A Genomic Map of Climate Adaptation in Arabidopsis thaliana at a Micro-Geographic Scale

Léa Frachon, Claudia Bartoli, Sebastien Carrere, Olivier Bouchez, Adeline Chaubet, Mathieu Gautier, Dominique Roby, Fabrice Roux

▶ To cite this version:

Léa Frachon, Claudia Bartoli, Sebastien Carrere, Olivier Bouchez, Adeline Chaubet, et al.. A Genomic Map of Climate Adaptation in Arabidopsis thaliana at a Micro-Geographic Scale. Frontiers in Plant Science, Frontiers, 2018, 9, 10.3389/fpls.2018.00967. hal-02373598

HAL Id: hal-02373598
https://hal.archives-ouvertes.fr/hal-02373598
Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A Genomic Map of Climate Adaptation in Arabidopsis thaliana at a Micro-Geographic Scale

Léa Frachon1,2,†, Claudia Bartoli1†, Sébastien Carrère1, Olivier Bouchez4, Adeline Chaubet4, Mathieu Gautier5, Dominique Roby1 and Fabrice Roux1*

1 Laboratoire des Interactions Plantes-Microorganismes, Université de Toulouse, Institut National de la Recherche Agronomique, Centre National de la Recherche Scientifique, Castanet-Tolosan, France, 2 Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Naples, Italy, 3 Department of Systematic and Evolutionary Botany, University of Zurich, Zürich, Switzerland, 4 Institut National de la Recherche Agronomique, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France, 5 Centre de Biologie pour la Gestion des Populations, Institut National de la Recherche Agronomique, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Institut de Recherche pour le Développement, Montpellier SupAgro, Université de Montpellier, Montpellier, France

Understanding the genetic bases underlying climate adaptation is a key element to predict the potential of species to face climate warming. Although substantial climate variation is observed at a micro-geographic scale, most genomic maps of climate adaptation have been established at broader geographical scales. Here, by using a Pool-Seq approach combined with a Bayesian hierarchical model that control for confounding by population structure, we performed a genome–environment association (GEA) analysis to investigate the genetic basis of adaptation to six climate variables in 168 natural populations of Arabidopsis thaliana distributed in south-west of France. Climate variation among the 168 populations represented up to 24% of climate variation among 521 European locations where A. thaliana inhabits. We identified neat and strong peaks of association, with most of the associated SNPs being significantly enriched in likely functional variants and/or in the extreme tail of genetic differentiation among populations. Furthermore, genes involved in transcriptional mechanisms appear predominant in plant functions associated with local climate adaptation. Globally, our results suggest that climate adaptation is an important driver of genomic variation in A. thaliana at a small spatial scale and mainly involves genome-wide changes in fundamental mechanisms of gene regulation. The identification of climate-adaptive genetic loci at a micro-geographic scale also highlights the importance to include within-species genetic diversity in ecological niche models for projecting potential species distributional shifts over short geographic distances.

Keywords: Arabidopsis thaliana, Bayesian hierarchical model, climate change, genome–environment association analysis, local adaptation, Pool-Seq, spatial grain

INTRODUCTION

In the context of contemporary climate change, a major goal in evolutionary ecology is to understand and predict the ability of a given species to persist in presence of novel climate conditions (Bay et al., 2017). A lack of response of species to selection due to climate change would cause an erosion of biodiversity by disrupting ecosystems sustainably (Pecl et al., 2017).
Overall, species can adopt three non-exclusive responses to face the altered and fluctuating climate conditions (Aitken et al., 2008; Hansen et al., 2012; Pecl et al., 2017). Firstly, species can migrate to track current climate spatial shifts. This response can, however, be limited for (i) long-distance dispersal organisms because of the presence of multiple anthropogenic barriers (Ewers and Didham, 2006), and (ii) organisms with restricted dispersal as for example sessile plants lacking dispersal mechanism or disperser reward (Wang et al., 2016).

Secondly, organisms can rapidly acclimate to novel climate conditions via phenotypic plasticity, defined as the ability of a given genotype to produce different phenotypes when exposed to different environmental conditions (Fusco and Minelli, 2010). Despite its theoretical benefits to help natural populations to reach a new phenotypic optimum (Lande, 2009; Chevin et al., 2010), adaptive phenotypic plasticity is not as frequent as expected in nature because it can be constrained by several costs and limits (initially reviewed in DeWitt et al., 1998). For example, one of the main limits concerns the unreliability of environmental cues, leading to non-adaptive or mal-adaptive plastic responses (van Kleunen and Fischer, 2005). In the context of climate change, such unreliable cues can correspond to extreme climate events that fall outside the range of the climate conditions encountered by natural populations over their history (Orlowsky and Seneviratne, 2012).

Thirdly, over a longer term, organisms can adapt to novel climate conditions via genetic selection, provided that there is sufficient standing genetic variation or new genetic variation arising from either de novo mutations or immigration of climate-adapted alleles from nearby populations (Hoffmann and Sgrò, 2011; Bay et al., 2017). Predicting the response of species to climate change therefore requires the description of the genomic architecture (i.e., number of genes, allelic effects, locations across the genome) underlying climate adaptation. For this purpose, two major approaches can be adopted, based on either phenotype–genotype or ecology–genotype associations across the genome (Bergelson and Roux, 2010; Bay et al., 2017). Few genome-wide association mapping studies (GWAS) reported the genomic architecture associated with phenotypic traits potentially related to climate adaptation such as thermal sensitivity (Li et al., 2014). On the other hand, based on the assumption that each population is adapted to local environmental conditions, the most exploited approach corresponds to genome–environment association (GEA) analyses. In this case, a genome scan is performed to identify significant associations between genetic polymorphisms and environmental variables (Yoder et al., 2014; Abebe et al., 2015; Lasky et al., 2015; Rellstab et al., 2015; Hoban et al., 2016; Manel et al., 2016). Due to publicly available gridded estimates of climate and the development of next-generation sequencing (NGS) technologies, the number of GEA analyses performed on climate variables rapidly increased in the last few years (Bay et al., 2017). In most studies, the genomic architecture of climate adaptation was found to be highly polygenic, with hundreds to thousands of small-effect genetic variants scattered across the genome (Bay et al., 2017).

Most of the GEA analyses on climate were performed at large spatial scales (i.e., from several hundred to several thousand kilometers). However, substantial climate variation can also be observed at smaller spatial scales (from several tens of meters to several tens of kilometers), leading for example to sharp climate gradients in mountains (Manel et al., 2010; Kubota et al., 2015) or a mosaic of climatically optimal and suboptimal sites within the reach of gene flow among populations (Pluess et al., 2016). The complementarity of performing GEA analyses from continental to local geographical scales should shed light on the genetic bases underlying coarse-grained and fine-grained climate variation (Manel et al., 2010), which in turn would increase the reliability of predictions of response to climate change. In addition, as previously advised for adaptive phenotypic traits (Bergelson and Roux, 2010), working at a small geographical scale should reduce the limitations of GEA analyses often observed when working at larger geographical scales such as the confounding background produced by population structure, rare alleles and genetic/allelic heterogeneity. Finally, a fine-grained spatial scale is much more coherent with the mean distance of species migration (few km per decade; Chen et al., 2011).

