Fire frequency, as well as stress response and developmental gene control serotiny level variation in a widespread pioneer Mediterranean conifer, Pinus halepensis

Bastien Romero, Ivan Scotti, Bruno Fady, Anne Ganteaume

RESEARCH ARTICLE

Ecology and Evolution WILEY

Fire frequency, as well as stress response and developmental gene control serotiny level variation in a widespread pioneer Mediterranean conifer, *Pinus halepensis*

Bastien Romero[1] | Ivan Scotti[2] | Bruno Fady[2] | Anne Ganteaume[1]

[1]INRAE, Aix Marseille Univ, RECOVER, Aix-en-Provence, France

[2]INRAE, URFM, Avignon, France

Correspondence
Anne Ganteaume, INRAE, Aix Marseille Univ, RECOVER, Aix-en-Provence, France.
Email: anne.ganteaume@inrae.fr

Abstract

Many plants undergo adaptation to fire. Yet, as global change is increasing fire frequency worldwide, our understanding of the genetics of adaptation to fire is still limited. We studied the genetic basis of serotiny (the ability to disseminate seeds exclusively after fire) in the widespread, pioneer Mediterranean conifer *Pinus halepensis* Mill., by linking individual variation in serotiny presence and level to fire frequency and to genetic polymorphism in natural populations. After filtering steps, 885 single nucleotide polymorphisms (SNPs) out of 8000 SNPs used for genotyping were implemented to perform an in situ association study between genotypes and serotiny presence and level. To identify serotiny-associated loci, we performed random forest analyses of the effect of SNPs on serotiny levels, while controlling for tree size, frequency of wildfires, and background environmental parameters. Serotiny showed a bimodal distribution, with serotinous trees more frequent in populations exposed to fire in their recent history. Twenty-two SNPs found in genes involved in stress tolerance were associated with the presence-absence of serotiny while 37 found in genes controlling for flowering were associated with continuous serotiny variation. This study shows the high potential of *P. halepensis* to adapt to changing fire regimes, benefiting from a large and flexible genetic basis of trait variation.

KEYWORDS
cone serotiny, fire regime, genetic association, phenotypic plasticity, *Pinus halepensis*, single nucleotide polymorphisms

TAXONOMY CLASSIFICATION
Ecological genetics, Evolutionary ecology

1 | INTRODUCTION

Wildfires represent a highly selective disturbance for plant species (Bond & van Wilgen, 1996; Pausas & Keeley, 2009). In fire-prone ecosystems, such as in the Mediterranean regions, several plant species have developed adaptations to fire. The specific types of adaptation depend on fire regime, which is defined, among other things, by fire frequency and intensity (Bradshaw et al., 2011; Keeley

et al., 2011; Ne'eman et al., 2012). Three types of fire adaptive strategies have been reported in the literature (Pausas, 2015a). Among these, fire-tolerant species survive a fire while fire-embracer species base their strategy on postfire regeneration via resprouting (e.g., *Quercus* ilex L.) or via producing seedlings (e.g., *Pinus halepensis*). Fire-tolerant species are adapted to fire regimes characterized by frequent low- to medium-intensity surface fires (He et al., 2019; Murphy et al., 2013). These species (e.g., *Pinus sylvestris* L.) possess traits that mitigate fire impacts (e.g., thick bark, self-pruning), enabling them to survive after fire (Fernandes et al., 2008; Keeley, 2012). By contrast, fire-embracer species are adapted to fire regimes characterized by infrequent, high-intensity crown fires leading to the death of the extant adult cohort (He et al., 2019). Fire-embracer species that are obligate seeders present fire-related traits enhancing crown flammability such as an aerated crown structure and dead fuel retention that create fuel continuity between the ground and the canopy, allowing the fire to spread vertically (Keeley, 2012; Ne'eman et al., 2004). This strategy releases propagules from competition for nutrients, water, and light by canopy vegetation during postfire regeneration.

The effects of fire regime on the phenotypic variation of traits are quite well studied (Engber & Varner, 2012; Enright et al., 2014; Pausas, 2015a, 2015b), but only a few studies found clear relationships between traits related to flammability and fire frequency (e.g., Pausas et al., 2012). Fire has shaped genetic diversity in species of fire-prone environments (Fady, 2012) and within species and populations, the variability of fire-related traits may be adaptive and response to changing fire regimes (Pausas, 2015a). The consequences of intra-population heritable variation in fire adaptive traits on population fitness depend on the fire regime.

Among fire-related traits, serotiny of pine cones presents a clear evolutionary response to changes in fire frequency (Keeley et al., 2011) and is evolutionary correlated with other fire-related traits (Schwilk & Ackerly, 2001). Therefore, in this study, we used serotiny as an indicator for multi-trait fire phenotypes in *Pinus halepensis* (see Budde et al., 2014 for *Pinus pinaster* Ait.). This functional trait is observed in fire-embracer species and allows the plant to create a canopy seed bank, storing seeds in serotinous cones that can remain closed for several years until the occurrence of a fire event (He et al., 2012; Lamont et al., 1991; Pausas, 2015b). The heat shock induced by the high temperatures reached during a fire, for instance, triggers cone opening and seed release in a nutrient-enriched environment with reduced competition. Therefore, serotiny increases fitness by increasing the success rate of seedlings after a fire (Causley et al., 2016; Hernández-Serrano et al., 2013; Keeley et al., 2011; Pausas, 2015b).

The proportion of serotinous vs nonserotinous cones (i.e., serotiny level) can strongly vary among but also within species (Climent et al., 2014; Hernández-Serrano et al., 2013; Peeler & Menges, 2018). An increase in serotiny level with fire frequency has been reported in many species. For instance, Ripa et al. (2020) found that serotiny level in *P. contorta* was higher in postfire recruited trees and could vary over a single generation after fire, suggesting a fast

response to selection. The evolutionary response of serotiny to fire selective pressure rests on its variability among and within populations (Keeley et al., 2011; Lamont et al., 2020). Serotiny level varies widely among species, but several studies reported patterns in its distribution. For example, for maritime pine *Pinus pinaster*, the number of trees without serotinous cones was high compared with trees expressing serotiny (Budde et al., 2014; Castellanos et al., 2015). However, serotiny data observed by Budde et al. (2017) for the Aleppo pine (*Pinus halepensis*) was following a normal distribution.

Heritability of serotiny has been assessed in several studies concluding that serotiny could be considered as highly heritable with a simple genetic control (one locus with two alleles, i.e., Mendelian control) because changes appeared within solely a few generations (as observed in *Pinus contorta* and *Pinus banksiana*; Teich, 1970; Perry & Lotan, 1979; Wymore et al., 2011). With the improvement of the genetic tools and the apparition of association studies between genotype and phenotype, later studies showed that 11 to 17 loci explained 50% to 29% of the phenotypic variation of serotiny in *P. contorta* (Parchman et al., 2012) and in natural populations of *P. pinaster* (Budde et al., 2014), respectively, underlining a polygenic process driving serotiny. These results suggest that serotiny is controlled by more than one locus and demonstrate the feasibility of genomic association studies in natural populations. These studies also showed that association studies are very promising to assess the link between genotype and plant traits (De La Torre et al., 2019; Feduck et al., 2015; Pausas, 2015b).

Aleppo pine is a common and widespread fire-embracer species of the central and western Mediterranean basin (Pausas, 2015a; Tapias et al., 2004; Wazen et al., 2020). In this species, serotiny level depends on fire regime, environmental conditions (e.g., soil nutrients, precipitations), and tree age, for instance (Cruz et al., 2019; He et al., 2012; Ne'eman et al., 2012). The identification of the genetic bases of serotiny in *P. halepensis* is important to assess the species' capacity to cope with expected future changes in fire regime (Dupuy et al., 2020). Castellanos et al. (2015), working on natural populations, found a heritability estimate of 10% while Hernández-Serrano et al. (2014) found a higher value (20%) in a common garden experiment. Finally, Pinosio et al. (2014) identified 541 genes as differentially expressed between two natural trees sampled in populations undergoing different fire regimes, which may or may not be suggestive of genetic differentiation of those genes between the two populations.

These previous studies showed that serotiny in *Pinus halepensis* was genetically controlled and paved the way for genomic association studies. Considering that this trait is variably expressed among fire regimes, populations, and ontogenetic stages, it is important to run such association studies using samples that cover a wide span of conditions and therefore of potential serotiny values and/or sources of genetic variation. Given that it is difficult to disentangle the genetic basis of serotiny itself because loci associated with serotiny might be associated with other correlated fire adaptive traits, the goal of this study was therefore to identify a variety of candidate genes in *Pinus halepensis* that could underlie

fire phenotypes. This work also aimed at understanding patterns of phenotypic variation in this trait when the fire history changes. To do so, we used an association approach in natural pine populations of southeastern France that vary in their fire history and in some environmental conditions.

