Original article

Understanding the legacy of widespread population translocations on the post-glacial genetic structure of the European beech, *Fagus sylvatica* L.

M. J. Sjölund[1,2], P. González-Díaz[1], J. J. Moreno-Villena[1,3], and A. S. Jump[1,4]

[1] Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK. [2] Current address: Science and Advice for Scottish Agriculture (SASA), Roddinglaw Road, Edinburgh, EH12 9FJ, UK. [3] Current address: Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK. [4] Centre de Recerca Ecològica i Aplicacions Forestals (CREAF), Campus UAB, Edifici C. E-08193, Bellaterra, Barcelona, Spain.

Corresponding author: M. J. Sjölund

Email: jennifer.sjolund@sasa.gsi.gov.uk

Address: Diagnostics, Wildlife and Molecular Biology (DWMB), Science and Advice for Scottish Agriculture (SASA), Roddinglaw Road, Edinburgh, EH12 9FJ, UK.

Running Header

Natural colonisation signals persist despite beech forest translocation

Word count (main text, abstract, references): 7175

ABSTRACT

Aim Human impacts have shaped species ranges throughout the Holocene. The putative native range of beech, *Fagus sylvatica*, in Britain was obscured by its late post-glacial arrival and subsequent extensive management. We sought to differentiate the interacting effects of post-glacial colonization and anthropic impacts on the current genetic structure and diversity of beech by contrasting phylogeographic signals from putatively natural and translocated populations.

Location Samples were obtained from 42 sites throughout Great Britain.

Methods Chloroplast and nuclear microsatellite marker data were interpreted alongside palynological, historical and anecdotal evidence. Genetic structure was analysed using individual-based Bayesian assignment methods and colonization history was analysed using an approximate Bayesian computation framework.

Results Phylogeographic patterns suggested contemporary forests originated from putative native south-eastern populations. High haplotypic diversity was found near the entry point of beech into Britain. Cryptic signals of isolation-by-distance persisted in the putative native range, together with higher levels of gene diversity in nuclear markers. Weak regional nuclear genetic structure suggested high levels of contemporary gene flow throughout the country.

Main conclusions Genetic patterns driven by natural colonization persist despite widespread anthropic intervention. Forests in northerly regions were established from forests in the putative native range, diminishing the credibility of any present boundary between the native and non-native range of beech in Britain.

Key words Anthropogenic, Britain, colonization, *Fagus sylvatica*, gene flow, microsatellites, phylogeography, post-glacial.

INTRODUCTION

The migration of species during the Holocene (Taberlet *et al.*, 1998) coincided with substantial human impact (Kalis & Merkt, 2003; Kalis *et al.*, 2003) shaping species contemporary ranges. Historic range limits have been examined from an ecological perspective and post-glacial plant migrations from a phylogeographic perspective (Comes & Kadereit, 1998; Hewitt, 2000; Jump *et al.*, 2010; Magri, 2010). However, many natural systems have been under profound and persistent anthropic influence, which has shaped species distributions and the genetic composition of their component populations (Bradshaw, 2004; Alessa & Iii, 2008; Schaberg *et al.*, 2008). The influence of forest management, through the selective removal of genotypes and the translocation of plant material, can impact genetic diversity and its spatial distribution within populations and across regions (Savolainen & Kärkkäinen, 1992; Bradshaw, 2004; Alessa & Iii, 2008; Schaberg *et al.*, 2008). However, despite widespread human influences, native forests might retain genetic signals from past population distribution due to natural regeneration of local stock in combination with the long life of individual trees (Petit & Hampe, 2006; Bradshaw, 2004), thereby allowing natural ranges to be detected.

Given its many uses, including timber, fuel and fodder, the European beech, *Fagus sylvatica* L., has experienced a long history of management (Nocentini, 2009; Read *et al.*, 2010; Packham *et al.*, 2012) including wide-scale historical translocations throughout Great Britain. Anecdotal, historical and palynological evidence suggest that the native range of beech was limited to south-east England (Rackham, 1980; Pott, 2000; Packham *et al.*, 2012). Despite this commonly held view, this species has been classified as native throughout mainland Great Britain, due to insufficient evidence available to accurately define the native range (Preston *et al.*, 2002). At the regional scale, the species range is believed to be under broad climatic control (Huntley *et al.*, 1989) and in Great Britain, beech has naturalised further north than its

79 presumed native range in the south-east.

80

81 Pollen records from Great Britain indicate that beech migrated into south-east England, with
82 its first establishment just before 3000 BP. Beech maintained a steady rate of spread of
83 100-200 m per year, whereas the majority of other tree species displayed a decrease in the
84 rate of spread. The relatively constant rate of spread suggests that beech had not reached its
85 natural climatic limit by 1000 BP (Birks, 1989). Human intervention might have manipulated
86 the species' range before it reached its climatic range limit (Watt, 1931; Packham *et al.*,
87 2012). Historical records suggests the existence of native populations that occur further north
88 than the predicted range as suggested by pollen evidence (Rackham, 1980). However,
89 confirmation of the species' presence in the pollen record does not directly indicate native
90 origin of forests that exist today. Beech therefore provides a valuable model for
91 understanding how human intervention through translocation of material might impact the
92 genetic structure and diversity of natural populations.

93

94 Studies that aim to determine the 'native status' of a species in a particular region have
95 hypothesised that introduced populations are genetically depauperate as they originate from
96 limited source propagules (Stone & Sunnucks, 1993; Fuentes-Utrilla *et al.*, 2014).
97 Consequently, we sought to determine the interacting effects of past migration and anthropic
98 impacts on the current genetic structure and diversity of beech by assessing if a
99 phylogeographic signal of natural colonization can be identified in its putative native range.
100 We used a combination of conservative, maternally-inherited chloroplast markers (Reboud &
101 Zeyl, 1994; Magri *et al.*, 2006) and highly variable nuclear markers to explore regional trends
102 in genetic variation in combination with historical and palynological data. Extensive sampling
103 was carried out in both the putative native and non-native range across Great Britain. We
104 hypothesised that; (1) patterns associated with natural colonization, such as spatial genetic

structure, should persist in the putative native range, and putative native sites will show high genetic similarity; (2) patterns driven by natural colonization should be absent in the putative non-native range due to extensive translocations of plant material by humans outside the putative native range; (3) putative non-native sites should display lower levels of haplotypic and genotypic diversity and greater inter-population divergence, compared to putative native sites, due to the high levels of genetic drift associated with the translocation of a limited number genotypes. We refer to sites as 'putative native' and 'putative non-native' as the *a priori* assigned origins of sites used in our study are purely for analytical purposes and should not be taken as precedent.

MATERIALS AND METHODS

Study species

Fagus sylvatica L. is a broadleaved, monoecious, primarily outcrossing tree, with pollen dispersed by wind and seeds dispersed by gravity and animals (Fig 1.; Wagner *et al.*, 2010; Packham *et al.*, 2012). It covers approximately 14 million ha, forming the dominant forest type in much of mainland Europe. With the exception of Great Britain, the distribution of beech is primarily climatically limited owing to the species' drought and late frost susceptibility (Watt, 1923; Peterken & Mountford, 1996). Beech trees reach approximately 300 years of age, commencing flowering typically between 40 to 80 years (Firbas & Losert, 1949).

Palaeobotanical and genetic data indicate central and northern European populations were colonized from source populations in southern France, eastern Alps-Slovenia-Isteria and potentially Moravia and southern Bohemia. Residual populations in the Iberian, Italian and Balkan refugia are believed to have expanded relatively late and did not contribute significantly to the colonization of central and northern Europe (Magri *et al.*, 2006). Recent

evidence found in Denmark suggests that post-glacial colonization was aided by occasional long-distance dispersal events, leading to the establishment of beech and other temperate tree species ahead of their main colonization fronts (Overballe-Petersen *et al.*, 2012) and might have contributed significantly to the observed rates of spread of the species (Feurdean *et al.*, 2013).

Study sites

Forty-two populations were sampled across Great Britain (Fig. 2). Using a combination of historical records, palynological, and anecdotal evidence, study sites were designated *a priori* as stands of putative native or putative non-native origin (see Table S1.1 in Appendix S1 in Supporting Information). In each site, leaf samples were collected from 20 mature trees within a 10 ha area, preferentially sampling the oldest trees determined by using diameter at breast height (DBH) as a proxy for age. All samples were geo-referenced using a GARMIN 62s handheld GPS (GARMIN, Southampton, UK).

