Full title: Genetic signatures of divergent selection in European beech (*Fagus sylvatica L.*) are associated with the variation in temperature and precipitation across its distribution range.

Running title: SNP variation throughout European beech distribution

Authors: Postolache D.[a1], Oddou-Muratorio S.[a*2,3], Vajana E. [4], Bagnoli F.[5], Guichoux E.[6], Hampe A.[6], Le Provost G.[6], Lesur I.[6,7], Popescu F.[1], Scotti I.[2], Piotti A.[5], Vendramin G.G.[5]

Authors' affiliations:

[1] National Institute for Research and Development in Forestry "Marin Drăcea", Bd. Eroilor 128, Voluntari, 077190 Ilfov, Romania

[2] INRAE, URFM, Avignon, France

[3] ECOBIOP Université de Pau et des Pays de l'Adour, E2S UPPA, INRAE, ECOBIOP, Saint-Pée-sur-Nivelle, France

[4] Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[5] Institute of Biosciences and Bioresources, National Research Council, Via Madonna del Piano 10, 50019 Sesto Fiorentino (Firenze), Italy

[6] Université de Bordeaux, INRAE, BIOGECO, F-33610 Cestas, France

[7] HelixVenture, F-33700 Mérignac, France

*corresponding author: sylvie.muratorio@inrae.fr

[a] co-first authors

Abstract

High genetic variation and extensive gene flow may help forest trees with adapting to ongoing climate change, yet the genetic bases underlying their adaptive potential remain largely unknown. We investigated range-wide patterns of potentially adaptive genetic variation in 64 populations of European beech (*Fagus sylvatica* L.) using 270 SNPs from 139 candidate genes involved either in phenology or in stress responses. We inferred neutral genetic structure and processes (drift and gene flow) and performed differentiation outlier analyses and gene-environment association (GEA) analyses to detect signatures of divergent selection.

Beech range-wide genetic structure was consistent with the species' previously identified postglacial expansion scenario and recolonization routes. Populations showed high diversity and low differentiation along the major expansion routes. A total of 52 loci were found to be putatively under selection and 15 of them turned up in multiple GEA analyses. Temperature and precipitation related variables were equally represented in significant genotype-climate associations. Signatures of divergent selection were detected in the same proportion for stress response and phenology-related genes. The range-wide adaptive genetic structure of beech appears highly integrated, suggesting a balanced contribution of phenology and stress-related genes to local adaptation, and of temperature and precipitation regimes to genetic clines. Our results imply a best-case scenario for the maintenance of high genetic diversity during range shifts in beech (and putatively other forest trees) with a combination of gene flow maintaining within-population neutral diversity and selection maintaining between-population adaptive differentiation.

Keywords: candidate gene, phenology, drought stress, divergence outlier, genotype-environment associations analyses, forest tree, local adaptation

Introduction

Local adaptation is pervasive in forest tree populations (Alberto, Aitken, Alía, González-Martínez, et al., 2013; Savolainen, Pyhäjärvi, & Knürr, 2007) and is expected to play a major role in their response to ongoing environmental changes (Fady et al., 2016). Local adaptation implies that some key adaptive traits are genetically differentiated among populations, and thus that individual populations could evolve differently to the same environmental stress due to their different genetic setup (Kawecki & Ebert, 2004). Most temperate tree species have developed their present-day geographical patterns of local adaptation following considerable a range expansion from their glacial refugia after the Last Glacial Maximum (19.5-26 kyr BP, Clark et al., 2009; de Lafontaine, Napier, Petit, & Hu, 2018). There is ample concern that population processes going along with this postglacial range expansion such as founder events, genetic drift, or allele surfing might have left lasting imprints that could compromise the correct identification of adaptive genetic variation in extant natural tree populations (de Villemereuil, Frichot, Bazin, François, & Gaggiotti, 2014; Hoban et al., 2016; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). Yet very few surveys of adaptive genetic variation in forest trees have to date assembled two prerequisites to properly account for species' postglacial population dynamics: (i) a rangewide perspective with a sampling that thoroughly replicates populations from different regions and (ii) independent palaeoecological evidence documenting the species' postglacial

range dynamics and expansion (de Lafontaine et al., 2018). Empirical research to elucidate the issue is urgently needed because knowledge on local adaptation is crucial for conceiving conservation and management practices to adapt forest tree species and their ecosystems to ongoing environmental change (Aitken & Whitlock, 2013; Fady et al., 2016; Oney, Reineking, O'Neill, & Kreyling, 2013).

Reciprocal transplant experiments have been the classic approach to investigate local adaptation in forest trees, highlighting that phenotypes generally match their environment in extant populations (Rehfeldt et al., 2002; Savolainen et al., 2007). Provenance trials have also shown that trees generally have high levels of phenotypic plasticity at adaptive traits, and high levels of genetic variability within populations (Alberto et al., 2013; Gárate-Escamilla et al., 2019). This combination of local adaptation (i.e., mean trait values close to the optimum) and high within-population variation at key adaptive traits (i.e., large variance around the means) indicates a high genetic load, which in turn can be an asset when facing a swift environmental change (Savolainen et al., 2007). The development of population genomics has provided complementary approaches to study local adaptation (Hoban et al., 2016; Lind, Menon, Bolte, Faske, & Eckert, 2018). Two approaches, in particular, have become widely used to identify loci involved in local adaptation for non model organisms: differentiation outlier analyses, which aim at identifying loci with disproportionate allele frequency differentiation among populations, and gene-environment association (GEA) analyses , which aim at identifying loci exhibiting significant correlations with ecological variables. A key strength of such approaches is that they are cost-efficient for targeting large numbers of populations across well-defined environmental gradients, and therefore for

90 investigating ecological hypotheses on the genomic basis of local adaptation (Capblancq et
91 al., 2020; Rellstab et al., 2016; Temunović et al., 2020).

92 Present-day patterns of adaptive phenotypic and genetic differentiation have built up
93 relatively quickly. It is only (at most) a few hundreds of generations ago that most temperate
94 forest tree species recolonized large areas becoming available after the Last Glacial
95 Maximum. Studies combining extensive surveys of fossil records (pollen and macro-remains)
96 and of population genetic variation have been able to provide detailed direct evidence of the
97 number and spatial location of glacial refugia, postglacial expansion routes, hybrid zones,
98 and the timing of expansion events (e.g., de Lafontaine, Amasifuen Guerra, Ducousso, &
99 Petit, 2014; Magri et al., 2006). This knowledge represents a highly valuable baseline
100 information for disentangling demographic effects from those of selection, which is a major
101 challenge for differentiation outlier and GEA analyses (de Villemereuil et al., 2014;Frichot,
102 Schoville, de Villemereuil, Gaggiotti, & François, 2015; Hoban et al., 2016; Rellstab et al.,
103 2015). Theoretical studies have shown that population expansions into new areas can go
104 along with repeated founder effects, increasing random fluctuations of allele frequencies,
105 possibly leading to the rise of neutral mutations to high frequencies ("allele surfing"), loss of
106 genetic diversity and strong spatial genetic structure (SGS) along the expansion axis (de
107 Lafontaine, Ducousso, Lefèvre, Magnanou, & Petit, 2013; Excoffier, Foll, & Petit, 2009;
108 Slatkin, 1993). Allele surfing, in particular, could be mistaken for the increase in allelic
109 frequency of a beneficial mutation propagated by selection (Paulose & Hallatschek, 2020;
110 Ruiz Daniels et al., 2018). As both allele surfing and selection typically affect only a subset of
111 loci in the genome, the former must be carefully considered when screening for the latter in
112 expanding populations. Alternatively, colonization by many individuals should result in high

113 genetic diversity at the colonization front and shallow SGS, particularly when founders
114 originate from a variety of source populations. The effective number of founders depends on
115 patterns of long-distance dispersal and on a variety of demographic processes and life-
116 history traits (Austerlitz, Mariette, Machon, Gouyon, & Godelle, 2000; Fayard, Klein, &
117 Lefèvre, 2009; Roques, Garnier, Hamel, & Klein, 2012). The actual relevance of these
118 processes during the postglacial range expansion of temperate forest trees and their
119 eventual traces in the present-day population genetic structures remain however under
120 investigated.

121 This study takes advantage of the outstandingly well known postglacial population
122 history of the European beech (*Fagus sylvatica* L.) to investigate the climate-associated
123 genetic variation across its distribution range. Beech putative glacial refugia and colonization
124 routes have been identified based on very detailed pollen records and genetic population
125 surveys with chloroplast and isozyme markers (Magri et al., 2006). Beech is known to be
126 highly sensitive to summer droughts (Aranda et al., 2015; Knutzen, Dulamsuren, Meier, &
127 Leuschner, 2017), and, to a lesser extent, to late frosts (Kreyling et al., 2014; Petit-Cailleux et
128 al., 2020). Genetic variation has been investigated at various climate-related phenological
129 traits (Gárate-Escamilla, Hampe, Vizcaíno-Palomar, Robson, & Benito Garzón, 2019; Gauzere,
130 Klein, Brendel, Davi, & Oddou-Muratorio, 2020; Gömöry & Paule, 2011; Kramer et al., 2017;
131 Vitasse, Delzon, Bresson, Michalet, & Kremer, 2009), physiological or morphological traits
132 (Bresson, Vitasse, Kremer, & Delzon, 2011; Hajek, Kurjak, von Wühlisch, Delzon, & Schuldt,
133 2016; Wortemann et al., 2011) and performance traits (Gárate-Escamilla et al., 2019).
134 Phenological traits such as budburst and leaf senescence show consistent patterns of genetic
135 variation across latitude or elevation at various spatial scales, with populations of higher

136 elevation or latitude flushing earlier than populations from low elevation or latitude in
137 common garden conditions (Gauzere et al., 2020; Gömöry & Paule, 2011; Vitasse et al.,
138 2009). Genetic variation for performance traits, such as growth and juvenile survival, also
139 shows a spatial structure, driven by spatial variations of maximal potential
140 evapotranspiration (Gárate-Escamilla et al., 2019). By contrast, other functional traits
141 involved in photosynthesis and transpiration are usually only weakly differentiated among
142 populations but show instead high within-population variation (Hajek et al., 2016). These
143 contrasting patterns of differentiation raise questions about the role of phenological and
144 physiological traits in local adaptation and the spatial scales at which it takes place. Only a
145 few published studies have so far used genetic approaches to investigate local adaptation,
146 and mostly at local to regional scales (Capblancq et al., 2020; Csilléry et al., 2014; Cuervo-
147 Alarcon et al., 2021; Krajmerová et al., 2017; Lalagüe et al., 2014; Müller, Seifert, &
148 Finkeldey, 2015; Pluess et al., 2016).

149 This study investigates range-wide patterns of adaptive genetic variation in beech
150 using 405 SNPs that are located in candidate genes involved either in budburst phenology
151 and dormancy regulation (Lalagüe et al. 2014; Lesur et al. 2015) or in response to stresses
152 (Lalagüe et al. 2014). We genotyped 446 individuals from 64 populations covering the entire
153 species range (Fig. 1) to address the three following questions: (Q1) What are the risks that
154 past population demography, including post-glacial recolonization, blur potential selective
155 imprints on genetic structure? We expect limited genetic drift and allele surfing in beech.
156 (Q2) Do certain loci show imprints of local adaptation? We expect genes related to
157 phenological traits to show stronger signals of selection compared to genes related to stress-
158 response traits. (Q3) if the spatial and climatic effects can be separated, what are the

159 respective impacts of temperature- versus precipitation-related variables on adaptive
160 differentiation? In line with Q2, we expected the temperature variables to stand out more
161 than precipitation variables.

162 ## Materials and Methods

163 ### *Sampling design*

164 European beech is a dominant broadleaved tree species of many lowland and
165 mountain forests across Europe, extending from Spain to the Carpathians and from Sicily to
166 southern Sweden. We sampled 446 adult trees in 64 populations across Europe (Fig. 1A),
167 thoroughly covering the geographical range and bioclimatic niche of beech (Fig. 1B) and
168 including all the major glacial refugia identified by Magri et al. (2006). Leaves were collected
169 from 4 to 10 (7 on average) haphazardly chosen dominant adult trees (at a minimal distance
170 of 40 m from each other) growing in native beech stands.

171 ### *SNP development, genotyping and filtering*

172 Nuclear DNA was extracted from 20-30 mg of dry leaf tissue per individual with the
173 DNeasy Plant Mini Kit (QIAGEN) following the manufacturer's instructions. DNA
174 concentration was measured on a ND-8000 NanoDrop spectrophotometer (Thermo
175 Scientific, Wilmington, USA). Samples were genotyped at 405 SNPs distributed in two
176 multiplex assays.

