Received: 28 October 2024 | Revised: 16 April 2025 | Accepted: 8 May 2025

DOI: 10.1002/ppp3.70042

Plants People Planet PPP Open Access

RESEARCH ARTICLE

Genetic diversity and population structure of pedunculate oaks (*Quercus robur*) in Wytham Woods

Tin Hang Hung | **Elias Formaggia** | **Lucy Morley** | **Keith Kirby** | **Roberto Salguero-Gómez** | **Ben C. Sheldon** | **John J. MacKay**

Department of Biology, University of Oxford, OX1 3RB, UK

Correspondence
Tin Hang Hung and John J. MacKay, Department of Biology, University of Oxford, OX1 3RB, United Kingdom.
Email: tin-hang.hung@biology.ox.ac.uk and john.mackay@biology.ox.ac.uk

Funding information
Oxford University Press John Fell Fund, Grant/Award Number: 0013514; Oxford University Press

Societal Impact Statement

Genetic diversity underpins forest resilience and adaptation, yet centuries of human activity have reshaped woodland ecosystems worldwide. We studied the population genetics of pedunculate oaks (*Quercus robur*) in Wytham Woods, UK, across stands with different ages and management histories. We found high within-stand diversity, low genetic differentiation among stands and signs of inbreeding in the youngest plantation. These findings suggest oak's adaptation and resilience to genetic erosion but highlight the result of intense management. Our results inform forest management, conservation policy and long-term ecological research by providing critical baseline data for monitoring keystone tree species across human-modified landscapes.

Summary

- Genetic diversity is fundamental for adaptation to changing environments. It is particularly important in forest trees because of their significant role in nature's contribution to people. However, forests have been significantly changed by human activities in the past centuries with potential effects on genetic diversity.
- This paper presents a population genetic study on the keystone producer, pedunculate oaks (*Quercus robur*) in Wytham Woods, UK, one of the most researched woodlands on Earth. We genotyped 210 trees and quantified levels of genetic diversity across stands with different histories and management regimes.
- We detected a weak population structure with the 218,567 SNPs, such that most genetic variation occurred within but not among stands, which included semi-natural woodland areas and plantations aged between 200 and 50 years ago. We also observed little difference in observed and expected heterozygosity among stand types, but detected signs of inbreeding in the youngest plantation. We discovered 161 genes (0.42%) that were under putative balancing selection (Tajima's $D > 2$), enriched in defence-related functions.
- We suggest that the life history traits of oak contribute to its resistance against genetic erosion. Preference for oaks as a timber tree and the tendency to use local

seed sources might have resulted in the homogeneous population structure. However, tree-to-tree differences may still harbour variation in putatively adaptive loci. Our study contributes crucial baseline information on the conservation and management of human-modified woodlands, in addition to supporting long-term ecological studies on many other species, which depend on this keystone oak species.

KEYWORDS
forest management, genetic diversity, genetic monitoring, land use change, natural history, population structure

1 | INTRODUCTION

Genetic diversity is fundamental for the adaptation of populations to changing environmental conditions, underpinning species' adaptability, persistence and evolution (Atkins & Travis, 2010; Frankham et al., 2019). Genetic diversity is particularly important for current populations, as new challenges emerge, such as pests, diseases, climate change, inbreeding and habitat degradation (Frankham, 2005; Holderegger et al., 2006). Understanding genetic diversity in forest trees is of broad importance, as they have a significant role in maintaining biodiversity (Stange et al., 2020), supporting community structure (Hughes et al., 2008) and stabilising ecosystem properties in natural environments (Ratnam et al., 2014a), in addition to sustaining productivity and resilience in forestry (Potter et al., 2017). Thus, genetic diversity in trees represents an essential aspect of nature's contribution to people (Des Roches et al., 2021).

Trees are among the most genetically diverse organisms on land (Savolainen & Pyhäjärvi, 2007). From an evolutionary perspective, they share unique life history traits and characteristics, including high levels of genetic diversity but usually low nucleotide substitution rates (Hamrick et al., 1992; Savolainen & Pyhäjärvi, 2007). Trees are predominantly outcrossing and thus maintain a high level of gene flow through efficient dispersal mechanisms (Koski, 1970), self-incompatibility (Williams & Savolainen, 1996) and severe inbreeding depression (Jiménez-Ramírez et al., 2021). Spatial analyses of population genetic structure in long-lived, woody trees reveal that most species have a higher within-population variation and a lower among-population variation compared to herbaceous and annual plants (Porth & El-Kassaby, 2014).

Forest trees are largely undomesticated but, human activities have influenced them with potential impacts on their adaptability. For example, forest management practices have modified tree density and demographics, connectivity and effective population size leading to effects on their genetic diversity (Ledig, 1992). In Europe, there is a long history of over-exploitation and loss of forest cover during medieval and early modern times (Bradshaw, 2004; Hosius et al., 2006). Forest cover has increased since the mid-20th century through both artificial and natural regeneration (Ratnam et al., 2014a). Simultaneously, human activities, such as species introduction, pollution and urbanisation, have exerted novel selective pressures with potentially long-term negative effects on fitness (Swain et al., 2007).

It has been speculated that forest plantations may have adverse genetic impacts on wild populations and pose genetic risks that are largely neglected (Fussi et al., 2016). Potential consequences could include loss of genetic variation, breakdown of adaptations, changes to genetic composition within populations and breakdown of population structure (Laikre et al., 2010). In England, the lack of genetic diversity in tree planting stocks and its potentially damaging impacts on forest health has led to renewed interest in understanding the genetic basis of resilience in trees (Forestry Climate Change Working Group, n.d.). However, a global synthesis of tree species has also shown cases where there is no significant erosion of genetic diversity despite widespread plantations (Ratnam et al., 2014b).

1.1 | Wytham Woods: a living laboratory

Wytham Woods are among the world's best-studied ecosystems and have been part of the UK Environmental Change Network since 1992, with other environmental and species population data continuously monitored for over 70 years (Sykes & Lane, 1996). Research conducted in Wytham Woods has contributed significantly to understanding wildlife populations and ecosystem dynamics, including the impact of climate variation on badgers (Macdonald et al., 2010), carbon dioxide fluxes in deciduous woodlands (Thomas et al., 2011) and climate-driven responses in great tits (Charmantier et al., 2008), among many others. One-third of its area is characterised as ancient woodland continuously wooded since at least 1,600 (Peterken, 1977) and traditionally managed by coppicing, where pedunculate oaks (*Q. robur*) are present as standard trees (Mihók et al., 2009). The rest consists of a mixture of self-sown stands and plantations with an oak component from c. 1810 through to the 1960s (Grayson & Jones, 1955; Savill et al., 2011).

Characterising the patterns of genetic diversity in Wytham Woods oaks is important for three main reasons (Kirby et al., 2014). First, oak has been the main timber tree in the Woods for centuries, but its importance is declining as a consequence of changes in land use and management regimes, which may impact its genetic diversity (Savill et al., 2011). Second, mature oak trees in the Woods display spatial variation in crown health (Cole et al., 2021) due to a form of chronic oak decline, in which the crown condition is seen to gradually decline over decades (Gagen et al., 2019a). Genetic differences may underpin the variation in susceptibility and resistance to chronic oak decline (Gagen et al., 2019b). Third, substantial and consistent tree-to-tree variation has been reported in tree phenology (Cole et al., 2021), which is likely due at least in part to genetic differences among individuals (Firmat et al., 2017). Variation in oak phenology has implications for the life cycles of birds and insects, which have been particularly well-studied at Wytham Woods (Cole et al., 2021). In contrast, the oaks themselves have remained comparatively understudied in this ecosystem.

The overarching aim of this study is to analyse the spatial pattern of genetic diversity across different stand types in *Q. robur* in Wytham Woods, which has adjacent semi-natural woodlands and plantations aged 200 to 50 years ago. We utilise a new reference genome assembled under the Darwin Tree of Life project and analyse 210 individual trees from five stand types across the Woods by using Genotyping by Sequencing (GbS), which has produced more reliable markers than RAD-seq (Ulaszewski et al., 2021). First, we analyse the population genetic structure. Second, we calculate and compare the genetic diversity among different stand types. Third, we scan for genome-wide outlier signals which may imply selection.

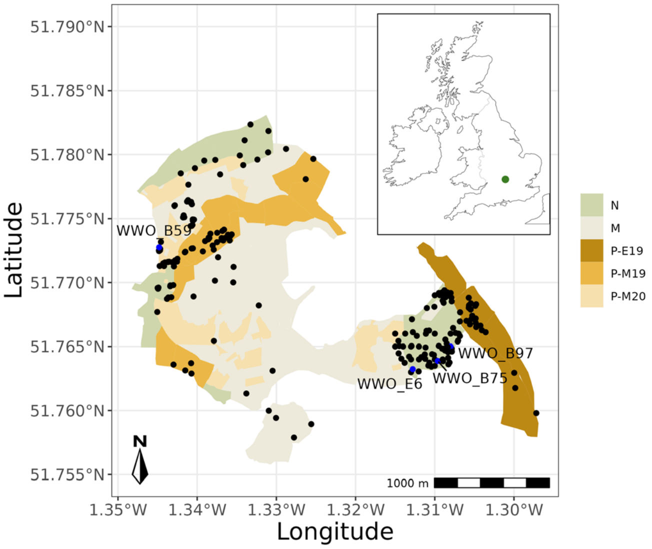

FIGURE 1 Map of Wytham Woods showing different woodland stand types (N = semi-natural woodland, M = mixed, P-E19 = early 19th-century plantation, P-M19 = mid-19th-century plantation, P-M20 = mid 20th-century plantation). The inset map at top right corner shows the approximate location of Wytham Woods. Black dots represent individual trees and blue dots represent trees that potentially arise from clonality.