2 | MATERIALS AND METHODS

2.1 | Study area and species

Sampling sites were located in southeastern France (Figure 1) and are characterized by a typical meso- to supra-Mediterranean climate (Köppen, 1900; mild and humid winters, hot and dry summers), calcareous soils (i.e. limestone), and elevation ranging from 100 to 600 meters. Southeastern France is affected by moderate fire frequency compared with other countries of southern Europe (4657 fires and more than 21,000 ha burned during the 2010–2020 period, according to the regional fire database Prométhée; www.promethee.com), with a strong spatial variability. In the study area, as in other Mediterranean regions, fire has shaped most ecosystems but assessing the historical number of recurrent fires on a location is difficult due to a long fire history. Fire databases, when available, can provide reliable information on fire frequency during the last six decades, at most. In the current work, the recent fire history (60 years) was reconstructed using the regional fire database Prométhée (www.promethee.com) that has recorded fires in southeastern France since 1973 and the georeferenced fire perimeters recorded in the database of the Direction Départementale des Territoires et de la Mer (DDTM Bouches du Rhône) available from 1961. When the fire perimeters were not recorded in the database, we used satellite images and aerial photos to outline fire scars (e.g., delineation of unburned areas and fire boundary adjustment were performed using QGIS software).

The most common species in the study area are *Quercus ilex*, *Quercus pubescens*, and *Pinus halepensis* (Quézel, 2000). The latter species is an obligate seeder, bearing serotinous cones, and is one of the most studied Mediterranean trees (Ne'eman et al., 2004). *P. halepensis* is also a pioneer species adapted to drought, poor soil, high temperatures, and crown fire regimes (Pausas, 2015a).
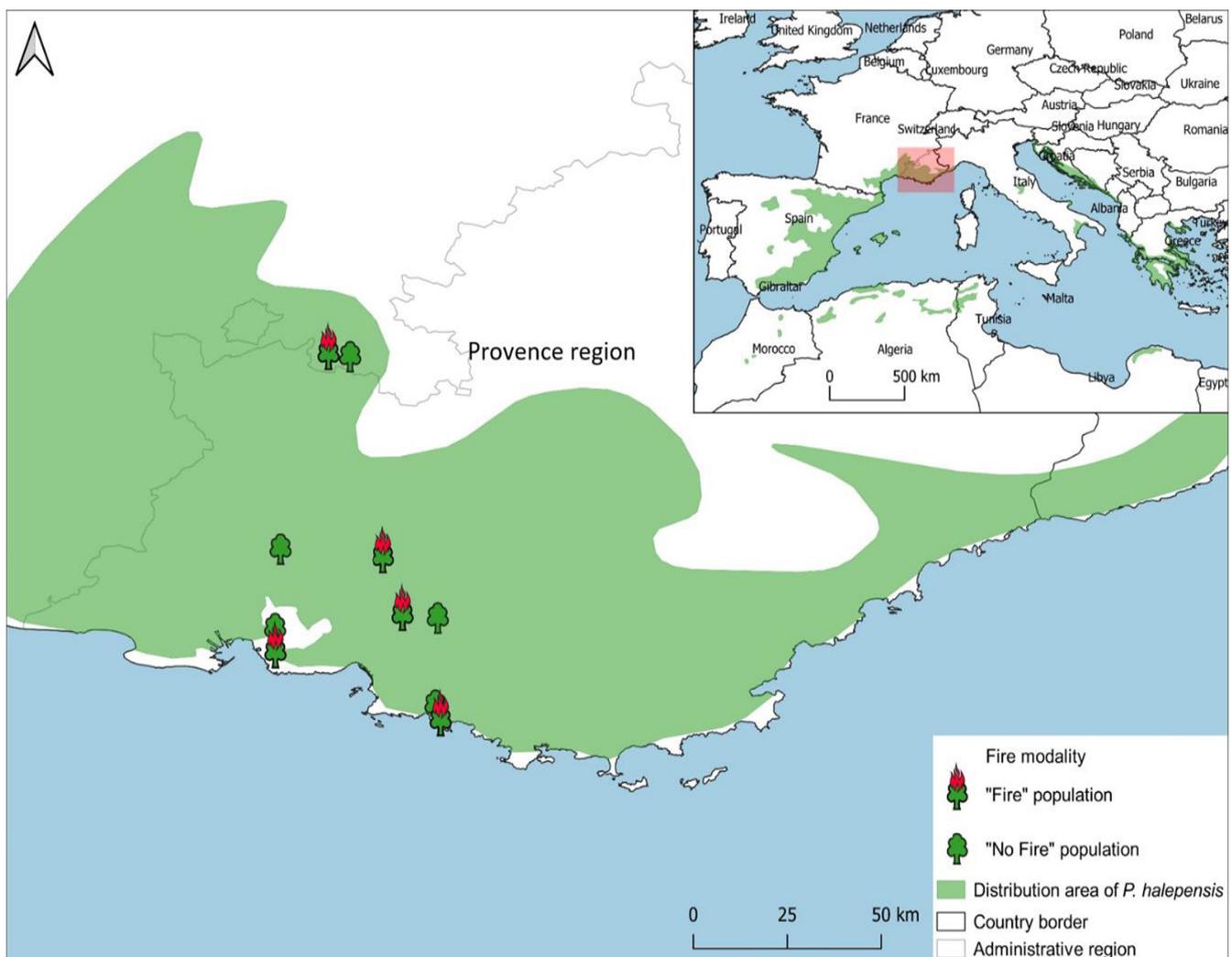
FIGURE 1 Map of the study area with "Fire" and "No-Fire" sampling sites and *Pinus halepensis'* distribution area in green.

2.2 | Plant sampling and serotiny level measurement

Overall, 10 populations of *P. halepensis* were selected according to two fire modalities, half located in areas that were fire-free from 1959 until 2018 ("No-Fire" modality populations, PHNF) and the other half in areas having suffered at least one fire ("Fire" modality populations, PHF) during the same time period. The pairs "Fire"–"No-Fire" were located between 5 and 25 km apart in order to minimize genetic divergence caused solely by drift (Lotterhos & Whitlock, 2015). The populations were sampled in sites where the past land use was the same in order to mitigate the impact of soil conditions on trait variation (See Romero & Ganteaume, 2020). In the current work, the sampling occurred in recent forests (as opposed to ancient forests that could have already been mapped in the 18th–19th century), mostly resulting from land abandonment that can provide adequate conditions for regeneration in the absence of fire. In each population, 19–20 mature and dominant trees, between 15- and 30-years-old (on average 22.03±3.1-years-old), were sampled. In areas undergoing fires during the past 60 years, we only sampled populations when at least 10 years had elapsed since the last fire event in order to be sure of the sexual maturity of the trees that have grown postfire (Santos del Blanco et al., 2010). Furthermore, in order to reduce environmental differences between populations, we selected sample sites with as homogeneous environmental conditions as possible (slope, exposure, elevation, and past land use; see Table S1). However, Romero and Ganteaume (2020) found a significant effect of these factors on trait variation, working on the same database as the one used in the current study. This effect has therefore to be accounted for in the analyses (see below).

Serotiny measurements were carried out following Budde et al. (2014). Using binoculars, we counted, on each tree, 20 mature cones (older than 3 years) on healthy branches, avoiding dominated trees and those with a diameter smaller than 10 cm. Measurements were performed during the summer (June–July 2018), at least 48 h after a rain event when occurred to avoid counting cones closed due to high relative humidity. Tree serotiny level was calculated as the number of closed cones divided by the total number of cones (on 192 trees).

Because of the apparent bimodal distribution of trait values (see Results), serotiny data were transformed in three different ways (Figure 2):

(i) Serotiny level data were recorded as "0" for trees without serotinous cones, and as "1" otherwise; this allowed the analysis of the genetic bases of the presence/absence of serotiny. This is the *"Binary"* ("B") data set.

(ii) Serotiny level data were log-transformed using the LogSt() function in the DescTools R package (Signorell, 2019), allowing for the Log-transformation of zero values. This is the *"Inclusive Quantitative"* ("IQ") data set.

(iii) Trees with serotiny = 0 were removed, and then transformation proceeded as in (ii). This is the *"Exclusive Quantitative"* ("EQ") data set.

The variation of serotiny level among populations and according to fire modality was assessed using an ANOVA, followed by a Fisher Least Significant Difference (LSD) test. As the data distribution was not normal, data were either log-transformed or tested using the Kruskal–Wallis test followed by Bonferroni post hoc tests.

To remove the possible effect of some covariables on serotiny level, in order to keep only the genetic effect in the GWAS, residues of multivariate generalized linear models accounting for all environmental and ontogenetic variables (Table S1) were extracted in two ways:

- We designed a linear model of the form: $\text{TransformedSerotiny}_x = \mu + \text{cofactor}_1 + \text{cofactor}_2 + \ldots \text{cofactor}_N + \varepsilon$; the best model was identified using variable selection as implemented in the step() function of the R package "stats" (R Core Team, 2013) with AIC (Table S2) ("GLM" residues)
- all cofactors were used to perform a Principal Component Analysis using the FactoExtra R package (Kassambara & Mundt, 2017), and then, the first three components were used in the linear model: $\text{TransformedSerotiny}_x = \mu + \text{PC}_1 + \text{PC}_2 + \text{PC}_3 + \varepsilon$ ("PCA" residues)

where $\text{TransformedSerotiny}_x$ is one of the "Binary," "Inclusive Quantitative," "Exclusive Quantitative" data sets, and cofactor$_i$ represents each cofactor listed in Table S1.

