New Phytologist

Research

Adaptive introgression as a driver of local adaptation to climate in European white oaks

Thibault Leroy[1,2], Jean-Marc Louvet[1], Céline Lalanne[1], Grégoire Le Provost[1], Karine Labadie[3], Jean-Marc Aury[3], Sylvain Delzon[1], Christophe Plomion[1] and Antoine Kremer[1]

[1]BIOGECO, INRA, Université de Bordeaux, 69 Route d'Arcachon, 33612, Cestas, France; [2]ISEM, CNRS, IRD, EPHE, University of Montpellier, Place Eugène Bataillon, 34095, Montpellier, France; [3]Genoscope, Institut de biologie François-Jacob, Commissariat à l'Energie Atomique (CEA), Université Paris-Saclay, Evry, France

Author for correspondence:
Antoine Kremer
Tel: +33 (0)5 57 12 38 32
Email: antoine.kremer@inra.fr

Received: *18 March 2019*
Accepted: *31 July 2019*

New Phytologist (2020) **226:** 1171–1182
doi: 10.1111/nph.16095

Key words: genetic clines, genome scans, genotype–environment associations, interspecific gene flow, local adaptation.

Summary

• Latitudinal and elevational gradients provide valuable experimental settings for studies of the potential impact of global warming on forest tree species. The availability of long-term phenological surveys in common garden experiments for traits associated with climate, such as bud flushing for sessile oaks (*Quercus petraea*), provide an ideal opportunity to investigate this impact.

• We sequenced 18 sessile oak populations and used available sequencing data for three other closely related European white oak species (*Quercus pyrenaica*, *Quercus pubescens*, and *Quercus robur*) to explore the evolutionary processes responsible for shaping the genetic variation across latitudinal and elevational gradients in extant sessile oaks. We used phenotypic surveys in common garden experiments and climatic data for the population of origin to perform genome-wide scans for population differentiation and genotype–environment and genotype–phenotype associations.

• The inferred historical relationships between *Q. petraea* populations suggest that interspecific gene flow occurred between *Q. robur* and *Q. petraea* populations from cooler or wetter areas. A genome-wide scan of differentiation between *Q. petraea* populations identified single nucleotide polymorphisms (SNPs) displaying strong interspecific relative divergence between these two species. These SNPs followed genetic clines along climatic or phenotypic gradients, providing further support for the likely contribution of introgression to the adaptive divergence of *Q. petraea* populations.

• Overall, the results indicate that outliers and associated SNPs are *Q. robur* ancestry-informative. We discuss the results of this study in the framework of the postglacial colonization scenario, in which introgression and diversifying selection have been proposed as essential drivers of *Q. petraea* microevolution.

Introduction

Evolutionary biologists are becoming increasingly fascinated by the tracking of adaptive genetic changes, as our understanding of paleoecology and genomics (Shapiro & Hofreiter, 2014) and climate reconstructions at various spatial and temporal scales (Mauri *et al.*, 2015) improve. Ultimately, assembling data relating to historical and genetic changes will increase our understanding of how, when, and how rapidly evolutionary shifts have enhanced adaptation. The recent report that introgression with Neanderthals or Denisovans increased the adaptation of modern Eurasian humans provides an emblematic example of a major evolutionary shift supported by historical and genomic evidence (Dannemann & Racimo, 2018). Adaptive shifts have also been predicted in nonmodel plants and animals that have repeatedly witnessed large-scale environmental changes over larger temporal scales due to Quaternary climatic oscillations (Dynesius & Jansson, 2000). Forest trees are relevant target species for explorations of evolutionary changes, as their past history and distribution can be reconstructed easily, owing to the availability of large quantities of fossil remains (Brewer *et al.*, 2017; Wagner *et al.*, 2018). Furthermore, understanding how trees adapt to changing environments has become a major topic of interest in theoretical and applied ecological genetics and genomics. On the one hand, there are concerns that trees may not cope with the velocity of ongoing climatic change; on the other, land managers and foresters are seeking methods of adaptive management (Lindner *et al.*, 2010). Adaptation can be triggered by new alleles originating from new mutations, neutral standing genetic variation, or adaptive introgression. Adaptation can be investigated empirically by assessing adaptive divergence in common garden experiments, or by tracking genetic changes over successive generations. The common garden approach has been widely

used for long-lived species, such as trees, and has revealed the existence of high levels of adaptive divergence between extant populations distributed along large geographic gradients (Alberto *et al.*, 2013a). Moreover, adaptive divergence is maintained between tree populations, despite extensive gene flow potentially causing pollen swamping effects that might constrain adaptation (Savolainen *et al.*, 2007). In many European tree species, pollen swamping results from intra- or interspecific gene flow, as many congeneric species live together in the same stands and interspecific hybridization is widespread. In contrast to the pollen swamping effect, it has also been suggested that gene flow may actually enhance adaptation by increasing the ability of populations to respond to natural selection (Kremer & Le Corre, 2012; Yeaman, 2015). It has also recently been suggested that interspecific gene flow may enhance adaptation by facilitating adaptive introgression (Suarez-Gonzalez *et al.*, 2018). Here, we explore the genomic footprints of adaptation variation in sessile oak (*Quercus petraea* (Matt.) Liebl.), a widespread species in Europe, and address the potential contribution of interspecific gene flow to local adaptation.

The geographic distribution of sessile oak extends from Spain to southern Scandinavian, and this species mostly grows in forests also containing the closely related pedunculate oak (*Quercus robur*), another white oak species with a range that extends farther north and east, to the Ural Mountains (Leroy *et al.*, 2020). At more southern latitudes, sessile oak also shares habitats with pubescent (*Quercus pubescens*) and Pyrenean (*Quercus pyrenaica*) oaks. Hybridization between these four European white oak species has been reported in sympatric stands (Curtu *et al.*, 2007; Lepais & Gerber, 2011). Furthermore, hybridization is thought to have played a major role in the expansion of *Q. petraea* populations during postglacial recolonization, as suggested by the widespread sharing of chloroplast genomes between *Q. petraea* and *Q. robur* in sympatric stands (Petit *et al.*, 2002). Interestingly, demographic inferences based on approximate Bayesian computation simulations support scenarios in which these four European white oak species came into secondary contact, at the onset of the last glacial period (Leroy *et al.*, 2017, 2020). A number of questions relating to this scenario remain unanswered. How much did introgression actually contribute to today's distribution of *Q. petraea*? Did introgression contribute to the adaptation of *Q. petraea* or was it a purely neutral process? Which adaptive alleles were introgressed, and from which other white oak species? We tackle these questions here by exploring imprints of historical population splits and admixture events through studies of whole-genome sequences from *Q. petraea* populations. We also performed genome scans to search for genomic footprints of divergence between *Q. petraea* populations, with either phenotypic data from common garden experiments or climatic data for the populations of origin. This experimental setup permitted us to test our earlier suggested hypothesis that introgression facilitated the postglacial expansion of *Q. petraea* following the tracks of *Q. robur*, a more pioneer species (Petit *et al.*, 2004).

Materials and Methods

Sampling and sequencing

We sampled eight *Q. petraea* populations (with up to 20 individuals per pool; Table 1) from lowlands to middle elevations in the Pyrenees (up to 1600 m) in southwest France. These eight populations are distributed along two neighboring elevational transects (the Ossau and Luz valleys) and are stands of natural origin (elevation of 100 m to 1600 m). We also collected data for 10 *Q. petraea* populations (with up to 25 individuals per pool; Table 1) growing in a large common garden experiment in Sillégny, eastern France, corresponding to a total of 116 sessile oak provenances in Europe (Saenz-Romero *et al.*, 2017). The populations selected were sampled along a latitudinal gradient extending from southern France to northern Germany (latitude 44°N to 53°N).

We also included a reference population of the four main white oak species in Europe (*Q. petraea*, *Q. robur*, *Q. pubescens*, *Q. pyrenaica*) corresponding to four oak stands located in southwest France analyzed in a companion paper (Leroy *et al.*, 2020). Samples of reference species populations came from the same geographic region. Such a sampling strategy does not bias species comparison, as shown by an earlier methodological study showing that species differentiation is only moderately impacted by the geographic origin of populations (Bodénès *et al.*, 1997). The *Q. petraea* reference population of Leroy *et al.* (2020) corresponded to a pool of 13 individuals from the same low-elevation population of the Luz valley (L1: Laveyron in Table 1). We used the two populations as pseudoreplicates, to check the accuracy of our pool-sequencing strategy.