177 The first assay of 165 SNPs previously developed by Lalagüe et al. (2014) was carried
178 out using Kompetitive Allele Specific PCR (KASP; He, Holme, & Anthony, 2014). Among those,
179 37 SNPs were located in 15 genes annotated as phenology-related in *Quercus petraea*. The
180 other 128 SNPs were located in 37 stress-related genes that were selected from different

sources (see Lalagüe et al., 2014 for details) : (1) literature on candidate genes involved in plant response to abiotic stress; (2) sequences of three proteins involved in cavitation resistance in beec); (3) annotated amplicons from *Q. petraea* and *Q. robur*.

The second assay targeted 240 SNPs in six multiplexes of 40 SNPs each. This assay was developed for this study from available genomic resources (Lesur et al., 2015) and included 104 SNPs located in 51 genes differentially expressed in quiescent buds (QB) as well as 116 SNPs located in 58 genes differentially expressed in swelling buds (SB). This assay also included 20 unrelated control SNPs located in 15 housekeeping genes. Genotyping was performed on a MassARRAY System (Agena Bioscience, USA) using the iPLEX Gold chemistry following Gabriel *et al.* (2009). Data analysis was performed with Typer Analyzer 4.0.26.75 (Agena Bioscience). We filtered out all monomorphic SNPs, as well as loci with a weak or ambiguous signal (i.e., displaying more than three clusters of genotypes or unclear cluster delimitation).

Raw variant data were filtered with *Plink* (v.1.9; Chang et al., 2015). Individuals and SNPs with >15% missing data were filtered out, together with variants with MAF<1%.

Linkage disequilibrium

Pairwise linkage disequilibrium (LD) estimates were obtained within each of the genetic clusters identified by genetic clustering analyses (see below) using the *LD()* function in the R package *genetics* (Warnes, Gorjanc, Leisch, & Man, 2021). The *p.adjust()* function was used to correct *p*-values for multiple testing with the Bonferroni method, and association between allele frequencies was deemed as statistically significant at a nominal significance threshold equal to 1×10^{-3}. An *ad hoc* algorithm was devised to iteratively identify

and remove the loci most frequently involved in pairwise significant tests, leading to cluster-specific lists of SNPs showing statistical evidence of linkage. The markers shared across all analyses were removed to obtain an LD-pruned version of the SNP dataset to be used in the analyses assuming linkage equilibrium among loci (*i.e.*, STRUCTURE and pcadapt, see below).

Population genetic structure

The Bayesian clustering analysis implemented in STRUCTURE (Pritchard, Stephens, & Donnelly, 2000) and the multivariate method implemented in the Discriminant Analysis of Principal Components (DAPC; Jombart, Devillard, & Balloux, 2010) were used to infer patterns of population structuring and admixture among beech populations. The main difference between the two methods is that STRUCTURE builds genetic clusters so to minimise the overall departures from HWE, whereas DAPC is based on maximizing the differentiation between inferred genetic clusters while minimizing variation within them.

STRUCTURE 2.3.4 (Pritchard et al., 2000) was run using default settings and parameter values, assuming the admixture model, and the putative number of different genetic clusters (K) ranging from one to 10. Each run consisted of 5×10^4 burn-in iterations and 1×10^5 data collection iterations. Ten independent runs were performed for each value of K. The average likelihood and ΔK statistics described in Evanno *et al.* (2005) were calculated for each K and used to identify the most-likely K-value. For informative values of K, distinct runs were averaged using CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) to obtain the final estimates of the membership coefficients (*q*-values) at individual and population levels. To comply with model assumptions, STRUCTURE was run first with the

224 complete SNPs dataset, and then with the LD-pruned version of the dataset (once LD
225 estimated within each cluster).

226 DAPC is based on a discriminant analysis (DA) of genetic data preceded by a few
227 analytical steps to meet its requirements, all implemented in the R package *adegenet*
228 (Jombart, 2008). Since DA requires *a priori* definition of clusters, K-means clustering of
229 principal components (PC) on individual allele frequencies was first used to identify both
230 group priors and the most likely number of genetic clusters. K-means was run on 150 PCs
231 with K ranging from one to 40, and the Bayesian Information Criterion (BIC) was used to
232 assess the best supported K-value. Then, as DA requires the variables to be uncorrelated and
233 fewer than the number of observations, we used a principal component analysis (PCA) and
234 the randomization approach implemented in the *a.score()* function in *adegenet* to select the
235 number of PCs optimizing the trade-off between power of discrimination and over-fitting.
236 Finally, the DA was run on 23 PCs extracted from the original dataset.

237 ***Basic diversity and differentiation statistics***

238 We computed allelic richness (A_r) and mean number of alleles per locus (N_a) using the
239 R package *diveRsity* (Keenan, McGinnity, Cross, Crozier, & Prodöhl, 2013); percentage of
240 polymorphic loci using GenAlEx 6.5 (Peakall & Smouse, 2012); observed (H_O) and expected
241 (H_E) heterozygosity, Wright's inbreeding coefficient (F_{IS}), and β_{WT} (a plot-specific index of
242 genetic differentiation relative to the entire pool; Weir and Goudet 2017) using the R
243 package *hierfstat* (Goudet, 2005). Parameters of genetic fixation (G_{ST}; Nei, 1977) and
244 differentiation (Jost's *D*; Jost, 2008) among populations/clusters were also calculated with
245 GenAlEx and their statistical significance assessed with 999 permutations.

246 ***Isolation by distance and barriers to gene flow***

247 We estimated spatial genetic structure (SGS) among populations and tested whether
248 geographic distance significantly shaped the patterns of genetic differentiation (as estimated
249 by $F_{ST}/1\text{-}F_{ST}$) using the software SpaGeDi 1.4c (Hardy and Vekemans 2002). To test for
250 isolation by distance (IBD), the ($F_{ST}/1\text{-}F_{ST}$) values were regressed on ln(d_{ij}), where d_{ij} is the
251 spatial distance between populations *i* and *j*. Then, we tested the regression slope (null
252 hypothesis: $blog_{FST} = 0$) using 5,000 permutations of genotypes over populations. These
253 analyses were run both on the 64 populations and within each cluster identified with DAPC.
254 For within-cluster analyses, we retained all the individuals successfully assigned to a given
255 cluster (i.e., those having a *q*-value above the nominal threshold of 0.6).

256 Spatial variation in genetic diversity and gene flow rates were estimated using
257 Estimated Effective Migration Surfaces (EEMS; Petkova et al., 2016). This method tests for
258 regional departures from the IBD model: areas where the decay of genetic differences across
259 geographical distance is higher than expected under an IBD model are considered as
260 suggestive of barriers to gene flow. A user-selected number of demes determines the
261 geographic grid size and possible migration paths between all populations, and the EEMS are
262 calculated by adjusting the migration rates so that the genetic differences obtained under a
263 stepping-stone model match as closely as possible the observed genetic differences. The
264 estimates are subsequently interpolated over the geographic space to provide a surface of
265 observed genetic dissimilarities. We ran the runeems_snps executable with 500,000 burn-in
266 MCMC steps and 2×10^6 subsequent iterations. To reduce the potential influence of grid size,
267 we averaged the results over nine independent runs with different numbers of demes (800,

1200 and 1600 with three repetitions each) and combined the results across the three independent analyses. We assessed convergence of runs, plotted geographic distance and genetic dissimilarity across demes, and generated effective diversity (q) and effective migration rates (m) surfaces using the R package *reemsplots* (Petkova et al., 2016).

Bioclimatic data

Each sampling site was characterized by a set of 19 bioclimatic variables extracted from the WorldClim database v.1.4 (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) with a grid cell resolution of 30-arc second (ca 1×1 km) using DIVA-GIS v.7.5. These bioclimatic variables represent annual trends (e.g., mean annual temperature, annual precipitation), seasonality (e.g., annual range in temperature and precipitation) and extreme or limiting climatic factors (e.g., temperature of the coldest and warmest month, and precipitation of the wettest and driest quarter). To calculate predictors in the following analyses, we used the mean values of these 19 bioclimatic variables (Table S1) over the period from 1950 to 2000.

To reduce the multidimensional bioclimatic data set to a few uncorrelated factors, we performed two PCAs using the R package *FactoMineR* (Lê, Josse, & Husson, 2008), one focusing on the temperature-related variables (BIO1 to 11), and the other focusing on the precipitation-related predictors (BIO12 to 19). The selected principal components of each PCA were used as individual climate variables in *lfmm* and *Sambada* analyses, and combined in matrices to represent the climatic structure in the variance partitioning analyses (see sections below).

Detection of signatures of selection

To detect loci carrying putative signatures of divergent selection (i.e., outliers), we used two differentiation outlier search methods, *pcadapt* (Luu, Bazin, & Blum, 2017) and the F_{ST}-based method by Martins et al. (2016) as implemented in *lea* (Frichot & François, 2015). Moreover, we used two genotype-environment associations analyses, *lfmm* (Frichot, Schoville, Bouchard, & François, 2013) and *Samβada* (Stucki et al., 2017); these approaches are detailed in Supplementary Online Appendix 2. Note that each method allows the correction of outlier detection for the confounding effects of population structure. For all four methods, *p*-values were corrected across multiple tests using the same local False Discovery Rate (FDR) algorithm.

Finally, we annotated the identified outliers. For the loci obtained from Lalagüe et al. (2014), we queried against The Arabidopsis Information Resource (TAIR 11) database using BlastX with an E-value cut-off of 10^{-5}. For the loci obtained from Lesur et al. (2015), the functional annotation of gene sequences containing outlier SNP was also reported (from their Table S2) .

Isolation by distance and environment

To evaluate the respective importance of IBD versus isolation by environment (IBE), we used a variance partitioning approach (Legendre, Fortin, & Borcard, 2015). We partitioned the explanatory power (as expressed by the adjusted R^2) of the climatic and spatial structures on the genetic structure. Genetic structure was obtained through a Principal Coordinate Analysis (PCA) on the matrix of pairwise population genetic distances as returned by GenAlEx (Peakall & Smouse, 2012). Principal coordinates explaining up to 80% of the total

311 variance were included in the response data table. The spatial structure was modelled by
312 distance-based Moran's eigenvector maps (dbMEM; Dray et al., 2006), as suggested by
313 Legendre et al. (2015), and estimated by the *mem()* function in the R package *adespatial*
314 (Dray et al., 2018). We retained only the statistically significant eigenvectors modelling
315 positive spatial autocorrelation and, therefore, describing global patterns *sensu* Jombart et
316 al. (2008). The climatic structure was summarized by temperature- or precipitation-related
317 synthetic climatic variables (see "Bioclimatic data" section above). We assessed the relative
318 contribution of climatic and spatial structure in explaining the genetic structure of
319 populations using the function *varpart()* of the R package *vegan* (Oksanen et al., 2019).
320 Significance of the variance components was calculated through an ANOVA-like permutation
321 test for redundancy analysis (RDA) and partial ReDundancy Analysis (pRDA) based on 10,000
322 permutations (Legendre & Legendre, 2012).

323 All statistical analyses were conducted using R 3.6.2 (R Core Team, 2019) unless
324 otherwise indicated.

325 ## Results

326 ### *SNP dataset*

327 Of the 405 SNPs from the two multiplex arrays, 135 were discarded from the analyses:
328 9 failed to amplify in all samples; 50 were monomorphic; 75 were of poor quality (visual
329 clustering inspection); and one had a call rate <85%. Sixteen individuals were discarded due
330 to a low call rate. The resulting dataset comprised 430 individuals and 270 SNPs
331 (Supplementary Online Appendix A1), with 3-10 individuals per population (6.7 on average).
332 The 270 SNPs included 150 SNPs in 93 phenology genes, 109 SNPs in 38 stress-related genes,

and 11 SNPs in 8 housekeeping genes. After removing linked markers, the LD-pruned version of the dataset comprised 212 SNPs.

Population genetic structure

Both the STRUCTURE and the DAPC analysis revealed an optimal grouping at K=3 (Fig. 2 and Fig. S1) with clear geographic boundaries among the three genetic clusters (as represented by the green, red and blue colors in Fig. 2) and admixture zones (areas with cross symbols in Fig. 2). Populations were considered as admixed when none of the population *q*-values exceeded the threshold of 0.60 for any inferred genetic cluster (Table S2).

The green (dominant in Northern Europe) and red (dominant in Western Europe) clusters were separated by a main genetic boundary extending from the northern Tyrrhenian coast to southern England, which was also an area of admixture. In southern Europe, populations from the Apennines clustered with those from the Balkan peninsula and Carpathian mountains, forming the blue cluster. In central-eastern Europe, the main boundary between the blue and green clusters was located between Slovenia and Croatia and between the Western and Eastern Carpathian mountains. The area between the southern Carpathian mountains and the Baltic sea represented the second largest admixture area detected. The G_{ST} and Jost's *D* pair-wise values among the three genetic clusters spanned from 0.024 to 0.025, and from 0.022 to 0.023, respectively.