2 | METHODS

2.1 | Study site, study species and sampling

Wytham Woods is a *c.*400 ha woodland in Oxfordshire, United Kingdom (~51.77° N, 1.32° W) (Savill et al., 2011). The soils are predominantly ill-drained Oxford Clay on the lower slopes where most of the oak occurs, while on the top of the hill shallow base-rich soils over limestone occur. In between is a narrow belt of more acidic sandy soil. The ancient woodland was historically managed as coppice-with-standards: about half has been left unmanaged since 1945, the 'semi-natural' stands, composed of native species that have not been planted, but have been managed in the past (Tansley, 1949). The rest has been replanted at various times including with stands of oak. Other woodland has developed on farmland, mainly former pasture over the last 200 years, through a mixture of planting and natural colonisation (Grayson & Jones, 1955; Savill et al., 2011).

Alongside the oaks, the woodland is dominated by European ash (*Fraxinus excelsior*), sycamore maple (*Acer pseudoplatanus*), birch (*Betula* spp.) and European beech (*Fagus aylvatica*) (Kirby et al., 2014). The mature oaks are almost exclusively *Q. robur*, except sessile oak (*Quercus petraea*), and northern red oak (*Quercus rubra*), and a few old Turkey oak (*Quercus cerris*), planted after 1945.

The oaks in the ancient semi-woodland are the relict standards from the coppice-with-standards management. They are mainly c. 150–200 years old and self-grown as there was relatively little planting of oaks before 1800. Three broad phases of subsequent plantings have been identified and mapped. The fifth Earl of Abingdon, who owned the Woods between 1799 and 1854, initiated a massive programme of restructuring and planting following the Enclosure Acts in 1814 and 1816 with some further planting by his successors during the later part of the 19th century (Grayson & Jones, 1955). In the 20th century, there was heavy felling during the World Wars and after WWII the University of Oxford's Department of Forestry organised widespread planting of oaks and other species to fill the gaps created and to expand the woodland extent. In some areas, the planted oaks are mixed with others that appear to be natural colonisation (Grayson & Jones, 1955; Savill et al., 2011).

We explored what genetic differences could be detected in these five types of oak stands of different origins across the Woods (Figure 1): (N) ancient semi-natural woodland where trees are self-sown; (M) mixture of self-sown and planted oaks; (P-E19) early 19th-century plantation; (P-M19) mid-to-late 19th-century plantation; and (P-M20) mid-20th-century plantation.

Foliage of *Q. robur* was sampled across 210 random mature trees from a broad range of semi-natural stands and plantations within the woods between June and September of 2023 (**NCBI BioSample records under BioProject PRJNA1135559**). Sample sizes were N = 84, M = 33, P-E19 = 27, P-M19 = 33, P-M20 = 33. For each tree, 3–5 healthy, non-senescent leaves were collected using either extendable secateurs or an arborist slingshot. Foliage for each sample was placed in a paper envelope, kept in a sealed plastic bag containing silica beads for desiccation and preservation, and then moved to a −20°C freezer at the end of the day.

The metadata of the samples can be found in Table S1.

2.2 | Library preparation and GbS

Dried foliage tissues were ground using a Retsch TissueLyser II (Qiagen, United Kingdom) at a frequency of 25 s^{-1}. Genomic DNA was isolated and purified by using the DNeasy Plant Mini Kit (Qiagen, United

Kingdom), according to the protocol with two slight modifications for effective lysis of plant tissues. First, a mix of 370 μl of Lysis Buffer and 30 μl of protease were used at the lysis step. Second, the incubation time was increased to 1 hour. DNA quantity and quality were assessed using NanoDrop 2000 (Thermo Scientific, United States) and Qubit dsDNA Broad Range assay, respectively (Thermo Scientific, United States).

Samples of ∼200 ng genomic DNA were normalised in 10 μl of TE buffer and sent to the Genomic Analysis Platform, Institute of Integrative and Systems Biology, Université Laval (Quebec, Canada) for the preparation of GbS libraries. DNA was digested with PstI (CTGCA↓G), NsiI (ATGCA↓T) and MspI (C↓CGG) according to the 3D-GbS protocol for library preparation (de Ronne et al., 2023). Sub-libraries were barcoded and pooled to equimolarity, then sequenced on an Illumina NovaSeq 6,000 system with a paired-end mode of 150 bp at the Génome Québec (Montreal, Canada).

2.3 | Variant calling

Raw reads were demultiplexed with Sabre 1.0 (Joshi, 2013) and trimmed with Cutadapt 1.18 (Martin, 2011) to remove the adaptors and sequences shorter than 50 bp. Reads were aligned against dhQueRobu3.1 (GCA_932294415.1), the reference genome of *Q. robur* from the Darwin Tree of Life project, using bwa-mem 2.2.1 (Md et al., 2019). The alignments were indexed and sorted using SAMtools 1.12 (Danecek et al., 2021). Variant pileup and calling were performed with bcftools 1.19 (Danecek et al., 2021). Variants were filtered with the proportion of missing data of 0.2 and MAF of 0.01 using VCFtools 0.1.16 (Danecek et al., 2011). Linkage disequilibrium was estimated using PLINK v1.90b6.21 (Purcell et al., 2007), and one SNP per pair was removed based on a r^2 threshold of 0.5 in a genomic window of 50 Kbp and a sliding window of 5 bp.

Estimates of heterozygosity values from SNPs are context dependent due to sample size and population differentiation (Schmidt et al., 2021; Sopniewski & Catullo, 2024). Thus, while the filtering above produces sufficient markers for assessing population structure, we also curated another set of SNPs for estimating 'autosomal heterozygosities' (Hohenlohe et al., 2010). Variants were filtered so that no missing data were allowed but included all monomorphic markers for unbiased estimation of genetic diversity.

2.4 | Estimation of inter-individual relationships

Pairwise relatedness of individual trees was assessed using VCFtools 0.1.16 (Danecek et al., 2011) based on the KING inference (--relatedness2). According to the manual, estimated kinship coefficients of > 0.354, 0.177–0.354, 0.0884–0.177 and 0.0442–0.0884 correspond to duplicate or monozygotic twin, 1st-degree, 2nd-degree and 3rd-degree relationships, respectively. First-degree relationship means no other members in the direct heredity between two individuals, such as parent and child. We also assessed the spatial correlation in the genetic relatedness among individuals by computing Spearman's rho between pairwise genetic relatedness and geographic distance, where geographic distance was calculated using sf 1.0–17 (Pebesma, 2018).

2.5 | Population genetic structure

We ran a sparse non-negative matrix factorisations (sNMF) algorithm in the LEA 3.10 R package (Frichot & François, 2015) to estimate the number of discrete genetic clusters (*K*) using 10 repetitions for each value of *K* from 1 to 10. This algorithm was considered more statistically robust to departures and violations from classical population genetic model assumptions and computationally efficient than likelihood-based approaches such as STRUCTURE and ADMIXTURE (Frichot et al., 2014). The optimal *K* was selected for the lowest cross-entropy.

We also cross-validated the population structure using principal component analysis (PCA) using PLINK v1.90b6.21. Eigenvalues were used to plot the scree plot to investigate the proportion of variance explained, and the first two eigenvectors (principal components) were then visualised.

We conducted an analysis of molecular variance (AMOVA) to partition the genetic variation within and among stand types using poppr 2.9.6 (Kamvar et al., 2014), with 1,000 replications for the Monte Carlo procedure.

2.6 | Genetic diversity among different stand types

We calculated the observed heterozygosity (H_O), the expected heterozygosity (H_E) and the inbreeding coefficient (F_{IS}) using hierfstat 0.5–11 (Goudet, 2005) with bootstrapping, for each stand type (N, M, P-E19, P-M19 and P-E20). We also assessed deviations from Hardy–Weinberg Equilibrium (HWE) for each SNP for each stand type using pegas 1.3 (Paradis & Barrett, 2010), with 1,000 replications for the Monte Carlo procedure. We also calculated and reported the autosomal aH_O, aH_E and aF_{IS} which consider all polymorphic and monomorphic sites for unbiased estimations.

2.7 | Identification and enrichment analysis of genes under putative selection

We identified genes that were potentially under selection using the Tajima's D statistics calculated by VCFtools 0.1.16 (Danecek et al., 2011). We retained only those with an absolute value of Tajima's D larger than 2. We visualised the results using a Manhattan plot. We also obtained Araport11 annotation (Cheng et al., 2017) for the gene models in dhQueRobu3.1 with BLASTx 2.11.0 + (Camacho et al., 2009). We also calculated the nucleotide diversity (π) for each gene using VCFtools 0.1.16.

We ranked the genes with descending values of Tajima's D value and performed enrichment analysis on these genes under putative selection with clusterProfiler 4.0 (Wu et al., 2021) on both Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG).

3 | RESULTS

3.1 | GbS results and processing

We obtained 448.10 M × 2 reads in paired-end mode from Illumina GbS, representing 135.33 Gbp in total, from 210 individual oaks. After demultiplexing and cleaning, we retained an average of 2.01 M (± 46.14 K) × 2 paired-end clean reads from each sample, whereas 14.53 M × 2 of sequences were unclassified (3.24%). More than 90% of these reads had quality scores ≥ 30, implying successful data cleaning. The read data were deposited in the NCBI Sequence Read Archive (SRA: SRR29822091–SRR29822300) under BioProject PRJNA1135559. After filtering, we obtained a final set of 218,567 SNPs (Data S1), which we utilised for pedigree estimations and analyses of population structure and genetic diversity.

3.2 | Pedigree estimation

Among the 21,945 (C_2^{210}) pairwise relationships between the 210 individual trees, most of the pairwise relatedness estimates were negative, with a mean of −0.092 (Figure 2a, Table S2). The vast majority of relationships (98.68%) were classified as unrelated, followed by 3° relationship (0.29%), 2° relationship (0.06%), 1° relationship (0.02%), 0° relationship (0.01%) (Figure 2b,c). We detected only two pairs (0.0091%) of clonal trees: WWO_B59 and WWO_E6, which shared a relatedness of 0.43, and also WWO_B75 and WWO_B97, which shared a relatedness of 0.37 (Figure 1).