DNA was extracted from individual trees with a cetyltrimethyl ammonium bromide DNA extraction protocol (Doyle & Doyle, 1987) (latitudinal gradient) or with the Invisorb Spin Plant Mini Kit (Startec Molecular; elevational gradient) according to the manufacturer's instructions (Startec Molecular GmbH, Berlin, Germany). DNA yields were evaluated with a NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) (elevational gradient) or with an Infinite F200 (Tecan Group Ltd, Männedorf, Switzerland), and DNA samples were mixed in equimolar amounts to obtain a single pool for each population. After trimming, the number of 2×100 bp paired-end reads retained for analyses ranged from 668 million to 998 million reads per pool, corresponding to a rough estimate of sequencing coverage of 180–270×, assuming an oak genome size of 740 Mb. The raw data have been deposited in the Sequence Read Archive: PRJEB32209.

Climatic and phenological data

Monthly mean climate data were obtained from WorldClim (Hijmans *et al.*, 2005) for the period 1950–2000. We downscaled these values to a spatial resolution of 100 m using locally weighted regressions and a finer resolution digital elevation model (100 m) before point overlay and accounted for topography as previously described (Zimmermann *et al.*, 2009; Dullinger *et al.*, 2012). We

Table 1 Geographic and climatic data for the *Quercus petraea* populations studied.

Code	Location	Elevation (m)	Latitude	Longitude	Temperature	Precipitation (mm yr^{-1})	Leaf unfolding	Sample size
Elevational gradient (French Pyrenees)								
L1	Laveyron, Luz Valley, France	131	43.75	−0.22	12.33	901	−1.333	20
L8	Chèze, Luz Valley, France	803	42.92	−0.03	9.27	914	0.817	20
L12	Gèdre, Luz Valley, France	1235	42.78	0.02	7.12	1016	1.011	20
L16	Péguères, Luz Valley, France	1630	42.87	−0.12	6.58	982	1.724	18
O1	Josbaig, Ossau Valley, France	259	43.22	−0.73	12.24	979	−1.309	20
O8	Le Hourcq, Ossau Valley, France	841	42.90	−0.43	9.16	933	−0.324	20
O12	Gabas, Ossau Valley, France	1194	42.88	−0.42	7.35	1031	0.036	20
O16	Artouste, Ossau Valley, France	1614	42.88	−0.40	5.16	1164	0.427	10
Latitudinal gradient								
9	Saint Sauvant, France	155	46.38	0.12	11.78	786	−0.166	25
97	Grésigne, France	310	44.04	1.75	12.05	791	−1.139	25
124	Killarney, Ireland	50	52.01	−9.50	9.96	1362	4.084	25
204	Bézanges, France	275	48.76	6.49	9.50	751	0.371	25
217	Bercé, France	165	47.81	0.39	10.65	698	0.434	25
218	Longchamp, France	235	47.26	5.31	10.59	801	−0.920	22
219	Tronçais, France	245	46.68	2.83	10.63	742	1.350	25
233	Vachères, France	650	43.98	5.63	10.22	797	−1.532	25
253	Göhrde, Germany	85	53.10	10.86	8.30	635	0.953	25
256	Lappwald, Germany	180	52.26	10.99	8.50	597	0.650	25

Date of leaf unfolding expressed as standardized values for common gardens (see the Materials and Methods section). Negative values indicate early flushing, and positive values indicate late flushing.

then calculated yearly average temperature and annual precipitations sums (Supporting Information Fig. S1; Table 1).

The date of leaf unfolding was recorded separately in two different common garden experiments. Saplings of populations sampled along the elevational gradient were transplanted to a common garden experiment in Toulenne in southwest France in spring 2007, and phenological observations were conducted over seven successive years (2009–2015) – see Firmat *et al.* (2017) for further details. Similarly, saplings of populations sampled along the latitudinal gradient were installed in a common garden located in the northeast of France in Sillegny in 1989 and 1993, and phenological observations were conducted in 2015 – see Firmat *et al.* (2017) and Torres-Ruiz *et al.* (2019) for further details. Population means were calculated in each common garden, and standardized values for each common garden were used to study genotype–phenotype associations across the two common gardens (Table 1). Earlier comparisons across different common garden experiments showed, indeed, that the date of leaf unfolding exhibited strong genetic stability; for example, high genetic correlations across sites (Derory *et al.*, 2010).

Mapping and single nucleotide polymorphism calling

We used the pipeline described by Leroy *et al.* (2020); see also https://github.com/ThibaultLeroyFr/GenomeScansByABC/tree/master/SNP_calling_filtering) to identify reliable single nucleotide polymorphisms (SNPs). In brief, we used BOWTIE2 v.2.1.0 to map sequencing data onto the v.2.3 oak haplome assembly (Plomion *et al.*, 2018), removed duplicates with PICARD v.1.106 (http://broadinstitute.github.io/picard/), and then used *snp-frequency-diff.pl* from the POPOOLATION2 suite (Koffler *et al.*, 2011) to select biallelic SNPs with at least 10 alternative alleles. Positions with a mean coverage of < 50× for any of the 18 populations and sites in the top 2% in terms of coverage per population were ignored. SNPs with a minimum allele frequency < 0.02 were filtered out to exclude most Illumina sequencing errors. A total of 37 062 111 SNPs was identified (Table S1).

Population splits and gene flow inferences

We used genome-wide allele frequency data derived from allele counts to build a population tree with TREEMIX v.1.12 (Pickrell & Pritchard, 2012) and to test for the presence of gene flow between populations. Given that *Q. petraea* populations are known to be frequently connected by gene flow with closely related species (Lepais & Gerber, 2011; Leroy *et al.*, 2017), we also evaluated the influence of interspecific gene flow on tree topology. We performed an analysis with a set of 1757 476 intragenic SNPs commonly detected in the 18 previously described sessile oak populations, and in the reference populations of four white oak species (including *Q. petraea*). Data from an outgroup (*Q. suber*, cork oak) described by Leroy *et al.* (2020) were included in the analysis. Variable numbers of migration nodes *m*, ranging from 0 to 13, were evaluated. For each fixed value of *m*, we performed 1000 replicated TREEMIX analyses with a python script from Michael G. Harvey (https://github.com/mgharvey/misc_python/blob/master/bin/TreeMix/treemix_tree_with_bootstraps.py). We then used SumTrees.py from the DENDROPY suite to generate a consensus tree with all the bootstrap values (Sukumaran & Holder, 2010). We used the total variation explained for each replicate in the R companion script of TREEMIX (plotting_funcs.R) as a judgment criterion for evaluation of the best number of migration nodes. We also evaluated the robustness of

the migration nodes, by counting the number of replicated analyses supporting the migration node.

We used the f_3 statistic calculated with THREEPOP v.0.1 from the TREEMIX suite (−k 1000) to determine whether a focal population *X* was the product of admixture between two populations *Y* and *Z*. Significant negative values of f_3 (*X*;*Y*,*Z*) were considered to indicate admixture – for details, see Reich *et al.* (2009) and Schaefer *et al.* (2016). The f_3 statistics were calculated for all possible three-population combinations, but we report only f_3 (*X*;*Y*,*Q.robur*) values, given the known particular contribution of *Q. robur* to the evolution of the populations investigated (see the Results section). No significant negative values were obtained for either f_3 (*X*;*Y*,*Q.pyrenaica*) or f_3 (*X*;*Y*,*Q.pubescens*).

