Heredity (2010) 105, 401–411

www.nature.com/hdy

ORIGINAL ARTICLE

Contrasting relations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks

J Derory[1,2], C Scotti-Saintagne[1,2,4], E Bertocchi[1,2], L Le Dantec[3], N Graignic[1,2], A Jauffres[1,2], M Casasoli[1,2], E Chancerel[1,2], C Bodénès[1,2], F Alberto[1,2] and A Kremer[1,2]

[1]*INRA, UMR1202 BIOGECO, Cestas, France;* [2]*Université de Bordeaux, UMR1202 BIOGECO, Cestas, France;* [3]*INRA, UPR419, Unité de Recherche sur les Espèces Fruitières, Villenave d'Ornon, France and* [4]*INRA, UMR745 ECOFOG, Kourou, French Guiana*

Nucleotide diversity was assessed within nine candidate genes (in total 4.6 kb) for the time of bud burst in nine sessile oak (*Quercus petraea*) populations distributed in central and northern Europe. The sampled populations were selected on the basis of their contrasting time of bud burst observed in common garden experiments (provenance tests). The candidate genes were selected according to their expression profiles during the transition from quiescent to developing buds and/or their functional role in model plants. The overall nucleotide diversity was large ($\pi_{tot} = 6.15 \times 10^{-3}$; $\pi_{silent} = 11.2 \times 10^{-3}$), but population differentiation was not larger than for microsatellites. No outlier single-nucleotide polymorphism (SNP), departing from neutral expectation, was found among the total of 125 SNPs. These results contrasted markedly with the significant associations that were observed between the candidate genes and bud burst in segregating populations. Quantitative trait loci (QTLs) for bud burst were identified for 13 year*site seasonal observations in a cloned mapping pedigree. Nineteen QTLs were detected, and QTLs located on linkage groups 2, 5 and 9 contributed repeatedly to more than 12% of the phenotypic variation of the trait. Eight genes were polymorphic in the two parents of the pedigree and could be mapped on the existing genetic map. Five of them located within the confidence intervals of QTLs for bud burst. Interestingly, four of them located within the three QTLs exhibiting the largest contributions to bud burst.
Heredity (2010) **105,** 401–411; doi:10.1038/hdy.2009.170; published online 9 December 2009

Keywords: nucleotide diversity; candidate gene; QTL; *Quercus petraea*

Introduction

Spatial distribution and abundance of long-lived species, such as forest trees, is closely related to phenological events, and especially bud burst (Chuine and Beaubien, 2001). Changes in the date of bud burst may modify the length of the growing season, the flowering time and reproductive success of trees. The fact that bud burst is a fitness-related trait is confirmed by the important clinal variation that is observed in provenance tests of many tree species (Savolainen *et al*., 2007). As the date of bud burst is mainly driven by temperature for temperate tree species (Chuine and Cour, 1999), global warming is expected to have a major impact on the phenology of trees. An increase in the length of the growing season has already been observed for trees in Europe and North America (Menzel and Fabian, 1999; Penuelas and Filella, 2001). Changes in the date of bud burst will modify exposures to late frost (Scheifinger *et al*., 2003), or to outbreaks of phytophagous insects (Van Asch and Visser, 2007). The adaptive response of trees to these rapid climatic changes will depend on the levels of genetic variation within natural populations. In this contribution, we explore the distribution of genetic variation of bud burst in European white oaks, by focusing on the trait variation itself and on the diversity of the underlying putative genes controlling the trait. Large latitudinal and altitudinal clinal variations were observed in provenance tests (Ducousso *et al*., 1996). In a recent contribution, we showed that the geographical pattern of bud burst variation was resulted from local adaptation in response to the warming following the last glaciation (Kremer *et al*., 2009). The clinal pattern observed along latitude and altitude is therefore the consequence of natural selection mediated by the temperature gradient, which has occurred during the last 15 000 years, and is not imprinted by the earlier differentiation existing among the refugial populations (Kremer *et al*., 2002)

The between-population variation was shown to be associated to the maintenance of important genetic variation within populations (Scotti-Saintagne *et al*., 2004). Previous quantitative trait locus (QTL) detection analysis suggested that the important genetic variation was likely due to a large number of genes, with rather small-to-moderate effects (Scotti-Saintagne *et al*., 2004). Technical limitations or reduced pedigree size in forest trees did not permit to implement QTL cloning. Indeed, positional cloning (Tanksley *et al*., 1995) and insertion

Correspondence: Dr A Kremer, INRA, UMR BIOGECO 1202, 69, rte d'Arachon, Cestas F-33610, France.
E-mail: antoine.kremer@pierroton.inra.fr
Received 18 August 2009; accepted 27 August 2009; published online 9 December 2009

mutagenesis (Bechtold *et al*., 1993) are not available for tree species. The research of the underlying genes residing within the QTLs was therefore oriented toward a candidate gene (CG) approach (Pflieger *et al*., 2001). In a recent study, accumulation of transcripts in apical buds was monitored during spring time in order to identify differentially expressed genes before and after bud flush of sessile oak, leading to the identification of expressional CGs that were either up- or downregulated during the transition from quiescent to developing buds (Derory *et al*., 2006). In this study, we compare the relations between the diversity of the CGs and the variation of the time of bud burst in natural and segregating populations in order to validate the expressional CGs. In natural populations, we assessed nucleotide diversity and differentiation in populations showing contrasting timing of bud burst. Molecular imprint of natural selection for bud burst was tested in two different ways: (i) by comparing the differentiation of CG sequences with control neutral markers (for example, microsatellites) and (ii) by comparing the differentiation of CG sequences between geographical and phenological groups of populations. In segregation populations, we refined the QTL position of the timing of bud burst by compiling observations across environments and years, and compared their position with the position of CG on the genetic map following earlier analysis conducted by Scotti-Saintagne *et al*. (2004) and Casasoli *et al*. (2006). Our overall objective was to strengthen the selection of CGs based on their variation in natural and segregation populations.

Materials and methods

Sequencing, genotyping and diversity analysis of microsatellites and CGs

Plant material: The study material used for the assessment of nucleotide polymorphisms in CGs and SSRs loci come from a common garden experiment (provenance test) installed in four different plantations in the Central Part of France (from the West to the East). Nine oak populations were sampled within the provenance tests comprising 112 in total (Ducousso *et al*., 1996). The plantations were installed between 1990 and 1995, and each provenance was represented by 10 repetitions of 24 trees in each test (Table 1). Bud burst was assessed in spring after 3 years of plantation following a grading system ranging between 0 (quiescent bud) to 5 (fully developed leaves and elongated shoot). Population differentiation (Q_{ST}) for bud burst, as assessed in these tests, amounted to 0.55 (Ducousso *et al*., 2005). The nine sampled populations were chosen according to their contrasting bud burst recorded in the tests (from early to late flushing) and their geographic distribution. Within this study, the populations were clustered in two different ways: according to the time of the bud burst as recorded in the provenance test (phenological groups) and according to their geographic origin (geographic groups) (Table 1). Within each of the nine sampled populations, six trees were selected at random for sequencing DNA of the CGs.

