Genomic basis of adaptation to a novel precipitation regime

Ahmed F. Elfarargi,[1] Elodie Gilbault,[2] Nina Döring,[1] Célia Neto,[1] Andrea Fulgione,[1] Andreas P. M. Weber,[3] Olivier Loudet,[2] Angela M. Hancock,[1,*]

[1]Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany.

[2]Université Paris-Saclay, INRAE, AgroParisTech, Institut Jean-Pierre Bourgin (IJPB), 78000 Versailles, France

[3]Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS), Heinrich Heine University, 40225 Düsseldorf, Germany

[*]Correspondence to: hancock@mpipz.mpg.de

Abstract

Energy production and metabolism are intimately linked to ecological and environmental constraints across the tree of life. In plants, which depend on sunlight to produce energy, the link between primary metabolism and the environment is especially strong. By governing CO_2 uptake for photosynthesis and transpiration, leaf pores, or stomata, couple energy metabolism to the environment and determine productivity and water use efficiency. Although evolution is known to tune physiological traits to the local environment, we lack knowledge of the specific links between molecular and evolutionary mechanisms that shape this process in nature. Here, we investigate the evolution of stomatal conductance and water use efficiency (WUE) in an *Arabidopsis* population that colonized an island with a montane cloud scrubland ecosystem characterized by seasonal drought and fog-based precipitation. We find that stomatal conductance increases and WUE decreases in the colonizing population relative to its closest outgroup population from temperate North Africa. Genome-wide association mapping reveals a polygenic basis of trait variation, with a substantial contribution from a nonsynonymous SNP in *MAP KINASE 12* (*MPK12* G53R), which explains 35% of the phenotypic variance in WUE in the island population. We reconstruct the spatially-explicit evolutionary history of *MPK12* 53R in the island and find that this allele increased in frequency in the population due to positive selection as *Arabidopsis* expanded into the harsher regions of the island. Overall, these findings show how adaptation shaped quantitative eco-physiological traits in a new precipitation regime defined by low rainfall and high humidity.

Keywords: *Arabidopsis thaliana*, local adaptation, seasonal drought, stomatal conductance, water use efficiency (WUE), *Mitogen-activated protein kinase 12* (*MPK12*).

Introduction

Matching physiological traits to the environment is crucial for survival and reproductive success across diverse life forms. Under directional selection, distributions of traits in a population are expected to shift towards their new optima due to differential fitness over evolutionary time (Falconer and Mackay 1983; Lynch and Walsh 1998), resulting in the matching of a population's physiology to its environment (Stearns 1989; de Jong 1993; Zera and Harshman 2001; Roff 2002; Shefferson et al. 2003; Morrison and Stacy 2014; Villellas and García 2018). In animals, observations that metabolism, body size and dimensions often vary with temperature are the basis of classic eco-physiological 'rules' (Bergmann 1847; Allen 1877; Ray 1960; Dikmen et al. 2013; Lafuente et al. 2018; Zhou et al. 2018). In plants, photosynthesis is the major mode of energy acquisition, and the interface between the environment and constraints on photosynthesis is crucial. Here, form and function predict economy of energy acquisition (Cowan 1986; Donovan and Ehleringer 1994), which in turn has been linked to spatial variation in selection pressures through associated physiological traits (Donovan and Ehleringer 1994; Wright et al. 2004; Díaz et al. 2016; Bjorkman et al. 2018). Overall, global distributions of traits involve optimization in the face of tradeoffs (Willi and Van Buskirk 2022).

In annual plants, flowering later can provide more time for accumulation of resources, resulting in a potential fitness benefit (Korves et al. 2007). However, in ecosystems with seasonal drought, growing quickly to reproduce before the dry season may be favored (Cohen 1970; Ludlow 1989). But such rapid growth requires high levels of photosynthesis, which relies on gas exchange through stomata, the pores on the surface of leaves. For photosynthesis to occur, stomata must be open to allow gas exchange, reducing water use efficiency and making the plant vulnerable to drying (Geber and Dawson 1990; Geber and Dawson 1997). Leaf water loss through stomata is especially high in environments where the vapor pressure deficit, i.e., the amount of air moisture relative to moisture-saturated air, is high (Will et al. 2013). Tuning the regulation of leaf pores, or stomata, is crucial for regulating the physiological trade-off between increasing energy production via photosynthesis and water loss at the leaf surface (Moreno-Gutiérrez et al. 2012; Kenney et al. 2014; Querejeta et al. 2018). Therefore, in a given environment, optimal stomatal aperture in natural populations depends on the availability of moisture through rainfall as well as the vapor pressure deficit.

While precipitation is often considered to be synonymous with rainfall, in many regions, plants rely heavily on 'horizontal' precipitation in the form of clouds or fog. These include the Lomas of Peru, fog deserts of Namibia, coastal western North American redwood forests and scrublands, and the seasonal montane cloud forests and scrublands of tropical Africa, Australia, and South America (Kerfoot 1968; Walter 1985; Stadtmüller 1987; Dawson 1998; Weathers 1999; Bruijnzeel et al. 2011; Karger et al. 2021). Such ecosystem types support a high proportion of Earth's biodiversity, especially its endemic species (Bruijnzeel and Hamilton 2017). Understanding how plants adapt to these ecosystems is important for preserving biodiversity and identifying effective approaches to improve sustainable agriculture in these critical regions.

Arabidopsis thaliana is the major molecular model plant as well as an important eco-evolutionary model (Alonso-Blanco and Koornneef 2000; Koornneef et al. 2004; Mitchell-Olds and Schmitt 2006; Verslues and Juenger 2011; Weigel 2012; Assmann 2013). Eurasian populations of *A. thaliana* have been extensively studied and used to understand the genetic bases of adaptation to local environments (Fournier-Level et al. 2011; Hancock et al. 2011; Lasky et al. 2012; Exposito-Alonso et al. 2018; Ferrero-Serrano and Assmann 2019) and of variation in a wide range of traits related to development timing, metabolite and elemental content, pathogen response, growth and drought response (e.g.,(Atwell et al. 2010; Brachi et al. 2010; Chan et al. 2010; Li et al. 2010; Filiault and Maloof 2012; Davila Olivas et al. 2017; Kalladan et al. 2017; Zan and Carlborg 2019; Wieters et al. 2021; Bhaskara et al. 2022; Gloss et al. 2022; Roux and Frachon 2022)). However, *A. thaliana* populations from North Africa (Brennan et al. 2014; Durvasula et al. 2017; Tabas-Madrid et al. 2018) and the Macaronesian archipelagos, including Madeira (Fulgione et al. 2018), the Canary Islands (Kranz and Kirchheim 1987), and the Cape Verde Islands (Fulgione et al. 2022; Tergemina et al. 2022), are mostly unstudied at the phenotypic level.

Here, we examine the evolution of stomatal conductance and water use efficiency (WUE) in an *A. thaliana* population that colonized the Cape Verde Islands (CVI). Islands can provide powerful systems for evolutionary analysis because they represent simplified 'natural laboratories' where evolution can be studied in isolation (Losos and Ricklefs 2009). Such systems provided the basis for the theory of evolution by natural selection (Wallace 1855; Darwin 1859) and have been used for elucidating classic cases of adaptive processes (Losos et al.

1997; Grant 1999). *A. thaliana* colonized CVI from temperate North Africa 5-7 kya through an extreme bottleneck that wiped out nearly all standing genetic variation (Fulgione et al. 2022) **(fig. 1)**. The CVI climate is defined by a short growing season with limited and highly variable rainfall. *A. thaliana* in Cape Verde is restricted to high altitude (> 950 m) north-facing slopes, where vegetation is bathed in moisture derived from humid trade-winds (Brochmann et al. 1997; Fulgione et al. 2022). The short growing season combined with high humidity creates an environment that differs substantially from the Mediterranean climate of the Moroccan Atlas Mountains, which supports the closest outgroup populations (fig. 4 and supplementary fig. S2 in (Fulgione et al. 2022)).

In this study, we find a shift in the phenotype distribution toward higher stomatal conductance and lower WUE in Cape Verde relative to North Africa. Using genome-wide association mapping, we characterize the trait architecture and identify a nonsynonymous variant (G53R) in the *MPK12* gene that explains a large proportion of the trait variation. We then reconstruct the historical spread of this variant across the island and find evidence that the derived allele facilitated local adaptation in the new tropical precipitation regime defined by limited rainfall and moisture delivered primarily through high air humidity.

Results

Stomatal conductance is higher and water use efficiency lower in CVI compared to Morocco

In CVI, rainfall is limited and unpredictable (**supplementary fig. S1A-B**), and water vapor pressure (specific humidity) is consistently high relative to Moroccan *A. thaliana* sites (**supplementary fig. S1C**). The median relative humidity across CVI sites during the growing season is 86.9% with lower and upper bounds of 65.5-95.9% (**supplementary fig. S2**).

We hypothesized that local adaptation may have acted to optimize performance in CVI *Arabidopsis* populations in response to the shift to higher humidity here. To investigate this possibility, we examined variation in water use efficiency (WUE; measured as carbon isotope discrimination - $\delta^{13}C$) and stomatal conductance (gas-exchange capacity) in well-watered (WW) and moderate water deficit (WD) conditions in 152 lines from the Cape Verde Island of Santo Antão and 24 representative Moroccan outgroup lines (**supplementary table S1**). In large-scale

phenotyping experiments, it is challenging to consistently apply drought stress conditions across pots because of spatial heterogeneity in drying rates. To deal with this, we used the high throughput Phenoscope platform that automatically circulates pots and adjusts watering several times per day based on pot weight, allowing experiments that would not be practical by manual procedures (Tisné et al. 2013).

We examined the effects of drought treatment and geographic origin on stomatal conductance and WUE. The WD condition led to an average of 40% less rosette growth at the end of the experiment compared to WW, indicating that the WD condition reduced growth rate on average. Average stomatal conductance was higher in the Santo Antão (CVI) population than in the Moroccan population in both watering conditions (WW: LMM, region fixed-effect estimate = 88.16 mmol m^{-2} s^{-1}, $P < 0.001$; WD: LMM, treatment fixed-effect estimate = -53.6 mmol m^{-2} s^{-1}, $P = 0.011$) (**supplementary table S2, fig. 2A**), and WUE was reduced in the Santo Antão population relative to the Moroccan outgroup population in both conditions (WW: LMM, region fixed-effect estimate = -0.44 ‰, $P = 0.003$; WD: LMM, treatment fixed-effect estimate = 1.6 ‰, $P < 0.001$) (**fig. 2B**). As expected, WUE was strongly negatively correlated with stomatal conductance across Santo Antão lines (Pearson correlation coefficient $R^2 = 0.23$, $P = 8.3 \times 10^{-10}$ and $R^2 = 0.28$, $P = 5.5 \times 10^{-11}$, for the WW and WD treatments, respectively; **supplementary fig. S3**). Overall, trait distributions shifted such that in the seasonally humid Santo Antão population, mean stomatal conductance was higher and mean water use efficiency lower than in the Moroccan population. The shifts in the distributions were similar across populations resulting in parallel reaction norms with consistent genetic differences in both treatments, which imply a simple genetic response, with no evidence of a genotype by environment (GxE) interaction (**supplementary fig. S4**).

Water use efficiency and stomatal conductance are moderately polygenic

The high trait variation we observed within the Santo Antão population suggested that genetic variation responsible for these traits may segregate there. The proportion of trait variance attributable to genetic variation, or heritability, provides information about the potential for genetic mapping within a natural population. We estimated heritability based on the proportion of the phenotypic variance explained by all genotyped SNPs, which is commonly referred to as "chip heritability" (Zhou and Stephens 2012; Zhou 2014). The estimated heritability was

moderate for stomatal conductance (0.45, 95% CI 0.29 to 0.60 for average stomatal conductance across conditions, 0.40, 95% CI 0.22 to 0.59 in WW, and 0.29, 95% CI 0.12 to 0.46 in WD) and high for WUE (0.82, 95% CI 0.75 to 0.88 for the average WUE, 0.30, 95% CI 0.14 to 0.47 for the drought response (the difference between WD and WW conditions), 0.81, 95% CI 0.73 to 0.87 in WW and 0.73, 95% CI 0.63 to 0.81 in WD). This discrepancy may imply that WUE is impacted less by uncontrolled environmental variation than stomatal conductance or that the genetic basis of stomatal conductance variation is more complex and not captured as well by additive genetic variance models. Moreover, stomatal conductance is an instantaneous measure and WUE measured as carbon isotope ratio is an integrated measure over the lifetime of the leaf, and thus may be expected to have higher heritability.