Genome scans Genome-wide scans were performed with the BAYPASS v.2.1 software package (Gautier, 2015). BAYPASS takes confounding demographic effects into account by estimating the covariance matrix of allele frequency between populations. The core model reports locus XtX, which is analogous to F_{ST} but explicitly corrected for this covariance matrix, thus accounting for neutral correlations of allelic frequencies (Gunther & Coop, 2013). For each SNP, BAYPASS can also compare models integrating population-specific covariables (here, temperature, precipitation, and date of leaf unfolding) by including a regression coefficient for the cline along the covariate gradient in the base model. All the covariates used here were scaled as described by Gautier (2015). Direct comparisons of the likelihood of models (Bayes factors, BFs) including and not including particular covariates were used to evaluate the genotype–environment association (GEA), defined as the association between changes in allele frequency and each population-specific climate variable (temperature, precipitation), and the genotype-phenotype association (GPA), defined as the association between changes in allele frequency and mean date of leaf unfolding. The identification of outliers (XtX or BFs) was based on a calibration procedure using pseudo-observed datasets (PODs), as discussed by Gautier (2015). For this step, we used the BayPass_utils.R script as a source for the simulation of 100 000 SNPs. We then ran BAYPASS again, to generate quantile values for XtX and BFs based on these PODs. Conservative and very conservative quantile values were used ('minor outlier' 0.999 for XtX and 0.9999 for BF; 'main outliers' 0.99999 for both XtX and BF) as thresholds to identify outliers. We then generated clusters of neighboring SNPs, by bulking SNPs separated by < 10 kb. We investigated only clusters of associated SNPs containing at least two outliers, including a 'main' outlier, to exclude the random associations that would be expected to occur in such a large dataset. We also investigated clusters with no 'main' outlier, provided that the cluster concerned contained at least five 'minor' outliers.

Manual gene annotations We first performed manual gene annotations based on protein blast searches ($P < 1 \times 10^{-5}$) against the *Arabidopsis* proteome. Gene functions were identified by manual inspection of all the best hits per candidate gene per region. In this study, functional annotation was achieved by performing extensive manual literature searches for each gene, using the strategy described by Leroy *et al.* (2020) rather than automatic approaches based on Gene Ontology-oriented methods. This approach was preferred to ensure that we obtained accurate and detailed information about gene function, supported by checked and traceable references.

Results

Population history and splits

We used TREEMIX to infer the evolutionary history of European white oaks, focusing, in particular, on the most likely sequence of population splits. This software infers the relationships between populations as a bifurcating tree. Interestingly, a strict drift model can explain most of the variance in relatedness between populations (mean over 1000 replicates: 0.926; median: 0.950; Fig. 1a). Looking at the *Q. petraea* populations (Figs 1b, S2), the consensus tree over the 1000 replicates supports basal splits and longer branches for populations at higher elevations than for lowland populations. The only exception was population #124 (Killarney, Ireland), which was at the base of all *Q. petraea* populations. In comparisons with all other *Q. petraea* populations and the *Q. robur* reference population, population #124 also presented significant negative values of the f_3 statistic (Fig. S3), suggesting that this population is probably an admixture between *Q. petraea* and *Q. robur*. Several significant negative f_3 values were also obtained for the L16 population (Péguères, France), for comparisons with several lowland populations (#9: Saint Sauvant, France; #97: Grésigne, France; and #233: Vachères, France; Fig. S3).

We then sequentially added migration events to the tree (Fig. S2). As expected, the variance in relatedness between populations was better explained as the number of migration events increased. At $m = 1$, the proportion of the variance explained increased (mean: 0.955; Fig. 1a) in two-thirds of replicated simulations due to a migration node between *Q. pubescens* and the L8 population (Chèze, France; 661/1000). Otherwise, TREEMIX inferred (219/1000) a migration event between *Q. robur* and population #124. The inferred tree remained inconsistent with the expected species tree (Leroy *et al.*, 2017, 2020), but this first node helped to explain most of variance generated by the relative positions of *Q. pubescens* and *Q. robur*. At $m = 4$ (Fig. 1c), the proportion of the variance explained again increased substantially (mean: 0.976) and the bootstrap replicate maximizing the likelihood was, for the first time, consistent with the expected species tree. Again, TREEMIX inferred a migration event between *Q. robur* and the Irish population (#124). It also provided support for candidate introgression events from *Q. robur* for populations at the highest elevation (i.e. from *Q. robur* to O16 (Artouste, France), and from *Q. robur* to the ancestral population of the modern L12 (Gèdre, France), L16 (Péguères, France), and #124 (Killarney, Ireland) populations; Fig. 1c).

Genome scans

For the identification of SNPs potentially subject to selection, and therefore displaying higher levels of differentiation than would be expected under a hypothesis of neutrality, we calculated the XtX statistic, an F_{ST}-like statistic explicitly accounting for population history (Gunther & Coop, 2013). We identified 761 554 SNPs deviating from neutral expectations (2.05% of all SNPs investigated, presenting XtX values above the 0.999 quantile threshold and referred to hereafter as 'minor outliers'), including 107 764 for which the evidence was strong (0.29%, with XtX values above the 0.99999 quantile threshold referred to hereafter as 'main outliers'; Table S1). These SNPs are distributed over all the chromosomes (outer circle, Fig. 2). Interestingly, XtX outliers were found to be strongly enriched in SNPs that were highly differentiated between species, particularly between *Q. robur* and *Q. petraea* (Fig. 3). More broadly, we observed a strong correlation between the intraspecific XtX value estimated for the 18 populations and the interspecific F_{ST} between *Q. robur* and *Q. petraea* ($P < 2.2 \times 10^{-16}$, $R^2 = 0.247$; Fig. S4).

Among the outlying SNPs detected with the XtX statistic, we identified GEA- and GPA-associated SNPs covarying with mean annual temperature, precipitation, or of the date of leaf unfolding recorded in common gardens (Figs 2, S5–S16). In total, we identified 5682 SNPs associated with at least one covariate. More precisely, we identified 1331 SNPs as outliers (including 216 main outliers) associated with temperature, 2932 as outliers (277 main outliers) associated with precipitation, and 1572 as outliers (125 main outliers) associated with leaf unfolding (Table S1). For the 153 SNPs involved in two different associations with the three covariates, the largest proportion of SNPs (143/153, 93%) was significantly associated with both temperature and leaf unfolding. The remaining set of common SNPs was found to be associated with both temperature and precipitation (6/153, 4%) or with both precipitation and leaf unfolding (4/153, 3%).

The significantly associated SNPs were highly enriched in SNPs strongly differentiated between *Q. petraea* and *Q. robur*, especially leaf unfolding- and temperature-associated SNPs (Fig. 4). We applied a binning procedure to group covariable-associated SNPs in close vicinity within the genome (< 10 kb apart). The SNPs were located in 780, 1617 and 1033 independent genomic regions for temperature, precipitation, and leaf unfolding, respectively (Table S1). No genomic regions matching the two GEA and GPA associations were detected, as expected given the near independence of the two climate covariables (Fig. S1C). In the vast majority of cases (87.1%, 76.0%, and 87.0% for temperature, precipitation and leaf unfolding, respectively), the genomic region consisted of a single associated SNP. Manual gene annotations were performed only for genes located in the 201 genomic regions containing at least two associated SNPs (see the Materials and Methods section for details). These regions accounted for a total of 2016 associated SNPs (mean: 10.2 associated SNPs/region; range: 2–137). These 201 regions included 13 that partially overlapped for leaf unfolding and temperature (Fig. 2, on chromosomes 1, 8, and 10 (see also Figs S5, S12, S14), and on scaffold Sc0000849), three that partially overlapped for precipitation and temperature (Fig. 2, on chromosomes 6, 9 and 10; see also Figs S10, S13, S14), and two that