Molecular analysis

DNA was obtained from the silica gel dried leaf samples and isolated using the QIAGEN DNeasy 96 Plant Kit (QIAGEN, Venlo, Netherlands). Out of 840 samples, 802 individuals were successfully genotyped using four Chloroplast DNA (CpDNA) microsatellite markers (ccmp4, ccmp7 (Weising & Gardner 1999), cmcs3, and cmcs12 (Sebastiani *et al.*, 2004)). Chloroplast markers were combined in one PCR multiplex, FSCplex, using 10ng of template DNA and the QIAGEN Type-it Microsatellite PCR Kit with the following primer concentrations, ccmp4 at 0.5 µM, ccmp7 at 0.5µM, cmcs3 at 3µM, and cmcs12 at 3µM. Annealing temperature was set to 55°C, with a total PCR reaction volume of 10µl. 837 individuals were successfully genotyped using 13 nuclear microsatellite markers (fs1-03, fs1-15, fs3-04, fs4-46, fcm5 (Pastorelli *et al.*, 2003), mfc7 (Tanaka *et al.*, 1999), mfs11 (Vornam *et al.*, 2004), sfc0007-2, sfc0018, sfc0036,

157 sfc1143, sfc1061, sfc1063 (Asuka *et al.*, 2004) processed in three multiplexes as detailed in
158 Sjölund & Jump (2015). A total of three chloroplast microsatellite loci were used, excluding
159 cmcs12 as it was monomorphic. Fragment analysis was performed on an ABI 3730 (Applied
160 Biosystems, Bleiswijk, Netherlands) and allele scoring on GeneMarker 2.4.0 (SoftGenetics,
161 State College, PA, U.S.A).
162
163 Scoring errors and null alleles in nuclear loci were checked using Micro-Checker (Van
164 Oosterhout *et al.*, 2004). Analyses presented exclude fs4-46, fcm5, and fs1-15 due to null
165 allele presence and use a total of 10 nuclear microsatellite loci. Gametic disequilibrium was
166 tested between pairs of nuclear loci using FSTAT 2.9.3.2 (Goudet, 1995), identifying
167 significant associations between loci by randomly associating genotypes at pairs of loci 1100
168 times, using a 5% nominal level after Bonferroni correction. The multilocus average error
169 rates were 0.0% for the 3 chloroplast loci and 0.4% for the 10 nuclear loci included in analysis.
170 The error rate per locus was calculated as the number of erroneously assigned loci over 45
171 repeated samples.
172
173 **Measuring genetic diversity and structure using nuclear and chloroplast markers**
174 Estimators of genetic diversity were prefixed with '*c*' for chloroplast and '*n*' for nuclear.
175 Multilocus estimates of haplotypic and genotypic diversity were mapped in ArcMap 10 (ESRI
176 software) against the pollen isochrones from Birks' (1989) map for the rational limit of beech
177 pollen. Chloroplast haplotypic diversity was measured as the number of haplotypes (cH_N) and
178 the number of private haplotypes (cH_P). Nuclear genetic diversity was measured as rarefied
179 allelic richness (nA_R) (Petit *et al.*, 1998), gene diversity corrected for sample size (nH_S) (Nei,
180 1978), and the inbreeding coefficient (nF_{IS}) with *P*-values derived from 10,000 permutations
181 of gene copies within individuals per site (Weir and Cockerham 1984), calculated in SPAGeDi
182 1.4b (Hardy & Vekemans, 2002), and rarefied private allelic richness (nA_P) calculated in ADZE

1.0 (Szpiech *et al.*, 2008). The minimum number of gene copies (*k*) used for rarefication analysis of nA_R and nA_P is 38. To map nuclear genetic differentiation, for each site we calculated the percentage of total sites that it was significantly differentiated from (i.e. percentage of differentiated sites, *nDS (%)*), based on nF_{ST} values (Weir & Cockerham, 1984). nF_{ST} values were obtained from pairwise tests of genetic differentiation not assuming Hardy-Weinberg, with significances determined for a 5% nominal level after Bonferonni correction in FSTAT 2.9.3.2 (Goudet, 1995).

To test for differences of nuclear-based measurements, nA_R, nH_S, nF_{IS} and nF_{ST} among *a priori* determined groups of putative native and putative non-native sites, we performed permutation tests using FSTAT 2.9.3.2 (Goudet, 1995). The difference between putative native and putative non-native groups in the remaining estimators nA_P and cN_H were tested using the non-parametric Mann-Whitney U test. To quantify the distribution of variation displayed by the maps of nuclear genetic diversity and chloroplast haplotypes (considering haplotype distances) we tested both marker sets in a hierarchical analysis of molecular variance (AMOVA) performed in ARLEQUIN 3.5.1.2 (Excoffier & Lischer, 2010) with sites grouped into potential stand origin, putative native or putative non-native.

Testing isolation-by-distance with nuclear markers

Subsets of sites with putative native and putative non-native origin were tested for isolation-by-distance following Rousset (1997) using nuclear markers in SPAGeDi 1.4b (Hardy & Vekemans, 2002) with significance of the slope of the correlation of the natural log of the linear spatial distance (Ln(Spatial Distance (km)) against $F_{ST}/(1-F_{ST})$ was computed after 10,000 permutations of sites among locations. IBD was analysed with and without the continental datasets to check whether a geographical cline in allelic frequencies was influencing clustering (Guillot *et al.*, 2009). To test for geographic gradients in genetic

diversity, we performed non-parametric corrected Spearman's Rank tests on all genetic diversity estimators and all sites, including separate tests on subsets of putative native and putative non-native sites.

Analysing population clusters

Individual-based Bayesian assignment methods were performed using data from nuclear loci in STRUCTURE 2.3.4. (Pritchard *et al.*, 2000) using the correlated allele frequency model (Falush *et al.*, 2007), the admixture ancestry model, and site location *a priori* (LOCPRIOR option) to improve the detection of weak population structure (Hubisz *et al.*, 2009). No stand origin was included *a priori* in cluster analysis. To examine relationships between British samples and putative colonization sources in continental Europe, we included nuclear microsatellite data from a subset of 150 samples collected from native beech forests in France, Germany and Italy as detailed in Sjölund & Jump (2015). Analysis without continental samples revealed a similar structure in Britain to that found with continental samples, therefore we present the data including continental samples to set the results in context. *K* was set from 1 to 20, with 10 runs each. Runs consisted of 500,000 Markov chain Monte Carlo (MCMC) iterations with a burn-in period of 100,000. To observe the consensus in number of clusters in the data, we plotted the log probability of the data (LnP(D)), identifying *K* where log likelihood values converged. Individual Q-matrices were computed in CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007), with graphics created in DISTRUCT 1.1 (Rosenberg, 2004). As results indicated cryptic genetic structure in the putative native range, we performed a subsequent analysis on a subset of the 17 putative native sites (i.e. using the putative stand origin *a priori*) under the same conditions as the main model and excluding continental samples to test for further population sub-structure.

Demographic history

We used an approximate Bayesian computation (ABC) framework implemented in DIYABC 2.1.0 (Cornuet *et al.*, 2010) to explore scenarios of demographic history that were likely to have generated current regional genetic structure of beech in Britain according to the results obtained from analyses of regional genetic structure, combined with palynological and historical information. Populations were grouped regionally according to Birks' (1989) isochrones, with British populations grouped into southern, central and northern populations, and populations in Italy, France and Germany grouped into a continental population (see Fig S1.1 in Appendix S1 in Supporting Information). Three scenarios were tested (Fig. 6): Scenario 1 assumes natural colonization of beech in southern and central populations from continental Europe, and northern populations originating from continental European stock (i.e. beech is established in northern Britain by anthropogenic translocation from the continent). Scenario 2 assumes natural colonization of southern and central populations from continental Europe, and northern populations originating from southern and central stock (i.e. northern populations are established from material derived from the putative native region). Scenario 3 assumes natural colonization of southern and central populations from continental Europe, and northern populations originating from an admixture of southern, central and continental populations (i.e. some northern populations are established from southern and central stock, the putative native region, and some from continental stock).