Next, we investigated the genetic architecture of the traits using a Bayesian sparse linear mixed model that allows for a mixture of large and infinitesimal genetic effects (Zhou et al. 2013). We found that seven loci explained 82% (95% CI: 75%-88%) of the genetic variance for average WUE and 68 loci explained 30% (95% CI: 14-47%) of the genetic variance for drought response of WUE (**fig. 3A, supplementary table S3-4**). Furthermore, we found that seven loci were predicted to have effects on WUE in WW and WD conditions (**supplementary fig. S5A, supplementary table S3-4**). For the genetic architecture of the average stomatal conductance and the drought response, we found about 39 and 44 loci have a major-effect, respectively (**supplementary fig. S6A, supplementary table S3-4**). In addition, about 39 and 53 loci were predicted to have major effects in WW and WD conditions, respectively (**supplementary fig. S7A, supplementary table S3-4**). We also examined the strength of genetic correlation between WUE and stomatal conductance, which reflects the average effect of pleiotropic action across all causal loci in both traits and helps to describe their complex relationships (van Rheenen et al. 2019). We observed a negative genetic correlation (Pearson correlation coefficient $R^2 = 0.12$, $P < 2.2\text{x}10^{-16}$ for the average traits, $R^2 = 0.28$, $P < 2.2\text{x}10^{-16}$ for the drought response of traits, $R^2 = 0.16$, $P < 2.2\text{x}10^{-16}$ for WW, and $R^2 = 0.55$, $P < 2.2\text{x}10^{-16}$ for WD) between both traits across Santo Antão lines. Overall, we found that genetic architecture was moderately complex for WUE and stomatal conductance and that a significant fraction of the genetic basis for the traits is shared between traits based on their genetic correlations.

A nonsynonymous variant in *MPK12* (G53R) explains a large proportion of trait variance

To identify specific loci underlying variation in the average traits, drought response of traits, and both conditions, we used a linear mixed model (LMM) approach that controls for population structure by including a relatedness matrix in the model (Zhou and Stephens 2014). For the average WUE based on $\delta^{13}C$ measurements, we detected a single Bonferroni significant peak on the end of chromosome 2 (**fig. 3B**) as well as in WW and WD conditions (**supplementary fig. S5B**). This peak contains a nonsynonymous variant (G53R) in the *Arabidopsis Mitogen-Activated Protein Kinase12* (*MPK12*; AT2G46070) gene, which was previously implicated in WUE in Cvi-0 x L*er*-0 RIL and NIL mapping populations (Juenger et al. 2005; Des Marais et al. 2014). Further, we found that one of the highest peaks contained the *MPK12* region in the drought response of WUE (**fig. 3B**). This variant explained 35% of the variation in the average WUE, 10% for drought response, 33% for WW, and 29% for WD in the Santo Antão population.

We identified several potentially interesting associations in addition to *MPK12* across the genome in the drought response of WUE. One of the highest peaks on chromosome 5 (drought response of WUE; **fig. 3B**) contains *PBL27* (AT5G18610), which encodes a receptor-like cytoplasmic kinase that is required to phosphorylate the SLOW ANION CHANNEL-ASSOCIATED HOMOLOG 3 (SLAH3) for anti-fungal immunity and chitin-induced stomatal closure. It has been shown that this signal transduction is independent of ABA-induced SLAH3 activation (Liu et al. 2019). Another genomic region on chromosome 5 comprises a downstream gene variant mapped to the *CNX1* gene. CNX1 catalyzes the final step of the synthesis of molybdenum cofactor (MoCo), a cofactor for multiple plant enzymes: abscisic acid (ABA), auxin, and nitrate (Porch et al. 2006). Another peak on chromosome 1 contained an upstream gene variant in *WRKY57*, a gene for which increased expression was previously shown to improve drought tolerance in *Arabidopsis* through increased ABA (Jiang et al. 2012). A peak at the top of chromosome 4 contained the well-known *FRIGIDA* (*FRI*) K232X variant, a major determinant of flowering time in *Arabidopsis thaliana* (Johanson et al. 2000; Gazzani et al. 2003; Shindo et al. 2005; Fulgione et al. 2022). Lovell and colleagues (Lovell et al. 2013) showed that the derived *FRI* allele pleiotropically confers a drought escape strategy through decreased flowering time, decreased WUE, and increased growth rate. We also identified an association

peak corresponding with a stop loss variant in the *ER-type Ca2+-ATPase 2* (*ECA2*) gene, which catalyzes the efflux of calcium from the cytoplasm. The cuticle mutant *eca2* that has an altered phenotype in cutin and wax showed a plant defense response to different biotic stresses, including biotrophic and necrotrophic pathogens and herbivory insects (Blanc et al. 2018; Aragón et al. 2021). Since some of these associations could arise due to partial linkage disequilibrium with the major effect variant at *MPK12*, we calculated the proportion of variance explained with and without *MPK12* G53R as a covariate. The PVE was reduced for *PBL27* (9% to 8%), *CNX1* (8% to 2%), and *WRKY57* (4% to 2%), unchanged for *ECA2* (6%) and the PVE increased for *FRI* (4% to 6%) with *MPK12* G53R as a covariate. Overall, these results support a moderately polygenic architecture for the drought response of WUE.

For stomatal conductance, GWAS revealed no Bonferroni significant results; however, the highest peaks in the average stomatal conductance (**supplementary fig. S6B**) as well as for both conditions separately (**supplementary fig. S7B**) contained the *MPK12* region. Here, the proportion of the genetic variance explained by *MPK12* G53R was 10% in the average stomatal conductance, 7% in the WW condition, and 12% in the WD condition. Plants from the natural population carrying the derived *MPK12* 53R allele had lower WUE and higher stomatal conductance than those carrying the ancestral G53 allele (**supplementary fig. S8A-B**). Taken together, our results support a central role for the *MPK12* G53R variant in trait variation in the natural CVI population.

Reconstructing the evolutionary history of variation in water use efficiency

Population structure in Santo Antão

As a first step toward reconstructing the evolutionary history of water use efficiency variation in Santo Antão, we examined the overall population structure of *A. thaliana* on the island. We found that the Santo Antão population could be divided into five major sub-populations based on results from principal component analysis (PCA) and neighbor-joining tree using LD-pruned genome-wide SNP variation (**fig. 4A-B, supplementary fig. S9A**). The sub-populations include Lombo de Figueira, Cova de Paúl, Ribeira de Poio, Pico da Cruz, and Espongeiro, which are hereafter referred to as Figueira, Cova, Ribeira, Pico, and Espongeiro. *A. thaliana* plants in Santo Antão tend to be found on rock outcrops and to be restricted to Northeast-facing slopes, where

they are exposed to humid northeasterly trade winds. This produces an east-west cline such that precipitation is highest and the growing season longest on the north-eastern side of the island, at the sites Figueira, Cova and Pico, and the growing season is shorter in the more western Ribeira and Espongeiro sites (Brochmann et al. 1997).

In the PCA, the Cova and Figueira sub-populations split on the first principal component axis, consistent with the previous finding that they represent the most ancestral variation in Santo Antão (Fulgione et al. 2022). Although the Ribeira sub-population lies geographically near the Espongeiro sub-population, it splits from Espongeiro on the second PC (**fig. 4B**). The third PC further distinguishes lines from within Espongeiro. Conversely, the two geographically separated sub-populations, Pico and Espongeiro, appear to be closely related despite their large geographic distance, suggesting recent spread and ongoing migration. These results are consistent with sub-population split times inferred previously (supplementary fig. S8 in (Fulgione et al. 2022)), and with results from sub-population topologies we inferred across 50-SNP genomic windows with *Twisst* (Martin and Van Belleghem 2017). The results showed the most common topology across the genome grouped Espongeiro and Pico, followed by Ribeira and then Figueira ((Cova, Figueira, Ribeira, (Espongeiro, Pico))) (**supplementary fig. S9B)**. Overall, these results support a deep split between Cova and Figueira, and a more recent expansion into the disjunct Espongeiro and Pico, with continuing gene flow between these regions.

Evolutionary history of genetic variation in water use efficiency

We next investigated the evolutionary history of the WUE trait in the Santo Antão natural population. We estimated the ages of loci associated with average WUE and found that the derived *MPK12* 53R allele was one of the first to arise. We estimated the age of *MPK12* 53R to be between 1.8 kya (time to the allele's most recent common ancestor; 95% CI: 0.87 – 2.4 kya) and 2.8 kya (based on allelic divergence; 95% CI: 2.2 – 3.1 kya) **(fig. 5A)**. Overall, our results are consistent with a model where the strong effect *MPK12* 53R variant arose early relative to other variants that impact WUE.

The *MPK12* G53R variant segregates at intermediate frequency (43%) in Santo Antão and exhibits structure across sub-populations **(fig. 5A-B, supplementary fig. S10)**. *MPK12* 53R is absent in the Figueira and Cova sub-populations, which represent the initial extent of the *A. thaliana* distribution in Santo Antão before expansion into the drier Espongeiro region at

approximately 3 kya (supplementary fig. S8 in (Fulgione et al. 2022)). The complete absence of *MPK12* 53R in the early-splitting Cova and Figueira, together with the age estimate for *MPK12* 53R, suggests that the allele likely arose after the split from these sub-populations. Among the more recently expanded sub-populations, the *MPK12* 53R allele varies in frequency across sites in an east-to-west gradient. The frequency of the derived allele is highest in the western-most sub-populations (Ribeira (90%) and Espongeiro (53%)) and lower in the moister eastern Pico region (29%).

To better understand the origin and historical spread of the *MPK12* 53R variant across the island, we examined the genealogical relationships between populations and individuals for the genomic region linked to this variant. The maximum likelihood topology for the 50-SNP window centered on the *MPK12* locus matched the major genome-wide topology (Cova, Figueira, Ribeira, (Espongeiro, Pico)) (**supplementary fig. S9B** and **supplementary fig. S11**). To examine the relationships at the scale of individual lines, we produced a marginal genealogical tree for the region using RELATE v1.1.4 (Speidel et al. 2019) (**fig. 5C**). The deepest branches of the derived *MPK12* 53R haplotype are found in the Ribeira and Espongeiro sub-populations, suggesting this allele first arose and rose to high frequency there. Clustering of individuals within the genealogical tree and the frequency distribution across the island suggest that *MPK12* 53R spread through multiple migrants into the Pico sub-population in the past few hundred years (**fig. 6D**). However, the *MPK12* 53R allele frequency has remained low in the moister Pico region. The allele frequency difference across populations suggested that *MPK12* 53R may be favored in the warmer, more exposed Ribeira/Espongeiro region where rapid growth to escape from drought would be most important.

Evidence for adaptive evolution at *MPK12* G53R

We next asked whether there was evidence the *MPK12* 53R allele was adaptive in Santo Antão. When an allele is driven quickly to high frequency in a population due to a partial selective sweep, a signature of an extended haplotype with reduced linked variation is expected (Hudson et al. 1994). Consistent with this, we identified an extended region of high haplotype homozygosity (EHH) (Sabeti et al. 2002) for the core derived *MPK12* 53R allele relative to the ancestral *MPK12* G53 allele (**fig. 6A-B**). To determine whether this locus is an outlier for haplotype homozygosity relative to the genome as a whole, we calculated the integrated

haplotype score (*iHS*) (Voight et al. 2006) across the genomes of the Santo Antão population. We found that *iHS* for the *MPK12* locus is extreme compared to the genome-wide distribution of haplotype homozygosity (|*iHS*|= 2.85, -$\log_{10}$(*p*-value) = 2.35) (**fig. 6C**).

We next used gene ontology enrichment to assess evidence of selection on traits based on the *iHS* results. Since the overall genetic variation in Santo Antão is low, the number of genes with *iHS* signals is also limited. Therefore, we did not expect to have high power in a gene ontology (GO) enrichment analysis. Still, we found a marginally significant enrichment for several biological processes. GO analysis revealed enrichment in genes regulating stomatal closure, abscission, osmotic stress, salicylic acid-mediated signaling, transcription elongation from RNA polymerase II promoter, regulation of DNA-templated transcription elongation, transition to flowering, auxin-activated signaling pathway, cellular response to an organic substance, and signal transduction (**supplementary fig. S12A-B, supplementary table S5**). Enrichments in stomatal closure, osmotic stress, salicylic acid signaling, response to organic substances, and signal transduction were largely driven by the same set of three genes. These included *MPK12*, the defense-related transcription factor *WRKY54*, and *AUXIN RESPONSE FACTOR 2* (*ARF2*). Enrichment of the transition to flowering category was also driven by three genes: *CLAVATA 2* (*CLV2*), *EMBRYONIC FLOWER 1* (*EMF1*), and *ARF2*. Overall, these results suggest that both flowering time and water balance may have been important selection pressures in the CVI population.