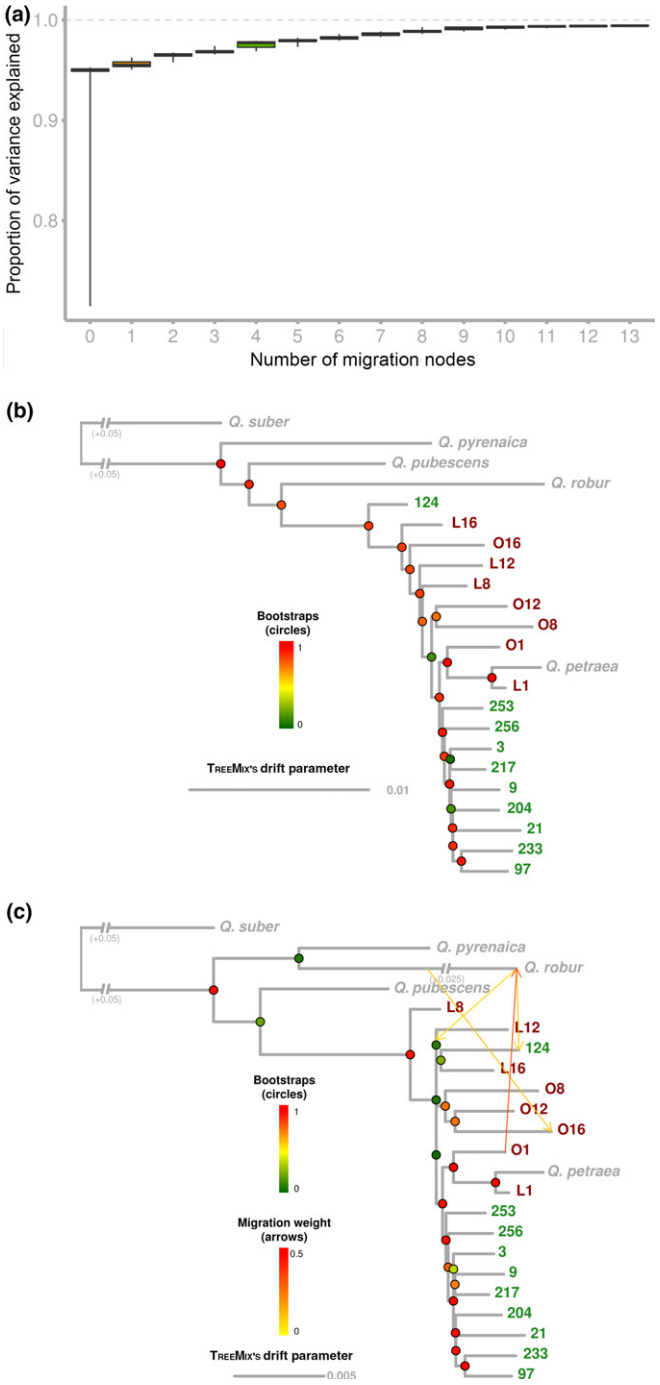

Fig. 1 Inference of splits and migration nodes from 1757 476 genic SNPs in TreeMix. (a) Boxplot of the proportion of the variance explained over 1000 replicates for 0–13 migration nodes. (b) Consensus TreeMix tree under a strict drift model ($m = 0$) for all sessile oak populations and reference populations for three other European white oak species (for details, see Leroy *et al*., 2020). Bootstrap values are shown in circles. (c) Consensus TreeMix tree for $m = 4$. Migration events correspond to the events inferred for the best case (inference with the highest likelihood among the 1000 replicates). Bootstrap percentages and migration weights are indicated by the corresponding color scales.

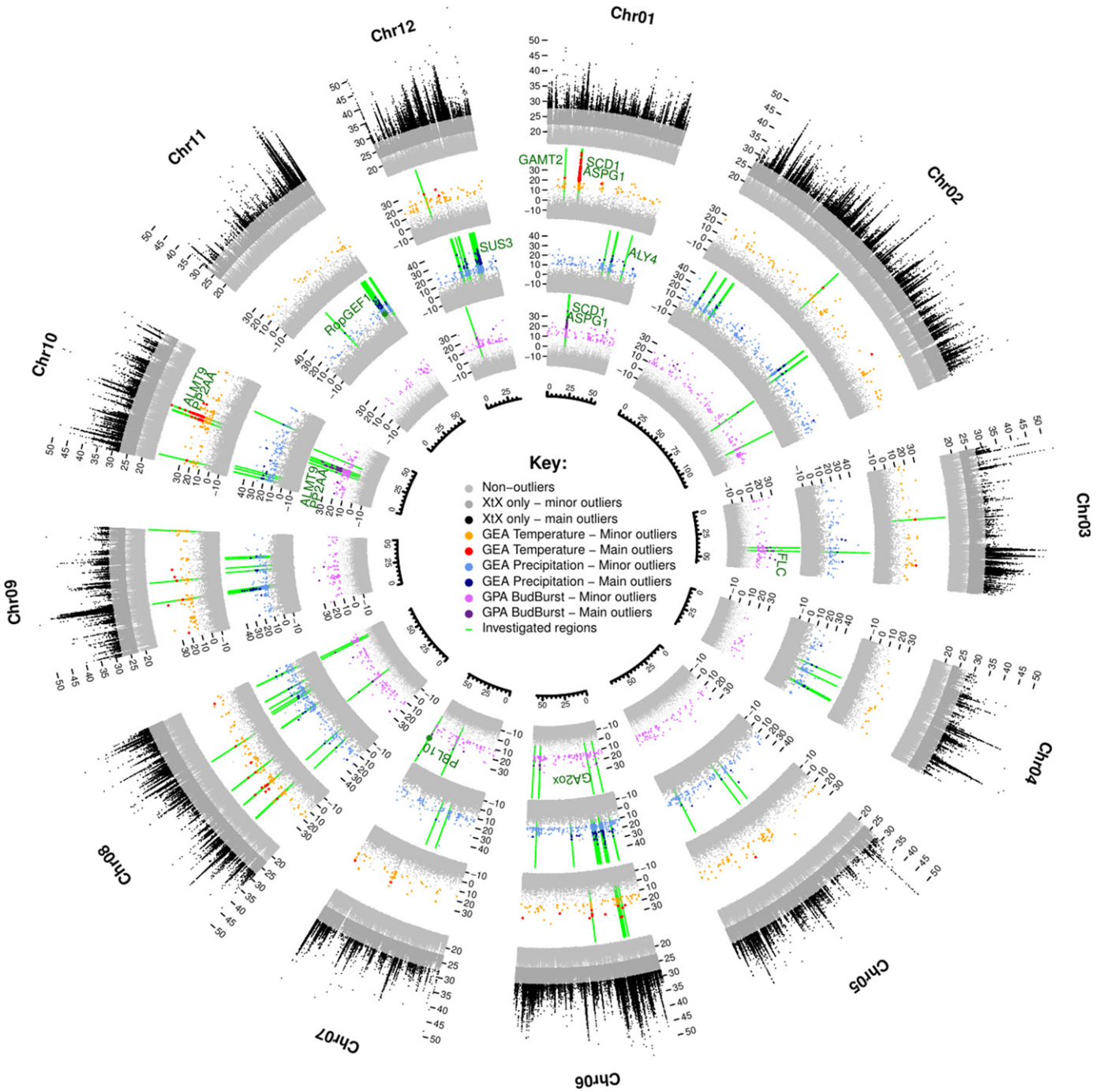

Fig. 2 Circular summary of the genome-wide scans for divergence, genotype–environment associations (GEAs) and genotype-phenotype associations (GPAs). From external to internal: divergence (XtX) and associations between temperature, precipitation, and the timing of leaf unfolding and allele frequencies. Colors highlight the significance of the single nucleotide polymorphisms (light: 'minor'; dark: 'main') evaluated with the calibration procedure described by Gautier (2015). Regions investigated for the gene annotation step are shown in green, and the genes discussed in the manuscript are indicated.

partially overlapped for precipitation and leaf unfolding (Fig. 2, on chromosomes 3 and 6; see also Figs S7, S10).

Gene annotations

Manual gene annotations were performed for the genes located within or close to these 201 regions (within 5 kb on either side) to exclude border effects. This led to the identification of 167 unique candidate genes. We found regulators of various growth and development processes in plants. As expected, we identified a set of genes involved in stomatal responses to water stress (e.g. *ALY4*, *ALMT9*, *PBL10*, *PP2AA*, *RopGEF1*, *SUS3*, *SCD1*). Concordantly, our genome scans also revealed some temperature-associated genes acting as regulators of the production of