Prior parameters for effective population sizes and timing of events were defined based on knowledge of beech colonization dynamics from palynological and historical information. Prior parameter distributions were uniform and bounded between 10-10^4 for the effective population size of southern, central and northern Britain, and 10-10^5 for continental Europe. Population divergence time priors ranged between 10-50 for *t1*, 25-100 for *t2* and 25-200 for *t3*, with the additional setting *t3*>*t2*, and *t2*> *t1*. Priors for admixture rates ranged between

0.01 and 0.99. Nuclear microsatellite loci were assumed to follow a Generalized Stepwise Mutation model and a uniform prior was assumed for the mean microsatellite mutation rate bounded between 10^{-3} and 10^{-4}. 36 summary statistics were used for the ABC analysis. Three single sample statistics were used (mean number of alleles, mean Nei's genetic diversity index (Nei, 1987) and mean allele size variance), and three between-sample statistics (mean allele size variance, F_{ST}, and mean index of classification (Rannala & Mountain, 1997; Pascual *et al.*, 2007). Following the methods outlined by Cornuet (2010), type I and type II errors were estimated to evaluate the power of the model. The selected scenario was used to estimate posterior distribution of demographic parameters.

RESULTS

Geographic trends in genetic diversity

Three variants were detected for chloroplast loci, ccmp4, ccmp7 and cmcs3. The number of haplotypes (cH_N) within sites ranged from one to four, with a total of seven haplotypes recorded (Table 1, Fig. 2; see Fig. S2.1 in Appendix S2 in Supporting Information). Haplotype (A) was present in all sites and was the dominant haplotype within sites. Haplotype diversity was highest in the putative native sites BLE (cH_N = 5) and LUL (cH_N = 3), the two most south-easterly British sites. Six out of seven haplotypes (A to F) were represented collectively in these two sites. Two private haplotypes, F and G, were present in sites BLE and DEV, respectively, in single individuals. Significant genetic structuring was found in chloroplast and nuclear markers between sites using AMOVA (Table 2), although no significant difference was found between groups of putative native and putative non-native sites. The remaining variation was present within individuals.

Multilocus estimates of genetic diversity were obtained for 10 nuclear loci, with an average number of 13.3 alleles per locus, and a maximum of five to 30 alleles per locus (Table 1 and

Fig. 3). All nuclear loci were under gametic equilibrium. Rarefied allelic richness (nA_R) varied from 4.58 to 7.19 with rarefied private allelic richness (nA_P) ranging from 0 to 0.156. Gene diversity estimates ranged from 0.606 to 0.750. Putative native site WYC and putative non-native MAB displayed a significant homozygote excess (WYC nF_{IS} = 0.088, P < 0.05; MAB nF_{IS} = 0.089, P < 0.05), whilst a heterozygote excess was found in sites BEE and CRA, both putative non-native sites (BEE nF_{IS} = -0.152, P < 0.001; CRA nF_{IS} = -0.086, P < 0.05). The percentage of significantly differentiated sites [*nDS (%)*] varied greatly, for example, site WYT was significantly differentiated to 4.9% of sites, whilst site BEE was significantly differentiated to 95.1% of all sites.

Differences between groups of putative native and non-native sites

Significantly higher levels of gene diversity (nH_S) were found in putative native sites, compared to putative non-native sites; H_S: putative native 0.708, putative non-native 0.690 (P < 0.05) (Table 1). The general trend of lower gene diversity in sites outside of the putative native range can be seen in the map for gene diversity in Fig. 3. No significant differences were found for other estimators; nA_R: putative native 6.225 v 6.252 putative non-native (P = 0.87), nA_P: 0.053 v 0.037 (U(40) = 245, Z = 0.84, P = 0.41), nF_{IS}: 0.022 v -0.005 (P = 0.07), nF_{ST}: 0.024 v 0.021 (P = 0.59), and cN_H: 1.5 v 1.4 (U(40) = 208, Z = -0.155, P = 0.89).

Isolation-by-distance

Significant isolation-by-distance was found in Britain amongst putative native sites (refer to Fig. 4C: slope = 0.0085, R^2 = 0.10, P < 0.05). IBD was not significant amongst putative non-native sites, nor was it significant in all sites combined with and without the continental subset (refer to Fig. 4, A: British sites only (slope = 0.0011, R^2 < 0.01, P = 0.22), B: British and continental sites (slope = 0.0032, R^2 = 0.03, P = 0.06), D: putative non-native sites (slope = 0.0030, R^2 = 0.02, P = 0.09). Within the putative native sites, there was a significant reduction

313 in haplotype number following an east to west gradient (rho = 0.70, $P < 0.01$; Fig. 2). This
314 effect disappeared when all sites were analysed together (rho = 0.18, $P = 0.241$), and was not
315 found in putative non-native sites alone (rho = -0.02, $P = 0.90$). Preliminary analysis revealed
316 no significant correlations between nuclear genetic diversity estimators and geographic
317 variables (i.e. latitude and longitude) overall sites and within subsets of putative native and
318 putative non-native sites (see Table S2.1 in Appendix S2 in Supporting Information).
319
320 **Detection of further regional structuring using nuclear markers**
321 Mean log-likelihood values for each of the STRUCTURE runs on samples from Britain including
322 the continental samples, gradually ceased to converge after $K = 2$, after which they began to
323 plateau (see Fig. S3.1 in Appendix S3 in Supporting Information). Examination of Q-matrices
324 indicated that $K = 3$ provided meaningful biological clusters that followed a regional
325 distribution (Fig. 5). Although values of Ln P(D) for $K = 3$ varied more than $K = 2$, assignment
326 of individuals to clusters were congruent between runs (see Fig. S3.2 in Appendix S3 in
327 Supporting Information).
328
329 Several sites throughout Britain contained highly admixed individuals, whereas individuals
330 from continental sites displayed homogenous levels of admixture among individuals within
331 site. Organising the Q-matrix according to Birks' (1989) isochrones in approximate geographic
332 order revealed some consistency in cluster assignment between neighbouring sites (Fig. 5).
333 There appeared to be a geographic gradient in the continental clusters with the predominant
334 cluster being blue in ITA, to grey in FRA and GER, with the introduction of the red cluster in
335 GER. TAN appears to have the highest assignment to the blue cluster out of any other sites in
336 Britain. Individuals from Britain were largely assigned to the grey and red clusters. The red
337 cluster appeared to be associated with sites within the putative native range and the putative
338 non-native sites in the south-west (DEV, GOL, HEM, and BRI). Similar patterns were observed

for K = 2 and K = 3 when analysing a subset of British samples alone and a subset of putative native sites, both revealing no further population sub-structuring (see Fig. S3.3 and Fig. S3.4 in Appendix S3 in Supporting Information).

Model-based inference using ABC analyses indicated that scenario 2 was supported with maximum probability compared to around zero support for scenarios 1 and 3 (Fig. 6). This is indicative of a gradual colonization of beech from continental Europe and expanding northwards into and through Britain, with northern populations being derived from the putative native range of beech in Britain. Confidence in scenario choice was high, with low error rates (type 1 = 0.001–0.054, type 2 = 0–0.054). For scenario 2, median values for the effective population size were 6830, 4920, 6270 and 25100 for *N1* (northern Britain), *N2* (central Britain), *N3* (southern Britain) and *N4* (Continental Europe). The median values for divergence events corresponded to 16.9, 26.5 and 170 number of generations for *t1*, *t2* and *t3*.

DISCUSSION

Analysing genetic data according to palynological and historical evidence allowed us to tease apart the post-glacial history of beech in Britain despite the prolonged human impacts on this species. Patterns of past colonization dynamics persisted in putative native sites, while phylogeographic patterns and modelled demographic history support the hypothesis that throughout Britain, beech forests are largely derived from native stock originating from trees that colonized the island during the Holocene.

