Genetic basis and adaptive relevance of drought response in Cape Verde *Arabidopsis*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Ahmed Elfarargi

aus Kairo, Ägypten

Köln

2023

Die vorliegende Arbeit wurde am Max-Planck-Institut für die Pflanzenzüchtungsforschung in Köln in in der Arbeitsgruppe von (Dr. Angela M. Hancock) angefertigt.

The work described in this dissertation was conducted under the supervision of Dr. Angela M. Hancock at the Max Planck Institute for Plant Breeding Research.

Erster Referent und Prüfer: Dr. Angela M. Hancock

Zweite Referentin und Prüferin: Prof. Dr. Ute Höcker

Vorsitzende der Prüfungskommission: Prof. Dr. Christine Heim

Tag der mündlichen Prüfung: 11.07.2023

Summary

Climate change is predicted to impact precipitation patterns, leading to shorter growing seasons and increased susceptibility to drought in many regions worldwide. These changes significantly threaten plant populations and may result in ecosystem desertification. Understanding the mechanisms enabling species to adapt to such changes is crucial for effective conservation strategies and developing resilient crop varieties. Plants cope with drought through various strategies, including avoidance, escape, and drought tolerance, which can be canalized or plastic, depending on the genetic and environmental context. Understanding the balance between canalization and plasticity is essential for predicting plant responses to future climate change. Here, we investigated the genetic architecture of drought adaptation in natural Cape Verdean *Arabidopsis thaliana* populations.

In Chapter One, we reviewed the impact of climate change on precipitation patterns and its consequences on plant populations, including increased susceptibility to drought and extinction risk. We discussed various strategies plants employ to cope with drought, such as avoidance, escape, and drought tolerance. We also discussed the importance of understanding the balance between canalization and plasticity for predicting plant responses to future climate changes. We also highlighted the significance of genetic adaptations in enabling species to adapt and persist in rapidly changing environments and the potential insights gained from studying *A. thaliana* populations on the Cape Verde Islands (CVI), which have experienced rapid adaptation and evolutionary rescue in response to drought-prone climates.

In Chapter Two, we investigated the evolution of stomatal conductance and water use efficiency (WUE) in an *A. thaliana* population that colonized an island with a montane cloud scrubland ecosystem characterized by seasonal drought and fog-based precipitation. We found 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 revealed 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. Furthermore, we reconstructed the spatially-explicit evolutionary history of *MPK12* 53R on the island and demonstrated that this allele increased in frequency due to positive selection as *A. thaliana* expanded into harsher regions of the island. The findings showed how adaptation shaped quantitative eco-physiological traits in a new precipitation regime defined by low rainfall and high humidity.

In Chapter Three, we examined the genetic architecture of variation in growth rate, leaf color, and stomatal patterning in response to precisely controlled water conditions among CVI *A. thaliana* populations. Genome-wide association mapping analyses revealed that moderately complex genetic architectures with roles for several major effect variants underlie variation in these traits. Furthermore, we found that several identified genes through genetic mapping have pleiotropic functions for complex traits underlying drought stress, highlighting the intricate nature of plant adaptation to these challenging conditions.

In conclusion, this work presents a comprehensive analysis of the mechanisms and genetic basis of plant adaptation to drought stress, focusing on the natural *A. thaliana* populations in the CVI islands. Understanding these mechanisms is critical for predicting species distribution and adaptive responses to drought stress. Furthermore, our findings expand our knowledge of how drought adaptation results from numerous genetic variants, suggesting polygenic adaptation, and reveal that new mutations arise frequently enough to potentially facilitate rapid adaptation in colonizing populations. Lastly, these findings enrich our understanding of plant responses to drought and provide valuable insights for developing effective conservation strategies and resilient crop varieties.

Zusammenfassung

Der Klimawandel wird voraussichtlich die Niederschlagsmuster beeinflussen, was in vielen Regionen weltweit zu kürzeren Vegetationsperioden und erhöhter Anfälligkeit für Dürre führt. Diese Veränderungen stellen eine erhebliche Bedrohung für Pflanzenpopulationen dar und können zur Versteppung von Ökosystemen führen. Das Verständnis der Mechanismen, die es den Arten ermöglichen, sich an solche Veränderungen anzupassen, ist entscheidend für wirksame Erhaltungsstrategien und die Entwicklung widerstandsfähiger Nutzpflanzensorten. Pflanzen bewältigen Dürre durch verschiedene Strategien, einschließlich Vermeidung, Flucht und Dürretoleranz, die je nach genetischem und Umweltkontext kanalisiert oder plastisch sein können. Das Verständnis des Gleichgewichts zwischen Kanalisierung und Plastizität ist entscheidend für die Vorhersage von Pflanzenreaktionen auf den zukünftigen Klimawandel. In dieser Arbeit untersuchten wir die genetische Architektur der Dürreadaptation in natürlichen Kapverdischen *Arabidopsis thaliana*-Populationen.