Further, a test of cross-population extended haplotype homozygosity (XP-EHH) (Sabeti et al. 2007) showed the derived haplotype in Espongeiro is highly differentiated with elevated haplotype homozygosity compared to the ancestral Cova sub-population (|XP-EHH|= 2.16, -$\log_{10}$(*p*-value)= 1.52) (**supplementary fig. S13A**) as well as between Espongeiro and Pico (|XP-EHH|= 2.03, -$\log_{10}$(*p*-value)= 1.37) (**supplementary fig. S13B**). We estimated a selection coefficient of 4% for the variant based on the inferred allele frequency trajectory from the local inferred genealogy (**fig. 5C, supplementary fig. S14, supplementary table S6**). To control for population growth during this timeframe, we estimated the selection coefficient for *MPK12* 53R against the previously inferred trajectory of historical population size (N_e) using whole-genome trees (Fulgione et al. 2022). Overall, these findings are consistent with positive selection acting on the derived *MPK12* allele in populations that expanded into the harsher western Ribeira/Espongeiro region of the island.

Environmental correlation analysis can provide further evidence for local adaptation based on statistical associations between climate variables and genetic variants (Hancock et al. 2011; Lasky et al. 2012). To determine whether local adaptation to climate might have shaped the frequency of *MPK12* 53R allelic variation across populations, we conducted a partial redundancy analysis (RDA). RDA links genomic variation to environmental predictors while accounting for geographic population structure by including geographic distance as a model covariate. We found a significant association between climate and genomic variation overall (P=0.001; R^2= 0.34; adjusted R^2= 0.313) and applied a stepwise model-building algorithm (*ordistep*) to determine which bioclimatic variables (**supplementary table S7**) best explained the spatial distribution of the genetic data. Five environmental variables (BIO5: Max Temperature of Warmest Month, BIO11: Mean Temperature of Coldest Quarter, BIO13: Precipitation of Wettest Month, BIO17: Precipitation of Driest Quarter, BIO19: Precipitation of Coldest Quarter) explained a large proportion of the variance across populations (P= 0.001; R^2= 0.15; adjusted R^2= 0.127) (**supplementary fig. S15**; **supplementary table S8-S9**). Espongeiro and Ribeira sub-populations separated from Figueira, Cova, and Pico on the first RDA axis, which was associated with the temperature variables (BIO5 and BIO11) (**supplementary fig. S15A**), whereas variation that separated sub-populations Figueira, Cova, and Pico loaded on the second RDA and was associated with the precipitation variables (BIO13, BIO17, and BIO19) (**supplementary fig. S15A**). We next examined the loadings by SNP (**supplementary fig. S15B**; **supplementary table S10**) to determine whether the *MPK12* 53R SNP variant was correlated with the partial RDA loadings. We found that *MPK12* G53R was an outlier in the RDA1 SNP loadings and its distribution was most strongly predicted by BIO5, the maximum temperature of warmest month (**fig. 6D**; **supplementary fig. S15C**). This suggests that the *MPK12* 53R variant is adaptive in the warmest microclimates in Santo Antão, in Espongeiro and Ribeira, where the growing seasons are shortest and the need for increased photosynthesis and faster growth may be strongest.

Finally, we asked whether the population genetic evidence we found for positive selection translated to a reproductive advantage in an experimental setting. To determine whether *MPK12* G53R was associated with differential fitness in a simulated Santo Antão environment, we used fitness data (total number of seeds produced) from plants we propagated in a growth chamber set to simulate humidity, air and soil temperature, soil chemistry and precipitation, photoperiod, and

light availability of an Espongeiro site in Santo Antão (Fulgione et al. 2022). We observed that plants carrying the derived *MPK12* 53R variant produced more seeds than plants with the ancestral *MPK12* G53 variant (negative binomial generalized linear model (GLM), *MPK12* 53R allele fixed-effect estimate= 0.76, $P = 0.00332$, **supplementary fig. S16**; **supplementary table S11**). Since we previously found that flowering time was strongly associated with fitness in the CVI-simulated environment (Fulgione et al. 2022) and because flowering time and water use efficiency have been implicated in drought avoidance (Mooney et al. 1976; Geber and Dawson 1990; Donovan and Ehleringer 1992; Geber and Dawson 1997; McKay et al. 2003; Heschel and Riginos 2005; Sherrard and Maherali 2006), we also examined the effect of *MPK12* G53R while controlling for *FRI* K232X. In a GLM with a negative binomial transformation of seed number, the signal for *MPK12* G53R on fitness was reduced but still highly significant (GLM, *MPK12* 53R allele fixed-effect estimate= 0.7, $P = 0.00519$) (**supplementary table S12**), indicating the *MPK12* 53R variant increases fitness independently from *FRI* 232X under CVI (Espongeiro) conditions.

Discussion

We examined the evolution of stomatal conductance and water use efficiency (WUE) in an *A. thaliana* population that colonized a novel precipitation regime. We found that average stomatal conductance increased and water use efficiency decreased in the humid Cape Verde Island (CVI) population relative to the North African outgroup. We found that trait architecture was polygenic, with an important contribution from a nonsynonymous variant (G53R) in *Mitogen-Activated Protein Kinase 12* (*MPK12*), which explained 35% of the trait variance in WUE in the Santo Antão island population. We found evidence that the derived *MPK12* 53R variant is evolving under positive selection based on its association with temperature across the island (**fig. 6D; supplementary fig. S15B-C**) and on a haplotype-based signature of selection in the genomic region (**fig. 6A-C**). Finally, we found that the derived *MPK12* 53R variant conferred higher fitness than the ancestral *MPK12* G53 variant in plants grown in CVI conditions (**supplementary fig. S16**). Overall, our findings reveal evidence that the *MPK12* 53R variant helped facilitate local adaptation in the island of Santo Antão, where 'horizontal' precipitation, or fog, is an important contributor to total precipitation.

Our findings are also relevant in the context of understanding how plants adapt to seasonal drought and the importance of physiological tradeoffs more generally. Plants use different strategies to maintain water balance (Klein 2014; Martínez-Vilalta et al. 2014; Skelton et al. 2015). Most plants are isohydric; they avoid reaching low water potential by closing their stomata during drought. However, in environments where humidity is reliably high and vapor pressure deficit is low, plants may be anisohydric, keeping their stomata open even when rainfall is limited. Rainfall in CVI is unpredictable, but trade winds provide a steady supply of high humidity to plants growing along the northeast-facing slopes during the short growing season (**supplementary figs. S1** and **S2**). In the humid regions of the island of Santo Antão in Cape Verde, where *A. thaliana* is found, the anisohydric strategy may be common. Our results indicate that *A. thaliana* populations here have evolved an anisohydric strategy in response to the humid environment.

This anisohydric strategy may provide other benefits. In drought-prone environments, plant populations may adapt by escaping drought (Levitt 1972; Ludlow 1989). When growing seasons are short, plant populations may maximize fitness by increasing stomatal conductance to increase rates of carbon gain through photosynthetic carbon assimilation and thus escape drought stress (Mooney et al. 1976; Geber and Dawson 1990; Donovan and Ehleringer 1992; Geber and Dawson 1997; McKay et al. 2003; Heschel and Riginos 2005; Sherrard and Maherali 2006). A drought-escape strategy appears to be strongly favored by selection in CVI (Fulgione et al. 2022), where growing seasons are short. More open stomata may enable higher levels of photosynthesis and faster growth, facilitating such a drought escape strategy. In this case, the high relative humidity may effectively reduce the tradeoff between photosynthesis and transpiration. Overall, our findings reveal a case where natural selection appears to have optimized carbon gain through increased stomatal aperture, facilitating drought escape in a natural population.

Although the genetic architecture of the traits studied here was moderately complex, we found that the *MPK12* 53R allele could explain a large proportion of the genetic variation in water use efficiency and stomatal conductance. Our finding that *MPK12 53R* underlies variation in stomatal conductance and water use efficiency in Cape Verde is consistent with previous evidence that this specific allele is important in water balance. Further, these previous findings

help to contextualize our results in the natural population. Prior work provides molecular evidence that MPK12 is important for sensing and responding to drought stress by regulating the stomatal guard cell response to abscisic acid (ABA), a key phytohormone involved in abiotic stress responses (Jammes et al. 2009; Montillet et al. 2013; Salam et al. 2013). Using QTL mapping and introgression, Juenger and colleagues identified the *MPK12* locus and subsequently validated the effect of the Cvi-0 *MPK12* allele on WUE (Juenger et al. 2005; Des Marais et al. 2014). Des Marais and colleagues (Des Marais et al. 2014) further showed that *MPK12* impacts guard cell size and behavior, and their work suggested that the CVI *MPK12* allele causes an altered response to vapor pressure deficit and abscisic acid-induced inhibition of stomatal opening. Additional analysis showed that the functional *MPK12* allele is involved in CO_2 signaling and that the CVI *MPK12* allele has an impact that is comparable to a complete loss of function (Jakobson et al. 2016). Finally, our finding that variation in *MPK12* impacts fitness in CVI conditions is interesting in the context of previous work (Campitelli et al. 2016) demonstrating that *MPK12* variation was associated with variation in fitness components in response to a combination of drought and competition. Our study focused on the CVI natural population supports these previous results and connects variation in the *MPK12* gene to ecology and evolution in the natural environment.

Our results provide the potential for crop improvement in sustainable agriculture. In regions of the world where horizontal precipitation is an important source of moisture, technological approaches have been developed to collect fog for agricultural use (Klemm et al. 2012; Schemenauer, Bignell, et al. 2016; Schemenauer, Zanetta, et al. 2016). However, these are difficult to maintain and their usefulness is thus limited. A more direct approach to exploit horizontal precipitation in agricultural improvement could potentially be achieved by breeding crops with increased stomatal aperture that can better use this available resource. Future work could apply the results of studies that identify such adaptive genetic variation in local wild populations to increase crop productivity in challenging conditions. Our results suggest that breeding crops with reduced activity of MPK12 or its homologues could increase crop productivity in tropical agricultural systems, where vertical precipitation is limited and horizontal precipitation is an important component of total precipitation.

There are several open questions that could be addressed in future research. We proposed that photosynthetic efficiency should be increased in plants with increased stomatal conductance (and

decreased WUE), in particular in those that carry the derived *MPK12* variant. This hypothesis could be explicitly tested in the future in a controlled study of photosynthetic efficiency. Further, while we have no specific evidence that *A. thaliana* from CVI is able to absorb water directly through the leaves, there is mounting evidence from diverse species that foliar water uptake through the leaf surface is a common strategy in humid environments where vertical precipitation is limited (Burkhardt et al. 2012; Berry et al. 2018; Binks et al. 2020). Further, there is evidence that plants in cloud forests may be especially susceptible to climate change (McDowell et al. 2008; McDowell et al. 2011). These hypotheses could be tested in future research in controlled lab-based experiments as well as in field experiments in CVI.

Although the traits studied here are polygenic, our findings revealed that one variant in *MPK12* explained a substantial fraction of the trait variation. Understanding how adaptation has occurred in specific cases can inform models and predictions of how populations might generally adapt to novel environments. Although we have information about functional loci and variants from QTL mapping studies in a range of species, it has only rarely been possible to connect results from QTL studies back to the ecology of the relevant natural population. This study serves as an example where it was possible to reconstruct the evolutionary history of a functional variant as it arose and spread across the landscape. Further, studies such as this one can inform models that aim to predict how species adapt as the environment changes or expand their ranges into more severe climates.

Materials and Methods

Study populations

In this work, we used the previously released whole-genome short-read data for 189 individuals collected from 26 different locations in Santo Antão (**fig. 1**; **supplementary table S1**) (ENA: PRJEB39079 (ERP122550)) (Fulgione et al. 2022) and 61 Moroccan lines (Durvasula et al. 2017) for genetic analysis. We used the SHORE pipeline (https://github.com/HancockLab/CVI) for SNP discovery and variant calling. The variant call format (VCF) file (EVA: PRJEB44201 (ERZ1886920)) (Fulgione et al. 2022) was filtered to minimize SNP calling bias and to retain only high-quality SNPs: (1) retain only bi-allelic SNPs; (2) convert heterozygous sites to missing data to mask possible false positives; (3) retain variants with coverage greater than 3 and base quality greater than 25. All maps were conducted using R v. 3.4.4 (R Development Core Team

2008) and the *ggmap* (Kahle and Wickham 2013) and *ggplot2* (Wickham 2016) libraries were used for plotting.