By combining the three data sets with our two types of linear models, we obtained six residue tables, named after the combination of data sets and linear models: "B-GLM," "IQ-GLM," "EQ-GLM," "B-PCA," "IQ-PCA," and "EQ-PCA." These six residue data sets were used in all association analyses; no other data set was used for this purpose.

2.3 | Genotyping, background population structure, data filtering, and genotype imputation

Healthy and mature needles were sampled on the 192 trees chosen for serotiny genotype measurements and put in tagged individual paper bags. In the laboratory, samples were oven-dried for 4 days at 30°C and then stocked with silica gel in order to avoid rehydration before DNA extraction. DNA isolation was performed using DNeasy® 96 Plant Kits from QIAGEN and genotyping was carried out at the Institute of Applied Genomics (IGA technologies services, Udine, Italy) based on 8000 SNPs, genotyped using the SPET technology (Scaglione et al., 2019). The target regions were chosen within coding sequences based on Pinosio et al. (2014).

Prior to analyses, the SNP data set was filtered as follows: (i) We removed variants with more than 20% missing data; (ii) we created subsets with the remaining variants so that no pair with linkage disequilibrium (LD) higher than 0.5 was left in each contig. To do this, we first computed LD using *vcftools v. 0.1.13* (Danecek et al., 2011), and then, we iteratively removed one variant from

each pair of high LD variant pairs, until LD < 0.5 for all pairs, using an ad-hoc R script.

Basic descriptors of the data set were obtained with *vcftools v. 0.1.13* (Danecek et al., 2011) and with the R package *"hierfstat" v. 0.5-7* (Goudet, 2005). Genetic population structure was checked using (i) a principal component analysis (PCA) performed with the R package *"adegenet"* (Jombart, 2008; Jombart & Ahmed, 2011) and (ii) the Bayesian clustering method implemented in *fastSTRUCTURE 2.3.4* (Raj et al., 2014). K values ranging from 1 to 10 were implemented prior to obtaining values that explained population structure. Population pairwise divergence was estimated using Weir & Cockerham's F_{ST} (Weir & Cockerham, 1984) and the R package *"hierfstat" v. 0.5-7* (Goudet, 2005).

Prior to genomic association analyses, we removed variants with a minor allele frequency (MAF) lower than 0.1 and we imputed missing genotypes using *Beagle 4.1* (Browning & Browning, 2016; Heer et al., 2018) without using a reference sequence. After the two filtering steps, 885 SNPs out of 8000 remained.

2.4 | Relatedness matrix and heritability estimates

Patterns of relatedness must be checked and accounted for in the genomic association analyses. We obtained genetic relatedness matrices following three methods. First, Ritland's kinship (Ritland, 1996a) was computed from the SNP genotype data. Then, a genome-wide relatedness matrix (GRM) was computed according to the GCTA (Yang et al., 2011) definition, as computed by the *snpgdsGRM()* function in the R package *"SNPrelate" v. 1.22.0* (Zheng et al., 2012) and finally as $X_a \times X_a^T$, where X_a is the centered matrix of genotypes, as suggested by the manual of the R package *"mlmm. gwas"* (Bonnafous et al., 2018).

We used relatedness matrices and trait data to estimate heritability in two ways (see Table 1 for technical details):

a. from Ritland's kinship and the value of the actual variance of kinship (Ritland, 1996b), as estimated in SPAGeDi (Hardy & Vekemans, 2002, as maintained at https://github.com/
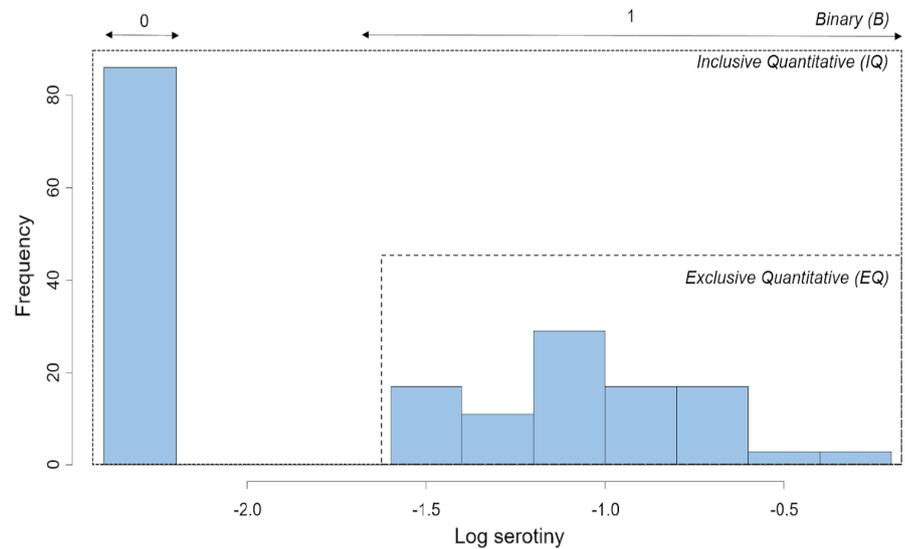
FIGURE 2 Histogram of distribution of the serotiny level (Log-transformed) with the details of the three different types of data used for association analyses ("Binary" (B), "Exclusive Quantitative" (EQ), and "Inclusive Quantitative" (IQ)).

TABLE 1 Heritability estimates obtained for the six data sets (B-GLM; B-PCA; IQ-GLM; IQ-PCA; EQ-PCA, and EQ-GLM), along with their lower and upper margins (Ritland's estimation) and confidence intervals (Bayesian estimation).

Data set	Lower margin	Central value	Upper margin	Lower credible interval	Central value	Upper credible interval
B-GLM	0.0021	0.55	2.6	0.049	0.28	0.54
B-PCA	0.037	0.79	3.5	0.18	0.45	0.67
IQ-GLM	−0.16	0.70	3.9	0.17	0.43	0.70
IQ-PCA	−0.15	0.84	4.8	0.27	0.56	0.80
EQ-GLM	−0.45	−0.0046	1.7	0.00	0.15	0.64
EQ-PCA	−0.44	0.090	2.0	0.00	0.11	0.57
B (GLM cofactors)	☐		☐	0.14	0.41	0.64
B (PCA cofactors)	☐		☐	0.0010	0.17	0.56
IQ (GLM cofactors)	☐		☐	0.0010	0.17	0.38
IQ (PCA cofactors)	☐		☐	0.0010	0.31	0.76
EQ (GLM cofactors)	☐		☐	0.0010	0.19	0.65
EQ (PCA cofactors)	☐		☐	0.0010	0.18	0.58

reedacartwright/spagedi), we computed Ritland's heritability (Ritland, 1996b). We used ad-hoc R scripts to compute upper, central, and lower heritability estimates for serotiny (taken as residuals of multiple regressions, see above). Following this goal, first, the standard deviation of the numerator (covariance of trait-based similarity and kinship) was obtained by bootstrap using the *bootstrap()* function in the R package *"bootstrap" v. 2019.6* (Tibshirani & Efron, 1993). Then, the standard deviation of the denominator (actual variance of relatedness) was obtained from SPAGeDi along with its central value (see above) and finally, the heritability, calculated from the ratio, was obtained from the central values ($m_{numerator}$, $m_{denominator}$; central estimator) and from the pairs of values ($m_{numerator} - sd_{numerator}$, $m_{denominator} + sd_{denominator}$; lower estimate) and ($m_{numerator} + sdm_{numerator}$, $m_{denominator} - sd_{denominator}$; upper estimate).

b. from the genome-wide relatedness matrix (GRM) and traits (taken either as uncorrected data (data sets "B," "IQ," and "EQ") or as residuals of multiple regressions, see above), Bayesian estimates of heritability (median and 95% credible intervals) from generalized linear mixed models were computed using the function h2.jags() of the R package *"gap" v. 1.2.2* (Zhao, 2008). When using uncorrected data, the same variables implemented as cofactors in the GLM were used as cofactors in heritability estimates.

It is worth noting that heritability estimates were additionally provided by the MLMM approach used for GWAS (see below).

2.5 | Genetic association analysis

The six data sets ("B-GLM," "IQ-GLM," "EQ-GLM," "B-PCA," "IQ-PCA," and "EQ-PCA") were used for running "random forest" analyses (machine learning algorithm; Breiman, 2001; Brieuc et al., 2018) to identify associations between SNPs and serotiny. The analyses were run with imputed missing data (see above), even if this process can induce a bias in the data (Heer et al., 2018). We carried out these analyses using two different R packages, *Boruta* (Kursa & Rudnicki, 2010) and *VSURF* (Genuer et al., 2015), for the six data sets described above, for a total of 12 association tests.