Im ersten Kapitel haben wir den Einfluss des Klimawandels auf Niederschlagsmuster und dessen Folgen für Pflanzenpopulationen untersucht, einschließlich erhöhter Anfälligkeit für Dürre und Aussterberisiko. Wir diskutierten verschiedene Strategien, die Pflanzen zur Bewältigung von Dürre einsetzen, wie Vermeidung, Flucht und Dürretoleranz. Wir erörterten auch die Bedeutung des Verständnisses des Gleichgewichts zwischen Kanalisierung und Plastizität für die Vorhersage von Pflanzenreaktionen auf zukünftige Klimaveränderungen. Darüber hinaus betonten wir die Bedeutung genetischer Anpassungen, die es Arten ermöglichen, sich an sich schnell verändernde Umgebungen anzupassen und darin zu bestehen, und die potenziellen Erkenntnisse, die sich aus der Untersuchung von *A. thaliana*-Populationen auf den Kapverdischen Inseln (CVI) ergeben, die eine rasche Anpassung und evolutionäre Rettung in Reaktion auf dürreanfällige Klimata erfahren haben

In Kapitel zwei untersuchten wir die Evolution der stomatären Leitfähigkeit und Wasserbenutzungseffizienz (WUE) in einer *A. thaliana*-Population, die eine Insel mit einem montanen Wolkenscrubland-Ökosystem besiedelte, das durch saisonale Dürre und nebelbasierte Niederschläge gekennzeichnet ist. Wir stellten fest, dass die stomatäre Leitfähigkeit in der besiedelnden Population im Vergleich zu ihrer nächsten ausgruppierenden Population aus dem gemäßigten Nordafrika zunimmt und die WUE abnimmt. Die genomweite Assoziationskartierung ergab eine polygene Basis der Merkmalsvariation, mit einem erheblichen Beitrag von einer nicht-synonymen SNP in *MAP KINASE 12* (*MPK12* G53R), die 35% der phänotypischen Varianz in WUE in der Inselbevölkerung erklärt. Darüber hinaus rekonstruierten wir die räumlich explizite evolutionäre Geschichte von *MPK12* 53R auf der Insel und zeigten, dass dieses Allel aufgrund positiver Selektion an Frequenz zunahm, als *A. thaliana* in härtere Regionen der Insel expandierte. Die Ergebnisse zeigten, wie die Anpassung quantitative ökophysiologische Merkmale in einem neuen Niederschlagsregime prägte, das durch geringe Niederschlagsmengen und hohe Luftfeuchtigkeit definiert ist.

In Kapitel drei untersuchten wir die genetische Architektur der Variation in Wachstumsrate, Blattfarbe und stomatärem Muster in Reaktion auf präzise kontrollierte Wasserbedingungen unter CVI *A. thaliana*-Populationen. Genomweite Assoziationskartierungsanalysen ergaben, dass mäßig komplexe genetische Architekturen mit Rollen für mehrere Haupteffektvarianten der Variation in diesen Merkmalen zugrunde liegen. Darüber hinaus stellten wir fest, dass mehrere durch genetische Kartierung identifizierte Gene pleiotrope Funktionen für komplexe Merkmale aufweisen, die der Dürrebelastung zugrunde liegen, und betonten so die komplexe Natur der Pflanzenanpassung an diese herausfordernden Bedingungen.

Abschließend präsentiert diese Arbeit eine umfassende Analyse der Mechanismen und genetischen Grundlagen der Pflanzenanpassung an Dürrebelastung, mit Schwerpunkt auf den natürlichen *A. thaliana*-Populationen auf den Kapverdischen Inseln (CVI) Das

Verständnis dieser Mechanismen ist entscheidend für die Vorhersage von Artverbreitung und adaptiven Reaktionen auf Dürrebelastung. Darüber hinaus erweitern unsere Ergebnisse unser Wissen darüber, wie Dürreadaptation aus zahlreichen genetischen Varianten resultiert, was auf polygene Anpassung hindeutet, und zeigen, dass neue Mutationen häufig genug auftreten, um möglicherweise eine schnelle Anpassung in kolonisierenden Populationen zu erleichtern. Schließlich bereichern diese Erkenntnisse unser Verständnis der Pflanzenreaktionen auf Dürre und liefern wertvolle Einblicke für die Entwicklung effektiver Erhaltungsstrategien und widerstandsfähiger Nutzpflanzensorten.