Spatial patterns of genetic diversity and differentiation

Genetic diversity and differentiation estimates are detailed in Table S2. The percentage of polymorphic SNPs within a population (%polloc) ranged from 62% to 87 % (mean %polloc

355 = 77%), corresponding to an allelic richness ranging from 1.48 to 1.65 (mean A_r = 1.58).
356 Observed and expected heterozygosities per population ranged, respectively, from 0.262 to
357 0.347 and from 0.248 to 0.327 (mean H_o = 0.305 and mean H_e = 0.300), indicating a small
358 heterozygote excess (mean F_{IS} = -0.025). All genetic diversity indices showed the same
359 patterns of variation across the range (Fig. 3 a, b), where southeastern populations had
360 below-average values, whereas western populations showed above-average values. The
361 indices of genetic diversity (H_e, %polloc) also varied among clusters (Fig. S2). Populations
362 assigned to red and green clusters had a higher H_e than populations assigned to the blue
363 cluster (p = 0.003 and 0.074 respectively), while the red cluster populations showed higher
364 %polloc values than those of the green and blue clusters ($p < 10^{-3}$).

365 Spatial patterns of genetic diversity were consistent when assessed independently
366 with SNPs from the two different arrays (Fig. S3). Although other array-specific
367 representation problems may occur, such a finding rules out a major distortion due to
368 ascertainment bias.

369 Genetic differentiation of each population from the entire gene pool ranged from -
370 0.034 to 0.20 (mean θ_{WT} = 0.05). Patterns of θ_{WT} variation across Europe were opposed to
371 diversity patterns (Fig. 3c), with southeastern populations characterized by higher average
372 values than western populations. Accordingly, θ_{WT} of the red and green cluster was higher
373 than θ_{WT} of the blue cluster (p = 0.005 and 0.091 respectively) (Fig. S2 c).

374 Genetic differentiation between all populations pairs revealed a significant signal of
375 IBD (Fig. 4a). Pairwise $F_{ST}/1-F_{ST}$ values increased with increasing geographic distance (blog =
376 0.027, $p < 0.001$). This IBD signal was also observed within the blue and green clusters (Table
377 S3, Fig. S4), but not within the red cluster.

The EEMS analyses highlighted several barriers to gene flow corresponding to biogeographical barriers (Fig. 4b). The first barrier separated the UK and Scandinavia from the European mainland, and corresponded to the English Channel, the Baltic sea and plains in Western Germany and North-Western France. A second barrier corresponded to the Alps (Northern Italy, Austria) and extended to Slovakia and the Carpathians in the east. Weaker barriers to gene flow also occurred in Southern Italy and the Balkans.

Spatial patterns of effective diversity estimated with EEMS (Fig. 4c) partially contrasted with the maps of H_e and %polloc (Fig. 3). Indeed, areas of higher-than-average diversity were found with EEMS in Western Europe (UK, Pyrenees mountains, Germany) but also in Central and South-Eastern Europe. Diversity was lower than average in Spain, Southern Italy, and in an area from Eastern Scandinavia to Poland.

Climate data analysis

The first three principal components of the temperature-focussed PCA (Temp1, Temp2, Temp3) were retained, and accounted for 90.9% of the total variance of the dataset (Online Appendix A3). Temp1 is an axis of mean temperatures, opposing hot (southern) to cold (northern) climates. Temp2 can be interpreted as an axis of climate continentality, opposing climates with strong versus weak variation of temperatures among years and seasons (e.g. continental *vs.* oceanic climates). Temp3 can be interpreted as an axis of climate xericity, opposing climates with a high diurnal range and the wettest season corresponding to the coldest months (i.e., mediterranean climates) to climates with a low diurnal range and the wettest season corresponding to the warmest months (i.e. temperate mesic climates).

The first three principal components of the precipitation-focussed PCA (Precip1, Precip2, Precip3) were also retained, and accounted for 98.6% of the total variance of the dataset. Precip1 is a precipitation abundance axis, opposing wet (Great-Britain, Northern Italy) to dry climates (Greece, Spain). Precip2 is a precipitation variability axis, opposing climates with strong (Greece, Italy) versus weak (France) variation of precipitation. Precip3 captures the coupling between precipitation and seasonal temperatures, opposing climate where high precipitation occurs during the vegetation period (Poland, Romania) to those where high precipitation occurs in winter (Greece, Italy).

Selection signatures

pcadapt : The first two PCs were retained to represent population structure in *pcadapt* analysis based on the Cattle's rule (Online Appendix A3). One candidate SNP under selection (0.3%) was identified after controlling for FDR (Table S4).

lea: The lowest cross-entropy criterion value was found at K = 3. Five SNPs (1.85%) were identified as potentially under divergent selection after controlling for FDR and after calibrating *p*-values using the calculated genomic inflation factor ($\lambda = 6.0$; Table S4).

lfmm: After controlling for FDR, 46 SNPs (17%) were found to be associated with temperature or precipitation-related climatic variables (60 significant associations in total, Table S5): seven SNPs showed correlations with Temp1, 13 with Temp2 and 11 SNPs with Temp3 ; four SNPs showed correlations with Precip1, nine with Precip2, and 16 with Precip3.

Samβada : Thirty-two genotypes at 22 SNPs (8.1%) were associated with temperature and precipitation-related variables after controlling for FDR. In particular, four loci showed

421 correlations with Precip1 and 11 with Precip3, while 5 SNPs were associated withTemp1,
422 seven with Temp2 and five with Temp3 (Table S6).

423 *Overlapping signatures of selection and ontology of genes bearing outlier:* Overlapping
424 signatures of selection from population divergence and GEA analyses were detected at
425 genes 154_1 and QB_c10512 (Table 1, Fig. 5). The GEA analyses shared signatures of local
426 adaptation in 12 additional genes (13 common SNPs). The population divergence analyses
427 also shared signatures of local adaptation at one additional gene. Finally, 29 and six outlier
428 SNPs were detected by *lfmm* and *Samβada* alone, respectively.

429 While our panel of 270 SNPs included 67.9% of putatively phenology-related, 26.4% of
430 stress-related and 5.7% of control-related genes, respectively, outliers included 61.4% of
431 phenology-related, 34.1% of stress-related, and 4.5% of control genes (Table 1). Hence,
432 there was no difference of category (stress, phenology, control) among initial and outlier
433 genes ($\chi^2 = 0.99$, $p = 0.61$).

434 ***Isolation by distance and environment***

435 The variance partitioning and partial RDA analyses revealed a greater effect of spatial
436 structure than of climatic structure on the spatial distribution of genetic variation (Figure S5,
437 Table S7). Considering the 218 putatively neutral SNPs only, the climatic structure alone
438 explained ~1% of variance in the genetic structure, which is not statistically different from
439 the null expectation of 0% variance explained ($F_{6,52} = 1.17$, $p = 0.16$). On the contrary, the
440 contribution of the spatial structure alone was much larger ($R^2 = 16\%$) and its effect was
441 statistically significant ($F_{5,52} = 4.12$, $p < 0.001$). The joint effect of climatic and spatial
442 structures contributed significantly in explaining the genetic structure ($R^2 = 25\%$). In this

443 case, such a joint effect is not equivalent to a standard interaction term, and relates to the
444 intrinsic covariation of climatic and spatial effects.

445 When considering the set of 52 SNPs putatively under selection, the contribution of
446 the climatic structure to the genetic structure was significant (R^2 = 3%; $F_{6,52}$ = 1.65, p =
447 0.006), but still smaller than that of the spatial structure (14%; $F_{5,52}$ = 4.008, p < 0.001). The
448 joint effects of climatic and spatial structures explained 32% of the genetic structure. Thus,
449 the variance contributed by climatic and spatio-climatic structures combined was higher for
450 the 52 outliers (~35%) than for the supposedly neutral loci (~26%; Fig. S5). Moreover, the
451 temperature and precipitation components of the sole climatic effects explained a similar
452 and statistically significant amount of variance (1.5% and 1.2%, respectively; p < 0.001; Table
453 S7B) in the supposedly adaptive genetic structure. The variance contributed by shared
454 spatial and temperature structures (21%) was higher than that contributed by shared spatial
455 and precipitation structures (6%).

456 ## Discussion

457 Our study supported our first expectation: we observed weak founder effects in beech,
458 which indicates that past population demography is not likely to blur the detection of
459 selection signatures. On the contrary, our second expectation was not met as we identified
460 loci with the signature of divergent selection as often in genes involved in phenology as in
461 genes involved in stress response. And counter to our third expectation, temperature and
462 precipitation related variables were equally represented in the significant genotype-climate
463 associations. Overall, our results suggest a balanced contribution of traits related with
464 phenology and with stress responses to local adaptation in beech.

Impact of past recolonization history on genetic diversity

Our results revealed a clear spatial disjunction between three main gene pools. A first pool (blue cluster) corresponds to the area harbouring beech glacial refugia in Southeastern Europe. Its geographical distribution matches well with results of genetic and palaeoecological studies showing that lineages from these refugia expanded only as far as the Northern Apennines and Central Carpathians (Leonardi & Menozzi, 1995; Magri, 2008; Magri et al., 2006). A second gene pool (red cluster) includes the Southwest European glacial refugia located on the Iberian Peninsula and in southern France, where beech persisted throughout several glacial cycles (de Lafontaine et al., 2014) and was even more abundant than in the southeastern refugia during the middle and upper Pleistocene (Magri et al., 2006). A third gene pool (green cluster) mostly corresponds to recently recolonized areas in Northern Europe (Sjölund, González-Díaz, Moreno-Villena, & Jump, 2017) and likely originates from glacial refugia located in the Eastern Alps-Slovenia and in Slovakia-Moravia (Magri, 2008). Boundaries between the three clusters were associated with strong admixture, which would be consistent with the relatively high number of recolonization routes known for beech as compared to other tree species (de Lafontaine et al. 2014; Magri, 2008).

Consistently with previous studies based on allozymes or microsatellites (Comps, Gomory, Letouzey, Thiebaut, & Petit, 2001; de Lafontaine et al., 2013), spatial patterns of SNP genetic diversity did not reflect signals of founder effects resulting from the post-glacial expansion. In particular, the northern populations (green cluster) showed values of Nei's heterozygosity (H_e) and genetic differentiation (β_{WT}) similar to those of the southwestern populations (red cluster), while the southeastern populations (blue cluster) showed lower H_e

and higher b_{WT}. Hence, diversity across the 270 studied SNPs was higher, and differentiation lower, both in recently recolonized areas, and in areas where beech was more abundant in the past. These patterns likely result from the combination of several processes and life-history traits specific to trees in general and beech in particular: first, the long juvenile phase of forest trees strongly attenuates founder effects during colonization in a diffusive dispersal model (Austerlitz et al., 2000). Moreover, long-distance pollen dispersal is frequent in beech (Gauzere, Klein, & Oddou-Muratorio, 2013; Piotti et al., 2012), which is expected to increase the number of founders and the mixing of genes from distant sources, resulting in a rapid increase of genetic diversity after the initial colonization (Fayard et al., 2009; Lander, Klein, Roig, & Oddou-Muratorio, 2021; Paulose & Hallatschek, 2020). Finally, beech is one of the tree species that recolonised northern Europe the latest, and the factors that limited its ability to migrate probably also contributed to its retaining a high level of diversity along the expansion front (Roques et al., 2012; Saltré et al., 2013)

Range-wide spatial genetic structure (SGS) was statistically significant but weak. The strongest signal was found in the southeastern genetic cluster. This is consistent with the theoretical work of Slatkin (1993) on IBD, who showed that a species having restricted dispersal should exhibit SGS if enough time has elapsed after establishment. Since the south-eastern European populations (blue cluster) have undergone a relatively early and short-distance post-glacial expansion (Magri et al., 2006), they would have had the longest time for the establishment of SGS.

Altogether, our results hence agree on the absence of a marked signature of genetic drift and allele surfing in beech due to recolonization. It appears therefore warranted to assume that our analysis of genetic signatures of local adaptation is little burdened with such

511 sources of uncertainty, in line with previous studies on temperate forest trees that have
512 explicitly tested for such effects (Eckert et al., 2010; Ruiz Daniels et al., 2018; Temunović et
513 al., 2020).