CGs and microsatellite loci: Nucleotide polymorphisms were assessed in partial fragments of 9 CGs and 15 microsatellites loci (Table 2). CGs were selected according to differential expression of expressed sequence tags (ESTs) before and after bud flush as assessed by cDNA macroarray experiments and real-time reverse-transcriptase PCR (Derory *et al*., 2006) or according to their functional role as assessed in model plants (Table 2). EST sequences were blasted in gene databases such as NCBI (http://www.ncbi.nlm.nih.gov/BLAST/), in order to search for homologs in oak species (Derory *et al*., 2006). A gene product was assigned to each EST, on the basis of sequence similarity to proteins with known function in the trEMBL and Swiss-prot databases, using BLASTX with an *e*-value $\leqslant 10^{-10}$ (http://www.expasy.ch/sprot/). The fragments sequenced in this study corresponded to the nine following genes. Galactinol synthase (GALA), Dof Affecting Germination 2 (DAG2), Alpha-amylase/subtilisin inhibitor (ASI) and metal-nicotianamine transporter (YSL1) were drastically downregulated during the transition from quiescent to developing buds (Derory *et al*., 2006), and thus constitute relevant expressional CGs for molecular signals regulating bud burst transition. The functional role of these genes was investigated in studies conducted on other plants. Indeed, GALA was reported to be involved

Table 1 Descriptive data of the sampled populations

Name	*Country*	*Longitude (°)*	*Latitude (°)*	*Altitude (m)*	*Bud burst score in provenance tests*[a]				*Phenological group*[b]	*Geographical group*[c]
					Test 1	*Test 2*	*Test 3*	*Test 4*		
Mölln	Germany	10.75	53.62	36	0.58	1.18	0.91	1.48	L	N
Sprakensehl	Germany	10.6	52.8	115	0.71	1.11	1.02	1.64	L	N
Cochem	Germany	7.05	50.08	400	1.86	2.17	2.02	3.22	E	C
La Neuville-en-Hez	France	2.33	49.4	70	2.15	1.87	1.93	3.05	–[d]	–[d]
Johanneskreuz	Germany	7.83	49.4	460	1.15	–[e]	1.13	2.14	I	C
Saint Jean	France	6.72	48.82	227	1.02	1.23	1.00	1.51	L	S
Westhoffen	France	7.45	48.6	400	2.77	2.7	2.63	3.21	E	S
Haslach	France	7.38	48.55	265	2	1.81	1.86	2.88	I	S
Klostermarienberg	Austria	16.57	47.41	310	3.28	3.8	4.15	4.08	–[d]	–[d]

[a]Results available in Ducousso *et al*. (1996).
[b]Phenological groups: E, early flushing; I, intermediate; L, late flushing.
[c]Geographical groups: S, Southern; C, Central; N, Northern.
[d]Populations Klostermarienberg and La Neuville-en-Hez not assigned to geographical nor phenological groups because of their eccentric location.
[e]Population Johanneskreuz not present in test 2.

Table 2 Summary data of candidate genes

Candidate gene	*Annotation*	*Coding (bp)*	*Non-coding (bp)*	*Amplified region*	*No. of SNPs*	*No. of singletons*	*Size of indels (bp)*	*Reference for PCR primer pairs*	*Accession no.*
GALA	Galactinol synthase	210	295	E/I/E	18	7	28	Derory *et al*. (2006)	CR327918
ASI	Alpha-amylase/ subtilisin inhibitor	327	0	E	19	3	3/1	Derory *et al*. (2006)	CR627933
DAG2	Dof Affecting Germination 2	360	0	E	2	0		Derory *et al*. (2006)	CR627781
H3	Histone H3	126	248	E/3′-UTR	5	2	2/1 (3 ×)	Casasoli *et al*. (2006)	CR627839
PM23	Seed maturation protein PM23	390	714	E/I/E/I/E	41	3	12/5/4/3 (3 ×)/ 2 (2 ×)/1 (4 ×)	Derory *et al*. (2006)	CR627986
AUX-REP	Auxine-repressed protein	259	188	E/I/E	7	3	2/1	Casasoli *et al*. (2006)	CR627732
YSL1	Metal-nicotianamine transporter YSL1	315	0	E	8	0		Derory *et al*. (2006)	CR627947
GA3	GA3-β-hydroxylase	342	310	E/I/E	3	0	6	Forward: 5′-TCCCTAAGCGTATGTGGTCC-3′ Reverse: 5′-CCCTCTCGTGTTGTTTTGGT-3′	AJ006453
GA20	GA20-oxydase	425	129	E/I/E	4	0	27	Forward: 5′-TCAGCTCAACCAGACTCATC-3′ Reverse: 5′-TTGAAAGTGCCATGAAGGTG-3′	AJ420192
Total		2754	1884		107	18	23		

Abbreviations: E, exon; I, intron; indel, insertion–deletion; UTR, untranslated region.

in tolerance to drought, high salinity and cold (Pukacka and Wojkiewicz, 2002; Taji *et al*., 2002) in *Arabidopsis thaliana*. DAG2 was shown to act as a transcription factor specifically involved in the maternal control of seed germination (Gualberti *et al*., 2002). In buds, as in seeds, it may potentially act on dormancy release. ASI is expressed in germinating seeds of rice, barley and wheat; the protein is multifunctional and is involved in hydrolysis of starch metabolism and in seed defense against pathogens (Furtado *et al*., 2003; Nielsen *et al*., 2004). In oaks, ASI was highly expressed in buds only at the quiescent stage, suggesting that hydrolysis of storage starch or glycogen is repressed in the quiescent bud (Derory *et al*., 2006). The observed reduction in the expression of this gene on bud swelling would indicate the onset of starch mobilization at this developmental stage. Lastly, three CGs putatively encoding for an auxin-repressed protein (AUX-REP), a seed maturation protein (PM23) and a histone (H3) were also selected because of their downregulation in the expression study from quiescent to developing buds. Additional CGs were selected based on their functional role known from other plants. As gibberellins are known to be implied in bud flushing (Or *et al*., 2000; Falusi and Calamassi, 2003), we included Gibberellin 20-oxidase (GA20) and Gibberellin 3-β-hydroxylase (GA3) (Perez-Flores *et al*., 2003; Calvo *et al*., 2004; Israelsson *et al*., 2004), which are involved in the synthesis of GA1, the active form of gibberellins in plants.

In addition, we selected 15 microsatellites loci as 'control markers' that would picture the existing geographic structure for neutral markers (Mariette *et al*., 2002). Genotyping for microsatellites was done on the same individuals as for CGs. The following microsatellite loci were scored on the sampled trees: QrZAG11, QrZAG39, QrZAG96, QrZAG112, QpZAG110, QrZAG5, QrZAG7, QrZAG20, QrZAG65, QrZAG87, QpZAG9, QpZAG15, QpZAG46, QpZAG36 and MSQ13 (Dow *et al*., 1995; Steinkellner *et al*., 1997; Kampfer *et al*, 1998).

DNA extraction and PCR amplification: Total genomic DNA was extracted from buds using a slightly modified protocol of Saghaï-Maroof *et al*. (1984). Primers and PCR conditions to amplify CGs were those used either in Derory *et al*. (2006) or in Casasoli *et al*. (2006). Primer pairs for GA3 and GA20 were designed from published sequences in GenBank (accession numbers AJ006453 and AJ420192, respectively) and their characteristics are presented in Table 1. The amplified fragments of CG covered the domain of the protein or extended to the 3′-UTR (untranslated region) gene region (Table 2).

Microsatellite loci were amplified in four different sets using multiplex PCRs as described in Lepais *et al*. (2006) for two sets of primers and two other sets (MH Pemonge, unpublished results) were used and comprised QpZAG9, QpZAG15, QpZAG46 and QpZAG36 and MSQ13, respectively (Dow *et al*., 1995; Steinkellner *et al*., 1997).