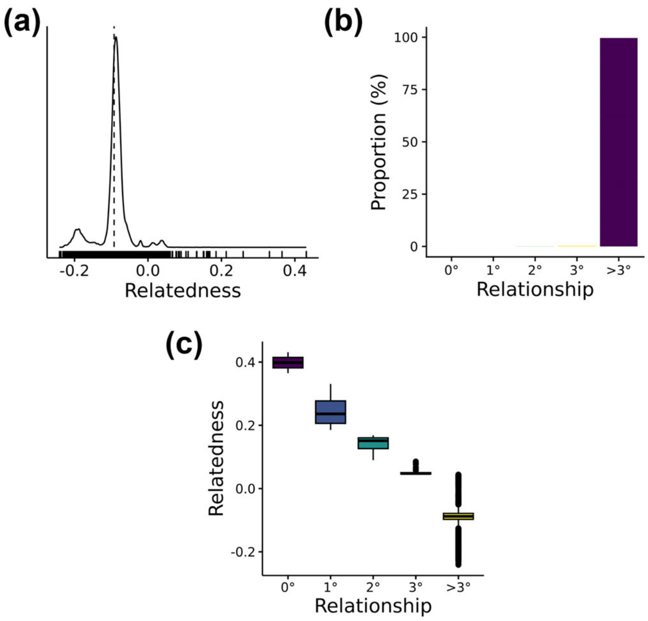

FIGURE 2 **(a)** Density distribution of pairwise relatedness among individual trees. **(b)** Proportion of pairwise relatedness among different categories of relationships, where relatedness coefficients of > 0.354, 0.177–0.354, 0.0884–0.177, 0.0442–0.0884, < 0.0442 correspond to duplicate or monozygotic twin (0°), 1st-degree (1°), 2nd-degree (2°), 3rd-degree (3°) and unrelated (>3°) relationships, respectively. **(c)** Boxplot of pairwise relatedness among different categories of relationships.

Three trees, 20,729, P6B and B199, had a slightly positive average relatedness with all other trees, of 0.0383, 0.0306 and 0.0118, respectively (Figure S1). All other trees had a negative averaged relatedness with others as expected. There was a very weak negative relationship between the pairwise geographic distance and genetic relatedness (Spearman's rho = −0.053, $P = 3.33\text{e}{-15}$), suggesting little spatial correlation (Figure S2).

3.3 | Population genetic structure

Cross-entropy analysis revealed that $K = 3$ was the most informative clustering in explaining the population genetic structure using the sNMF analysis (Figure 3a), while $K = 2$ and 4 were closely similar. The cross-entropy escalated quickly after $K \geq 5$. However, a complementary PCR-based approach revealed no clear spatial genetic structure, as the scree plot decayed quickly and remained flat (Figure 3b,c). We visualised the optimal $K = 3$ sNMF pattern in Figure 3d and the raw ancestry coefficients can be found in Table S1. Figure S3 visualised the $K = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ scenarios. As K increased, the genetic clustering broke down for only a small portion of trees, with the remaining large portion unexplained, revealing a very unbalanced hierarchical population structure.

This very weak population genetic structure indicated that there was very little genetic differentiation between the stand types. Only 1.34% of the genetic variation (df = 4, SS = 97481.14, MS = 24370.29) was partitioned between stands, and the other 98.66% (df = 200, SS = 3193578.25, MS = 15967.89) was between trees within stands. The genetic variance components were statistically significant ($P = 0.001$, Monte Carlo test, $n = 1{,}000$).

3.4 | Genetic diversity within different stand types

The levels of observed and expected heterozygosity for oak trees in Wytham Woods were as follows: mean $H_O = 0.120$, mean $H_E = 0.133$, mean $aH_O = 0.105$ and mean $aH_E = 0.102$. There was no substantial difference in H_O or aH_O among different stand types (Figure 4a,d), where M19 plantation had slightly higher values than the others of 0.123 and 0.106, followed by semi-natural woodland (0.121 and 0.104), mixed plantation (0.120 and 0.105), E19 plantation (0.118 and 0.104) and M20 plantation (0.117 and 0.106). There was also no substantial difference in H_E or aH_E among different stand types (Figure 4b,e), where M19 plantation again had slightly higher values of 0.137 and 0.101, and was followed by M20 plantation

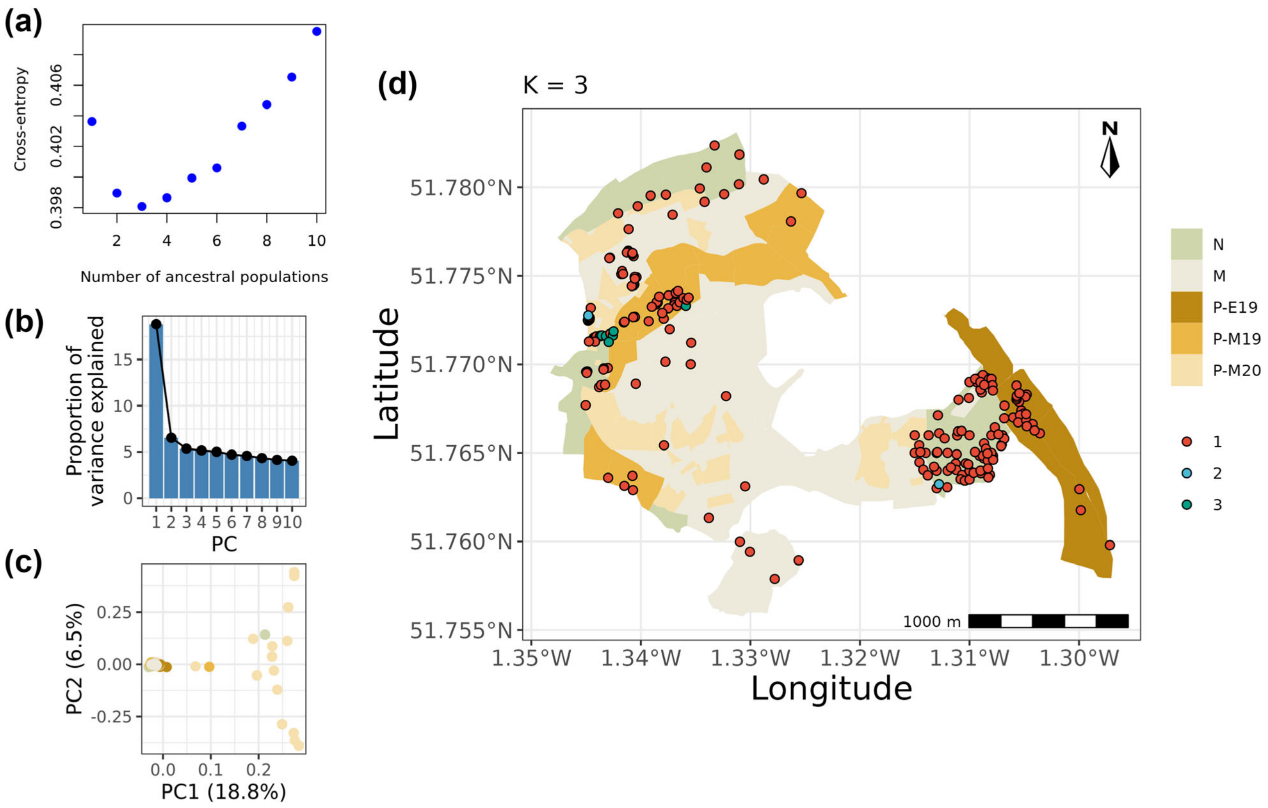

FIGURE 3 Population genetic structure of pedunculate oaks in Wytham Woods across different stand types and different numbers of clusters K predicted by sparse non-negative matrix factorisation (sNMF). **(a)** Cross entropy with increasing K. **(b)** Scree plot of SNP-based PCA. **(c)** Population structure from SNP-based PCA visualised using the first two principal components. **(d)** The optimal K = 3 was determined with the lowest cross-entropy value.

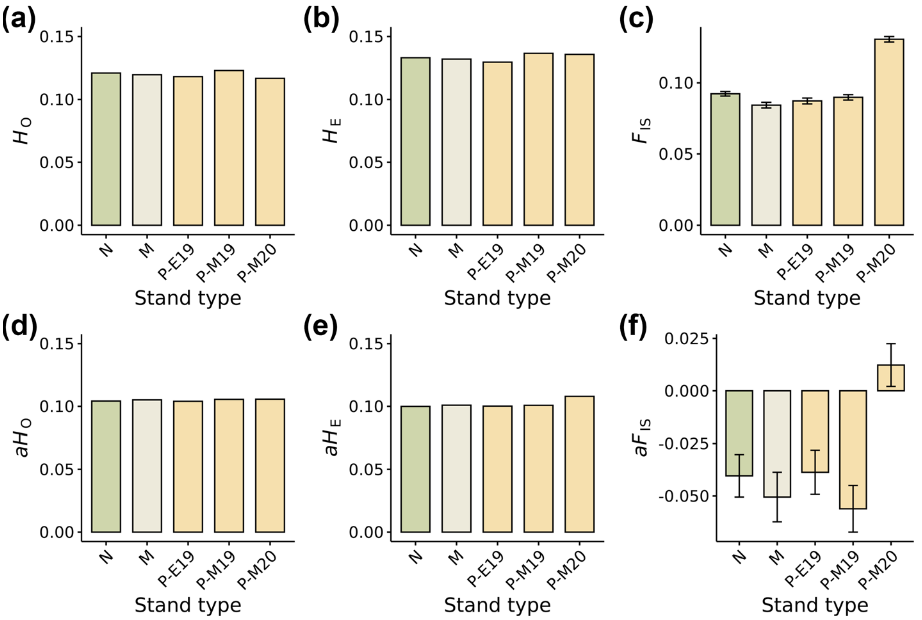

FIGURE 4 **(a)** Observed heterozygosity (H_O), **(b)** expected heterozygosity (H_E) and **(c)** inbreeding coefficient (F_{IS}) of all loci and **(d–f)** their autosomal estimates across different stand types: semi-natural (N), mixed (M), early-19th-century (P-E19), mid-to-late-19th-century (P-M19) and mid-20th-century plantations (P-M20). For inbreeding coefficients, the error bar denoted 95% confidence interval determined with bootstrapping.

