Molecular Footprints of Local Adaptation in Two Mediterranean Conifers

Delphine Grivet,[1] Federico Sebastiani,[2,3] Ricardo Alía,[1,4] Thomas Bataillon,[5] Sara Torre,[3] Mario Zabal-Aguirre,[1] Giovanni G. Vendramin,[3] and Santiago C. González-Martínez*[,1,4]

[1]Department of Forest Ecology and Genetics, Center of Forest Research, CIFOR-National Institute for Agriculture and Food Research and Technology (INIA), Madrid, Spain
[2]Department of Agricultural Biotechnology, Genexpress, University of Florence, Sesto Fiorentino, FI, Italy
[3]Plant Genetics Institute, Division of Florence, National Research Council, Sesto Fiorentino, FI, Italy
[4]Sustainable Forest Management Research Institute, University of Valladolid-INIA, Madrid, Spain
[5]Bioinformatics Research Center, Aarhus University, Aarhus C, Denmark

***Corresponding author:** E-mail: santiago@inia.es.

Associate editor: Naoki Takebayashi

Abstract

This study combines neutrality tests and environmental correlations to identify nonneutral patterns of evolution in candidate genes related to drought stress in two closely related Mediterranean conifers, *Pinus pinaster* Ait. and *P. halepensis* Mill. Based on previous studies, we selected twelve amplicons covering six candidate genes that were sequenced in a large sample spanning the full range of these two species. Neutrality tests relatively robust to demography (DHEW compound test and maximum likelihood multilocus Hudson–Kreitman–Aguadé test) were used to detect selection events at different temporal scales. Environmental associations between variation at candidate genes and climatic variables were also examined. These combined approaches detected distinct genes that may be targeted by selection, most of them specific to only one of the two conifers, despite their recent divergence (<10 Ma). An exception was *4-coumarate: CoA ligase*, a gene involved in the production of various important secondary products that appeared to play a role in local adaptation processes of both pines. Another remarkable result was that all significant environmental correlations involved temperature indices, highlighting the importance of this climatic factor as a selective driver on Mediterranean pines. The ability to detect natural selection at the DNA sequence level depends on the nature and the strength of the selection events, on the timescale at which they occurred, and on the sensitivity of the methods to other evolutionary forces that can mimic selection (e.g., demography and population structure). Using complementary approaches can help to capture different aspects of the evolutionary processes that govern molecular variation at both intra- and interspecific levels.

Key words: natural selection, neutrality tests, environmental associations, candidate genes, drought stress, *4cl*.

Introduction

Identifying candidate genes underlying genetic differences for adaptive traits can help to understand how species have adapted to their environment and to predict how they will respond to future climatic changes, which is a special concern in regions such as the Mediterranean Basin, where a substantial decrease in precipitation and a pronounced warming is expected in the near future (Petit et al. 2005; Giorgi and Lionello 2008). The ability to detect the footprints of natural selection in population DNA sequence samples depends on the nature and the strength of the selection events (Nielsen 2005), on the evolutionary scale at which they occur (Zhai et al. 2009) and on the sensitivity of the methods to other evolutionary forces that can mimic selection (e.g., demography, population structure; Biswas and Akey 2006).

Positive selection, which drives the increase in frequency of advantageous mutations, is of particular interest because it underlies local adaptation. Recent or ongoing positive selection events can be detected using methods based on polymorphism within species (Biswas and Akey 2006; Zhai et al. 2009). Recently, powerful methods based on site-frequency spectrum (SFS) and haplotype-frequency spectrum that are relatively insensitive to demography or population structure have been developed (Zeng, Shi, et al. 2007). Detecting selection events over a wider evolutionary scale (e.g., recurrent positive selection or ancient selective sweeps) can be efficiently done using methods contrasting within species polymorphism with among species divergence, as these methods are less sensitive to assumptions regarding recombination or demography. The Hudson–Kreitman–Aguadé (HKA) neutrality test (Hudson et al. 1987) and its extensions (e.g., Wright and Charlesworth 2004) are among the most powerful tests to detect positive selection (Zhai et al. 2009). In addition, the action of natural selection can be reflected in statistical associations between genetic and environmental data (Manel et al. 2003; Hancock et al. 2008; Coop et al. 2009). Several studies have shown that environmental heterogeneity influences the distribution of genetic diversity across plant populations (see Savolainen et al. 2007 for forest trees; Nakazato et al. 2008; Montesinos et al. 2009). Detecting clinal genetic variation along environmental gradients can thus provide evidence for the action of selection (Gram and Sork 2001; Vasemagi and Primmer

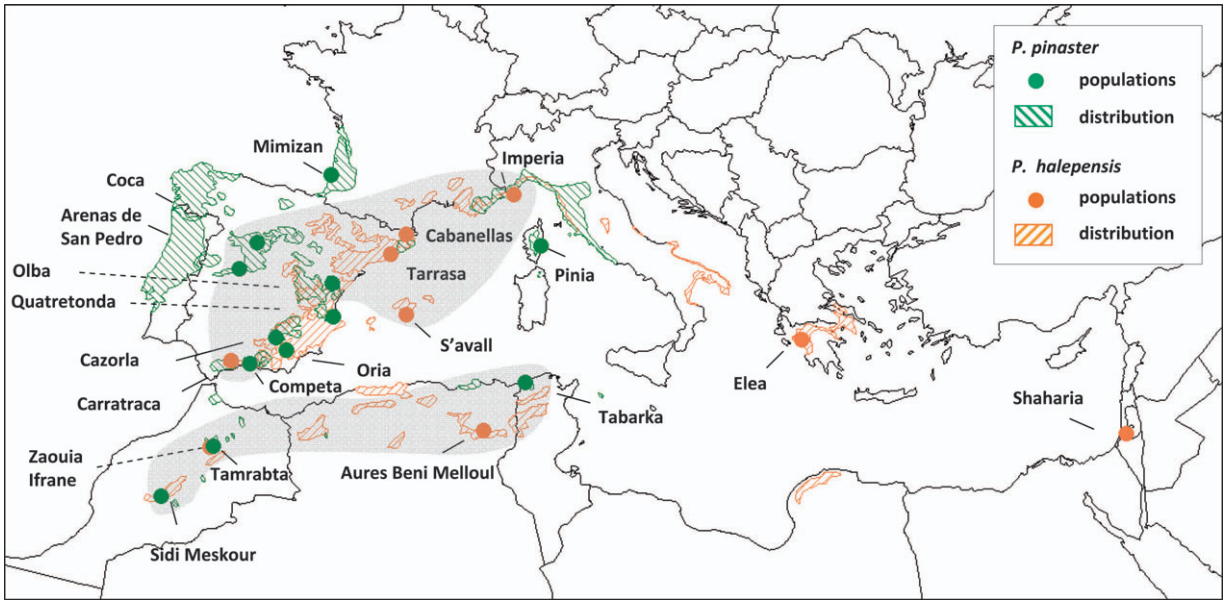

FIG. 1. Location of the 12 *P. pinaster* and the 9 *P. halepensis* populations and the distribution of the two species. Gray areas represent the two groups (western Mediterranean and North African) defined by chloroplast SSRs (see text for details).

2005; Parisod and Christin 2008). Approaches that detect adaptive clinal variation are very attractive because they can directly link environmental gradients with genotypes and phenotypes. However, detecting environmental associations depends on the spatial scale examined and the associations have to be controlled for selectively neutral processes that can also generate clines (e.g., Storz 2002; Hancock et al. 2008; Coop et al. 2009; Keller et al. 2009). Using environmental association approaches in conjunction with neutrality tests helps to capture different aspects of the evolutionary processes that govern molecular variation.

In the present study, we use the combined strategy outlined above to investigate patterns of polymorphism for a set of candidate genes related to drought stress in maritime pine (also known as cluster pine, *Pinus pinaster* Ait.) and in Aleppo pine (*P. halepensis* Mill.). *P. pinaster* is an economically important pine native to the occidental Mediterranean Basin and the European Atlantic front (fig. 1). This species is ecologically versatile, growing in a variety of substrates, in a wide range of elevations and under a range of Mediterranean and Atlantic climate regimes (semiarid to very humid; table 1). *P. halepensis*, a pine closely related to maritime pine, is one of the most abundant and widespread pine species of the Mediterranean Basin (fig. 1). This species, which also has a wide ecological breadth, is well adapted to dry conditions as well as high-intensity fire regimes (Tapias et al. 2004), which makes it a species of choice for afforestation in xeric Mediterranean areas. These two conifers, because they have extensive and partially overlapping distributions across the Mediterranean Basin, may possibly display adaptation to similar sets of contrasted environmental conditions, that is, regional climate types and spatial variability. Among the key environmental factors to *Pinus* distribution in Mediterranean climates, temperature, and rainfall have been shown to constitute good biological descriptors (Richardson and Bond 1991; Soto et al. 2010). Additionally, several common garden studies have documented differences in drought tolerance and/or gene expression among Mediterranean pine provenances (for *P. pinaster*, see Nguyen and Lamant 1989; Costa et al. 1998; Chambel et al. 2007; Aranda et al. 2009 and for *P. halepensis*, see Atzmon et al. 2004; Sathyan et al. 2005; Voltas et al. 2008). Overall, these studies point to Mediterranean pines as an interesting system to study local adaptation mediated by abiotic stress response. In that context, studying allelic (single nucleotide polymorphism [SNP] allele and haplotype) frequency in drought-related candidate genes along temperature or precipitation clines can help to improve our understanding of pine adaptation to environmental heterogeneity.

Although *P. pinaster* and *P. halepensis* share various ecological requirements and may have responded to their environment in a similar manner, they present quite contrasted biogeographic and demographic histories that can also have a strong effect on the pattern of variation of the candidate genes examined. *P. pinaster* has a long presence in the western Mediterranean Basin surviving the last glaciations in multiple refugia located in southeastern Spain, northern Africa, and the Atlantic coast of Portugal (Bucci et al. 2007). In contrast, *P. halepensis* would have undergone long-range colonization of the western Mediterranean from ancient populations located in Greece and Turkey (Bucci et al. 1998; Morgante et al. 1998; Grivet et al. 2009). In addition, the dynamics of *P. halepensis* populations throughout the Mediterranean Basin highly depend on forest fires. This species-specific information on population dynamics has to be taken into account when looking at the molecular footprint of selection.

Table 1. Spatial Coordinates and Climatic Variables for *Pinus pinaster* and *P. halepensis* Populations.