Phenoscope drought experiment and phenotyping

Trait measurement was performed using the high throughput phenotyping Phenoscope platform (https://phenoscope.versailles.inra.fr/) as previously described (Tisné et al. 2013). Santo Antão (Fulgione et al. 2022) (n=152) and Moroccan (n=24) *A. thaliana* lines (Brennan et al. 2014) (**supplementary table S1**) were grown under standard environmental conditions (8-h day/16-h night, 21°C day/17°C night, 65% relative humidity, and 230 µmol m^{-2} s^{-1} light intensity). For each trait, two independent replicate experiments were performed. In each experiment, two replicates per genotype and two watering conditions were used. The first was a 'well-watered' (WW) condition in which pots were provided with 60% of the maximum soil water content (SWC; 4,6g H_2O g^{-1} dry soil) not limiting for vegetative rosette growth. The second condition was a 'water-deficit condition' (WD) in which pots were provided with 25% SWC (1,4g H_2O g^{-1} dry soil). Plants were propagated on peat moss plugs, then selected for homogeneous germination and transferred onto the Phenoscope table eight days later, i.e., 8 days after sowing (DAS). On the Phenoscope, SWC reached 60% for control-treated plants at 12 DAS and 25% for moderate-drought-treated plants at 16 DAS. At 32 DAS, the whole rosette of two replicates for each genotype per treatment were collected, ground, and analyzed for carbon isotope discrimination ($\delta^{13}C$) as an estimate of water use efficiency (WUE). Isotope discrimination analysis was conducted at CEPLAS Plant Metabolism and Metabolomics Laboratory, Heinrich Heine University Düsseldorf (HHU) as described previously (Gowik et al. 2011). In short, dried plant material was ground to a fine powder and analyzed using an Isoprime 100 isotope ratio mass spectrometer coupled to an elemental analyzer (ISOTOPE cube; Elementar Analysensysteme, Hanau, Germany) following the manufacturer's recommendations. The carbon isotope ratio is expressed as ‰ against the Vienna Pee Dee Belemnite (VPDB) standard.

We measured leaf stomatal conductance using a leaf Porometer (SC-1, Decagon Devices, Pullman, WA, United States). According to the manual guide, the Porometer device was calibrated before measurements with a 100% humidity filter paper as a reference. It was challenging to measure the rosette leaves directly due to their reduced size and the small area of the SC-1 Porometer leaf clamp. Therefore, a fully developed leaf per line and per treatment of

each genotype was examined immediately after detachment. The measurements were performed across several days (29-32 DAS) around mid-day.

Phenotype data analysis

Differences in the phenotype distributions were evaluated using both parametric and non-parametric tests. For conducting Wilcoxon rank sum tests, we used *wilcox.test* in the *stat_compare_means* function ("*ggpubr*" package (Kassambara 2020)). We also used linear models to test fixed effects of treatment, geographic region, and their interaction on the measured phenotypic traits. For this, we used the R package *lme4* (Bates et al. 2014) to run the following model for each phenotype.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

Where Y_{ijk}: represents the phenotypic value; μ: the overall mean; α_i: the effect of the treatment; β_j: the effect of the geographic region; γ_{ij}: the interaction between treatment and region; ε_{ijk}: the residuals.

Correlations between phenotypes in both treatments are Pearson correlations calculated in R using the *cor.test* function. We evaluated the significance of correlations with the *t*-test implemented in the *cor.test* function.

We obtained the individual data for the total seed number (as a proxy of fitness) from (Fulgione et al. 2022). Since no block effect was detected in the simulated CVI conditions experiment, we used the median per genotype across replicates as phenotype. We tested the effects of the *MPK12* 53R derived variant on fitness using generalized linear models (GLM) (R function *glm*). To correct for over-dispersion of seed number, we used a negative binomial GLM using the "*glm.nb*()" function in the "*MASS*" v.7.3-54 package in R.

Quantitative genetic analyses

We estimated heritability for traits in this study based on the proportion of the phenotypic variance explained by all genotyped SNPs, which is commonly referred to as "chip heritability" (Zhou and Stephens 2012; Zhou 2014). To perform the association analysis, we first filtered out indels and non-biallelic SNPs from the VCF. We considered only SNPs with read coverage DP>=3 and quality GQ>=25. We then applied a 5% cutoff for the minor allele frequency (MAF).

Subsequently, we carried out the association analyses between genomic variants and stomatal conductance and WUE as traits using the univariate linear mixed model implemented in GEMMA (Zhou and Stephens 2012), separately for well-watered (WW) and water deficit (WD) conditions as well as the average for each trait across both conditions and the drought response (difference between conditions: WW-WD).

According to (Shim et al. 2015) and based on the GEMMA outputs, we calculated the proportion of variance in each trait explained by a given SNP (PVE) using the following equation:

$$PVE = \frac{2\hat{\beta}^2 MAF(1-MAF)}{2\hat{\beta}^2 MAF(1-MAF) + (se(\hat{\beta}))^2 2NMAF(1-MAF)}$$

where $\hat{\beta}$ is the effect size estimate, $se(\hat{\beta})$ is the standard error of effect size for the SNP, MAF is the minor allele frequency for the SNP, and N is the sample size.

To infer the genetic architecture of the traits, we used a polygenic GWA Bayesian Sparse Linear Mixed Model (BSLMM) implemented in GEMMA (Zhou and Stephens 2012), which models the polygenic architecture as a mixture of large and small effects. BSLMM accounts for the relatedness among individuals by including a genomic kinship matrix as a random effect in the model. Furthermore, the approach accounts for the LD between SNPs by inferring locus effect sizes (β) while controlling for other variants included in the model. Using this approach, we modelled two effect hyperparameters: a basal effect (β), which captures small-effect loci that contribute to the studied trait, and an additional effect (γ), which captures a subset of loci with the most potent effects. To estimate the effects of all SNPs, the sparse effect size for each locus was calculated by multiplying (β) by (γ). We listed the variants with the highest sparse effects on the studied trait.

We then investigated the genetic correlations between traits using the multivariate model in GEMMA (Zhou and Stephens 2012; Zhou 2014). Accordingly, we conducted the correlations between the effect sizes of all loci (β) for each trait through Pearson correlations calculated in R using the *cor.test* function. We evaluated the significance of correlations with the *t*-test implemented in the *cor.test* function.

Population structure analysis

In a pre-processing step before population structure analysis, we used PLINK v1.9 to prune our SNP sets for linkage disequilibrium by removing any variables with correlation coefficients (r^2) greater than 0.1 across windows of 50 Kb with a step size of 10 bp. Then, we removed variants with missing data by setting the parameter –geno to 0.

To conduct the principal component analysis (PCA), we used the --pca option in PLINK v1.9 (Purcell et al. 2007) . We produced the whole-genome neighbor-joining tree in R v3.3.4 (R Development Core Team 2008) using the packages "*APE*" v5.5 (Paradis and Schliep 2019), and "*adegenet*" v2.1.4 (Jombart 2008). To evaluate the relationships between the five Santo Antão sub-populations and visualize how the tree topology changes across the genome we used a phylogenetic weighting approach, *Twisst* (Martin and Van Belleghem 2017). This method uses maximum likelihood topology inference across genomic windows to produce a distribution of topology weightings (Martin and Van Belleghem 2017). Starting with our LD-pruned data set, we converted our data to '.geno' format using the script 'parseVCF.py' (https://github.com/simonhmartin/genomics_general/tree/master/VCF_processing), and we obtained the maximum likelihood trees in sliding windows of 50 SNPs using the script 'phyml_sliding_windows.py' (https://github.com/simonhmartin/genomics_general/tree/master/phylo). Then, we ran *Twisst* on the complete set of inferred trees for the five Santo Antão sub-populations (**Fi;** Lombo de Figueira, **Co;** Cova de Paúl, **Ri;** Ribeira de Poio, **Pi;** Pico da Cruz, and **Es;** Espongeiro) to calculate the exact weighting of each local window. We used the Cova de Paúl sub-population as an outgroup in this step. To plot the topologies, we used the R v3.3.4 (R Development Core Team 2008) and the "*APE*" package (Paradis and Schliep 2019).

Inferring the genealogical history of *MPK12* G53R

We used RELATE v1.1.4 (Speidel et al. 2019) to infer the genealogical trees for the derived *MPK12* 53R allele variant (Chr2:18947614). We used bcftools v1.9 (Li 2011) to filter the VCF file for quality, to remove non-biallelic SNPs, to remove fixed sites, and to filter out missing data with the command: <bcftools view -m2 -M2 -v snps –min-ac=1 -i 'MIN(FMT/DP)>3 & MIN(FMT/GQ)>25 & F_MISSING=0'>. Within RELATE, we used the command RelateFileFormats (using --mode ConvertFromVcf) to convert the VCF file into haplotype and

sample files. We ran RELATE under a haploid model for chromosome 2 (using –mode All) and we defined parameters as follows. For the mutation rate, we corrected the estimate for *A. thaliana* of $7x10^{-9}$ derived from (Ossowski et al. 2010) for the percent missing data in 1 Mb sliding windows every 50 kb across the entire genome ($2.245x10^{-9}$ for *MPK12* 53R variant). For the recombination map, we corrected a published map based on crosses (Salomé et al. 2012) for the outcrossing rate of 5% estimated in natural populations (Bomblies et al. 2010). For coalescence rates, we used the genome-wide rates inferred previously in (Fulgione et al. 2022) for the Santo Antão population. We set the generation time to one year. To produce genealogical trees for *MPK12* 53R variant with confidence intervals for the estimated ages based on 200 samples from the MCMC (derived using SampleBranchLengths.sh --format a, and using default settings), we used the script TreeViewSample.sh, with 10*N steps (N is the number of haplotypes) and 1000 burn-in iterations.

Climatic variables

We retrieved data for the 19 bioclimatic variables (**supplementary table S7**) commonly used to study the pattern of species distribution and the water vapor pressure (humidity) from the WorldClim global climate version 2 (Fick and Hijmans 2017) (https://worldclim.org/data/worldclim21.html), at a resolution of 30 seconds (~$1km^2$). We also obtained site-specific data for the accumulated rainfall amount during the growing season and aridity index from CHELSA (Karger et al. 2017) (**supplementary table S7**), which we used for a comparison between the climate of collection sites in Santo Antão and Moroccan sites (**supplementary fig. S1**). We extracted the climatic variable values for the specific geographical coordinates for each sampling location in Santo Antão and Morocco using the "*raster*" package in R (Hijmans et al. 2015). The shift of the climate variable distribution between Santo Antão and Morocco was tested using a two-tailed Wilcoxon rank sum test with the function *wilcox.test* implemented in the *stat_compare_means* function ("*ggpubr*" package (Kassambara 2020)).

Redundancy analysis (RDA): linking genomic variation to environment predictors

We used the redundancy analysis (RDA) approach implemented in the R package '*vegan*' v. 2.5-7 (Oksanen et al. 2020) to investigate the relative contributions of the bioclimatic variables and

spatial distribution of *MPK12* G53R across the Santo Antão landscape. RDA uses multiple regression to model matrices of explanatory variables (X and Y), in which X represents a set of environmental variables and Y represents a dependent matrix of genotypic data. It links genomic variation to environment predictors while accounting for geographic population structure by including geographic distance as a model covariate. Genotype data from a set of genome-wide LD-pruned SNPs (*n*= 8,475) and environmental data (**supplementary table S7**) were analyzed by running the full model. We used analyses of variance (ANOVA with 1,000 permutations) to assess the significance of each environmental variable within the RDA model. Then we used a stepwise permutational ordination method using the ordination step "*ordistep*" function in the R package '*vegan*' v. 2.5-7 (Oksanen et al. 2020) with 1,000 permutations to evaluate the environmental parameters and identify the model that best describes the spatial distribution of the genotype data. This function selects variables to build the 'optimal' model with the highest adjusted coefficient of determination (R_{adj}^2) and removes the non-significant variables one at a time using permutation tests.

Evolutionary history of *MPK12* 53R

- *Evidence of positive selection*

To detect signatures of positive selection in the Santo Antão population, we used three haplotype-based methods: the extended haplotype homozygosity (*EHH*) (Sabeti et al. 2002), the integrated haplotype score (*iHS*) (Voight et al. 2006), and cross-population EHH (*XP-EHH*) (Sabeti et al. 2007) implemented in the R package '*rehh*' version 2.0.2 (Gautier et al. 2017) in R. For *iHS* and *XP-EHH*, scores were transformed for each SNP into two-sided *p* values: p*iHS* = − log10[1–2|Φ(*iHS*)-0.5|] and *pXP-EHH* = − log10[1–2|Φ(*XP-EHH*)-0.5|] where Φ(x) represents the Gaussian cumulative distribution function. We used the default parameters for all analyses.