Publications

Ahmed F Elfarargi, Elodie Gilbault, Nina Döring, Célia Neto, Andrea Fulgione, Andreas P M Weber, Olivier Loudet, Angela M. Hancock, **Genomic basis of adaptation to a novel precipitation regime**, *Molecular Biology and Evolution*, 40, no. **3** (2023): msad031, https://doi.org/10.1093/molbev/msad031

Author 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 δ^{13}C analysis. N.K., A.P.M.W and his lab members were responsible for δ^{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.

Additional published works not featured in this thesis

1- Andrea Fulgione, Célia Neto, **Ahmed F. Elfarargi**, Emmanuel Tergemina, Shifa Ansari, Mehmet Göktay, Herculano Dinis, Nina Döring, Pádraic J. Flood, Sofia Rodriguez-Pacheco, Nora Walden, Marcus A. Koch, Fabrice Roux, Joachim Hermisson, Angela M. Hancock, **Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages**. *Nat Commun* **13**, 1461 (2022). https://doi.org/10.1038/s41467-022-28800-z

Author contributions: Conceptualization: J.H., A.M.H.; Methodology: A.F., C.N., A.M.H; Software: A.F.; Investigation, validation, and data curation: C.N., A.F., **A.F.E.**, E.T., M.G., N.W., N.D., A.M.H.; Formal analysis: S.R., A.F., N.W., J.H., S.A., **A.F.E.**, E.T., A.M.H.; Resources: H.D., C.N., E.T., P.J.F., **A.F.E.**, A.F., F.R., A.M.H.; Writing-first draft: C.N., A.F., A.M.H.; Writing-reviewing and editing: all authors; Project administration: A.M.H.; Supervision and funding acquisition: M.K., F.R., J.H., and A.M.H.

2- Emmanuel Tergemina, **Ahmed F. Elfarargi**, Paulina Flis, Andrea Fulgione, Mehmet Göktay, Célia Neto, Marleen Scholle, Pádraic J. Flood, Sophie-Asako Xerri, Johan Zicola, Nina Döring, Herculano Dinis, Ute Krämer, David E. Salt, Angela M. Hancock, **A two-step adaptive walk rewires nutrient transport in a challenging edaphic environment**. *Science Advances* **20**, (2022) https://doi.org/10.1126/sciadv.abm9385

Author contributions: E.T. and A.M.H. conceived and designed the study. E.T. and N.D. performed the laboratory and greenhouse work. E.T., A.M.H., H.D., A.F., C.N., and **A.F.E.** collected samples. P.F. and D.E.S. conducted ionomics analysis of soil and plant material and contributed to its interpretation. P.J.F. generated the F2 population derived from Col-0 and F13-8. E.T. and J.Z. contributed to the image analysis. M.G. assembled de novo genomes. M.S. and E.T. performed experiments on plates; M.S. and U.K. contributed to its interpretation. E.T., **A.F.E.**, A.M.H., A.F., and S.-A.X. contributed to analyses and interpretation of results. E.T. and A.M.H. wrote the paper. All authors read and edited the manuscript.

Table of contents

Chapter 1
General introduction

Drought adaptation in a changing climate and insights from the Cape Verdean *Arabidopsis thaliana* lineage

Ahmed F. Elfarargi and Angela M. Hancock

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

1.1 Abstract

Ongoing climate change is predicted to impact precipitation timing and quantity, leading to reduced growing season lengths and increased susceptibility to drought in many regions worldwide. These changes are expected to increase extinction risk to plant populations and desertification of ecosystems. Understanding the mechanisms enabling species to adapt to these changes is crucial for effective conservation strategies. Within and between species, plants follow multiple alternative strategies to cope with drought, including avoidance (reducing water loss), escape (rapid growth and transition to flowering to avoid the dry period), and drought tolerance (maintaining physiological and metabolic functions under water-limited conditions). Depending on the genetic and environmental context, these strategies can be canalized or plastic. Understanding the balance between canalization and plasticity is essential for predicting plant responses to future climate change and developing resilient crop varieties. Genetic adaptations are crucial for populations facing extreme drought conditions, enabling species to adapt and persist in rapidly changing environments. Insufficient genetic variation within populations may lead to a high risk of extinction if they cannot adapt to new habitats. Studying *A. thaliana* populations on the Cape Verde Islands (CVI), which have undergone rapid adaptation and evolutionary rescue in response to the drought-prone climate, offers valuable insights into adaptive divergence.

