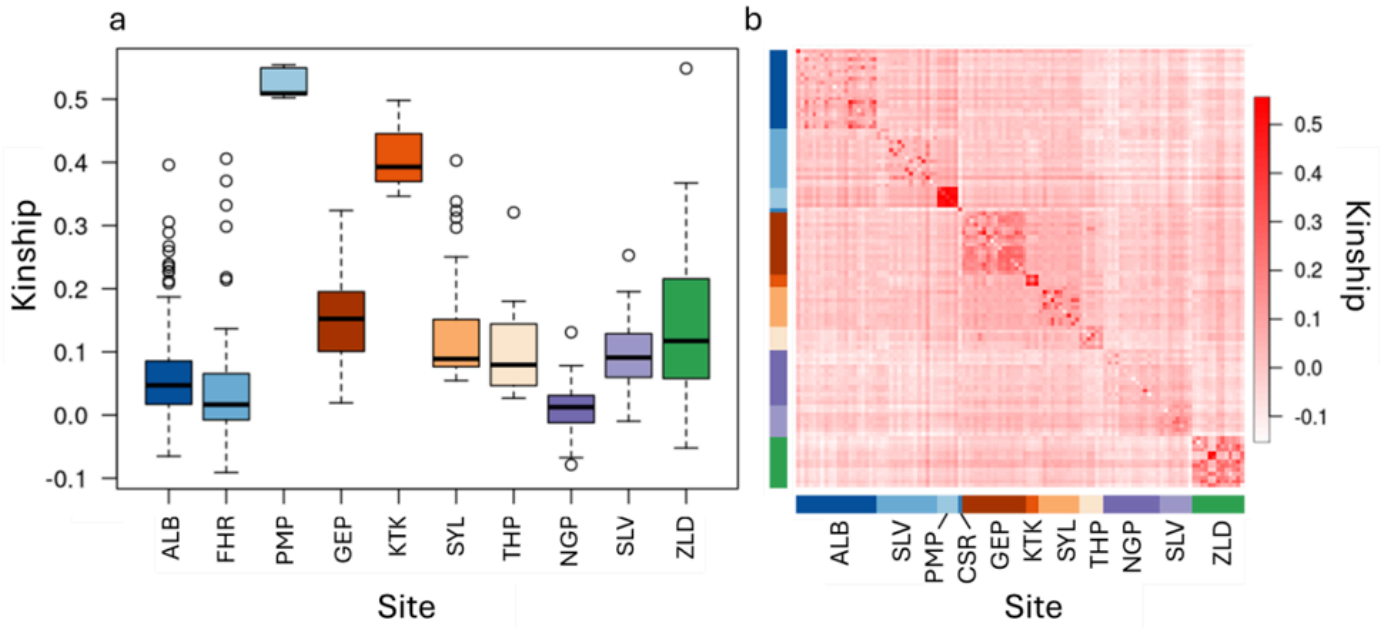

Figure 4

Kinship and inbreeding coefficients for reproductive age *S. maire* trees in the Greater Wellington Region . A) Boxplots representing the distribution of pairwise individual kinship within sampling sites. B) Heatmap depicting inbreeding and pairwise kinship of all individuals. Calculations were based on 171,172 SNPs filtered for linkage disequilibrium and missing information but not minor allele frequency. Inbreeding coefficients are displayed on the diagonal of the heatmap. Individuals are grouped into sites from which they were sampled, as depicted by multi-coloured bars alongside each heatmap.

