JOURNAL OF ECOLOGY - **Accepted Article** -

Genetic variation in the constitutive defensive metabolome and its inducibility are geographically structured and largely determined by demographic processes in maritime pine

López-Goldar, Xosé[1,2,4,5]; Villari, Caterina[2,3]; Bonello, Pierluigi[2]; Borg-Karlson, Anna Karin[4]; Grivet, Delphine[5,6]; Sampedro, Luis[1]; Zas, Rafael[1]

[1] Misión Biológica de Galicia, Consejo Superior de Investigaciones Científicas. Apdo. 28, 36080, Pontevedra, Spain.

[2] Department of Plant Pathology, The Ohio State University, Columbus, OH 43210, U.S.A.

[3] Daniel B. Warnell School of Forestry and Natural Resources, University of Georgia, 180 E Green Street, Athens, GA 30602, U.S.A

[4] Department of Chemistry, Ecological Chemistry Group, Royal Institute of Technology, 100 44 Stockholm, Sweden

[5] Department of Forest Ecology & Genetics, Forest Research Centre, INIA-CIFOR, Carretera de la Coruña km 7.5, 28040 Madrid, Spain.

[6] Sustainable Forest Management Research Institute, INIA-University of Valladolid, Avda. Madrid 44, 34004 Palencia, Spain.

Author for correspondence: Xosé López-Goldar, email: xlgoldar@mbg.csic.es

JOURNAL OF ECOLOGY - **Accepted Article** -

Received: 7 December 2018

Accepted: 22 February 2019

Published Online: 01 March 2019

Handling Editor: Richard Shefferson

Please cite this article as: López-Goldar X, Villari C, Bonello P, et al. Genetic variation in the constitutive defensive metabolome and its inducibility are geographically structured and largely determined by demographic processes in maritime pine. J Ecol. 2019;00:1–14.

https://doi.org/10.1111/1365-2745.13159

https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2745.13159

Abstract

1. Interspecific phenotypic variation in plant secondary metabolites (PSM) is often explained by biotic and abiotic factors. However, patterns of variation within species do not clearly fit the theoretical predictions. Exploring how genetics, environment and demographic processes shape such variation among and within populations is crucial for understanding evolution of PSM, particularly in long-lived plants such as forest trees.
2. Here we quantified genetic variation in PSM among and within populations, and explored drivers of local adaptation by studying the role of climate as source of population differentiation in PSM of maritime pine. Constitutive profile and concentrations of 63 PSM and their herbivory-associated inducibility were determined in the bark of 130 clonally replicated genotypes with known familial structure from ten populations covering the distribution range of the species. We compared neutral and quantitative population genetic differentiation of PSM (F_{ST} and Q_{ST}). Also, we accounted for population genetic structure and kinship among individuals when exploring climate-trait relationships.
3. We found large population differentiation and additive genetic variation in constitutive PSM. Many PSM were inducible, although very low genetic variation was observed with respect to their inducibility. Q_{ST}-F_{ST} comparisons suggest that differentiation of most diterpenes, monoterpenes, and phenolics can be explained by neutral demographic processes. Spatially heterogeneous selection across populations leading to local adaptation was only found for total constitutive sesquiterpenes and a few individual PSM. After accounting for population genetic structure, only the constitutive concentration of two sesquiterpenes showing signs of diversifying selection was predicted by climate, with decreasing concentrations along a growth-prone climatic gradient.
4. Synthesis. Evolutionary patterns of PSM depended on their chemical nature, with neutral differentiation governing most PSM. Evidence of local adaptation was only found for total constitutive sesquiterpenes and a few individual PSM. The low genetic variation in the inducibility of PSM suggests a conserved model of defensive induction in this species. Since population differentiation linked to past demographic history could lead to false positives of adaptive differentiation signals, accounting for the genetic relatedness among populations is required to infer the environmental determinants of intraspecific genetic variation in putatively adaptive traits such as plant defences.

Keywords

environmental gradients – F_{ST} – genetic variation – geographic structure – inducibility –*Pinus pinaster* – plant secondary metabolites (PSM) – Q_{ST}

Resumen

1. La variación fenotípica intraespecífica en metabolitos secundarios de plantas (MSP) se explica habitualmente por factores bióticos y abióticos. Sin embargo, los patrones de variación dentro de especies no se ajustan a las predicciones teóricas. Explorar cómo la genética, el ambiente y los procesos demográficos moldean dicha variación entre y dentro de las poblaciones es vital para comprender la evolución de los MSP, sobre todo en plantas de vida larga como los árboles.
2. En este trabajo cuantificamos la variación genética en MSP entre y dentro de poblaciones, y exploramos los factores causantes de adaptación local mediante el estudio del papel del clima como origen de la diferenciación de las poblaciones de pino marítimo en MSP. Se determinaron el perfil y concentración constitutiva de 63 MSP y su inducibilidad asociada a la herbivoría en la corteza de 130 genotipos replicados clonalmente, con estructura familiar conocida y procedentes de diez poblaciones, abarcando el rango de distribución de la especie. Comparamos la diferenciación genética de poblaciones a nivel neutral y cuantitativo de los MSP (F_{ST} y Q_{ST}). Además, consideramos la estructura genética entre poblaciones y el parentesco entre individuos al explorar las relaciones clima-carácter.
3. Encontramos una gran diferenciación entre poblaciones y variación genética aditiva en MSP constitutivos. Muchos MSP se indujeron, aunque observamos una variación genética limitada respecto a su inducibilidad. Las comparaciones Q_{ST}-F_{ST} sugieren que la diferenciación de muchos diterpenos, monoterpenos y fenoles pueden ser explicadas por procesos demográficos neutrales. Encontramos selección espacialmente heterogénea a través de las poblaciones, conducentes a adaptación local, solamente para sesquiterpenos constitutivos y algunos MSP individuales. Tras considerar la estructura genética entre poblaciones, sólo la concentración constitutiva de dos sesquiterpenos que mostraron signos de selección diversificadora apareció significativamente relacionada con el clima, con menor concentración a lo largo de un gradiente climático asociado a una mayor productividad primaria.
4. Síntesis. Los patrones evolutivos de los MSP dependieron de su naturaleza química, siendo la diferenciación neutral la dominante en la mayoría de MSP. Se encontraron evidencias de adaptación local solamente para los sesquiterpenos constitutivos y unos pocos MSP individuales. La escasa variación genética en la inducibilidad de los MSP sugiere un modelo de inducción defensiva conservado en esta especie. Debido a que la diferenciación entre poblaciones va unida a la historia demográfica y podría conducir a falsos positivos de señales de diferenciación adaptativa, es necesario tener en cuenta el parentesco genético entre poblaciones para inferir los determinantes ambientales de la variación genética intraespecífica en supuestos caracteres adaptativos como las defensas químicas.

Palabras clave

estructura geográfica – F_{ST} – gradientes ambientales – inducibilidad – metabolitos secundarios de plantas (MSP) – *Pinus pinaster* – Q_{ST} – variación genética

Introduction

Plant secondary metabolites (PSM) have received considerable attention from ecological and evolutionary perspectives, given their role as key modulators of plant interactions with other organisms (Agrawal 2010; Moore, Andrew, Kulheim & Foley 2013). During the last six decades, much effort has gone into explaining the reasons for the extreme interspecific phenotypic variation observed in PSM (Ehrlich & Raven 1964; Moles *et al.* 2011; Moreira *et al.* 2014; Anstett, Nunes, Baskett & Kotanen 2016). Abiotic and biotic stressors have both been recognized or postulated as drivers of interspecific differentiation of the defensive metabolome (Anstett *et al.* 2016; Mason *et al.* 2016). On the one hand, resource availability can influence the evolution of PSM by modulating the cost-benefit ratio of investing in plant defences. Theory predicts that resource-rich environments [i.e. warmer, wetter, with lower temperature seasonality, and higher availability of soil nutrients] would select for high growth rates at the expense of investing in plant defences, because tissue damage becomes less costly to replace [see for instance, Brenes-Arguedas, Coley and Kursar (2009); Pratt and Mooney (2013); Herms and Mattson (1992); Coley, Bryant and Chapin (1985)]. On the other hand, phenotypic variation in PSM may also arise from differences in biotic pressures across environments, with species living under high herbivore or pathogen pressure theoretically investing more in defences than species living in safer environments (Schemske, Mittelbach, Cornell, Sobel & Roy 2009). Existing evidence fairly supports these predictions, but contrasting responses have also been reported (Carmona, Lajeunesse & Johnson 2011; Moles *et al.* 2011; Anstett *et al.* 2016).

Within species, the phenotypic variation in profile and concentrations of PSM is also usually large, and is evident both among and within populations (Bernhardsson *et al.* 2013; Pratt & Mooney 2013; Agrawal *et al.* 2015). However, the extent to which abiotic and biotic drivers of interspecific differentiation in PSM can also explain intraspecific variation has received relatively less attention (Agrawal & Weber 2015; Hahn & Maron 2016). Moreover, the few available studies (see recent compilation by Hahn & Maron 2016) are mostly inconclusive. Such uncertainty may be due to the fact that theoretical predictions on the patterns of variation in plant defence were originally based on interspecific studies (see for instance, Coley *et al.* 1985) and are likely not valid for predicting within-species variation. In fact, at intraspecific scales, biotic pressure is often positively correlated with resource availability (Coley & Barone 1996; Brenes-Arguedas *et al.* 2009; Kooyers, Blackman & Holeski 2017); thus, the relative weight of biotic and abiotic factors in explaining the observed patterns may depend on the study system (Agrawal & Weber 2015). In addition, evolutionary factors such as demographic history and gene flow are often overlooked, in spite of expectations that they probably play a relevant role in intraspecific PSM variation (Bernhardsson *et al.* 2013; Pratt & Mooney 2013; Agrawal *et al.* 2015).

Another frequently overlooked component in plant defence theory is the phenotypic plasticity of PSM against biotic challenges, also known as inducibility (i.e. the ability to produce a plastic response to biotic challenges by modifying the concentration and profile of PSM). Inducibility constitutes a strategy that makes any plant defence more efficient against enemies, and thus is considered a selectable

trait (Agrawal 2005; Rasmann & Agrawal 2011). Current theory predicts that inter- and intraspecific variation in inducibility of plant defences should follow both abiotic and biotic environmental gradients (Hahn & Maron 2016). Defensive strategies based upon inducibility of defences are predicted to be favoured under low-resource availability, and in environments where herbivore pressure is low. However, very few studies have explored how patterns of population differentiation may also affect inducibility of plant defences (Rasmann *et al.* 2014; Więski & Pennings 2014; Agrawal *et al.* 2015), mostly showing inconclusive outcomes (Hahn & Maron 2016). This is especially true for long-lived plant species such as forest trees, for which induced defences are a particularly relevant strategy to cope with large temporal biotic and abiotic heterogeneity (Moore *et al.* 2013).

Exploring the genetic variation in inducibility of chemical defences is challenging. Inducibility of a given PSM should be determined as the difference between the basal constitutive level and the change due to the plastic response upon attack (Morris, Traw & Bergelson 2006; Moreira, Zas & Sampedro 2013). However, the mere fact of sampling plant tissues likely interfers with both measurements. In addition, for traits whose evaluation requires destructive sampling, consecutive measurements are impossible. Thus, genetic variation in inducibility of defences is usually studied by analysing the interaction between the genetic units (i.e. genotypes, families, populations) and the induction treatments (Bingham & Agrawal 2010; Sampedro, Moreira & Zas 2011a; Moreira *et al.* 2013). It follows that a proper quantitative genetic analysis of inducibility can only be achieved by using replicated clonal plants. To our knowledge, this approach has never been used for quantitative genetic studies in woody plants.

Pine trees are good model organisms to study intraspecific genetic variation of PSM in woody plants. They present a broad and diverse ensemble of carbon-based PSM, mainly terpenes and phenolics, with a well-known defensive role (Phillips & Croteau 1999; Witzell & Martín 2008). Terpenes and phenolics are C-based molecules that accumulate in pine tissues in large amounts (Sampedro, Moreira & Zas 2011b) and therefore impose a significant metabolic cost for their synthesis and accumulation (Mumm & Hilker 2006). Both terpenes and phenolics can be present constitutively or be induced once a triggering signal is perceived (Mumm & Hilker 2006; Eyles, Bonello, Ganley & Mohammed 2010). There is a large body of literature regarding the induction of defences in the genus *Pinus* (e.g. Heijari *et al.* 2005; Blodgett, Eyles & Bonello 2007; Moreira, Sampedro & Zas 2009; Sampedro *et al.* 2011b), but existing evidence accrues very little about the genetic basis of intraspecific PSM inducibility, and to what extent genetic variation might be structured and driven by geographical or environmental gradients, or is the product of demographic events.

In this study, we used maritime pine (*Pinus pinaster* Ait.) as a model to explore i) the inter- and intra- population genetic variation in constitutive PSM and their inducibility, ii) whether this genetic variation is structured across populations and to what extent population differentiation in PSM is driven by demographic neutral and/or adaptive processes, and iii) whether those adaptive processes acting on traits were associated to climatic conditions at the location of the populations. To do so, we conducted a common garden experiment using a hierarchical collection of ten populations covering the main distribution

range of the species, including families within populations and half-sibs within families. We produced clonal copies of each half-sib and we induced a defensive response in half of the plants with methyl jasmonate (MJ), a plant hormone involved in defensive signalling against chewing insects and necrotrophic fungi (Pieterse, Van der Does, Zamioudis, Leon-Reyes & Van Wees 2012). Single nucleotide polymorphisms (SNPs) were used to account for population structure resulting from demographic history and relatedness among individuals within and among populations, as they may generate spurious associations between quantitative traits and climate at the sites of origin of the populations. Based on the large within-population variation in other life-history traits (Santos-Del-Blanco, Climent, González-Martínez & Pannell 2012; Gaspar, Velasco, Feito, Alía & Majada 2013), we expected significant within-population genetic variation in PSM. Moreover, because of the contrasting abiotic environmental conditions across genetic clusters of populations (Serra-Varela et al. 2015), we hypothesised a large inter-population variation in both constitutive PSM and their inducibility, with populations coming from more stressful environments prioritizing constitutive defensive allocation over inducibility of PSM.

Materials and methods

Study system

Maritime pine is a model pine species present in a wide range of environments in southern Europe and northern Africa (Richardson & Rundel 1998) (Fig. 1). This species harbours large phenotypic variation for different life history traits across its distribution range (Tapias, Climent, Pardos & Gil 2004; Arrabal, Cortijo, De Simón, García Vallejo & Cadahía 2005; Santos-Del-Blanco *et al.* 2012), and large neutral genetic variation with strong spatial population structure (Vendramin, Anzidei, Madaghiele & Bucci 1998; Bucci *et al.* 2007; Jaramillo-Correa *et al.* 2015). Large genetic variation for several adaptive traits was also found within populations (e.g. Sampedro, Moreira, Llusia, Penuelas & Zas 2010, for defences; Santos-Del-Blanco *et al.* 2012, for reproduction), as usually happens for outcrossing wind-pollinated conifers (Savolainen, Pyhäjärvi & Knürr 2007). Genetic differentiation among populations is also evident in the chemical profiles of PSM, such as terpenes (Baradat & Marpeau 1988; Arrabal *et al.* 2005) and phenolics (Meijon *et al.* 2016).

Maritime pine has also been extensively studied to evaluate the plasticity of plant secondary compounds in response to biotic stress, using either MJ, real herbivory, or both (Moreira *et al.* 2009; Sampedro *et al.* 2011b; Zas *et al.* 2014). However, prior research has focused on very few compounds and/or on just constitutive metabolite levels. Moreover, the evolutionary forces driving population differentiation in constitutive PSM and their inducibility remain unexplored.

Plant material and greenhouse-controlled conditions

Plant material comprised a hierarchical collection of ten populations selected along a latitudinal gradient across the natural range of the species, consisting in five families within each population, and three half-sibs within each family (see also López-Goldar *et al.* 2018) (comprising 130 half-sib genotypes in total) (Fig. 1). Each half-sib was clonally propagated by recurrent early cutting as described in Majada *et al.* (2011). When cuttings were 2-years old they were relocated to an environmentally controlled

greenhouse in the MBG-CSIC (Pontevedra, Spain), where pines were allowed to acclimatise until the start of the induction treatments. A detailed description of how this collection was obtained and cultured before the experiment can be found in the Supporting Information (Methods S1).

Experimental design, induction treatments, and sampling

Plants were distributed following a split-plot design replicated in 3 blocks, with population as the whole plot factor (ten populations) and the factorial combination of family and MJ as the split factor (3-5 families × 2 levels of induction treatment). The experiment comprised a total of 260 plants from 10 populations, with 3-5 families per population, 3 half-sibs per family, and 2 clonal replicates per half-sib, one of which was induced with MJ and the other remained as a control.

Before induction, plants allotted to MJ treatment were transferred to a separate chamber to avoid cross-contamination with control plants. Then, MJ plants were sprayed with a solution of 25 mM of methyl-jasmonate (95%, Sigma-Aldrich, #39270-7) in deionized water with 2.5% ethanol (v/v). Control plants were treated just with the carrier solution. Both solutions were sprayed on the foliage to runoff [10.1 ± 0.5 mL per plant, mean ± standard error (s.e.)]. One month after the induction treatments were applied, plants were destructively harvested by cutting the stem aboveground, following the block design. Then, the needles were separated from the stem and two 2-cm-long stem subsamples were immediately collected from the middle portion of the plant, flash-frozen in liquid nitrogen, and stored at -80ºC. Both MJ concentration and the time elapsed between MJ application and sampling were based on previous studies (Moreira *et al.* 2009; Sampedro *et al.* 2011a).

Chemical analysis

For each stem subsample, the phloem and cortex tissues (hereafter 'bark') were carefully separated from the xylem, and ground in liquid nitrogen for chemical characterization of the PSM concentrations and profile. One of the stem subsamples was extracted in hexane for analysis of monoterpenes, sesquiterpenes, and diterpenes by GC-MS and GC-FID (Sampedro *et al.* 2010, with modifications). The other subsample was extracted in methanol for the analysis of phenolic compounds by UHPLC-MS and UHPLC-DAD (Villari *et al.* 2012). A detailed description of the analytical procedures and data processing prior to statistical analyses is provided in the Supporting Information (Methods S2).

We identified 118 PSM that were classified in the following chemical groups: monoterpenes, sesquiterpenes, diterpenes, flavonoids, hydroxycinnamic acids (HCAs), hydroxybenzoic acids (HBAs), lignans, and eugenols (Table S1, Table S2). For the purpose of the current study and to account for adequate sample size for the quantitative estimation of genetic variation, we only considered those compounds present in ≥ 75% of the samples in at least one of the levels of the induction treatment (control or MJ). Sixty-three chemical compounds were eventually selected (Table S1, Table S2) and grouped in seven groups of PSM according to their chemical nature (hereafter 'total PSM') (see also 'Data processing' section in Methods S2).

Sources of variation in PSM

Genetic variation in the constitutive concentration of total and individual PSM

was analysed in control plants by fitting a mixed model using restricted maximum likelihood in PROC MIXED (SAS v9.4) (Littell, Milliken, Stroup, Wolfinger & Schabenberger 2006), with population (P) and families within populations [F(P)] as fixed factors, and blocks (B) and the B×P interaction (i.e. the whole plots of the split-plot design) as random factors.

A second mixed model, which included both control and induced plants, was used to test for the effect of MJ induction. The model included MJ, P, F(P) and the interactions between them [MJ×P and MJ×F(P)] as fixed factors, and B, B×P and the genotype [MJ×F(P)×B] as random factors. The genotype was included as a random factor to account for the dependence between the two clonal replicates within each block. The MJ×P and MJ×F(P) interactions are informative with respect to the variation of the effect of MJ across populations and families.

The clonal structure of our experiment allowed for an accurate evaluation of the genetic basis of inducibility. Inducibility of PSM was measured for each genotype by computing the difference between the concentration of PSM in the MJ-induced and that in the corresponding control clonal replicate. The resulting genotype-based estimates of inducibility were analysed with a mixed model equivalent to that used for analysing the variation in constitutive defences.

For all mixed models, variables were square root- or log-transformed whenever residuals of the model deviated from normality after Kolmogorov-Smirnov tests. Heterogeneous residual variance models across MJ, P and MJ×P factors were fitted for each variable when appropriate. In order to analyse the 63 compounds without inflating Type I error due to multiple testing, p-value adjustments were performed after each mixed model analyses using the False Discovery Rate (FDR) for $p < 0.05$ (Benjamini & Hochberg 1995).

A multivariate analysis of variance (MANOVA) was performed on the 63 PSM with the PROC GLM in SAS, using the Wilk's lambda statistic, in order to test for multivariate effects of the main effects of MJ, P, F(P), and the interactions MJ×P and MJ×F(P) on the whole chemical profile. The multivariate information was further explored by summarizing the overall chemical profiles of the populations across induction treatments with a principal component analysis (PCA) using PROC PRINCOMP in SAS. All variables were standardized (mean = 0 and variance = 1) before the PCA, and the first two PCs (hereafter 'chemical indices', 'PC1' and 'PC2') were extracted. Genetic variation in constitutive and MJ-induced concentrations of PSM and their inducibility, summarized by the two chemical indices, were evaluated by using the mixed models described above.