514 **Genomic signatures of local adaptation along climatic gradients**

515 Two genes showed convergent signatures of selection using GEA and differentiation
516 outlier analyses. GEA detected more outliers than differentiation outlier analyses, with 12
517 genes showing convergent signatures of divergent selection using *lfmm* and *Sambada*. The
518 52 outliers identified with at least one of the methods displayed significant IBE patterns
519 (while the putatively neutral markers did not), consistently with the fact that allele
520 frequencies co-vary with climatic variables at the loci under selection.

521 According to GEA, 50 associations were attributable to the temperature variables and
522 42 to the precipitation variables. Among the temperature variables, associations with
523 climate continentality (22) were found more often than associations with mean temperature
524 (12) or climate xericity (16). This finding may indicate that the risk of late frosts could
525 represent a major constraint for the evolution of phenology-related traits in beech (Gauzere
526 et al., 2020; Kreyling et al., 2014). Among the precipitation variables, associations with the
527 coupling between precipitation and seasonal temperatures (25) were found more often than
528 associations with precipitation abundance (8) or variability (9). This result could be due to
529 genetic differentiation between locations where high precipitation occurs during the
530 vegetation period (coupling) versus those where a precipitation deficit occurs during the
531 vegetation period (decoupling). This would highlight the major role of low precipitation in
532 driving patterns of local adaptation, in agreement with the known sensitivity of beech to
533 drought (Aranda et al., 2015; Cuevo-Alarcon et al., 2021), and with the major role of maximal

534 potential evapotranspiration as a driver of genetic differentiation for growth and survival
535 (Gárate-Escamilla et al. 2019). Partial RDA analyses also indicate significant effects of
536 temperature and precipitation on the genetic structure at all the 52 outlier loci together,
537 even though the portion of genetic variance contributed by "pure" temperature or
538 precipitation effects was low in both cases (1%). Considering that phenology-related
539 candidate genes were slightly over-represented in our set of 270 SNPs, and assuming that
540 optimal values of phenological traits are likely to vary primarily with temperature and
541 photoperiod rather than with precipitation (Metcalf & Mitchell-Olds, 2009), the balanced
542 contribution of precipitation and temperature variables to the genetic-climate associations
543 suggests that range-wide local adaptation in beech is driven by traits related to various
544 climate components (see also Garate-Escamilla et al. 2019).

545 Although we cannot completely rule out some false positives, the large number of
546 outliers detected using GEA approaches is methodologically consistent. First, our sampling
547 design with 64 populations covering beech range proves to be appropriate to account for
548 steep ecological gradients, control for population structure, and ultimately optimize
549 statistical power in GEA (Selmoni, Vajana, Guillaume, Rochat, & Joost, 2020). Moreover, the
550 major post-glacial expansion axes of beech align with steep ecological gradients: the South-
551 to-North axis of expansion opposes hot to cold climates, while the axis from Central Europe
552 to Great Britain opposes continental to oceanic climates (Magri et al., 2006). Although the
553 populations sampled in this study may not fully capture these axes of climate variation, such
554 a configuration is expected to minimize the number of false positive with GEA, as compared
555 to the opposite case where ecological gradients are orthogonal to the expansion axis (Frichot
556 et al. 2015).

557

558 **Functional role of the genes under selection**

559 Among the 139 candidate genes investigated, eighteen showed signatures of divergent
560 selection with at least two methods. Eight of them had also shown signatures of divergent
561 selection in previous studies (Table 1 and Table S8). Among these "best candidates", the
562 outlier gene 92 encodes for ACO, an enzyme involved in the production of the plant
563 hormone ethylene, which regulates many plant developmental processes and stress
564 responses. This study found a putative signature of selection at the non-synonymous locus at
565 position 352, (coding for histidine or glutamine; Table 1), where the frequency of the
566 homozygous genotype TT decreased with drought stress (Fig. 6a). This locus was also
567 associated with annual and growth season temperatures in the study of Cuervo-Alarcon et
568 al. (2018). Two other loci within this gene (although non-coding or synonymous) were also
569 detected as outliers by Pluess et al. (2016), where their frequencies correlated with drought
570 indices. These two previous studies were conducted in Switzerland, along drought and
571 precipitation gradients using GEA approaches. Hence, our combined results suggest that the
572 ACO gene could be under divergent selection at various hierarchical scales across Europe.

573 Another interesting example is the outlier gene QB_c10512, which encodes for the
574 NAC domain-containing protein 72, a transcription factor responsive to desiccation. This
575 gene was found to be differentially expressed in beech quiescent buds by Lesur et al. (2015).
576 At synonymous position 206 (in a Leucine-coding codon), the frequency of the AA genotype
577 increased with drought stress (Fig. 6b). Two other variants within this gene (including a non-
578 synonymous one) were also detected as outliers by Cuervo-Alarcon et al. (2018), where their
579 frequencies also correlated with precipitation during the growing season. Moreover, the

gene QB13549, detected as an outlier by *pcadapt* and *lea*, also encodes for the NAC domain-containing protein 72. At position 857, allele frequency showed a strong variation from Eastern to Western Europe (Fig. 6c).

The signatures of divergent selection were not particularly enriched in genes related to phenology. This is a counter-intuitive result, as previous quantitative genetic approaches failed to detect divergent selection at various physiological traits related to drought stress, but did detect significant differentiation in phenological (Gauzere et al., 2020; Hajek et al., 2016), growth and survival traits (Gárate-Escamilla et al., 2019; Gauzere et al., 2020). This is likely because the stress-related genes genotyped in this study are involved in the response to multiple stresses varying across climate gradients. Moreover, our panel of candidate genes probably determines a larger number of stress-related traits than usually phenotyped in quantitative genetic approaches (Gauzere et al., 2020; Hajek et al., 2016). Another limitation of quantitative genetic approaches is that they are usually conducted on a few populations and do hence not adequately cover the range-wide diversity of stress gradients (see also Garate-Escamilla et al. 2019). This comparison illustrates the complementarity of quantitative genetic and molecular approaches to investigate local adaptation (Rudman et al., 2018).

Implications for conservation and management

The rates of expected (natural) species range shifts are likely to be insufficient for trees to track ongoing climate change (Saltré et al., 2013; Savolainen et al., 2007). In this context, there is an increasing interest in evolutionary-oriented management strategies, relying on the high genetic diversity observed within and among tree populations to adapt forest to ongoing climate change (Aitken & Whitlock, 2013; Lefèvre et al., 2014; Oney et al., 2013).

We showed that the spatial distribution of genetic diversity across beech range reflects both biogeographical history and adaptive processes. This has consequences for conservation, where the importance of maintaining adaptive genetic diversity - in addition to preserving as many lineages as possible - can not longer be overlooked (de Lafontaine et al. 2018; Ouborg, Pertoldi, Loeschcke, Bijlsma, & Hedrick, 2010; Shafer et al., 2015). While the conservation of lineages rely on the assessment of genetic boundaries, the conservation of adaptive diversity may require the identification of the relevant loci and the targeted conservation of specific alleles, genotypes or combinations thereof. Our results also have consequences for management, where knowledge of the genetic variants under selection, as combined with the estimation of their current spatial distribution and the prediction of future climate, may help inform decisions about assisted migration, perhaps under the form of an "enrichment" of existing stands with potentially favourable genotypes (Rellstab et al., 2016; Rochat, Selmoni, & Joost, 2021). This would however require a reliable validation of the adaptive meaning of our best candidates by independent proof, as well as the assessment of genotype×genotype interactions (to make sure there is no outbreeding depression) and of genotype×environment interactions (to avoid undesired, unforeseen under-performances of the introduced genotypes and their progeny in the new environments).

At the intersection of management and conservation lies the possibility to favour natural migration and regeneration dynamics, which could result in efficient mixing of genotypes in multiple environments, thus exposing them to natural selection and adaptive processes (Lefèvre et al., 2014). Here, the detailed analysis of barriers to gene flow is of the essence, to understand whether the barriers and "corridors" we detected have been caused

626 by geographical features or isolation by adaptation: while in the former case it may be
627 sensible to manage such barriers and corridors to shape gene flow, in the latter in may be
628 difficult - or even detrimental - to force the modification of gene flow patterns.
629
630

631 ## Acknowledgements

632 We thank Dalibor Ballian, Peter Zhelev and Aristotelis C. Papageorgiou for providing beech
633 samples from Bosnia and Herzegovina, Bulgaria and Greece, as well as Ilaria Spanu,
634 Mariaceleste Labriola, Catia Boggi and Adline Delcamp for technical assistance during lab
635 work. SOM and AH were supported by the French National Research Agency Biodiversa
636 project TipTree (ANR-12-EBID-0003); AP, FB and GGV supported by the Italian Ministry of
637 University and Research (FOE-2019) under the projects “Climate Change” (CNR
638 DTA.AD003.474) and "Green & Circular Economy" (CNR DTA.AD003.139); and AH by the
639 French National Research Agency Cluster of Excellence project COTE (ANR-10-LABX-45,
640 CLIMBEECH). The MassARRAY genotyping was performed at the Genome Transcriptome
641 Facility of Bordeaux (grants from EU FP7 through Trees4Future research infrastructure , from
642 the Conseil Régional d’Aquitaine n°20030304002FA and 20040305003FA, from the European
643 Union FEDER n°2003227 and from Investissements d'Avenir ANR-10-EQPX-16-01).

644 ## References

645 Aitken, S. N., & Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate
646 Change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. doi:
647 10.1146/annurev-ecolsys-110512-135747
648 Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., … Savolainen,
649 O. (2013). Potential for evolutionary responses to climate change – evidence from tree
650 populations. *Global Change Biology*, *19*(6), 1645–1661. doi: 10.1111/gcb.12181
651 Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., … Savolainen,
652 O. (2013). Potential for evolutionary responses to climate change—Evidence from tree

653 populations. *Global Change Biology*, *19*(6), 1645–1661. doi: 10.1111/gcb.12181
654 Aranda, I., Cano, F. J., Gascó, A., Cochard, H., Nardini, A., Mancha, J. A., … Sánchez-Gómez, D. (2015).
655 Variation in photosynthetic performance and hydraulic architecture across European beech
656 (*Fagus sylvatica* L.) populations supports the case for local adaptation to water stress. *Tree*
657 *Physiology*, *35*(1), 34–46. doi: 10.1093/treephys/tpu101
658 Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H., & Godelle, B. (2000). Effects of Colonization
659 Processes on Genetic Diversity: Differences Between Annual Plants and Tree Species.
660 *Genetics*, *154*, 1309–132113.
661 Bresson, C. C., Vitasse, Y., Kremer, A., & Delzon, S. (2011). To what extent is altitudinal variation of
662 functional traits driven by genetic adaptation in European oak and beech? *Tree Physiology*,
663 *31*(11), 1164–1174. doi: 10.1093/treephys/tpr084
664 Capblancq, T., Morin, X., Gueguen, M., Renaud, J., Lobreaux, S., & Bazin, E. (2020). Climate associated
665 genetic variation in *Fagus sylvatica* and potential responses to climate change in the French
666 Alps. *Journal of Evolutionary Biology*, jeb.13610. doi: 10.1111/jeb.13610
667 Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-
668 generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7.
669 doi: 10.1186/s13742-015-0047-8
670 Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., … McCabe, A. M. (2009).
671 The Last Glacial Maximum. *Science*, *325*(5941), 710–714. doi: 10.1126/science.1172873
672 Comps, B., Gomory, D., Letouzey, J., Thiebaut, B., & Petit, R. J. (2001). Diverging Trends Between
673 Heterozygosity and Allelic Richness During Postglacial Colonization in the European Beech.
674 *Genetics*, *157*, 389–397.
675 Csilléry, K., Lalagüe, H., Vendramin, G. G., González-Martínez, S. C., Fady, B., & Oddou-Muratorio, S.
676 (2014). Detecting short spatial scale local adaptation and epistatic selection in climate-
677 related candidate genes in European beech (*Fagus sylvatica*) populations. *Molecular*
678 *Ecology*, *23*(19), 4696–4708. doi: 10.1111/mec.12902
679 Cuervo-Alarcon, L., Arend, M., Müller, M., Sperisen, C., Finkeldey, R., & Krutovsky, K. V. (2018).
680 Genetic variation and signatures of natural selection in populations of European beech
681 (Fagus sylvatica L.) along precipitation gradients. *Tree Genetics & Genomes*, *14*(6), 84. doi:
682 10.1007/s11295-018-1297-2
683 Cuervo-Alarcon, L., Arend, M., Müller, M., Sperisen, C., Finkeldey, R., & Krutovsky, K. V. (2021). A
684 candidate gene association analysis identifies SNPs potentially involved in drought tolerance
685 in European beech (*Fagus sylvatica* L.). *Scientific Reports*, *11*(1), 2386. doi: 10.1038/s41598-
686 021-81594-w
687 de Lafontaine, G., Amasifuen Guerra, C. A., Ducousso, A., & Petit, R. J. (2014). Cryptic no more: Soil
688 macrofossils uncover Pleistocene forest microrefugia within a periglacial desert. *New*
689 *Phytologist*, *204*(3), 715–729. doi: 10.1111/nph.12833
690 de Lafontaine, G., Ducousso, A., Lefèvre, S., Magnanou, E., & Petit, R. J. (2013). Stronger spatial
691 genetic structure in recolonized areas than in refugia in the European beech. *Molecular*
692 *Ecology*, *22*(17), 4397–4412. doi: 10.1111/mec.12403
693 de Lafontaine, G., Napier, J. D., Petit, R. J., & Hu, F. S. (2018). Invoking adaptation to decipher the
694 genetic legacy of past climate change. *Ecology*, *99*(7), 1530–1546. doi: 10.1002/ecy.2382
695 de Villemereuil, P., Frichot, É., Bazin, É., François, O., & Gaggiotti, O. E. (2014). Genome scan methods
696 against more complex models: When and how much should we trust them? *Molecular*
697 *Ecology*, *23*(8), 2006–2019. doi: 10.1111/mec.12705
698 Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for
699 principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*(3–4),
700 483–493. doi: 10.1016/j.ecolmodel.2006.02.015
701 Eckert, A. J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S.
702 C., & Neale, D. B. (2010). Patterns of Population Structure and Environmental Associations to