To determine whether there was enrichment of specific functional gene sets in the tail of the distribution of *iHS* scores, we conducted gene ontology (GO) enrichment analysis. For this, we used the top 1% SNP variants (> 99% quantile based on the genome-wide empirical distribution) identified through the genome-wide *iHS* scores across the genomes of the Santo Antão population. Gene names were extracted based on the SNP position using the TAIR10 GFF3 gene annotation file through SNPEff (Cingolani et al. 2012). GO analysis was conducted using the

ShinyGO web tool (http://bioinformatics.sdstate.edu/go/) (Ge et al. 2020) (see all results in **supplementary table S5** and **supplementary fig. S12A-B**). After running the analysis, we checked that significant results were not driven by signals in clusters of genes. We did not find that any of the genes responsible for enrichments were located on the same chromosomes.

- *Inference of the selection coefficient*

To infer a selection coefficient based on the reconstructed historical frequency trajectory for the derived *MPK12* allele (Chr2:18947614) we used CLUES (Stern et al. 2019). CLUES uses importance sampling over trees generated in RELATE to produce a posterior distribution from which a frequency trajectory can be inferred. We obtained estimates of the posterior distributions of allele frequencies over time using 200 samples from the MCMC and a recessive model. We inferred the selection coefficient jointly across two-time bins (epochs) of 1.5 kya between the present day and the time in the past when the variant arose (0-1.5 and 1.5-3 kya) (**supplementary table S6**).

Acknowledgements

We thank Juliette de Meaux and Maria von Korff as well as members of the Hancock Lab, who provided helpful discussion and feedback. We are grateful to Ângela Moreno and Samuel Gomes at INIDA as well as the Parque Natural de Santo Antão for helpful advice and interactions during the course of this research. Dominik Brilhaus and Maria Graf provided technical assistance with the $\delta^{13}C$ analysis. This work was supported by Max Planck Society Funding and European Research Council (ERC) CVI_ADAPT 638810 to A.M.H. The International Max Planck Research School (IMPRS) Program "Understanding Complex Plant Traits using Computational and Evolutionary Approaches" provided partial support for A.F.E. The project benefited from the support of IJPB's Plant Observatory technological platforms. The IJPB benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007). The CEPLAS Metabolomics and Metabolism Laboratory is supported by a grant of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2048/1 – project ID 390686111. This research benefited from discussions in the context of the Kavli Institute for Theoretical Physics workshop ADAPT22 and was thus supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

Authors' Contributions

A.F.E. and A.M.H. conceived and designed the project. O.L. provided expertise for the design of the drought measurement (Phenoscope) experiment. A.F.E., O.L., and E.G. performed the Phenoscope drought experiment and collected sample materials for $\delta^{13}C$ analysis. N.K., A.P.M.W and his lab members were responsible for $\delta^{13}C$ measurements. A.F.E. conducted data preparation, statistical analyses and created figures. A.F.E. and A.M.H. contributed to analyses and interpretation of results. A.F.E., A.M.H., C.N., and A.F. collected samples. A.F.E. and A.M.H. wrote the manuscript with input from all authors.

Data Availability

All scripts for analyses and data visualization have been archived in the Github repository [https://github.com/HancockLab/CVI_WUE_MPK12_LocalAdapt].

References

Allen JA. 1877. The influence of physical conditions in the genesis of species. *Radical review* 1:108–140.

Alonso-Blanco C, Koornneef M. 2000. Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends in plant science* 5:22–29.

Aragón W, Formey D, Aviles-Baltazar NY, Torres M, Serrano M. 2021. Arabidopsis thaliana Cuticle Composition Contributes to Differential Defense Response to Botrytis cinerea. *Frontiers in Plant Science* [Internet] 12. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2021.738949

Assmann SM. 2013. Natural variation in abiotic stress and climate change responses in Arabidopsis: implications for twenty-first-century agriculture. *International Journal of Plant Sciences* 174:3–26.

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT. 2010. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature* 465:627–631.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B. 2014. lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7)[Computer software].

Bergmann C. 1847. Ober die Verhaltnisse der Warmeokonomie der Thiere zu ihrer Grosse C Bergmann - Gottinger Studien, 1847. *Gottinger Studien* 3:595–708.

Berry HL, Waite TD, Dear KBG, Capon AG, Murray V. 2018. The case for systems thinking about climate change and mental health. *Nature Clim Change* 8:282–290.

Bhaskara GB, Lasky JR, Razzaque S, Zhang L, Haque T, Bonnette JE, Civelek GZ, Verslues PE, Juenger TE. 2022. Natural variation identifies new effectors of water-use efficiency in Arabidopsis. *Proceedings of the National Academy of Sciences* 119:e2205305119.

Binks O, Coughlin I, Mencuccini M, Meir P. 2020. Equivalence of foliar water uptake and stomatal conductance? *Plant, Cell & Environment* 43:524–528.

Bjorkman AD, Myers-Smith IH, Elmendorf SC, Normand S, Rüger N, Beck PS, Blach-Overgaard A, Blok D, Cornelissen JHC, Forbes BC. 2018. Plant functional trait change across a warming tundra biome. *Nature* 562:57–62.

Blanc C, Coluccia F, L'Haridon F, Torres M, Ortiz-Berrocal M, Stahl E, Reymond P, Schreiber L, Nawrath C, Métraux J-P, et al. 2018. The Cuticle Mutant eca2 Modifies Plant Defense Responses to Biotrophic and Necrotrophic Pathogens and Herbivory Insects. *MPMI* 31:344–355.

Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D. 2010. Local-Scale Patterns of Genetic Variability, Outcrossing, and Spatial Structure in Natural Stands of Arabidopsis thaliana. *PLOS Genetics* 6:e1000890.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of Arabidopsis thaliana flowering time in nature. *PLoS genetics* 6:e1000940.

Brennan AC, Méndez-Vigo B, Haddioui A, Martínez-Zapater JM, Picó FX, Alonso-Blanco C. 2014. The genetic structure of Arabidopsis thaliana in the south-western Mediterranean range reveals a shared history between North Africa and southern Europe. *BMC Plant Biology* 14:17.

Brochmann C, Rustan ØH, Lobin W, Kilian N. 1997. The endemic vascular plants of the Cape Verde Islands, W Africa. *Sommerfeltia* 24:1–363.

Bruijnzeel LA, Hamilton LS. 2017. Decision Time Cloud Forests. Available from: http://www.unep.org/resources/report/decision-time-cloud-forests

Bruijnzeel LA, Kappelle M, Mulligan M, Scatena FN. 2011. Tropical montane cloud forests: state of knowledge and sustainability perspectives in a changing world. Cambridge University Press Available from: https://www.cambridge.org/core/books/tropical-montane-cloud-forests/tropical-montane-cloud-forests-state-of-knowledge-and-sustainability-perspectives-in-a-changing-world/D031D6D2C468B6464279174CBC391572

Burkhardt J, Basi S, Pariyar S, Hunsche M. 2012. Stomatal penetration by aqueous solutions – an update involving leaf surface particles. *New Phytologist* 196:774–787.

Campitelli BE, Des Marais DL, Juenger TE. 2016. Ecological interactions and the fitness effect of water-use efficiency: Competition and drought alter the impact of natural *MPK12* alleles in *Arabidopsis*.Enquist B, editor. *Ecology Letters* 19:424–434.

Chan EK, Rowe HC, Hansen BG, Kliebenstein DJ. 2010. The complex genetic architecture of the metabolome. *PLoS genetics* 6:e1001198.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* 6:80–92.

Cohen D. 1970. The expected efficiency of water utilization in plants under different competition and selection regimes. *Israel Journal of Botany* 19:50–54.

Cowan I. 1986. Economics of carbon fixation in higher plants. *Economics of carbon fixation in higher plants.*:133–170.

Darwin C. 1859. The Origin of Species by means of natural selection. London : J. Murray, 1859

Davila Olivas NH, Kruijer W, Gort G, Wijnen CL, van Loon JJ, Dicke M. 2017. Genome-wide association analysis reveals distinct genetic architectures for single and combined stress responses in Arabidopsis thaliana. *New Phytologist* 213:838–851.

Dawson TE. 1998. Fog in the California redwood forest: ecosystem inputs and use by plants. *Oecologia* 117:476–485.

Des Marais DL, Auchincloss LC, Sukamtoh E, Mckay JK, Logan T, Richards JH, Juenger TE. 2014. Variation in MPK12 affects water use efficiency in Arabidopsis and reveals a pleiotropic link between guard cell size and ABA response. *PNAS* 111:2836–2841.

Díaz S, Kattge J, Cornelissen JHC, Wright IJ, Lavorel S, Dray S, Reu B, Kleyer M, Wirth C, Colin Prentice I, et al. 2016. The global spectrum of plant form and function. *Nature* 529:167–171.

Dikmen S, Cole JB, Null DJ, Hansen PJ. 2013. Genome-Wide Association Mapping for Identification of Quantitative Trait Loci for Rectal Temperature during Heat Stress in Holstein Cattle. *PLOS ONE* 8:e69202.

Donovan L, Ehleringer J. 1992. Contrasting water-use patterns among size and life-history classes of a semi-arid shrub. *Functional Ecology*:482–488.

Donovan L, Ehleringer J. 1994. Water stress and use of summer precipitation in a Great Basin shrub community. *Functional Ecology*:289–297.

Durvasula A, Fulgione A, Gutaker RM, Alacakaptan SI, Flood PJ, Neto C, Tsuchimatsu T, Burbano HA, Picó FX, Alonso-Blanco C, et al. 2017. African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. *PNAS* 114:5213–5218.

Exposito-Alonso M, Vasseur F, Ding W, Wang G, Burbano HA, Weigel D. 2018. Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nature ecology & evolution* 2:352–358.

Falconer DS, Mackay TF. 1983. Quantitative genetics. Longman London, UK

Ferrero-Serrano Á, Assmann SM. 2019. Phenotypic and genome-wide association with the local environment of Arabidopsis. *Nature Ecology & Evolution* 3:274–285.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology* 37:4302–4315.

Filiault DL, Maloof JN. 2012. A genome-wide association study identifies variants underlying the Arabidopsis thaliana shade avoidance response. *PLoS genetics* 8:e1002589.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in Arabidopsis thaliana. *Science* 334:86–89.

Fulgione A, Koornneef M, Roux F, Hermisson J, Hancock AM. 2018. Madeiran Arabidopsis thaliana reveals ancient long-range colonization and clarifies demography in Eurasia. *Molecular Biology and Evolution* 35:564–574.

Fulgione A, Neto C, Elfarargi AF, Tergemina E, Ansari S, Göktay M, Dinis H, Döring N, Flood PJ, Rodriguez-Pacheco S, et al. 2022. Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. *Nat Commun* 13:1461.

Gautier M, Klassmann A, Vitalis R. 2017. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Molecular ecology resources* 17:78–90.

Gazzani S, Gendall AR, Lister C, Dean C. 2003. Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. *Plant Physiology* 132:1107–1114.

Ge SX, Jung D, Yao R. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36:2628–2629.

Geber MA, Dawson TE. 1990. Genetic variation in and covariation between leaf gas exchange, morphology, and development in Polygonum arenastrum, an annual plant. *Oecologia* 85:153–158.

Geber MA, Dawson TE. 1997. Genetic variation in stomatal and biochemical limitations to photosynthesis in the annual plant, Polygonum arenastrum. *Oecologia* 109:535–546.

Gloss AD, Vergnol A, Morton TC, Laurin PJ, Roux F, Bergelson J. 2022. Genome-wide association mapping within a local Arabidopsis thaliana population more fully reveals the genetic architecture for defensive metabolite diversity. *Philosophical Transactions of the Royal Society B* 377:20200512.

Gowik U, Bräutigam A, Weber KL, Weber APM, Westhoff P. 2011. Evolution of C4 Photosynthesis in the Genus Flaveria: How Many and Which Genes Does It Take to Make C4? *The Plant Cell* 23:2087–2105.

Grant PR. 1999. Ecology and Evolution of Darwin's Finches. Princeton University Press

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J. 2011. Adaptation to climate across the Arabidopsis thaliana genome. *Science* 334:83–86.

Heschel MS, Riginos C. 2005. Mechanisms of selection for drought stress tolerance and avoidance in Impatiens capensis (Balsaminaceae). *American Journal of Botany* 92:37–44.

Hijmans RJ, Van Etten J, Cheng J, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, Bevan A, Racine EB, Shortridge A. 2015. Package 'raster.' *R package* 734.