CG sequencing and SSR genotyping: PCR products were cloned into the pCR4-TOPO vector using a TA cloning kit from Invitrogen (Carlsbad, CA, USA). Clones were sequenced using a DYEnamic ET Dye Terminator Kit (Amersham Biosciences Inc., Little Chalfont, Buckinghamshire, UK) on a MegaBACE 1000 automated DNA Sequencer (Amersham Biosciences Inc.). Three different clones of each fragment were sequenced to ensure a high sequencing quality. Microsatellites loci were showed as described in Lepais *et al*. (2006) on a MegaBACE 1000 automated DNA Sequencer (Amersham Biosciences Inc.) using a combination of three different dyes.

SNP detection: The overall sequences per locus were aligned, and polymorphic sites were automatically identified using an informatic pipeline described by Le Dantec *et al*. (2004). Every polymorphic site was then manually verified with CodonCode Aligner v.1.5.1 (CodonCode Corporation, Dedham, MA, USA). In

order to distinguish true polymorphisms from scoring errors, each polymorphic site was visually checked on the chromatograms and further validated for Phred scores (quality threshold above 30).

Diversity and differentiation: For CGs, genetic diversity statistics, including the number of single-nucleotide polymorphisms (SNPs), insertion–deletions (indels), synonymous and replacement mutations, were calculated using DnaSP 4.00.5 software (Rozas *et al*., 2003). Two measures of DNA polymorphism were computed: π, the average number of pairwise nucleotide differences per site in the sample (Nei, 1987), and S, the number of segregating nucleotide sites. These parameters were computed with DnaSP considering SNPs and indels, at three different levels: the whole sequenced region, non-coding regions (including introns, 3′ and 5′-UTRs) and coding regions. The number of haplotypes and the haplotypic diversity (Hd) was also calculated using the DnaSP software (Rozas *et al*., 2003). Tajima's *D*-statistics (Tajima, 1989) and Fu and Li's *D*-statistics and F-statistics (Fu and Li, 1993) were estimated to detect deviations from neutrality using the DnaSP software. Differentiation indexes were estimated, on the whole sequence and at each polymorphic site as F_{ST} (Weir and Cockerham, 1984), using Arlequin ver 3.01 (Excoffier *et al*., 2005). We further tested whether SNPs deviated significantly from neutral expectations using the Bayesian approach for detecting outlier loci (Beaumont and Balding, 2004). The method is based on the posterior distribution of a locus-specific parameter (α_i) related to population differentiation. Population differentiation is modeled as the co-ancestry coefficient by $F_{ST_{ij}}/(1-F_{ST_{ij}})$, where $F_{ST_{ij}}$ is the probability that two randomly chosen alleles of locus i have a common ancestor within that the population j. The model accounts for different F_{ST} values for different subpopulations. Hence, $F_{ST_{ij}}$ is subdivided into a locus-specific effect (α_i), due to mutation or selection, and into a population-specific effect (β_j), such as population sizes and/or migration rates, and γ_{ij} an interaction effect between a population and a locus, for example, local adaptation of a given allele within a given population. The three components are estimated by logistic regression as: $\mathrm{Log}(F_{ST_{ij}}/(1-F_{ST_{ij}}) = \alpha_i + \beta_j + \gamma_{ij}$. Under neutral expectations, α-values are expected to be 0, whereas positive values would indicate directional selection and negative values balancing selection. At $P = 5\%$, α_i is significantly positive (directional selection) if its 2.5% quantile is positive, and is significantly negative (balancing selection) if its 97.5% quantile is negative. Similarly, a significant positive effect of β_j (reduced population size or gene flow) would be detected if 2.5% of its posterior distribution is positive, and negative (very large population size or gene flow) if its 97.5 quantile is negative. Posterior distributions of the parameters were computed with BAYESFST using MCMC simulations (Beaumont and Balding, 2004). The mean and standard deviation of prior distributions were set by default at 0 and 1 for α_i, and -2 and 1.8 for β_j, respectively. The Bayesian detection of outlier loci (Beaumont and Balding, 2004) was preferred to the frequentist approach (Beaumont and Nichols, 1996), as the latter is less robust when sample sizes are low and does not account for population effects on F_{ST} and is limited to the infinite allele mutation (and stepwise mutation model) model that is not appropriate for biallelic loci as SNPs (Eveno *et al*, 2008).

For SSRs, the following diversity and differentiation statistics were assessed: H_O (observed heterozygosity), H_S (mean population diversity), H_T (overall diversity), F_{ST} (differentiation) and F_{IS} (fixation index) (Nei, 1987). H_O, H_S and H_T were estimated using Fstat v2.9.3 software (Goudet, 2001), and F_{ST} and F_{IS} were estimated using Arlequin ver 3.01 (Excoffier *et al*., 2005).

Mapping CGs and detection of QTLs for bud burst

Plant material and assessments of bud burst: QTLs for bud burst were detected within the full-sibling family 3P*A4 comprising 278 full siblings (Scotti-Saintagne *et al*., 2004). Full-sibling seedlings were sown in spring 1995 and raised in a seed bed (field test 1) in the nursery of the INRA Research Station at Pierroton (near Bordeaux France). The full siblings were further vegetatively propagated during two successive campaigns in 1997 (6.2 cuttings per full sibling on average) and 1998 (10.6 cuttings per full sibling), and cuttings were installed in two experimental plantations in the fall of 1998 (field test 2) and spring 2000 (field test 3). Details about the two experimental plantations are given in Scotti-Saintagne *et al*. (2004). Both field tests 2 and 3 are located at the INRA Froot Tree Domain located at Bourran (near the city of Agen in the southwest of France).

Bud burst was recorded in 1999 in the seed bed (field test 1) and over 6 years in each of the two cutting plantations (2000, 2003, 2004, 2005, 2006 and 2007 for field test 2; and 2002, 2003, 2004, 2005, 2006 and 2007 for field test 3). In total 13 observations of bud burst were made over the three test sites. Depending on the number of cuttings available per full sibling, the number of observations per full sibling reached a maximum of 224 for each full sibling of the mapping pedigree. Assessments of bud burst consisted in recording bud development scores (ranging from 0 to 5) every 2–3 days starting when the first tree reached score 1. These data used for the detection of QTLs were the number of days necessary to reach stage 3.

QTL detection and mapping of CGs: Multiple Interval Mapping (Jansen and Stam, 1994; Zeng, 1994) was used to identify QTLs, their position (L in cM) and their contribution to the variance of the trait (percentage of the variation in %) using the MultiQTL V2.5 software package (http://www.multiqtl.com). Confidence intervals at 95% were estimated after 1000 bootstrap resampling (Visscher *et al*., 1996). The position of a QTL was defined as the range between the location of the QTL found with the observed data and the mean of 1000 bootstrap samples. Confidence intervals at 95% were estimated after 1000 bootstrap resampling and were positioned relative to the mean of the bootstrap samples.

The QTL detection was done separately on each parental map in two different steps. The one-QTL (per linkage group (LG)) model was first used and then followed by a model assuming two QTLs. When both models showed the existence of QTLs for a given LG, the two-QTL model was preferred to the one-QTL model when the presence of two QTLs was significant compared with the null hypothesis of no QTL

(H2 > H0) and with the null hypothesis of one QTL (H2 > H1). Otherwise, the one-QTL model was preferred.

To declare the presence of a QTL, type (I) error was set to 0.05 at the genome level. Type I errors obtained at the chromosome level (Doerge and Churchill, 1996) by MultiQTL were transformed at the genome level following the method of Scotti-Saintagne *et al*. (2004). An overall analysis using all 13 observations of bud burst was conducted using the multi-environment model. By bulking the information over all field tests and years, the multi-environment approach allows to improve the QTL detection by decreasing the environmental variance.