Association analyses were additionally carried out using the MLMM approach (Segura et al., 2012), which accounts for genetic relatedness, using the R package "mlmm.gwas" v. 1.0.6 (Bonnafous et al., 2018), with the integration of the additional mlmm_cof() function.

2.6 | Functional annotation

Genes containing serotiny-associated SNPs were annotated first by comparing them with *Arabidopsis thaliana*'s sequence database using BLAST (Basic Local Alignment Search Tool; Mount, 2007). Next, the best *A. thaliana* hits were annotated using the r package topGO (Alexa & Rahnenfuhrer, 2020), restricting the GO term search to the top hierarchical level ("biological process"). An enrichment test was

performed with *p*-values adjusted for multiple comparisons with the false discovery rate procedure (FDR) and a threshold of .05.

3 | RESULTS

3.1 | Phenotypic variation of serotiny level

Serotiny level (quantitative data) ranged from 0 to 60% overall and was highly variable among populations regardless of the fire modality (Figure 3), for both raw (ANOVA, Kruskal–Wallis test, KW = 161.73, *p*-value <.0001 with 33 significant comparisons according to the 95% Bonferroni test) and log-transformed data (ANOVA, F = 3.49, *p*-value <.0001 with 32 significant comparisons according to the Fisher Least Significant Difference (LSD) test). Taking into account only the data with serotiny level >0, the effect of the fire modality on the serotiny level was not significant (ANOVA, F = 0.84, p = .36). Furthermore, the number of trees expressing serotiny (i.e. with serotiny level>0) was significantly higher in the "Fire" populations compared with the "No-Fire" populations (Fisher test; *p*-value = .0002; Figure 4). Data were discontinuously distributed, with a gap between trees displaying zero and nonzero serotiny (see Figure 2).
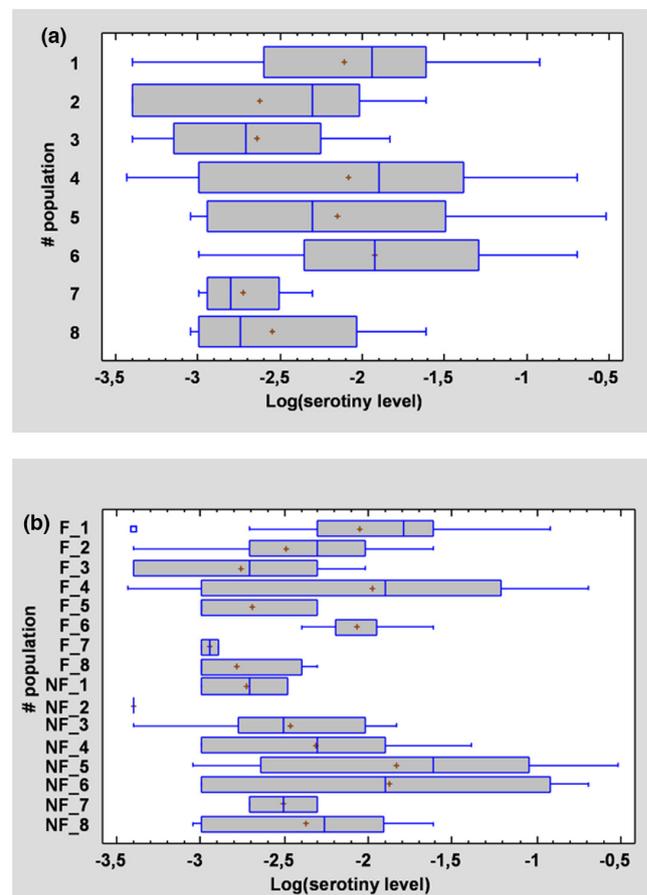
FIGURE 3 Variation of serotiny level among populations (a) and taking into account the fire modality (b) (F = Fire modality; NF = No-Fire modality).

3.2 | Genotype data and genetic structure of populations

Due to relatively high proportions of missing data and high levels of linkage disequilibrium among variants, we removed the vast majority of loci at the missing data and LD filtering steps. The final data set, prior to filtering for low-frequency variants (MAF < 0.1), contained 904 biallelic SNP variants, and after the MAF filter, only 885 SNPs were left. While this reduced considerably genomic coverage, the remaining loci are fully independent and high-quality, thus making the GWAS exercise more reliable, albeit very stringent. Average Nei's heterozygosity per SNP was 0.41 (standard deviation = 0.11) and there was a marked heterozygote excess (mean single-locus F_{IS} within-population: −0.44, standard deviation = 0.43).

Based on a principal component analysis, we chose to remove one outlier tree displaying a completely divergent genotypic structure (Figure 5a). After this step, the population structure was homogeneous, as indicated by the fact that the best number of clusters obtained with the Bayesian clustering analysis was $K = 1$. This suggests that there was not any confounding issue for association analyses, and therefore population structure was ignored in the subsequent analyses (Figure 5b). The estimation of pairwise population divergence (Weir and Cockerham F_{ST}) also showed a very shallow structure, with very small and similar values in all pairs (Table 2).

3.3 | Relatedness and heritability

The distribution of the values of relatedness was skewed towards the right in all cases (D'Agostino skewness test, skew of 0.30, 4.2,

and 4.4, respectively, for Ritland's, GCTA, and MLMM estimators; p-value < 10^{-15} in all cases), with a long, thin tail of excess values (Figure S1a). This suggests the presence of at least some small groups of tightly related individuals. Relatedness was not equally distributed among populations (Figure S1b); when comparing pairwise," No-Fire" numbers 5 and 6 (PHNF_5 and PHNF_6) had slightly higher relatedness than the matched fire-prone stands (for GCTA estimates: t-test, $t = -2.2038$, df = 319.78, p-value = .02 for stand 5; $t = -2.3775$, df = 247.17, p-value = .02 for stand; results are similar for the other estimators). The presence of unequally distributed relatedness values suggests that the kinship structure should be taken into account in GWAS analyses; however, the absolute numbers of pairs of individuals with very high kinship levels were small, suggesting that the bias may be relatively small. Heritability estimates were generally higher for the B and IQ type data sets and with the Ritland method, and lower in the EQ data set and with Bayesian estimations (Table 1). Central values varied between 0.28 and 0.84 for the G and IQ data sets but were much lowered when taking into account cofactors in the Bayesian analysis, down to 0.17–0.31 (with the exception of the B data set, for which introducing the cofactors increased the estimate). Estimates were close or equal to zero only for the EQ data set with the Ritland method (upper boundaries for the Ritland estimate went often beyond 1, while the Bayesian upper confidence interval limit was always lower than 1).

3.4 | Identification of markers associated with serotiny level variation

SNPs identified by analyzing PCA-based residuals (B-PCA, EQ-PCA, and IQ-PCA) and GLM-based residuals (B-GLM, EQ-GLM, and IQ-GLM) only partially overlapped, regardless of the algorithm used (Table 3, "Union" column). The MLMM analysis did not return any significant result after the most stringent criteria were applied (Bonferroni-corrected p-value <.05 for parametric tests, extended BIC (e-BIC) for likelihoods; Table S2).

Globally, taking into account the number of SNPs in both types of correction (GLM- and PCA-transformed), 24 significant SNPs in the *IQ* data set, and 15 in the *EQ* data set were associated with the variation of serotiny (both sets giving 37 different SNPs linked to the variation of serotiny). Twenty-two SNPs in the *B* data set were associated with the presence/absence of serotiny. Only one SNP was significant in both *IQ* and *B* data set analyses (i.e., Contig2340_2217) and another one was significant in all the *EQ* data set analyses (i.e., clc_contig_12713_224).

Twenty genes containing significant SNPs were associated with biological processes having specific functions possibly involved in serotiny. They were divided into three classes: "Flowering," "Stress tolerance," and/or "Reproduction." The list of loci driving the presence/absence of serotiny included genes involved in stress tolerance, while loci linked to the continuous variation of serotiny included flowering control genes (Table 4).
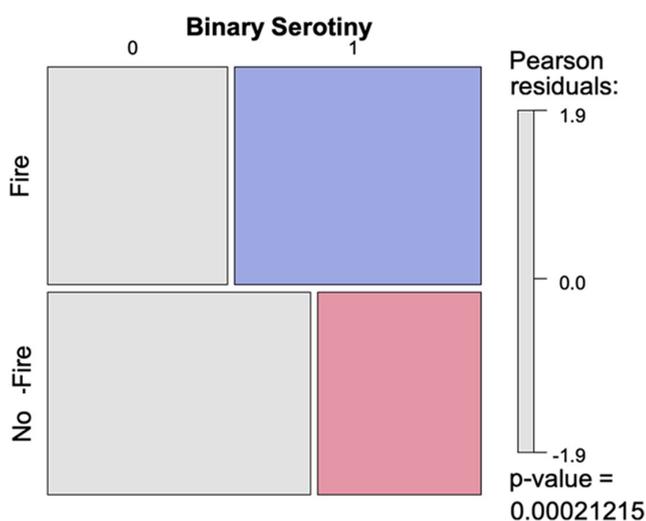
FIGURE 4 Mosaic of the binary serotiny difference between both fire modalities ("Fire" and "No-Fire") for trees expressing serotiny (1) or not (0). Fisher test p-value <.001 (in blue: presence of serotiny in Fire populations; in pink: presence of serotiny in No-Fire populations).