Hudson RR, Bailey K, Skarecky D, Kwiatowski J, Ayala FJ. 1994. Evidence for positive selection in the superoxide dismutase (Sod) region of Drosophila melanogaster. *Genetics* 136:1329–1340.

Jakobson L, Vaahtera L, Tõldsepp K, Nuhkat M, Wang C, Wang Y-S, Hõrak H, Valk E, Pechter P, Sindarovska Y, et al. 2016. Natural Variation in Arabidopsis Cvi-0 Accession Reveals an Important Role of MPK12 in Guard Cell CO2 Signaling. *PLoS biology* 14:e2000322.

Jammes F, Song C, Shin D, Munemasa S, Takeda K, Gu D, Cho D, Lee S, Giordo R, Sritubtim S, et al. 2009. MAP kinases MPK9 and MPK12 are preferentially expressed in guard cells and positively regulate ROS-mediated ABA signaling. *PNAS* 106:20520–20525.

Jiang Y, Liang G, Yu D. 2012. Activated Expression of WRKY57 Confers Drought Tolerance in Arabidopsis. *Molecular Plant* 5:1375–1388.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. 2000. Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. *Science* 290:344–347.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.

de Jong G. 1993. Covariances Between Traits Deriving From Successive Allocations of a Resource. *Functional Ecology* 7:75–83.

Juenger TE, Mckay JK, Hausmann N, Keurentjes JJ, Sen S, Stowe KA, Dawson TE, Simms EL, Richards JH. 2005. Identification and characterization of QTL underlying whole-plant physiology in Arabidopsis thaliana: δ13C, stomatal conductance and transpiration efficiency. *Plant, Cell & Environment* 28:697–708.

Kahle D, Wickham H. 2013. ggmap: Spatial Visualization with ggplot2. *The R Journal* 5:144–161.

Kalladan R, Lasky JR, Chang TZ, Sharma S, Juenger TE, Verslues PE. 2017. Natural variation identifies genes affecting drought-induced abscisic acid accumulation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 114:11536–11541.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific data* 4:1–20.

Karger DN, Kessler M, Lehnert M, Jetz W. 2021. Limited protection and ongoing loss of tropical cloud forest biodiversity and ecosystems worldwide. *Nat Ecol Evol* 5:854–862.

Kassambara A. 2020. ggpubr:"ggplot2" based publication ready plots (R package version 0.4. 0)[Computer software].

Kenney AM, McKay JK, Richards JH, Juenger TE. 2014. Direct and indirect selection on flowering time, water-use efficiency (WUE, δ^{13} C), and WUE plasticity to drought in *Arabidopsis thaliana*. *Ecology and Evolution* 4:4505–4521.

Kerfoot O. 1968. Mist precipitation on vegetation.

Klein T. 2014. The variability of stomatal sensitivity to leaf water potential across tree species indicates a continuum between isohydric and anisohydric behaviours. *Functional Ecology* 28:1313–1320.

Klemm O, Schemenauer RS, Lummerich A, Cereceda P, Marzol V, Corell D, van Heerden J, Reinhard D, Gherezghiher T, Olivier J, et al. 2012. Fog as a Fresh-Water Resource: Overview and Perspectives. *Ambio* 41:221–234.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in Arabidopsis thaliana. *Annual review of plant biology* 55:141–172.

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J. 2007. Fitness Effects Associated with the Major Flowering Time Gene FRIGIDA in Arabidopsis thaliana in the Field. *The American Naturalist* 169:E141–E157.

Kranz AR, Kirchheim B. 1987. Genetic resources in Arabidopsis. Arabidopsis Information Service

Lafuente E, Duneau D, Beldade P. 2018. Genetic basis of thermal plasticity variation in Drosophila melanogaster body size. *PLOS Genetics* 14:e1007686.

Lasky JR, Des Marais DL, McKAY JK, Richards JH, Juenger TE, Keitt TH. 2012. Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Molecular Ecology* 21:5512–5529.

Levitt J. 1972. Responses of plants to environmental stresses.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993.

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 107:21199–21204.

Liu Y, Maierhofer T, Rybak K, Sklenar J, Breakspear A, Johnston MG, Fliegmann J, Huang S, Roelfsema MRG, Felix G, et al. 2019. Anion channel SLAH3 is a regulatory target of chitin receptor-associated kinase PBL27 in microbial stomatal closure.Zhou J-M, Hardtke CS, Hoth S, editors. *eLife* 8:e44474.

Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457:830–836.

Losos JB, Warheitt KI, Schoener TW. 1997. Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature* 387:70–73.

Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK. 2013. Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Proceedings of the Royal Society B: Biological Sciences* 280:20131043.

Ludlow M. 1989. Strategies of response to water stress.

Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sinauer Sunderland, MA

Martin SH, Van Belleghem SM. 2017. Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. *Genetics* 206:429–438.

Martínez-Vilalta J, Poyatos R, Aguadé D, Retana J, Mencuccini M. 2014. A new look at water transport regulation in plants. *New Phytologist* 204:105–115.

McDowell N, Pockman WT, Allen CD, Breshears DD, Cobb N, Kolb T, Plaut J, Sperry J, West A, Williams DG, et al. 2008. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New Phytologist* 178:719–739.

McDowell NG, Beerling DJ, Breshears DD, Fisher RA, Raffa KF, Stitt M. 2011. The interdependence of mechanisms underlying climate-driven vegetation mortality. *Trends in Ecology & Evolution* 26:523–532.

McKay JK, Richards JH, Mitchell-Olds T. 2003. Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Molecular ecology* 12:1137–1151.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* 441:947–952.

Montillet J-L, Leonhardt N, Mondy S, Tranchimand S, Rumeau D, Boudsocq M, Garcia AV, Douki T, Bigeard J, Laurière C, et al. 2013. An Abscisic Acid-Independent Oxylipin Pathway Controls Stomatal Closure and Immune Defense in Arabidopsis. *PLOS Biology* 11:e1001513.

Mooney H, Ehleringer J, Berry J. 1976. High photosynthetic capacity of a winter annual in Death Valley. *Science* 194:322–324.

Moreno-Gutiérrez C, Dawson TE, Nicolás E, Querejeta JI. 2012. Isotopes reveal contrasting water use strategies among coexisting plant species in a Mediterranean ecosystem. *New Phytologist* 196:489–496.

Morrison K, Stacy E. 2014. Intraspecific divergence and evolution of a life-history trade-off along a successional gradient in Hawaii's Metrosideros polymorpha. *Journal of Evolutionary Biology* 27:1192–1204.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, et al. 2020. vegan: Community Ecology Package. Available from: https://CRAN.R-project.org/package=vegan

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010. The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis thaliana. *Science* [Internet]. Available from: https://www.science.org/doi/abs/10.1126/science.1180677

Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R.Schwartz R, editor. *Bioinformatics* 35:526–528.

Porch TG, Tseung C-W, Schmelz EA, Mark Settles A. 2006. The maize Viviparous10/Viviparous13 locus encodes the Cnx1 gene required for molybdenum cofactor biosynthesis. *The Plant Journal* 45:250–263.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 81:559–575.

Querejeta JI, Prieto I, Torres P, Campoy M, Alguacil MM, Roldán A. 2018. Water-spender strategy is linked to higher leaf nutrient concentrations across plant species colonizing a dry and nutrient-poor epiphytic habitat. *Environmental and Experimental Botany* 153:302–310.

R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. *Retrieved from http://www.R-project.org*.

Ray C. 1960. The application of Bergmann's and Allen's rules to the poikilotherms. *Journal of Morphology* 106:85–108.

van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR. 2019. Genetic correlations of polygenic disease traits: from theory to practice. *Nat Rev Genet* 20:567–581.

Roff DA. 2002. Life history evolution. Sinauer Associates Sunderland, MA

Roux F, Frachon L. 2022. A Genome-Wide Association study in Arabidopsis thaliana to decipher the adaptive genetics of quantitative disease resistance in a native heterogeneous environment. *PloS one* 17:e0274561.

Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419:832–837.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449:913–918.

Salam MA, Jammes F, Hossain MA, Ye W, Nakamura Y, Mori IC, Kwak JM, Murata Y. 2013. Two guard cell-preferential MAPKs, MPK9 and MPK12, regulate YEL signalling in Arabidopsis guard cells. *Plant Biology* 15:436–442.

Salomé PA, Bomblies K, Fitz J, Laitinen R a. E, Warthmann N, Yant L, Weigel D. 2012. The recombination landscape in Arabidopsis thaliana F2 populations. *Heredity* 108:447–455.

Schemenauer R, Bignell B, Makepeace T. 2016. Fog Collection Projects in Nepal: 1997 to 2016. In: p. 24–19.

Schemenauer R, Zanetta N, Rosato M, Carter Gamberini MV. 2016. The Tojquia, Guatemala Fog Collection Project 2006 to 2016.

Shefferson RP, Proper J, Beissinger SR, Simms EL. 2003. Life history trade-offs in a rare orchid: the costs of flowering, dormancy, and sprouting. *Ecology* 84:1199–1206.

Sherrard ME, Maherali H. 2006. The Adaptive Significance of Drought Escape in Avena barbata, an Annual Grass. *Evolution* 60:2478–2489.

Shim H, Chasman DI, Smith JD, Mora S, Ridker PM, Nickerson DA, Krauss RM, Stephens M. 2015. A Multivariate Genome-Wide Association Analysis of 10 LDL Subfractions, and Their Response to Statin Treatment, in 1868 Caucasians. *PLOS ONE* 10:e0120758.

Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C. 2005. Role of FRIGIDA and FLOWERING LOCUS C in Determining Variation in Flowering Time of Arabidopsis. *Plant Physiology* 138:1163–1173.

Skelton RP, West AG, Dawson TE. 2015. Predicting plant vulnerability to drought in biodiverse regions using functional traits. *Proceedings of the National Academy of Sciences* 112:5744–5749.

Speidel L, Forest M, Shi S, Myers SR. 2019. A method for genome-wide genealogy estimation for thousands of samples. *Nature Genetics 2019 51:9* 51:1321–1329.

Stadtmüller T. 1987. Cloud Forests in the Humid Tropics: A Bibliographic Review. Bib. Orton IICA / CATIE

Stearns SC. 1989. Trade-Offs in Life-History Evolution. *Functional Ecology* 3:259–268.

Stern AJ, Wilton PR, Nielsen R. 2019. An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data.Hernandez RD, editor. *PLOS Genetics* 15:e1008384.

Tabas-Madrid D, Méndez-Vigo B, Arteaga N, Marcer A, Pascual-Montano A, Weigel D, Xavier Pico F, Alonso-Blanco C. 2018. Genome-wide signatures of flowering adaptation to climate temperature: Regional analyses in a highly diverse native range of Arabidopsis thaliana. *Plant, cell & environment* 41:1806–1820.

Tergemina E, Elfarargi AF, Flis P, Fulgione A, Göktay M, Neto C, Scholle M, Flood PJ, Xerri S-A, Zicola J. 2022. A two-step adaptive walk rewires nutrient transport in a challenging edaphic environment. *Science Advances* 8:eabm9385.

Tisné S, Serrand Y, Bach L, Gilbault E, Ben Ameur R, Balasse H, Voisin R, Bouchez D, Durand-Tardif M, Guerche P, et al. 2013. Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *The Plant Journal* 74:534–544.

Verslues PE, Juenger TE. 2011. Drought, metabolites, and Arabidopsis natural variation: a promising combination for understanding adaptation to water-limited environments. *Current opinion in plant biology* 14:240–245.

Villellas J, García M. 2018. Life-history trade-offs vary with resource availability across the geographic range of a widespread plant. *Plant Biology* 20:483–489.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A Map of Recent Positive Selection in the Human Genome.Hurst L, editor. *PLoS Biology* 4:e72.

Wallace AR. 1855. On the law which has regulated theintroduction of new species. *Annals and Magazine of Natural History, 2nd Series*:184–196.

Walter H. 1985. Vegetation of the Earth and Ecological Systems of the Geo-biosphere. 3rd ed. Springer Berlin, Heidelberg Available from: https://doi.org/10.1007/978-3-642-96859-4

Weathers KC. 1999. The importance of cloud and fog in the maintenance of ecosystems. *Trends in Ecology & Evolution* 14:214–215.

Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant physiology* 158:2–22.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*.

Wieters B, Steige KA, He F, Koch EM, Ramos-Onsins SE, Gu H, Guo Y-L, Sunyaev S, de Meaux J. 2021. Polygenic adaptation of rosette growth in Arabidopsis thaliana. *PLoS genetics* 17:e1008748.