(0.136 and 0.108), semi-natural woodland (0.133 and 0.100), mixed plantation (0.132 and 0.101) and E19 plantation (0.130 and 0.100).

When using F_{IS}, we detected signs of inbreeding in all stand types, where F_{IS} differed significantly from zero (95% confidence interval, bootstrapping) (Figure 4c). The highest inbreeding was observed in the most recent M20 plantation, with F_{IS} of 0.130, followed by semi-natural population (0.0923), M19 plantation (0.0898), E19 plantation (0.0873) and mixed plantation (0.0843). However, when using aF_{IS}, we only detected signs of inbreeding in the most recent M20 plantation, with an aF_{IS} of 0.0123. Other stand types had aF_{IS} with 95% confidence intervals overlapped with zero (Figure 4f).

3.5 | Identification and enrichment analyses of genes under putative selection

We computed the Tajima's D for all 35,638 gene models and detected 161 genes that were under putative selection ($|D| > 2$). In

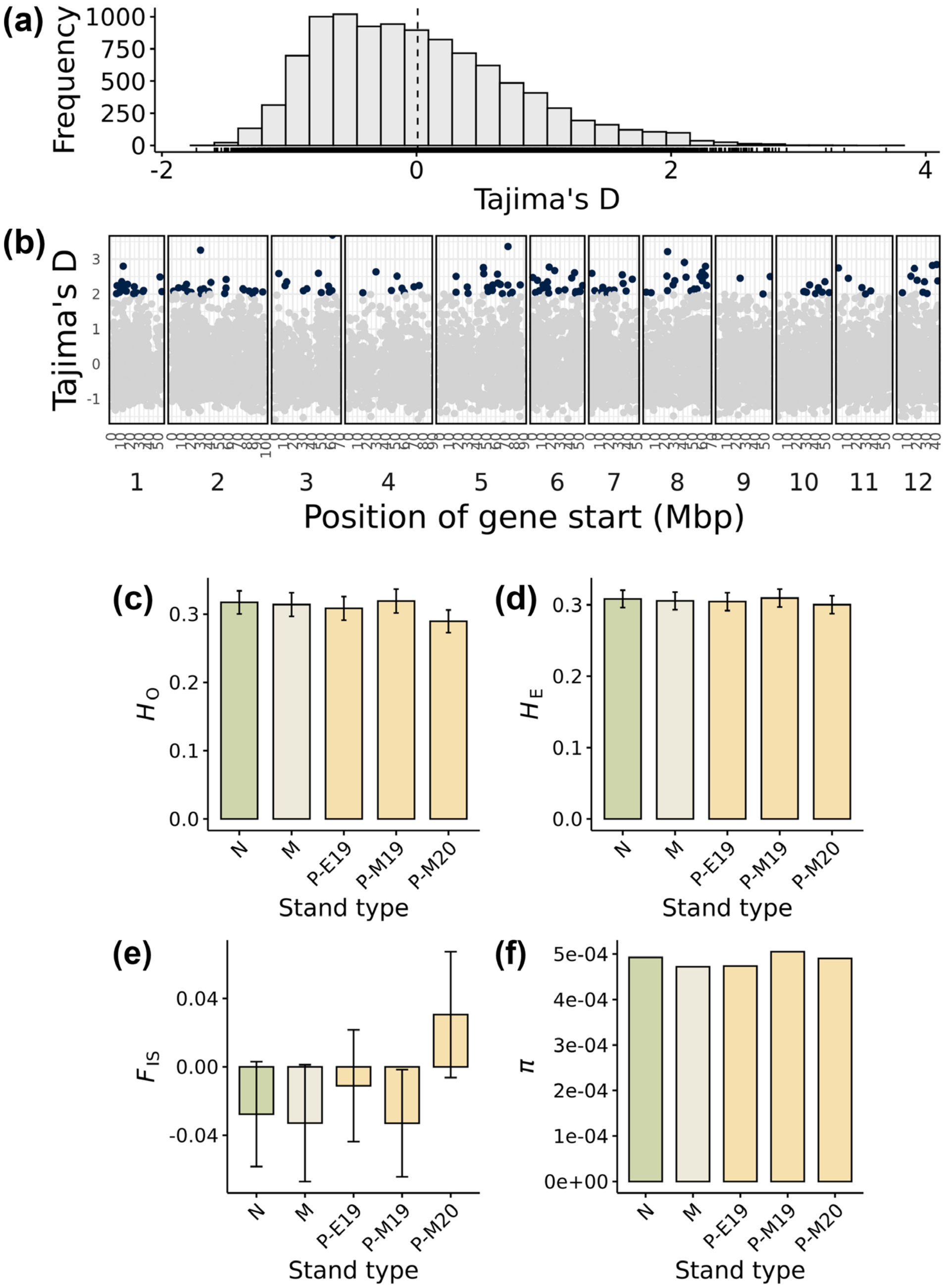

FIGURE 5 **(a)** Frequency distribution histogram of Tajima's D values across 35,638 genes. **(b)** Manhattan plot of Tajima's D values. Dark blue dots denote genes under putative selection ($|D| > 2$). **(c)** Observed heterozygosity (H_O), **(d)** expected heterozygosity (H_E) and **(e)** inbreeding coefficient (F_{IS}) and **(f)** nucleotide diversity (π) of genes under putative selection across different stand types: semi-natural (N), mixed (M), early-19th-century (P-E19), mid-to-late-19th-century (P-M19) and mid-20th-century plantations (P-M20). For the inbreeding coefficient, the error bar denoted a 95% confidence interval determined with bootstrapping.

particular, these 161 genes were under balancing selection ($D > 2$) while none of the genes were under directional selection ($D < -2$). Tajima's D ranged from -1.73 to 3.69, with a median of -0.0092 (Table S3).

The levels of observed and expected heterozygosity for the genes under putative selection indicated low to medium levels of genetic diversity but slightly higher than for all loci (mean $H_O = 0.310$, $H_E = 0.306$). In addition, there was considerable difference in H_O among different stand types (Figure 5c,f). It was highest in the M19 plantation with 0.319, followed by semi-natural woodland (0.318), mixed plantation (0.314), E19 plantation (0.309) and M20 plantation (0.290). The exact same ranking of values was observed for H_E (Figure 5d,g), where M19 had the highest value (0.309), then semi-natural woodland (0.309), mixed plantation (0.306), E19 plantation (0.305) and M20 plantation (0.300). We used nucleotide diversity (π) instead of autosomal heterozygosities to provide an unbiased estimate for per-gene statistics. π followed a similar trend where it is highest in the M19 plantation (0.00051), but followed by semi-natural woodland (0.000493), M20 plantation (0.000491), E19 plantation (0.000474) and mixed plantation (0.000472). When considering F_{IS}, we detected potential signs of inbreeding in the most recent M20 plantation with a positive F_{IS} of 0.0265, however, its 95% confidence interval overlapped with zero (Figure 5e).

We were able to assign an annotation to 34,815 out of 35,638 gene models (97.69%) of dhQueRobu3.1 with Araport11. This was followed by enrichment analysis of gene ontology (GO) which identified 61 Biological Process (BP) terms, 14 Cellular Component (CC) terms and 19 Molecular Function (MF) terms were over-represented among the genes under putative selection ($P < 0.05$, after Benjamini-Hochberg correction). Biological processes included protein phosphorylation (GO:0006468), phosphorylation (GO:0016310), response to bacterium (GO:0009617), defence response to bacterium (GO:0042742) and response to fungus (GO:0009620) (Figure 6a, Table S4). Cellular components included cell wall (GO:0005618), plant-type cell wall (GO:0009505), cell–cell junction (GO:0005911), plasmodesma (GO:0009506) and cell junction (GO:0030054). (Figure 6b, Table S5). Molecular functions included protein kinase activity (GO:0004672), protein serine/threonine kinase activity (GO:0004674), kinase activity (GO:0016301), phosphotransferase activity, alcohol group as acceptor (GO:0016773) and transferase activity, transferring phosphorus-containing groups (GO:0016772) (Figure 6c, Table S6).

We detected six enriched KEGG pathways among the genes under putative selection ($P < 0.05$, after Benjamini-Hochberg correction), including biosynthesis of secondary metabolites (ath01110), plant-pathogen interaction (ath04626), motor proteins (ath04814), phenylpropanoid biosynthesis (ath00940) and cysteine and methionine metabolism (ath00270) (Table S7).

4 | DISCUSSION

Forest trees face a range of challenges from changing environmental conditions and human activities which have been shown to alter their genetic diversity significantly. Understanding of the standing genetic diversity in pedunculate oaks (*Q. robur*) is thus essential to monitor and predict the adaptability of this species across the shifting climate of the Anthropocene. The present study constitutes the first population genetic study of pedunculate oaks in Wytham Woods and provides a historical perspective of management regimes that may underpin the differences among stand types.