		Spatial Variables			Climatic Variables							
Population	Country	Altitude (m)	Latitude (degrees)	Longitude (degrees)	AMT (°C)	TS	MTWM (°C)	MTCM (°C)	AP (mm)	PWM (mm)	PDM (mm)	PS
P. pinaster												
Arenas de San Pedro	Spain	733	40.194822	−5.116213	14.2	668.9	33.4	1.4	1,318	199	12	60.8
Cazorla	Spain	1,100	37.919675	−2.925765	11.5	650.2	30.6	−1.4	1,257	179	11	58.3
Coca	Spain	800	41.254705	−4.497827	12.3	655.9	31.2	−0.6	454	55	15	30.0
Cómpeta	Spain	903	36.834265	−3.953989	14.0	512.1	27.5	4.7	899	132	5	64.1
Olba	Spain	1,002	40.173309	−0.622966	12.4	600.7	28.3	0.6	509	63	24	31.5
Oria	Spain	1,223	37.531165	−2.351138	13.1	633.1	30.6	0.4	357	46	5	46.2
Quatretonda	Spain	435	38.971645	−0.358844	15.3	553.7	30.3	3.8	777	120	9	52.1
Mimizan	France	19	44.134167	−1.303056	13.2	490.9	24.9	3.1	1,235	149	62	24.3
Pinia	France	15	42.021083	9.464861	15.6	515.5	27.0	6.3	585	80	8	49.6
Tabarka	Tunisia	121	36.958397	8.703792	17.7	573.4	31.8	6.7	916	162	4	72.3
Tamrabta	Morocco	1,758	33.600000	−5.016667	11.8	665.3	31.3	−3.2	721	96	8	54.4
Sidi Meskour	Morocco	1,931	31.439375	−6.903864	11.4	711.3	32.5	−5.2	514	71	5	56.4
P. halepensis												
Cabanellas	Spain	210	42.235556	2.790000	14.8	575.5	27.4	3.9	713	97	35	26.0
Carratraca	Spain	650	36.841111	−4.834444	15.7	568.6	30.6	4.4	693	106	3	69.4
S'avall	Spain	10	39.287222	3.047778	16.8	525.9	28.7	6.8	566	92	5	54.7
Tarrasa	Spain	117	41.466667	2.100000	15.9	551.1	27.7	5.8	619	83	29	29.5
Imperia	Italy	109	43.900000	8.050000	15.1	550.5	27.5	4.6	804	107	18	38.5
Zaouia Ifrane	Morocco	1,512	33.570000	−5.140000	11.5	668.7	31.2	−3.8	849	123	8	57.4
Aures Beni Melloul	Algeria	936	35.166667	6.833333	13.4	657.2	31.2	0.1	375	53	10	36.1
Elea	Greece	155	37.766667	21.533333	16.7	592.0	30.8	5.4	808	150	6	77.1
Shaharia	Israel	236	31.600000	34.833333	20.0	496.4	31.8	8.0	397	100	0	116.0

NOTE.—AMT, annual mean temperature; TS, temperature seasonality (SD * 100); MTWM, maximum temperature of the warmest month; MTCM, minimum temperature of the coldest month; AP, annual precipitation; PWM, precipitation of the wettest month; PDM, precipitation of the driest month; PS, precipitation seasonality (coefficient of variation * 100).

Based on previous population genomic studies identifying drought tolerance genes potentially under selection (see Eveno et al. 2008; Grivet et al. 2009) as well as on gene expression studies identifying genes affected by drought stress (Watkinson et al. 2003; Rani et al. 2009), we examine here patterns of nucleotide variation in an ortholog set of targeted candidate genes in *P. pinaster* and *P. halepensis*. We sample a broad geographic range of populations to: 1) provide new evidence about positive selection acting on these genes, 2) identify the timescale at which selection events may have occurred, and 3) examine which environmental factors underlie molecular signatures of selection in both species. Our pluralistic approach provides insights on the adaptive strategy of two conifers that live under the same Mediterranean climate but present distinct demographic (re)colonization and life histories.

Materials and Methods

Study Species

Pinus pinaster Ait

Maritime pine populations have been assessed using various molecular markers: chloroplast and mitochondrial (e.g., Burban and Petit 2003; Bucci et al. 2007), as well as nuclear (e.g., Salvador et al. 2000; Eveno et al. 2008). In particular, chloroplast markers were able to identify different gene clusters related to the history of the species (Bucci et al. 2007). Because of its economic importance, various programs of genetic improvement have been developed in *P. pinaster*, in particular for the Atlantic provenances. Within this framework, several adaptive traits, such as growth, tolerance to drought and cold, and resistance to pests and diseases, have been the subject of genetic variability studies. Numerous approaches have been used to unravel the basis of quantitative traits in maritime pine: genetic (genetic cartography), physiologic (mechanisms implicated in traits), functional and structural genomic (candidate genes and proteins), and population genetic and genomic (genes under selection).

Pinus halepensis Mill

Aleppo pine genetic variability has been studied with both chloroplast and nuclear markers (Bucci et al. 1998; Morgante et al. 1998; Grivet et al. 2009). A recent study revealed that the pattern of polymorphism observed in some candidate genes related to drought tolerance in this species reflected long-range colonization and possibly natural selection during range expansion (Grivet et al. 2009). Various common garden studies have assessed Aleppo pine intraspecific variability in order to study the role of ecological factors in shaping adaptive strategies. Some experiments revealed adaptive variation to climate (total precipitation and dry season duration; Voltas et al. 2008) or reproductive features (Climent et al. 2008), highlighting thus the selective role of climate variables in determining population and family fitness in this species.

Sampling

Twelve populations of *P. pinaster* (77–122 individuals, depending on the gene; supplementary table S1, Supplementary Material online) and nine populations of *P. halepensis* (72–93 individuals; supplementary table S1, Supplementary Material online) were collected spanning the full range of each species (fig. 1 and table 1). Populations were selected considering not only spatial distribution but also environmental heterogeneity in both species, prioritizing populations that represent contrasted environments (table 1). Our sampling covers also different soil (siliceous, calcareous) and Mediterranean forest types. Finally, we included representations of the five traditional varieties or landraces described in maritime pine (Resch 1974): *mesogeensis* (e.g., Olba), *atlantica* (e.g., Mimizan), *corteensis* (e.g., Pinia), *maghrebiana* (e.g., Tamrabta), and *renoui* (e.g., Tabarka), as well as all genecological groups normally considered in Aleppo pine: Western Europe, Eastern Europe, and Northern Africa.

From each population, cones were collected from mother trees separated by at least 50 m without any phenotypic selection. Seeds from each mother tree were kept in individual paper bags and stored at 4 °C in a dry environment till DNA extraction (see below).

Spatial, Environmental, and Genetic Data

Spatial and Environmental Data

Three spatial variables were recorded for each population: altitude, latitude, and longitude (table 1). Eight climatic variables were considered: annual mean temperature (AMT), temperature seasonality (expressed as the standard deviation across months multiplied by 100; TS), maximum temperature of the warmest month (MTWM), minimum temperature of the coldest month (MTCM), annual precipitation (AP), precipitation of wettest month (PWM), precipitation of driest month (PDM), and precipitation seasonality (coefficient of variation multiplied by 100;PS). Climatic data for *P. pinaster*'s Iberian populations were obtained from a functional phytoclimatic model based on raw data from meteorological stations (Gonzalo 2007). Climatic data for *P. pinaster*'s non-Iberian populations as well as for all *P. halepensis*' populations were obtained from the WorldClim—Global Climate Data at 5 min resolution (Hijmans et al. 2005) (table 1). Graphical pairwise correlations between these 11 spatial and environmental variables are presented in supplementary figure S1, Supplementary Material online.

DNA Extraction and Candidate Gene Sequencing

Genomic DNA from *P. pinaster* (haploid) megagametophytes was extracted with a modified (Dellaporta et al. 1983) protocol. DNA extraction for *P. halepensis* was carried out as reported in Grivet et al. (2009).

Candidate genes related to drought tolerance were originally identified on the basis of functional studies performed in *P. taeda* and other conifers or derived from model species such as *Arabidopsis thaliana* as described elsewhere (González-Martínez et al. 2006; Eveno et al. 2008; Grivet et al. 2009; Wachowiak et al. 2009). Altogether the candidate genes selected for this study belong to three well known and relatively small multigene families: the ASR (*lp31-Pt* and *lp33-Pt*), dehydrin (*dhn2-Pp*, *dhn2-Ps*, and *dhn5-Ps*), and *4-coumarate: CoA ligase* (*4cl-Pt*) families. The ASR family is named after the *A* bscisic acid, *s* tress, and *r* ipening response (Frankel et al. 2006). These proteins are also induced by water deficit stress. The dehydrin gene, *dhn2-Pp*, previously described in maritime pine (Eveno et al. 2008), is not orthologous to the one described in Scots pine, that is, *dhn2-Ps* (Wachowiak et al. 2009).

Previously published primer pairs of these putative candidate genes were transferred from *P. tadea* (*Pt*), *P. sylvestris* (*Ps*), or *P. pinaster* (*Pp*) to either *P. pinaster* or *P. halepensis* or both, and three new primer pairs, which overlap with predescribed amplicons, were also designed in order to extend the coverage of these target genes (supplementary table S2, Supplementary Material online). Polymerase chain reaction (PCR) conditions for the two pines studied here are given in supplementary table S3, Supplementary Material online. Outgroup sequences for each amplicon were obtained from GenBank: *P. taeda* for amplicons *lp31-Pt*, *lp33-Pt*, *dhn2-Pp*, and *4cl-Pt*, and *P. sylvestris* for *dhn5-Ps* and one *dhn2-Ps* amplicon; or produced with our newly designed primers (*P. nigra* for the second *dhn2-Ps* amplicon) (supplementary table S2, Supplementary Material online). Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession numbers HM481479–HM483364.

PCR products were purified and sequenced from both ends following standard protocols (for *P. pinaster*, see Eveno et al. 2008; for *P. halepensis*, see Grivet et al. 2009). Multiple sequence alignments and manual adjustments were done using SeqMan v7 (DNASTAR Lasergene software) and BIOEDIT (http://www.mbio.ncsu.edu/BioEdit/page2.html).

Chloroplast Microsatellites

To control for associations between candidate genes and environmental data that could be due to neutral processes, we included variation of chloroplast microsatellites (the only neutral genetic markers available for the two species at the wide range scale) as a covariate in the multivariate logistic regressions (see Environmental Associations). We performed population-based principal component analysis on chloroplast markers and kept the principal components (PCs) that explained the majority of the overall inertia of the data. For *P. pinaster*, the three first PCs of the 16 most common haplotypes (determined on the basis of five chloroplast microsatellites)—accounting for 74% of the overall variation—were extracted from the lattice data set (430 records × 16 variables) of Bucci et al. (2007). For populations that did not match any of the 430 locations in this lattice, the mean of the closest 3–4 locations was computed (supplementary table S4, Supplementary Material online). For *P. halepensis*, PCs were computed in *R* (R Development Core Team) for each of the nine populations based on the 22 haplotypes determined from three chloroplast microsatellites (our unpublished data) and the three first PCs, which explain 82% of the overall variation, were used for

subsequent analysis (supplementary table S4, Supplementary Material online).