Neutral differentiation, isolation by distance and Q_{ST}-F_{ST} comparisons

The population neutral differentiation F_{ST} (Wright's F-statistics) and its 95% confidence interval were estimated using the program GDA v1.1 (Lewis & Zaykin 2001) for 1,745 SNPs from 153 unrelated individuals of the same 10 populations (8-26 individuals per population) [Jaramillo-Correa *et al.* (2015), see Methods S3]. Additionally, population pairwise F_{ST} was estimated with Arlequin v3.0 (Excoffier, Laval & Schneider 2005). Significance of the resulting genetic distances was tested by permutation analysis, while isolation by distance was tested by comparing genetic and geographical distance matrices through a Mantel test (see details in Methods S3). An

additional Mantel test was also used to estimate the correlation between genetic distance and chemical similarity matrices. Pairwise chemical similarities were estimated using the Euclidean distance between populations in the multivariate space derived from the PCA describe above.

To quantify the quantitative genetic differentiation for PSM, the Q_{ST} parameter for each chemical trait was estimated as the ratio between the amount of genetic variance among populations (σ^2_P) and the total genetic variance of the trait ($\sigma^2_P + 2 \cdot \sigma^2_A$ in diploid organisms, with σ^2_A as additive genetic variance). Standard errors and confidence intervals of Q_{ST} estimates were obtained by parametric bootstrap resampling (1,000 simulations) using the *bootMer* function of the *lme4* R package (Bates, Mächler, Bolker & Walker 2015). In order to test whether neutral variation, diversifying or stabilizing selection contributes to the differentiation in PSM among pine populations, a Q_{ST}–F_{ST} comparison analysis for each compound was performed using the *QstFstComp* package in R (Gilbert & Whitlock 2015) with 10,000 bootstraps, assuming neutrality for both the phenotypic trait and the genetic markers (i.e. the Q_{ST} equals the F_{ST}). Briefly, the method tests whether the observed Q_{ST}–F_{ST} difference for each trait is in the tail of the Q_{ST}–F_{ST} neutral null distribution. An observed Q_{ST}–F_{ST} difference in the lower tail is taken as evidence of spatially uniform stabilizing selection (i.e. $Q_{ST} < F_{ST}$), while a Q_{ST}–F_{ST} difference in the upper tail is taken as evidence of spatially diversifying selection on the trait (i.e. $Q_{ST} > F_{ST}$) (Fig. S4) (more details in Methods S3).

Relationships between climatic variables and variation in PSM

Average climate data at the location of each population were obtained for the period 1950-2000 from a regional climatic model (Gonzalo 2007) in the case of the seven Spanish populations, and from the Worldclim model (Hijmans, Cameron, Parra, Jones & Jarvis 2005) for the French and Moroccan populations. The model of Gonzalo (2007) is known to be more accurate in the Iberian Peninsula than Worldclim because it takes into consideration a denser network of meteorological stations (Jaramillo-Correa *et al.* 2015). The following climate variables were used: annual mean temperature (AMT), maximum temperature of the warmest month (MTWM), minimum temperature of the coldest month (MTCM), temperature seasonality (TS), annual precipitation (AP), and potential evapotranspiration [PET, calculated as described in Thorntwaite (1948)]. To minimize type I error in the correlation analyses, the 6 climatic variables were summarized into two main components by means of a PCA using PROC PRINCOMP in SAS. The two first PCs (hereafter 'climate indices') explained ~90% of the overall variation. Climate index 1 explained 53.4% of variance and was positively correlated with MTCM and AP, and inversely correlated with TS, suggesting a proxy of 'atlanticity' and 'mediterraneity' for higher and lower values of the index, respectively (Fig. S1a). Climate Index 2 (36.5% of variance) was positively correlated with AMT and PET (Fig. S1a).

To explore whether differentiation processes in PSM were due to adaptive responses to climate heterogeneity across the species, we performed across-populations correlations between those PSM that showed signs of diversifying selection (i.e. $Q_{ST} > F_{ST}$) and the climate indices. To avoid spurious correlations and false positives due

to the non-independence of the populations and individuals within the populations, the correlation analyses were carried out using population least square means corrected for population structure and relatedness among genotypes obtained with the mixed-model approach proposed by (Yu *et al.* 2006). To this end, the mixed models described above for constitutive concentrations and inducibility of PSM were fitted again including both the population structure matrix (*Q*) that assigns individuals to populations, and the kinship matrix (*K*) that describes the relatedness among individuals within and between populations. The *Q* and *K* matrices were constructed based on 126 SNPs successfully genotyped on an independent set of clonal replicates of the individual genotypes of the studied collection [see López-Goldar *et al.* (2018); also see methodological details in Methods S4]. Adjusted population least square means of each PSM in each defensive mode (i.e. constitutive and inducibility) were then used for the correlation analyses. Only genotypes with both phenotypic and genotypic data were used in these analyses.

Results

Neutral differentiation and population structure

The overall neutral genetic differentiation, F_{ST}, among populations was 0.1055 (with 95% confidence interval, CI = 0.1019-0.1090). All population pairwise F_{ST} were significantly different from zero, except two pairs that encompass the three Spanish Atlantic populations: Armayan (ARMY) and Cadavedo (CDVO), as well as Puerto de la Vega (PTOV) and Cadavedo (CDVO) (Table S3). Limited gene flow across space was evidenced by the Mantel test, which showed a significant and strong positive correlation between genetic and geographical distance matrices ($r = 0.75$; $P < 0.0001$, see Fig. S2). This isolation by distance pattern translated into strong population structure (Fig. 1).

Genetic variation in the constitutive concentration of PSM

Constitutive concentration of PSM showed large significant variation among populations for all total PSM groups, except for total eugenols and total lignans (Table 1). Variation in the constitutive concentration of PSM was particularly large for monoterpenes (3.1 – 7.7 $mg \cdot g^{-1}$ dry weight), sesquiterpenes (0.3 – 2.4 $mg \cdot g^{-1}$ dw), diterpenes (6.1 – 16.4 $mg \cdot g^{-1}$ dw), and flavonoids (0.3 – 1.7 $mg \cdot g^{-1}$ dw) (Fig. S3). Constitutive concentration of PSM groups also showed large variation among families within populations for total diterpenes, total HCAs and total lignans (Table 1). Considering the individual compounds, 42 out of 63 (67%) showed significant genetic variation among populations and 26 out of 63 (41%) showed significant genetic variation among families within populations (Table S4).

Total sesquiterpenes was the most differentiated group among populations, with a Q_{ST} significantly higher than the neutral F_{ST} (Table 1). Many individual sesquiterpenes also showed high Q_{ST} estimates, although only a few (α–copaene, germacrene D + phenethyl 2-methylbutyrate, α–muurolene, δ–cadinene and germacrene D–4–ol) showed Q_{ST} significantly higher than the neutral F_{ST} (Table S4, see also Fig. S4a).

Q_{ST} estimates were generally lower and more variable for individual monoterpenes, diterpenes, phenolics and fatty acids, with no significant differences between Q_{ST} and F_{ST} (Table S4, see also Fig. S4b). Departures from neutral variation were

only detected for the monoterpene β–pinene ($Q_{ST} > F_{ST}$), and for total HCA (Table 1), camphene and two individual phenolics [coumaric acid hexoside (Fig. S4c) and lignan xyloside derivative 2] ($Q_{ST} < F_{ST}$).

Induced response to methyl jasmonate (MJ)
MJ treatment had a significant effect on the concentration of all total PSM groups except total sesquiterpenes and total HCAs (Table S5). An overall increase in the concentration of most PSM groups was found after induction, except for flavonoids, which decreased (Table S5). Most individual monoterpenes and diterpenes responded significantly to induction (10 out of 14 and 8 out of 9, respectively), phenolics were fairly affected by MJ induction (11 out of 24) and we found almost no significant responses to MJ for individual sesquiterpenes (2 out of 14) (Table S5).

Genetic variation in the inducibility of PSM
No among-family genetic variation in inducibility was detected for any total PSM group. Among-population differentiation in inducibility was only significant for total eugenols (Table 1, Fig. S3f). When considering individual PSM, significant differences in inducibility among populations were detected for methyl eugenol, and among families within populations for linalool, germacrene D-4-ol and phenethyl isovalerate (Table S4). In all cases, Q_{ST} estimates for inducibility of PSM were fairly low, and did not differ significantly from neutral F_{ST} (Table 1, Table S4).

Multivariate analysis of constitutive concentrations of PSM and their inducibility
Significant genetic variation among and within pine populations was observed in the MANOVA of the PSM (Table S6). Variation of the whole profile of PSM was well summarized by PCA, with PC1 explained mostly by terpenes and a few phenolics, and PC2 mostly by phenolics and some terpenes (Fig. S5). Together, PC1 and PC2 accounted for 31.6% of the total variance. The populations clustered in the multivariate space defined by these two chemical indices in the constitutive state, following a marked geographical pattern, with three distinguishable groups: the Atlantic populations from northern Spain (PTOV, CDVO and ARMY) at one extreme, the southern edge represented by the Moroccan population TAMR at the opposite extreme, and all the other populations from Spain and France distributed in between (Fig. 2). Chemical distances among populations in the constitutive multivariate space paralleled neutral genetic distances among populations (Mantel test, $r = 0.57$, $P = 0.0002$).

Marked multivariate changes in the whole profile of PSM were also observed in response to MJ treatment, but no effects were detected for MJ×P and MJ×F(P) interactions at the multivariate level (Table S6). Most of the populations responded similarly to MJ-induction (Fig. 2), and the response was characterized by a general increase in the PC1 scores (terpenes) and a general decrease in PC2 scores (phenolics). Two populations (SCRI and PLEU), however, slightly deviated from this general pattern (Fig. 2).

Relationship between climate variables and constitutive and induced PSM
Without accounting for population structure (*Q* matrix) and kinship relationships (*K* matrix), population means of PSM that showed signs of diversifying selection appeared to follow marked climatic clines. Climate index 1 (precipitation and temperature stability) was a significant

predictor of the constitutive concentrations of total sesquiterpenes (Fig. 3), germacrene D + phenethyl 2–methylbutyrate, δ–cadinene and germacrene D–4–ol (Fig. S6). Populations from more Atlantic climates (higher values of climate index 1) tended to show greater concentrations of PSM in all cases.

However, after accounting for population genetic structure and relative kinship among individuals, the results were drastically different. The above reported climatic clines failed to hold for total constitutive sesquiterpenes (Fig. 3) and for most of the individual PSM (Fig. S6). The only significant climatic clines that emerged after correction were negative relationships between the constitutive concentration of α–muurolene and δ–cadinene with climate index 1 (Fig. S6d, e). Lower concentrations of these PSM were associated with growth-prone Atlantic climates, with higher precipitation and lower temperature seasonality. β–pinene and α–copaene showed no climatic clines for both uncorrected data and after accounting for population structure and relative kinship among individuals (Fig. S6a, b).

Discussion

Strong intraspecific genetic variation in constitutive PSM

We found large genetic variation among and within populations in the constitutive concentration of most PSM. Pine populations were grouped in three 'chemotypic' clusters, according to their constitutive multivariate PSM profile: North Atlantic Spain, the Moroccan region and all the remaining populations spanning Spain and France. With few exceptions, this clustering is fairly consistent with that reported for secondary metabolites (Baradat & Marpeau 1988), and to a lesser extent for neutral genetic markers in previous studies (Bucci *et al.* 2007; Jaramillo-Correa *et al.* 2015; Meijon *et al.* 2016). This similar spatial structure is supported by the positive correlation between chemical and neutral genetic distances. Moreover, quantitative trait differentiation among populations did not differ from neutral variation for most of the compounds (e.g. total monoterpenes and diterpenes, and most individual compounds). These results point out that demographic processes, rather than adaptation [as suggested in other studies using conifers (Baradat & Marpeau 1988; Keeling & Bohlmann 2006; Bohlmann 2008)], may be behind the phenotypic differentiation in PSM among populations,

Total sesquiterpenes, several individual sesquiterpenes, and β–pinene (a monoterpene) were the notable exceptions, since for these compounds quantitative genetic differentiation exceeded the neutral differentiation among populations. This result suggests that these PSM are experiencing effective spatially heterogeneous selection across populations (Whitlock 2008). Thus, abiotic and/or biotic environmental differences across the distribution range of the species seem to be exerting differential selection pressures on these chemical traits, contributing to their differentiation beyond the demographic processes associated with genetic drift and migration.

A contrasting case is that of total constitutive HCAs and several individual constitutive phenolics, which showed significantly lower quantitative trait differentiation than the neutral variation. These results suggest that stabilizing selection processes are operating on the constitutive accumulation of these compounds, with uniform selection pressures governing their evolution across

the whole distribution range (Whitlock 2008).

It should be noted, however, that the low precision in the estimation of the quantitative genetic differentiation (Q_{ST}) of PSM may be hampering the detection of deviations from the neutral genetic differentiation. This lack of precision is common to population genetic studies with a low number of populations and families (O'Hara & Merila 2005). Consequently, we cannot discard that some other PSM with no significant difference between quantitative and neutral differentiation could be, in fact, responding to natural selection.

Most individual compounds respond to MJ induction

Previous studies in pine and spruce species have demonstrated that MJ is an effective elicitor of chemical defences (e.g. Martin, Gerzhenson & Bohlmann 2003; Moreira *et al.* 2009), one that provides effective resistance to biotic attacks (e.g. Zas *et al.* 2014). In our study, responses to MJ were particularly strong for total monoterpenes, total diterpenes and total eugenols, but weak (and in many cases resulting in reductions) for sesquiterpenes and most phenolics. Consistently, most individual terpenes showed significant induction. The strongest response to MJ was found for the monoterpene linalool (> 60-fold increase), a volatile compound that has been shown to be highly inducible in pines and other conifers (Martin *et al.* 2003). This compound is known to be involved in attracting parasitoids that prey on lepidoptera eggs laid on host plants, and to directly reduce the activity of phytophagous insects in other plant species (Xiao *et al.* 2012). The concentration of most resin acids also sharply increased after MJ application, as previously reported elsewhere (Tomlin, Antonejevic, Alfaro & Borden 2000), which provides greater resistance against insect herbivores (López-Goldar *et al.* 2018).

About half of the individual phenolic compounds responded significantly to MJ, but displayed contrasting induction patterns, some with increased and others with decreased concentrations. This resulted in a lack of overall induction effect in the total concentration of phenolics in the phloem ($F_{1,73}$ = 0.0, P = 0.992), in agreement with previous studies (Moreira, Zas & Sampedro 2012; Villari *et al.* 2012). Smaller phenolic compounds may serve as precursors and are often incorporated into larger molecules (e.g. lignins) after induction; thus, it is possible that their concentrations decreases due to source-sink relationships (Witzell & Martín 2008). A phenolic compound particularly responsive to MJ was eugenol, mainly in its methylated form (> 15-fold increase). The role of eugenols has been scarcely studied in pines, but since they are involved as precursors in lignin biosynthesis pathways (Dudareva, Klempien, Muhlemann & Kaplan 2013), they might provide an enhanced source of substrate for lignin accumulation after biotic challenge. Moreover, these compounds have shown strong antifeedant activity against the pine-specialist herbivore *Hylobius abietis* (Borg-Karlson, Nordlander, Mudalige, Nordenhem & Unelius 2006).

Scarce intraspecific genetic variation in inducibility of PSM

This study contributes important advances in two areas: i) we provide an accurate quantification of the intraspecific genetic variation in the inducibility of chemical defences as well as ii) insights into its evolutionary basis. Our results indicated that genetic variation in the inducibility of PSM is mostly absent at both inter- and

intrapopulation scales. Total eugenols were the only total PSM group showing genetic variation in inducibility among populations, and only 4 out of the 63 individual PSM showed significant variation in inducibility among populations (methyl eugenol) and among families within populations (linalool, phenethyl isovalerate and germacrene D–4–ol). Previous results based on a larger set of families within a single population of maritime pine showed a similar large genetic variation in constitutive foliar volatile terpenes but reduced variation in their inducibility (Sampedro *et al.* 2010).

Additionally, the multivariate response to MJ was also similar across populations and families, with populations moving in a similar manner in the multivariate space from the constitutive to the MJ-induced defensive modes. Defensive chemical responses associated with the MJ-derived induction have been reported to be similar across phylogenetically distant pine species (Moreira, Sampedro, Zas & Pearse 2016). In light of our results, we can thus speculate that MJ-associated induction of maritime pine defences involves ancestral biosynthetic pathways or patterns of responses that are highly conserved. Alternatively, low variation among populations in inducibility of PSM may be indicative of stabilizing selection operating on these traits. However, the Q_{ST}–F_{ST} comparison suggested that this is not the case, even for eugenols, for which significant variation among populations was observed. Thus, our results of inducibility of PSM are consistent with a phylogenetically conserved and highly integrated pathways of responses to induction that became fixed prior to the expansion of the species after the last ice age bottlenecks (Gómez, Vendramin, González-Martínez & Alía 2005).

Climatic clines in PSM drastically changed after accounting for genetic relatedness

Populations largely differed in the constitutive concentrations of PSM and, without accounting for genetic relatedness, that variation showed marked climatic clines. Atlantic populations coming from more favourable environments for pine growth generally showed higher constitutive concentrations of PSM (e.g. total sesquiterpenes, total diterpenes and many individual PSM) than populations coming from more stressful Mediterranean environments. An initial glance at these patterns may lead the observer to infer that abiotic environmental variation across the natural range of the species is exerting selective pressures on constitutive PSM, as plant defence theory predicts (Hahn & Maron 2016). However, the phenotypic expression of the variation in PSM includes not only the adaptive variation that emerges from natural selection, but also the neutral variation that results from historical demographic processes (i.e. genetic drift and migration) (Desdevises, Legendre, Azouzi & Morand 2003). In maritime pine, neutral variation has been reported to be spatially structured as a result of past climatic and demographic changes (Bucci *et al.* 2007; Jaramillo-Correa *et al.* 2015). Hence, natural selection and neutral variation may be confounded when analysing climatic gradients in any putatively adaptive trait (Desdevises *et al.* 2003; Kooyers *et al.* 2017). This seems to be the case for most of the analysed constitutive PSM, as the Q_{ST}–F_{ST} analyses revealed that quantitative differentiation did not differ from neutral variation for most of them. Thus, for most PSM, the apparent associations between the phenotypic differentiation in PSM and climate are better explained by the current geographical patterns of neutral variation. In

addition, for those few PSM for which quantitative variation exceeded neutral variation (i.e. those subjected to diversifying selection), the significant and positive correlations between constitutive PSM and climatic variation disappeared completely after accounting for population structure and genetic relatedness among genotypes. Thus, when diversifying selection is operating, the results suggest that environmentally heterogeneous factors other than climate are the responsible of the adaptive differentiation processes. It is reasonable to hypothesize that biotic pressures could be among those factors. Several of the studied PSM that showed diversifying selection processes (α–copaene, germacrene D and germacrene D–4–ol, and β–pinene) have been reported to influence orientation, and to have antifeedant effect against insect herbivores (Petrakis, Roussis, Papadimitriou, Vagias & Tsitsimpikou 2005; Bede, Musser, Felton & Korth 2006; Lundborg, Nordlander, Bjorklund, Nordenhem & Borg-Karlson 2016; López-Goldar *et al.* 2018). Moreover, β–pinene is known to be highly responsive to biotic challenges (Sampedro *et al.* 2010), which is in agreement with biotic selective forces acting upon this PSM. However, inferring differences in the biotic environment across the distribution range of the species and along evolutionary time is highly complex and outside the scope of this study.

The sesquiterpenes α–muurolene and δ–cadinene were the only PSM subjected to diversifying selection that showed a significant climatic cline after accounting for population structure and genotypic relatedness. In both cases, the cline was negative (lower PSM concentrations in populations from more favourable environments) and drastically different from those obtained without corrections. These results indicate that besides the role of the singular demographic history of the species in the differentiation process of these two particular compounds, climate is exerting an important selective pressure, with unfavourable abiotic conditions selecting for greater concentration of these PSM. However, our findings should be interpreted with caution, as strong spatial population structure resulting from demographic histories parallels that of abiotic environmental variation in this species (Serra-Varela *et al.* 2015), making the separation between neutral and adaptive variation to climate difficult to disentangle (Desdevises *et al.* 2003; Xia *et al.* 2018).

Concluding remarks

Our findings show that the constitutive levels of most PSM in maritime pine are characterized by large and geographically structured genetic variation among populations, with diverse evolutionary forces acting on different types of compounds. For most of the analysed PSM, among-population variation paralleled neutral genetic variation. Thus, no evidence of adaptive processes was found, and neutral demographic processes alone could explain quantitative population differentiation for most PSM. Total and several individual sesquiterpenes and β–pinene were the only PSM whose population divergence in constitutive levels exceeded neutral differentiation. However, no adaptive climate clines were found for most of them, suggesting that other environmental factors not accounted for in this study (e.g. biotic pressure) might have regulated the differential adaptive evolution of these traits. The only exceptions were α–muurolene and δ–cadinene for which unfavourable climatic conditions selected for higher concentrations.