*Arabidopsis thaliana* is a widely distributed annual selfing plant species found in a large range of climate environments across its native range in Eurasia (Hoffmann, 2002). Given the main barochorous mode of seed dispersal of *A. thaliana*, seeds are dispersed over short distances (i.e., a few ten of centimeters, Wender et al., 2005), thereby limiting the potential to track current climate poleward shifts. Although adaptive plastic responses to climate change (in particular seasonal climate change) have been observed in *A. thaliana* (Fournier-Level et al., 2016), a growing number of studies also reported the importance of genetic selection underlying climate adaptation in *A. thaliana* (Frachon et al., 2017). Notably, reciprocal transplants performed at the European scale revealed that climate gradients likely play a major role in local adaptation of *A. thaliana* (Fournier-Level et al., 2011; Ågren and Schemske, 2012; Ågren et al., 2013; Wilczek et al., 2014). Furthermore, amongst plant species, *A. thaliana* pioneered the identification of climate-adaptive genetic loci at the genome-wide scale. A GEA analysis based on 948 Eurasian accessions succeeded to establish a genomic map of local adaptation to climate variation, which in turn successfully predicted the relative fitness of a subset of accessions grown together in a common garden in the north of France (Hancock et al., 2011). Local adaptation to climate also explains a substantial portion of genomic variation of *A. thaliana* at a regional scale (i.e., several hundred km; Méndez-Vigo et al., 2011; Lasky et al., 2012; Vidigal et al., 2016; Tabas-Madrid et al., 2018). However, studies reporting the genomic architecture of adaptation to various climate variables at a finer spatial grain (from several tens of meters to several tens of kilometers) are still scarce in *A. thaliana*, despite the identification of strong climate–phenotype associations either along sharp altitudinal gradients (Montesinos-Navarro et al., 2011; Luo et al., 2015; Günther et al., 2016; Tabas-Madrid et al., 2018) or in a mosaic of climatically contrasted sites (Brachi et al., 2013).
In this study, we aimed to establish a genomic map of climate adaptation in *A. thaliana* at a micro-geographic scale. We focused on a new set of 168 natural populations distributed homogeneously in the south-west of France (Bartoli et al., 2018), a geographical region under the influence of three contrasted climates (i.e., oceanic climate, Mediterranean climate and mountain climate). By using a Bayesian hierarchical model that control for confounding by population structure (Gautier, 2015), we conducted a GEA analysis between 1,638,649 SNPs and six fine-grained climate variables. Because most *A. thaliana* natural populations located in France are genetically diverse (Le Corre, 2005; Platt et al., 2010; Brachi et al., 2013), we obtained a representative picture of within-population genetic variation across the genome by adopting a Pool-Seq approach. We then searched for genome-wide signatures of selection on the SNPs the most associated with climate variation. Following Hancock et al. (2011) and Brachi et al. (2015), we therefore tested whether those top SNPs were enriched either for non-synonymous variants or in the extreme tail of a genome-wide spatial differentiation scan. We finally discussed the function of the main candidate genes.

**MATERIALS AND METHODS**

**Plant Material**

A field prospection in May 2014 allowed the identification of 233 *A. thaliana* natural populations in the Midi-Pyrénées region (south-west of France). In agreement with an important population turnover of natural populations observed in *A. thaliana* (Picó, 2012), individuals were present in only 168 populations (~72.1%) in late winter 2015 when the sampling campaign was performed (Supplementary Table S1). The average distance among the 168 populations was 100.6 km (median = 93.4 km, max = 265.2 km, min = 0 km, first quartile = 57.3 km, third quartile = 137.2 km). To our knowledge, no natural populations of *A. thaliana* have been previously sampled in this geographic region.

**Climate Characterization**

The 168 geo-localized populations were characterized for 20 biologically meaningful climate variables retrieved from the ClimateEU database (Table 1). Climate data has been generated with the ClimateEU v4.63 software package based on methodology described by Hamann et al. (2013). The grid resolution of the 20 climate variables (~1.25 arcmin, ~600 m) was smaller than the average distance among populations. Climate data were averaged across the 2003–2013 annual data. In addition, altitude was obtained from http://www.gps-coordinates.net. Following Hancock et al. (2011), the set of 21 climate variables was pruned based on the pairwise Spearman correlations of the variables (Supplementary Figure S1), by taking into account only variables that did not display a Spearman’s rho greater than 0.8. This step was performed to avoid inter-correlation between two variables. In cases where variables were strongly inter-correlated, we selected the variable with the most obvious link to the ecology of *A. thaliana*. The final set of six non-correlated climate variables considered in this study was composed by mean annual temperature, mean coldest month temperature and precipitations in winter, spring, summer, and autumn.

In order to compare the level of climate variation among the 168 populations of the Midi-Pyrénées region with the level of climate variation among natural populations of *A. thaliana* located in Europe, we first extracted data for the six climate variables (ClimateEU database) for 521 locations where *A. thaliana* have been collected and geo-localized in Europe, including 95 French locations (Hancock et al., 2011; Brachi et al., 2013). To visualize the climatic space encountered by *A. thaliana* at different geographical scales, we then performed a principal component analysis (PCA) based on the 689 locations by using the ade4 1.7-6 version package in the R environment (Chessel et al., 2004; Dray et al., 2017). Finally, the percentage of climatic variation in Europe observed among the 168 populations was calculated by dividing the extent of variation observed on the two first Principal Components (PC_climate) at the scale of the Midi-Pyrénées region by the extent of variation observed at the European scale.

**Spatial Grains of Climatic Variation**

To estimate the spatial grain of variation of each climate variable, a spectral decomposition of the spatial relationships among the 168 populations was first modeled with Principal Coordinates of Neighbor Matrices (PCNM), by running the pcnm() function implemented in the vegan package (R package version 2.3-5; Oksanen et al., 2016) using the Euclidean distance matrix based on the GPS coordinates of the 168 populations. This analysis allows decomposing the spatial structure among the sites under study into orthogonal PCNM components corresponding to successive spatial grains (Borcard and Legendre, 2002). The first PCNM components define a large spatial grain, while the last PCNM components correspond to finer grains (Borcard et al., 2004; Ramette and Tiedje, 2007). All PCNM components were then used as explanatory variables in a multiple linear regression on each climate variable. To account for multiple testing, a Benjamini–Hochberg procedure was performed for each climate variable across all PCNM components to control for a false discovery rate (FDR) at a nominal level of 5% (Benjamini and Hochberg, 1995).