Chloroplast microsatellites were also used to determine groups with similar evolutionary history to conduct neutrality tests (fig. 1). A previous study based on chloroplast microsatellites pointed at eight gene pools in *P. pinaster* (Bucci et al. 2007). Within the sampling used in the present study, Mimizan (Continental French lineage) and Pinia (Northern Italy and Corsican lineages) represent single lineages. Tabarka (Tunisia), Tamrabta (Moroccan Middle Atlas), and Sidi Meskour (Moroccan High Atlas) are all part of a North African lineage that combines western and eastern African origins. The rest of *P. pinaster* populations sampled are part of a wide western Mediterranean linage, except for Quatretonda and Oria that are considered marginal populations and may thus present a distinct evolutionary history (González-Martínez, Gómez, et al. 2007; Eveno et al. 2008). In *P. halepensis*, eastern Mediterranean populations (Elea in Greece and Shaharia in Israel) present a higher level of cpSSR diversity and substantial genetic differentiation from the rest and are considered single population lineages (Grivet et al. 2009); Zaouia Ifrane (Morocco) and Aures Beni Melloul (Algeria) form part of the North African group, whereas all four Spanish populations and the Italian population of Imperia belong to the western Mediterranean group. These regions (North Africa and western Mediterranean) define relatively homogeneous zones of the Mediterranean Basin in terms of soil and climate (Barbéro et al. 1998).

Statistical Analyses

Gene Diversity and Divergence

Number of segregating sites (S), nucleotide diversity statistics (θ_π, Tajima 1989; θ_w, Watterson 1975), number of haplotypes (K), and haplotypic diversity (H_e) were computed for both North African and western Mediterranean groups using scripts kindly provided by S.E. Ramos-Onsins (Department of Genetics, Faculty of Biology, University of Barcelona, Spain) and the program DnaSP v5 (Librado and Rozas 2009). Average divergence per site (K_{all}) for each geographical group was computed between each of the pines and the outgroup *P. taeda*, as well as between *P. pinaster* and *P. halepensis* themselves, using scripts also provided by S.E. Ramos-Onsins.

Neutrality Tests

All neutrality tests were performed at the regional level (see group definition above) considering all sequenced gametes in each amplicon, using one sequence from another pine species as outgroup when appropriate, and considering all substitutions as well as only silent sites (except for the McDonald–Kreitman [MK] test where all sites must be used, see below). First, we conducted neutrality tests based on within-species population genetic data with (to determine ancestral states) or without outgroup, including Tajima's D test (Tajima 1989), Fay and Wu' H normalized test (Zeng et al. 2006), both based on the SFS, Ewens–Watterson (EW) homozygosity test (Watterson 1978), and the DHEW neutrality test (Zeng, Shi, et al. 2007)—a compound test that combines the properties of D, H normalized, and EW tests—using scripts provided by K. Zeng (State Key Laboratory of Biocontrol and Key Laboratory of Gene Engineering of the Ministry of Education, Sun Yat-sen University, Guangzhou, China). Second, we used methods that compare patterns of polymorphism within species and divergence between species. The MK test, which compares the ratio of nonsynonymous and synonymous mutations between and within species (McDonald and Kreitman 1991), was computed using DnaSP v5 (Librado and Rozas 2009). Maximum likelihood multilocus Hudson–Kreitman–Aguadé (ML-HKA) neutrality tests were computed using the mlHKA program (Wright and Charlesworth 2004). This test is an extension of the HKA test that allows to test specific locus or groups of loci against a group of neutrally evolving genes for their levels of polymorphism and divergence. First, each amplicon was compared with all the rest and the selection parameter K was computed. The K parameter measures the degree to which diversity is increased or decreased with respect to divergence. Second, for significant tests, the analysis was repeated considering only amplicons from different genes in order to avoid potential linkage. All runs consisted in 100,000 cycles of the Markov chain; long runs of 1,000,000 cycles were also computed for significant tests as well as different starting values for the divergence time parameter (T). However, no differences were found in the results and only basic runs are presented here. Standard likelihood ratio tests (LRTs) were used to compare among models considering one amplicon under selection at a time and the null model of no selection (degrees of freedom equal one). The ML-HKA test was performed using different outgroups to gain insights on evolutionary times at which positive selective wave(s) may have taken place: 1) with *P. taeda* as outgroup for both *P. pinaster* and *P. halepensis* amplicons. In this case, the time frame would correspond to the split between the *Trifoliae* section (to which *P. taeda* belongs) and the *Pinus* section (to which the Mediterranean pines belong) in the Oligocene (~25 Ma; Gernandt et al. 2005, 2008). Three amplicons were discarded from these analyses because they had *P. sylvestris* or *P. nigra* as outgroups (two from *dhn2-Ps* and one from *dhn5-Ps*); 2) with *P. halepensis* as outgroup for *P. pinaster* amplicons and *P. pinaster* as outgroup for *P. halepensis* amplicons. The time frame, in this case, would correspond to the diversification of Mediterranean piness (in the Miocene, ~10 Ma; Gernandt et al. 2008). One amplicon was discarded from this series of tests (*dhn2-Ps_b*; see supplementary table S2, Supplementary Material online for amplicon nomenclature) due to bad alignment between species and, thus, the possible amplification of a paralogous fragment in *P. halepensis* (*dhn2-Ps_b* from *P. halepensis* had more similarity with the outgroup, *P. nigra*, than with its close relative *P. pinaster*).

Environmental Associations

We tested if SNP allele or haplotype frequencies at candidate loci correlated with climatic variables using logistic regression. We first carried out series of univariate logistic regressions to test for association between SNP/haplotype

frequencies and environmental variables using the program SAM (Joost et al. 2008). We considered a correlation as significant only when two LRTs (*G* and Wald tests) rejected the null hypothesis of no association between the genetic and the environmental variables (at the 5% level). A strict Bonferroni correction was applied to correct for multiple testing of univariate models.

In a second step, whenever a SNP/haplotype was found significantly correlated with an environmental variable, we performed a multivariate logistic regression using neutral marker PCs (see above) as covariates (together with the SNP and the environmental variables) to control for neutral processes (e.g., postglacial migrations and geographical isolation) that may have also generated clines in the absence of any local adaptation. Multiple logistic regressions were performed in *R* (R Development Core Team) using the glm function (assuming that allelic counts at SNPs were binomially distributed) for each SNP/haplotype separately. Model selection was carried out using the drop1 function and LRTs. Whenever the models compared were not nested (and LRTs were thus not appropriate), we used Akaike's Information Criterium. Tests for environmental associations were performed at the amplicon level (counting the SNPs located within overlapping regions only once) and including only populations with more than three gametes sampled (average number of gametes sampled per population of 7.98 in *P. pinaster* and 9.27 in *P. halepensis*, see supplementary table S1, Supplementary Material online).

Results

Gene Diversity and Divergence

Overall primer transfer rates across the two pine species were high (91.6% for *P. pinaster* and 83.3% for *P. halepensis*) and we obtained a total of 533 and 380 kbp of sequence in *P. pinaster* and *P. halepensis*, respectively. All sequences represented putative orthologs (determined on the basis of similarities with available sequences in GenBank and construction of phylogenies including multiple gene members of the target families in different conifers) except probably for *dhn2-Ps_b* in *P. halepensis* (see Methods). Maritime pine showed a higher level of gene diversity than Aleppo pine (table 2), and this result held when considering only the nine orthologous amplicons in the two species (data not shown). In total, 131 SNPs (24 nonsynonymous) were detected in *P. pinaster* (4–27 per amplicon) and 65 SNPs (12 nonsynonymous) in *P. halepensis* (1–21 per amplicon) (table 2). There were only four shared polymorphisms (i.e., 2% of polymorphic sites were shared) across species, with two of them (those within *4cl-Pt*) appearing only in one *P. pinaster* individual. Number of insertions/deletions was similar between the two species for the same locus, but these indels were not shared across species. None of them caused frame shifts. At the haplotype level, not only the number of haplotypes substantially differed between the two species (86 for *P. pinaster* vs. only 59 for *P. halepensis*) but also their frequency spectrum: *P. pinaster* had multiple common haplotypes along with some less frequent ones, whereas *P. halepensis* generally displayed one major haplotype together with some low-frequency ones (supplementary fig. S2, Supplementary Material online). Finally, there was also a contrasted pattern of gene diversity between the regional groups defined by chloroplast microsatellites in *P. pinaster*: The western Mediterranean group displayed higher nucleotide diversity than the North African group (average θ_w per population of 0.0039 vs. 0.0022). The SNP and haplotypes detected were used to conduct neutrality tests as well as to detect associations with environmental variables.

Divergence at the DNA sequence level was examined using two outgroups: *P. taeda*, a New World pine, and each of the Mediterranean pines as reference for the other (i.e., *P. halepensis* was used as outgroup for *P. pinaster* and vice versa). For both species, nucleotide divergence per site with *P. taeda* (K_{all} = 0.0397 for *P. pinaster*; K_{all} = 0.0498 for *P. halepensis*) was higher than with the other Mediterranean species (K_{all} = 0.0235 for *P. pinaster* vs. *P. halepensis*) except for *dhn2-Pp_b* in *P. pinaster* that displayed a slightly higher divergence between the Mediterranean pines than between each of them with *P. taeda*.

Neutrality Tests

Neutrality tests rejected the null neutral model for distinct genes depending on the species and the geographic group (tables 3 and 4). The *D* and H_{norm} tests identified some genes departing significantly from expectation under the neutral model that are not detected by the more robust DHEW test: *lp31-Pt_a* and *lp31-Pt_b* in the North African group for *P. pinaster*; *lp33-Pp* in the western Mediterranean group for *P. halepensis*. As shown below, these two tests are probably reflecting demographical events and not selective processes. More relevantly, two fragments of the *4cl-Pt* gene were identified as potential targets of selection by DHEW tests: *4cl-Pt_b* in the western Mediterranean group for *P. pinaster* and *4cl-Pt_d* in the North African group for *P. halepensis*. Considering all polymorphic sites or only silent sites did not change the qualitative results of the test (tables 3 and 4 and supplementary table S5, Supplementary Material online).