703 Aridity Across the Range of Loblolly Pine (*Pinus taeda* L., Pinaceae). *Genetics*, *185*(3), 969–
704 982. doi: 10.1534/genetics.110.115543
705 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the
706 software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. doi:
707 10.1111/j.1365-294X.2005.02553.x
708 Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic Consequences of Range Expansions. *Annual Review*
709 *of Ecology, Evolution, and Systematics*, *40*(1), 481–501. doi:
710 10.1146/annurev.ecolsys.39.110707.173414
711 Fady, B., Aravanopoulos, F. A., Alizoti, P., Mátyás, C., von Wühlisch, G., Westergren, M., … Zlatanov,
712 T. (2016). Evolution-based approach needed for the conservation and silviculture of
713 peripheral forest tree populations. *Forest Ecology and Management*, *375*, 66–75. doi:
714 10.1016/j.foreco.2016.05.015
715 Fayard, J., Klein, E. K., & Lefèvre, F. (2009). Long distance dispersal and the fate of a gene from the
716 colonization front. *Journal of Evolutionary Biology*, *22*(11), 2171–2182. doi: 10.1111/j.1420-
717 9101.2009.01832.x
718 Frichot, E, Schoville, S. D., de Villemereuil, P., Gaggiotti, O. E., & François, O. (2015). Detecting
719 adaptive evolution based on association with ecological gradients: Orientation matters!
720 *Heredity*, *115*(1), 22–28. doi: 10.1038/hdy.2015.7
721 Frichot, Eric, & François, O. (2015). LEA: An R package for landscape and ecological association
722 studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. doi: 10.1111/2041-210X.12382
723 Frichot, Eric, Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for Associations between
724 Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and*
725 *Evolution*, *30*(7), 1687–1699. doi: 10.1093/molbev/mst063
726 Gabriel, S., Ziaugra, L., & Tabbaa, D. (2009). SNP Genotyping Using the Sequenom MassARRAY iPLEX
727 Platform. *Current Protocols in Human Genetics*, *60*(1), 2.12.1-2.12.18. doi:
728 10.1002/0471142905.hg0212s60
729 Gárate-Escamilla, H., Hampe, A., Vizcaíno-Palomar, N., Robson, T. M., & Benito Garzón, M. (2019).
730 Range-wide variation in local adaptation and phenotypic plasticity of fitness-related traits in
731 *Fagus sylvatica* and their implications under climate change. *Global Ecology and*
732 *Biogeography*, *28*(9), 1336–1350. doi: 10.1111/geb.12936
733 Gauzere, J., Klein, E. K., Brendel, O., Davi, H., & Oddou-Muratorio, S. (2020). Microgeographic
734 adaptation and the effect of pollen flow on the adaptive potential of a temperate tree
735 species. *New Phytologist*, nph.16537. doi: 10.1111/nph.16537
736 Gauzere, J., Klein, E. K., & Oddou-Muratorio, S. (2013). Ecological determinants of mating system
737 within and between three *Fagus sylvatica* populations along an elevational gradient.
738 *Molecular Ecology*, *22*(19), 5001–5015. doi: 10.1111/mec.12435
739 Gömöry, D., & Paule, L. (2011). Trade-off between height growth and spring flushing in common
740 beech (Fagus sylvatica L.). *Annals of Forest Science*, *68*(5), 975–984. doi: 10.1007/s13595-
741 011-0103-1
742 Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular*
743 *Ecology Notes*, *5*(1), 184–186. doi: 10.1111/j.1471-8286.2004.00828.x
744 Hajek, P., Kurjak, D., von Wühlisch, G., Delzon, S., & Schuldt, B. (2016). Intraspecific variation in wood
745 anatomical, hydraulic, and foliar traits in ten European beech provenances differing in
746 growth yield. *Frontiers in Plant Science*, *7*. doi: 10.3389/fpls.2016.00791
747 He, C., Holme, J., & Anthony, J. (2014). *SNP genotyping: The KASP assay*. New York, NY: Humana
748 Press.
749 Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution
750 interpolated climate surfaces for global land areas. *International Journal of Climatology*,
751 *25*(15), 1965–1978. doi: 10.1002/joc.1276
752 Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C.

753 (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future
754 Directions. *The American Naturalist*, *188*(4), 379–397. doi: 10.1086/688018
755 Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers.
756 *Bioinformatics*, *24*(11), 1403–1405. doi: 10.1093/bioinformatics/btn129
757 Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new
758 method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94. doi:
759 10.1186/1471-2156-11-94
760 Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–
761 1241. doi: 10.1111/j.1461-0248.2004.00684.x
762 Knutzen, F., Dulamsuren, C., Meier, I. C., & Leuschner, C. (2017). Recent climate warming-related
763 growth decline impairs European beech in the center of its distribution range. *Ecosystems*,
764 *20*(8), 1494–1511. doi: 10.1007/s10021-017-0128-x
765 Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A
766 program for identifying clustering modes and packaging population structure inferences
767 across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. doi: 10.1111/1755-0998.12387
768 Krajmerová, D., Hrivnák, M., Ditmarová, Ľ., Jamnická, G., Kmeť, J., Kurjak, D., & Gömöry, D. (2017).
769 Nucleotide polymorphisms associated with climate, phenology and physiological traits in
770 European beech (Fagus sylvatica L.). *New Forests*, *48*(3), 463–477. doi: 10.1007/s11056-017-
771 9573-9
772 Kramer, K., Ducousso, A., Gömöry, D., Hansen, J. K., Ionita, L., Liesebach, M., … von Wühlisch, G.
773 (2017). Chilling and forcing requirements for foliage bud burst of European beech (Fagus
774 sylvatica L.) differ between provenances and are phenotypically plastic. *Agricultural and*
775 *Forest Meteorology*, *234–235*, 172–181. doi: 10.1016/j.agrformet.2016.12.002
776 Kreyling, J., Buhk, C., Backhaus, S., Hallinger, M., Huber, G., Huber, L., … Beierkuhnlein, C. (2014).
777 Local adaptations to frost in marginal and central populations of the dominant forest tree
778 *Fagus sylvatica* L. as affected by temperature and extreme drought in common garden
779 experiments. *Ecology and Evolution*, *4*(5), 594–605. doi: 10.1002/ece3.971
780 Lalagüe, H., Csilléry, K., Oddou-Muratorio, S., Safrana, J., de Quattro, C., Fady, B., … Vendramin, G. G.
781 (2014). Nucleotide diversity and linkage disequilibrium at 58 stress response and phenology
782 candidate genes in a European beech (*Fagus sylvatica* L.) population from southeastern
783 France. *Tree Genetics & Genomes*, *10*(1), 15–26. doi: 10.1007/s11295-013-0658-0
784 Lander, T. A., Klein, E. K., Roig, A., & Oddou-Muratorio, S. (2021). Weak founder effects but
785 significant spatial genetic imprint of recent contraction and expansion of European beech
786 populations. *Heredity*, *126*(3), 491–504. doi: 10.1038/s41437-020-00387-5
787 Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of*
788 *Statistical Software*, *25*(1). doi: 10.18637/jss.v025.i01
789 Lefèvre, F., Boivin, T., Bontemps, A., Courbet, F., Davi, H., Durand-Gillmann, M., … Pichot, C. (2014).
790 Considering evolutionary processes in adaptive forestry. *Annals of Forest Science*, *71*(7), 723–
791 739. doi: 10.1007/s13595-013-0272-1
792 Legendre, P., Fortin, M., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis?
793 *Methods in Ecology and Evolution*, *6*(11), 1239–1247. doi: 10.1111/2041-210X.12425
794 Legendre, P., & Legendre, L. F. (2012). *Numerical ecology*. Elsevier.
795 Leonardi, S., & Menozzi, P. (1995). Genetic variability of *Fagus sylvatica* L. in Italy: The role of
796 postglacial recolonization. *Heredity*, *75*(1), 35–44. doi: 10.1038/hdy.1995.101
797 Lesur, I., Bechade, A., Lalanne, C., Klopp, C., Noirot, C., Leplé, J.-C., … Le Provost, G. (2015). A unigene
798 set for European beech (*Fagus sylvatica* L.) and its use to decipher the molecular
799 mechanisms involved in dormancy regulation. *Molecular Ecology Resources*, *15*(5), 1192–
800 1204. doi: 10.1111/1755-0998.12373
801 Lind, B. M., Menon, M., Bolte, C. E., Faske, T. M., & Eckert, A. J. (2018). The genomics of local
802 adaptation in trees: Are we out of the woods yet? *Tree Genetics & Genomes*, *14*(2), 29. doi:

10.1007/s11295-017-1224-y

Luu, K., Bazin, E., & Blum, M. G. B. (2017). *pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77. doi: 10.1111/1755-0998.12592

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, *35*(3), 450–463. doi: 10.1111/j.1365-2699.2007.01803.x

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., … de Beaulieu, J.-L. (2006). A new scenario for the quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *The New Phytologist*, *171*(1), 199–221. doi: 10.1111/j.1469-8137.2006.01740.x

Martins, H., Caye, K., Luu, K., Blum, M. G. B., & François, O. (2016). Identifying outlier loci in admixed and in continuous populations using ancestral population differentiation statistics. *Molecular Ecology*, *25*(20), 5029–5042. doi: 10.1111/mec.13822

Metcalf, C. J. E., & Mitchell-Olds, T. (2009). Life history in a model system: Opening the black box with Arabidopsis thaliana. *Ecology Letters*, *12*(7), 593–600. doi: https://doi.org/10.1111/j.1461-0248.2009.01320.x

Müller, M., Seifert, S., & Finkeldey, R. (2015). A candidate gene-based association study reveals SNPs significantly associated with bud burst in European beech (*Fagus sylvatica* L.). *Tree Genetics & Genomes*, *11*(6), 116. doi: 10.1007/s11295-015-0943-1

Oksanen, J., Blanchet, F. G., & others. (2019). *Package 'vegan.'* Retrieved from https://github.com/vegandevs/vegan

Oney, B., Reineking, B., O'Neill, G., & Kreyling, J. (2013). Intraspecific variation buffers projected climate change impacts on *Pinus contorta*. *Ecology and Evolution*, *3*(2), 437–449. doi: 10.1002/ece3.426

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. (Kuke), & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics*, *26*(4), 177–187. doi: 10.1016/j.tig.2010.01.001

Paulose, J., & Hallatschek, O. (2020). The impact of long-range dispersal on gene surfing. *Proceedings of the National Academy of Sciences*, *117*(14), 7584–7593. doi: 10.1073/pnas.1919485117

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*(19), 2537–2539. doi: 10.1093/bioinformatics/bts460

Petit-Cailleux, C., Davi, H., Lefèvre, F., Garrigue, J., Magdalou, J.-A., Hurson, C., … Oddou-Muratorio, S. (2020). Comparing statistical and mechanistic models to identify the drivers of mortality within a rear-edge beech population. *BioRxiv 645747 Ver. 7 Peer-Reviewed and Recommended by PCI Ecology*, 645747. doi: 10.1101/645747

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. doi: 10.1038/ng.3464

Piotti, A., Leonardi, S., Buiteveld, J., Geburek, T., Gerber, S., Kramer, K., … Vendramin, G. G. (2012). Comparison of pollen gene flow among four European beech (Fagus sylvatica L.) populations characterized by different management regimes. *Heredity*, *108*(3), 322–331. doi: 10.1038/hdy.2011.77