**Genomic Characterization and Data Filtering**

For each population, a mean number of 16.5 plants (total number = 2,776 plants, median = 17 plants, max = 17 plants, min = 5 plants, Supplementary Table S1) were collected randomly in late February – early March 2015 and brought back to a cold frame greenhouse (no additional light or heating). In April 2015, leaf tissue was collected from approximately 16 plants per population for a total of 2,574 plants (min = 5 plants, max = 16 plants, mean = 15.32 plants, median = 16 plants, Supplementary Table S1). More precisely, a portion of a rosette leaf for each plant was placed in 96-well Qiagen S-block plates containing a
3 mm bead in each well and samples were stored at −80°C. Prior to DNA extraction, plates were put 30 s in liquid nitrogen and samples were then crushed by using Mixer Mill MM 300 Retsch® with 1 min of vibration at a frequency of 30 vibrations/s. Genomic DNA from the 2,574 plants was extracted as described in Brachet et al. (2013) and total DNA for each individual extraction was quantified by using a Quant-iT dsDNA Assay Kit with a qPCR ABI7900 machine. Individuals from each population were then used to constitute an equimolar pool. From 50 to 500 bp fragments were produced for each pool by using Covaris M220 Focused-ultrasonicator by following manufacturer instructions. The quality of the libraries was assessed with the Advanced Analytical Fragment Analyzer and libraries were then quantified by qPCR by using the Kapa Library quantification Kit. Samples were sequenced at the Get-PlaGe core facility (INRA of Toulouse) on an Illumina HiSeq 3000 sequencer by using a paired-end read length of 2 × 150 bp with the Illumina HiSeq3000 Reagents Kits. In particular, each Illumina lane was composed of a pool of 14 DNA samples. Twelve lines were used in total to sequence the DNAs (i.e., 14 A. thaliana populations were sequenced on each Illumina lane). Raw data for each population used in this study are available at the NCBI Sequence Read Archive (SRA) through the study accession SRP103198.

Raw reads were mapped on the reference genome Col-0 with glint tool (version 1.0.0c8.779) (Faraut and Courcelle, unpublished software) by using the following parameters: glint mappe –no-lc-filtering –best-score –mmis 5 –limin 80 –step 2. The mapped reads were filtered for proper pairs with SAMtools (v0.01.19; Li et al., 2009) (samtools view -f 0x01 0x02 (v0.01.19; Li et al., 2009)). A semistringent SNP-Calling across the genome was then performed for each population with SAMtools mpileup (Li et al., 2009) and VarScan mpileup2snp (Koboldt et al., 2012) software by using as parameters a minimum coverage (minimum read depth at a position to make a call) of five reads and a minimum variant allele frequency threshold of 0.00001. SNP-Pooling was then performed to obtain polymorphic sites across the pool of the 168 populations and SNP-Calling was inferred on the whole polymorphic sites as described above (VarScan mpileup2cnv; min coverage = 1). The bi-allelic positions were then selected. The mean and the median coverage to a unique position in the reference genome were ~26.3× and ~24.5×, respectively (min = 11.60×, max = 48.69×).

After bioinformatics analysis, the allele read count matrix (for both the reference and alternate alleles) was composed by 4,781,661 SNPs across the 168 populations. This data set was further filtered using custom scripts. First, SNPs without mapped reads in at least eight populations were removed (number of remaining SNPs = 3,798,406). Second, for each population, we calculated the relative coverage of each SNP as the ratio of its coverage to the median coverage (computed over all populations). The final set of six non-correlated climate variables considered in this study are in bold.

| Variable       | Description                                                                 | Source                      | Grid resolution* |
|----------------|------------------------------------------------------------------------------|-----------------------------|-----------------|
| Altitude       | Altitude (m)                                                                 | www.coordonnees-gps.fr      | –               |
| MAT            | Mean annual temperature (°C)                                                 | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| MWMT           | Mean warmest month temperature (°C)                                         | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| MCMT           | Mean coldest month temperature (°C)                                         | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| Tave_at        | Autumn (September–November) mean temperature (°C)                           | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| Tave_wt        | Winter [December (previous year)–February] mean temperature (°C)            | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| Tave_sp        | Spring (March–May) mean temperature (°C)                                    | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| Tave_sm        | Summer (June–August) mean temperature (°C)                                  | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| Tave_at        | Autumn (September–November) mean temperature (°C)                           | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| PPT_wt         | Winter precipitation (mm)                                                   | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| PPT_sp         | Spring precipitation (mm)                                                   | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| PPT_sm         | Summer precipitation (mm)                                                   | ClimateEU 1.25 arcmin       | 1.25 arcmin     |
| PPT_at         | Autumn precipitation (mm)                                                   | ClimateEU 1.25 arcmin       | 1.25 arcmin     |

*The final set of six non-correlated climate variables considered in this study are in bold.

The data for this study are available at the NCBI sequence read archive (SRA) through the study accession SRP103198.
the SNPs in the corresponding population). Because multiple gene copies in the 168 populations can map to a unique gene copy in the reference genome Col-0, we removed SNPs with a mean relative coverage across the 168 populations above 1.5 (number of remaining SNPs = 3,260,041). In addition, we removed SNPs with a standard deviation of allele frequency across the 168 populations below 0.004 (number of remaining SNPs = 3,248,168). Third, because genomic regions present in Col-0 can be absent in most of the 168 populations or genomic regions present in most of the 168 populations can be absent in Col-0, we removed SNPs with a mean relative coverage across the 168 populations below 0.5 (number of remaining SNPs = 3,172,313). Fourth, because of bias in GWA/GEA analysis due to rare alleles (Bergelson and Roux, 2010), we removed SNPs that were monomorphic in more than 90% of the populations, leading to a final data set consisting of read counts for 1,638,649 SNPs in the 168 populations.

**Genome–Environment Association Analysis**

Based on the 1,638,649 SNPs, whole genome scans for adaptive differentiation and association with climate variables were performed with BayPass 2.1 (Gautier, 2015). Dealing with Pool-Seq data, the underlying Bayesian hierarchical models explicitly account for the scaled covariance matrix of population allele frequencies (Ω) which make the analyses robust to complex demographic histories.