None of the MK tests showed any departure from the null hypothesis of an equal ratio of nonsynonymous to synonymous variation within and between species (data not shown). The ML-HKA tests based on silent sites (with *P. taeda* as outgroup) revealed uncoupled patterns of polymorphism and divergence in only one amplicon in *P. pinaster* in the western Mediterranean group (*dhn2-Pp_b*), whereas in *P. halepensis*, one amplicon in this same group (*lp33-Pp*) and one in the North African group (*4cl-Pt_b*) were significant (tables 3 and 4). These results still stood when analyses were restricted to only one amplicon per gene (in order to avoid potential linkage between fragments from the same gene) and when fragments potentially under selection were not used as control loci (data not shown). When all substitution sites were included in the analyses, three more genes were significant for *P. halepensis* in the western Mediterranean group (*dhn2-Ps_a*, *4cl-Pt_a*, and *4cl-Pt_b*) (supplementary table S5, Supplementary Material online). In contrast,

Table 2. Gene Diversity (all sites) for 12 Amplicons from Six Putative Candidate Genes across the 12 *Pinus pinaster* Populations and the nine *P. halepensis*.

	P. pinaster							*P. halepensis*						
Amplicon	*N*	*L*	*S*	θ_π	θ_w	*K*	H_e	*N*	*L*	*S*	θ_π	θ_w	*K*	H_e
lp31-Pt														
a	122	456	12	6.26	4.89	6	0.717	79	353	3	2.48	1.72	4	0.460
b	111	560	7	3.91	2.37	3	0.446	74	488	7	3.93	2.94	4	0.464
lp33-Pp	97	449	9	4.29	3.89	9	0.798	90	375	1	0.17	0.53	2	0.065
dhn2-Pp														
a	120	472	12	6.06	4.74	11	0.807	89	448	2	0.20	0.88	3	0.088
b	85	596	21	8.42	7.03	11	0.811	93	346	4	0.55	2.26	5	0.105
dhn2-Ps														
a	92	743	27	8.10	7.13	14	0.899	na	na	na	na	na	na	na
b	80	513	15	5.92	5.90	14	0.855	92	457	11	1.93	4.73	6	0.204
dhn5-Ps	77	449	4	4.00	1.81	5	0.735	na	na	na	na	na	na	na
4cl-Pt														
a	91	551	13	5.67	4.64	4	0.678	88	461	10	5.88	4.30	13	0.805
b	92	256	7	3.56	5.37	5	0.291	79	688	21	10.67	6.18	13	0.793
c	87	543	4	0.69	1.46	4	0.341	78	528	3	2.47	1.15	5	0.697
d	na	na	na	na	na	na	na	72	452	3	2.80	1.37	4	0.484
Total	1,054	5,588	131	—	—	86	—	834	4,596	65	—	—	59	
Mean	—	—	—	5.17	4.48	—	0.671	—	—	—	3.11	2.61	—	0.416

NOTE.—*N*, number of sequences; *L*, total analyzed length in base pair; *S*, polymorphic sites; θ_π, Tajima's nucleotide diversity $\times 10^{-3}$ (Tajima 1989), and θ_w: Watterson's nucleotide diversity (Watterson 1975) per site ($\times 10^{-3}$); *K*, number of haplotypes; H_e, haplotype diversity; na, amplicons that failed to transfer across species.

when Mediterranean pines were used as outgroup for each other, patterns of polymorphism and divergence either considering all sites (supplementary table S5, Supplementary Material online) or silent sites (tables 3 and 4) seemed to evolve neutrally and a similar trend of the maximum likelihood estimate of the selection parameter (ML-HKA's *K*) was observed in the two series of analyses.

Environmental Associations

Environmental associations were examined at both the SNP and the haplotype level, and, overall, similar associations were detected. Nonetheless, a few correlations differed, with some found only at the haplotype level, which highlights how these two levels complement each other as haplotypes may reflect interactions among linked mutations. Within the two species, all significant associations detected between SNPs or haplotypes and climatic variables involved temperature indices as environmental variables.

As many as 23 significant associations were initially found in *P. pinaster* (noncorrected model as provided by SAM; see Methods). Only three associations remained after integrating neutral marker PCs as covariates, two of them with spatial variables (one with altitude and one with latitude) and one (*4cl-Pt_c_*229) with the MTCM (fig. 2; supplementary table S6, Supplementary Material online). At the haplotype level, initially 32 associations were found in maritime pine but only six remained in the corrected model: two (*4cl-Pt_c_*A and *4cl-Pt_c_*B) with altitude and one (*dhn5-Ps_*C) with AMT, one (*dhn5-Ps_*E) with MTWM, and two (*4cl-Pt_c_*A and *4cl-Pt_c_*B) with MTCM (fig. 2; supplementary table S7, Supplementary Material online). In the case of *P. halepensis*, only one association was found by the corrected models, both at the SNP (between *4cl-Pt_c_*131 and TS) and at the haplotype (*4cl-Pt_c_*B and TS) level (fig. 2; supplementary tables S6 and S7, Supplementary Material online).

In summary, association analysis identified two loci exhibiting significant correlations with temperature indices: *4cl-Pt_c* that is associated with MTCM in *P. pinaster* and with TS in *P. halepensis* at both the SNP and the haplotype levels, and *dhn5-Ps* that is associated with AMT and MTWM at the haplotype level in *P. pinaster* only.

Discussion

In this study, we assessed the impact of natural selection on the same set of candidate genes related to drought tolerance in two widespread Mediterranean pine species. Our results revealed distinct selection patterns according to species, geographic regions, and loci. Below, we discuss these findings in the light of the history of each species and the specificities of each of the methods used to reveal footprints of selection.

Neutrality Tests

Neutrality tests examining selection events based on SFS and haplotype-frequency spectrum identified distinct genes potentially targeted by selection. The *D* and/or H_{norm} tests detected three loci departing from neutrality that were not detected by the DHEW compound tests. These results may be explained by the sensitivity of the *D* and H_{norm} tests to demographic factors and different degrees of background selection (Zeng et al. 2006). Especially, Tajima's *D* is sensitive to background selection and population growth, whereas the H_{norm} is more sensitive to population shrinkage and subdivision. In a previous study, in *P. halepensis* based on ten candidate genes, we have shown that this species has undergone historical bottlenecks and that the western populations of the species harbored some signatures of this

Table 3. Polymorphism, Divergence and Neutrality Tests (silent sites) in *Pinus pinaster* for the Phylogeographical Groups Defined with Chloroplast Microsatellites (see text for details).

Amplicon	*S*	*D*	H_{norm}	DHEW *P* value	ML-HKA's *K* (*P. taeda*)	K_s (*P. taeda*)	ML-HKA's *K* (*P. halepensis*)	K_s (*P. halepensis*)
Western Mediterranean[a]								
lp31-Pt								
a	8	2.172	−0.635	0.958	0.961	38.350	1.292	20.310
b	6	1.834	−1.380	0.913	0.698	44.250	0.720	24.360
lp33-Pp	6	1.133	0.911	0.984	0.696	60.060	1.839	15.990
dhn2-Pp								
a	7	0.269	−0.853	0.406	1.605	38.160	1.307	31.580
b	19	0.186	0.825	0.727	3.256*	31.940	1.975	39.720
dhn2-Ps								
a	15	0.631	0.703	0.562	ptna	ptna	phna	phna
b	8	−0.126	0.232	0.209	ptna	ptna	d	d
dhn5-Ps	2	1.960	0.136	0.907	ptna	ptna	phna	phna
4cl-Pt								
a	2	0.422	0.067	0.610	0.307	31.480	0.250	26.870
b	2	−1.041	−3.173*	0.014*	0.741	43.930	0.446	41.220
c	1	0.214	0.565	0.663	0.542	23.080	0.000	1.190
d	na	na	na	na	na	na	na	na
North Africa[b]								
lp31-Pt								
a	7	−0.276	−3.065*	0.336	0.878	39.300	0.952	22.280
b	6	−2.046*	−3.813	0.073	0.820	44.250	0.880	26.150
lp33-Pp	5	0.014	0.926	0.899	0.614	57.860	1.946	13.040
dhn2-Pp								
a	4	0.799	−0.554	0.615	0.868	36.610	0.792	28.620
b	13	1.095	0.361	0.718	2.036	32.780	1.362	33.600
dhn2-Ps								
a	6	0.190	−0.232	0.322	ptna	ptna	phna	phna
b	6	−1.116	−1.223	0.085	ptna	ptna	d	d
dhn5-Ps	2	0.688	−0.809	0.540	ptna	ptna	phna	phna
4cl-Pt								
a	8	2.797	0.411	0.995	1.279	39.160	0.442	28.880
b	3	2.266	0.090	0.982	1.303	48.190	1.017	34.730
c	2	0.201	0.387	0.494	1.039	25.240	0.000	3.410
d	na	na	na	na	na	na	na	na

NOTE.—*D*, Tajima's *D* (Tajima 1989); H_{norm}, Fay and Wu's normalized *H* (Zeng et al. 2006); DHEW, compound test (Zeng, Mano, et al. 2007); ML-HKA's *K*, selection parameter of the maximum likelihood multilocus Hudson–Kreitman–Aguadé neutrality tests (Wright and Charlesworth 2004); the species used as outgroup is given in parenthesis; *S*, number of segregating sites; K_s, average proportion of nucleotide differences between species per silent site $\times 10^{-3}$; na, amplicons that failed to transfer across species; ptna, *P. taeda* outgroup sequence not available; phna, *P. halepensis* outgroup sequence not available; d, discarded because possible paralog; *P* values (*P* neutrality test [neutral] $\leqslant$ *P* neutrality test [observed]): $0.01 < * < 0.05$.

[a]Western Mediterranean: Arenas de San Pedro, Cazorla, Coca, Cómpeta, and Olba.

[b]North Africa: Tabarka, Tamrabta, and Sidi Meskour.

demographic event (Grivet et al. 2009). In addition, based on the present data set and extensive coalescent simulation, we have found that observed values of *D* and H_{norm} for *P. pinaster* both in the western Mediterranean and in the North African groups reject the standard neutral model and suggest past bottlenecks in this species too (see supplementary fig. S3, Supplementary Material online). To test the robustness of the DHEW compound test, neutral coalescent simulations under realistic bottlenecks scenarios were further simulated in *P. pinaster* (1,000 coalescent simulations per scenario) and the significant values for this test were recorded. In all cases, the number of significant tests obtained was equal or lower than the number expected by chance (see supplementary fig. S3, Supplementary Material online), highlighting the robustness of the DHEW test in the presence of bottleneck events and geographic population structure of magnitudes similar to the ones found in maritime and Aleppo pines. In addition to our simulations, the DHEW test has been shown to be robust to recombination and to have high sensitivity to detect positive selection (Zeng, Mano, et al. 2007), as it combines powerful (and insensitive to recombination) haplotype-frequency spectrum tests (i.e., the EW test) with SFS statistics, such as the *D* test that maintains power to detect positive selection across a wide period of time (Zeng et al. 2006; Zhai et al. 2009). Only two ongoing events of selection were detected by the DHEW test, one in *P. pinaster* and one in *P. halepensis*. Both cases involved the *4cl-Pt* gene (albeit in different geographical regions and distinct amplicons), suggesting a potential role of *4cl* in local adaption in both pine species (see below).