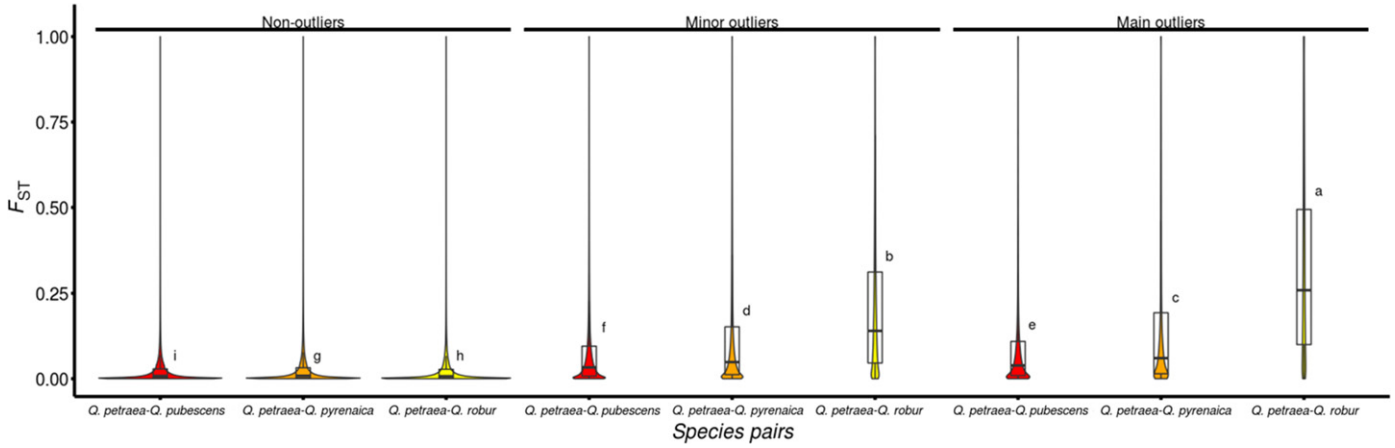

Fig. 3 Interspecific F_{ST} between the *Quercus petraea* reference pool and reference populations for three other oak species (for details, see Leroy *et al*., 2020) at XtX outlier loci. Tukey's honestly significant difference criterion at a significance level of 0.05 is reported. (a, b, h) *Quercus petraea*–*Quercus robur* differentiation; (c, d, g) *Q. petraea*–*Quercus pyrenaica* differentiation; (e, f, i): *Q. petraea*–*Quercus pubescens* differentiation).

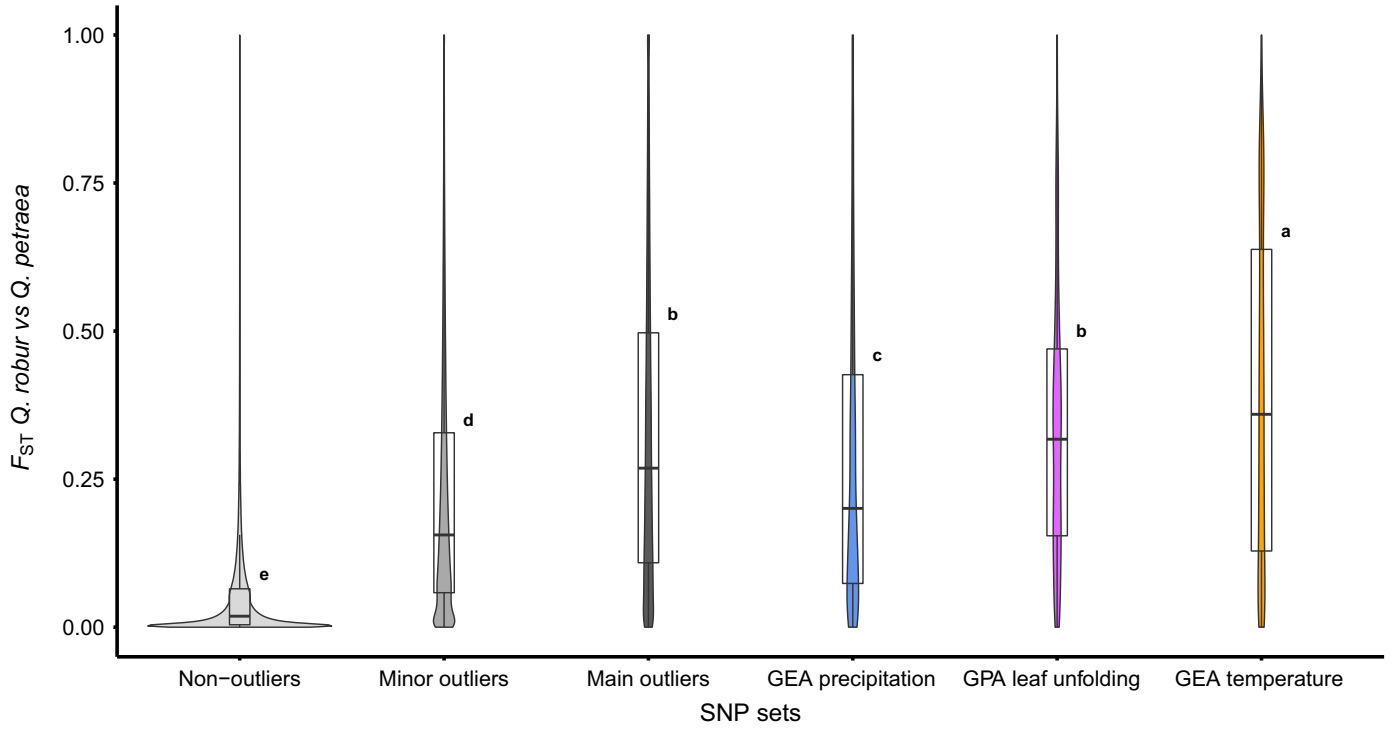

Fig. 4 Variation in interspecific F_{ST} between *Quercus robur* and *Quercus petraea* for nonoutlier and outlier single nucleotide polymorphisms (SNPs), including sets of associated with polymorphisms genotype–environment association (GEA) and genotype-phenotype association (GPA). Tukey's honestly significant difference criterion at a significance level of 0.05 is reported. SNPs for the GEA and GPA sets correspond to all associated SNPs (i.e. both the minor and main categories).

gibberellins (e.g. *GAMT2*, *ASPG1* or *GA2ox*). These genes also play important roles in various developmental processes, including seed dormancy (Shen *et al*., 2018).

We investigated the processes at work further, by generating ecological clines for each associated SNP (Notes S1). Most of the associated SNPs followed complex clines along the ecological gradients (i.e. nonlinear), highlighting the advantages of model implementation in BAYENV or BAYPASS rather than the use of linear or logistic model-derived methods (De Mita *et al*., 2013), but some SNPs displayed an almost linear change in allele frequency along the ecological gradient (Fig. 5). The clines were generally consistent with a continuum between the *Q. petraea* and *Q. robur* reference pools, with higher levels of *Q. robur*-like alleles in populations living in cooler and/or wetter environments, suggesting that interspecific introgression from *Q. robur* is an important source of adaptive variation for *Q. petraea* populations. Typical examples include the precipitation-associated SNPs and leaf-unfolding-associated SNPs located within two important genes

controlling stomatal responses, the *RopGEF1* and *PBL10* (=*APK1b*) genes (Fig. 5a and b respectively; Li & Liu, 2012; Elhaddad *et al.*, 2014).

Discussion

We assembled data for phenotypic surveys in common garden experiments, climatic data for the origin of the population, and whole-genome sequences to assess the genetic divergence of the timing of leaf unfolding between extant populations of *Q. petraea* at the genomic level. We drew inferences about population history by searching for footprints of population splits and admixture, and we conducted genome scans to explore the establishment of genetic divergence during the early Holocene in extant populations. We found that introgression from *Q. robur* had made a major contribution to population divergence of some locally adapted *Q. petraea* populations. This finding was supported by three major outcomes. First, we found that admixture events occurred mainly in *Q. petraea* populations from cooler and wetter climates. Second, the SNPs displaying the highest level of genetic differentiation between *Q. petraea* populations were also highly differentiated between *Q. petraea* and *Q. robur*. Third, some of the genes contributing to phenological divergence in *Q. petraea* displayed clinal variation consistent with the geographic variation of introgression. These findings echo earlier results in a Californian oaks showing that introgression follows environmental gradients. For example, the level of admixture in *Quercus wislizeni* was correlated with climatic variables, with more introgressed alleles from *Quercus parvula* and *Quercus agrifolia* in areas where summer temperatures were lower (Dodd & Afzal-Rafii, 2004). Conversely, associations with moisture gradients were revealed in a Californian white oak (*Quercus engelmannii*), with less introgression at drought-associated genes (Oney-Birol *et al.*, 2018).