(a)

(b)
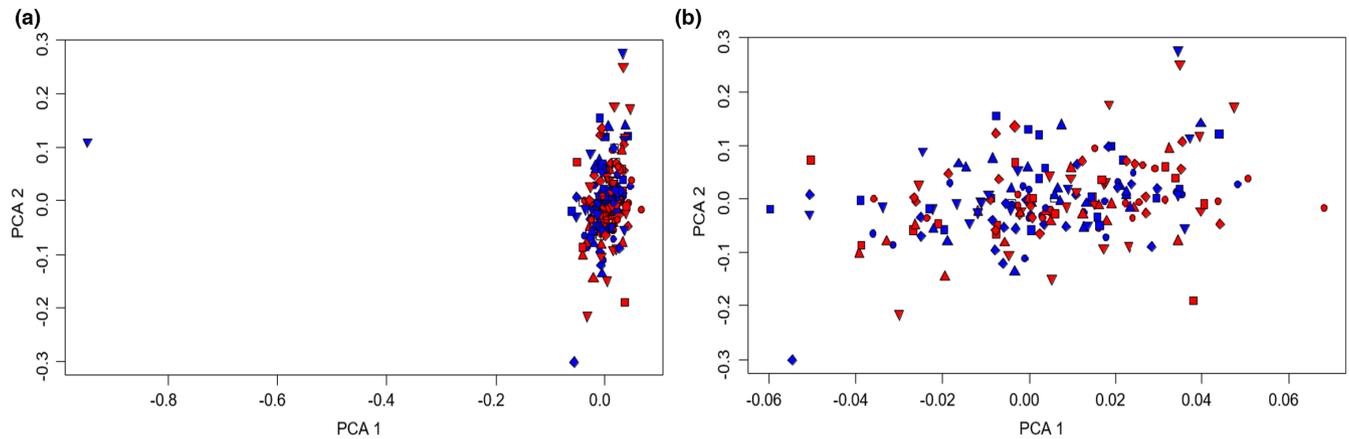
FIGURE 5 Principal component analysis biplots of the genetic structure of *Pinus halepensis* individuals according to the fire modality (a) data distribution on the two first components without removing the outlier, (b) data distribution after having removed the outlier ("Fire" populations in red and "No-Fire" populations in blue).

TABLE 2 Pairwise genetic divergence (Weir and Cockerham F_{ST}) between each *P. halepensis'* "Fire" (PHF) and "No-Fire" (PNHF) populations.

	PHF_1	PHF_4	PHF_5	PHF_6	PHF_8	PHNF_1	PHNF_4	PHNF_5	PHNF_6
PHF_4	5.43E-03								
PHF_5	4.44E-03	4.49E-04							
PHF_6	3.61E-03	3.14E-03	4.36E-04						
PHF_8	5.48E-03	1.65E-03	2.34E-03	1.55E-03					
PHNF_1	1.51E-03	5.48E-03	3.66E-03	4.38E-03	4.95E-03				
PHNF_4	7.35E-03	3.84E-03	1.65E-03	3.40E-03	6.42E-03	4.45E-03			
PHNF_5	7.44E-03	3.10E-03	1.38E-03	4.08E-03	3.85E-03	4.70E-03	1.06E-03		
PHNF_6	6.48E-03	6.52E-03	7.29E-04	3.98E-03	5.73E-03	8.88E-03	4.83E-03	5.47E-03	
PHNF_8	5.87E-03	3.12E-03	−8.32E-05	2.53E-03	2.87E-03	7.46E-03	3.21E-03	2.96E-03	2.90E-03

TABLE 3 Number of SNPs found for each type of data used (*Quantitative 1: IQ, Quantitative 2: EQ,* and *Binary: B*), type of serotiny data correction (*GLM-corrected* or *PCA-corrected*), and type of random forest used (Boruta or VSURF, or union of PCA-corrected and GLM-corrected, or UNION taking into account all of combination between models and correction).

	GLM.corrected		PCA.corrected		Union	UNION
	Boruta	VSURF	Boruta	VSURF		
IQ	3	14	11	14	24	
B	7	12	6	16	22	40
EQ	5	11	6	10	15	

4 | DISCUSSION

Our study provides new important information on the genetic bases of serotiny, one of the most studied fire-related traits in fire-embracer species such as *Pinus halepensis*. Using an in situ genomic association study, we identified several SNPs and a variety of candidate genes that could underlie fire phenotypes associated with the variation of serotiny level observed in the field, as loci associated with serotiny might be associated with other correlated fire adaptive traits. However, our results differed from previous ones

from the point of view of population structure. Indeed, population kinship structure can be affected by the population dynamics driven by fire events (Budde et al., 2017), with potentially closer kinship in fire-prone populations. Nevertheless, we did not detect such patterns in our populations. This could be due to a difference in fire intensity and frequency (fire regime), our study areas being less affected by fires than the ones of Budde et al. (2017), for instance. At the population level, the small levels of divergence are in agreement with previous studies using similar approaches (Scotti et al., 2023).

TABLE 4 Description of the SNPs associated with biological processes of interest.

SNP name	Model	Sequence	Gene name	Biological process
clc_contig_10191_90	EQ	AT5G01990	PILS6	Stress tolerance
clc_contig_1993_3310	EQ	AT2G06210	ELF8	Flowering
clc_contig_2167_2796	EQ	AT5G35750	HK2	Stress tolerance, Flowering
clc_contig_3226_4752	IQ, B	AT5G39710	EMB2745	Reproduction
clc_contig_5541_2946	IQ, B	AT1G55020	LOX1	Stress tolerance
clc_contig_6789_2014	B	AT5G13480	FY	Flowering, Reproduction
Contig1490_73	IQ	AT2G40890	CYP98A3	Stress tolerance
Contig2340_2217	IQ, B	AT4G34100	CER9	Stress tolerance
Contig2371_653	EQ	AT1G01040	DCL1	Flowering, Reproduction
Contig5893_417	IQ, B	AT1G33410	SAR1	Flowering
Contig6652_237	B	AT2G46660	CYP78A6	Reproduction (seed)
Contig6777_720	EQ	AT4G18780	IRX1	Stress tolerance
Phal_PtaS20492475_294	EQ	AT1G28520	VOZ1	Stress tolerance, Flowering
Phal_PtaS25086892_281	B	AT1G71860	PTP1	Stress tolerance
clc_contig_161_4027	EQ	AT5G49160	MET1	Flowering
clc_contig_2079_4430	IQ	AT4G29380	VPS15	Pollen germination
clc_contig_4658_2921	B	AT5G62790	DXR	Stress tolerance
clc_contig_9287_2067	EQ	AT5G05350	PLAC8	Stress tolerance
Contig1655_1637	B	AT3G43300	ATMIN7	Stress tolerance
Contig1165_2751	IQ	AT1G55250	HUB2	Flowering, Reproduction

Previous studies on serotiny suggested that, because trees of different pine species could be serotinous or not, this trait was driven by a simple genetic control involving only one gene (Givnish, 1981; Talluto & Benkman, 2013; Teich, 1970). More recently, serotiny was recognized as a continuous trait, driven by numerous genes involved in several functions (Budde et al., 2014; Hernández-Serrano et al., 2014; Lamont et al., 2020). In our study, the overall distribution of serotiny level was bimodal, with a large peak at zero and a long, flat continuous tail between 3% and 60% and a mean serotiny level (13 ± 11%) much lower than in other studies (e.g., Castellanos et al., 2015). Serotiny level can be driven by many factors, such as environmental conditions or tree age (i.e. ontogeny) and, of course, by fire frequency (Martín-Sanz et al., 2017; Ne'eman et al., 2004; Romero & Ganteaume, 2020). As fire frequency is lower in southeastern France than in some regions of Spain or Portugal (JRC PESETA IV, 2020), this could explain this difference in serotiny level. We also observed high variability in serotiny level among trees, agreeing with previous studies (Hernández-Serrano et al., 2013; Tapias et al., 2004).

"Fire" and "No-Fire" populations significantly differed regarding the presence/absence of serotiny but did not show any difference in the distribution of nonzero (quantitative) values of serotiny. This could be explained by the small difference between our fire modalities (i.e. 0, 1, or 2 fires), and by the fact that fire events have affected the populations only over the last few generations. Indeed, given the small number of generations since the start of the selection and the weak difference between fire regimes, selection could have had a stronger effect on population divergence at a threshold trait (such as presence/absence of serotiny) than at a continuous trait (such as level of serotiny).