Among the tests examining polymorphism within species and divergence among species, the MK test did not detect any loci under selection. Although this test is somewhat robust to demography, it seems less powerful in detecting positive selection than HKA type tests (Zhai et al. 2009). The ML-HKA tests for both silent and all sites, using

Table 4. Polymorphism, Divergence and Neutrality Tests (silent sites) in *Pinus halepensis* for the Phylogeographical Groups Defined with Chloroplast Microsatellites (see text for details).

Amplicon	*S*	*D*	H_{norm}	DHEW *P* value	ML-HKA's *K* (*P. taeda*)	K_s (*P. taeda*)	ML-HKA's *K* (*P. pinaster*)	K_s (*P. pinaster*)
Western Mediterranean[a]								
lp31-Pt								
a	2	1.689	−1.714	0.860	0.597	34.590	0.452	20.960
b	4	0.535	−0.948	0.501	0.889	48.280	1.380	23.090
lp33-Pp	1	−0.629	−4.343*	0.270	0.145*	69.100	0.361	14.020
dhn2-Pp								
a	0	nps	nps	nps	nps	55.770	nps	44.020
b	0	nps	nps	nps	nps	57.180	nps	38.220
dhn2-Ps								
a	na	na	na	na	na	na	na	na
b	2	−0.595	0.317	0.382	ptna	ptna	d	d
dhn5-Ps	na	na	na	na	na	na	na	na
4cl-Pt								
a	6	0.752	−1.222	0.599	2.236	43.260	1.618	22.500
b	16	2.386	−0.320	0.985	2.279	52.000	1.242	30.660
c	2	1.701	−1.305	0.867	1.350	31.970	2.976	9.440
d	3	2.324	−0.981	0.965	0.674	54.780	ppna	ppna
North Africa[b]								
lp31-Pt								
a	2	1.818	−1.053	0.898	0.819	34.140	1.123	12.710
b	2	1.882	0.602	0.909	0.530	49.000	0.924	22.700
lp33-Pp	0	nps	nps	nps	nps	69.410	nps	9.550
dhn2-Pp								
a	1	−0.774	0.289	0.313	0.329	56.280	0.525	19.310
b	1	−1.147	0.160	0.189	0.215	57.320	0.205	33.600
dhn2-Ps								
a	na	na	na	na	na	na	na	na
b	0	nps	nps	nps	ptna	ptna	d	d
dhn5-Ps	na	na	na	na	na	na	na	na
4cl-Pt								
a	4	1.306	−0.201	0.791	2.072	44.030	1.815	22.450
b	14	2.316	0.044	0.981	3.251*	50.760	1.951	31.290
c	2	1.077	0.221	0.834	1.665	27.720	3.309	5.200
d	3	−1.494	−6.584	0.009**	0.860	60.020	ppna	ppna

NOTE.—*D*, Tajima's *D* (Tajima 1989); H_{norm}, Fay and Wu's normalized *H* (Zeng et al. 2006); DHEW, compound(Zeng, Mano, et al. 2007); ML-HKA's *K*, selection parameter of the maximum-likelihood multilocus Hudson–Kreitman–Aguadé neutrality tests (Wright and Charlesworth 2004); the species used as outgroup is given in parenthesis; *S*, number of segregating sites; K_s: average proportion of nucleotide differences between species per silent site $\times 10^{-3}$; na, amplicons that failed to transfer across species; ptna, *P. taeda* outgroup sequence not available; ppna, *P. pinaster* outgroup sequence not available; d: discarded because possible paralog; nps, monomorphic fragment; *P* values (*P* neutrality test [neutral] $\leqslant$ *P* neutrality test [observed]): $0.001 < ** < 0.01 < * < 0.05$.

[a]Western Mediterranean: Cabanellas, Carratraca, Imperia, S'avall, and Tarrasa.

[b]North Africa: Aures Beni Melloul and Zaouia Ifrane.

P. taeda as outgroup, identified three amplicons with uncoupled levels of polymorphism and divergence in the two pines. Two genes (*dhn2-Pp_b* in *P. pinaster* in the western Mediterranean group and *4cl-Pt_b* in *P. halepensis* in the North African group) showed high levels of diversity compared with divergence (selection parameter higher than one), a pattern compatible with balancing selection. One other gene (*lp33-Pp* in the western Mediterranean group in *P. halepensis*) showed low diversity compared with divergence (selection parameter lower than one), which could reflect the transient reduction in variability occurring during a selective sweep. Three other genes showed uncoupled levels of polymorphism and divergence in *P. halepensis* when examining all sites but not when examining silent sites alone. Two of these genes (*4cl-Pt_a* and *4cl-Pt_b* in the western Mediterranean group) had an excess of nucleotide diversity, which could reflect recent balancing selection acting on nonsynonymous sites, whereas for one gene (*dhn2-Pp_a* in the western Mediterranean group) purifying selection reducing diversity in nonsynonymous sites could explain the observed pattern. Evidence for extensive purifying selection in conifers comes from an approximately 4-fold nucleotide diversity at synonymous compared with nonsynonymous sites in most species (see table 1 in González-Martínez et al. 2010), including maritime ($dN/dS = 0.169$) and Aleppo ($dN/dS = 0.344$) pines (this study). Although demographical events can result in a significant ML-HKA test, its multilocus nature combined with the HKA framework should produce a more robust test than those based on comparing different aspects of polymorphism at a single locus (Wright and Charlesworth 2004).

In order to gain power by increasing sample size while taking the specific history of each species into account, we performed the neutrality tests in groups of populations that present similar evolutionary history (i.e., North African and western Mediterranean groups). However, grouping

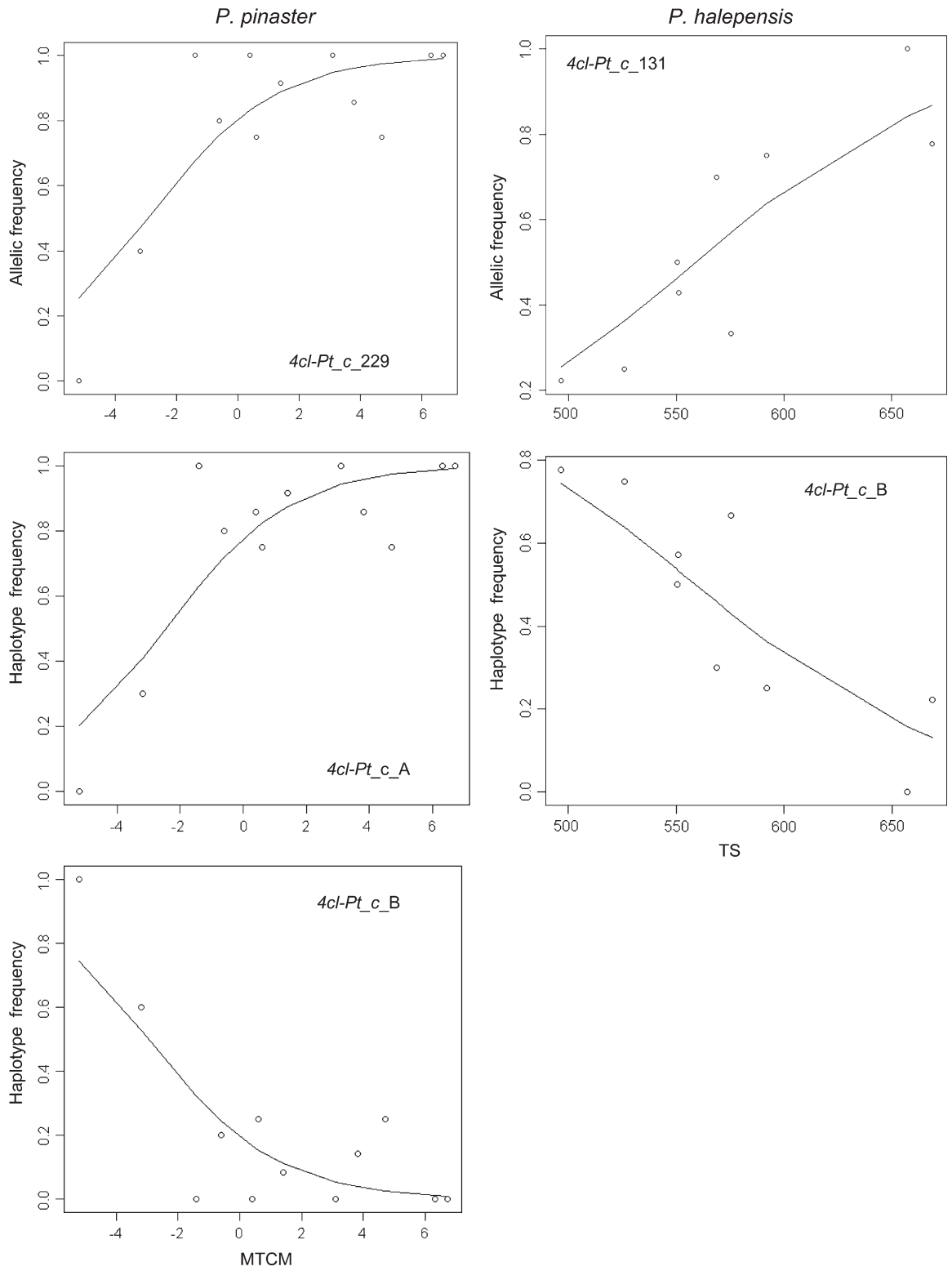

FIG. 2. Predicted and observed patterns of clinal variation at *4cl-Pt_c*. Sigmoid curves denote the clines in either allelic (SNP position given) or haplotypic (haplotypes represented by capital letters) frequencies under a logistic regression model. The logistic regression model uses a climatic variable (MTCM or TS) after controlling for neutral processes (i.e., using the PCs of neutral molecular markers as covariates) for *P. pinaster* and *P. halepensis*. MTCM: minimum temperature of the coldest month; TS: temperature seasonality.

populations may bias the output of the SFS-based neutrality tests as it can increase the proportion of singletons, leading to an excess of low frequency variants and thus to more negative *D* (Städler et al. 2009). In our study, populations were grouped according to homogeneous gene pools (see Bucci et al. 2007 for *P. pinaster* and Grivet et al. 2009