CGs were mapped using either SSCP (single-strand conformation polymorphism; Orita *et al*., 1989) or SNP variation (Casasoli *et al*., 2006). A sample of 57–135 full siblings of the mapping pedigree was genotyped for the ESTs of the CGs. MapMaker V.2 software (Lander *et al*., 1987) was used for linkage analysis using a LOD threshold of 6.0 as a grouping criterion and the Kosambi (1944) function to estimate genetic distances.

Results

Nucleotide diversity

For eight out of the nine studied loci, the sequence of 42–52 samples was obtained, except for GA20, where only 21 sequences were obtained. The regions analyzed covered a total number of 4.6 kb, corresponding to 2.7 kb of coding sequence and 1.9 kb of non-coding sequence (introns + 3′-UTR) (Table 2). The fragment sizes ranged from 315 bp for YSL1 to 1104 bp for PM23. A total of 107 SNPs, 18 singletons and 23 indels were detected. In total, 148 mutations (including SNPs, singletons and indels) were detected, suggesting that on average there is one mutation every 31 bp. The indels varied in size from 1 to 28 bp, and were mostly located in intronic regions. The average total nucleotide diversity (6.15×10^{-3}) varied from 1.09×10^{-3} for locus GA3 to 14.7×10^{-3} for locus GALA (Table 3). The level of diversity in non-coding regions (8.15×10^{-3}) is two-fold higher than in coding regions (4.81×10^{-3}). In coding sequences, synonymous regions were more variable than replacement regions (12.67×10^{-3} vs 2.40×10^{-3}), except for locus GA20. Interestingly, the ratio of replacement and silent nucleotide diversity was high for PM23 (0.51) and ASI (0.39), and notably different from the other seven genes. Average haplotypic diversity was 0.76, corresponding on average to 13.2 haplotypes per locus for the sample sizes analyzed (Table 3). With the exception of one locus (QrZag112), total genetic diversity of microsatellites varied between 0.70 and 0.94, and the mean number of alleles between 10 and 20 (data not shown), as previously reported in other genetic surveys (Mariette *et al*., 2002).

Population differentiation

The overall differentiation for CGs and microsatellites was extremely low (Tables 4a and 4b) and highly variable across loci due to the low sample size. There is a slightly higher differentiation for genes, mainly due to DAG2. Differentiation computed among geographical or phenological groups remained at extremely low levels, regardless if it was microsatellites or CGs. We explored in more details the comparison of geographical with phenological differentiation by computing F_{ST} at each single SNP for CGs and at each allele for the microsatellites, and comparing the distribution of F_{ST} values between the two categories of markers and clustering of populations.

The mean value of F_{ST} for SNPs and microsatellite alleles was located at 0 (Figure 1), and their distributions were slightly skewed to positive values. The overall range of distribution was larger in the case of SNPs than in the case of SSR alleles, as a result of the lower overall diversity of SNPs in comparison with SSR alleles. There is no trend of larger differentiation between phenological groups than between geographical groups. However, a detailed analysis of F_{ST} values of SNPs showed that 15 SNPs of GALA (out of 18) exhibited values larger than 0.07 among phenological groups and only 2 SNPs larger values than 0.07 among geographical groups. The Bayesian F_{ST} test did not show any significant SNP (locus-specific effect (α_i)) deviating from neutral expectations among the nine populations. Two different runs of BAYESFST were conducted. This first analysis was based

Table 3 Patterns of nucleotide variation

Locus	*N*	*SNP*												*No. of haplotypes*	*Hd*
		Total				*Non-coding*		*Coding*							
								Total		*Synonymous*		*Replacement*			
		S	π_{total}	π_{silent}	$\pi_{replacement}/\pi_{silent}$	*S*	π	*S*	π	*S*	π	*S*	π		
GALA	44	25	14.7	20.66	0.17	17	17.4	8	11.3	6	39.01	2	3.46	25	0.916
ASI	44	22	12.68	23.8	0.39	0	—	22	12.68	9	23.8	13	9.27	21	0.902
DAG2	42	2	1.56	5.43	0.06	0	—	2	1.56	1	5.43	1	0.34	3	0.528
H3	52	7	3.56	4.65	0.09	4	4.2	3	2.3	2	8.34	1	0.41	12	0.819
PM23	50	44	8.2	9.49	0.51	33	9.3	11	6.2	4	10.45	7	4.84	30	0.959
AUX-REP	49	10	2.93	5.24	0	9	7	1	0.2	1	0.68	0	0	9	0.642
YSL1	48	8	8.24	24.7	0.11	0	—	8	8.24	6	24.7	2	2.73	9	0.867
GA3	46	3	1.09	1.88	0	2	1.9	1	0.4	1	1.65	0	0	5	0.605
GA20	21	4	2.39	5.09	0.11	3	9.1	1	0.4	0	0	1	0.58	5	0.6
Total	396	125	6.15	11.22	0.16	68	8.15	57	4.81	30	12.67	27	2.4	13.22	0.76

Abbreviations: Hd, haplotypic diversity; π, nucleotide diversity ($\times 10^{-3}$); *S*, number of polymorphic sites; SNP, single-nucleotide polymorphism.
Silent sites include synonymous polymorphisms in coding regions and polymorphisms in non-coding regions.

Table 4a Population differentiation of candidate genes (F_{ST} values)

Candidate gene	*Overall F_{ST}*[a]	*Geo F_{ST}*[b]	*Phe F_{ST}*[c]
GALA	0.037	0.033	0.047
ASI	−0.026	**0.044**	−0.033
DAG2	0.180	0.084	−0.078
H3	−0.007	0.029	−0.038
PM23	−0.021	−0.003	0.017
AUX-REP	0.014	0.042	0.020
YSL1	0.062	0.008	0.002
GA3	0.012	0.066	0.010
GA20	0.124	−0.134	0.040
CG mean	0.042	0.019	−0.001

Abbreviation: CG, candidate gene.
The values in bold indicate significant F_{ST} value ($P<0.05$).
[a]Overall F_{ST}, differentiation among all nine populations.
[b]Geo F_{ST}, differentiation between geographical groups of populations (see Table 1 for composition of groups).
[c]Phe F_{ST}, differentiation between phenological groups of populations (see Table 1 for composition of groups).

Table 4b Population differentiation of microsatellites (F_{ST} values)

Locus	*Overall F_{ST}*[a]	*Geo F_{ST}*[b]	*Phe F_{ST}*[c]
QpZAG15	−0.012	0.001	−0.031
QpZAG46	0.013	−0.013	0.045
QpZAG9	0.023	0.018	−0.034
QpZAG110	0.016	−0.008	0.008
QrZAG11	−0.041	0.007	−0.008
QrZAG112	−0.029	**0.068**	0.004
QrZAG39	0.019	0.007	−0.041
QrZAG96	0.055	−0.038	0.013
MsQ13	−0.025	0.013	−0.028
QpZAG36	0.007	−0.001	0.027
QrZAG20	−0.005	0.012	−0.005
QrZAG5	0.030	**0.071**	0.020
QrZAG65	0.013	0.012	0.038
QrZAG7	0.050	0.000	−0.013
QrZAG87	0.039	0.021	0.034
SSR mean	0.013	0.011	0.012

[a]Overall F_{ST}, differentiation among all nine populations.
The values in bold indicate significant F_{ST} value ($P<0.05$).
[b]Geo F_{ST}, differentiation between geographical groups of populations (see Table 1 for composition of groups).
[c]Phe F_{ST}, differentiation between phenological groups of populations (see Table 1 for composition of groups).