Despite finding strong and significant inducibility for most PSM groups, very low genetic variation in their inducibility was detected, and no deviations from the neutral pattern of variation among populations were found. These results suggest a fixed ancestral pattern of defensive plastic responses associated to MJ signalling, and no current signs of adaptive evolution for the inducibility of PSM.

Our results highlight the need for adequate quantitative genetic analysis when testing hypotheses about environmental gradients in plant defensive traits. As demonstrated here, failing to account for the genetic population structure and genotypic relatedness may lead to spurious and/or erroneous results. We recommend that all future comparative ecology studies of plant defences within species make a stronger effort to disentangle adaptive from neutral differentiation (Q_{ST}–F_{ST} analysis) and include the estimation of population structure and relative kinship among individuals (*Q* and *K* matrices) when analysing environmental gradients. This would allow to adequately infer whether environmental and/or biotic factors act as drivers of PSM selection. Otherwise, neutral population differentiation linked to past demographic history could lead to false positive detection of adaptive differentiation signals.

Acknowledgements

We thank Dr. Juan Majada (SERIDA, Asturias, Spain) and Dr. Ricardo Alía (INIA-CIFOR, Madrid, Spain) for conceiving, and providing the plant material; José Carlos Rodríguez and Santiago Rodal (Viveros Figueirido, NORFOR, Pontevedra, Spain) for plant culture; César Cendán, María Luz Pato, and Helena Pazó for helping with plant sampling and chemical analyses. Also thank to Jean-Christophe Cocuron for helping with UHPLC-DAD-MS analyses, Santiago C. González-Martínez for help with SNP selection, Carmen García Barriga for DNA extraction, and Diana Barba Egido for SNPs editing. We thank two anonymous reviewers for constructive suggestions to improve an earlier version of this paper. This research was supported by the Spanish Government (MINECO/FEDER) grants FENOPIN (AGL2012-40151-C03-01), FUTURPIN (AGL2015-68274-C03-02-R), AdapCon (CGL2011-30182-C02-01/02) and by the Xunta de Galicia-GAIN grant (IN607A2016/013). This project has received funding from the European Union's Horizon 2020 Programme for Research & Innovation under grant agreement n°773383. XLG received financial support from Fundación Pedro Barrié de la Maza, and under the FPI Grant programs (MINECO).

Author contributions

LS and RZ designed the experiment, provided the reagents, performed the sampling, helped with the statistical and GC analyses and the interpretation of the results and improved the different versions of the manuscript; PB provided the infrastructure for UHPLC analyses; PB and CV provided know-how for the UHPLC analyses of phenolic compounds and helped with the interpretation of the associated data; AKBK provided the infrastructure and background for GC-MS analyses of terpenoid compounds and helped with the interpretation of the associated data; DG led the production of the SNPs database, constructed the *Q* and *K* matrices, and provided expertise on Q_{ST}-F_{ST} analyses; XLG performed the sampling, all the chemical analyses, most of the statistical analyses, produced the results, wrote the first

draft with RZ, LS and DG, and led the improvement of the manuscript along the peer-review process. PB, CV, and AKBK contributed equally to the final version of the manuscript. The authors declare that there is no conflict of interest.

Data accessibility

SNP database is publicly available at Zenodo repository (DOI: 10.5281/zenodo.1445313).

References

Agrawal, A.A. (2005) Natural selection on common milkweed by a community of specialized insect herbivores. *Evolutionary Ecology Research,* **7,** 17.

Agrawal, A.A. (2010) Current trends in the evolutionary ecology of plant defence. *Functional Ecology,* **25,** 420-432. DOI: 10.1111/j.1365-2435.2010.01796.x.

Agrawal, A.A., Hastings, A.P., Bradburd, G.S., Woods, E.C., Zust, T., Harvey, J.A. & Bukovinszky, T. (2015) Evolution of plant growth and defense in a continental introduction. *American Naturalist,* **186,** E1-E15. DOI: 10.1086/681622.

Agrawal, A.A. & Weber, M.G. (2015) On the study of plant defence and herbivory using comparative approaches: how important are secondary plant compounds. *Ecology Letters,* **18,** 985-991. DOI: 10.1111/ele.12482.

Anstett, D.N., Nunes, K.A., Baskett, C. & Kotanen, P.M. (2016) Sources of controversy surrounding latitudinal patterns in herbivory and defense. *Trends in Ecology & Evolution,* **31,** 789-802. DOI: 10.1016/j.tree.2016.07.011.

Arrabal, C., Cortijo, M., De Simón, B.F., García Vallejo, M.C. & Cadahía, E. (2005) Differentiation among five Spanish *Pinus pinaster* provenances based on its oleoresin terpenic composition. *Biochemical Systematics and Ecology,* **33,** 1007-1016. DOI: 10.1016/j.bse.2005.03.003.

Baradat, P. & Marpeau, A. (1988) Le pin maritime *Pinus pinaster* Ait : biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce. PhD Thesis, Université de Bordeaux I.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software,* **67**. DOI: 10.18637/jss.v067.i01.

Bede, J.C., Musser, R.O., Felton, G.W. & Korth, K.L. (2006) Caterpillar herbivory and salivary enzymes decrease transcript levels of Medicago truncatula genes encoding early enzymes in terpenoid biosynthesis. *Plant Molecular Biology,* **60,** 519-531. DOI: 10.1007/s11103-005-4923-y.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **57,** 289-300.

Bernhardsson, C., Robinson, K.M., Abreu, I.N., Jansson, S., Albrectsen, B.R. & Ingvarsson, P.K. (2013) Geographic structure in metabolome and herbivore community co-occurs with genetic structure in plant defence genes. *Ecology Letters,* **16,** 791-798. DOI: 10.1111/ele.12114.

Bingham, R.A. & Agrawal, A.A. (2010) Specificity and trade-offs in the induced plant defence of common milkweed *Asclepias syriaca* to two lepidopteran herbivores. *Journal of Ecology,* **98,** 1014-1022. DOI: 10.1111/j.1365-2745.2010.01681.x.

Blodgett, J.T., Eyles, A. & Bonello, P. (2007) Organ-dependent induction of systemic resistance and systemic susceptibility in *Pinus nigra* inoculated with *Sphaeropsis sapinea* and *Diplodia scrobiculata*. *Tree Physiology,* **27,** 511-517. DOI: 10.1093/treephys/27.4.511.

Bohlmann, J. (2008) Insect-induced terpenoid defenses in spruce. *Induced plant resistance to herbivory.* A. Schaller (Ed.). pp. 173–187. Springer Netherlands, Dordrecht, The Netherlands.

Borg-Karlson, A.K., Nordlander, G., Mudalige, A., Nordenhem, H. & Unelius, C.R. (2006) Antifeedants in the feces of the pine weevil *Hylobius abietis*: identification and biological activity. *Journal of Chemical Ecology,* **32,** 943-957. DOI: 10.1007/s10886-006-9050-y.

Brenes-Arguedas, T., Coley, P.D. & Kursar, T.A. (2009) Pests vs drought as determinants of plant distribution along a tropical rainfall gradient. *Ecology,* **90,** 1751-1761. DOI: 10.1890/08-1271.1.

Bucci, G., Gonzalez-Martinez, S.C., Le Provost, G., Plomion, C., Ribeiro, M.M., Sebastiani, F., Alia, R. & Vendramin, G.G. (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology,* **16,** 2137-2153. DOI: 10.1111/j.1365-294X.2007.03275.x.

Carmona, D., Lajeunesse, M.J. & Johnson, M.T.J. (2011) Plant traits that predict resistance to herbivores. *Functional Ecology,* **25,** 358-367. DOI: 10.1111/j.1365-2435.2010.01794.x.

Coley, P.D. & Barone, J.A. (1996) Herbivory and plant defenses in tropical forests. *Annual Review of Ecological Systems,* **27,** 305-335. DOI: 10.1146/annurev.ecolsys.27.1.305.

Coley, P.D., Bryant, J.P. & Chapin, F.S. (1985) Resource availability and plant antiherbivore defense. *Science,* **230,** 895-899. DOI: 10.1126/science.230.4728.895.

Desdevises, Y., Legendre, P., Azouzi, L. & Morand, S. (2003) Quantifying phylogenetically structured environmental variation. *Evolution,* **57,** 2647-2652. DOI: 10.1111/j.0014-3820.2003.tb01508.x.

Dudareva, N., Klempien, A., Muhlemann, J.K. & Kaplan, I. (2013) Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist,* **198,** 16-32. DOI: 10.1111/nph.12145.

Ehrlich, P.R. & Raven, P.H. (1964) Butterflies and plants: a study of coevolution. *Evolution,* **18,** 586-608. DOI: 10.1111/j.1558-5646.1964.tb01674.x.

EUFORGEN (2009) Distribution map of Maritime pine (*Pinus pinaster*) [WWW document].

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1,** 47-50. DOI: 10.1177/117693430500100003.

Eyles, A., Bonello, P., Ganley, R. & Mohammed, C. (2010) Induced resistance to pests and pathogens in trees. *New Phytologist,* **185,** 893-908. DOI: 10.1111/j.1469-8137.2009.03127.x.

Gaspar, M.J., Velasco, T., Feito, I., Alía, R. & Majada, J. (2013) Genetic variation of drought tolerance in *Pinus pinaster* at three hierarchical levels: A comparison of induced osmotic stress and field testing. *PLoS ONE,* **8,** e79094. DOI: 10.1371/journal.pone.0079094.

Gilbert, K.J. & Whitlock, M.C. (2015) Q_{ST}-F_{ST} comparisons with unbalanced half-sib designs. *Mol Ecol Resour,* **15,** 262-267. DOI: 10.1111/1755-0998.12303.

Gómez, A., Vendramin, G., González-Martínez, S.C. & Alía, R. (2005) Genetic diversity and differentiation of two Mediterranean pines along a latitudinal cline using chloroplast microsatellite markers. *Diversity and Distributions,* **11,** 257-263. DOI: 10.1111/j.1366-9516.2005.00152.x.

Gonzalo, J. (2007) Phytoclimatic analysis of the Spanish Peninsula: update and geostatistical analysis. PhD thesis. University of Valladolid.

Hahn, P.G. & Maron, J.L. (2016) A framework for predicting intraspecific variation in plant defense. *Trends in Ecology & Evolution,* **31,** 646-656. DOI: 10.1016/j.tree.2016.05.007.

Heijari, J., Nerg, A., Kainulainen, P., Viiri, H., Vuorinen, M. & Holopainen, J.K. (2005) Application of methyl jasmonate reduces growth but increases chemical defence and resistance againt *Hylobius abietis* in Scots pine seedlings. *Entomologia Experimentalis Et Applicata,* **115,** 117-124. DOI: 10.1111/j.1570-7458.2005.00263.x.

Herms, D.A. & Mattson, W.J. (1992) The dilemma of plants; to grow or to defend. *The Quarterly Review of Biology,* **67,** 54.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* **25,** 1965-1978. DOI: 10.1002/joc.1276.

Jaramillo-Correa, J.P., Rodriguez-Quilon, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Gere, P.H., Alia, R., Plomion, C., Vendramin, G.G. & Gonzalez-Martinez, S.C. (2015) Molecular proxies for climate maladaptation in a long-lived tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics,* **199,** 793-807. DOI: 10.1534/genetics.114.173252.

Keeling, C.I. & Bohlmann, J. (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytologist,* **170,** 657-675. DOI: 10.1111/j.1469-8137.2006.01716.x.

Kooyers, N.J., Blackman, B.K. & Holeski, L.M. (2017) Optimal defense theory explains deviations from latitudinal herbivory defense hypothesis. *Ecology,* **98,** 1036-1048. DOI: 10.1002/ecy.1731.

Lewis, P.O. & Zaykin, D. (2001) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R. & Schabenberger, O. (2006) *SAS System for Mixed Models,* 2nd edn. SAS Institute Inc., Cary, NC.

López-Goldar, X., Villari, C., Bonello, P., Borg-Karlson, A.K., Grivet, D., Zas, R. & Sampedro, L. (2018) Inducibility of plant secondary metabolites in the stem predicts genetic variation in resistance against a key insect herbivore in maritime pine. *Frontiers in Plant Science,* **9,** 1651. DOI: 10.3389/fpls.2018.01651.

Lundborg, L., Nordlander, G., Bjorklund, N., Nordenhem, H. & Borg-Karlson, A.K. (2016) Methyl Jasmonate-Induced Monoterpenes in Scots Pine and Norway Spruce Tissues Affect Pine Weevil Orientation. *Journal of Chemical Ecology,* **42,** 1237-1246. DOI: 10.1007/s10886-016-0790-z.

Majada, J., Martínez-Alonso, C., Feito, I., Kidelman, A., Aranda, I. & Alía, R. (2011) Minicuttings: an effective technique for the propagation of *Pinus pinaster* Ait. *New Forests,* **41,** 399-412. DOI: 10.1007/s11056-010-9232-x.

Martin, D., Gerzhenson, J. & Bohlmann, J. (2003) Induction of volatile terpene biosynthesis and diurnal emission by methyl jasmonate in foliage of Norway spruce. *Plant Physiology,* **132,** 1586-1599. DOI: 10.1104/pp.103.021196.

Mason, C.M., Bowsher, A.W., Crowell, B.L., Celoy, R.M., Tsai, C.-J. & Donovan, L.A. (2016) Macroevolution of leaf defenses and secondary metabolites across the genus *Helianthus*. *New Phytologist,* **209,** 1720-1733. DOI: 10.1111/nph.13749.

Meijon, M., Feito, I., Oravec, M., Delatorre, C., Weckwerth, W., Majada, J. & Valledor, L. (2016) Exploring natural variation of *Pinus pinaster* Aiton using metabolomics: Is it possible to identify the region of origin of a pine from its metabolites? *Molecular Ecology,* **25,** 959-976. DOI: 10.1111/mec.13525.

Moles, A.T., Wallis, I.R., Foley, W.J., Warton, D.I., Stegen, J.C., Bisigato, A.J., Cella-Pizarro, L., Clark, C.J., Cohen, P.S., Cornwell, W.K., Edwards, W., Ejrnaes, R., Gonzales-Ojeda, T., Graae, B.J., Hay, G., Lumbwe, F.C., Magana-Rodriguez, B., Moore, B.D., Peri, P.L., Poulsen, J.R., Veldtman, R., von Zeipel, H., Andrew, N.R., Boulter, S.L., Borer, E.T., Campon, F.F., Coll, M., Farji-Brener, A.G., De Gabriel, J., Jurado, E., Kyhn, L.A., Low, B., Mulder, C.P., Reardon-Smith, K., Rodriguez-Velazquez, J., Seabloom, E.W., Vesk, P.A., van Cauter, A., Waldram, M.S., Zheng, Z., Blendinger, P.G., Enquist, B.J., Facelli, J.M., Knight, T., Majer, J.D., Martinez-Ramos, M., McQuillan, P. & Prior, L.D. (2011) Putting plant resistance traits on the map: a test of the idea that plants are better defended at lower latitudes. *New Phytologist,* **191,** 777-788. DOI: 10.1111/j.1469-8137.2011.03732.x.

Moore, B.D., Andrew, R.L., Kulheim, C. & Foley, W.J. (2013) Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytologist*, **201,** 733-750. DOI: 10.1111/nph.12526.

Moreira, X., Mooney, K.A., Rasmann, S., Petry, W.K., Carrillo-Gavilán, A., Zas, R., Sampedro, L. & Novotny, V. (2014) Trade-offs between constitutive and induced defences drive geographical and climatic clines in pine chemical defences. *Ecology Letters*, **17,** 537-546. DOI: 10.1111/ele.12253.

Moreira, X., Sampedro, L. & Zas, R. (2009) Defensive responses of *Pinus pinaster* seedlings to exogenous application of methyl jasmonate: Concentration effect and systemic response. *Environmental and Experimental Botany*, **67,** 94-100. DOI: 10.1016/j.envexpbot.2009.05.015.

Moreira, X., Sampedro, L., Zas, R. & Pearse, I.S. (2016) Defensive Traits in Young Pine Trees Cluster into Two Divergent Syndromes Related to Early Growth Rate. *PLoS ONE*, **11,** e0152537. DOI: 10.1371/journal.pone.0152537.

Moreira, X., Zas, R. & Sampedro, L. (2012) Differential allocation of constitutive and induced chemical defenses in pine tree juveniles: a test of the optimal defense theory. *PLoS ONE*, **7,** e34006. DOI: 10.1371/journal.pone.0034006.

Moreira, X., Zas, R. & Sampedro, L. (2013) Additive genetic variation in resistance traits of an exotic pine species: little evidence for constraints on evolution of resistance against native herbivores. *Heredity*, **110,** 449-456. DOI: 10.1038/hdy.2012.108.

Morris, W.F., Traw, M.B. & Bergelson, J. (2006) On testing for a tradeoff between constitutive and induced resistance. *Oikos*, **112,** 102-110. DOI: 10.1111/j.0030-1299.2006.14253.x.

Mumm, R. & Hilker, M. (2006) Direct and indirect chemical defence of pine against folivorous insects. *Trends in Plant Science*, **11,** 351-358. DOI: 10.1016/j.tplants.2006.05.007.

O'Hara, R.B. & Merila, J. (2005) Bias and precision in Q_{ST} estimates: problems and some solutions. *Genetics*, **171,** 1331-1339. DOI: 10.1534/genetics.105.044545.

Petrakis, P.V., Roussis, V., Papadimitriou, D., Vagias, C. & Tsitsimpikou, C. (2005) The effect of terpenoid extracts from 15 pine species on the feeding behavioural sequence of the late instars of the pine processionary caterpillar *Thaumetopoea pityocampa*. *Behavioural Processes*, **69,** 303-322. DOI: 10.1016/j.beproc.2004.12.008.

Phillips, M.A. & Croteau, R.B. (1999) Resin-based defenses in conifers. *Trends in Plant Science*, **4,** 184-190. DOI: 10.1016/S1360-1385(99)01401-6.

Pieterse, C.M., Van der Does, D., Zamioudis, C., Leon-Reyes, A. & Van Wees, S.C. (2012) Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology*, **28,** 489-521. DOI: 10.1146/annurev-cellbio-092910-154055.

Pratt, J.D. & Mooney, K.A. (2013) Clinal adaptation and adaptive plasticity in *Artemisia californica*: implications for the response of a foundation species to predicted climate change. *Glob Chang Biol*, **19,** 2454-2466. DOI: 10.1111/gcb.12199.

Rasmann, S. & Agrawal, A.A. (2011) Latitudinal patterns in plant defense: evolution of cardenolides, their toxicity and induction following herbivory. *Ecology Letters*, **14,** 476-483. DOI: 10.1111/j.1461-0248.2011.01609.x.

Rasmann, S., Buri, A., Gallot-Lavallée, M., Joaquim, J., Purcell, J., Pellissier, L. & Heard, M. (2014) Differential allocation and deployment of direct and indirect defences by *Vicia sepium* along elevation gradients. *Journal of Ecology*, **102,** 930-938. DOI: 10.1111/1365-2745.12253.

Richardson, D.M. & Rundel, P.W. (1998) Ecology and biogeography of *Pinus*: an introduction. *Ecology and Biogeography of Pinus*. D.M. Richardson (Ed.). pp. 3-46. Cambridge University Press, Cambridge.

Sampedro, L., Moreira, X., Llusia, J., Penuelas, J. & Zas, R. (2010) Genetics, phosphorus availability, and herbivore-derived induction as sources of phenotypic variation of leaf volatile terpenes in a pine species. *Journal of Experimental Botany*, **61,** 4437-4447. DOI: 10.1093/jxb/erq246.

Sampedro, L., Moreira, X. & Zas, R. (2011a) Costs of constitutive and herbivore-induced chemical defences in pine trees emerge only under low nutrient availability. *Journal of Ecology,* **99,** 818-827. DOI: 10.1111/j.1365-2745.2011.01814.x.

Sampedro, L., Moreira, X. & Zas, R. (2011b) Resistance and response of *Pinus pinaster* seedlings to *Hylobius abietis* after induction with methyl jasmonate. *Plant Ecology,* **212,** 397-401. DOI: 10.1007/s11258-010-9830-x.

Santos-Del-Blanco, L., Climent, J., González-Martínez, S.C. & Pannell, J.R. (2012) Genetic differentiation for size at first reproduction through male versus female functions in the widespread Mediterranean tree *Pinus pinaster*. *Annals of Botany,* **110,** 1449-1460. DOI: 10.1093/aob/mcs210.

Savolainen, O., Pyhäjärvi, T. & Knürr, T. (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics,* **38,** 595-619. DOI: 10.1146/annurev.ecolsys.38.091206.095646.

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009) Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology, Evolution, and Systematics,* **40,** 245-269. DOI: 10.1146/annurev.ecolsys.39.110707.173430.