Forty SNPs were associated with serotiny, out of which 22 were associated with a binary component of serotiny variation (that is, with the trait being treated as 0 = no serotinous cones, 1 = fraction of serotinous cones above X%; remind that no tree had a proportion of serotinous cones between 0 and X%). While it seems counter-intuitive that a binary component of a trait has a polygenic control, this could be explained by the fact that the action of multiple genes is needed to trigger the development of serotiny, perhaps in a redundant way and through threshold effects. Notice, though, that serotiny is a complex, derived trait that may be driven by many other parameters, themselves under polygenic control, which may explain this result. Six significant genes for binary variation (B and IQ models) were associated with biotic and/or abiotic stress tolerance, such as temperature, drought, light, or pest attack, which were already identified in several previous studies (Brown et al., 2005; Chen et al., 2005; Feraru et al., 2019; Lü et al., 2012; Song et al., 2018; Xing et al., 2010). Budde et al. (2014) found eight loci in P. pinaster associated with serotiny level variation, also involved in stress tolerance linked to water deprivation or high temperatures. Three significant loci lie in genes controlling seed development, end of the dormancy, and embryo sac development (EMB2745, HUB2, and CYP78A6). Seeds from serotinous cones have indeed been shown to differ from those of nonserotinous cones for insulation against high

temperatures, which stimulates their germination compared with those in nonserotinous cones and is linked to a different enzymatic activity (Goubitz et al., 2002, 2004; Moya et al., 2013). However, no indication of gene association between herbivory resistance and serotiny emerged from our GO term analysis. Indeed, granivores can select for low serotiny, since serotiny exposes seeds longer to a risk of consumption (by squirrels that are quite frequent in the study area, for instance).

Quantitative variation in serotiny levels (i.e. EQ models) was also controlled by loci determining responses to water deprivation or high temperature (PILS6, IRX1, HK2, RBP-DR1) but was also driven by genes controlling flowering (i.e. ELF8, HK2, MET1, DCL1, and VOZ1), already identified in previous studies (Celesnik et al., 2013; Shiraya et al., 2008; Tsuzuki et al., 2014). ELF8 (Early flowering 8), DCL1 (Dicer-Like 1), MET1, and VOZ1 (Vascular plant One-Zinc finger) control flowering time (Kankel et al., 2003; Tsuzuki et al., 2014; Yu & Michaels, 2010). Their involvement in the determination of serotiny level suggests that developmental processes of reproductive organs are key components of the trait.

Contig2340_2217 was the only significant locus shared by the IQ and B models. This SNP is located in gene CER9, which could be involved in several important functions for the serotinous cone formation. One of these functions is cuticle development, affecting seed permeability (Sieber et al., 2000), water loss through stomata (Lu et al., 2012), resistance to pathogen attack, plant thermo-tolerance (Li et al., 2017), and flower opening (Chen et al., 2019). Other important functions are wax production (Rashotte et al., 2001), playing a role in seed thermic insulation and water permeability in serotinous cones (Keeley et al., 2011; Moya et al., 2008; Salvatore et al., 2010) and suberin production (Lee & Suh, 2015), also affecting thermal insulation of serotinous cones and seeds. Finally, biosynthesis of isoprenoids (also controlled by gene DXR, identified in B models), which are easily flammable compounds, and the building blocks of terpenes occurring in resin are also involved in serotinous cone formation.

The MLMM approach did not return any significant result. This may be due to differences in precision or stringency between methods, as also suggested by the smaller estimation of background genome-wide heritability obtained with MLMM than with other methods, also controlling for relatedness (Table 1, vs Table 2, Bayesian estimates). As a matter of fact, it is hard to say whether, conversely, the methods that have detected more signals are too liberal; it is well known that, with relatively small populations, allelic effects can be overestimated (Xu, 2003).

While our study has returned significant associations between sequence variants and serotiny, the genomic coverage of our SNP set is quite small. It represents a few percent of gene space—assuming that we represent slightly fewer than 1000 genes and that there are 20,000–30,000 unique genes in a pine genome—and a very small fraction of the very large pine genome. The fact that even just a handful of positive signals could be found with such a narrow sample of the potential variant space suggests that many genomic regions may control the trait.

CONCLUSIONS

We observed a higher proportion of serotinous trees in the "Fire" than in the "No-Fire" populations, notwithstanding the fact that populations have undergone only a few fire events over the last six decades. This suggests that population levels of serotiny can quickly respond to a change in fire frequency, in agreement with previous studies (Talluto & Benkman, 2013). This result confirms the strong relationship between fire and *P. halepensis*, and the ability of this species to grow under recurrent fire events with this key functional trait selected over generations.

Our heritability estimates confirm that serotiny is a moderately to highly heritable trait for *P. halepensis*. Our association analyses returned 40 genes associated with serotiny, approximately half for its presence/absence component and half for its quantitative component, supporting a polygenic control of the trait, contrary to what was suggested in several studies (Givnish, 1981; Talluto & Benkman, 2013; Teich, 1970) but agreeing with the conclusions of Budde et al. (2014) and Lamont et al. (2020). The observed bimodal trait distribution suggests separate genetic bases for the presence/absence of serotinous cones and serotiny level. The variety of gene functions controlling serotiny suggests that adaptation to fire-prone environments through serotiny involves chemical properties (terpene levels), developmental cues (development of reproductive organs), and stress resistance. This stresses the high potential of *P. halepensis* to adapt to fire through its fire-embracer strategy, benefiting from a large and flexible genetic basis of trait variation.

The ongoing climate crisis will increase fire risks and fire frequency in the Mediterranean basin (Dupuy et al., 2020). *P. halepensis* seems to be a good candidate to adapt to this change. Yet, if fires become too recurrent, then other components of serotiny as a trait, namely age to maturity and age of the first serotinous cone production, will become crucial, and will need to be closely scrutinized, to more precisely assess the species' promising resilience level.

AUTHOR CONTRIBUTIONS

Bastien Romero: Conceptualization (supporting); data curation (lead); formal analysis (equal); investigation (lead); methodology (supporting); visualization (equal); writing – original draft (lead); writing – review and editing (equal). **Ivan Scotti:** Formal analysis (lead); funding acquisition (equal); investigation (equal); methodology (equal); supervision (equal); validation (lead); visualization (equal); writing – original draft (equal); writing – review and editing (lead). **Bruno Fady:** Conceptualization (equal); funding acquisition (supporting); investigation (equal); supervision (equal); validation (lead); writing – original draft (supporting); writing – review and editing (supporting). **Anne Ganteaume:** Conceptualization (equal); data curation (supporting); funding acquisition (equal); investigation (supporting); project administration (lead); resources (supporting); supervision (equal); validation (supporting); visualization (supporting); writing – original draft (supporting); writing – review and editing (equal).

CONFLICT OF INTEREST STATEMENT

The authors know of no conflicts of interest associated with this publication.

DATA AVAILABILITY STATEMENT

Single Primer Enrichment Technology genotypes are available at https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/7COC1A. Data on serotiny is accessible at the Dryad Digital Repository: https://doi.org/10.5061/dryad.280gb5mth.

ORCID

Bastien Romero https://orcid.org/0000-0002-3197-5594
Bruno Fady https://orcid.org/0000-0003-2379-7617
Anne Ganteaume https://orcid.org/0000-0003-0444-1981

REFERENCES

Alexa, A., & Rahnenfuhrer, J. (2020). TopGO: Enrichment analysis for gene ontology. R package version 2.40.0. https://bioconductor.org/packages/release/bioc/html/topGO.html

Bond, W. J., & van Wilgen, B. W. (1996). Surviving fires—Vegetative and reproductive responses. In *Fire and plants* (pp. 34–51). Springer and Chapman & Hall.

Bonnafous, F., Fievet, G., Blanchet, N., Boniface, M.-C., Carrère, S., Gouzy, J., Legrand, L., Marage, G., Bret-Mestries, E., & Munos, S. (2018). Comparison of GWAS models to identify non-additive genetic control of flowering time in sunflower hybrids. *Theoretical and Applied Genetics*, 131, 319–332.

Bradshaw, S. D., Dixon, K. W., Hopper, S. D., Lambers, H., & Turner, S. R. (2011). Little evidence for fire-adapted plant traits in Mediterranean climate regions. *Trends in Plant Science*, 16, 69–76.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.

Brieuc, M. S., Waters, C. D., Drinan, D. P., & Naish, K. A. (2018). A practical introduction to random Forest for genetic association studies in ecology and evolution. *Molecular Ecology Resources*, 18, 755–766.

Brown, D. M., Zeef, L. A., Ellis, J., Goodacre, R., & Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *The Plant Cell*, 17, 2281–2295.

Browning, B. L., & Browning, S. R. (2016). Genotype imputation with millions of reference samples. *The American Journal of Human Genetics*, 98, 116–126.

Budde, K. B., González-Martínez, S. C., Navascués, M., Burgarella, C., Mosca, E., Lorenzo, Z., Zabal-Aguirre, M., Vendramin, G. G., Verdú, M., & Pausas, J. G. (2017). Increased fire frequency promotes stronger spatial genetic structure and natural selection at regional and local scales in *Pinus halepensis* Mill. *Annals of Botany*, 119, 1061–1072.