Keywords: climate change; evolution; adaptation; drought; *Arabidopsis thaliana*; natural variation.

1.2 Global climate change and distributional shifts

Global climate change is anticipated to have a significant impact on the geographic distribution of species worldwide, resulting in a sharp increase in the risk of extinction for up to 48% of all species (Williams *et al.*, 2008; Parmesan and Yohe, 2003; IPCC, 2022). Although the effects of climate change are expected to be local, certain patterns of vulnerability are widespread. For example, ecosystems in extreme climates, particularly arid regions, are at exceptionally high risk (Walther *et al.*, 2002; Diffenbaugh and Field, 2013), although the species that inhabit these areas may succeed after migration. Overall, range-edge and, in particular, island species are expected to be at the greatest risk of extinction (Channell and Lomolino, 2000; Foden *et al.*, 2013).

Predicting the extent to which species will be affected by climate change is complicated by the potential for ecological (e.g., distributional shifts and phenotypic plasticity) or evolutionary (e.g., genetic adaptation and gene flow) mechanisms to mitigate the consequences of climate change (Davis and Shaw, 2001; Aitken *et al.*, 2008). Elevation in extinction rates has been associated with drastic ecological perturbations and climate disruptions (Pimm et al., 2014). Thus, the capacity of species to respond to environmental stressors through adaptation or acclimation will determine their ability to cope with climate change.

Species and their populations, when strongly affected by changes in habitat suitability, such as rapid changes in temperature or precipitation, typically respond by either shifting their distribution range or adjusting their phenology to persist in a new suitable habitat (Parmesan, 2006). These rapid changes in temperature and water availability can reduce fitness in the natural populations, leading to decreased species abundance in the affected areas. The differential changes in species abundance may sometimes result in range shifts, contractions, or expansions (Lenoir and Svenning, 2015; Fei *et al.*, 2017). Many species will need to shift their ranges to survive, either by moving to more suitable habitats or adapting to the new conditions. However, this may not be possible for species with limited mobility, such as plants and slow-moving animals, leading to population declines or even

extinction. In response to a rapidly-changing climate, range shifts to suitable ecological niches have been observed in plants (Sturm *et al.*, 2001; Parmesan and Yohe, 2003; Hickling *et al.*, 2006; Beckage *et al.*, 2008; Morueta-Holme *et al.*, 2015; Fei *et al.*, 2017).

Overall, climate change is a major challenge to global biodiversity, with significant implications for ecosystem functions and services. Although the impacts of climate change are local, it is clear that extreme climatic conditions such as drought and heat pose significant risks to the survival of species (Thuiller *et al.*, 2005; Allen *et al.*, 2010). Species that inhabit isolated locations, including range-edges and islands, are particularly susceptible to extinction (Frankham, 1997; Caujapé-Castells *et al.*, 2010). Therefore, understanding the mechanisms that enable such species to respond and adapt to climate change is critical for developing effective conservation strategies to mitigate climate change's adverse effects on biodiversity.

1.3 Impacts of climate change on plants: Drought as a potent selective pressure

One of the most potent climate change-related selective pressures is the increase in drought conditions, which is anticipated to exacerbate water scarcity in many areas and lead to the loss of vegetation and habitat fragmentation. Drought substantially impacts the distribution of plant populations worldwide, mainly affecting their growth, fitness, and reproduction (Stebbins Jr, 1952; Bray, 1997; Gurevitch *et al.*, 2002). Over the last few decades, climatic data have shown an increase in the frequency, intensity, and duration of drought events worldwide. Climate change models have predicted future impacts of extreme drought events (Perkins *et al.*, 2012; Shukla *et al.*, 2019). To mitigate the impacts of drought-related distributional shifts, conservation efforts must focus on preserving and restoring habitat connectivity, as well as identifying areas that are likely to remain suitable for key species in the face of changing climate conditions. Given the significant challenges posed by drought conditions, it is crucial to understand the mechanisms that allow plants to persist under such stressful environments. In the following sections, we will explore the

diverse strategies and molecular pathways that plants employ to cope with drought, shedding light on the complex interplay between genetic and environmental factors underpinning plant adaptation and resilience.