Pluess, A. R., Frank, A., Heiri, C., Lalagüe, H., Vendramin, G. G., & Oddou-Muratorio, S. (2016). Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. *New Phytologist*, *210*(2), 589–601. doi: 10.1111/nph.13809

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). *Inference of Population Structure Using Multilocus Genotype Data*. *155*, 945–95915.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rehfeldt, G. E., Tchebakova, N. M., Parfenova, Y. I., Wykoff, W. R., Kuzmina, N. A., & Milyutin, L. I. (2002). Intraspecific responses to climate in *Pinus sylvestris* . *Global Change Biology*, *8*(9), 912–929. doi: 10.1046/j.1365-2486.2002.00516.x

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. doi: 10.1111/mec.13322

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A. R., Graf, R., … Gugerli, F. (2016). Signatures of local adaptation in candidate genes of oaks (*Quercus* spp.) with respect to present and future climatic conditions. *Molecular Ecology*, *25*(23), 5907–5924. doi: 10.1111/mec.13889

Rochat, E., Selmoni, O., & Joost, S. (2021). Spatial areas of genotype probability: Predicting the spatial distribution of adaptive genetic variants under future climatic conditions. *Diversity and Distributions*, (n/). doi: https://doi.org/10.1111/ddi.13256

Roques, L., Garnier, J., Hamel, F., & Klein, E. K. (2012). Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences*, *109*(23), 8828–8833. doi: 10.1073/pnas.1201695109

Rudman, S. M., Barbour, M. A., Csilléry, K., Gienapp, P., Guillaume, F., Hairston Jr, N. G., … Levine, J. M. (2018). What genomic data can reveal about eco-evolutionary dynamics. *Nature Ecology & Evolution*, *2*(1), 9–15. doi: 10.1038/s41559-017-0385-2

Ruiz Daniels, R., Taylor, R. S., Serra-Varela, M. J., Vendramin, G. G., González-Martínez, S. C., & Grivet, D. (2018). Inferring selection in instances of long-range colonization: The Aleppo pine (*Pinus halepensis*) in the Mediterranean Basin. *Molecular Ecology*, *27*(16), 3331–3345. doi: 10.1111/mec.14786

Saltré, F., Saint-Amant, R., Gritti, E. S., Brewer, S., Gaucherel, C., Davis, B. A. S., & Chuine, I. (2013). Climate or migration: What limited European beech post-glacial colonization?: European beech post-glacial migration. *Global Ecology and Biogeography*, *22*(11), 1217–1227. doi: 10.1111/geb.12085

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 595–619. doi: 10.1146/annurev.ecolsys.38.091206.095646

Selmoni, O., Vajana, E., Guillaume, A., Rochat, E., & Joost, S. (2020). Sampling strategy optimization to increase statistical power in landscape genomics: A simulation-based approach. *Molecular Ecology Resources*, *20*(1), 154–169. doi: https://doi.org/10.1111/1755-0998.13095

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., … Zieliński, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*(2), 78–87. doi: 10.1016/j.tree.2014.11.009

Sjölund, M. J., González-Díaz, P., Moreno-Villena, J. J., & Jump, A. S. (2017). Understanding the legacy of widespread population translocations on the post-glacial genetic structure of the European beech, *Fagus sylvatica* L. *Journal of Biogeography*, *44*(11), 2475–2487. doi: 10.1111/jbi.13053

Slatkin, M. (1993). *Isolation by Distance in Equilibrium and Non-Equilibrium Populations*. *47*(1), 264–279.

Stucki, S., Orozco-terWengel, P., Forester, B. R., Duruz, S., Colli, L., Masembe, C., … Joost, S. (2017). High performance computation of landscape genomic models including local indicators of spatial association. *Molecular Ecology Resources*, *17*(5), 1072–1089. doi: 10.1111/1755-0998.12629

Temunović, M., Garnier-Géré, P., Morić, M., Franjić, J., Ivanković, M., Bogdan, S., & Hampe, A. (2020). Candidate gene SNP variation in floodplain populations of pedunculate oak (*Quercus robur* L.) near the species' southern range margin: Weak differentiation yet distinct associations with water availability. *Molecular Ecology*, mec.15492. doi: 10.1111/mec.15492

Vitasse, Y., Delzon, S., Bresson, C. C., Michalet, R., & Kremer, A. (2009). Altitudinal differentiation in

growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*, *39*(7), 1259–1269. doi: 10.1139/X09-054

Warnes, G., Gorjanc, with contributions from G., Leisch, F., & Man, M. (2021). genetics: Population Genetics (Version 1.3.8.1.3). Retrieved from https://CRAN.R-project.org/package=genetics

Wortemann, R., Herbette, S., Barigah, T. S., Fumanal, B., Alia, R., Ducousso, A., … Cochard, H. (2011). Genotypic variability and phenotypic plasticity of cavitation resistance in Fagus sylvatica L. across Europe. *Tree Physiology*, *31*(11), 1175–1182. doi: 10.1093/treephys/tpr101

Data Accessibility

Raw SNPs genotypes, detailed information on SNPs and sampling locations, as well as R scripts for LD analyses, for differentiation outlier analyses (with *pcadapt*, and *lea*) and for GEA analyses (with *lfmm* and *Samβada*) will be made available at Portail Data INRAE: https://data.inrae.fr/, upon acceptance of the manuscript.

Author Contributions

Conceptualization: S.O.M., G.G.V., D.P., E.V., A.P. Population sampling: D.P., F.B., G.G.V, A.H and F.P. Selection of candidate genes, SNPs and associated bioinformatics tools: D.P., I.L., G.L.P, S.O.M, G.G.V. SNP genotyping: D.P. and E.G. Statistical analyses: D.P., S.O.M, E.V., A.P. Writing- Original Draft : D.P., S.O.M, E.V., A.P. All the authors reviewed, edited and approved the final manuscript. Funding acquisition: S.O.M. and G.G.V.

Figures legends

Figure 1: Distribution of the 64 studied populations (red dots) (a) in the geographical space, overlaid on beech distribution range (in grey) ; (b) in the bioclimatic niche defined by annual precipitations and temperature, with the grey colour intensity indicating increasing density of beech stands.

Figure 2: Spatial interpolates of the admixture coefficients estimated with STRUCTURE for K=3. Each color corresponds to one cluster and the colour intensity indicates the probability to belong to the cluster at a given position in space, based on spatial kriging of the individual q-matrix. Only areas belonging to beech distribution range are considered. Crosses indicate admixed populations, not assigned to a single cluster.

Figure 3: Estimates of diversity (He), percentage of polymorphic loci (%polloc) and genetic differentiation relative to the entire pool (β_{WT}) in the 64 studied populations, overlaid on beech distribution range (in grey).

Figure 4: Patterns of isolation by distance and barriers to gene flow among the 64 studied populations. (a) Spatial genetic structure as depicted by the variation of genetic differentiation against geographic distance (on a log-scale). The grey envelope represents expected $F_{ST}/1\text{-}F_{ST}$ values under complete spatial randomness and bars represent standard error at 95% level within each distance class. (b) Contour maps representing the posterior mean of effective migration (m) surface; populations in the blue areas are connected by higher migration rates than expected under isolation by distance (IBD) while the ones in the orange areas have lower migration rates than expected and are interpreted as migration barriers. In white areas, the effective migration surface is close to the one expected under IBD. (c) Contour maps representing the posterior mean of effective diversity (q) surface; populations in the orange (respectively blue) areas have lower-than-expected (respectively higher-than-expected) genetic diversity than the average. On maps (b) and (c), black dots represent the studied populations, aggregated per grid cell (with size proportional to the number of genotyped individuals)

Figure 5: Venn diagram of the private and common outliers identified by the different methods to detect signatures of divergent selection.

Figure 6: Variation in allelic/genotypic frequencies at three outlier SNPs (panel a: 92_352; b: QB_c10512; c: QB13549_857) across the studied geographic range. For panels (a) and (b), the graph on the left represents the predicted variation in genotype occurrence probability (GOP) across the environmental gradient as estimated by *Samβada*. Maps show the observed genotypic/allelic frequency superimposed on the environmental gradient (panels a and b) and spatial genetic structure (c).

Table 1. Outlier SNPs showing signature of divergent selection with at least one of the four methods (*pcadapt, lea, lfmm* and *Samβada*). We applied analysis-specific FDR cut-offs granting no expected false positives. For *lfmm* and *Samβada*, we report the climatic variable for which the genetic-environment association was found (T1-3: Temp1-3; P1-3: Precip 1-3). Converging selection signatures from at least two methods are underlined in grey. For each SNP, we code the study where it was described first (1: Lalagüe, et al. 2014 ; 2: Lesur, et al. 2015) and give the gene sequence (in Genbank for 1; in Lesur, et al. 2015 for 2). We finally provide for each candidate gene its category (stress-related, phenology-related or control genes), its annotation based on the homology with *Arabidopsis thaliana* sequence (with the TAIR ID and probability of matching E)

SNP	Method				SNP Resource		Cat	Annotation	TAIR_ID	E
	pcadapt	LEA	LFMM	Samβada	Study	Sequence name				
7_186	-	-	T3	-	1	JX406438.1	Stress	Xyloglucan endotransglucosylase/hydrolase 18 (ATXTH18)	AT4G30280.1	1.00E
19_206	-	-	P1	-	1	JX406440.1	Stress	pseudogene of Histone superfamily protein	AT1G75610.1	2.00E
21_243	-	-	P3,T1,T2	P3,T2	1	JX406442.1	Stress	S-adenosylmethionine synthetase 3 (SAMS3)	AT3G17390.1	4.00E
27_485	-	-	T3	-	1	JX406445.1	Stress	Trehalose-phosphate/synthase 7 (TPS7)	AT1G06410.1	5.00E
39_225	-	-	T1	-	1	JX406448.1	Stress	Potassium transporter (AtKT2p)	AT2G40540.1	2.00E
50_232	-	-	T3	T3	1	JX406449.1	Stress	C-repeat/dre binding factor 1 (ATCBF1)	AT4G25490.1	5.00E
52_1_246	-	-	T2	T2	1	JX406451.1	Stress	S-adenosyl-l-homocystein hydrolase 1 (SAHH1)	AT4G13940.1	1.00E
66_698	-	-	P1	P1	1	JX406455.1	Stress	S-adenosylmethionine decarboxylase (SAMDC)	AT3G02470.1	1.00 E
68_277	-	-	P3	-	1	JX406456.1	Stress	Glyceraldehyde-3-phosphate dehydrogenase c subunit (GAPC1)	AT3G04120.1	5.00E
92_352	-	-	P3,T2,T3	T3	1	JX406462.1	Stress	1-aminocyclopropane-1-carboxylate oxidase (ACO4)	AT1G05010.1	1.00E
129_685	-	-	P3	P3	1	JX406471.1	Pheno	Glyceraldehyde-3-phosphate dehydrogenase asubunit 2 (GAPA-2)	AT1G12900.1	4.00E
133_306	-	-	P1	P1,T2	1	JX406475.1	Pheno	NA	NA	NA
142_143	-	-	T2	-	1	JX406476.1	Pheno	Membrane protein CONTINUOUS VASCULAR RING (COV1)	AT2G20120.1	6.00E