Serra-Varela, M.J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. & Zimmermann, N.E. (2015) Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography,* **24,** 1302-1313. DOI: 10.1111/geb.12369.

Tapias, R., Climent, J., Pardos, J.A. & Gil, L. (2004) Life histories of Mediterranean pines. *Plant Ecology,* **171,** 53-68. DOI: 10.1023/B:VEGE.0000029383.72609.f0.

Thorntwaite, C.W. (1948) An approach toward a rational classification of climate. *Geographical Review,* **38,** 55-94. DOI: 10.2307/210739.

Tomlin, E.S., Antonejevic, E., Alfaro, R.I. & Borden, J.H. (2000) Changes in volatile terpene and diterpene resin acid composition of resistant and susceptible white spruce leaders exposed to simulated white pine weevil damage. *Tree Physiology,* **20,** 1087-1095. DOI: 10.1093/treephys/20.16.1087.

Vendramin, G.G., Anzidei, M., Madaghiele, A. & Bucci, G. (1998) Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics,* **97,** 456-463. DOI: 10.1007/s001220050917.

Villari, C., Battisti, A., Chakraborty, S., Michelozzi, M., Bonello, P. & Faccoli, M. (2012) Nutritional and pathogenic fungi associated with the pine engraver beetle trigger comparable defenses in Scots pine. *Tree Physiol,* **32,** 867-879. DOI: 10.1093/treephys/tps056.

Whitlock, M.C. (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology,* **17,** 1885-1896. DOI: 10.1111/j.1365-294X.2008.03712.x.

Więski, K. & Pennings, S. (2014) Latitudinal variation in resistance and tolerance to herbivory of a salt marsh shrub. *Ecography,* **37,** 763-769. DOI: 10.1111/ecog.00498.

Witzell, J. & Martín, J.A. (2008) Phenolic metabolites in the resistance of northern forest trees to pathogens — past experiences and future prospects. *Canadian Journal of Forest Research,* **38,** 2711-2727. DOI: 10.1139/x08-112.

Xia, H., Wang, B., Zhao, W., Pan, J., Mao, J.-F. & Wang, X.-R. (2018) Combining mitochondrial and nuclear genome analyses to dissect the effects of colonization, environment, and geography on population structure in Pinus tabuliformis. *Evolutionary Applications*. DOI: 10.1111/eva.12697.

Xiao, Y., Wang, Q., Erb, M., Turlings, T.C., Ge, L., Hu, L., Li, J., Han, X., Zhang, T., Lu, J., Zhang, G. & Lou, Y. (2012) Specific herbivore-induced volatiles defend plants and determine insect community composition in the field. *Ecology Letters,* **15,** 1130-1139. DOI: 10.1111/j.1461-0248.2012.01835.x.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S. & Buckler, E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, **38,** 203-208. DOI: 10.1038/ng1702.

Zas, R., Bjorklund, N., Nordlander, G., Cendan, C., Hellqvist, C. & Sampedro, L. (2014) Exploiting jasmonate-induced responses for field protection of conifer seedlings against a major forest pest, *Hylobius abietis*. *Forest Ecology and Management*, **313,** 212-223. DOI: 10.1016/j.foreco.2013.11.014.

Table 1. Summary of the mixed models testing the effects of population (P) and family within population [F (P)] on the constitutive concentration of plant secondary metabolite (PSM) groups and their inducibility, including multivariate chemical indices, in the bark of maritime pine. Estimates of population differentiation for each variable are shown (mean bootstrap Q_{ST} ± s.e.). Q_{ST} estimates significantly higher or lower than the neutral F_{ST} are indicated by an asterisk followed by a sign denoting the direction of the difference (↑: $Q_{ST} > F_{ST}$ and ↓: $Q_{ST} < F_{ST}$). Significant effects in each defensive mode, after adjustment for multiple testing correction using false discovery rate [FDR for $P \leq 0.05$, Benjamini and Hochberg (1995)], are highlighted in bold font.

PSM groups	*Constitutive*			*Inducibility*		
	Population $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*	*Population* $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*
Terpenes						
Total monoterpenes	**5.2**	1.5	0.28±0.21	0.5	0.8	0.04±0.13
Total sesquiterpenes	**22.3**	1.4	0.47±0.20* ↑	0.2	0.9	0.05±0.17
Total diterpenes	**4.4**	**1.8**	0.22±0.17	0.6	1.3	0.04±0.10
Phenolics						
Total flavonoids	**9.1**	1.7	0.09±0.11	3.5	1.3	0.03±0.07
Total HCAs	**5.7**	**2.4**	0.03±0.04* ↓	2.5	1.3	0.07±0.13
Total eugenols	2.1	1.4	0.04±0.10	**5.8**	1.6	0.17±0.16
Total lignans	1.9	**2.6**	0.03±0.05	1.0	1.5	0.03±0.11
Chemical index 1	**7.6**	**1.9**	0.30±0.20	0.3	1.0	0.04±0.14
Chemical index 2	**3.4**	**2.1**	0.13±0.11	2.0	1.3	0.13±0.05

HCAs: hydroxycinnamic acids.

† Denominator degrees of freedom for family: total eugenols = 50; total monoterpenes, sesquiterpenes and diterpenes = 55; total flavonoids, HCAs and lignans = 57.

Figures

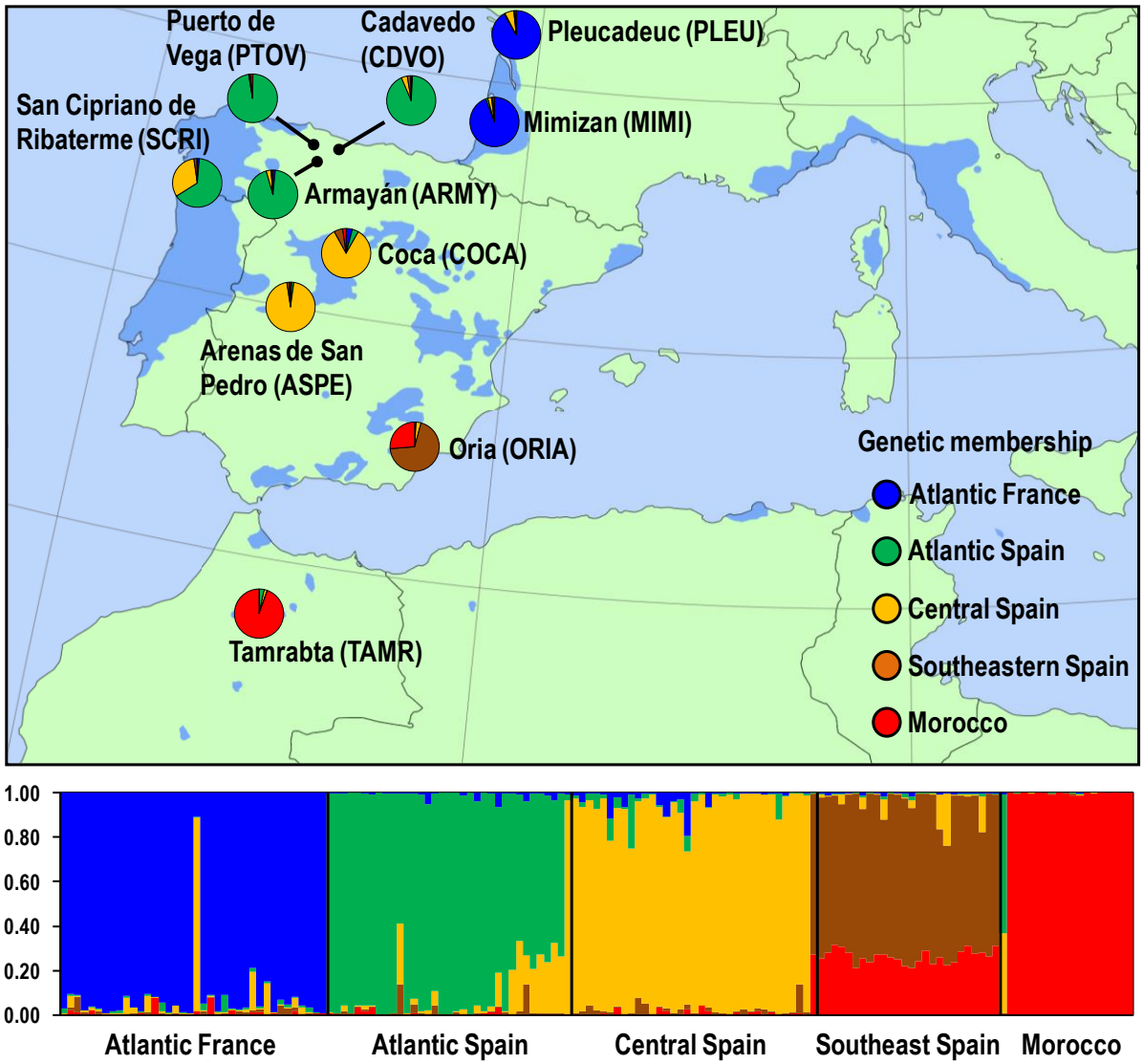

Fig. 1. Natural distribution of maritime pine (dark blue) and location of the ten studied populations (pie charts). Pie chart colours represent the average genetic membership of the populations to the neutral gene pools based on STRUCTURE analysis (bar plot below the geographic map) based on 1,745 SNPs genotyped in 153 randomly sampled and unrelated individuals (from Jaramillo-Correa *et al.* 2015). The four-letter population codes are used throughout the paper. Map modified from EUFORGEN (2009).

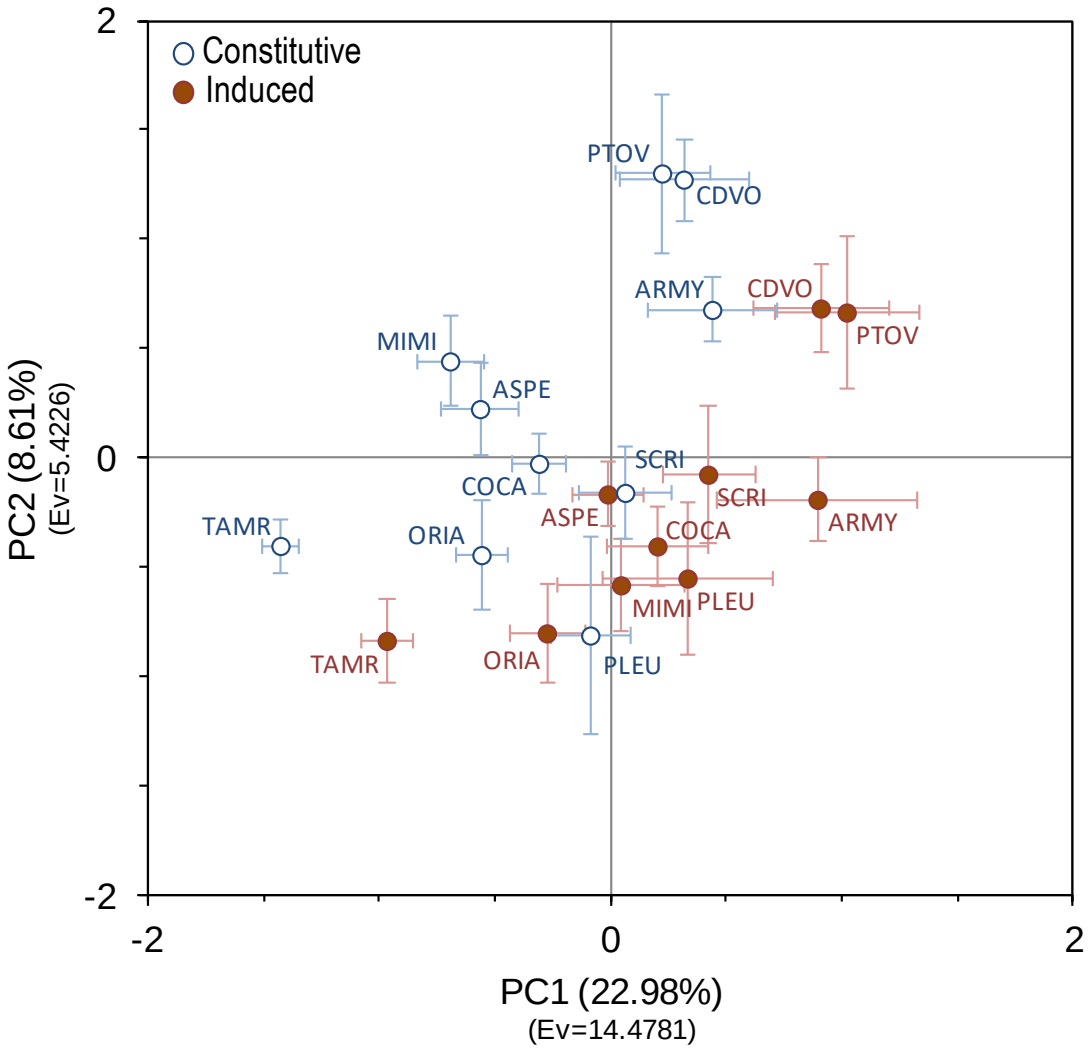

Fig. 2. Principal Component Analysis plot of the constitutive and induced concentration of 63 plant secondary metabolites in the bark of ten maritime pine populations grown in a common garden. PC1 was positively correlated with the concentration of most major terpenes and PC2 with those of most phenolics, some sesquiterpenes and a few minor monoterpenes (see Fig. S5). The corresponding explained variance (in %) and eigenvalues (Ev) for each PC are shown next to their axis in brackets. Each point represents the mean ± standard error of PC1 and PC2 for each population (N = 6 to 15 genotypes) in the multivariate space. Population codes are as described in Fig. 1.

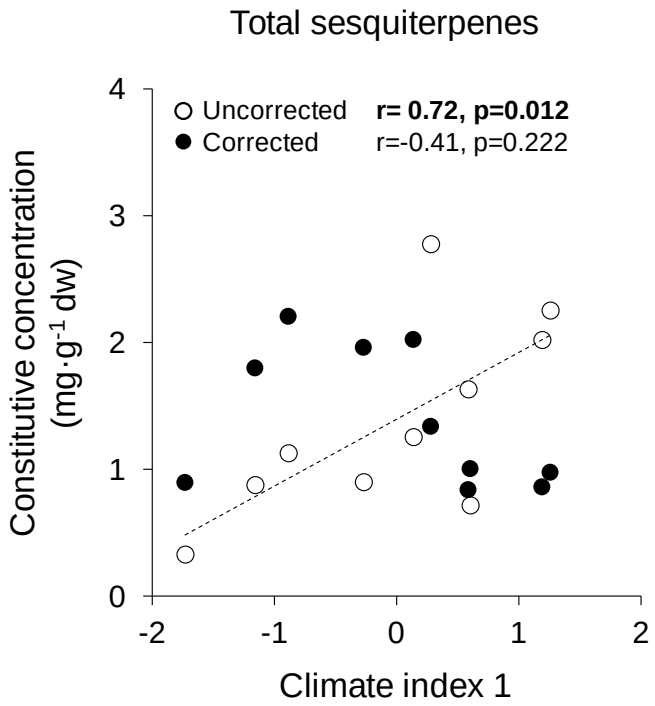

Fig. 3. Relationship between climate index 1 and the constitutive bark concentration of total sesquiterpenes, the only total PSM that showed quantitative genetic differentiation ($Q_{ST} > F_{ST}$), among 10 maritime pine populations. White circles represent uncorrected data whereas black circles represent the corrected population means after accounting for population structure and relative kinship among individuals in the corresponding mixed models. Higher values of climate index 1 correspond to more favourable environments (temperate, wetter and more isothermal); lower values of the index correspond to harsher environments (drier and with higher temperature seasonality). Each point represents the population least square means from uncorrected and corrected mixed models (N = 5–15 genotypes). For each series, significant *r* and *P* values ($P < 0.05$) are given in bold along with the trend line (dashed: uncorrected).

Supporting Information

Article title: Genetic variation in the constitutive defensive metabolome and its inducibility are geographically structured and largely determined by demographic processes in maritime pine.

Authors: López-Goldar, Xosé; Villari, Caterina; Bonello, Pierluigi; Borg-Karlson, Anna-Karin; Grivet, Delphine; Sampedro, Luis; Zas, Rafael.

The following Supporting Information is available for this article:

Methods S1. Selection, production and culture of the collection 'CLONAPIN Bank 1'.

Methods S2. Extraction, identification and quantification, and data processing of plant secondary metabolites (PSM).

Methods S3. Population neutral genetic differentiation, isolation by distance and Q_{ST}-F_{ST} comparisons.

Methods S4. Genotyping, construction and implementation of population structure (*Q*) and kinship (*K*) matrices in the mixed models using SNP data.

Fig. S1. Component loadings of the two first components extracted and multivariate distribution of the ten maritime pine populations using the six environmental variables from their locations of origin in a Principal Component Analysis.

Fig. S2. Plot of Mantel test showing isolation by distance among pairs of populations.

Fig. S3. Constitutive and methyl jasmonate-induced concentrations of total plant secondary metabolite groups in the bark of ten maritime pine populations.

Fig. S4. Examples of three simulated null distributions of $Q_{ST} - F_{ST}$ assuming neutrality of both the phenotypic trait and the molecular markers for three traits with contrasting quantitative differentiation.

Fig. S5. Component loadings of the two first components from the Principal Component Analysis of the constitutive and methyl jasmonate-induced concentration of 63 plant secondary metabolites in the bark of maritime pine.

Fig. S6. Relationships between climate index 1 and the constitutive bark concentrations of individual PSM that showed quantitative genetic differentiation among maritime pine populations.

Fig. S7. Plot of mean likelihood from STRUCTURE on 214 individuals genotyped for 126 SNPs.

Table S1. Identity of the 93 terpenoid compounds found in the bark of 2-year-old pine juveniles from ten natural populations of maritime pine by GC-MS analysis.

Table S2. Identity of the 25 phenolic compounds found in the bark of ten natural populations of 2-years old maritime pine juveniles by UHPLC-DAD-MS analysis.

Table S3. Population pairwise F_{ST}.

Table S4. Summary of the mixed models testing the effects of population (P) and family within population [F(P)] on the constitutive concentration of individual plant secondary metabolites (PSM) and their inducibility in the bark of 130 young maritime pine genotypes, and their estimates of population differentiation (mean bootstrap Q_{ST} ± s.e.).

Table S5. Effect of the methyl jasmonate induction (MJ) on the concentration of total and individual plant secondary metabolites (PSM) in the bark of ten maritime pine populations.

Table S6. Multivariate Analysis of Variance for the effect of methyl jasmonate induction (MJ), pine populations (P), families within populations [F(P)], as well as MJ×P and MJ×F(P) interactions on the concentration of plant secondary metabolites (PSM) in the bark of maritime pine.

Methods S1. Selection, production and culture of the collection 'CLONAPIN Bank 1'.

To obtain the collection, open-pollinated seeds were collected from five mother trees randomly selected in natural stands within each of the ten populations, spanning from France to Morocco (see Fig. 1), resulting in fifty half-sib families (i.e. known 'mother' and different 'fathers'). Selected mother trees were separated from each other by at least 50 m to minimize inbreeding (González-Martínez *et al.* 2006). Five seeds from each mother tree were sown and cultured in the facilities of SERIDA (Asturias, Spain), producing the complete ortet collection, which comprised 250 genotypes (10 populations $\times$ 5 families $\times$ 5 half-sibs) that were further clonally propagated as described in Majada *et al.* (2011). Then, three out of the five half-sibs per family (150 genotypes in total) were eventually used for this study. Due to low germination rate and/or rooting efficiency, only 130 out of the 150 genotypes were eventually included (belonging to 48 out of the original 50 families).

The cuttings were grown at SERIDA Agricultural Station (Grado, Asturias, Spain) in 200 mL containers containing peat and perlite (70:30 v/v), and later transported to Norfor Forest Nursery (Figueirido, Pontevedra, Spain) and transplanted to 2 L pots. In September 2012, 2-years old pine cuttings were transferred to an environmentally controlled greenhouse in the MBG-CSIC (Pontevedra, Spain), watered twice a week, fertilized with a slow release fertilizer (Granum, Soaga SL, Vilanova de Arousa, Spain, NPK 11-22-9) and supplemented with a foliar fertilizer (Fertimón, Progando SL, Galicia, Spain) until the start of the induction treatments in December 2012. During this period, mean daily temperatures and relative humidity in the greenhouse were 25.6 ± 0.06 ºC and 57.2 ± 0.08 % during the day and 18.2 ± 0.05 ºC and 69.0 ± 0.05 % during the night, respectively. Plants were maintained using a 12:12 hours of day:night photoperiod.

References:

González-Martínez, S.C., Burczyk, J., Nathan, R., Nanos, N., Gil, L. & Alía, R. (2006) Effective gene dispersal and female reproductive success in Mediterranean maritime pine (*Pinus pinaster* Aiton). *Molecular Ecology,* **15,** 4577-4588. DOI: 10.1111/j.1365-294X.2006.03118.x.