1.4 Mechanisms that allow plants to persist under drought conditions

Given the increasing frequency, intensity, and duration of drought events worldwide and the predicted future impacts of extreme drought events, plants face a significant challenge to survive and thrive under these changing conditions. However, plants have evolved adaptive mechanisms to cope with water scarcity and maintain their growth, fitness, and reproduction. These mechanisms include drought escape, avoidance, and tolerance strategies that rely on genetic, physiological, and morphological changes (Bray, 1997). These strategies are not mutually exclusive and may be observed in combination. The first strategy is drought escape, where plants flower early to avoid periods of water deficit during major growth and reproductive phases. The second strategy is drought avoidance, where plants maintain their water status by increasing water uptake and reducing water loss via transpiration by closing stomata. This results in a decrease in CO_2 availability in leaves and an increase in the ^{13}C isotope, which can be used as a measure of water use efficiency (WUE) (Tuberosa, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). Another measure of WUE is transpiration efficiency (TE), which is higher in C4 photosynthesis species than in C3 photosynthesis species (Tardieu, 2013; Assmann, 2013). Drought tolerance by osmotic adjustment (OA) is another strategy by which plants protect their cellular function under water stress by increasing intercellular solutes, such as proline, glycine betaine, and sugar (Verslues and Juenger, 2011; Tuberosa, 2012). The plant hormone abscisic acid (ABA) plays a crucial role in regulating OA and other drought tolerance mechanisms, such as stomatal closure and water uptake (Tuberosa, 2012). ABA also regulates the accumulation of proline and antioxidants that scavenge Reactive Oxygen Species (ROS) in plants experiencing abiotic stress (Sharma *et al.*, 2012; Cruz de Carvalho, 2008; Bhargava and Sawant, 2013).

These drought strategies can be either canalized or plastic in nature, depending on the specific plant species and their genetic and environmental context (Sultan, 2000; Nicotra *et al.*, 2010). Canalized strategies are more rigid and genetically determined, while plastic strategies allow for greater flexibility in response to environmental fluctuations (Waddington, 1942; Via *et al.*, 1995). Both approaches have advantages, with canalized strategies providing consistent responses to predictable stressors (Bradshaw, 1965) and plastic strategies offering adaptability to more variable conditions (Miner *et al.*, 2005). Understanding the balance between canalization and plasticity in drought strategies is essential for predicting plant responses to future climate change (Aitken *et al.*, 2008; Chevin *et al.*, 2010) and developing more resilient crop varieties (Tardieu, 2012; Ray *et al.*, 2013).

1.5 Plastic responses may become canalized over evolutionary time

Phenotypic plasticity is commonly the primary response mechanism to environmental changes, allowing species to cope with the rapidly-changing climate (Matesanz *et al.*, 2010; Merilä and Hendry, 2014). Furthermore, phenotypic plasticity has been found to have a critical role in facilitating the success of invasive species in a new environment expanding their range shifts (Geng *et al.*, 2007; Davidson *et al.*, 2011; Lande, 2015). In the long term, plastic responses can become canalized over evolutionary time as populations experience stabilizing selection in a consistent environment (Waddington, 1953). This process, referred to as genetic assimilation, involves the reduction of phenotypic plasticity and the fixation of formerly plastic traits as genetically determined (Waddington, 1953; Pigliucci *et al.*, 2006). Genetic assimilation can be advantageous when the environment remains stable, as canalized traits are less affected by environmental fluctuations and may increase fitness (Price *et al.*, 2003; Crispo, 2007). However, genetic assimilation can also lead to the loss of adaptive plasticity, potentially reducing the ability of populations to respond to future environmental changes (Lande, 2009; Chevin *et al.*, 2010). Thus, the balance between plasticity and canalization is critical for species adaptation and