Budde, K. B., Heuertz, M., Hernández-Serrano, A., Pausas, J. G., Vendramin, G. G., Verdú, M., & González-Martínez, S. C. (2014). In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (*Pinus pinaster*). *New Phytologist*, 201, 230–241.

Castellanos, M. C., González-Martínez, S. C., & Pausas, J. G. (2015). Field heritability of a plant adaptation to fire in heterogeneous landscapes. *Molecular Ecology*, 24, 5633–5642.

Causley, C. L., Fowler, W. M., Lamont, B. B., & He, T. (2016). Fitness benefits of serotiny in fire-and drought-prone environments. *Plant Ecology*, 217, 773–779.

Celesnik, H., Ali, G. S., Robison, F. M., & Reddy, A. S. (2013). *Arabidopsis thaliana* VOZ (vascular plant one-zinc finger) transcription factors are required for proper regulation of flowering time. *Biology Open*, 2, 424–431.

Chen, B., Piel, W. H., Gui, L., Bruford, E., & Monteiro, A. (2005). The HSP90 family of genes in the human genome: Insights into their divergence and evolution. *Genomics*, 86, 627–637.

Chen, X., Shi, L., Chen, Y., Zhu, L., Zhang, D., Xiao, S., Aharoni, A., Shi, J., & Xu, J. (2019). Arabidopsis HSP70-16 is required for flower opening under normal or mild heat stress temperatures. *Plant, Cell and Environment*, 42, 1190–1204.

Climent, J., Martín Sanz, R., Santos del Blanco, L., Chambel, M. R., & Notivol-Paíno, E. (2014). Genetic, environmental and ontogenetic effects on cone serotiny in Aleppo pine (*Pinus halepensis* Mill.). In *MEDPINE5. 5th international conference on Mediterranean pines*, Solsona, Spain. https://doi.org/10.13140/2.1.3638.5282

Cruz, O., García-Duro, J., Casal, M., Reyes, O. (2019). Role of serotiny on Pinus pinaster Aiton germination and its relation to mother plant age and fire severity. *iFores*, 12, 491–497. https://10.3832/ifor2968-012.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., & Sherry, S. T. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.

De La Torre, A. R., Puiu, D., Crepeau, M. W., Stevens, K., Salzberg, S. L., Langley, C. H., & Neale, D. B. (2019). Genomic architecture of complex traits in loblolly pine. *New Phytologist*, 221, 1789–1801.

Dupuy, J. L., Fargeon, H., Martin-StPaul, N., Pimont, F., Ruffault, J., Guijarro, M., Hernando, C., Madrigal, J., & Fernandes, P. (2020). Climate change impact on future wildfire danger and activity in southern Europe: A review. *Annals of Forest Science*, 77, 1–24.

Engber, E. A., & Varner, J. M. (2012). Patterns of flammability of the California oaks: The role of leaf traits. *Canadian Journal of Forest Research*, 42, 1965–1975.

Enright, N. J., Fontaine, J. B., Lamont, B. B., Miller, B. P., & Westcott, V. C. (2014). Resistance and resilience to changing climate and fire regime depend on plant functional traits. *Journal of Ecology*, 102, 1572–1581.

Fady, B. (2012). Biogeography of neutral genes and recent evolutionary history of pines in the Mediterranean Basin. *Annals of Forest Science*, 69, 421–428.

Feduck, M., Henry, P., Winder, R., Dunn, D., Alfaro, R., & Hawkes, B. (2015). The genetic basis of cone serotiny in Pinus contorta as a function of mixed-severity and stand-replacement fire regimes. *BioRxiv*, 023267.

Feraru, E., Feraru, M. I., Barbez, E., Waidmann, S., Sun, L., Gaidora, A., & Kleine-Vehn, J. (2019). PILS6 is a temperature-sensitive regulator of nuclear auxin input and organ growth in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 116, 3893–3898.

Fernandes, P. M., Vega, J. A., Jimenez, E., & Rigolot, E. (2008). Fire resistance of European pines. *Forest Ecology and Management*, 256, 246–255.

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2015). VSURF: An R package for variable selection using random forests. *The R Journal*, 7, 19–33.

Givnish, T. J. (1981). Serotiny, geography, and fire in the pine barrens of New Jersey. *Evolution*, 35(1), 101–123.

Goubitz, S., Nathan, R., Roitemberg, R., Shmida, A., & Ne'eman, G. (2004). Canopy seed bank structure in relation to: Fire, tree size and density. *Plant Ecology*, 173, 191–201.

Goubitz, S., Werger, M., & Ne'eman, G. (2002). Germination response to fire-related factors of seeds from non-serotinous and serotinous cones. *Plant Ecology*, 169, 195–204.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184–186.

Hardy, O. J., & 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.

He, T., Lamont, B. B., & Pausas, J. G. (2019). Fire as a key driver of Earth's biodiversity. *Biological Reviews*, 94, 1983–2010.

He, T., Pausas, J. G., Belcher, C. M., Schwilk, D. W., & Lamont, B. B. (2012). Fire-adapted traits of Pinus arose in the fiery cretaceous. *New Phytologist*, 194, 751–759.

Heer, K., Behringer, D., Piermattei, A., Bässler, C., Brandl, R., Fady, B., Jehl, H., Liepelt, S., Lorch, S., & Piotti, A. (2018). Linking dendroecology and association genetics in natural populations: Stress responses archived in tree rings associate with SNP genotypes in silver fir (*Abies alba* Mill.). *Molecular Ecology*, 27, 1428–1438.

Hernández-Serrano, A., Verdú, M., González-Martínez, S. C., & Pausas, J. G. (2013). Fire structures pine serotiny at different scales. *American Journal of Botany*, 100, 2349–2356.

Hernández-Serrano, A., Verdú, M., Santos-del-Blanco, L., Climent, J., González-Martínez, S. C., & Pausas, J. G. (2014). Heritability and quantitative genetic divergence of serotiny, a fire-persistence plant trait. *Annals of Botany*, 114, 571–577.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071.

Kankel, M. W., Ramsey, D. E., Stokes, T. L., Flowers, S. K., Haag, J. R., Jeddeloh, J. A., Riddle, N. C., Verbsky, M. L., & Richards, E. J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics*, 163, 1109–1122.

Kassambara, A., & Mundt, F. (2017). Factoextra: Extract and visualize the results of multivariate data analyses. R package version 1, 337–354.

Keeley, J. E. (2012). Ecology and evolution of pine life histories. *Annals of Forest Science*, 69(4), 445–453.

Keeley, J. E., Pausas, J. G., Rundel, P. W., Bond, W. J., & Bradstock, R. A. (2011). Fire as an evolutionary pressure shaping plant traits. *Trends in Plant Science*, 16, 406–411.

Köppen, W. (1900). Versuch einer Klassifikation der Klimate, vorzugsweise nach ihren Beziehungen zur Pflanzenwelt (Schluss). *Geographische Zeitschrift*, 6, 657–679.

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, 36, 1–13.

Lamont, B. B., Le Maitre, D., Cowling, R., & Enright, N. (1991). Canopy seed storage in woody plants. *The Botanical Review*, 57, 277–317.

Lamont, B. B., Pausas, J. G., He, T., Witkowski, E. T., & Hanley, M. E. (2020). Fire as a selective agent for both serotiny and nonserotiny over space and time. *Critical Reviews in Plant Sciences*, 39, 140–172.

Lee, S. B., & Suh, M. C. (2015). Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop species. *Plant Cell Reports*, 34, 557–572.

Li, L.-M., Lü, S.-Y., & Li, R.-J. (2017). The Arabidopsis endoplasmic reticulum associated degradation pathways are involved in the regulation of heat stress response. *Biochemical and Biophysical Research Communications*, 487, 362–367.

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular ecology*, 24, 1031–1046.

Lü, S., Zhao, H., Des Marais, D. L., Parsons, E. P., Wen, X., Xu, X., Bangarusamy, D. K., Wang, G., Rowland, O., & Juenger, T. (2012). Arabidopsis ECERIFERUM9 involvement in cuticle formation and maintenance of plant water status. *Plant Physiology*, 159, 930–944.

Lu, W., Tang, X., Huo, Y., Xu, R., Qi, S., Huang, J., Zheng, C., & Wu, C. (2012). Identification and characterization of fructose 1, 6-bisphosphate aldolase genes in Arabidopsis reveal a gene family with diverse responses to abiotic stresses. *Gene*, 503, 65–74.

Martín-Sanz, R. C., Callejas-Díaz, M., Tonnabel, J., & Climent, J. M. (2017). Maintenance costs of serotiny in a variably serotinous pine: The role of water supply. *PLoS One*, 12, e0181648.

Mount, D. W. (2007). Using the basic local alignment search tool (BLAST). *Cold Spring Harbor Protocols*, 2007, pdb.top17.