Since the large number of SNPs analyzed here, we adopted a sub-sampling procedure to estimate Ω. This procedure consisted in dividing the full data set into 32 sub-data sets, each containing 3.125% of the 1,638,649 SNPs (51,207 SNPs taken every 32 SNPs along the genome), that were further analyzed in parallel under the core model using default options for the Markov Chain Monte Carlo (MCMC) algorithm (except -npilot 15 -pilotlength 500 -burnin 2500). Pairwise comparisons of the 32 resulting covariance matrices confirmed that all estimates were consistent with highly correlated elements. In addition, the pairwise FMD distances (Fürstner and Moonen, 2003) had a narrow range of variation (from 2.04 to 2.24) with a mean value equal to 2.15.

These analyses carried out under the core model also provided the estimation of the XTX measure of differentiation for all the SNPs (combined over sub-data sets). For a given SNP, the XTX is defined as the variance of the standardized population allele frequencies (Günther and Coop, 2013; Gautier, 2015). This allows for a robust identification of highly differentiated allele frequencies, i.e., rescaled using \( \hat{X}_{i} \). The coordinates of the two first resulting Principal Components (PC\(_{\text{genomic}} \)) were then regressed against latitude and longitude of population \( i \) according to the following formula (PROC GLM procedure in SAS 9.3 SAS Institute Inc., Cary, NC, United States):

\[
\text{PC}_{\text{genomic}} \text{ coordinates}_1 = \text{latitude}_i + \text{longitude}_i
\]

\[
\text{PC}_{\text{genomic}} \text{ coordinates}_2 = \text{latitude}_i' \text{longitude}_i + \varepsilon_i
\]

Finally, genome-wide analysis of association with climate covariables were carried out under the AUX model (-auxmodel) parameterized with \( \Omega \). The underlying model consists in a linear regression of the SNP population allele frequencies with the population-specific covariables while accounting for the shared covariance structure of allele frequencies as captured by the matrix \( \Omega \). In the AUX model, a Bayesian (binary) auxiliary variable \( \delta \) is assigned to each SNP regression coefficient that indicates whether a specific SNP can be regarded as associated to the covariable (\( \delta = 1 \)) or not (\( \delta = 0 \)). By looking at the posterior mean of each SNP auxiliary variables (known as Posterior Inclusion Probability or PIP, in the model averaging literature), it is then straightforward to derive a Bayes Factor (BF) to compare models considering the SNP as associated or not (Gautier, 2015). Hence, the support for association of each SNP with each covariable \( k \) (i.e., a non-null regression coefficient \( \hat{\beta}_{ik} \) between SNP \( i \) allele frequencies and a covariable \( k \) ) was evaluated by computing BF measured in deciban units (dB) (Gautier, 2015).

Note that the AUX model allows to explicitly accounting for multiple testing issue by integrating over (and estimating) the unknown proportion of SNPs actually associated with a given covariable. Here, six climate variables were considered separately and standardized prior to analyses using the scalecov option. In practice, BF and the associated regression coefficients \( \hat{\beta}_i \) (mean of the product \( \delta \hat{\beta}_i \) over the corresponding posterior sampled values) between SNP allele frequencies and climate variation were estimated for each SNP by analyzing in parallel the 32 sub-data sets described above (but with the same matrix \( \Omega_1 \) using default options (except -npilot 15 -pilotlength 500 -burnin 2500) for MCMC.

**Testing the Power of GEA Analysis to Identify True Positives**

Because climate variation was significantly associated with the first PC\(_{\text{genomic}} \) (Supplementary Table S2), we performed a simulation study aiming at characterizing how the degree of correlation between a given environmental covariable and the major axis of population structure might affect the sensitivity of BayPass to detect associated SNPs. For this purpose, using the R function simulate.PCcorrelated.covariate that we specifically developed in this study (Supporting Information of Supplementary Material) and that will be integrated in a future release of the BayPass software, we first simulated 10 environmental covariable vectors \( Z_i \) (of length 168, the number of populations) with a Spearman’s coefficient correlation \( \rho_1 \) ranging from 0 to 0.9 (\( \rho_1 = 0.1(i-1) \) for \( i = 1, \ldots, 10 \)) with the first PC\(_{\text{genomic}} \). For each covariable vector, 25 data sets consisting of read counts for 10,000 SNPs including 100 associated SNPs (with a regression coefficient set at 0.1) were simulated under the BayPass inference model using the function simulate.baypass()
Enrichment in Signatures of Selection for Climatic Associations

For each climate variable, we tested whether SNPs with the highest BF were over-represented in the extreme tail of the BF distribution according to the methodology described in Brachi et al. (2015):

\[ FE_{\text{XtX}} = \frac{n_a/N}{\frac{S}{S_a}} \]

where \( n_a \) is the number of XtX in the 0.5% upper tail of the BF distribution, \( N_a \) is the number of SNPs in the 0.5% upper tail of the BF distribution that were also in the 0.5% upper tail of the BF distribution and \( N \) is the total number of SNPs tested genome-wide and \( N_a \) is the number of SNPs in the 0.5% upper tail of the BF distribution. Statistical significance of enrichment was assessed by running 10,000 null permutations based on the methodology described in Hancock et al. (2011).

Identification of Candidate Genes Associated With Climate Variation

A three-step procedure was adopted to identify candidate genes associated with climate variation. Firstly, we selected the 50 SNPs with the highest BF for each of the six climate variables, leading to a total of 300 SNPs. Secondly, for each climate variable, candidate regions were defined on the genomic regions supported by at least three top SNPs successively separated by less than 10 kb. This step led to the identification of 12 candidate regions, including one region detected for both mean annual temperature and mean coldest month temperature. The 12 candidate regions contain on average 3.6 SNPs (median = 4 SNPs, min = 3 SNPs, max = 6 SNPs) and have a mean length of 10.1 kb (median = 9.5 kb, min = 2.4 kb, max = 16.9 kb). Finally, using the TAIR 10 database we retrieved all the annotated genes located within or overlapping with the 13 candidate regions, leading to the identification of 34 annotated genes.

RESULTS

Climate Variation and Associated Spatial Grains

Here, we focused on 168 A. thaliana natural populations distributed in the Midi-Pyrénées region located in the south-west of France (Figure 1). We identified 82 Principal Coordinates of Neighbor Matrices (PCNM) components, suggesting a relatively homogeneous spatial distribution of the 168 populations across the sampling area (Figure 1).

---

3https://www.arabidopsis.org/
Using Pool-Seq data, we estimated within-population allele frequencies across the genome for a final number of 1,638,649 SNPs (i.e., one SNP every 72 bp). Based on SVD of the population covariance-variance matrix $\Omega_1$, we found that 96.4% of the genomic variation observed in the Midi-Pyrénées region was explained by the first PC_{genomic} (Supplementary Figure S3), reinforcing the pattern of strong population subdivision already observed in other French regions (Le Corre, 2005; Brachi et al., 2013). In addition, a weak geographic pattern along a south-west/north-east axis was observed for genomic variation (latitude: $t$ value = 4.734, $P = 4.73 \times 10^{-6}$, longitude: $t$ value = 4.417, $P = 1.81 \times 10^{-5}$, latitude $\times$ longitude: $t$ value = $-4.428$, $P = 1.73 \times 10^{-5}$, adjusted $R^2 = 10.5%$; Supplementary Figure S3).