Native origins of beech in Britain

The distribution of chloroplast haplotypes and modelled demographic history strongly support the colonization of northern beech populations from stock originating from the

putative native range of beech. The distribution of haplotypes in Britain matched the expected phylogeographic signal of postglacial colonization, with the highest number of haplotypes found in south-eastern sites, LUL and BLE (Table 1; Fig. 2) in close proximity to the purported entry point of beech migration into Britain (Birks 1989). All the haplotypes, except one (G), were represented in these sites, suggesting that the majority of Britain's beech forests were colonized from native stock. Similar south to north patterns in haplotype distribution have been found for other European tree taxa, including *Alnus*, *Quercus* and *Populus* (King & Ferris *et al.,* 1998; Petit *et al.,* 2002a; Cottrell *et al.,* 2005 along with genetic clines indicative of gradual recolonization of *Quercus robur* and *Quercus petraea* in Ireland (Kelleher *et al.*, 2004). The loss of haplotype diversity occurring away from proposed entry sites can be a result of founder effects induced by the progressive movement of the migration front (Excoffier *et al.*, 2009). The clinal pattern of haplotype diversity in *F. sylvatica* in Britain is in contrast with *Q. robur* and *Q. petraea*, which displayed a highly clumped distribution of haplotypes indicative of long-distance dispersal events (Cottrell *et al.*, 2002a) following the much earlier entry of these oaks into Britain (Birks 1989). There is little similar evidence of long-distance dispersal events in *F. sylvatica* in Britain. site BLE harboured one of the two private haplotypes found while haplotype (G) in site DEV may have arisen from a single base mutation in haplotype A or a long-distance dispersal event (Overballe-Petersen *et al.*, 2012).

The high number of haplotypes (cH_N = 7) is in contrast to that found by Magri *et al.* (2006) who report one haplotype throughout Britain. This is likely a consequence of our larger sample size, the significant partitioning of haplotype variation between sites, and polymorphism in cmcs3. Restricting analysis to ccmp4 and ccmp7 used by Magri *et al.* (2006) reduces haplotype number to four (data not shown) and matches the regional trend seen with all three loci, with diversity restricted to south-eastern sites. The average levels of

rarefied allelic richness using nuclear markers in Britain (nA_R = 6.25±0.08) were lower than those reported in studies using some of the same microsatellite markers, approximately ranging from 8.2 to 18.2 in other studies (Jump & Peñuelas, 2006; Buiteveld *et al.*, 2007; Sjölund & Jump, 2015). Sites in Britain also displayed lower levels of rarefied private allelic richness (nA_P = 0.044±0.007) compared to sites in continental Europe sampled by Sjölund and Jump (2015) where values for nA_P ranged between 1.51 and 2.36. Overall levels of gene diversity were similar to those found in other studies (nH_s = 0.696±0.004) (Jump & Peñuelas, 2006; Buiteveld *et al.*, 2007; Oddou-Muratorio *et al.*, 2008; Sjölund & Jump, 2015).

Genetic variation between groups of different a priori stand origins

Isolation-by-distance occurs when the genetic differentiation between individuals or populations increases with geographic distance (Wright, 1940). In plants, this is primarily a consequence of restricted gene flow via seed or pollen (Loveless & Hamrick, 1984). Although beech is assumed to show high levels of gene flow as a wind-pollinated tree, it displays significant structuring at local (Chybicki *et al.*, 2009; Jump *et al.*, 2012; Piotti *et al.*, 2013; Sjölund & Jump, 2015) and regional scales (Jump & Peñuelas, 2006; de Lafontaine *et al.*, 2013). In agreement with the significant genetic structuring found in natural populations of beech, putative native sites displayed a weak but significant trend of IBD based on nuclear loci, which was absent in putative non-native sites, with the addition of all sites obscuring the IBD signal (Fig. 4).

Populations of beech in France with relatively recent colonization histories displayed stronger IBD compared to southern refugial populations in France (de Lafontaine *et al.*, 2013). As beech only arrived in Britain around 3000 BP (Birks, 1989), IBD in the putative native range is likely driven by relatively recent colonization dynamics. In contrast, widespread

translocations are likely to have prevented the development of IBD between putative non-native populations, possibly due to the anthropic movement of plant material throughout the country. A similar result was found by Leonardi *et al.* (2012) in beech populations of different levels of fragmentation in central Italy. IBD in less fragmented non-marginal populations was obscured when populations from marginal, fragmented populations where included. This finding was interpreted as a result of intense genetic drift in fragmented populations.

We found a significant decrease in gene diversity (nH_S) in putative non-native sites, suggesting a reduction in genetic diversity due to founder effects, although the magnitude of this difference may have been limited by high gene flow associated with wind pollination. However, no significant difference was found for rarefied allelic richness between sites of different origins. Allelic richness is expected to be more sensitive to reductions in effective population sizes, as rare alleles, which do not contribute considerably to gene diversity, are more likely to be lost first (Nei *et al.*, 1975). Low levels of allelic richness throughout Britain suggest a significant proportion was lost during post-glacial colonization, probably due to founder effects. The lack of a pattern in allelic richness between putative native and putative non-native sites may be due to a lack of sensitivity of the analysis arising from the comparison of two already limited gene pools. In agreement with theoretical predictions, gene diversity throughout Britain displayed similarly high levels to that found on the continent (Jump & Peñuelas, 2006; Buiteveld *et al.*, 2007; Oddou-Muratorio *et al.*, 2008; Sjölund & Jump, 2015). Although putative native sites displayed the highest average private allelic richness (nA_P), this difference was not statistically significant.

Regional patterns of postglacial migration of beech into Britain

Within Britain, neighbouring sites displayed congruent levels of admixture with lower levels

443 of admixture in northern, putative non-native sites (Fig. 5). A gradient of admixture marked
444 the transition of continental regions to Britain, with continental sites assigned predominantly
445 to the blue and grey cluster. This is in contrast to that found in *Quercus* spp. (Petit *et al.,*
446 2002b) and *Populus nigra* (Cottrell *et al.,* 1997; Cottrell *et al.,* 2002b; Cottrell *et al.,* 2004),
447 where much longer periods of translocations have led to a decrease in genetic differentiation.
448 Individuals from the putative non-native site, TAN, may have arisen from translocation of
449 continental stock as it displayed the highest levels of assignment to the blue cluster in Britain,
450 similar to individuals in FRA. Individuals from the south of Britain displayed a relatively higher
451 probability of assignment to the red cluster, which was generally associated with the putative
452 native range, in addition to the south-west region, including putative non-native sites DEV,
453 GOL, HEM, and BRI. There appeared to be a transition between the assignment of individuals
454 to the red and grey cluster that occurred in proximity to the border for the 1000 BP
455 isochrone, with a tendency towards assigning individuals to the red cluster in regions pre-
456 1000 BP. Only one putative native site, CWw, occurred outside of the border of the 1000 BP
457 limit and was predominantly assigned to the red cluster. However, CWw is only 15km away
458 from the 1000 BP isochrones.
459
460 **Conclusion**
461 Using multiple data sources, we were able to identify signals of natural colonization in a
462 forest system that has been heavily impacted by humans. Our results build upon existing
463 palynological and historical evidence to substantially advance our understanding of the
464 contemporary genetic impact of past population translocation. Chloroplast markers suggest
465 that the majority of trees sampled were derived from putative native stock either through
466 natural regeneration or translocation, whilst nuclear markers confirm the persistence of
467 signals of the natural colonization of beech in Britain in putative native sites. Although cryptic
468 genetic signals of population expansion remain in the putative native range of beech, we

caution against using this evidence as a means to classify stand origins, since gene flow between neighbouring regions essentially blurs the borders of the native range. Warming climate is increasing the productivity of this species in more northerly parts of its range while decreasing growth and reproduction are predicted in the south. Given that our data suggest native origin for most of Britain's populations, it is paramount that climate-induced range shifts are considered in management plans and this species managed as a whole, irrespective of regional boundaries.

ACKNOWLEDGEMENTS

We thank all the landowners; T. Bimson for help with field work; L. Cavin for site selection assistance; P. Ruiz-Benito for GIS assistance; E. Tisdall, M. Abdelaziz Mohamed, G. Flint and L. Herridge for advice. This work was funded by the Natural Environment Research Council as part of the ERA-Net BiodivERsA Project 'European Beech Forests for the Future' (BEFOFU) [Grant NE/G002118/1].

REFERENCES

Alessa L. & Iii F.S.C. (2008) Anthropogenic biomes: a key contribution to earth-system science. *Trends in Biotechnology*, **23**, 529–531.

Asuka Y., Tani N., Tsumura Y., & Tomaru N. (2004) Development and characterization of microsatellite markers for *Fagus crenata* Blume. *Molecular Ecology Notes*, **4**, 101–103.