Majada, J., Martínez-Alonso, C., Feito, I., Kidelman, A., Aranda, I. & Alía, R. (2011) Minicuttings: an effective technique for the propagation of *Pinus pinaster* Ait. *New Forests,* **41,** 399-412. DOI: 10.1007/s11056-010-9232-x.

Methods S2. Extraction, identification and quantification, and data processing of plant secondary metabolites (PSM)

Sampling and extraction

One of the two flash-frozen stem subsamples from each individual plant was used for terpene analyses by gas chromatography-mass spectrometry (GC-MS) and gas chromatography-flame ionization detection (GC-FID), while the other stem subsample was utilized for phenolic analyses by ultra high performance liquid chromatography-mass spectrometry (UHPLC-MS) and UHPLC-diode array detection (UHPLC-DAD). Phloem was separated from the xylem by hand with a surgical knife on ice, cut into small pieces and ground in liquid nitrogen using mortar and pestle.

Extraction of terpenes was performed following Sampedro, Moreira, Llusia, Penuelas and Zas (2010) with modifications. Briefly, ca. 300 mg (fresh weight) of the ground tissue was extracted for 24 h in 1 ml hexane (HiperSolv Chromanorm #83992.320) in an ultrasonic bath at 25ºC using 0.05 $mg \cdot mL^{-1}$ pentadecane (Sigma-Aldrich, #76510) as internal standard. The extract was split into two vials, one used directly for analysis of volatile terpenes (mono and sesquiterpenes) whereas the other was dried under a flow of N_2 for the analysis of diterpene resin acids (hereafter 'resin acids'). The dried extract was diluted in methanol with heptadecanoic acid (Sigma-Aldrich, #H3500) as internal standard and tetramethylammonium hydroxide (TMAH, Sigma-Aldrich, #334901) was added as methylation agent in a proportion of 1:10 (TMAH:methanol, vol:vol). Volatile terpenes were identified and quantified by GC-MS, whereas resin acids (as their methyl esters) were identified by GC-MS and quantified by GC-FID. The pellet in the extraction vial was oven-dried and weighed, and results expressed in a dry weight basis.

Extraction of phenolics was performed as described in Villari *et al.* (2012), with modifications. Briefly, 100 mg fresh weight (fw) of the ground tissue was extracted twice (at 4 ºC in darkness for 24 h) in 500 µL of methanol (HiperSolv Chromanorm #152506X) containing 0.5 $mg \cdot mL^{-1}$ of resorcinol (Sigma-Aldrich, #398047) as internal standard. The extract was centrifuged at 16,000 rcf for 10 minutes each time, and the two supernatants were combined; the pellet in the extraction vial was oven-dried and weighed, and results were expressed in a dry weight basis. The non-polar resins in the combined extract were precipitated by adding 500 µL of deionized water (Milli-Q), followed by centrifugation at 16,000 rcf for 10 minutes. The supernatant was then transferred to another vial and the resin pellet was discarded. 800 µL of the clean extract were concentrated 8 times in a vacuum concentrator (Savant SPD2010 'Speedvac Concentrator', Thermo Scientific) and stored at -20 ºC until analysis.

Identification and quantification of volatile terpenes and resin acids

Identification and quantification of volatile terpenes were performed at KTH (Stockholm, Sweden) using a GC-MS in total ion count mode (TIC). The instrument used was a HP6890 GC equipped with a DB-5 capillary column (30 m, ID 0.25 mm, film thickness 0.25 µm, Agilent Technologies, CA, USA), coupled to a HP5973 mass spectrometer (Agilent Technologies, CA, USA) and using the G1701EA MSD ChemStation software (Agilent Technologies, CA, USA). A volume of 1 µl of each sample was injected in splitless mode,

using Helium as carrier gas. The oven temperature program was set at 40 ºC for 2 min, followed by a first temperature rise of 4 ºC·min^{-1} up to 200 ºC, then by a second temperature ramp of 10 ºC·min^{-1} up to 250 and maintained at this final temperature for 5 min. The injector temperature was set at 250 ºC.

The identification of each peak in the chromatogram was performed by comparing the retention times and mass spectra to that of available known standards (all from Fluka, Chemie AG, Buchs, Switzerland), to those in the NIST and Wiley Mass Spectral Libraries included in the analytical software, and from the correspondence between the calculated Kovat's Index (KI_C) from alkane series with those published in the literature (Adams 2007; KI_L) using the same column type.

Volatile compounds were quantified by using calibration curves prepared with seven available commercial authentic standards of those compounds present in the samples (Table S1). For those compounds whose standard was not available, relative quantification was done using the generated calibration curves from authentic standards of related compounds (Table S1). Internal standard was used if no related compounds were found. All calibration curves showed linear regressions with R^2>0.9999. Data handling for each chromatogram was performed as follows: peak integration was enabled between minutes 8 and 40, with a minimum detectable peak area of 500.000 area units, and peak width of 0.045 to avoid integrate noise spikes or badly integrate narrow peaks. Instrument calibration and consistency were evaluated every 40 samples by injecting a known concentration (100 µg/mL) of all terpene standards, check internal standard (50 µg/mL) and alkane series (~10 µg/mL) to ensure both peak signal variation and retention time shifts were under control during the same session. Variability of the calibration (measured as coefficient of variation) was below 12% for all standards.

Identification of resin acids were performed at KTH (Stockholm, Sweden) using the same GC-MS hardware and instrument parameters used for volatile terpenes. The oven program for resin acids was set at 152 ºC for 2 min, followed by a temperature ramp of 3 ºC·min^{-1} up to 260 ºC and maintained at this final temperature for 5 min. The identification of each present peak in the chromatogram was performed by comparing the retention times and mass spectra to those to available known standards (Sigma-Aldrich), to the NIST and Wiley Mass Spectral Libraries included in the analytical software and from the correspondence between the calculated Kovat's Index (KI_C) from alkane series with those published in the literature (Adams 2007; KI_L) using the same column type. Resin acids were quantified at Misión Biológica de Galicia (Pontevedra, Spain) using a GC-FID and TotalChrom Workstation v6.3.2 (Perkin Elmer, MA, USA) as analytical software. The instrument used was a Clarus 500 GC equipped with an Elite-5 capillary column (30 m, ID 0.25 mm, film thickness 0.25 µm, Perkin Elmer, MA, USA), coupled to a FID. All instrument parameters were configured identically as for previous GC-MS analysis. Hydrogen was used as carrier gas. FID temperature was set at 300ºC. Quantification of all resin acids was performed by preparing a calibration curve of authentic standard of abietic acid (Sigma-Aldrich) (Table S1). Individual compound concentration was expressed in mg·g^{-1} stem dry weight (dw). Data handling for each chromatogram was performed as follows: peak integration was enabled from minute 12 until the end of the run, with a minimum detectable peak area of 5.000 area units, and bunching factor of 1.0 to avoid integrate noise spikes or badly integrate narrow peaks. Instrument calibration and consistency was evaluated as described for volatile terpenes using abietic acid as external standard (100 µg/mL), check internal standard (50 µg/mL) and alkane series (µg/mL). Variability of the calibration (measured as coefficient of variation) was below 13% for all standards.

Identification and quantification of phenolic compounds by UHPLC-MS-DAD and UPLC-DAD

Identification and quantification of phenolic metabolites was performed following the procedure described by Raffa *et al.* (2017). Identification of phenolic metabolites was carried out at the Targeted Metabolomics Laboratory at the Ohio State University (Columbus, OH, USA) using a UHPLC 1290 (Agilent Technologies, CA, USA) coupled to a DAD 1260 (Agilent Technologies, CA, USA) in line with a hybrid Triple Quadrupole/Ion trap mass spectrometer 5500 (QTRAP, AB Sciex, MA, USA). DAD spectral data were recorded from 210 to 400 nm with phenolic compounds being detected at 280 nm. Mass spectrometry data were acquired in negative ion mode and processed using Analyst 1.6.1 software. The column used was an Acquity UPLC® BEH C_{18} (50 x 2.1 mm ID, 130Å, 1.7 µm particle size, Waters, MA, USA). Sample and column temperatures were set to 24 ºC and 50 ºC, respectively. The binary mobile phase consisted of 0.1% acetic acid in water (solvent A), and 0.1% acetic acid in methanol (Solvent B), with a constant flow rate of 0.42 ml·min^{-1}. The following linear gradient (cumulative run time (min), % solvent A) was used: 0.0, 93; 4.5, 85; 10.0, 70; 13.0, 10; 15.0, 0; 16.5, 0; 17.0, 93; 17.5, 93; 21.0, 93 (total run time 21 min). Phenolic compounds were identified by overlaying full scan mass chromatograms and DAD chromatogram traces at 280 nm to match the retention times of $[M-H]^-$ parent ions to λ_{max} of individual compounds. Phenolic compounds were identified based on negative ion fragmentation pattern, congruence of λ_{max}, retention time based on standards and on the literature (see Table S2).

Quantification of phenolic compounds was performed at the Department of Plant Pathology, College of Food, Agricultural and Environmental Sciences at the Ohio State University (Columbus, OH, USA) using an Acquity UPLC® H-class coupled to a DAD (Waters, MA, USA). DAD spectral data were recorded from 210 to 400 nm with phenolic compounds being detected and quantitated at 280 nm using Empower v3.0 software (Waters, MA, USA). Column type, sample and column temperatures, solvents and linear solvent gradient were identical to those used in the identification phase. λ_{max} from DAD spectral data and retention time from UPLC® were then compared with those obtained from the identification phase to assign the corresponding identity to the compounds when possible.

Calibration curves were prepared for phenolic quantification with eleven commercially available authentic standards (Apin, UK; Extrasynthése, France; and Sigma-Aldrich, Germany) related to those compounds present in the samples (Table S2). For those identified compounds whose standard was not available, relative quantification was done using the generated calibration curves from authentic standards of closely-related compounds (Table S2). If no related compounds were available, compounds were quantified as internal standard equivalents. All calibration curves showed linear regressions with R^2 > 0.999. Those compounds that could not be tentatively identified were labelled as 'unknown' peaks (Unk P#) and quantified as internal standard equivalents (Table S2). Individual compound concentration was expressed in mg·g^{-1} stem dw. In order to avoid processing noise spikes or badly integrated narrow peaks, the following processing thresholds were adopted: minimum detectable peak area of 10.000 area units and minimum peak width of 10 seconds. To ensure both peak signal variation and retention time shifts were consistent during the same session, instrument calibration was evaluated every 20 samples by injecting a known concentration (100 µg/mL) of all phenolic standards and check internal standard (50 ng/µL). Variability of the calibration (measured as coefficient of variation) was below 7% for all standards.

Data processing prior to statistical analyses

7 pairs and 1 trio of terpenes were impossible to separate individually during the data handling and were treated as single compounds (Table S1). From the selected compounds used for statistical analyses, individual PSM were summed up in their corresponding chemical groups as total PSM [total monoterpenes, total sesquiterpenes, total diterpenes, total flavonoids, total hydroxycinnamic acids (HCAs), total hydroxybenzoic acids (HBAs), total lignans, total eugenols, and total fatty acids]. The last two total PSMs (eugenols and fatty acids) were found in the GC-MS analyses and were classified as phenolic compounds and fatty acids in Table S1, respectively. Because the group of total HBAs has only one representative PSM (vanillic acid hexoside, Table S2), it was not considered in the analysis of total PSM. Also, total fatty acids were excluded from the analyses of total PSM. Jasmonates and their coeluting peaks were not included in the analyses because they were found only in the MJ-plants as a result of the application of the induction treatment.

References:

Adams, R.P. (2007) Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry, 4th edn. Allured Publishing Corporation, Carol Stream, USA.

Raffa, K.F., Mason, C.J., Bonello, P., Cook, S., Erbilgin, N., Keefover-Ring, K., Klutsch, J.G., Villari, C. & Townsend, P.A. (2017) Defence syndromes in lodgepole - whitebark pine ecosystems relate to degree of historical exposure to mountain pine beetles. Plant Cell Environ, 40, 1791-1806. DOI: 10.1111/pce.12985.

Sampedro, L., Moreira, X., Llusia, J., Penuelas, J. & Zas, R. (2010) Genetics, phosphorus availability, and herbivore-derived induction as sources of phenotypic variation of leaf volatile terpenes in a pine species. Journal of Experimental Botany, 61, 4437-4447. DOI: 10.1093/jxb/erq246.

Villari, C., Battisti, A., Chakraborty, S., Michelozzi, M., Bonello, P. & Faccoli, M. (2012) Nutritional and pathogenic fungi associated with the pine engraver beetle trigger comparable defenses in Scots pine. Tree Physiol, 32, 867-879. DOI: 10.1093/treephys/tps056.

Methods S3. Population neutral genetic differentiation, isolation by distance and Q_{ST}-F_{ST} comparisons

In order to avoid biases due to effects of sampling close relatives (i.e. half-sib families within populations, as in our study) (Wang 2018), population neutral differentiation F_{ST} was estimated using 1,745 neutral SNP obtained from 153 randomly (unrelated) sampled individuals within the same populations as in our study (8-26 individuals per population) [data obtained from Jaramillo-Correa *et al.* (2015)]. Wright's F-statistics (F_{ST}) was estimated using the program GDA (Lewis & Zaykin 2001). This estimate corresponds to the neutral molecular analog of the previously computed estimates of genetic differentiation among populations for phenotypic traits (Q_{ST}). Confidence interval for F_{ST} was obtained by 1,000 bootstrap simulations in GDA with resampling over loci.

In order to explore the geographical structure of the neutral differentiation among populations, population pairwise F_{ST} was estimated with Arlequin (Excoffier, Laval & Schneider 2005). Significance of the genetic distances was tested by permuting the individuals between the populations with 1,000 permutations and significance level of α=0.05. Additionally, a Mantel test was performed with Genepop v. 4.7.0 (Rousset 2008) in order to test for isolation by distance among populations (see Fig. S2). Pairwise genetic distance matrix (F_{ST} /(1 - F_{ST})) and geographic distance matrix were used as input to test the null hypothesis that there is no spatial correlation between genetic samples, with 10,000 permutations of samples between geographical locations.

In order to test whether neutral variation, directional or stabilizing selection are contributing to the differentiation in PSM among pine populations, the Q_{ST} estimate for each trait was compared to the mean neutral F_{ST} estimated upon the 1,745 neutral SNPs from Jaramillo-Correa *et al.* (2015) (see above). Traits with Q_{ST} significantly higher than the mean are inferred to be subjected to spatially heterogeneous divergent selection while $Q_{ST} < F_{ST}$ would be indicative of stabilizing selection, and $Q_{ST} = F_{ST}$ would reflect neutral evolution associated to demographic processes (O'Hara & Merila 2005; Whitlock 2008). Q_{ST}–F_{ST} comparisons were done with the *QstFstComp* package in R (Gilbert & Whitlock 2015) with 10,000 simulations. This package is based on the method developed by Whitlock and Guillaume (2009), expanding it to half-sib designs with known mothers and different fathers, and allowing for unbalanced data sets, as occurred in the present study. By parametric simulations and bootstrap, the method predicts a null distribution of the quantity Q_{ST}–F_{ST} under the null hypothesis that both the quantitative trait and the neutral markers show neutral differentiation (i.e. the Q_{ST} equals the F_{ST}). To test for departures from this null hypothesis of neutral differentiation, the method test whether the observed Q_{ST}–F_{ST} is in the tail of the neutral null distribution. An observed Q_{ST}–F_{ST} difference in the lower tail is taken as evidence of spatially uniform stabilizing selection, while a Q_{ST}–F_{ST} difference in the upper tail is taken as evidence of spatially divergent selection on the trait (Fig. S4). This comparison is essential to rule out genetic drift as an alternative mechanism for phenotypic divergence among populations.

References:

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1,** 47-50. DOI: 10.1177/117693430500100003.

Gilbert, K.J. & Whitlock, M.C. (2015) Q_{ST}-F_{ST} comparisons with unbalanced half-sib designs. *Mol Ecol Resour,* **15,** 262-267. DOI: 10.1111/1755-0998.12303.

Jaramillo-Correa, J.P., Rodriguez-Quilon, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Gere, P.H., Alia, R., Plomion, C., Vendramin, G.G. & Gonzalez-Martinez, S.C. (2015) Molecular proxies for climate maladaptation in a long-lived tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics,* **199,** 793-807. DOI: 10.1534/genetics.114.173252.

Lewis, P.O. & Zaykin, D. (2001) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

O'Hara, R.B. & Merila, J. (2005) Bias and precision in Q_{ST} estimates: problems and some solutions. *Genetics,* **171,** 1331-1339. DOI: 10.1534/genetics.105.044545.

Rousset, F. (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour,* **8,** 103-106. DOI: 10.1111/j.1471-8286.2007.01931.x.

Wang, J. (2018) Effects of sampling close relatives on some elementary population genetics analyses. *Mol Ecol Resour,* **18,** 41-54. DOI: 10.1111/1755-0998.12708.

Whitlock, M.C. (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology,* **17,** 1885-1896. DOI: 10.1111/j.1365-294X.2008.03712.x.

Whitlock, M.C. & Guillaume, F. (2009) Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. *Genetics,* **183,** 1055-1063. DOI: 10.1534/genetics.108.099812.

Methods S4. Genotyping, construction and implementation of population structure (*Q*) and kinship (*K*) matrices in the mixed models using SNP data.

In order to estimate whether the correlations with climate are influenced by the population structure and relatedness among genotypes, the correlation analyses were carried out using population least square means corrected for population neutral structure (*Q* matrix) and relatedness among genotypes (*K* matrix) using the mixed-model approach proposed by (Yu *et al.* 2006). To this end, genotypic data of the same individuals used for quantitative trait analyses is needed.

We used publicly available SNP data (see below) from individuals of the same collection as in our study (López-Goldar *et al.* 2018). Briefly, an independent set of clonal replicates of the original 250 genotypes of the collection 'CLONAPIN-Bank 1' (see Methods S1) was genotyped for 200 single nucleotide polymorphisms (SNPs) selected from two SNP arrays developed in maritime pine (Chancerel *et al*. 2011; Plomion *et al*. 2016). The selected SNPs included 50 SNPs considered as neutral, 50 SNPs considered as adaptive based on previous studies, and 100 SNPs selected in genes involved in water stress, pathogen and wounding responses, and wood formation. Needles from 220 genotypes out of the original 250 available in CLONAPIN-Bank 1, coming from the 10 populations (16-26 individuals per population) were silicagel-dried and DNA was extracted using the DNeasy Plant Mini Kit (Qiagen). SNP genotyping was performed by LGC Genomics using the KASP™ genotyping technology (http://www.lgcgenomics.com; LGC, Teddington, UK). After SNP quality control, a total of 126 SNPs and 220 genotypes were available (DOI: 10.5281/zenodo.1445313) for further use.

Population structure matrix (*Q*) was built using the Bayesian cluster analysis available in STRUCTURE v. 2.3.4 (Pritchard, Stephens & Donnelly 2000) with the following parameters: admixture model on correlated allele frequencies; burn-in of 10,000 steps followed by 100,000 iterations; number of clusters (*K*) set from 1 to 10; 6 runs were performed for each *K*. The number of genetic groups (*K*) for each marker was explored following Pritchard and Wen (2003), using STRUCTURE HARVESTER (Earl & vonHoldt 2011), which plots the mean L(*K*) and its variance over the runs for each *K* value.

The curve plateaued at *K*=5 (see Fig. S7), and this optimal *K* was in agreement with that from a study analysing a higher sample size in terms of individuals, populations and SNPs (Jaramillo-Correa *et al.* 2015). Despite the SNPs used here included both neutral and putatively adaptive SNPs, the resulting population structure accurately resembled that obtained in the previous study using a higher sampling size (Jaramillo-Correa *et al.* 2015). These results suggest that the genetic signal obtained with the 126 SNPs reflected well the historical and demographic processes of the species, and consequently these SNPs can be used to build the *Q* and *K* matrices. Similarity across runs with the same *K* was calculated with CLUMPP (Jakobsson & Rosenberg 2007), and the resulting membership coefficients for each individual formed the *Q*-matrix.

The kinship matrix (*K*) was constructed from the SNP dataset based on Loiselle, Sork, Nason and Graham (1995) kinship coefficients using SPAGeDi (Hardy & Vekemans 2002). Negative genetic covariances between individuals (i.e., individuals that are less related than random individuals) were set to zero as in Yu *et al.* (2006).

Eight individual genotypes of the 220 previously genotyped (see above) showed inconsistencies between their resulting genetic membership in the *Q* and *K* data and their original labelling to population and family, and were consequently removed prior the analyses. Furthermore, four additional individual

genotypes from families not considered in the original design of the collection of 250 genotypes (see Methods S1) were not included in the analyses. Finally, six individuals were discarded because they displayed a high rate of missing data (more than 60%). Hence, *Q* and *K* matrices finally comprised 202 genotypes and are publicly available at Zenodo repository (DOI: 10.5281/zenodo.1445313). Given that in the quantitative trait analyses we used only 130 out of the original 250 genotypes of the collection (CLONAPIN Bank 1, see Methods S1), and that *Q* and *K* data were not available for all these 130 genotypes, only those genotypes sharing data of both matrices and the concentration of PSM were subsequently used in the analyses.