To identify the genomic regions associated with climate variation, we then performed a genome-wide scan for association with the six climate variables, using a Bayesian hierarchical model that includes a population covariance matrix accounting for the neutral covariance structure across population allele frequencies. For each climate variable, we estimated the regression coefficients between SNP allele frequencies and climate variation ($\beta_i$) and evaluated the support for association (non-null $\beta_i$) of the association between a given SNP and a climate variable with a Bayes factor (BF). By applying this method, we identified neat and strong peaks of association for most of the climate variables (Figure 3 and Supplementary Figure S4). Accordingly, as illustrated for the mean annual temperature and the winter precipitations, standardized allele frequencies variation of the most significant SNPs strongly overlapped with climate variation (Figure 4).

Because climate variation was significantly associated with the first PC_{genomic} (absolute Spearman’s correlation coefficients ranging from 0.148 to 0.701; Supplementary Table S2), we evaluated to which extent the level of correlation of the analyzed covariables with this axis can have affected the performance of GEA analyses. We analyzed 250 data sets simulated under the inference model and each data set consisted of 10,000 SNPs including 100 SNPs associated with a given environmental covariable out of 10 simulated ones (25 data sets per covariable) that displayed different levels of correlation with the first PC_{genomic} (ranging from 0 to 0.9 with an increasing step of 0.1). The simulation model was calibrated to obtain data sets closely mimicking our observed climate data, both in terms of the sample design (number of populations, haploid sample size of the pools, read coverage, across population allele frequencies) and population structure (summarized by the matrix $\Omega$). As expected, performances of the model for detecting environmental covariable – SNPs associations decreased (i.e., for a given power both the false discovery and false positive rates increased) when the correlation of the considered environmental covariable with the first PC_{genomic} increased (Figure 5). For instance, ROC and PR AUC were minimal for Spearman’s $\rho = 0.9$ (ROAuc = 0.813 and PRAuc = 0.236) (Figure 5). Interestingly, up to correlation values equal to 0.5 (i.e., beyond the range of our climate variables of interest), performances of the model remained similar, with ROC AUC ranging from 0.936 (Spearman’s $\rho = 0.4$) to 0.953 (Spearman’s $\rho = 0.1$) and PR AUC ranging from 0.629 (Spearman’s $\rho = 0.2$) to 0.701 (Spearman’s $\rho = 0.1$) (Figure 5).
Altogether, these results suggest that a large fraction of the SNPs the most associated with climate variation correspond to true positives.

**Signatures of Selection**

To test whether our results detect true signals of adaptation, we first tested whether the top SNPs (i.e., the 0.5% upper tail of the BF distribution) were differentially enriched for six categories of genetic variants (intragenic, intergenic, UTR, intron, synonymous and replacement). For the variables mean annual temperature, mean coldest month temperature and winter precipitations, the highest significant fold enrichment was observed for replacement SNPs (Figure 6). For the variable summer precipitations, the highest significant fold enrichment was detected for SNPs located in UTR regions (Figure 6). No significant enrichment was detected for spring and autumn precipitations variables after a correction for multiple comparisons (Figure 6).

We then tested whether the top SNPs (i.e., the 0.5% upper tail of the BF distribution) significantly overlapped with the 0.5% upper tail of the XtX distribution (i.e., the 0.5% most overly differentiated SNPs among populations). For all climate variables with the exception of winter precipitations, climate-related SNPs were significantly enriched in the 0.5% upper tail of the XtX distribution ($P < 0.001$), with a fold-enrichment ranging from 1.70 for mean coldest month temperature to 3.14 for spring precipitations.

**Identification of Candidate Genes**

To identify candidate genes associated with climate variation, we retrieved all the annotated genes located within or overlapping with 12 candidate regions (each supported by at least three top SNPs), including one region at the beginning of chromosome 5 detected for both mean annual temperature and mean coldest month temperature (Supplementary Table S3). By considering a list of 34 candidate genes, a literature survey identified different
functions encoded by these genes that can be classified in two main categories: (i) a large number of proteins \((n = 16)\) involved in the regulation of gene expression, including RNA binding proteins and transcriptional factors, but also miRNAs or transposable elements; (ii) genes involved in abiotic (or biotic) stress response \((n = 6)\), including genes involved in the regulation of salt stress, cold/high temperature stress, light/dark treatment, or pathogen attack such as ACR8, RZ-1C, AN1 like zing finger protein, MAF1/FLM or extensin-like protein (ELP). Other candidate genes correspond to diverse general functions, including developmental regulators, or to unknown functions.

**DISCUSSION**

**Climate Adaptation in a Patchy Climate Environment**

In agreement with the influence of three contrasted climates (i.e., oceanic climate, mountain climate, and Mediterranean climate) in the south-west of France, up to 24% of climate variation across European locations inhabited by *A. thaliana* was observed in the Midi-Pyrénées region. The presence of mountains in the south and north-east in our sampling area likely explained the steep temperature gradients observed in this study. On the other hand, we observed a mosaic of contrasted precipitation regimes. Therefore, the different spatial grains between temperature and precipitations lead to rugged climate landscapes over very short geographical scales, which in turn better match with the small distance of seed and pollen dispersal expected for a barochorous and selfing plant species such as *A. thaliana*.

In comparison with studies performed at larger geographical scales (Hancock et al., 2011; Lasky et al., 2012), we identified very neat and strong peaks of association with climate variation. Such a pattern is similar to a previous GWAS in *A. thaliana* reporting that the significance level of association peaks for phenological traits potentially related to climate adaptation (such as flowering time) was stronger based on regional or local accessions than worldwide or European accessions (Brachi et al., 2013). Two non-exclusive hypotheses can be suggested to explain the strong SNP-climate associations detected in our study. First, because the genetic variability of most natural populations of *A. thaliana* has long been considered to be low (likely due to its selfing rate close to 98%; Platt et al., 2010), most genome–climate association studies performed at the European or regional scale have been based on very few accessions per population, i.e., ~1.7 accession per population (Hancock et al., 2011; Lasky et al., 2012). However, recent studies challenged this view, and many natural populations have been described to be highly genetically variable at both neutral SNPs and polymorphisms associated within natural phenotypic variation (Le Corre, 2005; Picó et al., 2008; Lundemo et al., 2009; Montesinos et al., 2009; Bomblies et al., 2010; Platt et al., 2010; Kronholm et al., 2012; Samis et al., 2012; Brachi et al., 2013; Huard-Chauveau et al., 2013; Karasov et al., 2014; Luo et al., 2015; Debieu et al., 2016; Frachon et al., 2017). Since pool sequencing has been demonstrated a cost-effective method to infer demography and to identify genetic markers underlying local adaptation in several plant and animal species (Schlötterer et al., 2014), we obtained a representative picture of within-population genetic variation across the genome by sequencing pools of ~16 individuals from each population. Second, in agreement with previous studies reporting that global effects of
the demographic evolutionary forces in A. thaliana should be limited at a small geographical scale (Nordborg et al., 2005; Platt et al., 2010), we observed that genomic variation among the 168 natural populations is weakly correlated to geographic variation. Such a pattern likely alleviated the limitations of GEA analyses often observed at larger geographical scales such as confounding background produced by population structure, rare alleles and genetic/allelic heterogeneity.