Birks H.J.B. (1989) Holocene isochrone maps and patterns of tree-spreading in the British Isles. *Journal of Biogeography*, **16**, 503–540.

Bradshaw R.H.W. (2004) Past anthropogenic influence on European forests and some possible genetic consequences. *Forest Ecology and Management*, **197**, 203–212.

493 Buiteveld J., Vendramin G.G., Leonardi S., Kamer K., & Geburek T. (2007) Genetic diversity
494 and differentiation in European beech (*Fagus sylvatica* L.) stands varying in
495 management history. *Forest Ecology and Management*, **247**, 98–106.

496 Chybicki I.J., Trojankiewicz M., Oleksa A., Dzialuk A., & Burczyk J. (2009) Isolation-by-distance
497 within naturally established populations of European beech (*Fagus sylvatica*). *Botany*,
498 **87**, 791–798.

499 Comes H.P. & Kadereit J.W. (1998) The effect of Quaternary climatic changes on plant
500 distribution and evolution. *Trends in Plant Science*, **3**, 432–438.

501 Cornuet J.M., Ravigné V., & Estoup A. (2010) Inference on population history and model
502 checking using DNA sequence and microsatellite data with the software DIYABC (v1.0).
503 *BMC Bioinformatics*, **11**, 401.

504 Cottrell J.E., Forrest G.I., & White I.M.S. (1997) The use of RAPD analysis to study diversity in
505 British black poplar (*Populus nigra* L. ssp. *betulifolia* (Pursch) W. Wettst. (Salicaceae)) in
506 Great Britain. Watsonia, **21**, 305–312.

507 Cottrell J., Munro R.C., Tabbener H.E., Gillies A.C.M, Forrest G.I., Deans J.D., & Lowe A.J.
508 (2002a) Distribution of chloroplast DNA variation in British oak (*Quercus robur* and *Q.*
509 *petraea*): the influence of postglacial colonisation and human management. *Forest*
510 *Ecology and Management*, **156**, 181–195.

511 Cottrell J.E., Tabbener H.E., & Forrest G.I. (2002b) Distribution of variation in British black
512 poplar: role for human management. In: van Dam B.C., Bordács S. (Ed.), Genetic
513 Diversity in River Populations of European Black Poplar-implications for Riparian Eco-
514 system Management. Proceedings of an International Symposium, Szekza´rd, Hungary,
515 May 2001. Str. 73-84.

516 Cottrell J.E., Krystufek V., Tabbener H.E., Milner A.D., Connolly T., Sing L., Fluch S., Burg K.,
517 Lefévre F., Achard R., Bordács S., Gebhardt K., Vornam B., Smulders M.J.M., Vanden
518 Broeck A.H., Van Slycken J., Storme V., Boerjan W., Castiglione S., Fossati T., Alba N.,
519 Agundez D., Maestro C., Notivol E., Bovenschen J., & Van Dam B.C. (2005) Postglacial
520 migration of *Populus nigra* L.: lessons learnt from chloroplast DNA. *Forest Ecology and*
521 *Management*, **206**, 71-90.

522 Excoffier L., Foll M., & Petit R.J. (2009) Genetic consequences of range expansions. *Annual*
523 *Review of Ecology, Evolution, and Systematics*, **40**, 481–501.

524 Excoffier L. & Lischer H.E.L. (2010) Arlequin suite ver 3.5: a new series of programs to perform
525 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*,
526 **10**, 564–567.

527 Falush D., Stephens M., & Pritchard J.K. (2007) Inference of population structure using
528 multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*,
529 **7**, 574–578.

530 Feurdean A., Bhagwat S.A., Willis K.J., Birks H.J.B., & Lischke H. (2013) Tree migration-rates:
531 narrowing the gap between inferred post-glacial rates and projected rates. *PLoS*
532 *Genetics*, **8**, 1–7.

533 Firbas F., Losert H. (1949) *Spät- und nacheiszeitliche Waldgeschichte Mitteleuropas närdlich*
534 *der Alpen.* Fischer, Germany.

535 Fuentes-Utrilla P., Venturas M., Hollingsworth P.M., Squirrell J., Collada C., Stone G.N., & Gil L.
536 (2014) Extending glacial refugia for a European tree: genetic markers show that Iberian
537 populations of white elm are native relicts and not introductions. *Heredity*, **112**, 105–
538 113.

539 Goudet J. (1995) FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of*
540 *Heredity*, **86**, 485–486.

541 Guillot G., Leblois R., Coulon A., & Frantz A.C. (2009) Statistical methods in spatial genetics.
542 *Molecular Ecology*, **18**, 4734–4756.

543 Hardy O.J. & Vekemans X. (2002) Spagedi: a versatile computer program to analyse spatial
544 genetic structure at the individual or population levels. *Molecular Ecology Notes,* **2**,
545 618–620.

546 Hewitt G.M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

547 Hubisz M.J., Falush D., Stephens M., & Pritchard J.K. (2009) Inferring weak population
548 structure with the assistance of sample group information. *Molecular Ecology*
549 *Resources*, **9**, 1322–1332.

550 Huntley A.B., Bartlein P.J., Prentice I.C., Journal S., Nov N., Huntley B., Bartlein P.J., & Road S.
551 (1989) Climatic control of the distribution and abundance of Beech (*Fagus* L.) in Europe
552 and North America. *Journal of Archaeological Science*, **16**, 551–560.

553 Jakobsson M. & Rosenberg N.A. (2007) CLUMPP: a cluster matching and permutation
554 program for dealing with label switching and multimodality in analysis of population
555 structure. *Bioinformatics*, **23**, 1801–1806.

556 Jump A.S., Cavin L., & Hunter P.D. (2010) Monitoring and managing responses to climate
557 change at the retreating range edge of forest trees. *Journal of Environmental*
558 *Monitoring*, **12**, 1791–1798.

559 Jump A.S. & Peñuelas J. (2006) Genetic effects of chronic habitat fragmentation in a wind-
560 pollinated tree. *Proceedings of the National Academy of Sciences of the United States of*
561 *America*, **103**, 8096–8100.

562 Jump A.S., Rico L., Coll M., & Peñuelas J. (2012) Wide variation in spatial genetic structure
563 between natural populations of the European beech (*Fagus sylvatica*) and its
564 implications for SGS comparability. *Heredity*, **108**, 633–639.

565 Kalis A.J. & Merkt J. (2003) Environmental changes during the Holocene climatic optimum in
566 central Europe - human impact and natural causes. Quaternary Science Reviews, **22**, 33–
567 79.

568 Kelleher C.T., Hodkinson T.R., Kelly D.L., & Douglas G.C. (2004) Characterisation of chloroplast
569 DNA haplotypes to reveal the provenance and genetic structure of oaks in Ireland.
570 *Forest Ecology and Management*, **189**, 123–131.

571 King, A.R., & Ferris, C. (1998). Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn.
572 *Molecular ecology*, **7**, 1151-1161.

573 De Lafontaine G., Ducousso A., Lefèvre S., Magnanou E., & Petit R.J. (2013) Stronger spatial
574 genetic structure in recolonized areas than in refugia in the European beech. *Molecular*
575 *Ecology*, **22**, 4397–4412.

576 Loveless M.D. & Hamrick J.L. (1984) Ecological determinants of genentic structure in plant
577 populations. *Annual Review of Ecology and Systematics*, **15**, 65–95.

578 Magri D. (2010) Persistence of tree taxa in Europe and Quaternary climate changes.
579 *Quaternary International*, **219**, 145–151.

580 Magri D., Vendramin G.G., Comps B., Dupanloup I., Geburek T., Gömöry D.S., Litt T., Paule L.,
581 Roure J.M., & Tantau I. (2006) A new scenario for the Quaternary history of European
582 beech populations: palaeobotanical evidence and genetic consequences. *New*
583 *Phytologist*, **171**, 199–221.

584 Nei M. (1978) Estimation of average heterozygosity and genetic distance from a small
585 number of individuals. *Genetics*, **89**, 583–590.

586 Nei M. (1987) *Molecular evolutionary genetics.* Columbia University Press, USA.

587 Nocentini S. (2009) Structure and management of beech (*Fagus sylvatica* L.) forests in Italy.
588 *iForest - Biogeosciences and Forestry*, **2**, 105–113.