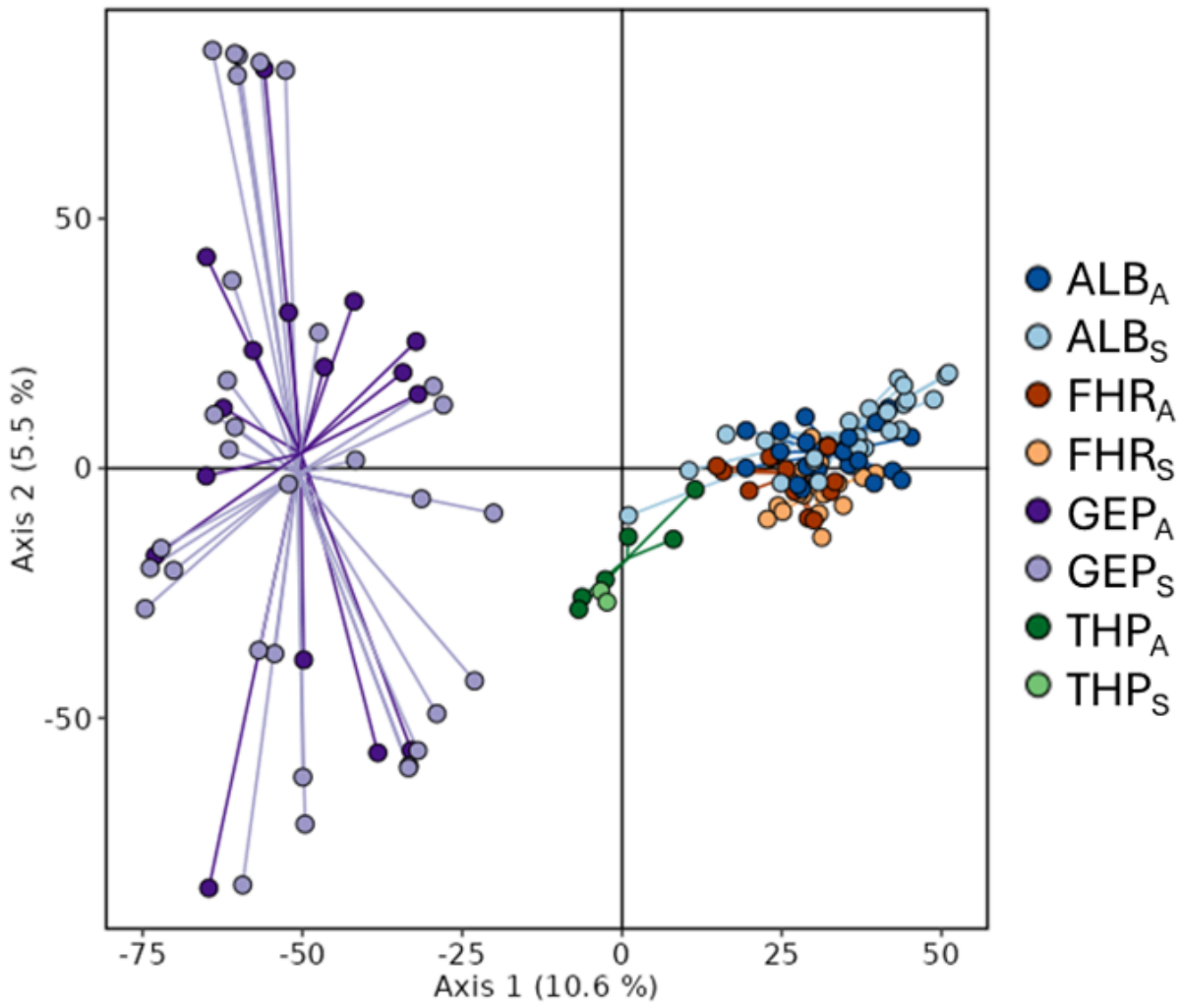

Figure 5

PCA of adult and seedling *S. maire* SNPs coloured by sampling site and cohort., symbols denote individual trees. Analyses were based on 53,633 SNPs at a minor allele frequency of 0.05, filtered for linkage disequilibrium and without missing information. Individual points represent single trees coloured according to sampling sites. Subscripts depict whether the point is a seedling (S) or adult (A).

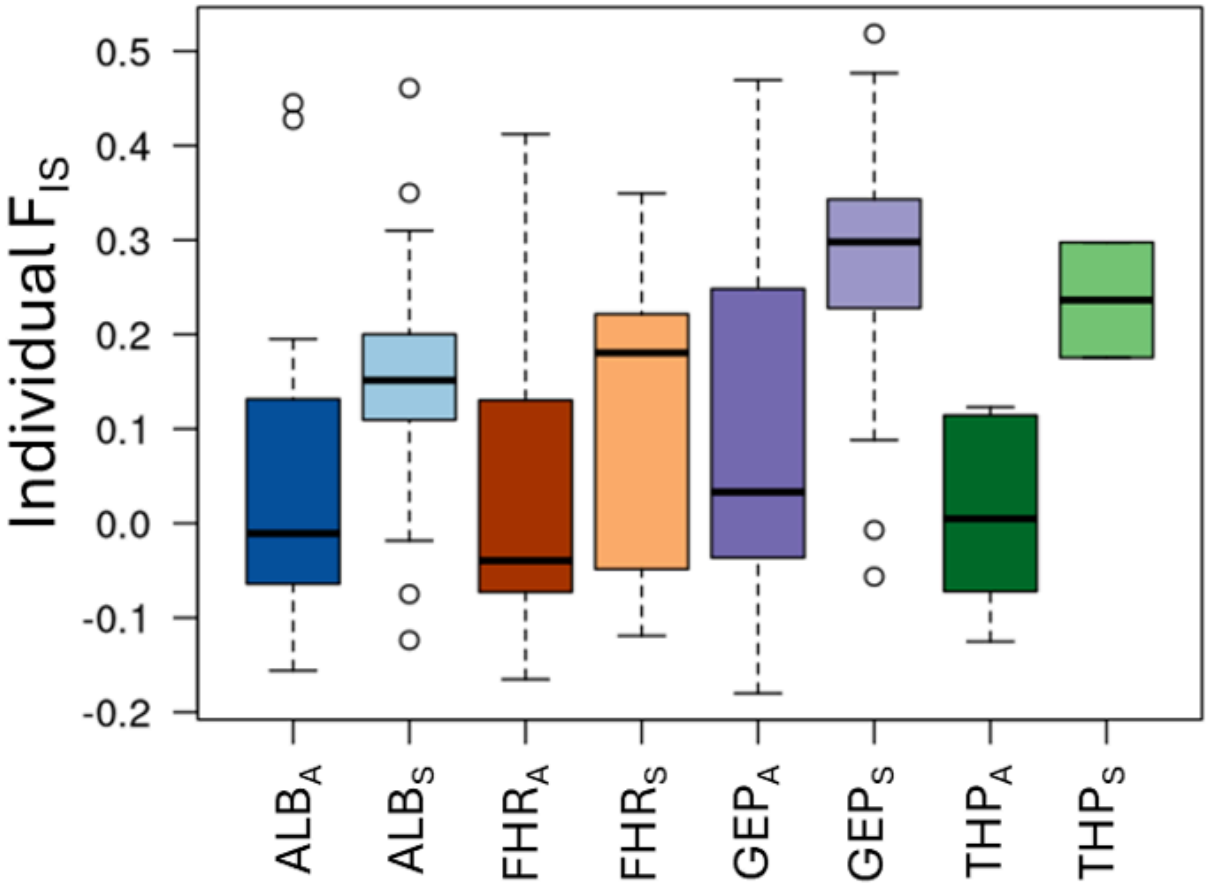

Figure 6

The distribution of individual inbreeding values across adult and seedling cohorts of S. maire for each sampling site. F_{IS} was calculated per individual using Weir and Goudet (2017)'s beta-estimators. Subscripts represent whether a data is shown for seed (S) or adult (A) trees.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppleMatLandscapeEcol.docx