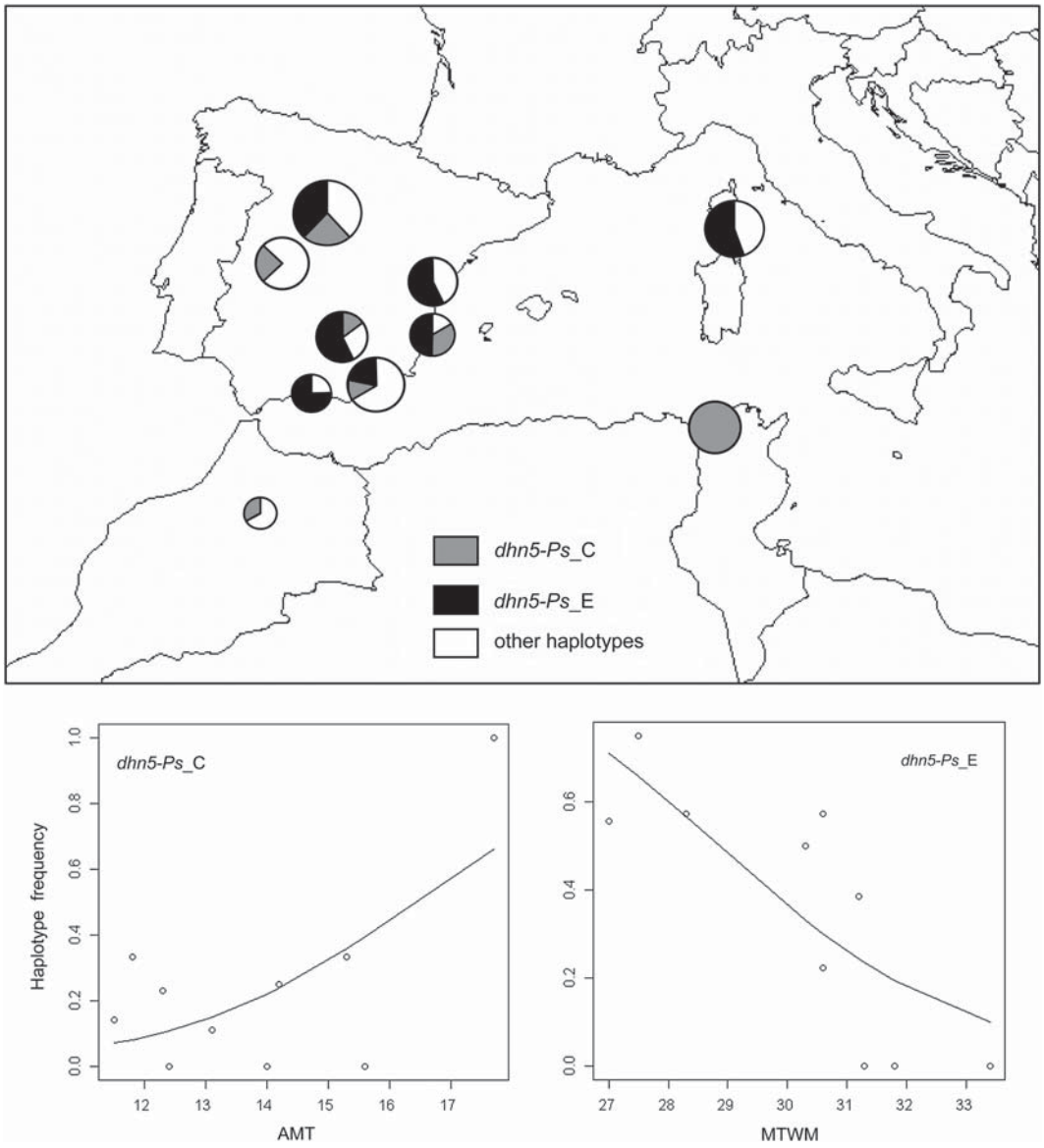

FIG. 3. Upper section: *dhn5-Ps* haplotype distribution showing the proportion of haplotypes C (in gray), E (in black), and all others (in white) in *P. pinaster*. The size of the pie is proportional to the sample size at each location. Lower section: expected (sigmoid curves) and observed covariation in haplotype frequencies of *dhn5-Ps*_C and *dhn5-Ps*_E and climatic variables. Expected curves were obtained under a logistic regression model including a single environmental variable (ATM or MTWM) after controlling for neutral processes (i.e., using the PCs of neutral molecular markers as covariates) for *P. pinaster*. AMT: annual mean temperature; MTWM: maximum temperature of the warmest month.

for *P. halepensis*), the consequence of which should be that no substantial differences in the number of rare variants at the population and group levels are found. The extensive gene flow normally found in conifers with large and continuous distribution would also support this approach. Nevertheless, to check that our results were not biased because of population grouping, we performed neutrality tests at the population level too and found that values of *D* test at the group level were not more negative than at the single population level (data not shown). Moreover, results of the H_{norm} and DHEW tests should not be affected by the accumulation of singletons as the first is based on intermediary and high-frequency variants, whereas the second integrates all variants of the SFS. Thus, our analyses appear reasonably robust to population grouping.

Selective events appear to have affected distinct genes in *P. pinaster* and in *P. halepensis* despite their overlapping environment and close phylogenetic relationship, a result that may be connected to the different histories of the two pines that have likely resulted in different selective pressures. In *P. pinaster*, selective events were detected only within the western Mediterranean group, whereas in *P. halepensis*, footprints of selection were identified both in the western Mediterranean and in the North African groups. These results point to different geographical selection pressures that may have led to the process of regional adaptation of the pines (see Barbéro et al. 1998 for a description of the different Mediterranean environmental zones; and Gómez and Zamora 2000; Nakazato et al. 2008; Montesinos et al. 2009 for some examples on adaptive variation across

heterogeneous environments). It is also noticeable that *P. pinaster* had lower levels of nucleotide variation within the North African range. This fact can be attributed to population history and/or interpreted in the light of more extreme environmental conditions constraining population sizes and/or resulting in stronger selection in this range.

Environmental Associations

Environmental associations identified two loci that were correlated with temperature, suggesting the importance of this climatic variable as selective agent (Saxe et al. 2001; Jump et al. 2006). One of these genes was common for both pines (*4cl-Pt* correlated with MTCM in *P. pinaster* and with TS in *P. halepensis*), whereas the other gene association was only significant in *P. pinaster* (*dhn5-Ps* correlated with AMT and MTWM). Some of the correlations between genetic and climatic data in *P. pinaster* were due basically to the extreme values of a few North African populations (Sidi Meskour and Tamrabta for *4cl-Pt_c*; Tabarka for *dhn5-Ps _C*; see figs. 2 and 3) and may not represent true adaptive responses to environmental gradients but local adaptation to particular environments or genetic drift due to population isolation (see references in Alleaume-Benharira et al. 2006; Rosenblum et al. 2007). In contrast, the other associations (*4cl-Pt_c* in *P. halepensis* and *dhn5-Ps_E* in *P. pinaster*) showed more robust patterns. For instance, *dhn5-Ps _E* tended to be absent from populations characterized by the highest MTWM (Arenas de San Pedro in central Spain, Tamrabta in Moroccan Middle Atlas, and Tabarka in coastal Tunisia) regardless of their geographical location (table 1 and fig. 3).

Environmental correlations with allelic variation can be spuriously inflated by neutral processes that may also generate genetic clines, such as population history or population genetic structure. This is particularly true for forest trees that may have followed a postglacial colonization pathway overlapping temperature and rainfall clines. Here, to control for confounding associations between candidate genes and environmental data, we included variation of chloroplast microsatellites as a covariate in the multivariate logistic regressions. In *P. pinaster*, the use of cpSSRs has allowed a more accurate description of genetic structure (Bucci et al. 2007) than previous studies based on biochemical markers, such as terpenes or allozymes (e.g., Baradat and Marpeau-Bezard 1988; Salvador et al. 2000). Recently, nuclear molecular data for a geographically limited set of populations (nuSSRs and SNPs) (Eveno et al. 2008; our unpublished results) have identified similar gene pools as Bucci et al. (2007). More extensive sampling with highly polymorphic biparentally inherited nuclear markers could, in principle, reveal a more complex spatial genetic pattern affecting the outcome of some of the multivariate logistic regressions reported in this study. However, because the available chloroplast data set strongly reflects the species' history, it is expected to "correct" for the presence of overall neutral genetic gradients. Accordingly, including these markers as covariates identified a substantial fraction of the correlations initially retained as significant as false positives (83.6% in *P. pinaster* and 85.7% in *P. halepensis*) (supplementary tables S6 and S7, Supplementary Material online).

The Dehydrin Gene Family

Neutrality tests and environmental associations both point to dehydrins as potential targets of natural selection in *P. pinaster*. Some dehydrins were also suggested to be under selection in a study based on detection of outlier loci in this species (Eveno et al. 2008). Dehydrins displayed also non-neutral patterns of nucleotide diversity in *P. sylvestris* populations showing divergence for cold tolerance in Europe (Wachowiak et al. 2009) and were associated with carbon isotope discrimination (and, thus, potentially with drought tolerance) in *P. taeda* (González-Martínez et al. 2008). Altogether these studies suggest the involvement of dehydrins in the adaptive response of pines to abiotic stress.

Dehydrins are part of a relatively small multigene family of intracellular stabilizers that plays a major role in cell protection against desiccation. These proteins are produced in response to any type of stress that causes dehydration at the cellular level, such as cold, drought, or salinity (Close 1997). Changes in dehydrin gene expression have been reported in response to drought and/or cold stress in many plants, such as cowpea (Ismail et al. 1999), barley (Suprunova et al. 2004), wheat (Lopez et al. 2001), apple (Wisniewski et al. 2008), tomato (Weiss and Egea-Cortines 2009), and blueberry (Panta et al. 2001), whereas transgenic experiments confirmed the role of dehydrin genes in enhancing tolerance to drought or freezing stress in plants (e.g., in *Arabidopsis*, Puhakainen et al. 2004). In conifers, at least eight dehydrin genes have been identified in *P. sylvestris* (Joosen et al. 2006) and *Picea abies* (Yakovlev et al. 2008), and dehydrin expression has been shown to increase under wounding, cold, and drought stress (Richard et al. 2000; Watkinson et al. 2003). Thus, it is not surprising to find in our study significant associations between dehydrins and temperatures, as critical low temperatures can cause tissue injury while high temperatures accompany dehydration, both stimulating the accumulation of dehydrins (Lewitt 1980; Ingram and Bartels 1996; Rizhsky et al. 2002). Although our results, and others in *P. pinaster*, suggest that selection may have acted on some of the dehydrins tested, further genetic association and functional studies are necessary to confirm the role of dehydrins in local adaptation of this Mediterranean pine.