4.1 | Small difference in genetic diversity among stand types

We detected minor differences in genetic diversity among the five different stand types of Wytham Woods comprised of semi-natural woodland and plantations of different ages. We only observed a significant signal of inbreeding ($aF_{IS} > 0$) in the most recent plantations established mid-20th century (P-M20). The P-M20 trees possessed an excess of homozygotes ($aH_O < < aH_E$) but nonetheless showed the

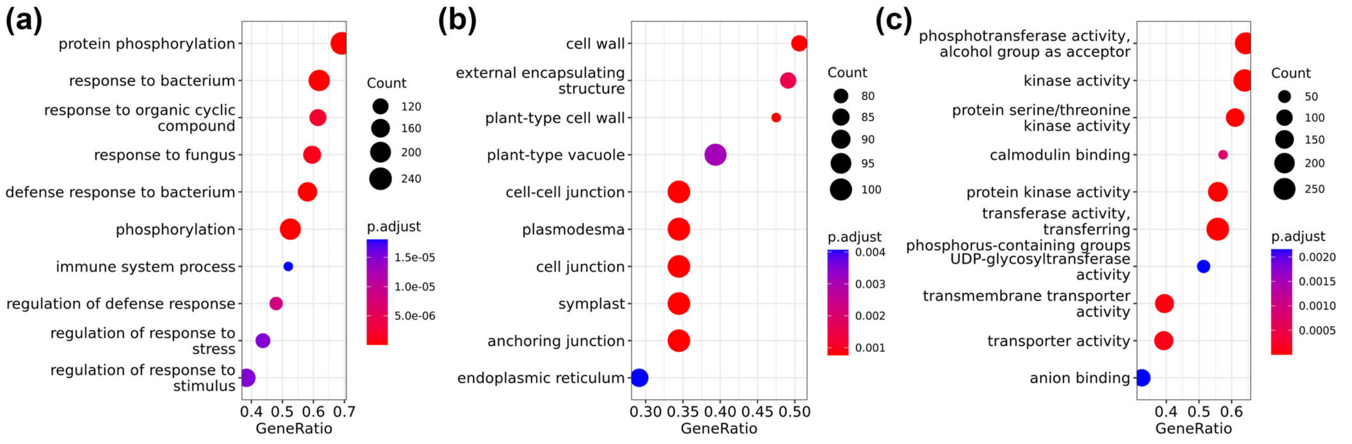

FIGURE 6 Dot plots of top 10 enriched gene ontology (GO) terms associated with negative Tajima's D values in categories of **(a)** biological processes, **(b)** cellular components and **(c)** molecular functions, whereas gene models of dhQueRobu3.1 were annotated with Araport11.

highest aH_O and aH_E among all stand types. There are two potential reasons that may explain these observations. First, it may be a classic example of Wahlund effect, where trees (or subpopulations) with different allelic frequencies are combined during the plantation process, thus maintaining a comparable level of genetic diversity as other subpopulations but showing an excess of homozygotes (Duminil et al., 2009). Second, oak trees only mature and start to produce acorns regularly at 40 to 50 years old (Woziwoda et al., 2019). Therefore, the germplasm that was used for establishing P-M20 is likely to come from the youngest trees at the time with limited reproductive capacity and low levels of gene flow (Ruiz-González et al., 2023).

It has been widely speculated that forest management practices can have significant effects on genetic diversity and effective population size (Ratnam et al., 2014b). However, global synthesis of research has shown that although genetic diversity can be adversely affected, such as in coppiced oak species (*Quercus* spp.) in the Mediterranean region which has fragmented populations (Ortego et al., 2010; Valbuena-Carabaña et al., 2008), no decline in genetic diversity was found between semi-natural woodlands and plantations in several cases, such as European beech (*Fagus sylvatica*) (Rajendra et al., 2014), black spruce (*Picea mariana*) (Rajora & Pluhar, 2003), lodgepole pine (*Pinus contorta*)(S. E. Macdonald et al., 2001), Douglas fir (*Pseudotsuga menzeisii*) (Neale, 1985). These species share several important life history traits in common with pedunculate oak including long generation time, longevity and effective gene flow due to wind pollination, which might have contributed to the resistance to genetic erosion (Ratnam et al., 2014b).

Our study also highlighted the importance of standardised methods for the quantification of heterozygosity. At first glance, the genetic diversity of pedunculate oaks in Wytham Woods (mean $H_O = 0.120$ and mean $H_E = 0.133$) may seem lower than previous reports in the same species, such as in the eastern range of Russia (mean $H_E = 0.669$–0.773) (Semerikova et al., 2023), the southern range of Serbia (mean $H_E = 0.740$–0.793) (Kesić et al., 2021) or Croatia (0.195–0.962) (Popović et al., 2023). However, such differences are to be expected when contrasting genome-wide SNP marker data as in the present study, and more targeted marker sets, which are usually selected as polymorphic, as reported in oaks including either isozymes (Gömöry et al., 2001), microsatellites (or simple sequence repeats SSR) (Kesić et al., 2021; Popović et al., 2023; Semerikova et al., 2023) or small sets (< 100) of SNP markers (Degen et al., 2021). The heterozygosity levels cannot be compared directly because of the difference in the marker types. For example, SSRs harbour exceptionally high mutation rates and polymorphism levels when compared to SNPs (Hansson & Westerberg, 2002), and higher estimates of heterozygosity across several species (Fischer et al., 2017; Sunde et al., 2020; Zimmerman et al., 2020).

There are a few tree species that have been assessed with both genome-wide SNP analyses and SSRs for genetic diversity. A GbS study on Norway spruce reports a H_O of 0.237 and a H_E of 0.271 (Korecký et al., 2021) for SNPs but the same authors report a mean H_O of 0.585 and a mean H_E of 0.722 when using SSR (Bínová et al., 2020). Similarly, work in Siamese rosewood reported a mean H_O of 0.56 and a mean H_E of 0.57 with SSR, but a mean H_O of 0.18 and a mean H_E of 0.18 using RAD-seq SNP (Hartvig et al., 2020). Even when considering 130 SNP loci in Russian populations of mixed *Q. robur* and *Q. petraea*, H_O was between 0.259 and 0.291 and H_E was between 0.219 and 0.272 (Blanc-Jolivet et al., 2020), providing a strong contrast with SSR findings in the same geographic area.

4.2 | Weak population structure and history of Wytham Woods

The nature and intensity of woodland management has changed considerably in Wytham Woods, over the last few hundred years as the coppice management of the 18th to early 19th centuries was gradually replaced by high forest and plantations in the 20th century (Grayson & Jones, 1955).

However, we found only weak population structure in Wytham oaks, such that most of the genetic variation occurs between individual trees but not among stand types. First, the lack of difference in genetic diversity across stands dating from different periods may reflect a tendency to use local seed sources. Estates often had their own nurseries and might collect acorns from their best stands (Whittet et al., 2017). Second, despite the overwhelming preference for oak as a timber tree, there may only be a strong impact at loci controlling economic traits, maintaining the overall homogeneity of genetic makeup (Finkeldey & Ziehe, 2004). The weak structure was also reflected in the pedigree analysis such that most individuals were completely unrelated. The exception of two clonal pairs might likely arise from artificial propagation materials in plantation efforts.

The signs of inbreeding in younger stands absent in older stands in Wytham Woods reflect a potential mechanism for the progressive elimination of inbred trees. Despite oaks being largely outcrossing, the potential high frequency of pollination at short distances (Vranckx et al., 2014) and the homogeneity of population structure might still contribute to potential inbreeding in saplings in younger stands. A similar pattern was also observed in an outcrossing dominant tree species *Castanopsis sieboldii*, where it was suggested that pronounced inbreeding depression would result in mortality or reduced growth, thus the level of inbreeding would decay in subsequent generations (Nakanishi et al., 2015).

4.3 | Balancing selection on resistance-related genes

Genomic scans can offer important insights into natural selection and local adaptation. Given the weak population structure and limited forest area, the outlier test based on the empirical distribution of F_{ST} and departure from Hardy–Weinberg equilibrium might fail to detect loci under selection (Waples & Gaggiotti, 2006; Wigginton et al., 2005). Therefore, we utilised the Tajima's D which analysed the site frequency spectrum of alleles across the population of Wytham Woods.

When Tajima's D is negative, it suggests an excess of low-frequency alleles, indicative of population expansions or selective sweeps from positive natural selection (Kawecki et al., 2021). Conversely, positive Tajima's D reflects an excess of intermediate alleles from balancing selection which may indicate an adaptive advantage in a fluctuating environment (Ryynänen & Primmer, 2004).

Our study revealed significantly enriched gene ontology terms and pathways that include genes related to pathogen response and immune reactions, which shared negative Tajima's D values consistent with balancing selection. The most highly enriched pathway, biosynthesis of secondary metabolites, which produce diverse flavonoids, terpenoids, alkaloids and phenolics, critical for plant defence (Zaynab et al., 2018). Pathogens and hosts often engage in a rapid evolutionary arms race driven by strong selection, such that hosts evolve resistance and pathogens evolve means to evade or tolerate those resistance traits (Bishop et al., 2000). Balancing selection particularly operates on the host's resistance and its genomic architecture across the tree of life (Bourgeois et al., 2021; Cornetti et al., 2024; Delph & Kelly, 2014; Koenig et al., 2019). In plants, balancing selection in resistance-related genes is observed in both model and crop species, such as in *Arabidopsis* and its relative *Capsella rubella* (Koenig et al., 2019; Q. Wu et al., 2017), rice (Gladieux et al., 2024), grass (Zhang et al., 2019) and field mustard (O'Hara et al., 2021). Balancing selection in resistance-related genes is also observed in forest trees (Van der Hoorn et al., 2002), which may reflect their exposure to a large diversity of pathogens during their lifetime (Naidoo et al., 2019).

4.4 | Implications for conservation of oaks and future studies in Wytham Woods

Intraspecific variation, as a major component of biodiversity, has received relatively little attention when formulating conservation plans for wild species and ecosystem services management (Heuertz et al., 2023; Laikre, 2010). The genetic consequences of land use change and human-induced climate change in many species and populations still remain poorly characterised (Aguilar et al., 2008; Ratnam et al., 2014b). The recent Kunming–Montreal Global Biodiversity Framework, from the Convention on Biological Diversity meeting in 2022 (COP15), sets out new guidelines to monitor, manage and report on the genetic status of species in local, regional and global programmes, where intraspecific genetic variation is a key indicator (Hoban et al., 2024; Mastretta-Yanes et al., 2024).

There is a very limited number of population genetic studies on oaks within the UK. They either do not report genetic diversity metrics (Nocchi et al., 2022) or use an incomparable one, such as chloroplast DNA (Cottrell et al., 2002) or microsatellite markers (Cottrell et al., 2003). As discussed above, the choice of markers for estimating genetic diversity can greatly influence the results. We, therefore, advise future studies to use unbiased parameters, such as autosomal heterozygosities (Hohenlohe et al., 2010), which consider both monomorphic and polymorphic sites, so that we can continue to monitor and compare genetic diversity meaningfully, especially when they are important parameters in forest conservation and management (Weeks et al., 2011).