Q and *K* matrices were then incorporated to the mixed model resulting in the following expression:

$$y = X\boldsymbol{\beta} + Q\mathbf{v} + Z\mathbf{u} + \mathbf{e} \quad (1)$$

Equation (1) shows an extended version of a typical mixed model, where **y** is a vector of phenotypic observations, X**β** represents those fixed effects other than the population structure; **β** is a vector of fixed effects other than population group effects; **v** is a vector of population effects; Q is a matrix from STRUCTURE relating **y** to **v**; Z**u** represents the random effects; **u** is a vector of polygene background effects; **e** is a vector of residual effects; and *X* and *Z* are incidence matrices of ones and zeroes relating **y** to **β** and **u**, respectively. The variances of the random effects are assumed to be Var(**u**) = $2KV_g$, and Var(**e**) = RV_R, where *K* is an $n \times n$ matrix of relative kinship coefficients that define the degree of genetic covariance between a pair of individuals; *R* is an $n \times n$ matrix in which the off-diagonal elements are 0 and the diagonal elements are the reciprocal of the number of observations for which each phenotypic data point was obtained; V_g is the genetic variance; and V_R is the residual variance (Yu *et al.* 2006).

References

Chancerel, E., Lepoittevin, C., Provost, G.L., Lin, Y.C., Jaramillo-Correa, J.P., Eckert, A.J., Wegrzyn, J.L., Zelenika, D., Boland, A., Frigerio, J.M., Chaumeil, P., Garnier-Géré, P., Boury, C., Grivet, D., González-Martínez, S.C., Rouzé, P., Van de Peer, Y., Neale, D.B., Cervera, M., Kremer, A. & Plomion, C. (2011) Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. BMC Genomics, 12. DOI: 10.1186/1471-2164-12-368.

Earl, D.A. & vonHoldt, B.M. (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources,* **4,** 359-361. DOI: 10.1007/s12686-011-9548-7.

Hardy, O. & Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes,* **2,** 618-620. DOI: 10.1046/j.1471-8278.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics,* **23,** 1801-1806. DOI: 10.1093/bioinformatics/btm233.

Jaramillo-Correa, J.P., Rodriguez-Quilon, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Gere, P.H., Alia, R., Plomion, C., Vendramin, G.G. & Gonzalez-Martinez, S.C. (2015) Molecular proxies for climate maladaptation in a long-lived tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics,* **199,** 793-807. DOI: 10.1534/genetics.114.173252.

Loiselle, B., Sork, V., Nason, J. & Graham, C. (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany,* **82,** 1420-1425. DOI: 10.1002/j.1537-2197.1995.tb12679.x.

López-Goldar, X., Villari, C., Bonello, P., Borg-Karlson, A.K., Grivet, D., Zas, R. & Sampedro, L. (2018) Inducibility of plant secondary metabolites in the stem predicts genetic variation in resistance

against a key insect herbivore in maritime pine. *Frontiers in Plant Science,* **9,** 1651. DOI: 10.3389/fpls.2018.01651.

Plomion, C., Bartholome, J., Lesur, I., Boury, C., Rodriguez-Quilon, I., Lagraulet, H., Ehrenmann, F., Bouffier, L., Gion, J.M., Grivet, D., de Miguel, M., de Maria, N., Cervera, M.T., Bagnoli, F., Isik, F., Vendramin, G.G. & Gonzalez-Martinez, S.C. (2016) High-density SNP assay development for genetic analysis in maritime pine (Pinus pinaster). Mol Ecol Resour, 16, 574-587. DOI: 10.1111/1755-0998.12464.

Pritchard, J., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155,** 945-959.

Pritchard, J. & Wen, W. (2003) Documentation for STRUCTURE Software: Version 2. Chicago: University of Chicago press.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S. & Buckler, E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics,* **38,** 203-208. DOI: 10.1038/ng1702.

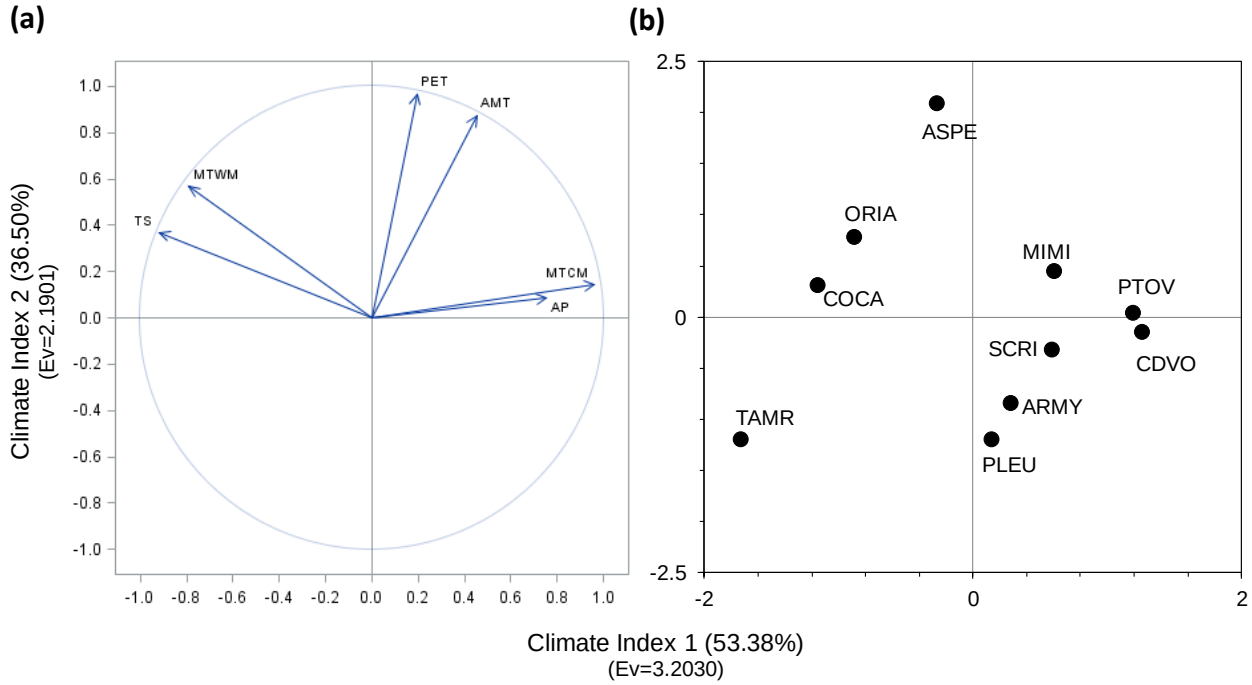

Fig. S1. Component loadings of the two first components extracted from a PCA summarizing the information of six bioclimatic variables at the site of origin of ten maritime pine populations (a), and plot of those populations in the multivariate space defined by PC1 and PC2 (b). The corresponding explained variance (in %) and eigenvalues (Ev) for each PC are also shown on each axis. Each arrow on the left panel represents the correlation between the variable and the two PCs. AP = annual precipitation; MTCM = minimum temperature of the coldest month; AMT = annual mean temperature; PET = evapotranspiration potential; MTWM = maximum temperature of the warmest month; TS = temperature seasonality. Climate index 1 is strongly and positively correlated with AP and MTCM and negatively correlated with MTWM and TS, whereas Climate index 2 is positively correlated with AMT and PET.

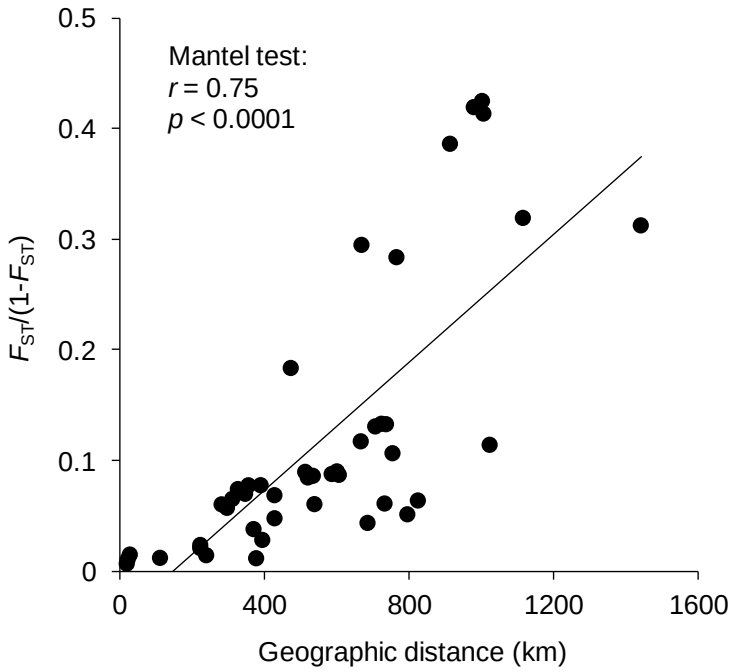

Fig. S2. Plot of Mantel test showing isolation by distance among pairs of populations, where pairwise genetic distance matrix (F_{ST} / (1 - F_{ST})) was regressed over geographic distance.

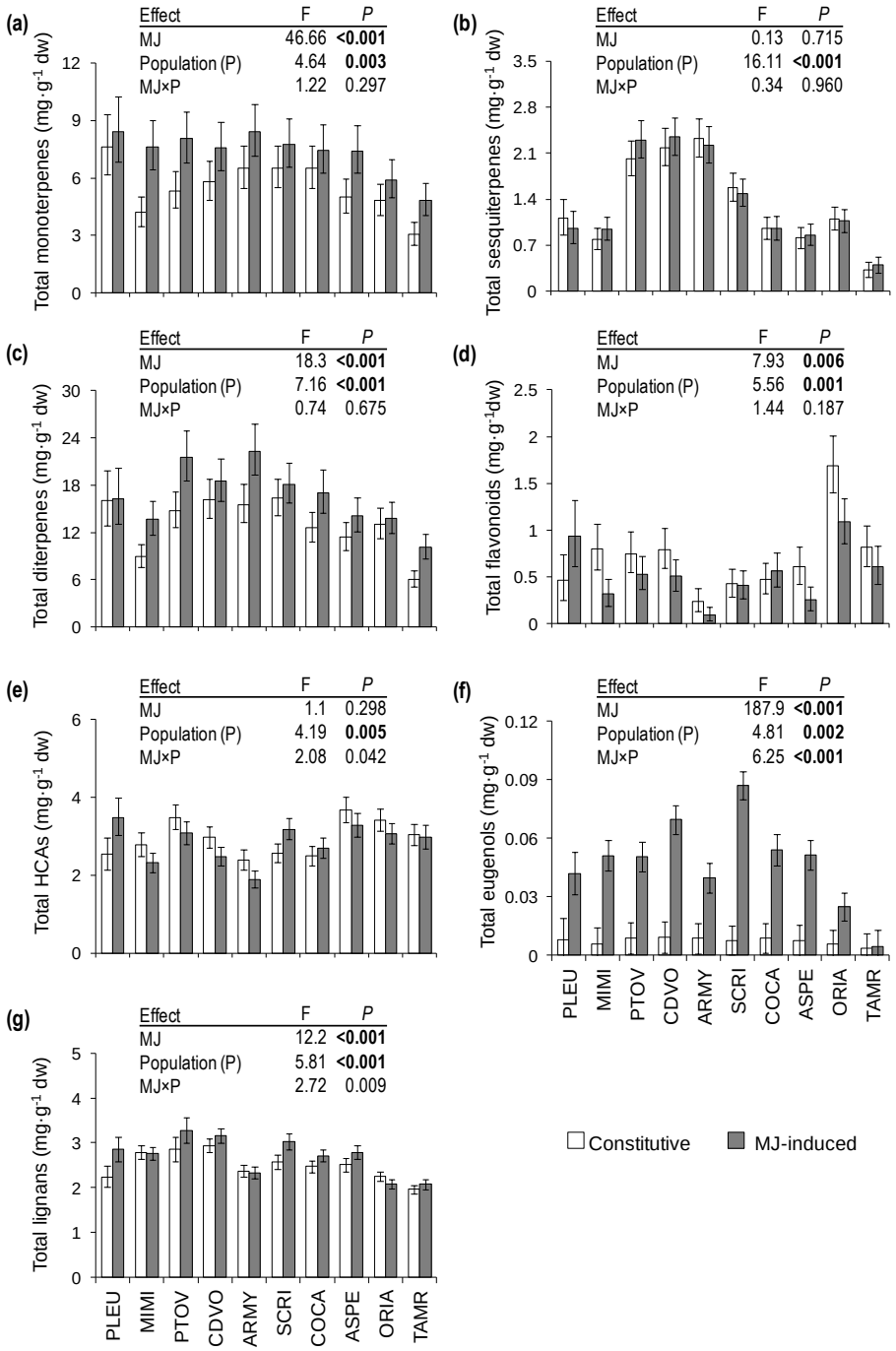

Fig. S3. Constitutive and methyl jasmonate-induced (MJ-induced) concentrations of total plant secondary metabolite (PSM) groups (panels a-c, terpenes; d-g, phenolics) in the bark of ten maritime pine populations, summarizing the within species genetic variation, grown in a common garden. Companion tables show the effects of the induction treatment with MJ, population (P) and their interaction (MJ×P) for each total PSM group. Significant p-values are highlighted in bold after adjustment using False Discovery Rate (FDR) for $P < 0.05$ (Benjamini & Hochberg 1995). Bars represent the least square mean of each of the induction treatments (N=6–15 genotypes per population). Error bars are the standard error of the mean. HCAs = hydroxycinnamic acids. Populations in the X axis are sorted from North (left) to South (right). See Fig. 1 for population codes.

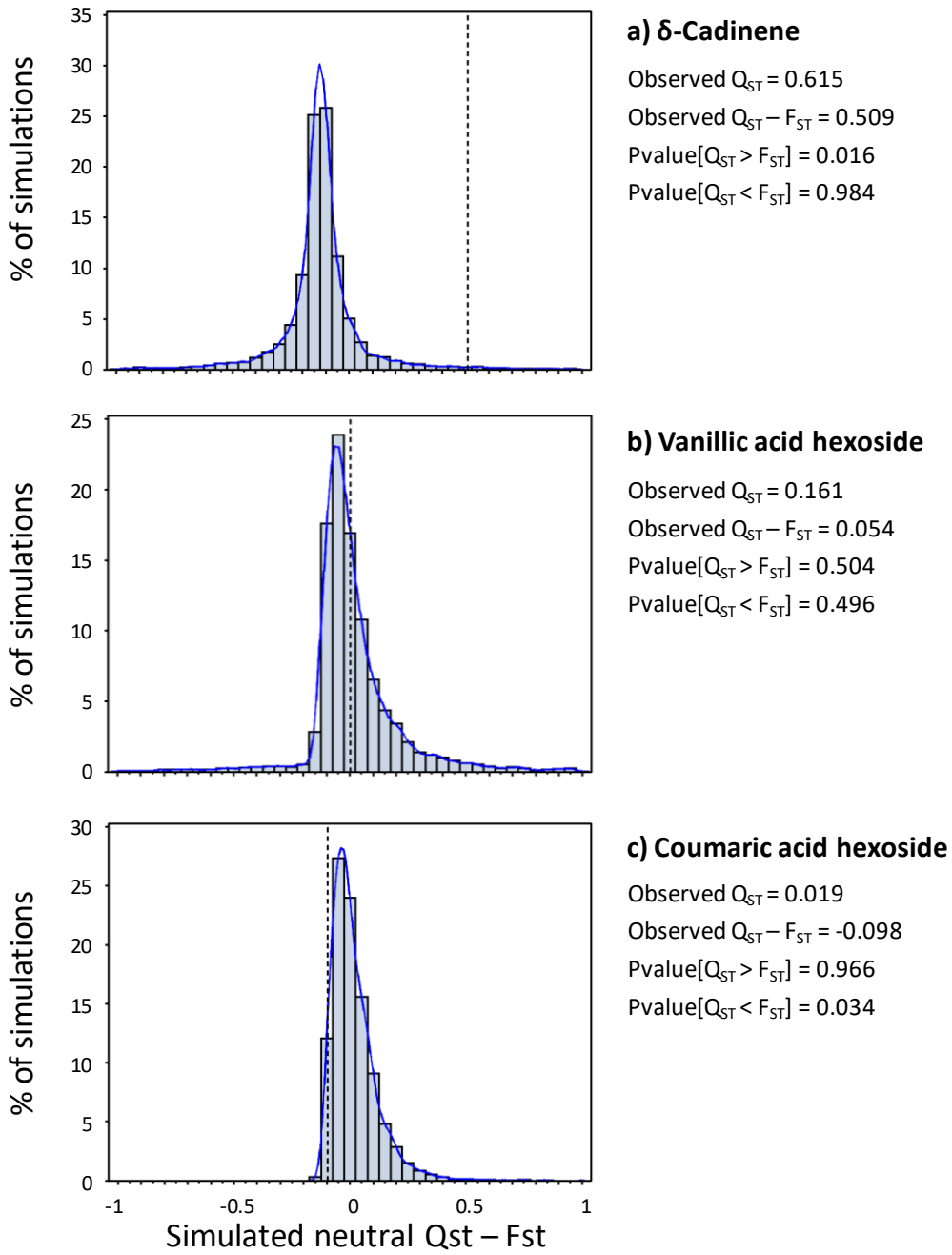

Fig. S4. Examples of three simulated null distribution of Q_{ST} - F_{ST} assuming neutrality of both the phenotypic trait and the molecular markers for three traits with contrasting quantitative differentiation: a) δ-Cadinene, with a Q_{ST} significant higher than the $F_{ST,}$ b) Vanillic acid hexoside, with no significant differences between Q_{ST} and F_{ST}, and c) Coumaric acid hexoside, with a Q_{ST} significant lower than the $F_{ST.}$ Observed Q_{ST}, and probability levels for testing the relative position of the observed Q_{ST} - F_{ST} difference (dotted vertical line) are shown next to each panel.

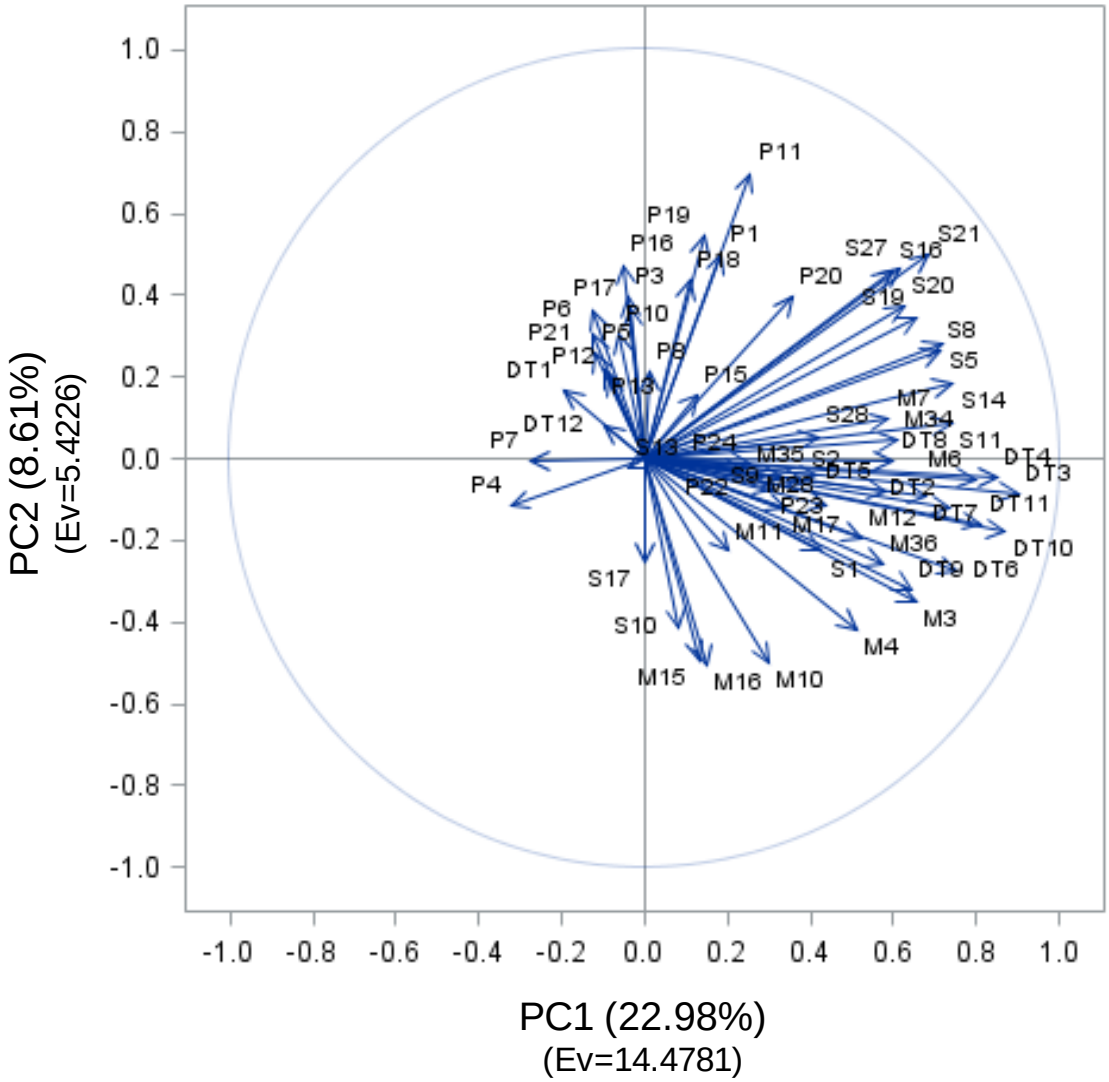

Fig. S5. Component loadings of the two first components from the Principal Component Analysis of the constitutive and methyl jasmonate-induced concentration of 63 plant secondary metabolites detected and quantified in the bark of maritime pine. The corresponding explained variance (in %) and eigenvalues (Ev) for each PC are also shown on each axis. Each arrow represents the correlation between the variable and the two corresponding components. M = monoterpene, S = sesquiterpene, DT = diterpene, P = phenolic. PC1 was strongly and positively correlated with most terpenes and negatively with some phenolics and fatty acids, whereas PC2 was positively correlated with most phenolics and some sesquiterpenes, and negatively correlated with some monoterpenes, sesquiterpenes and diterpenes.