589 Oddou-Muratorio S., Vendramin G.G., Buiteveld J., & Fady B. (2008) Population estimators or
590 progeny tests: what is the best method to assess null allele frequencies at SSR loci?
591 *Conservation Genetics*, **10**, 1343–1347.

592 Van Oosterhout C., Hutchinson W.F., Wills D.P.M., & Shipley P. (2004) Micro-Checker:
593 software for identifying and correcting genotyping errors in microsatellite data.
594 *Molecular Ecology Notes*, **4**, 535–538.

595 Overballe-Petersen M. V., Nielsen a. B., Hannon G.E., Halsall K., & Bradshaw R.H. (2012) Long-
596 term forest dynamics at Gribskov, eastern Denmark with early-Holocene evidence for
597 thermophilous broadleaved tree species. *The Holocene*, **23**, 243–254.

598 Packham J.R., Thomas P.A., Atkinson M.D., & Degen T. (2012) Biological flora of the British
599 Isles: *Fagus sylvatica*. *Journal of Ecology*, **100**, 1557–1608.

Pascual M., Chapuis M.P., Mestres F., Balanyà J., Huey R.B., Gilchrist G.W., Serra L., & Estoup A. (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology*, **16**, 3069–3083.

Pastorelli R., Smulders M.J.M., Wastende V., Vosman B., Giannini R., Vettori C., & Vendramin G.G. (2003) Characterization of microsatellite markers in *Fagus sylvatica* L . and *Fagus orientalis* Lipsky. *Molecular Ecology Notes*, **96**, 76–78.

Peterken G.F. & Mountford E.P. (1996) Effects of drought on beech in Lady Park Wood, an unmanaged mixed deciduous woodland. *Forestry*, **69**, 125-136.

Petit R.J., Mousadik A.E.L., & Pons O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Petit J.R., Csaikl U.M., Bordács S., Burg K., Coart E., Cottrell J., Van Dam, B., Deans, J.D., Dumolin-Lapegues S., Fineschi S., Finkeldey R., Gillies A., Glaz I., Goicoechea P.G., Jensen J. S., König A. O., Iowe A.J., Madsen S.F., Matyás C., Munro R.C., Olalde, M., Pemonge M.H., Popescu F., Slade D., Tabbener H., Taurchini D., De Vries, S.G.M., Ziegenhagen G., & Kremer A. (2002a) Chloroplast DNA variation in European white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, **156**, 5–26.

Petit J.R., Brewer S., Bordács S., Burg K., Cheddadi R., Coart E., Cottrell J., Csaikl U.M., Van Dam B., Deans J.D., Espinel S., Fineschi S., Finkeldey R., Glaz I., Goicoechea P. G., Jensen J.S., König A.O., Lowe A.J., Madsen S.F., Mátyás C., Munro R.C., Popescu F., Slade D., Tabbener H., De Vries, S.G.M., Ziegenhagen B., De Beaulieu J-L., & Kremer A. (2002b) Identification of refugia and post-glacial colonisation routes of European

white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49-74.

Petit, R. J., & Hampe, A. (2006). Some Evolutionary Consequences of Being a Tree. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 187–214.

Piotti A., Leonardi S., Heuertz M., Buiteveld J., Geburek T., Gerber S., Kramer K., Vettori C., & Vendramin G.G. (2013) Within-population genetic structure in beech (*Fagus sylvatica* L.) stands characterized by different disturbance histories: does forest management simplify population substructure? *PloS One*, **8**, e73391.

Pott R. (2000) Palaeoclimate and vegetation - long-term vegetation dynamics in central Europe with particular reference to beech. *Phytocoenologia*, **30,** 285–333.

Preston C.D., Pearman D., Dines T. (2002) *New atlas of the British & Irish flora.* Oxford University Press, UK.

Pritchard J.K., Stephens M., & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–59.

Rannala B. & Mountain J. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9197–9201.

Reboud X. & Zeyl C. (1994) Organelle inheritance in plants. *Heredity*, **72**, 132–140.

Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

642 Savolainen O., Kärkkäinen K. (1992) Effect of forest management on gene pools. *New Forests*,
643 **6**, 329–345.

644 Schaberg P.G., DeHayes D.H., Hawley G.J., & Nijensohn S.E. (2008) Anthropogenic alterations
645 of genetic diversity within tree populations: Implications for forest ecosystem resilience.
646 *Forest Ecology and Management*, **256**, 855–862.

647 Sebastiani F., Carnevale S., & Vendramin G.G. (2004) A new set of mono- and dinucleotide
648 chloroplast microsatellites in Fagaceae. *Molecular Ecology Notes*, **4**, 259–261.

649 Sjölund M.J. & Jump A.S. (2015) Coppice management of forests impacts spatial genetic
650 structure but not genetic diversity in European beech (*Fagus sylvatica* L.). *Forest Ecology*
651 *and Management*, **336**, 65–71.

652 Stone G.N. & Sunnucks P. (1993) Genetic consequences of an invasion through a patchy
653 environment - the cynipid gallwasp Andricus. *Molecular Ecology*, **2**, 251–268.

654 Szpiech Z.A., Jakobsson M., & Rosenberg N.A. (2008) ADZE: a rarefaction approach for
655 counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498–2504.

656 Taberlet P., Fumagalli L., Wust-Saucy A.G., & Cossons J.F. (1998) Comparative phylogeography
657 and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

658 Tanaka K., Tsumura Y., & Nakamura T. (1999) Development and polymorphism of
659 microsatellite markers for *Fagus crenata* and the closely related species , *F . japonica*.
660 *Theoretical and Applied Climatology*, **99**, 11–15.

661 Vornam B., Decarli N., & Gailing O. (2004) Spatial distribution of genetic variation in a natural
662 beech stand (*Fagus sylvatica* L.) based on microsatellite markers. *Conservation Genetics*,
663 **5**, 561–570.

664 Wagner S., Collet C., Madsen P., Nakashizuka T., Nyland R.D., & Sagheb-Talebi K. (2010) Beech
665 regeneration research: From ecological to silvicultural aspects. *Forest Ecology and*
666 *Management*, **259**, 2172–2182.

667 Watt A.S. (1923) On the ecology of British beechwoods with special reference to their
668 regeneration. *Journal of Ecology*, **11**, 1–48.

669 Watt A.S. (1931) Preliminary observations on Scottish beechwoods. Introduction and part I.
670 *Journal of Ecology*, **19**, 137–157.

671 Weir B.S. & Cockerham C.C. (1984) Estimating F-statistics for population structure. *Evolution*,
672 **38**, 1358–1370.

673 Weising K. & Gardner R. C. (1999). A set of conserved PCR primers for the analysis of simple
674 sequence repeat polymorphisms in chloroplast genomes of dicotyledonous
675 angiosperms. *Genome*, **42**, 9-19.

676 Wright S. (1940) Breeding structure of populations in relation to speciation. *The American*
677 *Naturalist*, **74**, 232–248.

678

679

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Stand origins for study sites

Appendix S2 Chloroplast and nuclear genetic data

Appendix S3 STRUCTURE analysis output

Data Accessibility

All microsatellite and GPS data for this study are available at DataSTORRE: Stirling Online Repository for Research Data: http://hdl.handle.net/11667/90

Biosketch

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

Author contributions: MJS and ASJ designed the research. MJS conducted field-based work. MJS, JJM, and PGD conducted lab-based work. MJS and PGD conducted data analysis. ASJ supervised the research project. MJS, ASJ, PGD, and JJM wrote the manuscript.

Editor: Jim Provan

Tables

Table 1 Genetic diversity estimates obtained from nuclear (*n*) and chloroplast (*c*) markers for beech in Great Britain. Data from chloroplast (*c*) markers include; *cN,* no. of samples; cH_N, no. of haplotypes, † identifies private haplotypes (cH_P). Data from nuclear (*n*) markers include; *nN,* no. of samples; nA_R, rarefied allelic richness; nA_P, rarefied private allelic richness; nH_S, gene diversity; nF_{IS}, inbreeding coefficient; and, *nDS%,* percentage of significantly differentiated sites. Mean±SE is given per group (Putative *Native* and *Non-native)* and overall for genetic diversity estimators. Significant *P*-values are indicated as * $P < 0.05$, *** $P < 0.001$.