The 4cl Gene

There was converging evidence both from neutrality tests (DHEW in *P. pinaster* and both DHEW and ML-HKA in *P. halepensis*) and from environmental associations that the *4cl* gene may be under selection in the two Mediterranean pines studied. Preliminary results on phenotypical association corrected by neutral gradients in *P. pinaster* also point at *4cl-Pt*. Indeed, multiple regressions reveal significant associations between polymorphism in *4cl-Pt_c* at the SNP and haplotype levels and total height in populations growing in sites characterized by dry and intermediate humidity (supplementary table S8, Supplementary Material online). The *4cl* family has been extensively studied in plants, where

it is encoded by four to five genes in the fully sequenced genomes of *Arabidopsis*, rice, and poplar (reviewed in Souza et al. 2008). The *4cl* gene is involved, among other processes, in the production of basic enzymes of the phenylpropanoid metabolism that are important metabolites acting as protectants against biotic and abiotic stresses (Rani et al. 2009). The *4cl* gene also encodes key enzymes in the biosynthesis of lignin and several studies have demonstrated its involvement in plant growth (e.g., Yun et al. 2005, 2009; Wagner et al. 2009). Thus, we expect that changes in *4cl* function would have significant repercussions on tree physiology and morphology. Implication of *4cl* in pine morphology and physiology has been shown in gene association studies (for *P. taeda* see González-Martínez, Wheeler, et al. 2007) as well as in gene suppression studies (for *P. radiata* see Wagner et al. 2009).

Neutrality Tests Versus Environmental Correlations

The different approaches used in this study suggested different loci under selection within each of the two Mediterranean pine species. This fact has to be connected to the specificities of the statistical methods used, whose performance has been recently studied under various demographic and recombination scenarios (Zeng, Mano, et al. 2007; Zeng, Shi, et al. 2007; Ramírez-Soriano et al. 2008; Zhai et al. 2009). Within-species SFS and haplotype-frequency spectrum-based methods are suitable to detect ongoing or recently fixed selective sweeps—EW tending to be more powerful around the time when a selected mutation reaches fixation (Zeng, Shi, et al. 2007). In contrast, tests comparing levels of (within species) polymorphism and (among species) divergence are able to detect the cumulative effects of positive selection events over a wider evolutionary scale. Within this category, the ML-HKA test is suitable to quantify the amount of selection in the genome and, due to its multilocus nature, it is expected to be more robust to changes in population size such as population bottlenecks and expansion than traditional tests (Wright and Charlesworth 2004). Simulations have shown that the HKA test is relatively insensitive to change of divergence time as most of its power comes from the transient reduction in variability occurring during selective sweeps (Zhai et al. 2009). However, we only obtained significant tests when using the less closely related outgroup (i.e., a New World pine, *P. taeda*). This points to either lower power when divergence time from the outgroup is low or, alternatively, to selection events that took place before the split of the two Mediterranean pines considered. In our study, none of the EW neutrality tests were significant (data not shown), whereas both the DHEW and the ML-HKA tests revealed distinct genes potentially under selection. As a consequence, potential selective events detected by the neutrality tests assessed in this study would correspond to either ongoing or relatively old selective events (i.e., before the diversification of the Mediterranean pines in the Miocene, $\sim$ 10 Ma; Gernandt et al. 2008).

Last generation tests, such as DHEW or ML-HKA, are powerful for detecting positive selection and are relatively insensitive to other evolutionary forces, but still do not integrate recombination rate, a factor that can produce substantial biases (Nielsen et al. 2007; Ramírez-Soriano et al. 2008), or provide insights on selection drivers. Methods based on correlation between genetic and environmental data are appealing because they aim at understanding species-specific adaptations and processes that connect an organism to its environment by looking at allelic frequencies or the genetic structure of populations (Foll and Gaggiotti 2006; Joost et al. 2008). These correlation methods are also more appealing than methods based on detection of outliers (e.g., Eveno et al. 2008) as they target particular selection drivers and provide a hypothesis testing framework. In addition, they are well suited for high-throughput genotyping even for nonmodel species (Namroud et al. 2008; Eckert et al. 2010). However, correlation methods are not without limitations, as it may be challenging to find the environmental factors that are relevant for each species adaptation, and associations have to be controlled for historical and demographical processes, in particular for allelic clines produced by postglacial migrations in Europe and the Americas.

Together, neutrality tests and environmental association approaches complement each other by looking at different evolutionary scales and types of selection. In our study, they detected a relatively high number of genes showing nonneutral patterns of evolution, a result that can be attributed to a selection of candidate genes based on earlier studies (González-Martínez et al. 2006; Eveno et al. 2008; Grivet et al. 2009). This fact supports a candidate gene approach with targeted genes, at least for organisms that have large genomes (e.g., conifers), which, so far, preclude dense genome-wide sampling.

Conclusion

Our pluralistic approach revealed a dynamic action of natural selection in space and time for maritime and Aleppo pines (Felsenstein 1976; see Vasemagi and Primmer 2005 for other examples). Selection events along with environmental associations were detected; some of these events differ between the two species reflecting individual histories (recolonization, demography, and adaptation), whereas others are shared, which translates partly as a common history of these closely related and partially sympatric Mediterranean pines. The simultaneous search for patterns of selection in two closely related species can help to understand the evolutionary forces responsible for adaptive responses and thus provides an effective way of assessing the degree of local adaptation, a key factor to integrate in future management and conservation strategies (Wright and Gaut 2005).

Supplementary Material

Supplementary tables S1–S8 and figures S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgements

We are grateful to C. García-Barriga for her assistance in producing sequences of *P. pinaster* candidate genes, to C. Collada, M.T. Cervera, M.A. Guevara, G. Gill, G.R. Brown, and D.B. Neale for providing some of the primers, as well as to A.I. de-Lucas and C. Ordoñez for supplying some of the *P. pinaster* seeds. We thank J. Gonzalo for kindly providing the climatic data from his functional phytoclimatic model, as well as S.E. Ramos-Onsins for his scripts to compute the diversity parameters, and K. Zeng for his scripts to compute the DHEW neutrality test. Thanks are extended to M. Robson, M.A. Zavala, and J. Climent for helpful discussion on the study. We are also thankful to A.J. Eckert for thoughtful comments on the manuscript, N. Takebayashi and two anonymous reviewers for constructive comments and suggestions, and P.C. Grant for revising the English. The EUFORGEN program (Bioversity International) provided the distribution maps used in figure 1. This work was supported by the EU EVOLTREE Network of Excellence, the Collaborative Project on "Conservation of Forest Genetic Resources" between the Spanish Ministry of Environment and National Institute for Agriculture and Food Research and Technology (AEG06-054), and Projects CGL2008-05289-C02-02/BOS (VaMPiro) and EUI2008-03713 (BiodivERsA-LinkTree) from the Spanish Ministry of Science and Innovation, the later through the FP6 BiodivERsA Eranet. T.B. acknowledges support from the Danish Council for Independent Research through a Steno fellowship.

References

Alleaume-Benharira M, Pen IR, Ronce O. 2006. Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *J Evol Biol.* 19:203–215.

Aranda I, Alía R, Ortega U, Dantas AK, Majada J. 2009. Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from *Pinus pinaster* populations. *Tree Genet Genomes*. 6:169–178.

Atzmon N, Moshe Y, Schiller G. 2004. Ecophysiological response to severe drought in *Pinus halepensis* Mill. trees of two provenances. Plant Ecol. 171:15–22.

Baradat PH, Marpeau-Bezard A. 1988. Le pin maritime, *Pinus pinaster* Ait. Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce [PhD thesis]. Bordeaux (France): University of Bordeaux I.

Barbéro M, Loisel R, Quézel P, Richardson DM, Romane F. 1998. Pines of the Mediterranean Basin. In: Richardson DM, editor. Ecology and biogeography of *Pinus*. Cambridge (UK): Cambridge University Press. p. 153–170.

Biswas S, Akey JM. 2006. Genomic insights into positive selection. *Trends Genet*. 22:437–446.

Bucci G, Anzidei M, Madaghiele A, Vendramin GG. 1998. Detection of haplotypic variation and natural hybridization in halepensis-complex pine species using chloroplast simple sequence repeat (SSR) markers. *Mol Ecol*. 7:1633–1643.

Bucci G, González-Martínez SC, Le Provost G, Plomion C, Ribeiro MM, Sebastiani F, Alía R, Vendramin GG. 2007. Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Mol Ecol*. 16:2137–2153.

Burban C, Petit RJ. 2003. Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Mol Ecol*. 12:1487–1495.

Chambel MR, Climent J, Alía R. 2007. Divergence among species and populations of Mediterranean pines in biomass allocation of seedlings grown under two watering regimes.*Ann For Sci*. 64: 87–97.

Climent J, Prada MA, Calama R, Chambel MR, De Ron DS, Alía R. 2008. To grow or to seed: ecotypic variation in reproductive allocation and cone production by young female Aleppo pine (*Pinus halepensis*, Pinaceae).*Am J Bot*. 95:833–842.

Close T. 1997. Dehydrins: a commonality in the response of plants to dehydration and low temperature. Physiol Plant. 100:291–296.

Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW, Pritchard JK. 2009. The role of geography in human adaptation. PLoS Genet. 5(6):e1000500.

Costa P, Bahrman N, Frigerio JM, Kremer A, Plomion C. 1998. Water-deficit-responsive proteins in maritime pine. Plant Mol Biol. 38:587–596.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA minipreparation. Version II. Plant Mol Biol Rep. 1:19–21.

Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB. 2010. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* 185:969–982.

Eveno E, Collada C, Guevara MA, et al. (11 co-authors). 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.

Felsenstein J. 1976. Theoretical population-genetics of variable selection and migration. *Annu Rev Genet*. 10:253–280.

Foll M, Gaggiotti O. 2006. Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174:875–891.

Frankel N, Carrari F, Hasson E, Iusem ND. 2006. Evolutionary history of the Asr gene family. *Gene* 378:74–83.

Gernandt DS, Lopez GG, Garcia SO, Liston A. 2005. Phylogeny and classification of *Pinus*. *Taxon* 54:29–42.

Gernandt DS, Magallon S, Lopez GG, Flores OZ, Willyard A, Liston A. 2008. Use of simultaneous analyses to guide fossil-based calibrations of *Pinaceae* phylogeny. *Int J Plant Sci*.169:1086–1099.