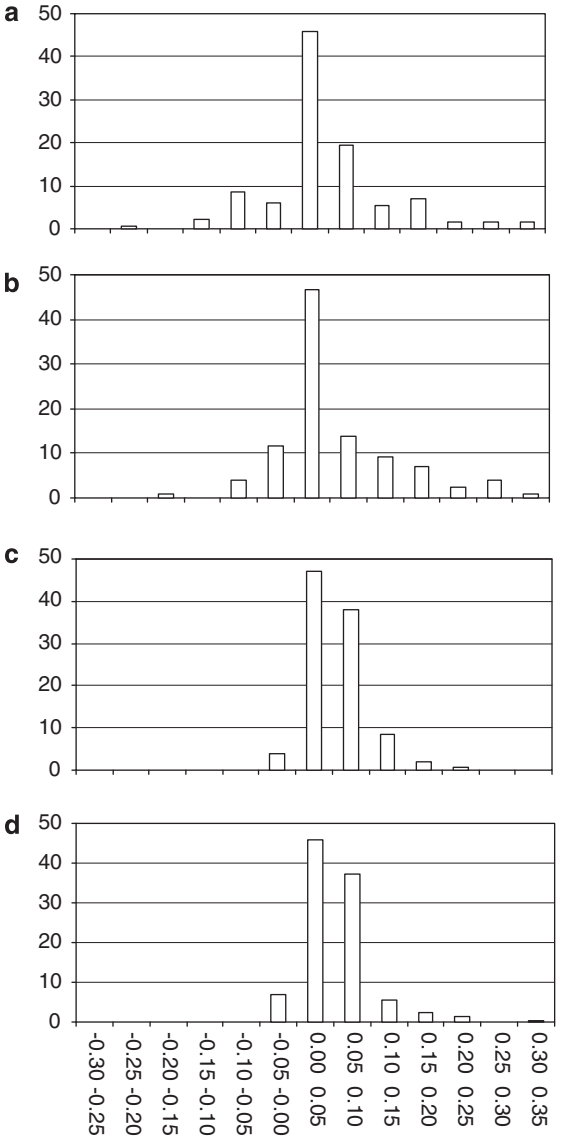

Figure 1 Distribution of F_{ST} values. (**a**) Candidate genes. Differentiation among geographic groups. (**b**) Candidate genes. Differentiation among phenological groups. (**c**) Microsatellites. Differentiation among geographic groups. (**d**) Microsatellites. Differentiation among phenological groups.

on 10 000 drawings from the posterior distribution, and the second on 2000 drawings. Both came to the same conclusion that none of the SNPs did show any significant deviation from neutral expectations. The Bayesian F_{ST} test was then further carried out across the phenological and geographical groups using the same methodology (two independent analysis). No SNP showed either positive or negative significant deviations. However, BAYESFST runs indicated significant negative population-specific effects, for example, for which 97.5 of the posterior distribution was lower then the prior mean −2. Significant negative values of β_i were obtained for all populations except Cochem, Klostermarienberg and Mölln. We further did the computation by changing the prior standard deviation of β_j to 1 and 3, and obtained the same results. None of the analysis did indicate any significant γ-effect (an interaction between a gene and a population).

As the F_{ST} analysis strongly suggested a very low population differentiation, we considered the overall sample of trees as belonging to one single random mating population. Under these hypotheses, we calculated Tajima's *D*-statistic and Fu and Li's *D*-statistics and F-statistics. None of these tests showed any significant deviation from neutral expectations (data not shown). As an example, among the nine genes, four showed positive Tajima's *D*-test, four negative and one close to zero, but none was significant. AUX-REP and ASI were the two genes showing Tajima's *D*-tests closest to significance and both were negative, suggesting a potential directional selection.

QTL mapping of bud burst

The QTL detection resulted in 19 QTLs located on 11 female LGs and as many on 11 male LGs (Table 5). As not all the QTLs are located in homologous regions in the

Table 5 Distribution of QTLs of bud burst

Linkage group	*Number of QTLs*	*LOD score*	*PEV*[a] *min–max*	*QTL1*[b]	*QTL2*[b]
1 F	2	40	1.6–11.8	17.7	38.9
1 M	2	68	0.6–12.3	4.2	47.0
2 F	2	178	4.7–16.1	11.1	43.6
2 M	2	100	0.3–15.0	21.5	101.35
3 F	2	58	0.2–13.3	13.7	23.2
4 F	2	51	0.9–12.0	23.2	40.0
4 M	2	75	1.2–7.9	0	75.6
5 F	2	62	1.0–10.9	23.9	44
5 M	2	101	1.8–15.4	19.0	52.3
6 F	2	36	0.0–6.2	27.2	55.3
6 M	2	51	1.5–21.2	23.7	49.5
7 F	2	64	1.8–10.4	41.3	79.4
7 M	1	18	0–3.4	18.8	
8 M	2	83	2.5–11.0	7.4	31.9
9 F	1	32	0.9–6.7	8.3	
9 M	2	195	1.0–31.1	8.7	25.2
10 F	2	86	1.3–10.5	25.0	70.4
10 M	2	70	0.8–9.7	12.8	35.2
11 F	1	33	0.7–5.3	62.9	—
11 M	1	33	0.2–9.4	54.3	—
12 F	1	13	0–3.0	18.7	—
12 M	1	45	0.5–7.8	0.6	—

Abbreviations: F, female; M, male; LOD; PEV, percentage of variation; QTL, quantitative trait locus.
[a]PEV explained by the QTLs (minimum and maximum value assessed over 13 site*year observations).
[b]Position of QTL (in cM).

male and female LGs, one could conclude that there are more than 19 QTLs controlling bud burst in oaks. Standard deviations of QTL positions were substantially reduced in comparison with the previous analysis (Scotti-Saintagne *et al*., 2004; Casasoli *et al*., 2006), as a result of the multiyear and multisite observations taken into account by the multi-environment option. On average, the standard deviation amounted to 5 cM and reached <0.02 cM in the extreme cases (LG2).

We further monitored the distribution of the contribution of the 38 QTLs across all 13 observations (sites*-years) representing a total of 286 QTL detections, by estimating the percentage of the variation of the mean clonal value explained by the QTL for each observation (Figure 2). The distribution of the percentage of the variation values follows an L-shaped curve. Of the 286 detections, 265 (about 90%) corresponded to QTLs contributing for 12% to the variation of bud burst. The tail of the distribution, for example, the 21 detections with contributions larger than 12%, comprised mainly QTLs located at three LGs (8 detections for LG2, 6 detections for LG9 and 3 detections for LG5). Eight CGs were polymorphic in the two parents of the pedigree and could be mapped on the existing genetic map. Five of them located within the confidence intervals of QTLs for bud burst (using the multi-environment option, Figure 3). Interestingly, these five were located within QTL regions exhibiting the largest contribution to bud burst: LG9 (ASI), LG2 (YSL1) and LG5 (AUX-REP, GA3).

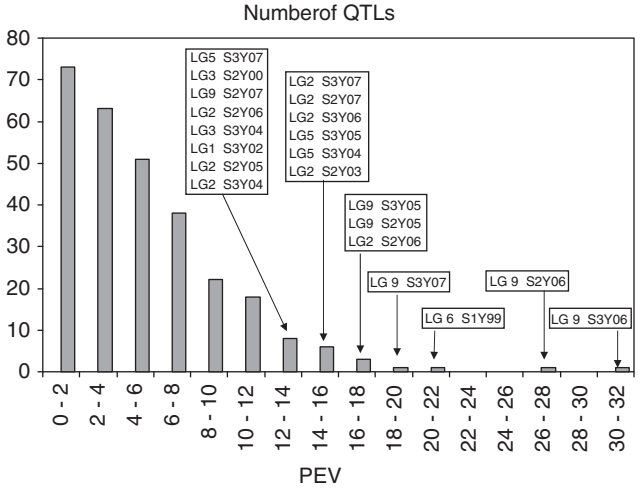

Figure 2 Distribution of the percentage of the variation (PEV) values (percentage of variation explained by the quantitative trait loci (QTLs)) for all QTLs. Tags for QTLs located at the tail of the distribution indicate the linkage groups where the QTL is located and the field Test and Year of the bud burst assessment. For example, LG5 T3Y07 stands for Linkage Group5 (LG5), field Test3 Year 2007.