Importantly, following methodologies previously developed in A. thaliana to identify environment-adaptive genetic loci at the European scale (i.e., climate and herbivore resistance; Hancock et al., 2011; Lasky et al., 2012, 2018; Brachi et al., 2015), we found that the SNPs the most associated with climate were significantly enriched in likely functional variants (i.e., non-synonymous variants) and/or in the extreme tail of spatial differentiation among populations. These clear signatures of selection suggest that climate is an important driver of adaptive genomic variation in A. thaliana at a micro-geographic scale. Although studies reporting the identification of climate adaptive genetic loci at a small geographical scales are still scarce (Manel et al., 2010; Kubota et al., 2015; Günther et al., 2016; Pluess et al., 2016), there is mounting genomic evidence that micro-geographic adaptation to climate is more widespread than is commonly assumed.

### Overrepresentation of Genes Involved in Regulatory Mechanisms in the Plant Functions Involved in Local Adaptation to Climate

The identification of candidate genes associated with climate variation suggests that regulation of gene expression is a key factor of climate adaptation. Transcription factors (TFs), RNA binding proteins and epigenetic mechanisms were the major functions uncovered by our study. These findings are in agreement with global gene expression studies demonstrating that plant response to environmental constraints relies on a number of distinct transcriptional responses operating spatially, temporally and in combination with other signals like hormones (Coolen et al., 2016). The $RZ-1C$ protein identified here in association with the mean annual temperature and the mean coldest month temperature, was shown to belong to a unique group of GRPs (Glycine-rich RNA binding Proteins) with...
FIGURE 5 | Comparison of the performances of BayPass for the identification of SNPs associated with population-specific covariable with varying correlation coefficients with the first axis of genetic variation PC_genomic. For each of the 10 covariable considered with correlation coefficients ranging from Spearman’s rho = 0 to Spearman’s rho = 0.9 (with an increasing step of 0.1) with the first PC_genomic, the plotted average ROC (Upper) and PR (Lower) curves result from the analyses of the 25 simulated data sets consisting of 10,000 SNPs including 100 associated SNPs (β = 0.1). The simulation model was calibrated to obtain data sets closely mimicking our observed data both in terms of the sample design (number of populations, haploid sample size of the pools, read coverage, across population allele frequencies) and population structure (summarized by the matrix Ω). The corresponding area under the curve (AUC) are reported in the box legend of each plot (r, Spearman’s rho).

Potential roles in increasing cold tolerance in Arabidopsis (Kim et al., 2010). These proteins have been recently shown to control gene splicing via interaction with SR (Serine/arginine-Rich) proteins, and consequently expression of many genes, including developmental regulators. Among them, the MADS-box transcription factor FLOWERING LOCUS C (FLC) is a direct target of RZ-1C that both promotes its splicing and represses its transcription (Wu et al., 2016). In the same vein, the MADS box transcription factor gene FLOWERING LOCUS M (FLM/MAF1) is involved in temperature dependent regulation of flowering through FLM splicing changes and nonsense-mediated mRNA decay in response to elevated temperature (Posé et al., 2013; Sureshkumar et al., 2016). miR172b, also found here in relation with the mean annual temperature and the mean coldest
FIGURE 6 | Enrichment analysis of different classes of sites in the 0.5% upper tail of the BF distribution of six climate variables (A: mean annual temperature, B: mean coldest month temperature, C: spring precipitation, D: summer precipitation, E: autumn precipitation, and F: winter precipitation). Enrichments shown are relative to the proportion of each class of SNPs in the genome overall. Large dots represent the enrichment values for each class of sites. Gray dots represent 10,000 null permutations of site categories. The horizontal dashed line shows the expected enrichment under the null hypothesis of no enrichment. Enrichments that are significant relative to permutations are denoted by asterisks. ns: non-significant, *0.05 > P > 0.01, ** 0.01 > P > 0.001, *** P < 0.001. Red asterisks indicate significant enrichments after a false discovery rate (FDR) correction at the nominal level of 5%. intra, intragenic SNPs (i.e., SNPs in transposable element gene); inter, intergenic SNPs; UTR, SNPs in 5’ and 3’ untranslated transcribed regions; intron, intronic SNPs; synon, synonymous SNPs; replac, replacement SNPs (amino-acid changing SNPs and stop codon gained SNPs).

month temperature, is a particularly interesting candidate as it controls transition of germinating seedlings from heterotrophic to autotrophic growth, and consequently the post-germination developmental arrest checkpoint under diverse stress (Zou et al., 2013). Found in relation to spring precipitations, ACT domain repeat 8 (ACR8) belongs to the ACR gene family which is differentially regulated by diverse abiotic stress (salt, cold, light/dark) and plant hormones. Interestingly, ACR proteins bind to specific ligands (amino acids or small ligands) and are thought to exert their regulatory functions through this binding and/or by sensing environmental conditions (Hsieh and Goodman, 2002). These examples illustrate the importance of diverse regulatory processes (metabolic and hormonal signals, RNA splicing, gene expression, protein-protein interactions, ...) in adaptation to climate variables. Interestingly, we also identified some transposable elements (known to be related with gene regulation) as candidate genes. This is in agreement with a recent work demonstrating that the composition and the activity of the Arabidopsis mobilome vary greatly among accessions (Quadrana et al., 2016) and that loci controlling adaptive responses to the environment are the most frequent transposition targets observed.