	Sites	*cN*	*cHN*	*nN*	nA_R	nA_P	nH_S	nF_{IS}	*nDS (%)*
Native		**321**	**1.5±0.2**	**340**	**6.22±0.13**	**0.053±0.13**	**0.707±0.006***	**0.022±0.010**	**46.1±6.4**
	FEL	17	2	20	5.17	0.095	0.651	0.057	75.6
	BED	18	1	20	5.95	0.025	0.722	0.059	17.1
	SEC	17	1	20	6.06	0.008	0.717	-0.040	56.1
	WYC	20	1	20	7.12	0.077	0.694	0.088*	9.8
	LAD	20	1	20	6.72	0.110	0.736	0.033	22.0
	BUC	20	1	20	5.33	0.000	0.682	0.010	87.8
	CWe	20	1	20	6.64	0.146	0.750	0.055	85.4
	CWw	20	1	20	6.00	0.000	0.710	-0.037	85.4
	MON	19	1	20	6.71	0.156	0.716	-0.013	56.1
	GRE	20	1	20	5.98	0.004	0.712	0.038	26.8
	BUR	19	2	20	6.80	0.109	0.713	-0.003	22.0
	SAV	16	1	20	6.09	0.001	0.687	0.041	43.9
	LUL	17	3	20	6.09	0.020	0.727	-0.054	29.3
	FRI	20	1	20	6.51	0.087	0.679	0.007	26.8
	BLE	18	4†	20	6.82	0.027	0.721	0.030	24.4
	WEA	20	2	20	5.86	0.007	0.703	0.070	56.1
	DEN	20	1	20	5.97	0.024	0.706	0.031	58.5
Non-native		**481**	**1.4±0.1**	**497**	**6.25±0.11**	**0.037±0.008**	**0.690±0.006***	**-0.006±0.010**	**32.9±4.5**
	APP	19	2	19	6.68	0.065	0.682	0.004	14.6
	BAR	19	2	19	6.50	0.001	0.702	-0.012	17.1
	BEE	19	2	20	4.58	0.000	0.606	-0.152***	95.1
	BRI	20	1	20	6.21	0.080	0.663	-0.065	58.5
	CAR	18	1	19	5.92	0.000	0.685	0.057	26.8
	CLE	20	1	20	6.30	0.014	0.695	0.014	17.1
	CRA	20	1	20	5.29	0.000	0.706	-0.086*	75.6
	DEV	16	2†	20	6.30	0.038	0.691	-0.066	22.0
	DRU	20	1	20	6.44	0.042	0.692	-0.012	39.0
	DUN	19	2	20	7.16	0.103	0.684	0.005	24.4
	ECC	20	2	20	6.32	0.025	0.700	0.036	56.1
	GEL	20	1	20	6.32	0.013	0.700	0.014	17.1
	GOL	20	1	20	5.49	0.048	0.672	0.034	36.6
	HEM	20	1	20	5.86	0.001	0.672	-0.042	70.7
	KIN	20	1	20	5.87	0.013	0.675	-0.005	34.1
	MAB	20	1	20	6.14	0.011	0.701	0.089*	9.8
	PLO	17	2	20	6.04	0.013	0.693	0.019	24.4
	STR	20	1	20	6.67	0.099	0.692	-0.012	12.2
	TAL	15	2	20	6.28	0.122	0.667	0.018	17.1
	TAN	20	1	20	7.19	0.077	0.747	0.065	51.2
	TON	20	1	20	6.23	0.000	0.680	0.023	24.4
	TWO	20	1	20	6.33	0.014	0.717	0.003	39.0
	WAL	20	1	20	6.52	0.033	0.708	-0.069	7.3
	WYT	20	2	20	6.79	0.103	0.700	-0.005	4.9
	YEL	19	1	20	6.84	0.000	0.722	-0.012	26.8
Overall		**802**	**1.4±0.1**	**837**	**6.25±0.08**	**0.044±0.007**	**0.696±0.004**	**0.004±0.008**	**38.7±3.8**

Table 2 Hierarchical analysis of molecular variance (AMOVA) for chloroplast and nuclear markers for beech in Great Britain. The degrees of freedom (*df*), percentage of variation explained by each level (Variation (%)), and the relevant F-statistic are presented with significant *P*-values indicated as *** $P < 0.001$.

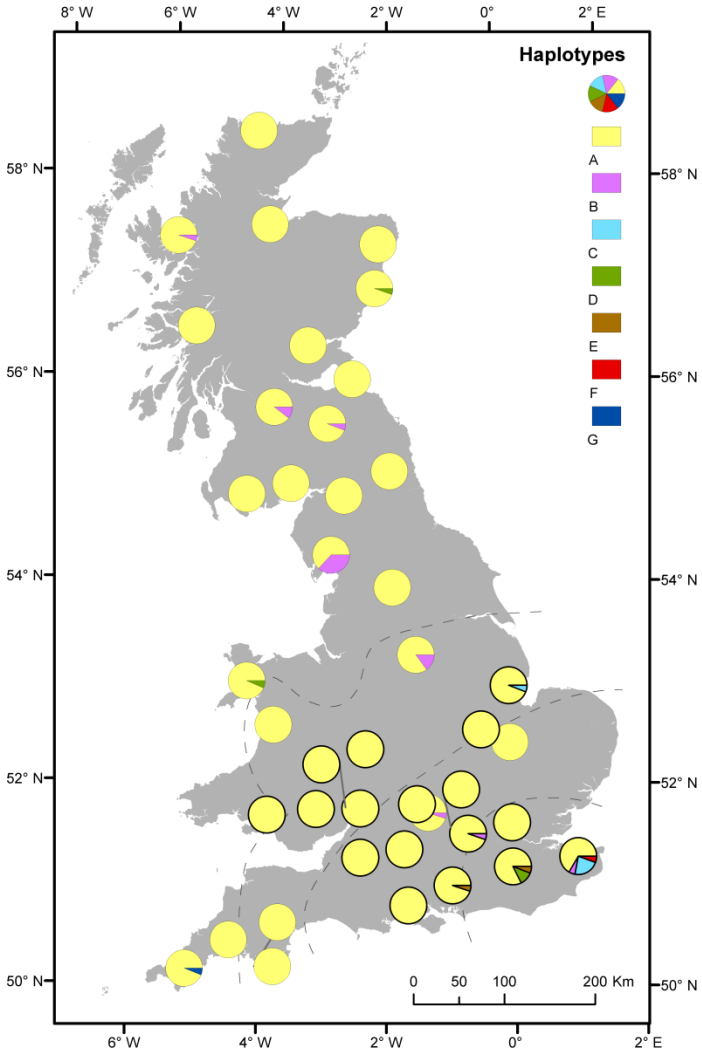
	Chloroplast			Nuclear		
Levels	*df*	Variation (%)	*F*-statistic	*df*	Variation (%)	*F*-statistic
Among groups	1	-0.22	<0.001	1	0.01	0.000
Among sites within groups	40	12.59	0.005***	40	2.25	0.022***
Within sites	763	87.63	0.032***	795	97.74	0.023***

Figures

Fig. 1 Beech in the putative non-native range in Scotland, Great Britain (site CRA, N57.5777 W4.1435). Photo courtesy of M. J. Sjölund.