Discussion

Diversity of CGs of bud burst vs diversity of neutral markers

This study is the second reporting on nucleotide diversity in genes of oaks. Indeed, a recent paper on a Asian oak species (Quang *et al*., 2008), and using genes that were obtained from the same cDNA library than ours (Casasoli *et al*., 2006; Derory *et al*, 2006), made a population survey in 11 genes of *Quercus crispula*. Interestingly, the overall diversity was of similar magnitude in their study ($\pi = 6.93 \times 10^{-3}$) than in ours ($\pi = 6.15 \times 10^{-3}$). These figures are lower than those previously reported on *Populus tremula* (11.1×10^{-3}) (Ingvarsson, 2005), but higher than in pine species, *Pinus taeda* (3.98×10^{-3}) (Brown *et al*., 2004), *Pinus pinaster* (2.41×10^{-3} in Pot *et al*., 2005; 5.51×10^{-3} in Eveno *et al*., 2008) or *Pinus sylvestris* (1.4×10^{-3}) (Dvornyk *et al*., 2002), or other conifers (Gonzalez-Martinez *et al*., 2006; Savolainen and Pyhäjärvi, 2007). Larger diversity in broad-leaves than in conifers is also observed when the comparative analysis is conducted at the level of silent polymorphism: the level of diversity in oaks ($\pi_{silent} = 11.2 \times 10^{-3}$) is higher than earlier reports in pines ($\pi_{silent} = 7.7 \times 10^{-3}$ in *P. sylvestris*, Wachowiak *et al*., 2009; $\pi_{silent} = 8.6 \times 10^{-3}$ in *P. pinaster*, Eveno *et al*., 2008) Although the number of genes is still low, our results confirm earlier findings obtained with other marker systems, suggesting that oak is highly variable species (Kremer and Petit, 1993; Mariette *et al*., 2002). As oaks are outcrossing species exhibiting extensive pollen flow, they compose large populations that contribute to maintain high levels of diversity. However, these results should be taken with caution, as this overall picture overshadows the very large variation observed among genes (15-fold variation from the lower to the largest polymorphic gene) and within genes (between silent and replacement regions). Genes showing the largest diversity were GALA and ASI, and they also showed among the largest ratio of replacement vs silent diversity (Table 3). These two genes are downregulated during the transition between quiescent to developing buds (Derory *et al*., 2006), and their functional role has been investigated in annual plants in germinating seeds, suggesting that similar metabolic pathways may be

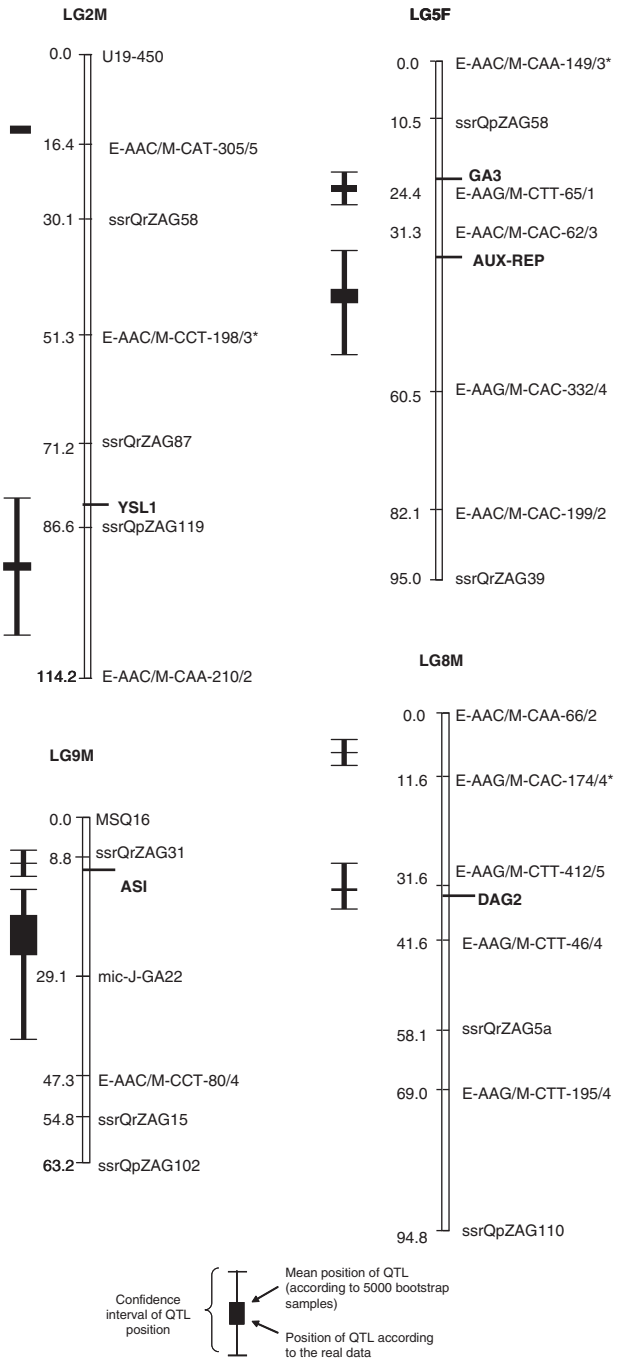

Figure 3 Distribution of quantitative trait loci (QTLs) of bud burst and candidate genes along linkage groups. This figure represents the linkage groups for which QTLs were detected that collocate with at least one of the eight mapped candidate genes. LG2M, linkage group2 male.

active in developing buds and germinating seeds. In *A. thaliana* (Taji *et al*., 2002), galactinol synthase contributes to the accumulation of galactinol under abiotic stress conditions (drought, salinity and cold) and acts as osmoprotectant of the seed. In barley, the expression of ASI in germinating seeds indicates a potential role in defense against pathogens of the developing seed and embryo. The gene is also involved in the inhibition of starch hydrolysis in the peripheral tissues of the seed (Furtado *et al*., 2003). Therefore, both genes appear to have a multifunctional role, which may contribute to the maintenance of a larger diversity due to balancing selection in natural conditions.