Historical introgression from *Q. robur* into *Q. petraea*

The gene flow events inferred by TreeMix suggest that admixture events, mostly involving *Q. robur*, underlie the genetic differentiation of current *Q. petraea* populations over wide geographic gradients. Contemporary hybridization between other closely related white oaks and *Q. petraea* has been reported in empirical contemporary gene flow studies (Curtu *et al.*, 2007; Salvini *et al.*, 2009; Lepais & Gerber, 2011), but we found evidence only for an evolutionary footprint of admixture with *Q. robur* in our sampled populations of *Q. petraea*. Today, *Q. petraea* and *Q. robur* are frequently found in sympatric stands across Europe, due to their common pattern of postglacial colonization dynamics. As suggested by the widespread sharing of chloroplast haplotypes between these two closely related species when present in the same stands (Petit *et al.*, 2002), hybridization followed by recurrent backcrossing between *Q. petraea* and *Q. robur* has been an essential mechanism in *Q. petraea* expansion. Similarly, hybridization and subsequent backcrossing between the pioneer (resident) species *Q. robur* and the late successional (invading) species *Q. petraea* made it possible for *Q. petraea* to migrate with *Q. robur* (Petit *et al.*, 2004; Guichoux *et al.*, 2013). Most of the admixture events inferred from the TreeMix model support directional introgression from *Q. robur* to *Q. petraea* (Fig. 1), as predicted by theory in the case of this invasion scenario (Currat *et al.*, 2008).

The ultimate outcome of hybridization followed by unidirectional backcrossing is the restoration of *Q. petraea* within *Q. robur* stands. As suggested by artificial backcross experiments performed with breeding populations of other closely related European white oak species (Diskin *et al.*, 2006), the genome of the invading species (here *Q. petraea*) should be regenerated in a limited number of generations, ultimately maintaining a limited imprint of the introgression from *Q. robur*. However, our results show that genes introgressed from *Q. robur* are maintained in

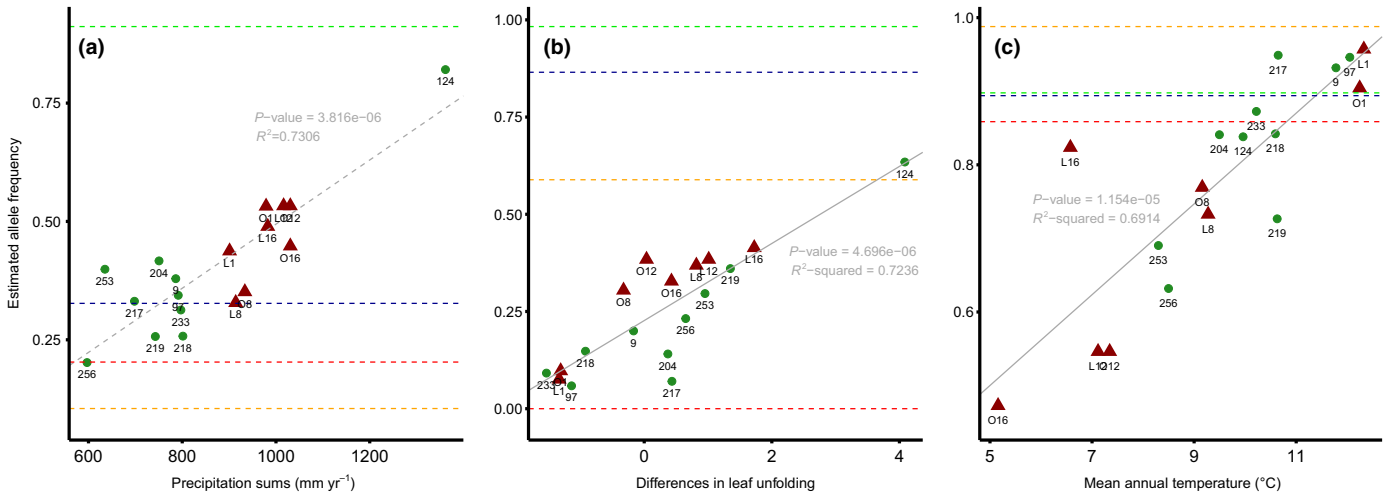

Fig. 5 Near-linear genetic clines along ecological gradients. The clines shown are for the three single nucleotide polymorphisms (SNPs) detected by BayPass: (a) the precipitation-associated SNP Chr11:47276412, (b) the leaf-unfolding-associated SNP Chr07:51928762, and (c) the temperature-associated SNP Chr07:28987726. Best linear regressions for each associated SNP are shown in gray. Red, blue, green, and orange dotted lines correspond to allele frequencies estimated in the *Quercus petraea*, *Quercus pubescens*, *Quercus robur*, and *Quercus pyrenaica* reference populations, respectively. Clines for all other associated SNPs are shown in Supporting Information Notes S1.

Q. petraea populations, particularly those from northern latitudes or higher elevations. Previous studies based on Bayesian clustering methods have already highlighted the occurrence of more admixture at higher elevations (Alberto *et al.*, 2010). Introgressed genes may still be present if admixture occurred recently, if there is a demographic imbalance between the invader and resident populations during initial contact (Currat *et al.*, 2008), or if these populations are subject to directional selection. Studies of pollen remains have indicated that temperate deciduous oaks have been present in sampled areas since *c.* 10 000 yr BP in northern Germany (Alberto *et al.*, 2010; Giesecke, 2016) and the Pyrénées (Jalut *et al.*, 1992; Reille & Lowe, 1993) and since *c.* 8000 yr BP in Ireland (Kelleher *et al.*, 2004). This timescale would be long enough for completion of the regeneration process and the eradication of introgression. We cannot rule out the possibility that demographic imbalance between these two species facilitated the introgression and maintenance of *Q. robur* alleles in *Q. petraea*, but the detection of introgression principally at higher elevations and latitudes suggests that some introgressed alleles were most probably maintained by selection, increasing the degree of differentiation between *Q. petraea* populations (Fig. 4). Interestingly, our results also show that genes highly differentiated between *Q. petraea* populations are also highly differentiated between *Q. petraea* and *Q. robur* (Figs 4, 5), suggesting their involvement in either adaptive divergence or reproductive barriers between the two species (Leroy *et al.*, 2020). These two species display subtle differences in soil preferences when present in the same forest (Timbal & Aussenac, 1996; Eaton *et al.*, 2016), but they also have different climate responses, as suggested by their allopatric (temperature- and precipitation-dependent) distributions at the edge of their ranges. *Quercus robur* extends farther north (up to Finland) and east (up to the Ural Mountains) – see Leroy *et al.* (2020) – and has a higher frequency in wetter climates (Eaton *et al.*, 2016). We suspect that the introgression of *Q. robur* alleles for genes involved in these differential responses may have contributed to the expansion of *Q. petraea* populations to higher elevations in the Pyrenees and wetter climates in Ireland.

Genomic and genetic clines shaped by introgression

Association studies identified additional elements relating to the probable contribution of introgression to the adaptive divergence of *Q. petraea* populations. First, we recovered the clinal phenotypic gradient of leaf unfolding with temperature variations (Fig. S1B), as observed in previous common garden experiments (Ducousso *et al.*, 1996; Vitasse *et al.*, 2009; Alberto *et al.*, 2011,2013b; Firmat *et al.*, 2017), with populations from cooler climates (higher latitudes or elevations) flushing later than populations from warmer climates. Second, we found that the genes displaying clinal variations of allelic frequency along a temperature gradient were enriched in genes differentiated between the *Q. petraea* and *Q. robur* species. This raises questions about whether the genes introgressed from *Q. robur* also contributed to the later flushing of *Q. petraea* at higher elevations. No significant interspecific differences in leaf unfolding are observed between these two closely related species in common garden experiments or *in situ* (Kleinschmit, 1993; Jensen & Hansen, 2008; Wilkinson *et al.*, 2017), although there is a slight trend towards earlier flushing in *Q. petraea* (Kleinschmit, 1993). However, *Q. robur* is well known to display ecotypic differentiation, with the recognition of late-flushing populations that have been attributed subspecies status as *Q. robur* var. *tardiflora*, mostly in eastern Europe, and flush almost a month later than *Q. robur* (Wesolowski & Rowinski, 2008; Utkina & Rubtsov, 2017). Similar extremely late-flushing populations have also been reported in western Europe for *Q. robur* (Riedacker, 1968), but no such phenological differentiation has been reported for *Q. petraea*.