148_1_1411	-	-	P3	-	1	JX406478.1	Pheno	S phase kinase-associated protein 1 (SKP1)	AT1G75950.1	7.00E-1
150_2_924	-	-	P3,T1,T2,T3	P3,T2	1	JX406479.1	Pheno	Auxin response factor 6 (ARF6)	AT1G30330.2	2.00E-1
154_1_715	-	□	P3	-	1	JX406480.1	Stress	Pectin methylesterase 39 (PME39)	AT4G02300.1	2.00E-1
154_1_845	-	□	P3,T2	T1	1					
154_1_251	-	□	-	T1	1					
154_1_390	-	-	-	T1	1					
154_2_371	-	-	P3,T2,T3	T2	1	JX406481.1	Stress	Pectin methylesterase 3 (PME3)	AT3G14310.1	6.00E-1
155_2_911	-	-	P2	-	1	JX406482.1	Stress	Polygalacturonase 2 (PG2)	AT1G70370.1	2.00E-1
62_1_148	-	-	P3	-	1	JX406489.1	Stress	Heat shock protein 70 (HSP70)	AT3G12580.1	2.00E-1
91_2_1441	-	-	P3,T1	P3,T1	1	JX406491.1	Stress	Catalase 2 (CAT2)	AT4G35090.1	6.00E-1
91_2_57	-	-	P3	-	1	JX406491.1				
134_2_834	-	-	-	P3	1	JX406493.1	Pheno	Metallothionein 2a (MT2A)	AT3G09390.1	2.00E-1
QB_c10460-202	-	-	-	T1	2	c10460	Pheno	Na	na	na
QB_c10512-206	-	□	P3,T2,T3	T3	2	c10512	Pheno	NAC domain-containing protein 72, Responsive to desication 26 (ANAC72)	AT4G27410.2	7.00E-1
QB_c10517-414	-	-	P2,P3	-	2	c10517	Pheno	Hypothetical protein	AT4G02040.1	
QB_c10517-841	-	-	T1	-	2	c10517				
SB_c5654-1048	-	-	T3	-	2	c5654	Pheno	Glucose-methanol-choline (GMC) oxidoreductase family protein	AT5G51950.1	6.00E-1
QB_c6167-1062	-	-	T2	-	2	c6167	Pheno	Leucine-rich repeat (LRR) family protein	AT1G33590.1	3.00E-1
SB_c6451-300	-	-	-	T3	2	c6451	Pheno	Translocase of the inner membrane 9 (TIM9)	AT3G46560.1	6.00E-1
QB_c7172-467	-	-	-	P1	2	c7172	Pheno	Aba-hypersensitive germination 3 (AHG3)	AT3G11410.1	1.00E-1
SB_c7640-125	-	-	P2	-	2	c7640	Pheno	Leucine-rich repeat (LRR) family protein	AT3G19320.1	
SB_c968-1354	-	-	P2	-	2	c968	Pheno	Picloram Resistant30 (PIC30)	AT2G39210.1	
SB_c968-192	-	-	P2	-	2	c968				
SB_c968-719	-	-	P2	-	2	c968				
SB_c968-935	-	-	P2	-	2	c968				
QB_c13130-798	-	-	T3	-	2	c13130	Pheno	Ndr1/hin1-like	AT2G35980.1	3.00E-1
QB_c13152-130	-	-	P1,T3	-	2	c13152	Pheno	Na	na	na
ctrlfagus_c13215-830	-	-	T2	T2	2	c13215	Ctrl	Glyceraldehyde-3-phosphate dehydrogenase c subunit 1 (GAPC1)	AT3G04120.1	9.00E-1

SB_c13339-608	-	-	P2	P3	2	c13339	Pheno	Fasciclin-like arabinogalactan family protein (FLAs)	AT5G16920.1	3.00E-1
QB_c13406-208	-	-	P2	-	2	c13406	Pheno	Seed storage 2S albumin superfamily protein	AT2G37870.1	
SB_c13429-427	-	-	-	P3	2	c13429	Pheno	ARABINOGALACTAN PROTEIN 30 (AGP30), Pistil-specific extensin-like protein precursor	AT2G33790.1	9.00E-1
QB_c13549-857	□	□	-	-	2	c13549	Pheno	NAC domain-containing protein 72, Responsive to desication 26 (ANAC72)	AT4G27410.2	
SB_c13643-626	-	-	P3	-	2	c13643	Pheno	Pathogenesis-related thaumatin superfamily protein	AT1G19320.1	3.00E-1
QB_c15642-205	-	-	T1	-	2	c15642	Pheno	Embryonic cell protein 63 (ECP63),	AT2G36640.1	2.00E-1
SB_c15868-233	-	-	P3,T1	-	2	c15868	Pheno	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	AT3G52130.1	1.00E-1
QB_c15913-724	-	-	T2	-	2	c15913	Pheno	Low-temperature-induced 65 kda protein (LTI65), responsive to dessication	AT5G52300.2	1.00E-1
QB_c15913-902	-	-	T2	T2	2	c15913	Pheno	Low-temperature-induced 65 kda protein (LTI65), responsive to dessication	AT5G52300.2	1.00E-1
ctrlfagus_c15935-232	-	-	T3	-	2	c15935		Tubulin alpha-2 chain (TUA2)	AT1G50010.1	9.00E-1
QB_c170171-048	-	-	T2	-	2	c17017	Pheno	Low-temperature-induced 65 kda protein (LTI65), responsive to dessication	AT5G52300.2	2.00E-1

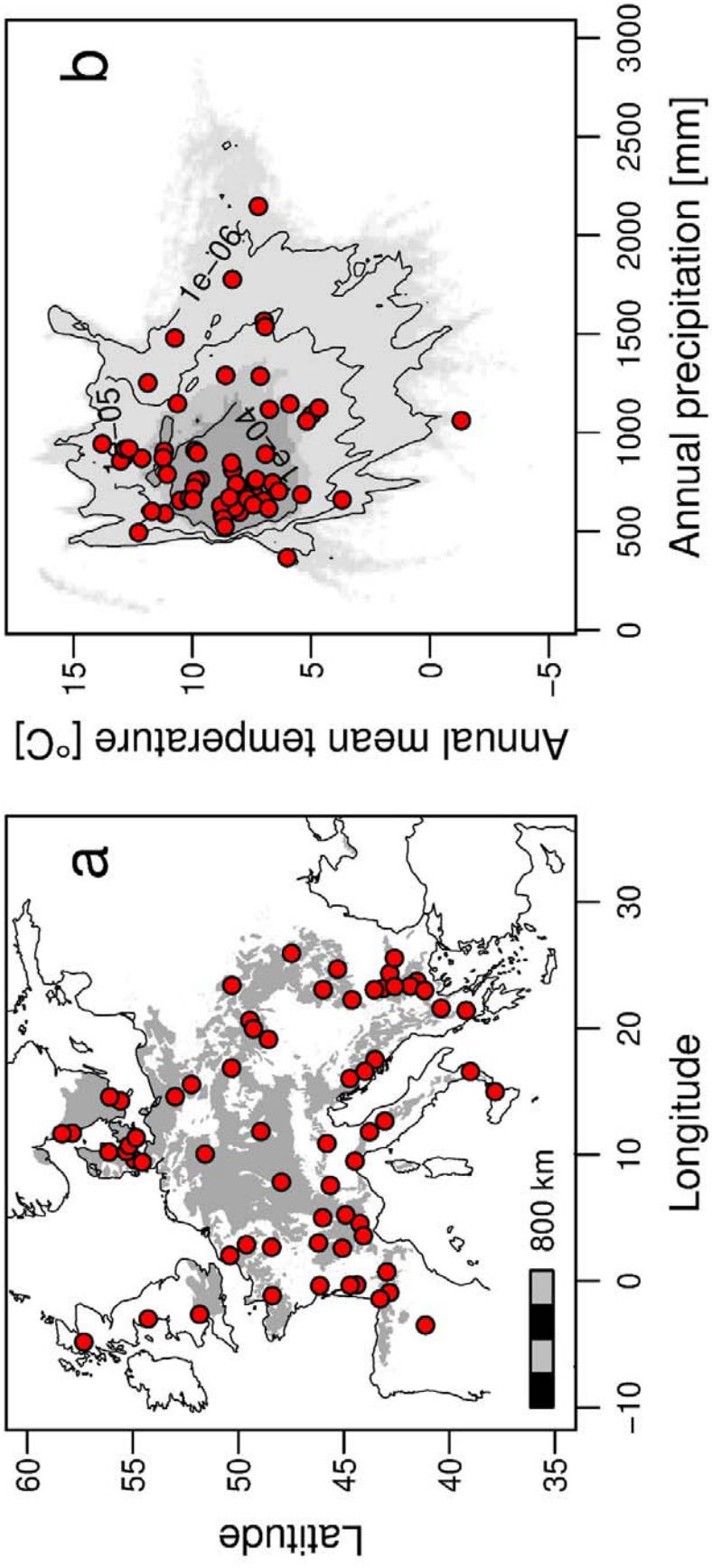
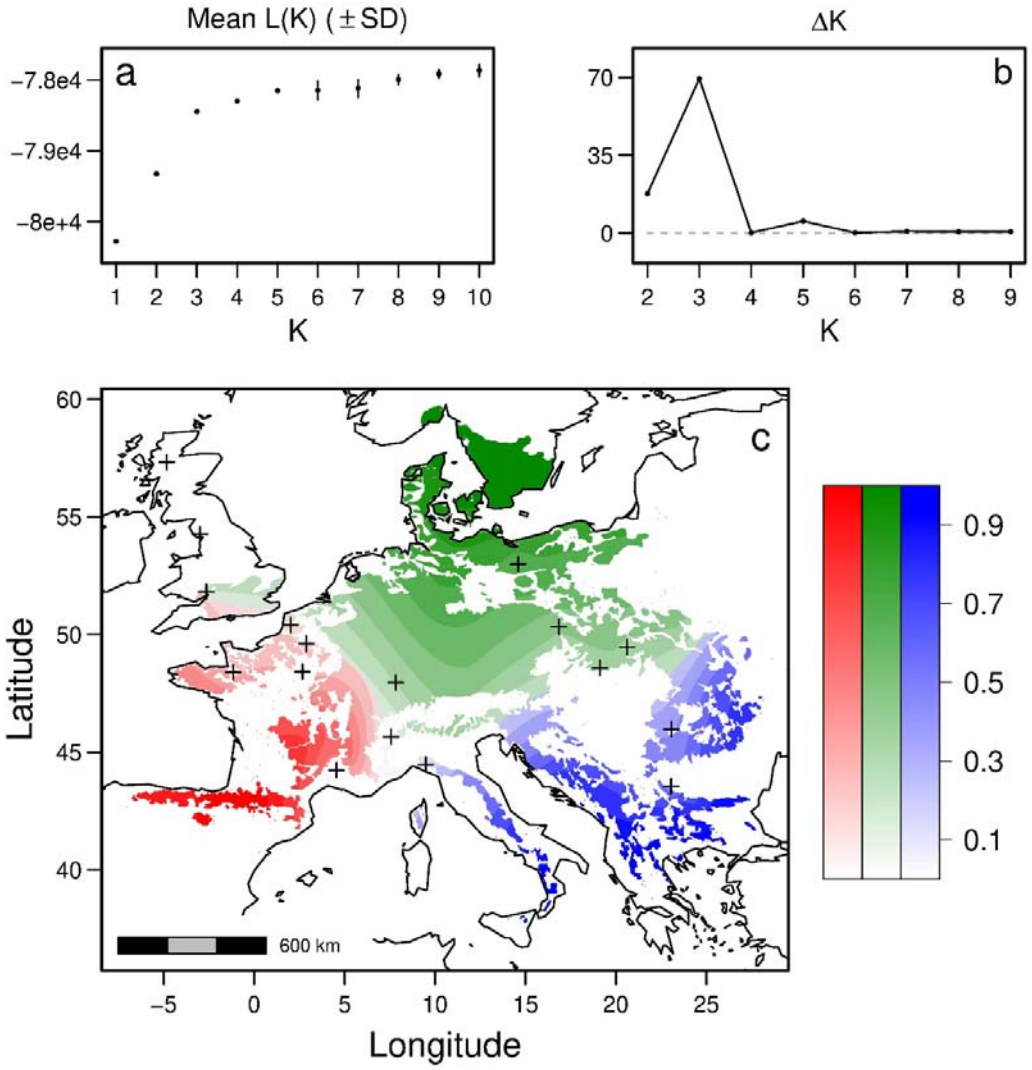