Will RE, Wilson SM, Zou CB, Hennessey TC. 2013. Increased vapor pressure deficit due to higher temperature leads to greater transpiration and faster mortality during drought for tree seedlings common to the forest–grassland ecotone. *New Phytologist* 200:366–374.

Willi Y, Van Buskirk J. 2022. A review on trade-offs at the warm and cold ends of geographical distributions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377:20210022.

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JH, Diemer M. 2004. The worldwide leaf economics spectrum. *Nature* 428:821–827.

Zan Y, Carlborg Ö. 2019. A polygenic genetic architecture of flowering time in the worldwide Arabidopsis thaliana population. *Molecular biology and evolution* 36:141–154.

Zera AJ, Harshman LG. 2001. The Physiology of Life History Trade-Offs in Animals. *Annual Review of Ecology and Systematics* 32:95–126.

Zhou X. 2014. Gemma user manual. *Univ. Chicago, USA*.

Zhou X, Carbonetto P, Stephens M. 2013. Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *PLOS Genetics* 9:e1003264.

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* 44:821–824.

Zhou X, Stephens M. 2014. efficient multivariate linear mixed model algorithms for genome-wide association studies. 11:407.

Zhou Z, Li M, Cheng H, Fan W, Yuan Z, Gao Q, Xu Y, Guo Z, Zhang Y, Hu J. 2018. An intercross population study reveals genes associated with body size and plumage color in ducks. *Nature Communications* 9:1–10.

Figure Legends

Figure 1. Collection locations of *A. thaliana* Santo Antão, CVI (blue, n=189) and Morocco (green, n=61). The arrow indicates the colonization of the Cape Verde Islands from Morocco approximately 5-7 kya.

Figure 2. Phenotypic variation in (A) stomatal conductance (gs) and (B) water use efficiency (WUE) for Santo Antão, CVI and Moroccan *A. thaliana* populations in well-watered (WW) and water deficit (WD) conditions. The line in the center of the boxplots represents the median, the box edges represent the 25^{th} and 75^{th} percentiles (lower and upper bound, respectively), and the whiskers represent the 95% CI. The WUE is measured as carbon isotope discrimination ($\delta^{13}C$), and the carbon isotope ratio is expressed per mil, ‰. The *P*-values for the Mann-Whitney-Wilcoxon test are shown.

Figure 3. Genome-wide association (GWA) mapping of water use efficiency (WUE**)**. (A) Polygenic modeling of the average WUE across the well-watered (WW) and water deficit (WD) conditions (left) and the drought response of WUE (the difference between both conditions: WW-WD) (right) in the Santo Antão *A. thaliana* population using BSLMM. The y-axis represents the posterior inclusion probability (PIP) and the size of symbol denotes the effect size. (B) Genome-wide association mapping of average WUE (left) and its drought response (right) using LMM. The horizontal dashed line corresponds to the Bonferroni significance threshold at $\alpha = 0.05$. In A and B, points represent SNPs along the five chromosomes. The red point at chromosome 2 represents the *MPK12* G53R variant (a substitution of arginine for glycine at amino acid position 53).

Figure 4. Population structure of *A. thaliana* sub-populations in Santo Antão**.** (A) Geographical distribution of sub-populations across Santo Antão. Images of representative sites for the five sub-populations in Santo Antão show the diversity of habitats that A. thaliana occupies. (B) PCA of genome-wide SNPs showing clustering within Santo Antão. Colors for sub-populations in 'B' match those used in 'A'. Abbreviations: Figueira: Lombo de Figueira; Cova: Cova de Paúl; Ribeira: Ribeira de Poio; Pico: Pico da Cruz.

Figure 5. Evolutionary history of water-use efficiency variation. (A) Estimated allele ages (inferred in RELATE) versus allele frequencies of variants with major effects estimated from GWA mapping of average water use efficiency (WUE). Color denotes variant effect size from

BSLMM and shape denotes predicted impact from gene annotation. (B) Spatial distribution of *MPK12* G53R in Santo Antão. Pie charts show the frequency of *MPK12* alleles, with size representing the number of individuals per sampling location. (C) Marginal genealogical tree estimated in RELATE for *MPK12* G53R. Yellow denotes the derived allele. (D) A model of the origin and spread of the *MPK12* G53R variant based on the genealogical inference in (C). Abbreviations: Figueira: Lombo de Figueira; Cova: Cova de Paúl; Ribeira: Ribeira de Poio; Pico: Pico da Cruz.

Figure 6. Signature of a partial selective sweep at *MPK12* G53R. (A) The decay of extended haplotype homozygosity (EHH) and (B) bifurcation analysis of the ancestral (violet) and derived (yellow) alleles for *MPK12* G53R. The grey dotted line marks the position of the focal *MPK12* SNP. The width of the lines in (B) represents the frequency of haplotypes bearing the ancestral and derived *MPK12* variant. (C) Genome-wide integrated haplotype score (*iHS*) analysis for the Santo Antão population. Horizontal solid line represents the significance threshold applied to detect the outlier SNPs ($-\log_{10}(P) = 2$) and the horizontal dashed line represents the 1% tail based on the genome-wide empirical distribution. The *MPK12* 53R variant is marked with a dark red point. (D) Geographic distribution of Santo Antão *A. thaliana* individuals overlaid on the max temperature of warmest month (BIO5). Abbreviations: Figueira: Lombo de Figueira; Cova: Cova de Paúl; Ribeira: Ribeira de Poio; Pico: Pico da Cruz.

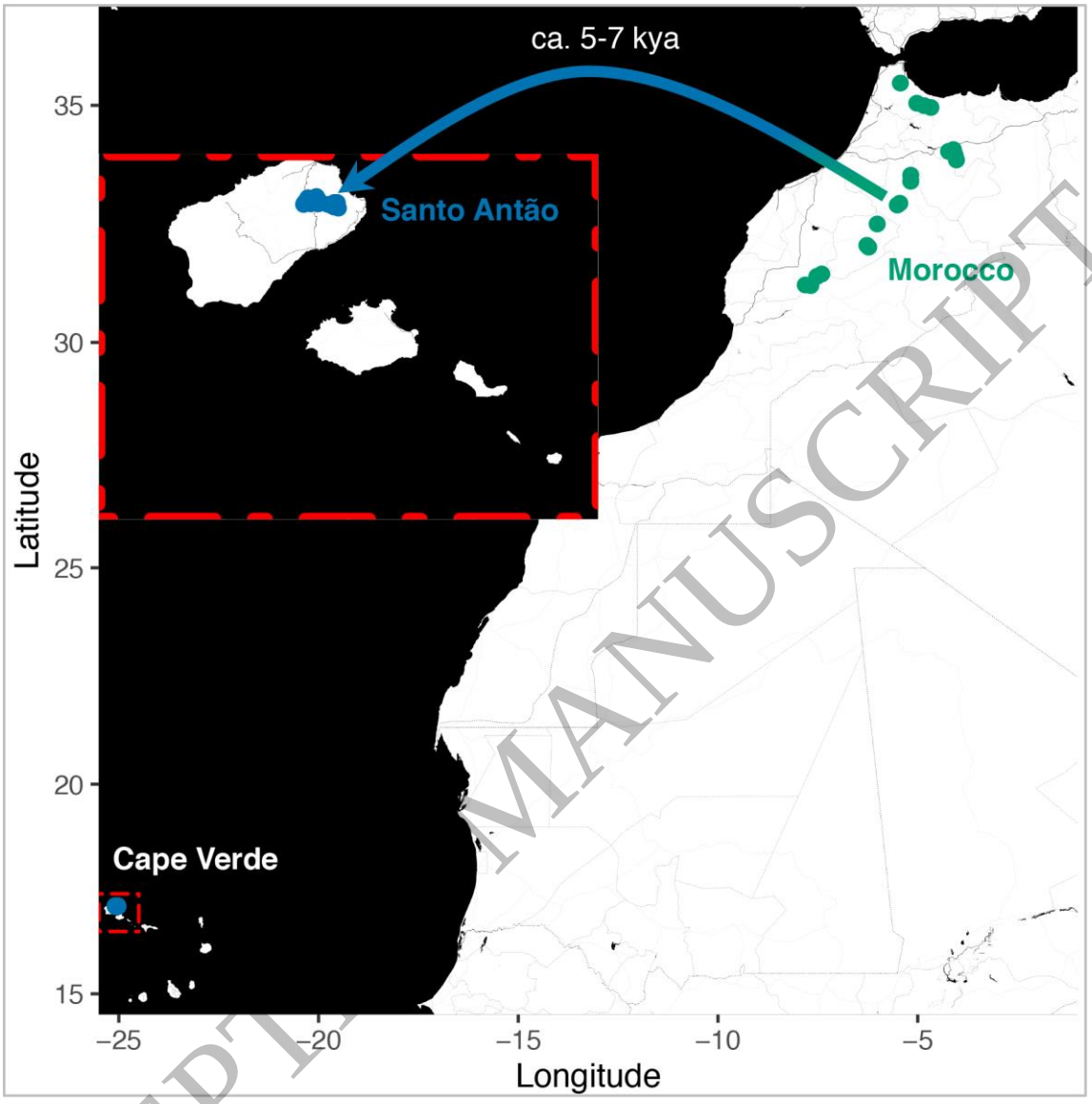

Figure 1
151x152 mm (.80 x DPI)

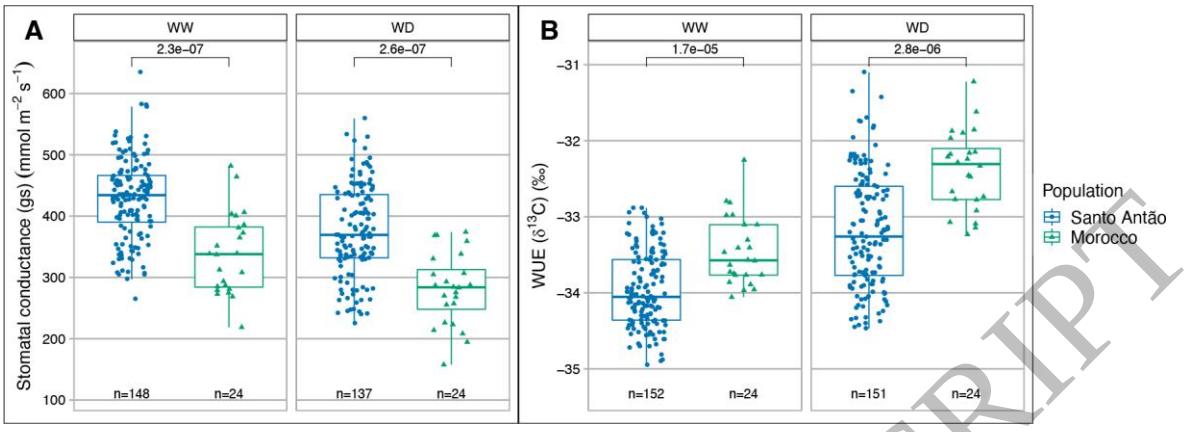

Figure 2
154x55 mm (.80 x DPI)

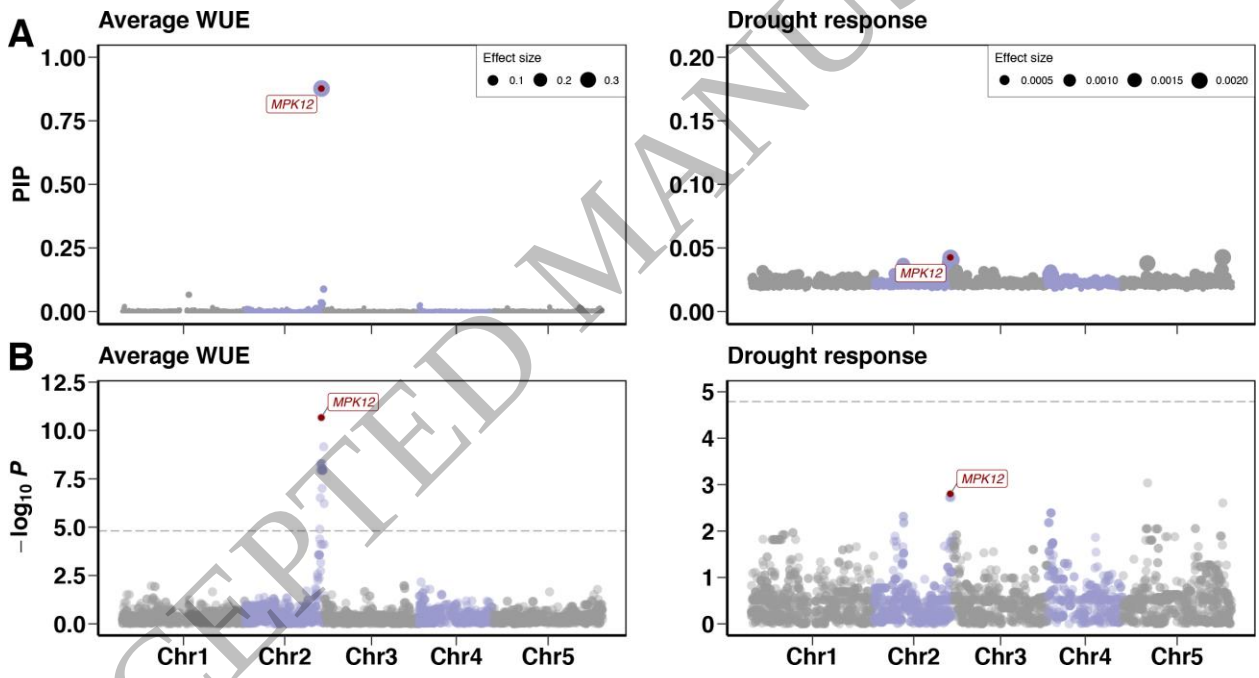

Figure 3
165x86 mm (.80 x DPI)