Differentiation of CGs vs trait differentiation

Overall population differentiation of expressional and functional CGs of bud burst was of the same level than genomic microsatellites. Furthermore, differentiation among phenological groups was not larger than among geographic groups, and the distributions of the single SNP F_{ST} values for the two subdivisions (phenological vs geographical) overlapped completely. Lastly, F_{ST} values of SNPs did not deviate significantly from neutral expectations based on the Bayesian F_{ST} test. Our results are in line with recent data obtained in various species, where differentiation of CGs was compared with that of neutral markers and of the target trait (in *P. pinaster*, Eveno *et al*., 2008; in *P. sylvestris*, Pyhäjärvi *et al*., 2008; in *Picea abies*, Heuertz *et al*., 2006; and in *P. tremula*, Hall *et al*., 2007; Luquez *et al*., 2007). In all these case studies, the mean differentiation of genes was of the same level than that of the neutral markers and far less than differentiation of the target trait. Recent theoretical developments suggested that this discrepancy may be related to the multi-locus structure of a complex trait. Indeed, Latta (2003) and Le Corre and Kremer (2003) showed that phenotypic differentiation among populations (Q_{ST}) is driven by two major components: covariances among allelic effects at the different loci controlling the trait and variances of allelic effects at each locus. Only the latter component is dependent on F_{ST} of the genes, whereas the former is generated by allelic associations between loci. Interestingly, these authors also showed that the contribution of these components to the phenotypic differentiation of the trait is unbalanced under a wide range of evolutionary scenarios. For tree species, exhibiting extensive gene flow and existing in large populations, differentiation at the trait level is mostly created by allelic associations, rather than changes in allelic frequencies. In other words, diversifying selection, creating population differentiation, is capturing first beneficial allelic associations distributed among the loci contributing to the trait, before modifying allele frequencies at these genes. Hence, strong phenotypic differentiation may coexist with very low differentiation at the genes controlling the trait. Furthermore, the discrepancy is even inflated when the trait is controlled by a large number of loci. Indeed, more loci offer more opportunities for allelic associations to build up. We identified in this study at least 19 QTLs (Table 5) that may be contributing to the trait in only one full-sibling cross. There might be more QTLs in natural populations, suggesting that the high heritability of bud burst in oaks (Scotti-Saintagne *et al*., 2004) may result from the summing of allelic effects over many loci, hence increasing the relative contribution of allelic associations to the overall differentiation of the bud burst. An interesting extension of this study would be to test whether intergenic allelic associations among the CGs were generated by diversifying selection and were responsible for the large phenotypic differentiation that was

observed in the provenance tests as suggested by simulations by Latta (1998) and Le Corre and Kremer (2003). Intergenic allelic associations could be detected by calculating multi-locus F_{ST} taking into account intergenic disequilibria (Kremer *et al*., 1997) but would require larger population samples than those used in this study.

Searching for CGs of bud burst

Our investigations lead to paradoxical conclusions about the link between CGs diversity and bud burst variation. On the one hand, diversity and differentiation statistics show that expressional and functional CGs actually behave as neutral markers, whereas in the same time the collocation of CGs and QTLs was confirmed with stronger confidence in comparison with earlier experiments (Casasoli *et al*., 2006). Mapping positions of five CGs were located within the confidence interval of the three strongest QTLs. There are at least four interpretations to these contrasting results: (1) expressional and functional CGs that were selected for this study are not related to the variation of bud burst; (2) QTL mapping and collocation of CGs is still imprecise; (3) as suggested in the previous paragraph, intergenic allelic associations may be the main component of the phenotypic differentiation of bud burst; (4) the causal mutation of bud burst variation is located outside the genomic region that was explored in this study. With regard to the first and second interpretations, we recently monitored the level of expression of ASI that is located within the confidence interval of the strongest QTL of bud burst within the same mapping population and identified a strong eQTL (expression QTL) located at the same spot, suggesting that the level of expression of ASI is correlated to earliness of flushing (data not published), and reinforcing its putative feature as CGs. Yet, other genes located in the same region may as well be the main source of variation of bud burst. The mean range of the confidence interval of a QTL in this study varies from 2 to 20 cM (Figure 3). Given that the physical size of the oak genome amounts to $\sim$740 Mb/C and the genetic size to 1200 cM (Kremer *et al*., 2007), 1 cM roughly represents 600 000 bp. Hence, the confidence interval of a QTL may comprise from several tens to hundreds genes, among which the gene causing the observed variation of bud burst is. Lastly, we explored nucleotide diversity within a limited region of the different CGs. Although the targeted region derived from EST sequences overlaps the active domain of the corresponding protein (Table 2), no sequence data were available within the promoting regions. As linkage disequilibrium decays quite rapidly in oaks (Quang *et al*., 2008), chances for detecting any causal association locate several hundred base pairs away from the recorded mutation are extremely low.

In conclusion, our results comparing nucleotide diversity of CGs and their collocation with QTL on the genetic maps lead to implement additional investigations in two different directions to confirm their role in bud burst variation. First the exploration of nucleotide diversity needs to be extended across the full length of the genes, including the promoter regions. Second, SNP frequencies should be monitored in a larger set of populations in order to assess for intergenic allelic associations among populations that may account for the overall population differentiation of CGs.

Conflict of interest

The authors declare no conflict of interest

Acknowledgements

We thank Alexis Ducousso for providing the data of the provenance test of *Q. petraea* and his advice for selecting the populations. We acknowledge JM Louvet for his help during the field collections. This work was funded by the European Union (FP5 project TREESNIPS: QLK3-CT-2002-01973).

References

Beaumont MA, Nichols RA (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond Ser Biol Sci* **263**: 1619–1626.

Beaumont MA, Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* **13**: 969–980.

Bechtold N, Ellis J, Pelletier G (1993). In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C R Acad Sci III* **316**: 1194–1199.

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004). Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proc Natl Acad Sci USA* **101**: 15255–15260.

Calvo AP, Nicolas C, Nicolas G, Rodriguez D (2004). Evidence of a cross-talk regulation of a GA 20-oxidase (FsGA20ox1) by gibberellins and ethylene during the breaking of dormancy in *Fagus sylvatica* seeds. *Physiol Plant* **120**: 623–630.

Casasoli M, Derory J, Morera-Dutrey C, Brendel O, Porth I, Guehl JM *et al*. (2006). Comparison of quantitative trait loci for adaptive traits between oak and chestnut based on an expressed sequence tag consensus map. *Genetics* **172**: 533–546.

Chuine I, Beaubien EG 2001)). Phenology is a major determinant of tree species range. *Ecol Lett* **4**: 500–510.

Chuine I, Cour P (1999). Climatic determinants of budburst seasonality in four temperate-zone tree species. *New Phytol* **143**: 339–349.

Derory J, Leger P, Garcia V, Schaeffer J, Hauser MT, Salin F *et al*. (2006). Transcriptome analysis of bud burst in sessile oak (*Quercus petraea*). *New Phytol* **170**: 723–738.

Doerge RW, Churchill GA (1996). Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**: 285–294.

Dow BD, Ashley MV, Howe HF (1995). Characterization of highly variable (GA/CT)*n* microsatellites in the bur oak, *Quercus macrocarpa*. *Theor Appl Genet* **91**: 137–141.

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

Ducousso A, Louvet JM, Jarret P, Kremer A (2005). Geographic variations of sessile oak in French provenance tests. In: Rogers R, Ducousso A, Kanazashi A(eds). *Proceedings of the Joint Meeting of IUFRO working groups genetic of oaks and improvement and silviculture of oaks* FFPRI (Forestry and Forest Products Research Institute) Scientific Meeting Report 3 Tsukuba: Japan. pp 128–138.

Dvornyk V, Sirvio A, Mikkonen M, Savolainen O (2002). Low nucleotide diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. *Mol Biol Evol* **19**: 179–188.

Eveno E, Collada C, Guevara MA, Léger V, Soto A, Díaz L *et al*. (2008). Contrasting patterns of selection at *Pinus pinaster* Ait. Drought stress candidate genes as revealed by genetic differentiation analyses. *Mol Biol Evol* **25**: 417–437.

Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.

Falusi R, Calamassi C (2003). Dormancy of *Fagus sylvatica* L. buds. III. Temperature and hormones in the evolution of dormancy in one-node cuttings. *Plant Biosyst* **137**: 185.

Fu YX, Li WH (1993). Statistical tests of neutrality of mutations. *Genetics* **133**: 693–709.

Furtado A, Henry R, Scott K, Meech S (2003). The promoter of the ASI gene directs expression in the maternal tissues of the seed in transgenic barley. *Plant Mol Biol* **52**: 787–800.

Gonzalez-Martinez SC, Krutovsky K, D. Neale DB (2006). Forest tree population genomics and adaptive evolution. *New Phytol* **170**: 227–238.