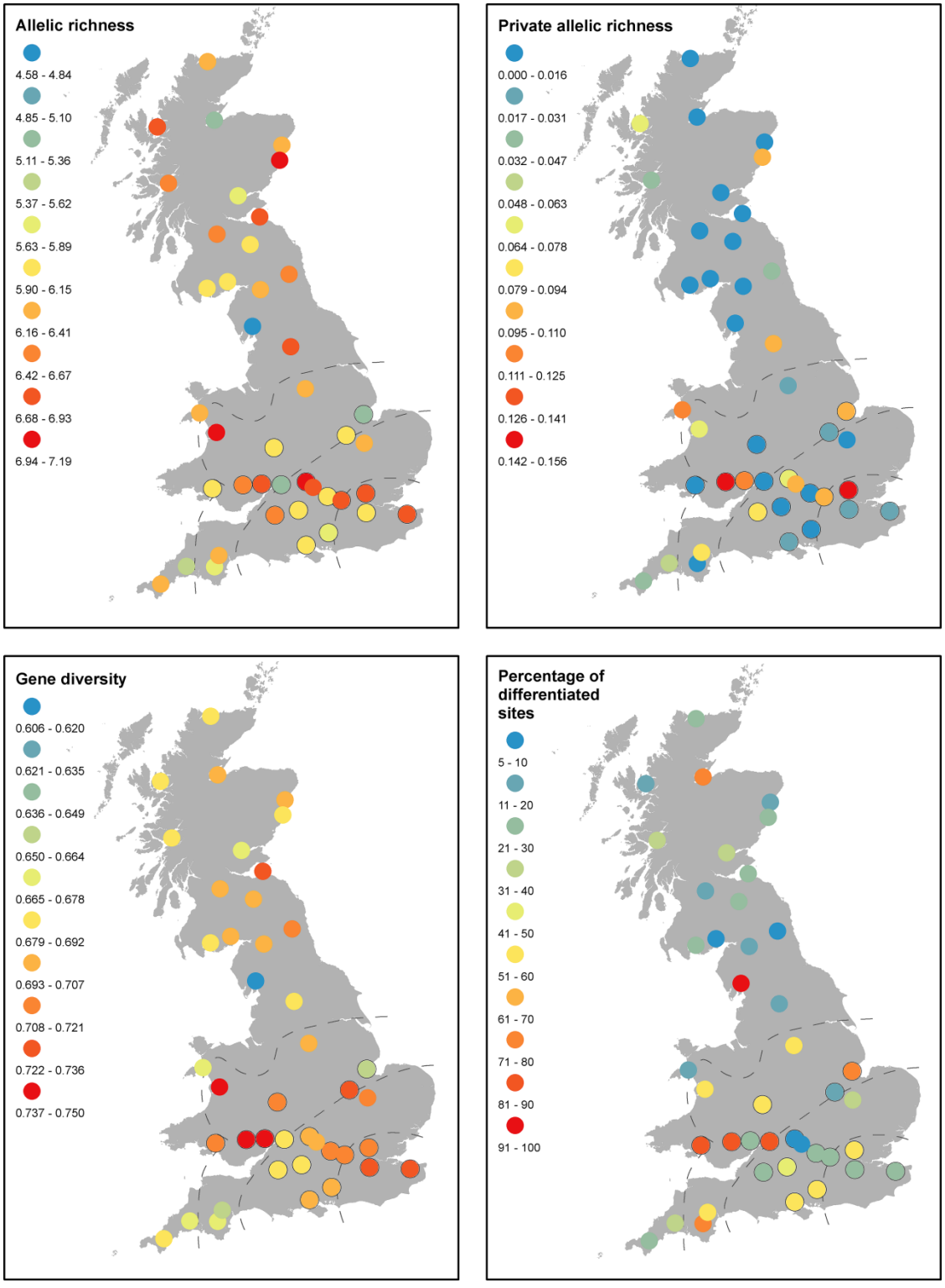
816 **Fig. 2 Geographical variation of chloroplast haplotypic diversity of beech in Great Britain.** A
817 total of seven haplotypes are displayed against Birks' (1989) isochrones. Putative native sites
818 are outlined.

819



820

821 **Fig. 3 Geographical variation of estimators of beech nuclear genetic diversity in Great**
822 **Britain.** Estimators include rarefied allelic richness (nA_R), rarefied private allelic richness (nA_P),
823 gene diversity (nH_S), and the percentage of significantly differentiated sites (*nDS%*). Putative
824 native sites are outlined, with Birks' (1989) isochrones for *F. sylvatica* redrawn as broken
825 lines.


Allelic richness
4.58 - 4.84
4.85 - 5.10
5.11 - 5.36
5.37 - 5.62
5.63 - 5.89
5.90 - 6.15
6.16 - 6.41
6.42 - 6.67
6.68 - 6.93
6.94 - 7.19
Private allelic richness
0.000 - 0.016
0.017 - 0.031
0.032 - 0.047
0.048 - 0.063
0.064 - 0.078
0.079 - 0.094
0.095 - 0.110
0.111 - 0.125
0.126 - 0.141
0.142 - 0.156
Gene diversity
0.606 - 0.620
0.621 - 0.635
0.636 - 0.649
0.650 - 0.664
0.665 - 0.678
0.679 - 0.692
0.693 - 0.707
0.708 - 0.721
0.722 - 0.736
0.737 - 0.750
Percentage of differentiated sites
5 - 10
11 - 20
21 - 30
31 - 40
41 - 50
51 - 60
61 - 70
71 - 80
81 - 90
91 - 100

826

827

828

Fig. 4 Comparison of isolation-by-distance analyses amongst beech populations in putative native and putative non-native regions in Great Britain and continental Europe. Sites included in the analysis are as follows; A) British sites only; B) British and continental sites; C) putative native sites; D) putative non-native sites.

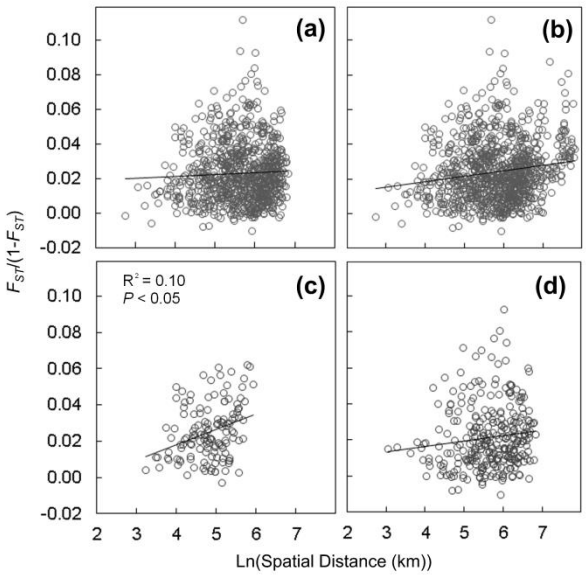
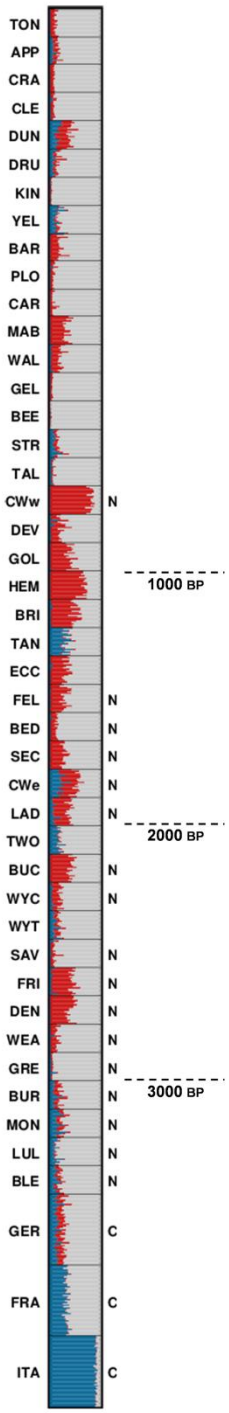


Fig. 5 Regional genetic structure of beech in Great Britain. Three clusters are shown in blue, red, and grey. Each horizontal bar represents an individual with the proportions of its genetic make-up assigned probabilistically to each of the three clusters. Sites are ordered on an approximate geographical gradient by ordering sites following Birks' (1989) isochrones to reflect the putative migration route of Beech into Britain. Continental samples are situated at the bottom of the figure, with a general northward trend to the top of the graph. Stand history of the sites are indicated on the right of the Q-matrix, with continental sites labelled C, putative native sites N, and putative non-native sites left blank. Approximate borders of the isochrones are indicated by a dashed line with years in BP, with site codes on the left.



851

852

853

854

855

Fig. 6 Comparison of modelled colonization scenarios of beech in Great Britain. There are four assumed populations of beech: South (SGB), Central (CGB) and North (NGB) populations of Great Britain, and continental Europe (EUR). All scenarios assume colonization from EUR towards SGB and CGB and either assumes NGB originates from EUR (scenario 1 - a); from natural colonization of SGB and CGB (scenario 2 - b); or from the admixture of both EUR and British CGB and SGB (scenario 3 - c). *t1*, *t2*, and *t3* correspond to the divergence times in generations.