persistence in changing environments. Plants, as sessile organisms, can adjust their physiology, morphology, and reproduction to confront environmental changes through phenotypic plasticity (Zhang *et al.*, 2014; Bercovich *et al.*, 2019). However, whether or not phenotypic plasticity is adaptive is still unclear. There are contrasting types of plasticity, adaptive and non-adaptive (Ghalambor *et al.*, 2007). According to the fitness optimum, phenotypic plasticity could be adaptive, maladaptive, or neutral for each individual in a population. The distinctive way of individual phenotype differentiation across multiple environments can be known as a reaction norm. Accordingly, phenotypic plasticity can lead to evolutionary changes through genetic accommodation, which increases plasticity for genotypes that facilitate fitness optimization, or genetic assimilation, where the stability of the new environment results in the loss of plasticity. Alternatively, genetic compensation may occur if plasticity is maladaptive in the new environment, leading to a reduction in plasticity to resist phenotypic change (Kelly, 2019). Understanding the relationship between phenotypic plasticity, canalization, and evolutionary adaptation is essential for predicting species responses to climate change and developing effective conservation strategies.

1.6 Genetic adaptation and evolutionary rescue in populations facing extreme drought

Genetic adaptation and evolutionary rescue are crucial mechanisms for populations facing extreme drought conditions, enabling species to adapt and persist in rapidly changing environments. Insufficient genetic variation within populations can lead to a high risk of extinction if they cannot adapt to new habitats, reducing fitness and survival rates (Frankham, 2005). Evolutionary rescue, while demonstrating the potential for populations to adapt to novel environments, may only effectively prevent extinction for mildly affected populations with high natural densities (Burger and Lynch, 1995; Orr and Unckless, 2014). A study on the model species *A. thaliana* has shown that the distance between a species' historically native range and experimental gardens can cause lagging in

adaptive evolution, resulting in local maladaptation that impacts fitness and persistence (Wilczek *et al.*, 2014).

Phenotypic plasticity may increase the probability of evolutionary rescue through increased persistence times (Chevin and Lande, 2010). This process is essential for maintaining genetic variation and adaptive potential under strong selective pressures, such as those imposed by extreme drought conditions (Jump and Peñuelas, 2005). Moreover, studies have shown that evolutionary rescue can be promoted through gene flow, which provides new genetic variation for adaptation in maladapted populations (Aitken and Whitlock, 2013; Chevin and Hoffmann, 2017). This influx of novel genetic variants can facilitate adaptation to local environmental conditions by introducing adaptive alleles that were not previously present, ultimately increasing the likelihood of evolutionary rescue (Garant *et al.*, 2007; Rieseberg and Burke, 2001). In conclusion, genetic adaptation and evolutionary rescue are critical mechanisms for populations facing extreme drought conditions.

1.7 Genetic load and adaptive dynamics

The spatial distribution of species' genetic diversity is shaped by past range expansions and contractions, resulting in genetic drift that can impact adaptive dynamics (Hewitt, 2000; Excoffier *et al.*, 2009). Range-edge populations, in particular, are expected to experience genetic drift, leading to increased accumulation rates of deleterious mutations, known as genetic load. This genetic load can be driven by either reduction in effective population size in non-suitable habitats or multiple historical bottlenecks during range expansion and can negatively influence the fitness and persistence of individuals in these populations (Lynch *et al.*, 1995; Excoffier and Ray, 2008; Excoffier *et al.*, 2009; Peischl *et al.*, 2013). However, locally adapted populations that expand from their species' native environment to a novel range edge with a harsh climate may often have adapted to their new harsher environments, resulting in new adaptive signals in the genome (Colautti and Barrett, 2013; Savolainen *et al.*, 2013).

1.8 *Arabidopsis thaliana* populations as a model for drought response in widespread and endemic species