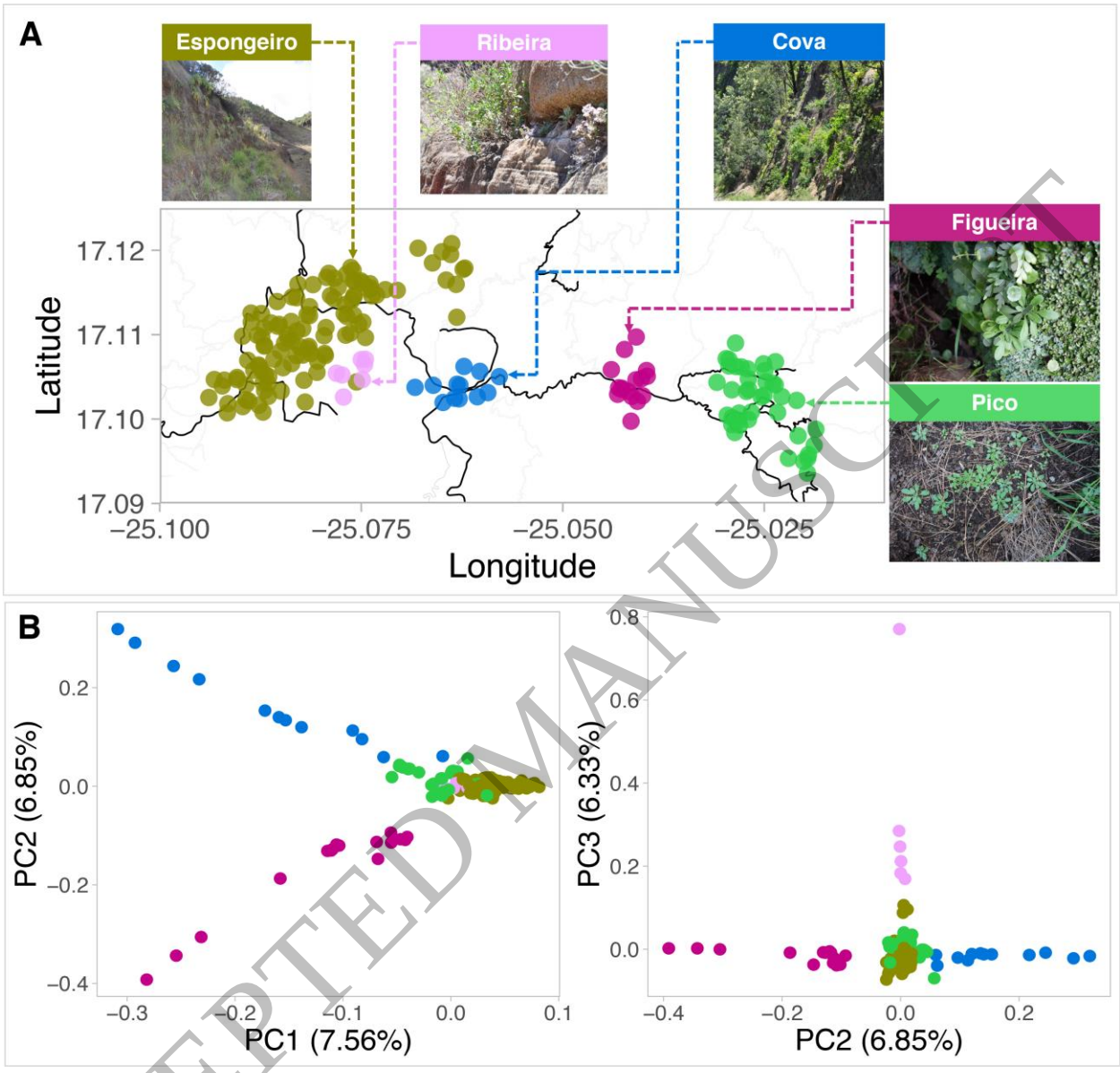

Figure 4
165x157 mm (.80 x DPI)

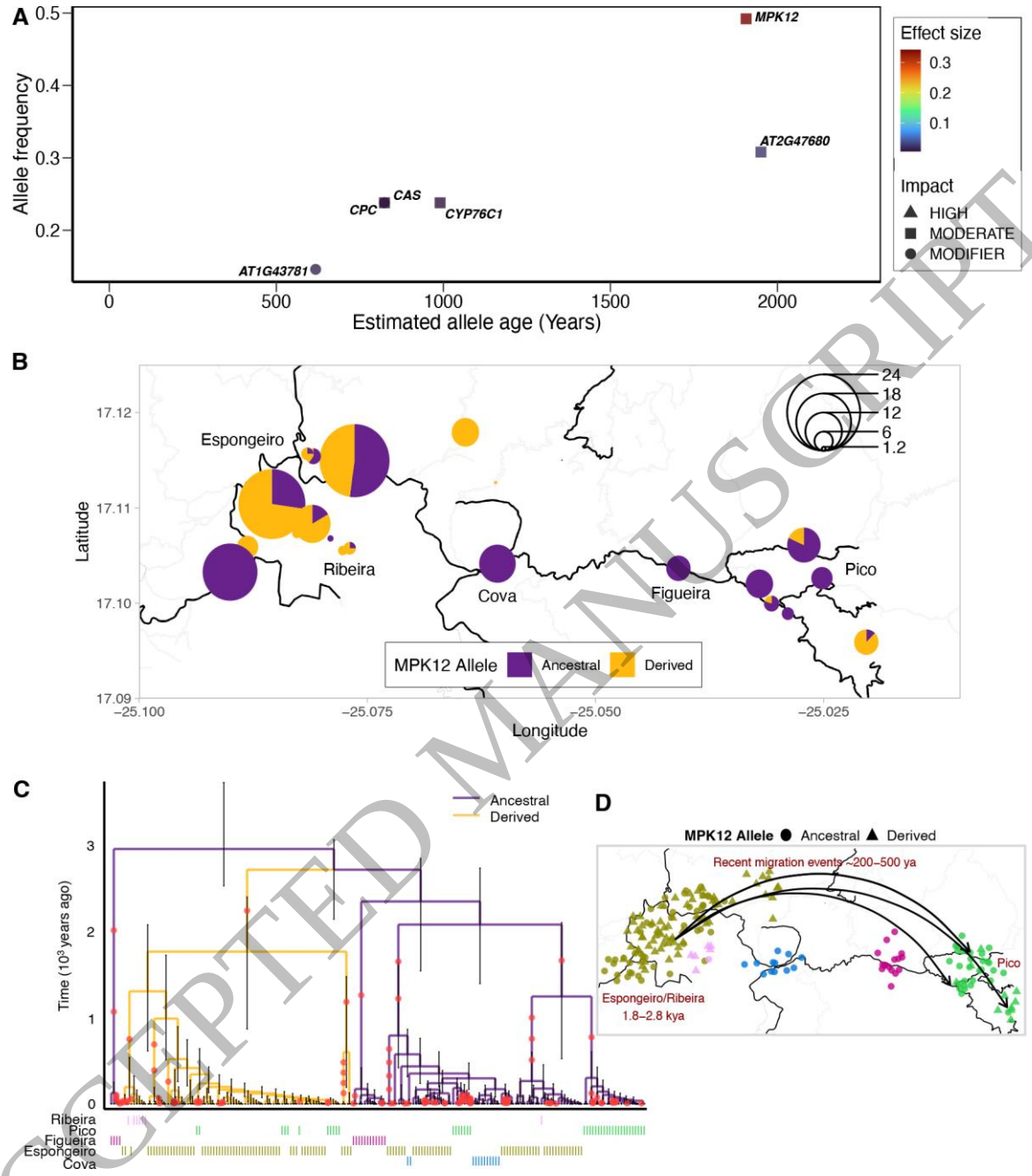

Figure 5
151x174 mm (.80 x DPI)

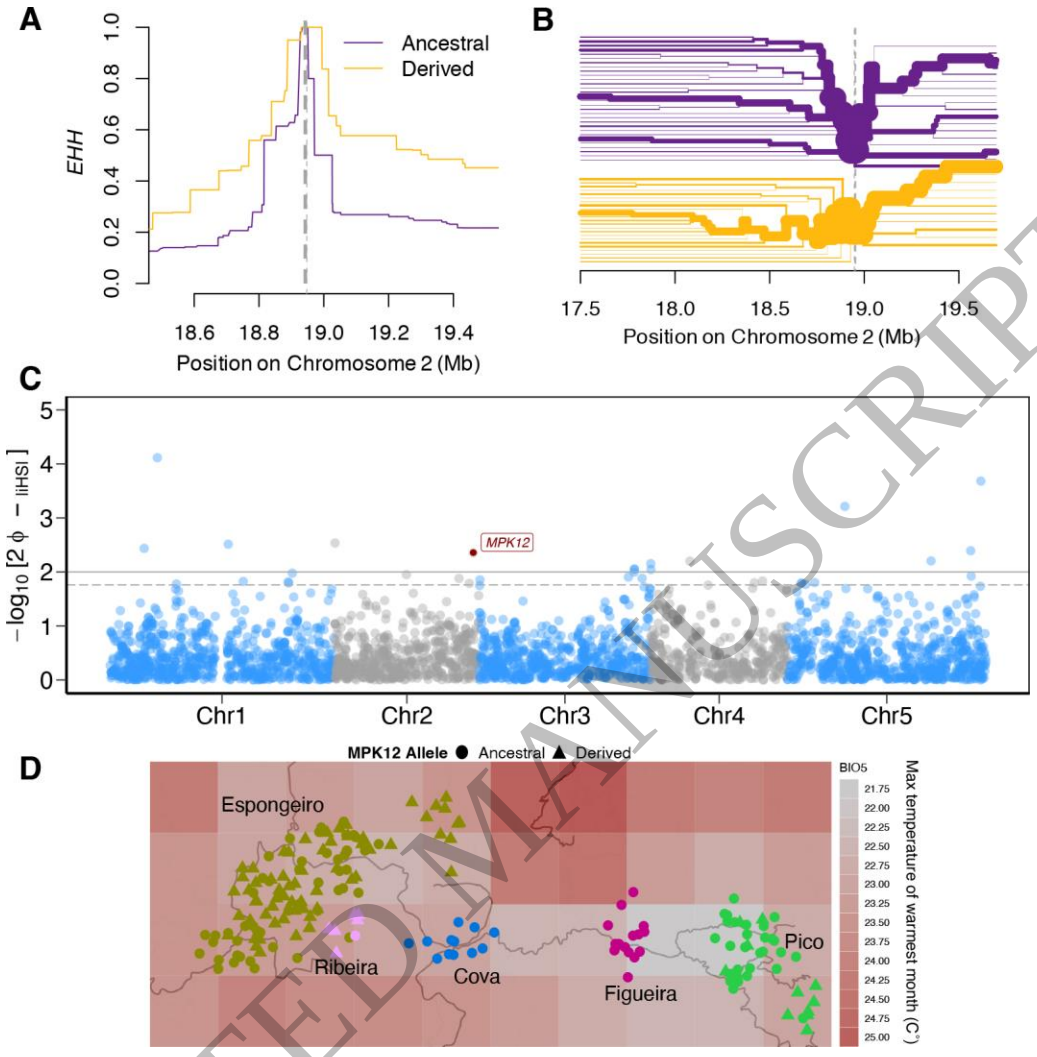

Figure 6
135x138 mm (.80 x DPI)