Goudet J (2001). Fstat, a program to estimate and test gene diversities and fixation indices. Version 2.9.3.2 (Available at http://www2.unil.ch/popgen/softwares/fstat.htm).

Gualberti G, Papi M, Bellucci L, Ricci L, Bouchez D, Camilleri C *et al.* (2002). Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects the germination of *Arabidopsis* seeds. *Plant Cell* **14**: 1253–1263.

Hall D, Luquez V, Garcia VM, St Onge KR, Jansson S, Ingvarsson PK (2007). Adaptive population differentiation in phenology across a latitudinal gradient in European Aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes, and phenotypic traits. *Evolution* **61**: 2849–2860.

Heuertz M, De Paoli E, Källmann T, Larsson H, Jurman I, Morgante M *et al.* (2006). Multilocus patterns of nucleotide diversity, linkage disequilibrium, and demographic history of Norway Spruce. *Genetics* **174**: 2095–2105.

Ingvarsson PK (2005). Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). *Genetics* **169**: 945–953.

Israelsson M, Mellerowicz E, Chono M, Gullberg J, Moritz T (2004). Cloning and overproduction of gibberellin 3-oxidase in hybrid Aspen trees. Effects on gibberellin homeostasis and development. *Plant Physiol* **135**: 221–230.

Jansen RC, Stam P (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**: 1447–1455.

Kampfer S, Lexer C, Glössl J, Steinkellner H (1998). Characterization of (GA)n microsatellite loci from *Quercus robur*. *Hereditas* **129**: 183–186.

Kosambi DD (1944). The estimation of map distance from recombination values. *Ann Eugen* **12**: 172–175.

Kremer A, Petit RJ (1993). Gene diversity in natural populations of oak species. *Ann Sci Forest* **50**: 186s–203s.

Kremer A, Zanetto A, Ducousso A (1997). Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics* **145**: 1229–1241.

Kremer A, Kleinschmit J, Cottrell J, Cundall EP, Deans JD, Ducousso A *et al.* (2002). Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history and selection on genetic diversity in European oaks? *Forest Ecol Manage* **156**: 75–87.

Kremer A, Le Corre V, Petit R J, Ducousso A (2009). Historical and contemporary dynamics of adaptive differentiation in European oaks. In: DeWoody A, Bickham J, Michler C, Nichols K, Rhodes G, Woeste K (eds). *Molecular Approaches in Natural Resource Conservation*. Cambridge University Press (in press).

Kremer A, Casasoli M, Barreneche T, Bodénès C, Sisco P, Kubisiak T *et al.* (2007). Fagaceae: comparative genetic mapping in Fagaceae. In: Kole CR (ed). *Genome Mapping and Molecular Breeding*, Vol. 5. Forest Trees: Springer-Verlag, Berlin. pp 161–187.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE *et al.* (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**: 174–181.

Latta RG (1998). Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am Nat* **151**: 283–292.

Latta RG (2003). Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytol* **161**: 51–58.

Le Corre V, Kremer A (2003). Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* **164**: 2005–2019.

Le Dantec L, Chagne D, Pot D, Cantin O, Garnier-Gere P, Bedon F *et al.* (2004). Automated SNP detection in expressed sequence tags: statistical considerations and application to Maritime Pine sequences. *Plant Mol Biol* **54**: 461–470.

Lepais O, Léger V, Gerber S (2006). Short note: high throughput microsatellite genotyping in oak species. *Silvae Genet* **55**: 238–240.

Luquez V, Hall D, Albrectsen B, Karlsson J, Ingvarsson PK, Jansson S (2007). Natural phenological variation in Aspen (*Populus tremula*): The Swedish Aspen Collection. *Tree Genet Genomes* **4**: 279–292.

Mariette S, Cottrell J, Csaikl UM, Goikoechea P, Konig A, Lowe AJ *et al.* (2002). Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q-petraea* (Matt.) Liebl. and *Q-robur* L. stands. *Silvae Genet* **51**: 72–79.

Menzel A, Fabian P (1999). Growing season extended in Europe. *Nature* **397**: 659.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University, Press: New York, USA.

Nielsen PK, Bonsager B, Fukuda K, Svensson B (2004). Barley α-amylase/subtilisin inhibitor: structure, biophysics and protein engineering. *Biochim Biophys Acta* **1696**: 157–164.

Or E, Belausov E, Popilevsky I, Ben Tal Y (2000). Changes in endogenous ABA level in relation to the dormancy cycle in grapevines grown in a hot climate. *J Hortic Sci Biotechnol* **75**: 190–194.

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA* **86**: 2766–2770.

Penuelas J, Filella I (2001). Phenology: responses to a warming world. *Science* **294**: 793–795.

Perez-Flores L, Carrari F, Osuna-Fernandez R, Rodriguez MV, Enciso S, Stanelloni R *et al.* (2003). Expression analysis of a GA 20-oxidase in embryos from two sorghum lines with contrasting dormancy: possible participation of this gene in the hormonal control of germination. *J Exp Bot* **54**: 2071–2079.

Pflieger S, Lefebvre V, Causse M (2001). The candidate gene approach in plant genetics: a review. *Mol Breed* **7**: 275–291.

Pot D, McMillan L, Echt C, Le Provost G, Garnier-Gér P, Cato S *et al.* (2005). Nucleotide variation in genes involved in wood formation in two pines species. *New Phytol* **167**: 101–112.

Pukacka S, Wojkiewicz E (2002). Carbohydrate metabolism in Norway maple and sycamore seeds in relation to desiccation tolerance. *J Plant Physiol* **159**: 273–279.

Pyhäjärvi T, García-Gil R, Knürr T, Mikkonen M, Wachowiak W, Savolainen O (2008). Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics* **177**: 1713–1724.

Quang ND, Ikeda S, Harada K (2008). Nucleotide variation in *Quercus crispula* Blume. *Heredity* **101**: 166–174.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984). Ribosomal DNA Spacer-Length Polymorphisms in Barley: Mendelian Inheritance, Chromosomal Location, and Population Dynamics. *Proc Natl Acad Sci USA* **81**: 8014–8018.

Savolainen O, Pyhäjärvi T (2007). Genomic diversity in forest trees. *Curr Opin Plant Biol* **10**: 1–6.
Savolainen O, Pyhäjärvi T, Knürr T (2007). Gene flow and local adaptation in trees. *Annu Rev Ecol Evol Syst* **38**: 595–619.
Scheifinger H, Menzel A, Koch E, Peter C (2003). Trends of spring time frost events and phenological dates in Central Europe. *Theor Appl Climatol* **74**: 41–51.
Scotti-Saintagne C, Bodénès C, Barreneche T, Bertocchi E, Plomion C, Kremer A (2004). Detection of quantitative trait loci controlling bud burst and height growth in *Quercus robur* L. *Theor Appl Genet* **109**: 1648–1659.
Steinkellner H, Lexer C, Turetschek E, Glossl J (1997). Conservation of (GA)n microsatellite loci between *Quercus* species. *Mol Ecol* **6**: 1189–1194.
Tajima F (1989). Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M *et al*. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J* **29**: 416–426.
Tanksley SD, Ganal MW, Martin GB (1995). Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends Genet* **11**: 63–68.
Van Asch M, Visser ME (2007). Phenology of forest caterpillars and their host trees: the importance of synchrony. *Annu Rev Entomol* **52**: 37–55.
Visscher PM, Thompson R, Haley CS (1996). Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**: 1013–1020.
Wachowiak W, Balk PA, Savolainen O (2009). Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Trees Genet Genomes* **5**: 117–132.
Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
Zeng ZB (1994). Precise mapping of quantitative trait loci. *Genetics* **136**: 1457–1468.