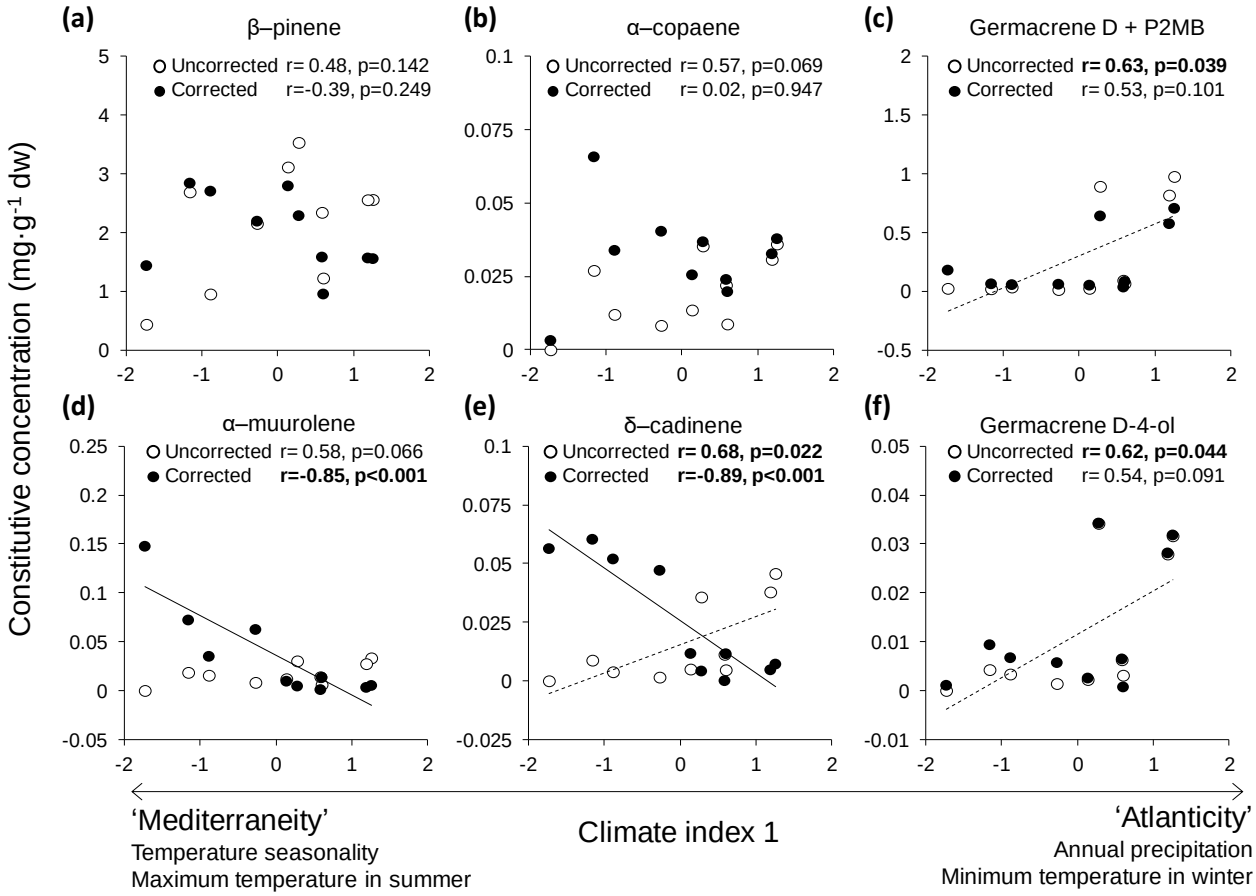

Fig. S6. Relationships between climate index 1 and the constitutive bark concentrations of individual PSM that showed quantitative genetic differentiation ($Q_{ST} > F_{ST}$) among 10 maritime pine populations. White circles represent uncorrected data whereas black circles represent the corrected population means after accounting for population structure and relative kinship among individuals in the corresponding mixed models. Higher values of climate index 1 correspond to more favourable environments (temperate, wetter and more isothermal); lower values of the index correspond to harsher environments (drier and with higher temperature seasonality). Each point represents the population least square means from uncorrected and corrected mixed models (N = 5–15 genotypes). For each series, significant *r* and *P* values ($P < 0.05$) are given in bold along with the trend line (dashed: uncorrected; solid: corrected). P2MB = Phenethyl 2-methylbutyrate.

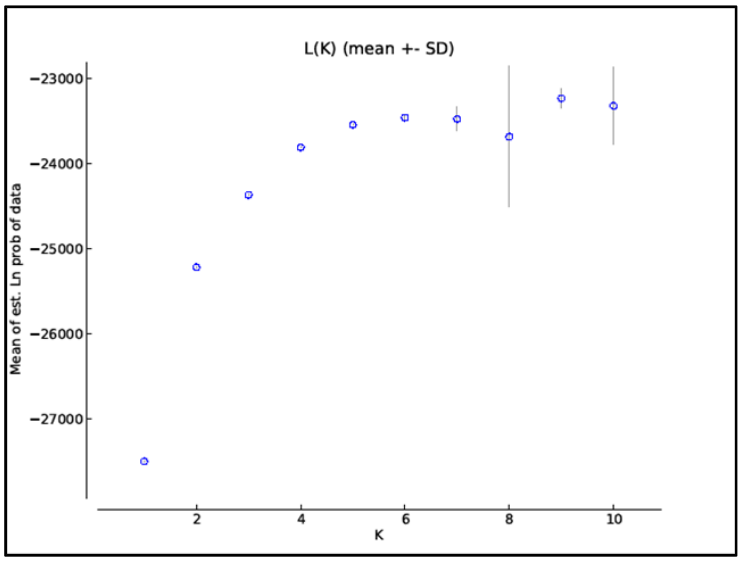

Fig. S7. Plot of mean likelihood L(*K*) from STRUCTURE on 214 (220 – 6 individuals with more than 60% of missing data) individuals genotyped for 126 single nucleotide polymorphisms (SNPs). Error bars represent the variance per *K* value.

Table S1. Identity of the 93 terpenoid compounds found in the bark of 2-year-old pine juveniles from ten natural populations of maritime pine by GC-MS analysis. Compounds are shown grouped by terpene chemical species and sorted by retention time (RT, in minutes). Compounds were identified by comparing the retention times and mass spectra to those from available known standards (Sigma-Aldrich), to the NIST and Wiley Mass Spectral Libraries and from the correspondence between the calculated Kovat's Index (KI_C) with those published in the literature (Adams, 2007; KI_L). Assignment of standard equivalent was based on similarity in retention time and chemical structure between the compound in the sample and the standard used for quantification. Compounds present in ≥ 75% of the samples are highlighted in bold and were subsequently used in the data analyses.

Code	Chemical compound	RT (min)	KI_C	KI_L	Standard equivalent
	Monoterpenes				
M1	Tricyclene	9.783	921	921	α-Pinene
M2	α-Thujene	9.985	926	924	α-Pinene
M3	**α-Pinene**	10.270	934	932	α-Pinene
M4	**Camphene**	10.754	947	946	β-Pinene
M5	Sabinene	11.694	973	969	β-Pinene
M6	**β-Pinene**	11.833	977	974	β-Pinene
M7	**Myrcene**	12.370	991	988	Limonene
M8	α-Phellandrene	12.833	1004	1002	α-Terpinene
M9	3-Carene	13.059	1010	1008	α-Pinene
M10	**α-Terpinene**	13.295	1016	1014	α-Terpinene
M11	**ρ-Cymene**	13.583	1024	1020	α-Terpinene
M12	**Limonene**	13.760	1028	1024	Limonene
	β-Phellandrene	13.760	1028	1025	
M13	*cis*-β-Ocimene	14.123	1038	1032	Limonene
M14	*trans*-β-Ocimene	14.502	1048	1044	Limonene
M15	**γ-Terpinene**	14.894	1059	1054	α-Terpinene
M16	**α-Terpinolene**	16.005	1088	1086	α-Terpinene
M17	**Linalool**	16.419	1099	1095	Pentadecane
M18	Solusterol	16.627	1105	1102	Pentadecane
M19	α-Campholene aldehyde	17.374	1125	1122	Pentadecane
M20	Norpinone	17.748	1136	1135	β-Pinene
M21	*trans*-pinocarveol	17.877	1139	1135	β-Pinene
	cis-Verbenol	17.946	1141	1137	
M22	*trans*-Verbenol	18.085	1145	1140	α-Pinene

Code	Chemical compound	RT (min)	KI_C	KI_L	Standard equivalent
M23	exo-methyl Camphenilol	18.214	1149	1145	α-Pinene
M24	*trans*-Pinocamphone	18.639	1160	1158	β-Pinene
M25	Pinocarvone	18.712	1162	1160	β-Pinene
M26	Borneol	18.873	1167	1165	α-Pinene
M27	Terpinen-4-ol	19.274	1178	1174	α-Terpinene
M28	**α-Terpineol**	19.759	1191	1186	α-Terpinene
M29	Myrtenol	19.968	1197	1194	α-Pinene
	Myrtenal	19.975	1197	1195	
M30	Nopol	20.216	1204	1278	α-Pinene
M31	Verbenone	20.379	1208	1204	β-Pinene
M32	*trans*-Carveol	20.758	1219	1215	Limonene
M33	Cuminic aldehyde	20.845	1222	1238	Limonene
M34	**Methyl thymol**	21.306	1235	1232	α-Terpinene
M35	**Piperitone**	21.997	1254	1249	Pentadecane
	Linalyl acetate	22.022	1256	1254	
M36	**Bornyl acetate**	23.100	1287	1284	α-Pinene
M37	*trans*-Pinocarvyl acetate	23.560	1291	1298	β-Pinene
M38	ρ-vinyl guaiacol [P]	24.009	1314	1309	Limonene
	Sesquiterpenes				
S1	**α-Cubebene**	25.303	1353	1345	Isolongifolene
	α-Longipinene	25.333	1355	1350	
S2	**Eugenol [P]**	25.455	1358	1356	Pentadecane
S3	Cyclosativene	25.894	1372	1369	Pentadecane
S4	α-Ylangene	26.011	1375	1373	Isolongifolene
S5	**α-Copaene**	26.157	1380	1374	Isolongifolene
S6	Geranyl acetate [ME]	26.279	1383	1379	Pentadecane
S7	β-Bourbonene	26.458	1389	1387	Pentadecane
S8	**β-Cubebene**	26.619	1394	1387	Isolongifolene
	Sativene	26.640	1394	1390	
	β-Elemene	26.663	1395	1389	
S9	**Methyl eugenol [P]**	26.959	1404	1403	Pentadecane
S10	**Longifolene**	27.153	1411	1407	Isolongifolene
S11	***trans*-β-Caryophyllene**	27.592	1425	1417	β-Caryophyllene
S12	β-Gurjunene	27.871	1434	1431	Isolongifolene
S13	***trans*-Isoeugenol [P]**	28.407	1451	1448	Pentadecane

Code	Chemical compound	RT (min)	KI_C	KI_L	Standard equivalent
S14	**α-Humulene**	28.642	1459	1452	α-Humulene
S15	α-Amorphene	29.334	1481	1483	Isolongifolene
S16	**Germacrene D**	29.525	1487	1484	β-Caryophyllene
	Phenethyl 2-methylbutyrate *	29.554	1488	1486	
S17	**Phenethyl isovalerate ***	29.687	1493	1490	Pentadecane
S18	Bicyclosesquiphellandrene	29.832	1497	-	Isolongifolene
S19	**α-Muurolene**	30.048	1504	1500	Isolongifolene
S20	**γ-Cadinene**	30.518	1520	1513	Isolongifolene
S21	**δ-Cadinene**	30.739	1528	1522	Isolongifolene
S22	Zonarene	30.816	1531	1528	Isolongifolene
S23	Cadina-1,4-diene	31.018	1537	1533	Isolongifolene
S24	α-Cadinene	31.173	1543	1537	Isolongifolene
S25	Elemol	31.488	1553	1548	Isolongifolene
S26	Citronellyl propionate [ME]	32.158	1576	-	Pentadecane
S27	**Germacrene D-4-ol**	32.298	1581	1574	β-Caryophyllene
S28	**Caryophyllene oxide**	32.531	1589	1582	β-Caryophyllene
S29	Longiborneol	32.957	1604	1599	β-Caryophyllene
	Diterpene resin acids and fatty acids				
DT1	**Oleic acid C18:1**	14.971	2105	-	Heptadecanoic acid
DT2	**Pimaric acid**	19.161	2244	2237	Abietic acid
DT3	**Sandaracopimaric acid**	19.707	2261	2256	Abietic acid
DT4	**Unk DT1**	20.623	2288	-	Abietic acid
DT5	**Isopimaric acid**	21.156	2303	2297	Abietic acid
DT6	**Levopimaric**	21.589	2314	2306	Abietic acid
	Palustric	21.589	2314	-	
DT7	**Arachidic acid C20:0**	22.281	2331	-	Heptadecanoic acid
	Unk DT2	22.281	2331	-	
DT8	**Dehydroabietic acid**	22.802	2344	2341	Abietic acid
DT9	**8,12-abietadien-18-oic acid**	23.492	2362	-	Abietic acid
DT10	**Abietic acid**	24.617	2390	2385	Abietic acid
DT11	**Neoabietic acid**	26.882	2442	2443	Abietic acid
DT12	**Behenic acid C22:0**	31.057	2536	-	Heptadecanoic acid
	Jasmonates				

Code	Chemical compound	RT (min)	KI_C	KI_L	Standard equivalent
S30	***trans*-Methyl jasmonate**	33.834	1635	-	*trans*-Methyl jasmonate
S31	***cis*-Methyl jasmonate**	34.265	1651	1648	*cis*-Methyl jasmonate
	α-Muurolol[†]	34.265	1652	1644	
S32	**Methyl epijasmonate**	35.089	1680	1678	Methyl epijasmonate

Compounds were coded according to the chemical analysis group where they were detected during the runs. M#, S# and DT# refer to monoterpene, sesquiterpene and diterpene chemical analysis groups, respectively. Compounds with no code were considered co-eluted with the previous peak in the table and treated hereafter as a single compound in the data analysis. RT, retention time (min); KI_C, calculated Kovat's Index using n-alkane series; KI_L, Kovat's Index extracted from the literature (Adams 2007) using the same column type, DB-5 (5%-phenyl methylpolysiloxane).

[†] α-Muurolol is a sesquiterpene, but coeluted with *cis*-methyl jasmonate during the GC-MS runs.

[P] Aromatic compound (phenylpropanoid). ρ-vinyl guaiacol, eugenol, methyl eugenol and *trans*-isoeugenol are phenolic compounds found in the GC-MS runs.

* Aromatic ester. Phenethyl 2-methylbutyrate and phenethyl isovalerate are not terpenes but are present as volatile compounds in pine species (Petrakis, Roussis, Papadimitriou, Vagias & Tsitsimpikou 2005).

[ME] Monoterpene ester.

References:

Adams, R.P. 2007. Identification of Essential oil components by Gas Chromatography/Mass Spectrometry. 4th Edition. Allured Publishing Corporation, Illinois, USA. 803 p.

Petrakis, P.V.; Roussis, V.; Papadimitriou, D.; Vagias, C.; Tsitsimpikou, C. 2005. The effect of terpenoid extracts from 15 pine species on the feeding behavioural sequence of the late instars of the pine processionary caterpillar *Thaumetopoea pityocampa*. Behavioural Processes, 69:303-322.

Table S2. Identity of the 25 phenolic compounds found in the bark of ten natural populations of 2-year-old maritime pine juveniles by UHPLC-DAD-MS analysis. Compounds are sorted by retention time (RT, in minutes). Tentative identification was based on *m/z*, MS fragmentation and UV absorbance maxima compared with those obtained from literature or standards when available. Assignment of standard equivalent was based on similarity in retention time and chemical structure between the compound in the sample and the standard used for quantification. Compounds present in ≥ 75% of the samples are highlighted in bold and were subsequently used in the data analyses.

Code	Assigned identity	RT (min)	UV λ max (nm)	$[M-H]^-$ (m/z)	Main ESI-MS fragments	References	Standard equivalent
P1	**Vanillic acid hexoside**	1.867	253.4, sh 290.9	329	167, 108, 152	c, f	Vanillic acid
P2	Procyanidin trimer	2.657	279.0	865	577, 407, 289, 425, 451, 125, 161	d	Procyanidin B2
P3	**Coumaric acid hexoside**	2.799	294.6	325	119, 163, 152	d, e, f	Coumaric acid
P4	**Coumaroylquinic acid**	3.066	310.7	337	163, 119, 191, 93, 155, 173	a, b, d	Coumaric acid
P5	**Taxifolin derivative 1**	3.583	285.4	465	286, 276, 125, 177, 304, 153	c, f	Taxifolin
P6	**Ferulic acid hexoside**	4.006	290.3, sh 316	(711) 355	193, 134, 178, 149, 119	e, f	Ferulic acid
P7	**Unk P1**	4.152	291.5				Resorcinol
P8	**Unk P2**	4.434	280.5				Resorcinol
P9	Taxifolin derivative 2^S	5.113	284.2				Taxifolin
P10	**Unk P3**	5.922	297.7				Resorcinol
P11	**Lignan hexoside derivative 1**	7.094	280.0	507	315, 327, 300, 345, 255	c, f	Pinoresinol
P12	**Lignan hexoside derivative 2**	7.388	280.0	507	315, 327, 300, 345, 255, 283	c, f	Pinoresinol
P13	**Unk P4**	7.436	287.8				Resorcinol
P14	Ferulic acid	7.676	321, sh 218				Ferulic acid
P15	**Unk P5**	7.749	288.5				Resorcinol
P16	**Lignan xyloside derivative 1**	7.910	278.0	495	167, 179, 327, 146, 121, 315, 221, 345	c, f	Pinoresinol
P17	**Lignan xyloside derivative 2**	8.063	278.0	495	167, 179, 165, 327, 221, 146,	c, f	Pinoresinol

Code	Assigned identity	RT (min)	UV λ max (nm)	$[M-H]^-$ (m/z)	Main ESI-MS fragments	References	Standard equivalent
					121		
P18	**Lignan xyloside derivative 3**	8.113	278.0	495	167, 179, 327, 221, 146, 121	c, f	Pinoresinol
P19	**Unk P6**	9.092	278.0				Resorcinol
P20	**Lignan deoxyhexoside**	9.637	280.0	491	315, 327, 300, 312	c, f	Pinoresinol
P21	**Unk P7**	10.447	263.3				Resorcinol
P22	**Unk P8**	12.758	263.3				Pinoresinol
P23	**Unk P9**	13.307	244.3				Resorcinol
P24	**Unk P10**	14.082	273.1				Pinocembrin
P25	Unk P11	14.287	277.4, sh 326.1				Resorcinol

sh: shoulder

S: tentative identity based on spectra.

Unk: unknown compound

References: a) Chen, Whitehill, Bonello and Poland (2011), J. Chem. Ecol.; b) Kammerer, Carle and Schieber (2004), Rapid. Comm. Mass. Spec.; c) Karonen *et al.* (2004), J. Agri. Food. Chem.; d) Lin and Harnly (2007), J. Agri. Food. Chem.; e) Pan and Lundgren (1996), Phytochem.; f) Wallis *et al.* (2011), For. Pathol.

References:

Chen, Y.; Whitehill, J.; Bonello, P.; Poland, T. 2011. Differential response in foliar chemistry of three ash species to emerald ash borer adult feeding. J. Chem. Ecol., 37: 29-39.