Giorgi F, Lionello P. 2008. Climate change projections for the Mediterranean region. *Glob Planet Change* 63:90–104.

Gómez JM, Zamora R. 2000. Spatial variation in the selective scenarios of *Hormathophylla spinosa* (Cruciferae). *Am Nat*. 155:657–668.

González-Martínez SC, Dillon S, Garnier-Géré P, et al. (16 co-authors). Forthcoming 2010. Patterns of nucleotide diversity and association mapping. In: Kole C, editor. Genomics of conifers. Enfield (NH): Science Publishers, Inc.

González-Martínez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB. 2006. DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics* 172:1915–1926.

González-Martínez SC, Gómez A, Carrión JS, Agúndez D, Alía R, Gil L. 2007. Spatial genetic structure of an explicit glacial refugium of maritime pine (*Pinus pinaster* Aiton) in southeastern Spain. In: Weiss S, Ferrand N, editors. Phylogeography in southern European refugia: evolutionary perspectives on the origins of European biodiversity. London: Springer. p. 257–269.

González-Martínez SC, Huber D, Ersoz E, Davis JM, Neale DB. 2008. Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* 101:19–26.

González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB. 2007. Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* 175:399–409.

Gonzalo J. 2007. Phytoclimatic analysis of the Spanish Peninsula. Update and geostatistical analysis [PhD thesis]. Palencia (Spain): University of Valladolid.

Gram WK, Sork VL. 2001. Association between environmental and genetic heterogeneity in forest tree populations. *Ecology* 82:2012–2021.

Grivet D, Sebastiani F, González-Martínez SC, Vendramin GG. 2009. Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. *New Phytol.* 184:1016–1028.

Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G, Di Rienzo A. 2008. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet*. 4(2):e32.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 25:1965–1978.

Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153–159.

Ingram J, Bartels D. 1996. The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol. 47:377–403.

Ismail AM, Hall AE, Close TJ. 1999. Allelic variation of a dehydrin gene cosegregates with chilling tolerance during seedling emergence. *Proc Natl Acad Sci U S A*. 96:13566–13570.

Joosen RVL, Lammers M, Balk PA, Bronnum P, Konings MCJM, Perks M, Stattin E, Van Wordragen MF, van der Geest AHM. 2006. Correlating gene expression to physiological parameters and environmental conditions during cold acclimation of *Pinus sylvestris*, identification of molecular markers using cDNA microarrays. *Tree Physiol.* 26:1297–1313.

Joost S, Kalbermatten M, Bonin A. 2008. Spatial analysis method (SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. *Mol Ecol Resour*. 8:957–960.

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J. 2006. Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Mol Ecol.* 15:3469–3480.

Keller SR, Sowell DR, Neiman M, Wolfe LM, Taylor DR. 2009. Adaptation and colonization history affect the evolution of clines in two introduced species. *New Phytol.* 183:678–690.

Lewitt J. 1980. Responses of plants to environmental stresses. Vol. I. Chilling, freezing and high temperature stresses. New York: Academic Press.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.

Lopez CG, Banowetz G, Peterson CJ, Kronstad WE. 2001. Differential accumulation of a 24-kd dehydrin protein in wheat seedlings correlates with drought stress tolerance at grain filling. *Hereditas* 135:175–181.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol.* 18:189–197.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the ADH locus in *Drosophila*. *Nature* 351:652–654.

Montesinos A, Tonsor SJ, Alonso-Blanco C, Picó FX. 2009. Demographic and genetic patterns of variation among populations of *Arabidopsis thaliana* from contrasting native environments. *PLoS One* 4:e7213.

Morgante M, Felice N, Vendramin GG. 1998. Analysis of hypervariable chloroplast microsatellites in *Pinus halepensis* reveals a dramatic genetic bottleneck. In: Karp A, Isaac PG, Ingram DS, editors. Molecular tools for screening biodiversity. London, UK: Chapman and Hall. p. 407–412.

Nakazato T, Bogonovich M, Moyle LC. 2008. Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. *Evolution* 62:774–792.

Namroud MC, Beaulieu J, Juge N, Laroche J, Bousquet J. 2008. Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Mol Ecol.* 17:3599–3613.

Nguyen A, Lamant A. 1989. Variation in growth and osmotic regulation of roots of water-stressed maritime pine (*Pinus pinaster* Ait) provenances. *Tree Physiol.* 5:123–133.

Nielsen R. 2005. Molecular signatures of natural selection. *Annu Rev Genet.* 39:197–218.

Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. 2007. Recent and ongoing selection in the human genome. *Nat Rev Genet.* 8:857–868.

Panta GR, Rieger MW, Rowland LJ. 2001. Effect of cold and drought stress on blueberry dehydrin accumulation. *J Horticult Sci Biotechnol.* 76:549–556.

Parisod C, Christin PA. 2008. Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). *New Phytol.* 178:436–447.

Petit RJ, Hampe A, Cheddadi R. 2005. Climate changes and tree phylogeography in the Mediterranean. *Taxon* 54:877–885.

Puhakainen T, Hess MW, Mäkelä P, Svensson J, Heino P, Palva ET. 2004. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Mol Biol.* 54:743–753.

Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics* 179:555–567.

Rani A, Singh K, Sood P, Kumar S, Ahuja PS. 2009. p-Coumarate:CoA ligase as a key gene in the yield of catechins in tea [*Camellia sinensis* (L.) O. Kuntze]. *Funct Integr Genomics* 9:271–275.

Resch T. 1974. Essai de distinction des races majeurs de *Pinus pinaster*. *Ann Rech For Maroc* 14:91–102.

Richard S, Morency MJ, Drevet C, Jouanin L, Séguin A. 2000. Isolation and characterization of a dehydrin gene from white spruce induced upon wounding, drought and cold stresses. *Plant Mol Biol.* 43:1–10.

Richardson DM, Bond WJ. 1991. Determinants of plant-distribution—evidence from pine invasions. *Am Nat.* 137:639–668.

Rizhsky L, Liang HJ, Mittler R. 2002. The combined effect of drought stress and heat shock on gene expression in tobacco. *Plant Physiol.* 130:1143–1151.

Rosenblum EB, Hickerson MJ, Moritz C. 2007. A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution* 61:2971–2985.

Salvador L, Alía R, Agúndez D, Gil L. 2000. Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. *Theor Appl Genet.* 100:89–95.

Sathyan P, Newton RJ, Loopstra CA. 2005. Genes induced by WDS are differentially expressed in two populations of Aleppo pine (*Pinus halepensis*). *Tree Genet Genomes* 1:166-173.

Savolainen O, Pyhäjärvi T, Knürr T. 2007. Gene flow and local adaptation in trees. *Annu Rev Ecol Evol Syst.* 38:595–619.

Saxe H, Cannell MGR, Johnsen B, Ryan MG, Vourlitis G. 2001. Tree and forest functioning in response to global warming. *New Phytol.* 149:369–399.

Soto A, Robledo-Arnuncio JJ, González-Martínez SC, Smouse PE, Alía R. 2010. Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. *Mol Ecol.* 19:1396–1409.

Souza CD, Barbazuk B, Ralph SG, Bohlmann J, Hamberger B, Douglas CJ. 2008. Genome-wide analysis of a land plant-specific acyl: coenzymeA synthetase (ACS) gene family in *Arabidopsis*, poplar, rice and *Physcomitrella*. *New Phytol.* 179:987–1003.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P. 2009. The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* 182:205–216.

Storz JF. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Mol Ecol*. 11:2537–2551.

Suprunova T, Krugman T, Fahima T, Chen G, Shams I, Korol A, Nevo E. 2004. Differential expression of dehydrin genes in wild barley, *Hordeum spontaneum*, associated with resistance to water deficit. *Plant Cell Environ*. 27:1297–1308.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.

Tapias R, Climent J, Pardos JA, Gil L. 2004. Life histories of Mediterranean pines. *Plant Ecol*. 171:53–68.

Vasemagi A, Primmer CR. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Mol Ecol*. 14:3623–3642.

Voltas J, Chambel M, Prada M, Ferrio J. 2008. Climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine grown in common-garden tests. *Trees Struct Funct*. 22:759–769.

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.). *Tree Genet Genomes* 5:117–132.

Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, Torr K, Koch G, Schmitt U, Ralph J. 2009. Suppression of 4-Coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. *Plant Physiol*. 149:370–383.

Watkinson JI, Sioson AA, Vasquez-Robinet C, et al. (14 co-authors). 2003. Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed loblolly pine. *Plant Physiol*. 133:1702–1716.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Poppul Biol*. 7:256–276.

Watterson GA. 1978. The homozygosity test of neutrality. *Genetics* 88:405–417.

Weiss J, Egea-Cortines M. 2009. Transcriptomic analysis of cold response in tomato fruits identifies dehydrin as a marker of cold stress. *J Appl Genet*. 50:311–319.

Wisniewski M, Bassett C, Norelli J, Macarisin D, Artlip T, Gasic K, Korban S. 2008. Expressed sequence tag analysis of the response of apple (*Malus* × domestica 'Royal Gala') to low temperature and deficit. *Physiol Plant*. 133:298–317.

Wright SI, Charlesworth B. 2004. The HKA test revisited: a Maximum-Likelihood-Ratio test of the standard neutral model. *Genetics* 168:1071–1076.

Wright SI, Gaut BS. 2005. Molecular population genetics and the search for adaptive evolution in plants. *Mol Biol Evol*. 22:506–519.

Yakovlev IA, Asante DKA, Fossdal CG, Partanen J, Junttila O, Johnsen O. 2008. Dehydrins expression related to timing of bud burst in Norway spruce. *Planta* 228:459–472.

Yun MS, Chen WJ, Deng F, Yogo Y. 2005. Differential properties of 4-coumarate: CoA ligase related to growth suppression by chalcone in maize and rice. *Plant Growth Regul*. 46:169–176.

Yun MS, Chen WJ, Deng F, Yogo Y. 2009. Selective growth suppression of five annual plant species by chalcone and naringenin correlates with the total amount of 4-coumarate: coenzyme A ligase. *Weed Biol Manag*. 9:27–37.

Zeng K, Fu Y-X, Shi S, Wu C-I. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* 174:1431–1439.

Zeng K, Mano SH, Shi SH, Wu CI. 2007. Comparisons of site- and haplotype-frequency methods for detecting positive selection. *Mol Biol Evol*. 24:1562–1574.

Zeng K, Shi S, Wu CI. 2007. Compound tests for the detection of hitchhiking under positive selection. *Mol Biol Evol*. 24:1898–1908.

Zhai WW, Nielsen R, Slatkin M. 2009. An investigation of the statistical power of neutrality tests based on comparative and population genetic data. *Mol Biol Evol*. 26:273–283.