In conclusion, we found that most outliers and associated loci are *Q. robur* ancestry-informative, and that introgression from *Q. robur* to *Q. petraea* is correlated to the overall clinal variation for some key adaptive traits, including leaf-unfolding-related traits in particular. Our results are, however, only based on statistical associations and do not necessarily imply causality. Complex demographic scenarios, as well as linked selection, are known to produce footprints that mimic local selection (e.g. Fraïsse *et al.*, 2018; Simon & Duranton, 2018). Causal inference methods based on hybrids of backcrossed seedlings or gene modification experiments will be needed to discover the genes/polymorphisms that matter for adaptation. In addition to shedding light on the evolutionary processes at work, this study identified several key oak genes involved in the stomatal behavior as potentially involved in adaptive introgression, providing support for the general view that stomatal regulation makes a major contribution to the maintenance of homeohydry in land plants (Brodribb & McAdam, 2017). Other leaf-unfolding-associated genes would be expected to have key roles, with pleiotropic effects. A typical example is provided by Flowering Locus C (FLC). FLC antagonizes the gibberellin pathway and downregulates flowering (Deng *et al.*, 2011; Li *et al.*, 2015). The effects of FLC are known to be suppressed by vernalization, suggesting that FLC plays a major role in seasonal and developmental timing.

Acknowledgements

This research was funded by the European Research Council under the European Union's Seventh Framework Programme (TREEPEACE project, FP/2014-2019: ERC Grant Agreement no. 339728) and by the French ANR (GENOAK project, 11-BSV6-009-021). We thank the Genotoul Bioinformatics Platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) and the Biogenouest BiRD core facility (Université de Nantes) for providing computing and storage resources. We thank the staff of the Experimental Units of INRA Nancy (UEFL, Unité Expérimentale Forestière de Lorraine) and INRA Toulenne (UE 0393, INRA, Domaine des Jarres, 33210-Toulenne) for their contribution during field phenological assessments and sampling. We also thank Jorge A. P. Paiva for providing access to *Q. suber* data, Nick Zimmerman for providing climate data, Mathieu Gautier for providing advice on how to make the best use of BayPass, and Quentin Rougemont for fruitful discussions. TL also thanks the project coordinator Benoit Nabholz (BirdIslandGenomic project, ANR-14-CE02-0002) for his support and feedback. We

are grateful to Nicolas Bierne and to the three anonymous reviewers for their helpful comments.

Author contributions

TL designed and performed the research, analyzed the data, and drafted the manuscript. J-ML and SD were involved in the acquisition of the climate and common garden data. CL performed the DNA extractions and equimolarly pooled DNA. GLP, JML and CP were involved in the sampling. KL and J-MA generated the sequencing data. CP and AK contributed to the design of the research, interpretation, and drafted the manuscript.

ORCID

Jean-Marc Aury https://orcid.org/0000-0003-1718-3010
Sylvain Delzon https://orcid.org/0000-0003-3442-1711
Karine Labadie https://orcid.org/0000-0001-7467-8509
Thibault Leroy https://orcid.org/0000-0003-2259-9723
Christophe Plomion https://orcid.org/0000-0002-3176-2767

References

Alberto F, Bouffier L, Louvet JM, Lamy JB, Delzon S, Kremer A. 2011. Adaptive responses for seed and leaf phenology in natural populations of sessile oak along an altitudinal gradient. *Journal of Evolutionary Biology* **24**: 1442–1454.

Alberto F, Niort J, Derory J, Lepais O, Vitalis R, Galop D, Kremer A. 2010. Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. *Molecular Ecology* **19**: 2626–2639.

Alberto FJ, Aitken SN, Alia R, Gonzalez-Martinez SC, Hanninen H, Kremer A, Lefevre F, Lenormand T, Yeaman S, Whetten R *et al.* 2013a. Potential for evolutionary responses to climate change evidence from tree populations. *Global Change Biology* **19**: 1645–1661.

Alberto FJ, Derory J, Boury C, Frigerio JM, Zimmermann NE, Kremer A. 2013b. Imprints of natural selection along environmental gradients in phenology-related genes of *Quercus petraea*. *Genetics* **195**: 495–512.

Bodénès C, Labbé T, Pradère S, Kremer A. 1997. General vs local differentiation between two closely related white oak species. *Molecular Ecology* **6**: 723–724.

Brewer S, Giesecke T, Davis BAS, Finsinger W, Wolters S, Binney H, de Beaulieu JL, Fyfe R, Gil-Romera G, Kuhl N *et al.* 2017. Late-glacial and Holocene European pollen data. *Journal of Maps* **13**: 921–928.

Brodribb TJ, McAdam SAM. 2017. Evolution of the stomatal regulation of plant water content. *Plant Physiology* **174**: 639–649.

Currat M, Ruedi M, Petit RJ, Excoffier L. 2008. The hidden side of invasions: massive introgression by local genes. *Evolution* **62**: 1908–1920.

Curtu AL, Gailing O, Finkeldey R. 2007. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evolutionary Biology* **7**: e218.

Dannemann M, Racimo F. 2018. Something old, something borrowed: admixture and adaptation in human evolution. *Current Opinion in Genetics & Development* **53**: 1–8.

De Mita S, Thuillet AC, Gay L, Ahmadi N, Manel S, Ronfort J, Vigouroux Y. 2013. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology* **22**: 1383–1399.

Deng WW, Ying H, Helliwell CA, Taylor JM, Peacock WJ, Dennis ES. 2011. FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **108**: 6680–6685.

Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A, Casasoli M, Chancerel E, Bodénès C, Alberto F *et al.* 2010. Contrasting correlations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity* **104**: 438–448.

Diskin M, Steiner KC, Hebard FV. 2006. Recovery of American chestnut characteristics following hybridization and backcross breeding to restore blight-ravaged *Castanea dentata*. *Forest Ecology and Management* **223**: 439–447.

Dodd R, Afzal-Rafii Z. 2004. Selection and dispersal in a multispecies oak hybrid zone. *Evolution* **58**: 261–269.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* **19**: 11–15.

Ducousso A, Guyon JP, Kremer A. 1996. Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (*Quercus petraea* (Matt) Liebl). *Annales des Sciences Forestières* **53**: 775–782.

Dullinger S, Gattringer A, Thuiller W, Moser D, Zimmermann NE, Guisan A, Willner W, Plutzar C, Leitner M, Mang T *et al.* 2012. Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Climate Change* **2**: 619–622.

Dynesius M, Jansson R. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences, USA* **97**: 9115–9120.

Eaton E, Caudullo G, Oliveira S, de Rigo D. 2016. *Quercus robur* and *Quercus petraea* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz J, de Rigo D, Caudullo G, Houston D, Urrant T, Mauri A, eds. *European atlas of forest trees*. Brussels, Belgium: Publication Office of the European Union, 160–163.

Elhaddad NS, Hunt L, Sloan J, Gray JE. 2014. Light-induced stomatal opening is affected by the guard cell protein kinase APK1b. *PLoS ONE* **9**: e97161.

Firmat C, Delzon S, Louvet JM, Parmentier J, Kremer A. 2017. Evolutionary dynamics of the leaf phenological cycle in an oak metapopulation along an elevation gradient. *Journal of Evolutionary Biology* **30**: 2116–2131.

Fraïsse C, Haguenauer A, Gerard K, Weber AAT, Bierne N, Chenuil A. 2018. Fine-grained habitat-associated genetic connectivity in an admixed population of mussels in the small isolated Kerguelen Islands. *bioRxiv*: 239244.

Gautier M. 2015. Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* **201**: 1555–1579.

Giesecke T. 2016. Did thermophilous trees spread into central Europe during the Late Glacial? *New Phytologist* **212**: 15–18.

Guichoux E, Garnier-Gere P, Lagache L, Lang T, Boury C, Petit RJ. 2013. Outlier loci highlight the direction of introgression in oaks. *Molecular Ecology* **22**: 450–462.

Gunther T, Coop G. 2013. Robust identification of local adaptation from allele frequencies. *Genetics* **195**: 205–220.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Jalut G, Monserrat Marti J, Fontugne M, Delibrias G, Vilaplana JM, Julia R. 1992. Glacial to interglacial vegetation changes in the northern and southern Pyrénées: deglaciation, vegetation cover and chronology. *Quaternary Sciences Reviews* **11**: 449–482.

Jensen JS, Hansen JK. 2008. Geographical variation in phenology of *Quercus petraea* (Matt.) Liebl and *Quercus robur* L. oak grown in a greenhouse. *Scandinavian Journal of Forest Research* **23**: 179–188.

Kelleher CT, Hodkinson TR, Kelly DL, Douglas GC. 2004. Characterisation of chloroplast DNA haplotypes to reveal the provenance and genetic structure of oaks in Ireland. *Forest Ecology and Management* **189**: 123–131.