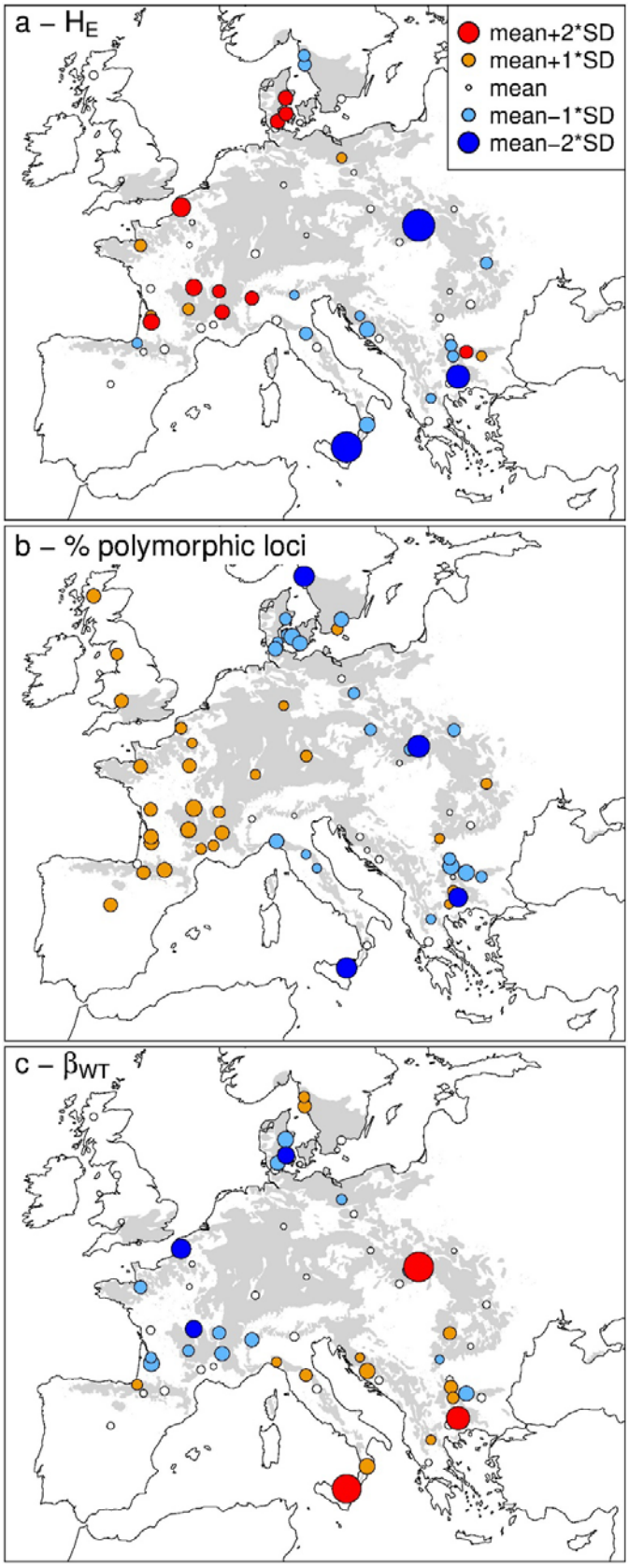
Figure 1



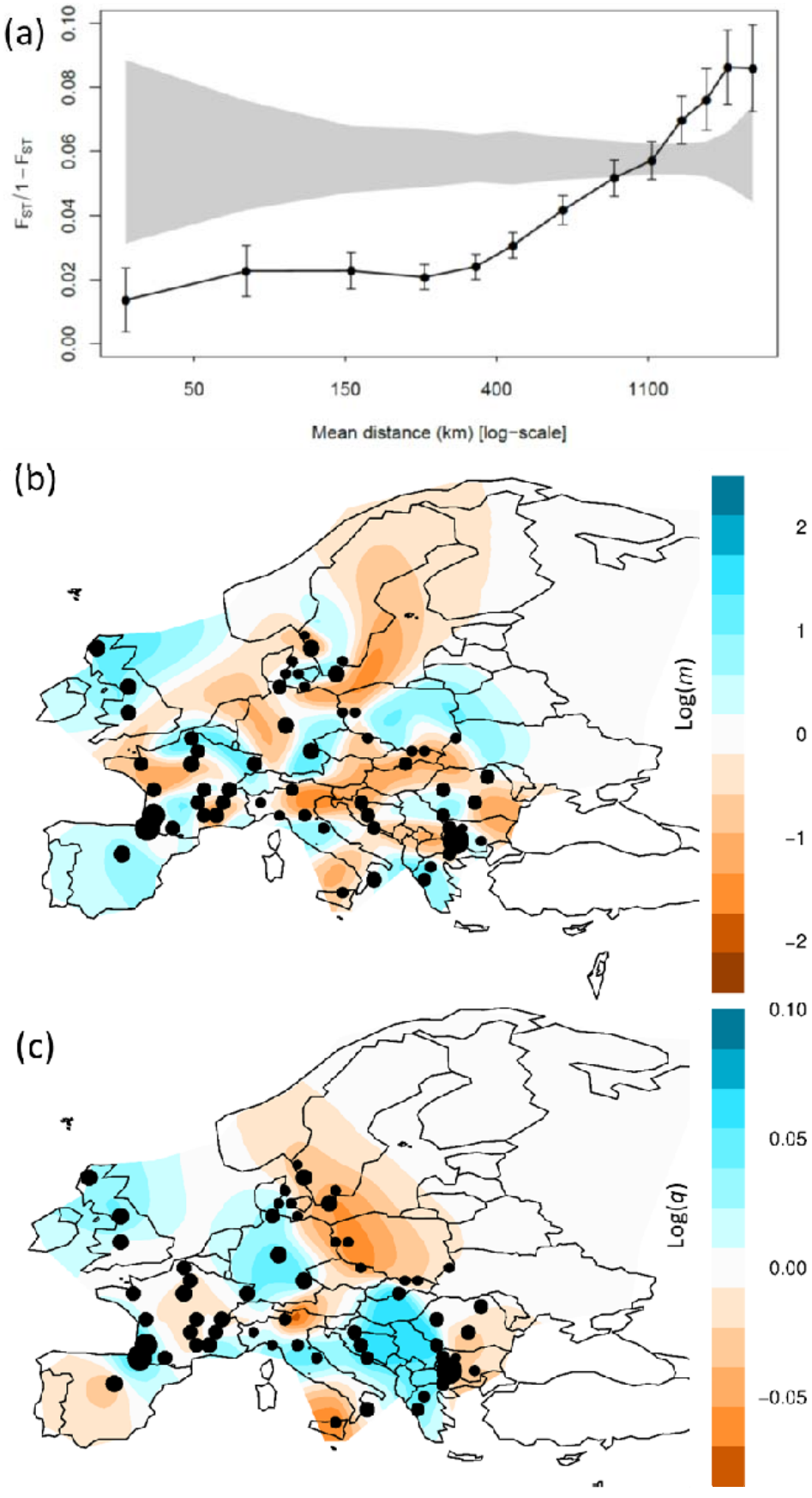
966
967

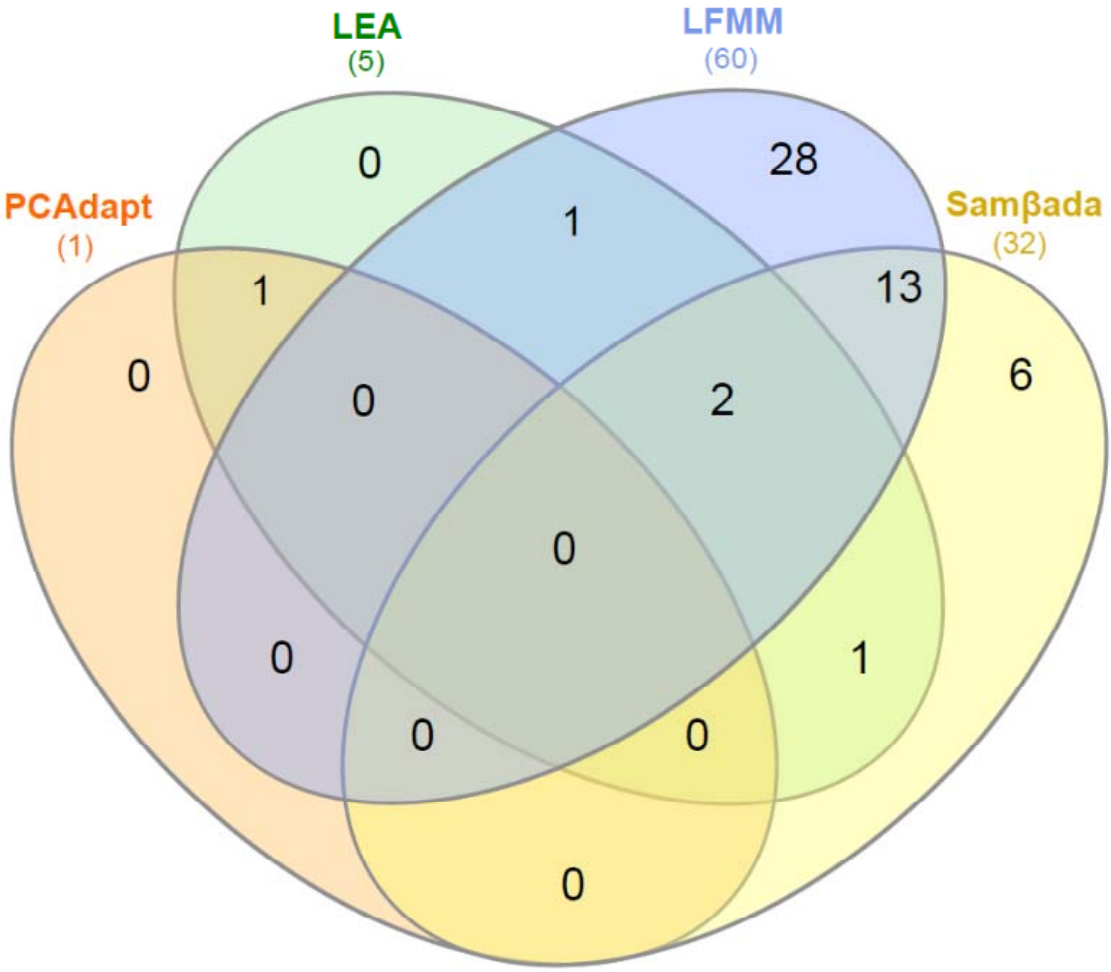
Figure 2

968



969 **Figure 3**

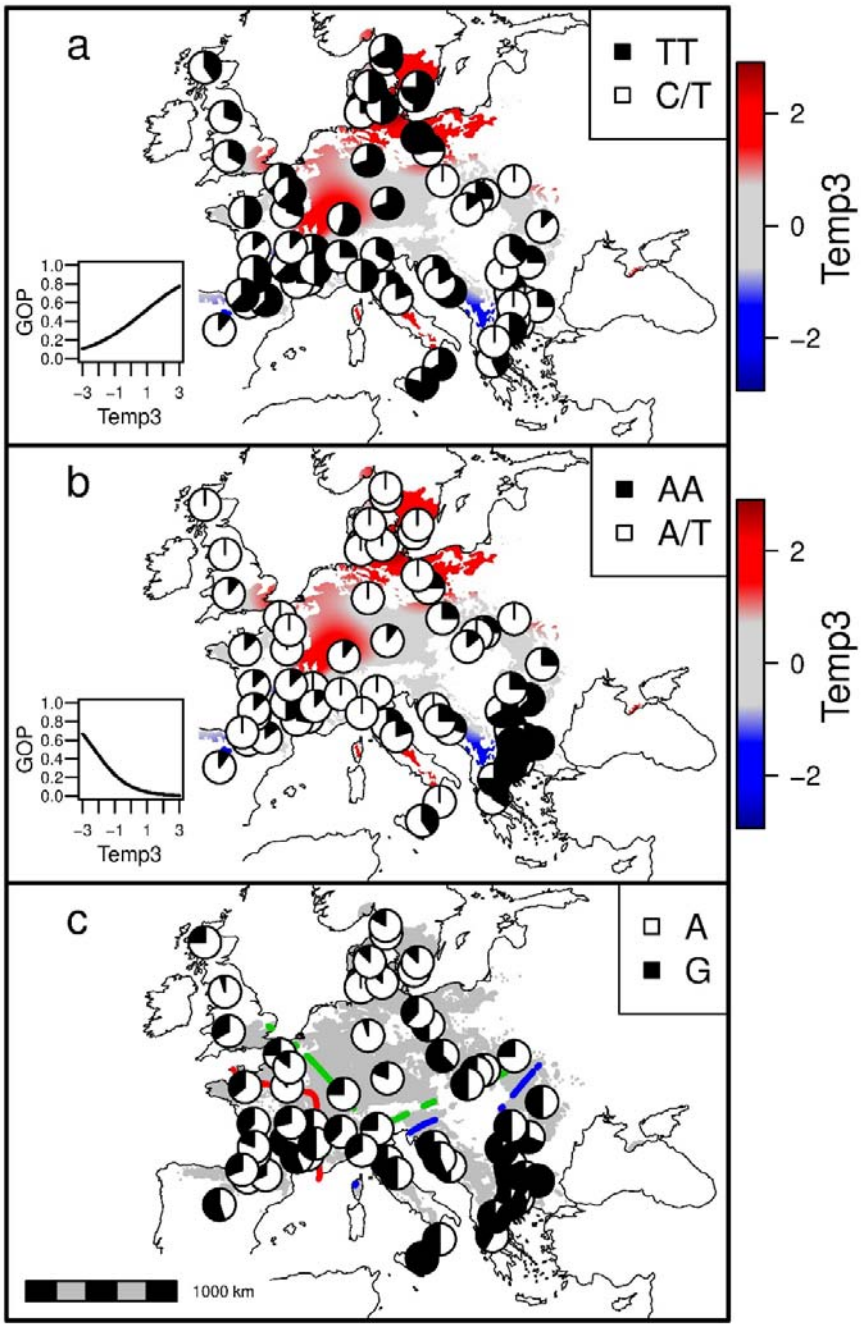




971
972

Figure 5

973



974 **Figure 6**

975