Moya, D., De las Heras, J., Salvatore, R., Valero, E., & Leone, V. (2013). Fire intensity and serotiny: Response of germination and enzymatic activity in seeds of *Pinus halepensis* Mill. from southern Italy. *Annals of Forest Science*, 70, 49–59.

Moya, D., Saracino, A., Salvatore, R., Lovreglio, R., de Las Heras, J., & Leone, V. (2008). Anatomic basis and insulation of serotinous cones in *Pinus halepensis* Mill. *Trees*, 22, 511–519.

Murphy, B. P., Bradstock, R. A., Boer, M. M., Carter, J., Cary, G. J., Cochrane, M. A., Fensham, R. J., Russell-Smith, J., Williamson, G. J., & Bowman, D. M. (2013). Fire regimes of Australia: A pyrogeographic model system. *Journal of Biogeography*, 40, 1048–1058.

Ne'eman, G., Goubitz, S., & Nathan, R. (2004). Reproductive traits of *Pinus halepensis* in the light of fire–a critical review. *Plant Ecology*, 171, 69–79.

Ne'eman, G., Lev-Yadun, S., & Arianoutsou, M. (2012). Fire-related traits in Mediterranean basin plants. *Israel Journal of Ecology and Evolution*, 58, 177–194.

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, 21, 2991–3005.

Pausas, J. G. (2015a). Evolutionary fire ecology: Lessons learned from pines. *Trends in Plant Science*, 20, 318–324.

Pausas, J. G. (2015b). Bark thickness and fire regime. *Functional Ecology*, 29, 315–327.

Pausas, J. G., Alessio, G. A., Moreira, B., & Corcobado, G. (2012). Fires enhance flammability in *Ulex parviflorus*. *New Phytologist*, 193, 18–23.

Pausas, J. G., & Keeley, J. E. (2009). A burning story: The role of fire in the history of life. *Bioscience*, 59, 593–601.

Peeler, J. L., & Menges, E. S. (2018). Effects of fire history, tree age, and canopy seed bank size on serotiny of Ocala sand pine (*Pinus clausa* var. *clausa*) in Florida scrub. *Florida Scientist*, 81, 3–11.

Perry, D. A., & Lotan, J. E. (1979). A model of fire selection for serotiny in lodgepole pine. *Evolution*, 33(3), 958–968.

Pinosio, S., González-Martínez, S. C., Bagnoli, F., Cattonaro, F., Grivet, D., Marroni, F., Lorenzo, Z., Pausas, J. G., Verdú, M., & Vendramin, G. G. (2014). First insights into the transcriptome and development of new genomic tools of a widespread circum-Mediterranean tree species, *Pinus halepensis* Mill. *Molecular Ecology Resources*, 14, 846–856.

Quézel, P. (2000). Taxonomy and biogeography of Mediterranean pines (*Pinus halepensis* and *P. brutia*). In G. Néeman & L. Trabaud (Eds.), *Ecology, biogeography and management of Pinus halepensis and P. brutia forest ecosystems in the Mediterranean Basin* (pp. 1–12). Backhuys Publishers.

R Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, 197, 573–589.

Rashotte, A. M., Jenks, M. A., & Feldmann, K. A. (2001). Cuticular waxes on eceriferum mutants of *Arabidopsis thaliana*. *Phytochemistry*, 57, 115–123.

Ripa, R. R., Franzese, J., Premoli, A. C., & Raffaele, E. (2020). Increased canopy seed-storage in post-fire pine invaders suggests rapid selection mediated by fire. *Basic and Applied Ecology*, 45, 76–85.

Ritland, K. (1996a). _Multilocus mating system program (MLTR), version 1.1._ Department of Botany, University of Toronto.

Ritland, K. (1996b). Estimators for pairwise relatedness and individual inbreeding coefficients. _Genetics Research, 67,_ 175–185.

Romero, B., & Ganteaume, A. (2020). Does recent fire activity impact fire-related traits of _Pinus halepensis_ Mill. and _Pinus sylvestris_ L. in the French Mediterranean area? _Annals of Forest Science, 77,_ 1–19.

Salvatore, R., Moya, D., Pulido, L., Lovreglio, R., López-Serrano, F., De las Heras, J., & Leone, V. (2010). Morphological and anatomical differences in Aleppo pine seeds from serotinous and non-serotinous cones. _New Forests, 39,_ 329–341.

Santos del Blanco, L., Zas, R., Notivol-Paíno, E., Chambel, M. R., Majada, J., & Climent, J. (2010). Variation of early reproductive allocation in multi-site genetic trials of Maritime pine and Aleppo pine. _Forest Systems, 19_(3), 381–392.

Scaglione, D., Pinosio, S., Marroni, F., di Centa, E., Fornasiero, A., Magris, G., Scalabrin, S., Cattonaro, F., Taylor, G., & Morgante, M. (2019). Single primer enrichment technology as a tool for massive genotyping: A benchmark on black poplar and maize. _Annals of Botany, 124_(4), 543–551.

Schwilk, D. W., & Ackerly, D. D. (2001). Flammability and serotiny as strategies: Correlated evolution in pines. _Oikos, 94,_ 326–336.

Scotti, I., Lalagüe, H., Oddou-Muratorio, S., Scotti-Saintagne, C., Ruiz Daniels, R., Grivet, D., Lefevre, F., Cubry, P., Fady, B., González-Martínez, S. C., Roig, A., Lesur-Kupin, I., Bagnoli, F., Guerin, V., Plomion, C., Rozenberg, P., & Vendramin, G. G. (2023). Common microgeographical selection patterns revealed in four European conifers. _Molecular Ecology, 32_(2), 393–411.

Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. _Nature Genetics, 44,_ 825–830.

Shiraya, T., Sato, S., Kato, T., Tabata, S., & Iwasaki, T. (2008). Arabidopsis VIP6/ELF8, the homolog of CTR9 component of the transcriptional complex PAF1, is essential for plant development. _Plant Biotechnology, 25,_ 447–455.

Sieber, P., Schorderet, M., Ryser, U., Buchala, A., Kolattukudy, P., Métraux, J.-P., & Nawrath, C. (2000). Transgenic Arabidopsis plants expressing a fungal cutinase show alterations in the structure and properties of the cuticle and postgenital organ fusions. _The Plant Cell, 12,_ 721–737.

Signorell, A. (2019). DescTools: tools for descriptive statistics. R package, v0. 99.30.

Song, S., Xu, Y., Huang, D., Miao, H., Liu, J., Jia, C., Hu, W., Valarezo, A. V., Xu, B., & Jin, Z. (2018). Identification of a novel promoter from banana aquaporin family gene (MaTIP1; 2) which responses to drought and salt-stress in transgenic _Arabidopsis thaliana. Plant Physiology and Biochemistry, 128,_ 163–169.

Talluto, M. V., & Benkman, C. W. (2013). Landscape-scale eco-evolutionary dynamics: Selection by seed predators and fire determine a major reproductive strategy. _Ecology, 94,_ 1307–1316.

Tapias, R., Climent, J., Pardos, J. A., & Gil, L. (2004). Life histories of Mediterranean pines. _Plant Ecology, 171,_ 53–68.

Teich, A. (1970). Cone serotiny and inbreeding in natural populations of _Pinus banksiana_ and _Pinus contorta. Canadian Journal of Botany, 48,_ 1805–1809.

Tibshirani, R. J., & Efron, B. (1993). An introduction to the bootstrap. _Monographs on Statistics and Applied Probability, 57,_ 1–436.

Tsuzuki, M., Takeda, A., & Watanabe, Y. (2014). Recovery of dicer-like 1-late flowering phenotype by miR172 expressed by the noncanonical DCL4-dependent biogenesis pathway. _RNA, 20,_ 1320–1327.

Wazen, N., Garavaglia, V., Picard, N., Besacier, C., & Fady, B. (2020). Distribution maps of twenty-four Mediterranean and European ecologically and economically important forest tree species compiled from historical data collections. _Annals of Silvicultural Research, 44_(2), 95–101. https://doi.org/10.12899/asr-1933

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. _Evolution, 38,_ 1358–1370.

Wymore, A. S., Keeley, A. T., Yturralde, K. M., Schroer, M. L., Propper, C. R., & Whitham, T. G. (2011). Genes to ecosystems: Exploring the frontiers of ecology with one of the smallest biological units. _New Phytologist, 191,_ 19–36.

Xing, S., Salinas, M., Höhmann, S., Berndtgen, R., & Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. _The Plant Cell, 22,_ 3935–3950.

Xu, S. (2003). Theoretical basis of the Beavis effect. _Genetics, 165_(4), 2259–2268.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. _The American Journal of Human Genetics, 88,_ 76–82.

Yu, X., & Michaels, S. D. (2010). The Arabidopsis Paf1c complex component CDC73 participates in the modification of FLOWERING LOCUS C chromatin. _Plant Physiology, 153,_ 1074–1084.

Zhao, J. H. (2008). Gap: Genetic analysis package. _Journal of Statistical Software, 23,_ 1–18.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. _Bioinformatics, 28,_ 3326–3328.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.