CONCLUSION

In agreement with the increasing number of phenotypic studies reporting micro-geographic adaptation (Richardson et al., 2014), climate variation appears as an important driver of adaptive genomic variation in A. thaliana at a fine spatial grain. This result reinforces the need to choose mapping populations according to the spatial scale of ecological variation at which species are adapted (Bergelson and Roux, 2010). In addition, the identification of climate-adaptive genetic loci at a micro-geographic scale highlights the importance to include within-species genetic diversity in ecological niche models for projecting potential species distributional shifts over short geographic distances (Valladares et al., 2014). The over-representation of genes involved in regulatory mechanisms in the plant functions associated with climate variation is a common pattern at different geographical scales in A. thaliana. The candidate genes identified in this study undoubtedly constitute key candidate genes for functional analysis, thereby providing an exciting opportunity to dissect the molecular bases of climate adaptation at a short geographic scale. The most significant associated SNP identified here for the mean cold month temperature in the
**REFERENCES**

Abebe, T. D., Naz, A. A., and Léon, J. (2015). Landscape genomics reveal signatures of local adaptation in barley (Hordeum vulgare L.). *Front. Plant Sci.* 6:813. doi: 10.3389/fpls.2015.00813

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., et al. (2016). The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 44, W3–W10. doi: 10.1093/nar/gkw343

Ägren, J., and Schemske, D. W. (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. *New Phytol.* 194, 1112–1122. doi: 10.1111/j.1469-8137.2012.04112.x

Ägren, J., Oakley, C. G., McKay, J. K., Lovell, J. T., and Schemske, D. W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* 110, 21077–21082. doi: 10.1073/pnas.1311773110

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., and Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol. Appl.* 1, 95–111. doi: 10.1111/j.1752-4571.2007.00013.x

Baroti, C., Frachon, L., Barret, M., Rigal, M., Huard-Chauveau, C., Mayjonode, B., et al. (2018). In situ relationships between microbiota and potential pathobiota in Arabidopsis thaliana. *ISME J.* doi: 10.1038/s41396-018-0152-7 [Epub ahead of print].

Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., et al. (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *Am. Nat.* 189, 463–473. doi: 10.1086/691233

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.

Bergelson, J., and Roux, F. (2010). Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. *Nat. Rev. Genet.* 11, 867–879. doi: 10.1038/nrg2896

Bomblies, K., Yant, L., Latinen, R. A., Kim, S. T., Hollister, J. D., Warthmann, N., et al. (2010). Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of Arabidopsis thaliana. *PloS Genet.* 6:e1000890. doi: 10.1371/journal.pgen.1000890

Borcard, D., and Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Modell.* 153, 51–68. doi: 10.1016/S0304-3800(01)00501-4

Borcard, D., Legendre, P., Avois-Jacquet, C., and Tuomisto, H. (2004). Dissecting the spatial structure of ecological data. *Ecology* 85, 1826–1832. doi: 10.1890/03-3111

Brachi, B., Villoutreix, R., Faure, N., Hautekéte, N., Piquot, Y., Pauwels, M., et al. (2013). Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in Arabidopsis thaliana. *Mol. Ecol.* 22, 4222–4240. doi: 10.1111/mec.12396

**AUTHOR CONTRIBUTIONS**

LF, CB, and FR planned and designed the research. LF and FR conducted fieldwork. LF performed the climate database searches. CB and FR performed the DNA extraction. CB, OB, and AC generated the sequencing data. CB and SC performed the bioinformatics analysis. LF and MG performed the genome–environment association analysis. LF, DR, and FR performed the enrichment analyses and the identification of candidate genes. LF, CB, MG, DR, and FR wrote the manuscript. All authors contributed to the revisions.

**FUNDING**

This work was funded by the Région Midi-Pyrénées (CLIMARES project) and the LABEX TULIP (ANR-10-LABX-41 and ANR-11-IDEX-0002-02).

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.00967/full?supplementary-material
Fusco, G., and Minelli, A. (2010). Phenotypic plasticity in development and evolution: facts and concepts. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365, 547–556. doi: 10.1098/rstb.2009.0267

Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. Genetics 201, 1555–1579. doi: 10.1534/genetics.115.181453

Grau, J., Grosse, I., and Keilwagen, J. (2015). PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics 31, 2595–2597. doi: 10.1093/bioinformatics/btv153

Günther, T., and Koop, G. (2013). Robust identification of local adaptation from allele frequencies. Genetics 195, 205–220. doi: 10.1534/genetics.113.152462

Karasov, T. L., Kniskern, J. M., Gao, L., DeYoung, B. J., Ding, J., Dubiella, U., Günther, T., and Coop, G. (2013). Robust identification of local adaptation from genome. J. Biogeogr. 40, 1311–1329. doi: 10.1111/j.1365-294X.2012.05709.x

Kubota, S., Iwasaki, T., Hanada, K., Nagano, A. J., Fujiyama, A., Toyoda, A., et al. (2011). Adaptation to climate across the North Italian Alps. Mol. Ecol. 20, 3574–3592. doi: 10.1111/j.1365-294X.2010.02722.x

Lasky, J. R., Forster, B. R., and Reimherr, M. (2018). Coherent synthesis of genomic associations with phenotypes and home environments. Mol. Ecol. Resour. 18, 91–106. doi: 10.1111/1755-0998.12714

Le Corre, V. (2005). Variation at two flowering time genes within and among populations of Arabidopsis thaliana: comparison with markers and traits. Mol. Ecol. 14, 4181–4192. doi: 10.1111/j.1365-294X.2005.02722.x

Li, Y., Handsaker, B., Wyssoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Luo, Y., Widmer, A., and Karrenberg, S. (2015). The roles of genetic drift and natural selection in quantitative trait divergence along an altitudinal gradient in Arabidopsis thaliana. Heredity 114, 220–228. doi: 10.1038/hdy.2014.89

Manel, S., Poncet, B. N., Legendre, P., Gugerli, F., and Holderegger, R. (2010). Common factors drive adaptive genetic variation at different spatial scales in Arabis alpina. Mol. Ecol. 19, 3824–3835. doi: 10.1111/j.1365-294X.2010.04716.x

Montesinos, A., Persier, C., Pratlong, M., Abi-Rached, I., Paganini, J., Pontarotti, P., et al. (2016). Genomic resources and their influence on the detection of the signal of positive selection in genome scans. Mol. Ecol. 25, 170–184. doi: 10.1111/mec.13468

Montesinos-Navarro, A., Wig, J., Picó, F. X., and Tenor, S. J. (2017). Arabidopsis thaliana populations show clinal variation in a climatic gradient associated with altitude. New Phytol. 189, 282–294. doi: 10.1111/nph.14379.x

Montesinos, A., Torrens, S. J., Alonso-Blanco, C., and Picó, F. X. (2009). Demographic and genetic patterns of variation among populations of Arabidopsis thaliana from contrasting native environments. PLoS One 4:e27213. doi: 10.1371/journal.pone.0007213

Nordborg, M., Hu, T. T., Ishino, Y., Jhaeveri, J., Toomajian, C., Zheng, H., et al. (2005). The pattern of polymorphism in Arabidopsis thaliana. PLoS Biol. 3:e196. doi: 10.1371/journal.pbio.0030196

Oksanen, J., Blanchet, F. G.,Kindt, R.,Legendre, P.,Minchin, P. R., O’Hara, R. B., et al. (2016). vegan: Community Ecology Package. R Package Version 2.3-5.