Kammerer, D.; Carle, R.; Schieber, A., 2004: Characterization of phenolic acids in black carrots (*Daucus carota* ssp *sativus* var. atrorubens Alef.) by high-performance liquid chromatography / electrospray ionization mass spectrometry. Rapid Comm. Mass Spec. 18, 1331–1340.

Karonen, M.; Hamalainen, M.; Nieminen, R.; Klika, K. D.; Loponen, J.; Ovcharenka, V. V.; Moilanen, E.; Pihlaja, K., 2004: Phenolic extractions from the bark of *Pinus sylvestris* L. and their effects on inflammatory mediators nitric oxide and prostaglandin E2. J. Agri. Food Chem. 52, 7532–7540.

Lin, L.; Harnly, J. 2007. A screening method for the identification of glycosylated flavonoids and other phenolic compounds using a standard analytical approach for all plant materials. J. Agric. Food Chem., 55: 1084-1096.

Pan, H.; Lundgren, L. N., 1996: Phenolics from the inner bark of *Pinus sylvestris*. Phytochemistry 42, 1185–1189.

Wallis, C., Eyles, A., Chorbadjian, R. A., Riedl, K., Schwartz, S., Hansen, R., Cipollini, D., Herms, D. A. and Bonello, P. 2011. Differential effects of nutrient availability on the secondary metabolism of Austrian pine (*Pinus nigra*) phloem and resistance to *Diplodia pinea*. Forest Pathology, 41: 52–58.

Table S3. Population pairwise F_{ST}. All the genetic distances were significantly greater than zero ($p < 0.05$) except that of CDVO-ARMY and PTOV-CDVO (not significant = NS).

Populations	**ARMY**	**ASPE**	**CDVO**	**COCA**	**MIMI**	**ORIA**	**PLEU**	**PTOV**	**SCRI**	**TAMR**
ARMY	0									
ASPE	0.069	0								
CDVO	0.010^{NS}	0.065	0							
COCA	0.057	0.011	0.054	0						
MIMI	0.078	0.057	0.082	0.045	0					
ORIA	0.115	0.072	0.117	0.064	0.105	0				
PLEU	0.080	0.048	0.080	0.041	0.011	0.102	0			
PTOV	0.014	0.072	0.006^{NS}	0.061	0.079	0.117	0.082	0		
SCRI	0.020	0.036	0.014	0.027	0.057	0.096	0.060	0.023	0	
TAMR	0.295	0.228	0.298	0.221	0.242	0.155	0.238	0.292	0.278	0

Table S4. Summary of the mixed models testing the effects of population (P) and family within population [F(P)] on the constitutive concentration of individual plant secondary metabolites (PSM) and their inducibility in the bark of 130 young maritime pine genotypes, and their estimates of population differentiation (mean bootstrap Q_{ST} ± s.e.). Q_{ST} estimates significantly higher or lower than the neutral F_{ST} are indicated by an asterisk followed by a sign denoting the direction of the difference (↑: $Q_{ST} > F_{ST}$ and ↓: $Q_{ST} < F_{ST}$). Significant F-values are given in bold after adjustment for multiple testing correction using false discovery rate (FDR) at p ≤ 0.05 in each defensive mode (Benjamini & Hochberg 1995).

PSM	*Constitutive*			*Inducibility*		
	Population	*Family*	*Q_{ST} (± s.e.)*	*Population*	*Family*	*Q_{ST} (± s.e.)*
	$F_{9,18}$	$F_{38,}$†		$F_{9,18}$	$F_{38,}$†	
Monoterpenes						
α-Pinene	1.7	1.2	0.13±0.19	0.4	0.7	0.03±0.11
Camphene	**4.3**	**2.7**	0.03±0.04* ↓	0.7	0.9	0.08±0.19
β-Pinene	**17**	1.4	0.49±0.21* ↑	0.6	0.8	0.05±0.15
Myrcene	**6.9**	1.1	0.44±0.24	0.9	1	0.05±0.14
α-Terpinene	**6.9**	1.7	0.11±0.11	2.5	1.8	0.09±0.11
ρ-Cymene	**5.4**	**2.7**	0.04±0.05	1.5	1.4	0.04±0.12
β-Phellandrene + Limonene	1.3	**2.1**	0.04±0.07	0.3	0.6	0.08±0.17
γ-Terpinene	**6**	**2**	0.03±0.04	2.3	1.4	0.10±0.11
α-Terpinolene	**5.5**	**1.8**	0.05±0.06	2.3	1.1	0.12±0.14
Linalool	0.4	1.2	0.04±0.13	2.2	**2.5**	0.04±0.08
α-Terpineol	1.2	1.4	0.03±0.08	0.8	0.6	0.09±0.17
Methyl thymol	**6.9**	1.1	0.40±0.24	0.6	0.9	0.06±0.16
Piperitone + Linalyl acetate	2	**2.5**	0.09±0.13	0.4	0.5	0.05±0.14
Bornyl acetate	2.2	1.1	0.07±0.14	0.6	1.4	0.04±0.10
Sesquiterpenes						
α-Cubebene + α-Longipinene	**5.2**	1.6	0.13±0.13	0.5	1.3	0.04±0.12
α-Copaene	**8.6**	1.3	0.41±0.22* ↑	0.1	1.3	0.03±0.10
β-Cubebene + Sativene + β-Elemene	**8.7**	1.3	0.39±0.23	0.2	1.1	0.04±0.13
Longifolene	**13.7**	1.6	0.12±0.14	0.4	0.8	0.05±0.17
trans-β-Caryophyllene	**9.6**	1.4	0.24±0.19	0.5	1.2	0.05±0.14
α-Humulene	**11.2**	1.4	0.34±0.20	0.3	0.8	0.04±0.14
Germacrene D + Phenethyl 2-methylbutyrate	**46**	**40.3**	0.54±0.19* ↑	0.8	0.7	0.06±0.18

PSM	*Constitutive*			*Inducibility*		
	Population $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*	*Population* $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*
Phenethyl isovalerate	**8.1**	**2.1**	0.18±0.15	2.5	**3.9**	0.24±0.23
α-Muurolene	**10.5**	1.4	0.42±0.21* ↑	0.3	0.7	0.05±0.15
γ-Cadinene	**7.6**	1.5	0.36±0.21	0.2	1.4	0.04±0.10
δ-Cadinene	**15**	1.2	0.62±0.21* ↑	0.5	1.1	0.04±0.13
Germacrene D-4-ol	**34.4**	**1.8**	0.52±0.19* ↑	0.5	**21.5**	0.04±0.12
Caryophyllene oxide	**8**	2.9	0.13±0.12	2.7	0.9	0.06±0.16
Diterpenes						
Pimaric acid	**4.1**	1.5	0.11±0.14	1.1	1.1	0.04±0.12
Sandaracopimaric acid	**4.3**	**2.1**	0.22±0.17	0.6	1.3	0.05±0.12
Unk DT1	**4.8**	1.5	0.33±0.21	0.3	1.2	0.06±0.17
Isopimaric acid	1.3	1.3	0.07±0.13	1	0.8	0.10±0.19
Levopimaric + Palustric	**3.2**	**2.8**	0.08±0.08	0.5	1.3	0.05±0.11
Dehydroabietic acid	**3.7**	1.4	0.25±0.20	1.2	0.9	0.11±0.17
8,12-abietadien-18-oic acid	**3.3**	**3.1**	0.06±0.06	0.6	1.2	0.05±0.12
Abietic acid	**5.3**	1.4	0.34±0.23	0.9	1.7	0.03±0.08
Neoabietic acid	**5.5**	**1.8**	0.24±0.18	0.5	1.3	0.04±0.11
Phenolics						
Vanillic acid hexoside	2.9	1.6	0.16±0.18	0.7	1.2	0.05±0.13
Coumaric acid hexoside	**5.5**	**2.5**	0.02±0.03* ↓	2	1.2	0.06±0.14
Coumaroylquinic acid	2.8	1.5	0.10±0.13	1.5	1	0.09±0.17
Taxifolin derivative 1	**7.4**	1.5	0.12±0.15	4.1	1.3	0.05±0.11
Ferulic acid hexoside	**4.8**	**2.2**	0.04±0.06	2.4	1.2	0.07±0.13
Unk P1	**3.7**	1.3	0.15±0.17	1.3	0.8	0.06±0.15
Unk P2	0.5	1.6	0.03±0.09	2	1.1	0.03±0.10
Unk P3	2.7	1.3	0.16±0.19	1	0.9	0.07±0.16
Lignan hexoside derivative 1	**5.3**	**4**	0.26±0.19	2.5	1.8	0.08±0.16
Lignan hexoside derivative 2	1.3	**2.3**	0.05±0.08	1	2	0.03±0.08
Unk P4	2.6	1.1	0.14±0.18	1.8	2.4	0.07±0.17
Unk P5	**4.8**	**1.9**	0.01±0.11	1.4	0.6	0.06±0.17
Lignan xyloside derivative 1	**18.4**	**133**	0.04±0.11	1.9	1.4	0.04±0.11
Lignan xyloside derivative 2	1.9	1.6	0.04±0.07* ↓	3.3	1.7	0.07±0.15
Lignan xyloside derivative 3	1.6	**2.2**	0.02±0.05	1.8	1.1	0.13±0.19

PSM	*Constitutive*			*Inducibility*		
	Population $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*	*Population* $F_{9,18}$	*Family* $F_{38,}$†	Q_{ST} *(± s.e.)*
Unk P6	**7.6**	1.8	0.22±0.15	1.1	0.9	0.06±0.15
Lignan deoxyhexoside	**9.2**	**2.6**	0.25±0.14	2.2	0.8	0.12±0.19
Unk P7	1.7	1.7	0.05±0.12	1	0.7	0.05±0.16
Unk P8	**3.2**	1.7	0.12±0.15	1.5	1.2	0.08±0.18
Unk P9	**5.9**	1.5	0.13±0.14	1.5	1.1	0.09±0.17
Unk P10	**4.7**	1.2	0.33±0.24	2.1	1.3	0.07±0.17
Eugenol	0.9	1.4	0.06±0.14	3.1	1.2	0.04±0.08
Methyl eugenol	2.7	**2.4**	0.05±0.07	**7.2**	1.3	0.29±0.20
trans-Isoeugenol	1.5	1.6	0.03±0.11	1.3	0.6	0.08±0.18
Fatty acids						
Oleic acid C18:1	2.7	**2.2**	0.03±0.05* ↓	1.2	1.1	0.05±0.15
Unk DT2 + Arachidic acid C20:0	0.8	**2.4**	0.03±0.06	0.3	0.8	0.08±0.18
Behenic acid C22:0	2.1	1.4	0.12±0.16	1.9	0.9	0.20±0.23

† Denominator degrees of freedom for Family factor for constitutive and inducibility of defences: monoterpenes, sesquiterpenes, diterpenes and fatty acids = 55; phenolics = 57.

References:

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **57,** 289-300.

Table S5. Effect of the methyl jasmonate induction (MJ) on the concentration of total and individual plant secondary metabolites (PSM) in the bark of ten maritime pine populations. Summary of the mixed model (F and P values), mean concentrations (± s.e.; $mg \cdot g^{-1}$ dry weight) of each PSM in control and MJ-induced plants and the associated fold change in the concentration are shown. Significant p-values are given in bold after adjustment for multiple testing correction using false discovery rate (FDR) at $p \leq 0.05$ (Benjamini & Hochberg 1995). HCAs = hydroxycinnamic acids.

PSM	*MJ-induction (MJ)*		*Mean concentration ± s.e. ($mg \cdot g^{-1}$dry weight)*		*Fold change*
	$F_{1,}$†	*P*	*Control*	*MJ-induced*	
PSM groups					
Terpenes					
Total monoterpenes	46.66	**<0.001**	5.54 ± 0.41	7.34 ± 0.31	**0.33**
Total sesquiterpenes	0.13	0.715	1.32 ± 0.09	1.36 ± 0.07	0.03
Total diterpenes	18.3	**<0.001**	13.09 ± 1.36	16.53 ± 0.66	**0.26**
Phenolics					
Total flavonoids	7.93	**0.006**	0.28 ± 0.03	0.21 ± 0.03	**-0.25**
Total HCAs	1.1	0.298	2.93 ± 0.11	2.84 ± 0.09	-0.03
Total eugenols	187.9	**<0.001**	0.01 ± <0.00	0.05 ± <0.00	**5.37**
Total lignans	12.2	**<0.001**	2.49 ± 0.08	2.70 ± 0.05	**0.09**
Chemical index 1	22.88	**<0.001**	-0.28 ± 0.12	0.21 ± 0.07	**1.74**
Chemical index 2	21.15	**<0.001**	0.22 ± 0.10	-0.15 ± 0.09	**-1.66**
Individual PSM					
Monoterpenes					
α-Pinene	38.86	**<0.001**	2.12 ± 0.17	2.87 ± 0.14	**0.35**
Camphene	136.2	**<0.001**	0.16 ± 0.03	0.33 ± 0.01	**1.11**
β-Pinene	38.07	**<0.001**	2.15 ± 0.14	2.85 ± 0.12	**0.33**
Myrcene	1.78	0.186	0.45 ± 0.03	0.40 ± 0.03	-0.1
α-Terpinene	16.02	**<0.001**	<0.01	<0.01	**0.31**
ρ-Cymene	1.71	0.196	<0.01	<0.01	0.2
β-Phellandrene + Limonene	43.34	**<0.001**	0.18 ± 0.02	0.26 ± 0.02	**0.51**
γ-Terpinene	0.27	0.606	0.01 ± 0.00	0.01 ± 0.00	0.09
α-Terpinolene	1.93	0.17	0.05 ± 0.00	0.06 ± 0.01	0.26
Linalool	429.19	**<0.001**	<0.01	0.01 ± 0.00	**64.4**
α-Terpineol	74.57	**<0.001**	<0.01	0.01 ± 0.00	**4.58**
Methyl thymol	13.68	**<0.001**	<0.01	<0.01	**0.48**
Piperitone + Linalyl acetate	74.33	**<0.001**	<0.01	0.01 ± 0.00	**3.18**
Bornyl acetate	156.67	**<0.001**	0.09 ± 0.02	0.24 ± 0.01	**1.64**

PSM	*MJ-induction (MJ)*		*Mean concentration ± s.e. (mg·g^{-1}dry weight)*		*Fold change*
	$F_{1,}$†	*P*	*Control*	*MJ-induced*	
Sesquiterpenes					
α-Cubebene + α-Longipinene	0.42	0.521	0.02 ± 0.00	0.02 ± 0.00	-0.05
α-Copaene	0.06	0.815	0.02 ± 0.00	0.02 ± 0.00	-0.01
β-Cubebene + Sativene + β-Elemene	1.73	0.193	0.01 ± 0.00	0.01 ± 0.00	-0.08
Longifolene	0.33	0.57	0.08 ± 0.02	0.09 ± 0.01	0.05
trans-β-Caryophyllene	0.01	0.906	0.50 ± 0.03	0.51 ± 0.03	0.01
α-Humulene	0.06	0.8	0.11 ± 0.01	0.10 ± 0.01	-0.02
Germacrene D + Phenethyl 2-methylbutyrate	0.68	0.412	0.30 ± 0.03	0.31 ± 0.02	0.05
Phenethyl isovalerate	63.56	**<0.001**	0.01 ± 0.00	0.02 ± 0.00	**0.85**
α-Muurolene	0.13	0.715	0.02 ± 0.00	0.02 ± 0.00	0.03
γ-Cadinene	1.09	0.299	0.07 ± 0.01	0.06 ± 0.01	-0.05
δ-Cadinene	1.23	0.272	0.02 ± 0.00	0.01 ± 0.00	-0.1
Germacrene D-4-ol	0.5	0.481	0.01 ± 0.00	0.01 ± 0.00	-0.09
Caryophyllene oxide	7.19	**0.009**	0.01 ± 0.00	0.02 ± 0.00	**0.77**
Diterpenes					
Pimaric acid	11.14	**0.001**	1.06 ± 0.08	1.29 ± 0.06	**0.22**
Sandaracopimaric acid	16.07	**<0.001**	0.32 ± 0.03	0.41 ± 0.02	**0.28**
Unk DT1	13.78	**<0.001**	0.74 ± 0.08	0.93 ± 0.04	**0.26**
Isopimaric acid	11.18	**0.001**	0.50 ± 0.10	0.72 ± 0.04	**0.43**
Levopimaric + Palustric	13.39	**<0.001**	2.37 ± 0.34	2.98 ± 0.15	**0.26**
Dehydroabietic acid	12.13	**<0.001**	0.89 ± 0.12	1.18 ± 0.07	**0.33**
8,12-abietadien-18-oic acid	9.63	**0.003**	0.37 ± 0.03	0.42 ± 0.02	**0.16**
Abietic acid	28.72	**<0.001**	4.65 ± 0.39	6.10 ± 0.22	**0.31**
Neoabietic acid	3.66	0.06	3.28 ± 0.32	3.80 ± 0.23	0.16
Phenolics					
Vanillic acid hexoside	0.23	0.633	0.30 ± 0.01	0.30 ± 0.01	-0.01
Coumaric acid hexoside	0.68	0.412	0.80 ± 0.05	0.78 ± 0.04	-0.03
Coumaroylquinic acid	43.65	**<0.001**	0.20 ± 0.01	0.14 ± 0.01	**-0.29**
Taxifolin derivative 1	8.95	**0.004**	0.45 ± 0.04	0.35 ± 0.04	**-0.21**
Ferulic acid hexoside	0.15	0.701	1.91 ± 0.06	1.89 ± 0.06	-0.01
Unk P1	31.44	**<0.001**	1.49 ± 0.11	0.94 ± 0.11	**-0.37**
Unk P2	1.04	0.31	0.37 ± 0.01	0.36 ± 0.01	-0.03
Unk P3	1.47	0.229	0.51 ± 0.03	0.48 ± 0.02	-0.05
Lignan hexoside derivative 1	3.02	0.087	0.68 ± 0.02	0.72 ± 0.02	0.06
Lignan hexoside derivative 2	8.45	**0.005**	0.14 ± 0.02	0.21 ± 0.02	**0.52**

PSM	*MJ-induction (MJ)*		*Mean concentration ± s.e. ($mg \cdot g^{-1}$dry weight)*		*Fold change*
	$F_{1,}$†	*P*	*Control*	*MJ-induced*	
Unk P4	23.36	**<0.001**	0.36 ± 0.04	0.21 ± 0.03	**-0.42**
Unk P5	3.75	0.057	0.61 ± 0.05	0.70 ± 0.04	0.16
Lignan xyloside derivative 1	0.21	0.648	0.14 ± 0.01	0.14 ± 0.01	-0.03
Lignan xyloside derivative 2	4.53	0.037	0.51 ± 0.02	0.55 ± 0.02	0.07
Lignan xyloside derivative 3	18.55	**<0.001**	0.37 ± 0.01	0.42 ± 0.01	**0.13**
Unk P6	2.41	0.125	0.42 ± 0.02	0.45 ± 0.02	0.07
Lignan deoxyhexoside	7.59	**0.007**	0.60 ± 0.01	0.64 ± 0.02	**0.07**
Unk P7	0.54	0.467	0.39 ± 0.02	0.38 ± 0.02	-0.04
Unk P8	10.9	**0.002**	0.21 ± 0.02	0.25 ± 0.01	**0.19**
Unk P9	41.22	**<0.001**	0.18 ± 0.01	0.26 ± 0.01	**0.45**
Unk P10	1.23	0.272	0.20 ± 0.05	0.23 ± 0.03	0.14
Eugenol	38.36	**<0.001**	<0.01	<0.01	**1.4**
Methyl eugenol	364.88	**<0.001**	<0.01	0.04 ± <0.00	**15.74**
trans-Isoeugenol	2.96	0.09	<0.01	<0.01	0.19
Fatty acids					
Oleic acid C18:1	239.3	**<0.001**	2.45 ± 0.16	0.82 ± 0.15	**-0.66**
Unk2 + Arachidic acid C20:0	4.74	0.033	0.96 ± 0.10	1.10 ± 0.06	0.14
Behenic acid C22:0	1.13	0.292	0.52 ± 0.03	0.49 ± 0.05	-0.05

† Denominator degrees of freedom for MJ factor: monoterpenes, sesquiterpenes, diterpenes, and fatty acids = 70; phenolics = 73.

References:

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **57,** 289-300.

Table S6. Multivariate Analysis of Variance for the effect of methyl jasmonate induction (MJ), pine population (P), family within population [F(P)], as well as MJ×P and MJ×F(P) interactions on the concentration of plant secondary metabolites (PSM) in the bark of maritime pine. Significant p-values ($p < 0.05$) are given in bold. Sixty-three PSM [those present in more than 75% of the cases in at least one of the induction treatments (control or MJ)] were included in the analysis.

Effect	Wilk's Lambda	*F*	P-value
Induction (MJ)	0.0760203	16.4	**<0.001**
Population (P)	0.0000012	5.0	**<0.001**
Family [F(P)]	0.0000000	1.7	**<0.001**
MJ×P	0.0109837	0.9	0.887
MJ×F (P)	0.0000009	0.7	1.000