Kleinschmit J. 1993. Intraspecific variation of growth and adaptive traits in European oak species. *Annales des Sciences Forestieres* S166–S185.

Koffler R, Pandey RV, Schlötterer C. 2011. Popoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**: 3435–3436.

Kremer A, Le Corre V. 2012. Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* **108**: 375–385.

Lepais O, Gerber S. 2011. Reproductive patterns shape introgression dynamics and species succession within the European white oak species complex. *Evolution* **65**: 156–170.

Leroy T, Rougemont Q, Dupouey J-L, Bodénès C, Lalanne C, Belser C, Labadie K, Le Provost G, Aury J-M, Kremer A *et al.* 2020. Massive postglacial gene flow between European white oaks uncovered genes underlying species barriers. *New Phytologist* **226**: 1183–1197.

Leroy T, Roux C, Villate L, Bodenes C, Romiguier J, Paiva JAP, Dossat C, Aury JM, Plomion C, Kremer A. 2017. Extensive recent secondary contacts between four European white oak species. *New Phytologist* **214**: 865–878.

Li M, An F, Li W, Ma M, Feng Y, Zhang X, Guo H. 2015. DELLA proteins interact with FLC to repress flowering transition. *Journal of Integrative Plant Biology* **58**: 642–655.

Li ZX, Liu D. 2012. ROPGEF1 and ROPGEF4 are functional regulators of ROP11 GTPase in ABA-mediated stomatal closure in *Arabidopsis*. *FEBS Letters* **586**: 1253–1258.

Lindner M, Maroschek M, Netherer S, Kremer A, Barbati A, Garcia-Gonzalo J, Seidl R, Delzon S, Corona P, Kolstrom M *et al.* 2010. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest Ecology and Management* **259**: 698–709.

Mauri A, Davis BAS, Collins PM, Kaplan JO. 2015. The climate of Europe during the Holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation. *Quaternary Science Reviews* **112**: 109–127.

Oney-Birol S, Fitz-Gibbon S, Chen J-M, Gugger PF, Sork VL. 2018. Assessment of shared alleles in drought-associated candidate genes among southern California white oak species (*Quercus* sect. *Quercus*). *BMC Genetics* **19**: e88.

Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A. 2004. Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**: 151–164.

Petit RJ, Csaikl UM, Bordacs S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapegue S, Fineschi S *et al.* 2002. 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.

Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics* **8**: e1002967.

Plomion C, Aury JM, Amselem J, Leroy T, Murat F, Duplessis S, Faye S, Francillonne N, Labadie K, Le Provost G *et al.* 2018. Oak genome reveals facets of long lifespan. *Nature Plants* **4**: 440–452.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. *Nature* **461**: 489–494.

Reille M, Lowe JJ. 1993. A re-evaluation of the vegetation history of the eastern Pyrenees (France) from the end of the last glacial to the present. *Quaternary Science Reviews* **12**: 47–77.

Riedacker A. 1968. Le chêne de juin à l'étang de Brin. *Revue Forestière Française* 487–492.

Saenz-Romero C, Lamy JB, Ducousso A, Musch B, Ehrenmann F, Delzon S, Cavers S, Chalupka W, Dagdas S, Hansen JK *et al.* 2017. Adaptive and plastic responses of *Quercus petraea* populations to climate across Europe. *Global Change Biology* **23**: 2831–2847.

Salvini D, Bruschi P, Fineschi S, Grossoni P, Kjaer ED, Vendramin GG. 2009. Natural hybridisation between *Quercus petraea* (Matt.) Liebl. and *Quercus pubescens* Willd. within an Italian stand as revealed by microsatellite fingerprinting. *Plant Biology* **11**: 758–765.

Savolainen O, Pyhajarvi T, Knurr T. 2007. Gene flow and local adaptation in trees. *Annual Review of Ecology Evolution and Systematics* **38**: 595–619.

Schaefer NK, Shapiro B, Green RE. 2016. Detecting hybridization using ancient DNA. *Molecular Ecology* **25**: 2398–2412.

Shapiro B, Hofreiter M. 2014. A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. *Science* **343**: e1236573.

Shen WZ, Yao X, Ye TT, Ma S, Liu X, Yin XM, Wu Y. 2018. *Arabidopsis* aspartic protease ASPG1 affects seed dormancy, seed longevity and seed germination. *Plant and Cell Physiology* **59**: 1415–1431.

Simon A, Duranton M. 2018. Digest: demographic inferences accounting for selection at linked sites. *Evolution* **72**: 1330–1332.

Suarez-Gonzalez A, Lexer C, Cronk QCB. 2018. Adaptive introgression: a plant perspective. *Biology Letters* **14**: doi: 10.1098/rsbl.2017.0688.

Sukumaran J, Holder MT. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**: 1569–1571.

Timbal J, Aussenac G. 1996. An overview of ecology and silviculture of indigenous oaks in France. *Annals of Forest Science* **53**: 649–661.

Torres-Ruiz JM, Kremer A, Carins-Murphy MR, Brodribb TJ, Lamarque LJ, Truffaut L, Bonne F, Ducousso A, Delzon S. 2019. Genetic differentiation in functional traits among European sessile oak populations. *Tree Physiology*, in press. doi: 10.1101/584847.

Utkina IA, Rubtsov VV. 2017. Studies of phenological forms of pedunculate oak. *Contemporary Problems of Ecology* **10**: 804–811.

Vitasse Y, Delzon S, Dufrene E, Pontailler JY, Louvet JM, Kremer A, Michalet R. 2009. Leaf phenology sensitivity to temperature in European trees: do within-species populations exhibit similar responses? *Agricultural and Forest Meteorology* **149**: 735–744.

Wagner S, Lagane F, Seguin-Orlando A, Schubert M, Leroy T, Guichoux E, Chancerel E, Bech-Hebelstrup I, Bernard V, Billard C *et al.* 2018. High-throughput DNA sequencing of ancient wood. *Molecular Ecology* **27**: 1138–1154.

Wesolowski T, Rowinski P. 2008. Late leaf development in pedunculate oak (*Quercus robur*): an antiherbivore defence? *Scandinavian Journal of Forest Research* **23**: 386–394.

Wilkinson M, Eaton EL, Morison JIL. 2017. Variation in the date of budburst in *Quercus robur* and *Q. petraea* across a range of provenances grown in southern England. *European Journal of Forest Research* **136**: 1–12.

Yeaman S. 2015. Local adaptation by alleles of small effect. *American Naturalist* **186**: S74–S89.

Zimmermann NE, Yoccoz NG, Edwards TC, Meier ES, Thuiller W, Guisan A, Schmatz DR, Pearman PB. 2009. Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences, USA* **106**: 19723–19728.

Supporting Information

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

Fig. S1 Climatic and phenological data for the *Quercus petraea* populations studied.

Fig. S2 Best-fit Treemix trees for various numbers of migration edges.

Fig. S3 *F3*-statistics for admixture between *Q. petraea* and *Q. robur*.

Fig. S4 Intra- and interspecific differentiation between *Q. robur* and *Q. petraea*.

Fig. S5 Scans for divergence, GEA and GPA associations for SNPs on chromosome 1.

Fig. S6 Scans for divergence, GEA and GPA associations for SNPs on chromosome 2.

Fig. S7 Scans for divergence, GEA and GPA associations for SNPs on chromosome 3.

Fig. S8 Scans for divergence, GEA and GPA associations for SNPs on chromosome 4.

Fig. S9 Scans for divergence, GEA and GPA associations for SNPs on chromosome 5.

Fig. S10 Scans for divergence, GEA and GPA associations for SNPs on chromosome 6.

Fig. S11 Scans for divergence, GEA and GPA associations for SNPs on chromosome 7.

Fig. S12 Scans for divergence, GEA and GPA associations for SNPs on chromosome 8.

Fig. S13 Scans for divergence, GEA and GPA associations for SNPs on chromosome 9.

Fig. S14 Scans for divergence, GEA and GPA associations for SNPs on chromosome 10.

Fig. S15 Scans for divergence, GEA and GPA associations for SNPs on chromosome 11.

Fig. S16 Scans for divergence, GEA and GPA associations for SNPs on chromosome 12.

Notes S1 Genetic clines along ecological gradients for GEA and GPA-associated SNPs.

Table S1 Number of SNPs used or outliers identified in this study.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

See also the Editorial on this article by Plomion & Martin, **226**: 943–946.