By combining studies of the genetic diversity in oaks in Wytham Woods with its known history, we are able to suggest possible mechanisms for how changes in land use and forestry practices affect genetic diversity. In return, the many continuing, long-term ecological studies in the Woods may benefit from a better understanding of the genetic variation in a keystone species. Moreover, we believe that our study of the genetic diversity of oak trees in Wytham Woods may be relevant for other co-occurring species, opening new research avenues that can benefit both conservation and productive forestry.

AUTHOR CONTRIBUTIONS

T.H.H.: designed the study, conducted the bioinformatic analyses and drafted the manuscript. E.F.: collected and processed the samples and revised the manuscript. L.M.: collected and processed the samples and revised the manuscript. K.K.: processed the spatial data and revised the manuscript. R.S.: designed the study, collected the samples and revised the manuscript. B.C.S.: designed the study, collected the samples and revised the manuscript. J.J.M.: designed the study, collected the samples and revised the manuscript.

ACKNOWLEDGEMENTS

This work was supported by funding to T.H.H., R.S., B.C.S., and J.J.M. from the Oxford University Press John Fell Fund (0013514). We also thank Sienna Rattigan for assistance in the fieldwork.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The read data from Illumina sequencing were deposited in the NCBI Sequence Read Archive (SRA: SRR29822091–SRR29822300) under BioProject PRJNA1135559, along with associated BioSample records.

ORCID

Tin Hang Hung https://orcid.org/0000-0001-9853-2053
Roberto Salguero-Gómez https://orcid.org/0000-0002-6085-4433
Ben C. Sheldon https://orcid.org/0000-0002-5240-7828
John J. MacKay https://orcid.org/0000-0002-4883-195X

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*(24), 5177–5188. https://doi.org/10.1111/J.1365-294X.2008.03971.X

Atkins, K. E., & Travis, J. M. J. (2010). Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology*, *266*(3), 449–457. https://doi.org/10.1016/J.JTBI.2010.07.014

Bínová, Z., Korecký, J., Dvořák, J., Bílý, J., Zádrapová, D., Jansa, V., & Lstibůrek, M. (2020). Genetic structure of Norway spruce ecotypes studied by SSR markers. *Forests*, *11*(1), 110. https://doi.org/10.3390/F11010110

Bishop, J. G., Dean, A. M., & Mitchell-Olds, T. (2000). Rapid evolution in plant chitinases: Molecular targets of selection in plant-pathogen coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(10), 5322–5327. https://doi.org/10.1073/PNAS.97.10.5322

Blanc-Jolivet, C., Bakhtina, S., Yanbaev, R., Yanbaev, Y., Mader, M., Guichoux, E., & Degen, B. (2020). Development of new SNPs loci on *Quercus robur* and *Quercus petraea* for genetic studies covering the whole species' distribution range. *Conservation Genetics Resources*, *12*(4), 597–600. https://doi.org/10.1007/S12686-020-01141-Z/TABLES/2

Bourgeois, Y., Fields, P. D., Bento, G., & Ebert, D. (2021). Balancing selection for pathogen resistance reveals an intercontinental signature of red queen coevolution. *Molecular Biology and Evolution*, *38*(11), 4918–4933. https://doi.org/10.1093/MOLBEV/MSAB217

Bradshaw, R. H. W. (2004). Past anthropogenic influence on European forests and some possible genetic consequences. *Forest Ecology and Management*, *197*(1–3), 203–212. https://doi.org/10.1016/J.FORECO.2004.05.025

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B., & Sheldon, B. C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, *320*(5877), 800–803. https://doi.org/10.1126/SCIENCE.1157174/SUPPL_FILE/CHARMANTIER.SOM.PDF

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2017). Araport11: A complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal*, *89*(4), 789–804. https://doi.org/10.1111/tpj.13415

Cole, E. F., Regan, C. E., & Sheldon, B. C. (2021). Spatial variation in avian phenological response to climate change linked to tree health. *Nature Climate Change*, *11*(10), 872–878. https://doi.org/10.1038/s41558-021-01140-4

Cornetti, L., Fields, P. D., Du Pasquier, L., & Ebert, D. (2024). Long-term balancing selection for pathogen resistance maintains trans-species polymorphisms in a planktonic crustacean. *Nature Communications*, *15*(1), 5333. https://doi.org/10.1038/s41467-024-49726-8

Cottrell, J. E., Munro, R. C., Tabbener, H. E., Gillies, A. C. M., Forrest, G. I., Deans, J. D., & Lowe, A. J. (2002). Distribution of chloroplast DNA variation in British oaks (Quercus robur and Q. Petraea): The influence of postglacial colonisation and human management. *Forest Ecology and Management*, *156*(1–3), 181–195. https://doi.org/10.1016/S0378-1127(01)00642-9

Cottrell, J. E., Munro, R. C., Tabbener, H. E., Milner, A. D., Forrest, G. I., & Lowe, A. J. (2003). Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. *Forest Ecology and Management*, *176*(1–3), 287–303. https://doi.org/10.1016/S0378-1127(02)00289-X

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/BIOINFORMATICS/BTR330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), 1–4. https://doi.org/10.1093/GIGASCIENCE/GIAB008

de Ronne, M., Légaré, G., Belzile, F., Boyle, B., & Torkamaneh, D. (2023). 3D-GBS: A universal genotyping-by-sequencing approach for genomic selection and other high-throughput low-cost applications in species with small to medium-sized genomes. *Plant Methods*, *19*(1), 1–10. https://doi.org/10.1186/S13007-023-00990-7/TABLES/3

Degen, B., Yanbaev, Y. A., Ianbaev, R. Y., Bakhtina, S. Y., Gabitova, A. A., & Tagirova, A. A. (2021). Genetic diversity and differentiation of northern populations of pedunculate oak based on analysis of new SNP markers. *Russian Journal of Genetics*, *57*(3), 374–378. https://doi.org/10.1134/S1022795421030054/TABLES/1

Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, *201*(1), 45–56. https://doi.org/10.1111/NPH.12441

Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, *5*(5), 574–582. https://doi.org/10.1038/s41559-021-01403-5

Duminil, J., Hardy, O. J., & Petit, R. J. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, *9*(1), 1–14. https://doi.org/10.1186/1471-2148-9-177 /FIGURES/4

Finkeldey, R., & Ziehe, M. (2004). Genetic implications of silvicultural regimes. *Forest Ecology and Management*, *197*(1–3), 231–244. https://doi.org/10.1016/J.FORECO.2004.05.036

Firmat, C., Delzon, S., Louvet, J. M., Parmentier, J., & Kremer, A. (2017). Evolutionary dynamics of the leaf phenological cycle in an oak metapopulation along an elevation gradient. *Journal of Evolutionary Biology*, *30*(12), 2116–2131. https://doi.org/10.1111/JEB.13185

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger, R., & Widmer, A. (2017). Estimating genomic diversity and population differentiation - an empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. *BMC Genomics*, *18*(1), 1–15. https://doi.org/10.1186/S12864-016-3459-7/FIGURES/4

Forestry Climate Change Working Group. n.d. *Action plan for climate change adaptation of forests, woods and trees in England*. Retrieved 7 July 2024, from https://planthealthy.org.uk/assets/downloads/action-plan-for-climate-change-adaptation.pdf

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/J.BIOCON.2005.05.002

Frankham, R., Ballou, J. D., Ralls, K., Eldridge, M., Dudash, M. R., Fenster, C. B., Lacy, R. C., & Sunnucks, P. (Eds.). (2019). Inbreeding and loss of genetic diversity increase extinction risk. In *A Practical Guide for Genetic Management of Fragmented Animal and Plant Populations* (pp. 31–48). Oxford University Press. https://doi.org/10.1093/OSO/9780198783411.003.0003

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210X.12382

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973–983. https://doi.org/10.1534/GENETICS.113.160572/-/DC1/GENETICS.113.160572-3.PDF

Fussi, B., Westergren, M., Aravanopoulos, F., Baier, R., Kavaliauskas, D., Finzgar, D., Alizoti, P., Bozic, G., Avramidou, E., Konnert, M., & Kraigher, H. (2016). Forest genetic monitoring: An overview of concepts and definitions. *Environmental Monitoring and Assessment*, *188*(8), 493. https://doi.org/10.1007/S10661-016-5489-7

Gagen, M., Matthews, N., Denman, S., Bridge, M., Peace, A., Pike, R., & Young, G. (2019a). The tree ring growth histories of UK native oaks as a tool for investigating chronic oak decline: An example from the Forest of Dean. *Dendrochronologia*, *55*, 50–59. https://doi.org/10.1016/J.DENDRO.2019.03.001

Gagen, M., Matthews, N., Denman, S., Bridge, M., Peace, A., Pike, R., & Young, G. (2019b). The tree ring growth histories of UK native oaks as a tool for investigating chronic oak decline: An example from the Forest of Dean. *Dendrochronologia*, *55*, 50–59. https://doi.org/10.1016/J.DENDRO.2019.03.001

Gladieux, P., van Oosterhout, C., Fairhead, S., Jouet, A., Ortiz, D., Ravel, S., Shrestha, R. K., Frouin, J., He, X., Zhu, Y., Morel, J. B., Huang, H., Kroj, T., & Jones, J. D. G. (2024). Extensive immune receptor repertoire diversity in disease-resistant rice landraces. *Current Biology*, *34*(17), 3983–3995.e6. https://doi.org/10.1016/J.CUB.2024.07.061

Gömöry, D., Yakovlev, I., Zhelev, P., Jedináková, J., & Paule, L. (2001). Genetic differentiation of oak populations within the Quercus robur/Quercus petraea complex in central and Eastern Europe. *Heredity*, *86*(5), 557–563. https://doi.org/10.1046/j.1365-2540.2001.00874.x

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical f-statistics. *Molecular Ecology Notes*, *5*(1), 184–186. https://doi.org/10.1111/J.1471-8286.2004.00828.X

Grayson, A. J., & Jones, E. W. (1955). *Notes on the history of the Wytham estate with special reference to the woodlands*. Imperial Forestry Institute.