Oworoski, B., and Seneviratne, S. I. (2012). Global changes in extreme events: regional and seasonal dimension. Clim. Change 110, 669–696. doi: 10.1007/s10584-011-0222-9

Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., et al. (2017). Biodiversity redistribution under climate change: impacts on ecosystems and human well-being. Science 355,eaav9214. doi: 10.1126/science.aa9214

Picó, F. X. (2012). Demographic fate of Arabidopsis thaliana cohorts of autumn- and spring-germinated plants along an altitudinal gradient. J. Ecol. 100, 1009–1018. doi: 10.1111/j.1365-2745.2012.01979.x

Picó, F. X., Mendez-Vigo, B., Martinez-Zapater, J. M., and Alonso-Blanco, C. (2008). Natural genetic variation of Arabidopsis thaliana is geographically structured in the Iberian Peninsula. Genetics 180, 1009–1021. doi: 10.1530/gene.107.089581

Platt, A., Horton, M., Huang, Y. S., Li, Y., Anastasio, A. E., Mulyati, N. W., et al. (2010). The scale of population structure in Arabidopsis thaliana. Mol. Ecol. 19, 3824–3835. doi: 10.1111/j.1365-294X.2010.04716.x

Pluess, A. R., Frank, A., Heiri, C., Lalagué, H., Vendramin, G. G. and Oddou-Muratorio, S. (2016). Genome–environment association study suggests local adaptation to climate at the regional scale in Fagus sylvatica. New Phytol. 210, 589–601. doi: 10.1111/nph.13809

Posé, D., Verhage, L., Ott, F., Yant, L., Mathieu, J., Angenent, G. C., et al. (2013). Temperature-dependent regulation of flowering by antagonistic FLM variants. Nature 503, 414–417. doi: 10.1038/nature12633

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. isbn=978-3-900051-07-0.

Roldán, J., Roldán-Luján, N., and Domínguez, M. (2016). Selective Trait Evolution and the Analysis of Plant Domestication. In: Plant Domestication Research: From the Laboratory to the Field. (eds S. Díaz-Caycedo, P. Aguirre and P. Zuloaga). Springer. isbn=978-3-319-30445-0.
Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddoh, J. A., et al. (2016). The Arabidopsis thaliana mobilome and its impact at the species level. eLife 5:e15716. doi: 10.7554/eLife.15716

Ramette, A., and Tiedje, M. (2007). Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. Proc. Natl. Acad. Sci. U.S.A. 104, 2761–2766. doi: 10.1073/pnas.0610671104

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. Mol. Ecol. 24, 4348–4370. doi: 10.1111/mec.13322

Richardson, J. L., Urban, M. C., Bolnick, D. I., and Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. Trends Ecol. Evol. 29, 165–176. doi: 10.1016/j.tree.2014.01.002

Samis, K. E., Murren, C. J., Bossdorf, O., Donohue, K., Fenster, C. B., Malmberg, R. L., et al. (2012). Longitudinal trends in climate drive flowering time clines in North American Arabidopsis thaliana. Ecol. Evol. 2, 1162–1180. doi: 10.1002/ece3.262

Schlötterer, C., Tohler, R., Kofler, R., and Nolte, V. (2014). Sequencing pools of individuals - mining genome-wide polymorphism data without big funding. Nat. Rev. Genet. 15, 749–763. doi: 10.1038/nrg3803

Sureshkumar, S., Dent, C., Seleznev, A., Tasset, C., and Balasubramanian, S. (2016). Nonsense-mediated mRNA decay modulates FLM-dependent thermosensory flowering response in Arabidopsis. Nat. Plants 2, 16055. doi: 10.1038/nplants.2016.55

Tabas-Madrid, D., Méndez-Vigo, B., Arteaga, N., Marcet, A., Pascual-Montana, A., Weigel, D., et al. (2018). Genome-wide signatures of flowering adaptation to climate temperature: regional analyses in a highly diverse native range of Arabidopsis thaliana. Plant Cell Environ. doi: 10.1111/pce.13189

Valladares, F., Matesanz, S., Guilhaumon, F., Araujo, M. B., Balaguer, L., Benito-Garzon, M., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. Ecol. Lett. 17, 1351–1364. doi: 10.1111/ele.12348

van Kleunen, M., and Fischer, M. (2005). Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytol. 166, 49–60. doi: 10.1111/j.1469-8137.2004.01296.x

Vidigal, D. S., Marques, A. C., Willems, L. A. J., Buijs, G., Méndez-Vigo, B., Hilhorst, H. W. M., et al. (2016). Altitudinal and climatic associations of seed dormancy and flowering traits evidence adaptation of annual life cycle timing in Arabidopsis thaliana. Plant Cell Environ. 39, 1737–1748. doi: 10.1111/pce.12734

Wang, Z., Wang, L., Liu, Z., Li, Y., Liu, Q., and Liu, B. (2016). Phylogeny, seed trait, and ecological correlates of seed germination at the community level in a degraded sandy grassland. Front. Plant Sci. 7:1532. doi: 10.3389/fpls.2016.01532

Wender, N. J., Polisetty, C. R., and Donohue, K. (2005). Density-dependent processes influencing the evolutionary dynamics of dispersal: a functional analysis of seed dispersal in Arabidopsis thaliana (Brassicaceae). Am. J. Bot. 95, 960–971. doi: 10.3732/ajb.92.6.960

Wilczek, A. M., Cooper, M. D., Korves, T. M., and Schmitt, J. (2014). Lagging adaptation to warming climate in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U.S.A. 111, 7906–7913. doi: 10.1073/pnas.1406314111

Wu, Z., Zha, D., Lin, X., Miao, J., Gu, L., Deng, X., et al. (2016). RNA binding proteins RZ-1B and RZ-1C play critical roles in regulating pre-mRNA splicing and gene expression during development in Arabidopsis. Plant Cell 28, 55–73. doi: 10.1105/tpc.15.00949

Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D., and Tiffin, P. (2014). Genomic signature of adaptation to climate in Medicago truncatula. Genetics 196, 1263–1275. doi: 10.1534/genetics.113.159319

Zou, Y., Wang, Y., Wang, L., Yang, L., Wang, R., and Li, X. (2013). miR172b controls the transition to autotrophic development inhibited by ABA in Arabidopsis. PLoS One 8:e64770. doi: 10.1371/journal.pone.0064770

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Frachon, Bartoli, Carrère, Bouchez, Chaubet, Gautier, Roby and Roux. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.