Hamrick, J. L., Godt, M. J. W., & Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. In W. T. Adams, S. H. Strauss, D. L. Copes, & A. R. Griffin (Eds.), *Population genetics of Forest trees* (pp. 95–124). Springer. https://doi.org/10.1007/978-94-011-2815-5_7

Hansson, B., & Westerberg, L. (2002). On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology*, *11*(12), 2467–2474. https://doi.org/10.1046/J.1365-294X.2002.01644.X

Hartvig, I., So, T., Changtragoon, S., Tran, H. T., Bouamanivong, S., Ogden, R., Senn, H., Vieira, F. G., Turner, F., Talbot, R., Theilade, I., Nielsen, L. R., & Kjær, E. D. (2020). Conservation genetics of the critically endangered Siamese rosewood (*Dalbergia cochinchinensis*): Recommendations for management and sustainable use. *Conservation Genetics*, *21*(4), 1–16. https://doi.org/10.1007/s10592-020-01279-1

Heuertz, M., Carvalho, S. B., Galindo, J., Rinkevich, B., Robakowski, P., Aavik, T., Altinok, I., Barth, J. M. I., Cotrim, H., Goessen, R., González-Martínez, S. C., Grebenc, T., Hoban, S., Kopatz, A., McMahon, B. J., Porth, I., Raeymaekers, J. A. M., Träger, S., Valdecantos, A., … Garnier-Géré, P. (2023). The application gap: Genomics for biodiversity and ecosystem service management. *Biological Conservation*, *278*, 109883. https://doi.org/10.1016/J.BIOCON.2022.109883

Hoban, S., da Silva, J. M., Hughes, A., Hunter, M. E., Kalamujić Stroil, B., Laikre, L., Mastretta-Yanes, A., Millette, K., Paz-Vinas, I., Bustos, L. R., Shaw, R. E., Vernesi, C., Genetics, the C. for C, Funk, C., Grueber, C., Kershaw, F., MacDonald, A., Meek, M., Mittan, C., … Segelbacher, G. (2024). Too simple, too complex, or just right? Advantages, challenges, and guidance for indicators of genetic diversity. *BioScience*, *74*(4), 269–280. https://doi.org/10.1093/BIOSCI/BIAE006

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in Threespine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*(2), e1000862. https://doi.org/10.1371/JOURNAL.PGEN.1000862

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, *21*(6), 797–807. https://doi.org/10.1007/S10980-005-5245-9 /METRICS

Hosius, B., Leinemann, L., Konnert, M., & Bergmann, F. (2006). Genetic aspects of forestry in the Central Europe. *European Journal of Forest Research*, *125*(4), 407–417. https://doi.org/10.1007/S10342-006-0136-4 /METRICS

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*(6), 609–623. https://doi.org/10.1111/J.1461-0248.2008.01179.X

Jiménez-Ramírez, A., Grivet, D., & Robledo-Arnuncio, J. J. (2021). Measuring recent effective gene flow among large populations in Pinus sylvestris: Local pollen shedding does not preclude substantial long-distance pollen immigration. *PLoS ONE*, *16*(8), e0255776. https://doi.org/10.1371/JOURNAL.PONE.0255776

Joshi N. A. 2013 Sabre - A barcode demultiplexing and trimming tool for FastQ files https://github.com/najoshi/sabre

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/PEERJ.281/TABLE-6

Kawecki, T. J., Erkosar, B., Dupuis, C., Hollis, B., Stillwell, R. C., & Kapun, M. (2021). The genomic architecture of adaptation to larval malnutrition points to a trade-off with adult starvation resistance in drosophila. *Molecular Biology and Evolution*, *38*(7), 2732–2749. https://doi.org/10.1093/MOLBEV/MSAB061

Kesić, L., Cseke, K., Orlović, S., Stojanović, D. B., Kostić, S., Benke, A., Borovics, A., Stojnić, S., & Avramidou, E. V. (2021). Genetic diversity and differentiation of Pedunculate oak (Quercus robur L.) populations at the southern margin of its distribution range—Implications for conservation. *Diversity*, *13*, 371. https://doi.org/10.3390/D13080371

Kirby, K. J., Bazely, D. R., Goldberg, E. A., Hall, J. E., Isted, R., Perry, S. C., & Thomas, R. C. (2014). Changes in the tree and shrub layer of Wytham Woods (southern England) 1974–2012: Local and national trends compared. *Forestry: an International Journal of Forest Research*, *87*(5), 663–673. https://doi.org/10.1093/FORESTRY/CPU026

Koenig, D., Hagmann, J., Li, R., Bemm, F., Slotte, T., Nueffer, B., Wright, S. I., & Weigel, D. (2019). Long-term balancing selection drives evolution of immunity genes in *capsella*. *eLife*, *8*, e43606. https://doi.org/10.7554/ELIFE.43606

Korecký, J., Čepl, J., Stejskal, J., Faltinová, Z., Dvořák, J., Lstibůrek, M., & El-Kassaby, Y. A. (2021). Genetic diversity of Norway spruce ecotypes assessed by GBS-derived SNPs. *Scientific Reports 2021 11:1*, *11*(1), 23119. https://doi.org/10.1038/s41598-021-02545-z

Koski, V. (1970). A study of pollen dispersal as a mechanism of gene flow in conifers. *Metsätieteellisen Tutkimuslaitoksen Julkaisuja*, *70*(4), 1–78.

Laikre, L. (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics*, *11*(2), 349–354. https://doi.org/10.1007/S10592-009-0037-4/TABLES/1

Laikre, L., Schwartz, M. K., Waples, R. S., & Ryman, N. (2010). Compromising genetic diversity in the wild: Unmonitored large-scale release of plants and animals. *Trends in Ecology & Evolution*, *25*(9), 520–529. https://doi.org/10.1016/J.TREE.2010.06.013

Ledig, F. T. (1992). Human impacts on genetic diversity in Forest ecosystems. *Oikos*, *63*(1), 87. https://doi.org/10.2307/3545518

Macdonald, D. W., Newman, C., Buesching, C. D., & Nouvellet, P. (2010). Are badgers 'under the weather'? Direct and indirect impacts of climate variation on European badger (Meles meles) population dynamics. *Global Change Biology*, *16*(11), 2913–2922. https://doi.org/10.1111/J.1365-2486.2010.02208.X

Macdonald, S. E., Thomas, B. R., Cherniawsky, D. M., & Purdy, B. G. (2001). Managing genetic resources of lodgepole pine in west-Central Alberta: Patterns of isozyme variation in natural populations and effects of forest management. *Forest Ecology and Management*, *152*(1–3), 45–58. https://doi.org/10.1016/S0378-1127(00)00616-2

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*(1), 10–12. https://doi.org/10.14806/EJ.17.1.200

Mastretta-Yanes, A., Suárez, S., Jordan, R., Hoban, S., Da Silva, J. M., Castillo-Reina, L., Heuertz, M., Ishihama, F., Köppä, V., Laikre, L., Macdonald, A. J., Mergeay, J., Paz-Vinas, I., Segelbacher, G., Knapps, A., Rakoczy, H., Weiler, A., Atsaves, A., Cullmann, K., … de México, C. C. (2024). Guideline materials and documentation for the genetic diversity indicators of the monitoring framework for the Kunming-Montreal global biodiversity framework. *Biodiversity Informatics*, *18*, 24–27. https://doi.org/10.17161/BI.V18I.22332

Mihók, B., Kenderes, K., Kirby, K. J., Paviour-Smith, K., & Elbourn, C. A. (2009). Forty-year changes in the canopy and the understorey in Wytham Woods. *Forestry: an International Journal of Forest Research*, *82*(5), 515–527. https://doi.org/10.1093/FORESTRY/CPP021

Naidoo, S., Slippers, B., Plett, J. M., Coles, D., & Oates, C. N. (2019). The road to resistance in forest trees. *Frontiers in Plant Science*, *10*, 418403. https://doi.org/10.3389/FPLS.2019.00273 /BIBTEX

Nakanishi, A., Yoshimaru, H., Tomaru, N., Miura, M., Manabe, T., & Yamamoto, S. i. (2015). Inbreeding depression at the sapling stage and its genetic consequences in a population of the outcrossing dominant tree species, *Castanopsis sieboldii*. *Tree Genetics & Genomes*, *11*(3), 1–10. https://doi.org/10.1007/S11295-015-0890-X/FIGURES/3 /FIGURES/3

Neale, D. B. (1985). Genetic implications of shelterwood regeneration of Douglas-fir in Southwest Oregon. *Forest Science*, *31*(4), 995–1005. https://doi.org/10.1093/FORESTSCIENCE/31.4.995

Nocchi, G., Brown, N., Coker, T. L. R., Plumb, W. J., Stocks, J. J., Denman, S., & Buggs, R. J. A. (2022). Genomic structure and diversity of oak populations in British parklands. *Plants*, *People*, *Planet*, *4*(2), 167–181. https://doi.org/10.1002/PPP3.10229

O'Hara, N. B., Franks, S. J., Kane, N. C., Tittes, S., & Rest, J. S. (2021). Evolution of pathogen response genes associated with increased disease susceptibility during adaptation to an extreme drought in a *Brassica rapa* plant population. *BMC Ecology and Evolution*, *21*(1), 1–11. https://doi.org/10.1186/S12862-021-01789-7/FIGURES/3

Ortego, J., Bonal, R., & Muñoz, M. (2010). Genetic consequences of habitat fragmentation in long-lived tree species: The case of the Mediterranean holm oak (*Quercus ilex*, L.). *Journal of Heredity*, *101*(6), 717–726. https://doi.org/10.1093/JHERED/ESQ081

Paradis, E., & Barrett, J. (2010). Pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/BIOINFORMATICS/BTP696

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Peterken, G. F. (1977). Habitat conservation priorities in British and European woodlands. *Biological Conservation*, *11*(3), 223–236. https://doi.org/10.1016/0006-3207(77)90006-4

Popović, M., Katičić Bogdan, I., Varga, F., Šatović, Z., Bogdan, S., & Ivanković, M. (2023). Genetic diversity in peripheral Pedunculate oak (Quercus robur L.) provenances—Potential climate change Mitigators in the Center of Distribution despite challenges in natural populations. *Forests*, *14*(12), 2290. https://doi.org/10.3390/F14122290/S1

Porth, I., & El-Kassaby, Y. A. (2014). Assessment of the genetic diversity in Forest tree populations using molecular markers. *Diversity*, *6*(2), 283–295. https://doi.org/10.3390/D6020283

Potter, K. M., Jetton, R. M., Bower, A., Jacobs, D. F., Man, G., Hipkins, V. D., & Westwood, M. (2017). Banking on the future: Progress, challenges and opportunities for the genetic conservation of forest trees. *New Forests*, *48*(2), 153–180. https://doi.org/10.1007/S11056-017-9582-8 /FIGURES/2

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rajendra, K. C., Seifert, S., Prinz, K., Gailing, O., & Finkeldey, R. (2014). Subtle human impacts on neutral genetic diversity and spatial patterns of genetic variation in European beech (Fagus sylvatica). *Forest Ecology and Management*, *319*, 138–149. https://doi.org/10.1016/J.FORECO.2014.02.003

Rajora, O. P., & Pluhar, S. A. (2003). Genetic diversity impacts of forest fires, forest harvesting, and alternative reforestation practices in black spruce (*Picea mariana*). *Theoretical and Applied Genetics*, *106*(7), 1203–1212. https://doi.org/10.1007/S00122-002-1169-9 /FIGURES/1

Ratnam, W., Rajora, O. P., Finkeldey, R., Aravanopoulos, F., Bouvet, J. M., Vaillancourt, R. E., Kanashiro, M., Fady, B., Tomita, M., & Vinson, C. (2014a). Genetic effects of forest management practices: Global synthesis and perspectives. *Forest Ecology and Management*, *333*, 52–65. https://doi.org/10.1016/J.FORECO.2014.06.008

Ratnam, W., Rajora, O. P., Finkeldey, R., Aravanopoulos, F., Bouvet, J. M., Vaillancourt, R. E., Kanashiro, M., Fady, B., Tomita, M., & Vinson, C. (2014b). Genetic effects of forest management practices: Global synthesis and perspectives. *Forest Ecology and Management*, *333*, 52–65. https://doi.org/10.1016/J.FORECO.2014.06.008

Ruiz-González, H., Raggio, M., Labrín-Sotomayor, N. Y., Ferrer, M. M., & Peña-Ramírez, Y. J. (2023). Genetic variability of *Tabebuia rosea* (Bignoniaceae) from plantations and remnant populations in the Mayan Forest. *Forests*, *14*(10), 2006. https://doi.org/10.3390/F14102006/S1 /S1

Ryynänen, H. J., & Primmer, C. R. (2004). Distribution of genetic variation in the growth hormone 1 gene in Atlantic salmon (Salmo salar) populations from Europe and North America. *Molecular Ecology*, *13*(12), 3857–3869. https://doi.org/10.1111/J.1365-294X.2004.02370.X

Savill, P., Perrins, C., Kirby, K., & Fisher, N. (Eds.). (2011). *Wytham Woods: Oxford's Ecological Laboratory* (pp. 1–288). Oxford University Press. https://doi.org/10.1093/ACPROF:OSOBL/9780199605187.001.0001

Savolainen, O., & Pyhäjärvi, T. (2007). Genomic diversity in forest trees. *Current Opinion in Plant Biology*, *10*(2), 162–167. https://doi.org/10.1016/J.PBI.2007.01.011

Schmidt, T. L., Jasper, M. E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*(10), 1888–1898. https://doi.org/10.1111/2041-210X.13659

Semerikova, S. A., Tashev, A. N., & Semerikov, V. L. (2023). Genetic diversity and history of pedunculate oak *Quercus robur* L. in the east of the range. *Russian Journal of Ecology*, *54*(5), 423–438. https://doi.org/10.1134/S1067413623050089 /FIGURES/4

Sopniewski, J., & Catullo, R. A. (2024). Estimates of heterozygosity from single nucleotide polymorphism markers are context-dependent and often wrong. *Molecular Ecology Resources*, *24*(4), e13947. https://doi.org/10.1111/1755-0998.13947

Stange, M., Barrett, R. D. H., & Hendry, A. P. (2020). The importance of genomic variation for biodiversity, ecosystems and people. *Nature Reviews Genetics 2020 22:2*, *22*(2), 89–105. https://doi.org/10.1038/s41576-020-00288-7

Sunde, J., Yıldırım, Y., Tibblin, P., & Forsman, A. (2020). Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics*, *11*(513177), 218. https://doi.org/10.3389/FGENE.2020.00218 /BIBTEX

Swain, D. P., Sinclair, A. F., & Hanson, J. M. (2007). Evolutionary response to size-selective mortality in an exploited fish population. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1613), 1015–1022. https://doi.org/10.1098/RSPB.2006.0275

Sykes, J. M., & Lane, A. M. J. (1996). In J. M. Sykes & A. M. J. Lane (Eds.), *The United Kingdom environmental change network: Protocols for standard measurements at terrestrial sites*. The Stationery Office.

Tansley, A. G. (1949). *The British islands and their vegetation*. Cambridge University Press.

Thomas, M. V., Malhi, Y., Fenn, K. M., Fisher, J. B., Morecroft, M. D., Lloyd, C. R., Taylor, M. E., & McNeil, D. D. (2011). Carbon dioxide fluxes over an ancient broadleaved deciduous woodland in southern England. *Biogeosciences*, *8*(6), 1595–1613. https://doi.org/10.5194/BG-8-1595-2011

Ulaszewski, B., Meger, J., & Burczyk, J. (2021). Comparative analysis of SNP discovery and genotyping in Fagus sylvatica L. and Quercus robur L. using RADseq, GBS, and ddRAD methods. *Forests*, *12*(2), 222. https://doi.org/10.3390/F12020222

Valbuena-Carabaña, M., González-Martínez, S. C., & Gil, L. (2008). Coppice forests and genetic diversity: A case study in Quercus pyrenaica Willd. From Central Spain. *Forest Ecology and Management*, *254*(2), 225–232. https://doi.org/10.1016/J.FORECO.2007.08.001

Van Der Hoorn, R. A. L., De Wit, P. J. G. M., & Joosten, M. H. A. J. (2002). Balancing selection favors guarding resistance proteins. *Trends in Plant Science*, *7*(2), 67–71. https://doi.org/10.1016/S1360-1385(01)02188-4

Vasimuddin, M., Misra, S., Li, H., & Aluru, S. (2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. *Proceedings - 2019 IEEE 33rd International Parallel and Distributed Processing Symposium, IPDPS 2019*, 314–324. https://doi.org/10.1109/IPDPS.2019.00041

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

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419–1439. https://doi.org/10.1111/J.1365-294X.2006.02890.X

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., Byrne, M., Coates, D. J., Eldridge, M. D. B., Sunnucks, P., Breed, M. F., James, E. A., & Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: A genetic perspective. *Evolutionary Applications*, *4*(6), 709–725. https://doi.org/10.1111/j.1752-4571.2011.00192.x

Whittet, R., Cavers, S., Cottrell, J., & Ennos, R. (2017). Seed sourcing for woodland creation in an era of uncertainty: An analysis of the options for Great Britain. *Forestry: an International Journal of Forest Research*, *90*(2), 163–173. https://doi.org/10.1093/FORESTRY/CPW037

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A note on exact tests of Hardy-Weinberg equilibrium. *American Journal of Human Genetics*, *76*(5), 887–893. https://doi.org/10.1086/429864

Williams, C. G., & Savolainen, O. (1996). Inbreeding depression in conifers: Implications for breeding strategy. *Forest Science*, *42*(1), 102–117. https://doi.org/10.1093/FORESTSCIENCE/42.1.102

Woziwoda, B., Dyderski, M. K., Kobus, S., Parzych, A., & Jagodziński, A. M. (2019). Natural regeneration and recruitment of native Quercus robur and introduced Q. Rubra in European oak-pine mixed forests. *Forest Ecology and Management*, *449*, 117473. https://doi.org/10.1016/J.FORECO.2019.117473

Wu, Q., Han, T. S., Chen, X., Chen, J. F., Zou, Y. P., Li, Z. W., Xu, Y. C., & Guo, Y. L. (2017). Long-term balancing selection contributes to adaptation in *Arabidopsis* and its relatives. *Genome Biology*, *18*(1), 1–15. https://doi.org/10.1186/S13059-017-1342-8/FIGURES/5 /FIGURES/5

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovations*, *2*(3), 100141. https://doi.org/10.1016/J.XINN.2021.100141

Zaynab, M., Fatima, M., Abbas, S., Sharif, Y., Umair, M., Zafar, M. H., & Bahadar, K. (2018). Role of secondary metabolites in plant defense against pathogens. *Microbial Pathogenesis*, *124*, 198–202. https://doi.org/10.1016/J.MICPATH.2018.08.034

Zhang, Y., Guo, M., Shen, J., Song, X., Dong, S., Wen, Y., Yuan, X., & Guo, P. (2019). Comparative genomics analysis in grass species reveals two distinct evolutionary strategies adopted by R genes. *Scientific Reports 2019 9:1*, *9*(1), 10735. https://doi.org/10.1038/s41598-019-47121-8

Zimmerman, S. J., Aldridge, C. L., & Oyler-Mccance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics*, *21*(1), 1–16. https://doi.org/10.1186/S12864-020-06783-9 /FIGURES/6

SUPPORTING INFORMATION

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

How to cite this article: Hung, T. H., Formaggia, E., Morley, L., Kirby, K., Salguero-Gómez, R., Sheldon, B. C., & MacKay, J. J. (2025). Genetic diversity and population structure of pedunculate oaks (*Quercus robur*) in Wytham Woods. *Plants, People, Planet*, 1–14. https://doi.org/10.1002/ppp3.70042