Arabidopsis thaliana is a plant species that has been extensively studied and is considered a model organism for genetics and molecular biology (Meinke *et al.*, 1998). As an annual species, it has adapted to open, arid habitats susceptible to seasonal droughts (Ruppert *et al.*, 2015) and exhibits reproductive success directly linked to interannual environmental fluctuations (Segrestin *et al.*, 2018). *A. thaliana* exhibits many traits of a "perfect weed" and shares traits with invasive species, such as being an annual herb, self-fertilizing, and highly reproductive (Weigel and Mott, 2009). Its extensive distribution encompasses Europe, the Balkans, Central Asia, China, parts of Africa, North America, and Argentina (Hoffmann, 2002; Hoffmann, 2005; He *et al.*, 2007; Stock *et al.*, 2015; Alonso-Blanco *et al.*, 2016; Durvasula *et al.*, 2017; Kasulin *et al.*, 2017; Exposito-Alonso, Becker, *et al.*, 2018). Local adaptation has been observed throughout *A. thaliana's* range (Fournier-Level *et al.*, 2011; Hancock *et al.*, 2011; Ågren and Schemske, 2012; Savolainen *et al.*, 2013; Lee *et al.*, 2017). Field studies and climate parameter correlations have identified numerous genomic regions linked to local climate conditions, with several SNPs contributing to optimizing plant performance under local environmental conditions (Hancock *et al.*, 2011; Fournier-Level *et al.*, 2011; Lasky *et al.*, 2012). Due to its wide distribution, *A. thaliana* has been used to study eco-evolutionary dynamics in response to drought and rising temperatures (McKay *et al.*, 2003; Hausmann *et al.*, 2005; McKay *et al.*, 2008; Verslues and Juenger, 2011; Juenger, 2013; Kenney *et al.*, 2014; Clauw *et al.*, 2015; Kalladan *et al.*, 2017; Monroe *et al.*, 2018; Exposito-Alonso, Vasseur, *et al.*, 2018; Exposito-Alonso *et al.*, 2019; Marchadier *et al.*, 2019; Van Dooren *et al.*, 2020). This has enabled researchers to identify which *A. thaliana* populations are at evolutionary risk and likely to experience more negative selection in the future, influencing their fitness and persistence (Exposito-Alonso *et al.*, 2019).

An intriguing case of incipient endemic island species is found in *A. thaliana* populations on the Cape Verde Islands (CVI). Our recent research indicates that these populations, which colonized the islands from North Africa approximately 5-7 kya, have undergone

rapid adaptation and evolutionary rescue in response to the drought-prone climate (Fulgione *et al.*, 2022). This adaptation aligns with the adaptive walk hypothesis (Gillespie, 1983; Gillespie, 1984; Orr, 1998), as the initial colonists relied on de novo mutations that occurred in parallel at *FRI* and *FLC* genes. These mutations facilitated adaptation to the extreme habitat, which differed significantly from their previous optimal environmental conditions in North Africa (Fulgione *et al.*, 2022). The Cape Verdean *A. thaliana* populations exhibit distinct genetic and phenotypic divergence from mainland populations, likely driven by unique environmental conditions and geographic isolation (Fulgione *et al.*, 2022). These findings highlight the importance of genetic adaptation and evolutionary rescue for populations facing extreme drought conditions. Studying these island populations can offer valuable insights into adaptive divergence processes and the formation of new endemic species in response to changing environmental conditions. To fully understand and capture these adaptive processes, high-throughput phenotyping platforms are essential. These platforms enable the automatic and frequent collection of traits for a large number of plants, providing precision and consistency in applying drought stress conditions, which are critical for investigating the complex mechanisms underlying adaptation and resilience.

1.9 Phenoscope: A platform for precision drought stress phenotyping

Phenotyping poses a significant challenge for quantitative genetic studies because traits must be measured accurately for large numbers of plants. In the context of drought stress studies, it is challenging to consistently apply drought stress conditions across plants in large-scale phenotyping experiments because of spatial heterogeneity in drying rates. High-throughput phenotyping platforms using imaging techniques to measure a battery of traits related to plant performance can be particularly beneficial in this situation, as they allow for the automatic and frequent collection of traits for a large number of plants. The Phenoscope is a platform that automates the circulation of pots and adjusts watering multiple times daily based on pot weight, making experiments possible that would be

impractical with manual methods (Tisné *et al.*, 2013). In this thesis, we use the Phenoscope platform, which offers significant advantages, particularly its precision and consistency in applying drought stress conditions, reducing the impact of spatial heterogeneity in drying rates. Moreover, its automated and high-throughput capabilities streamline the phenotyping process, enabling us to collect comprehensive, accurate, and reliable data that are crucial for understanding the genetic basis of complex traits and enhancing our ability to predict plant responses to drought stress. Overall, these non-destructive imaging techniques allow for the efficient collection of phenotyping data for many plants, proving invaluable for GWA mapping of dynamic traits.

1.10 Objectives and scope of this thesis

In this thesis, we investigated the genetic basis of drought adaptation in the Cape Verdean *Arabidopsis thaliana* lineage. Specifically, the thesis explores how *A. thaliana* populations adapt to changes in precipitation regimes and extreme environmental conditions, such as prolonged dry periods and high humidity, and how this adaptation is encoded in the plant's genetic makeup.

In **Chapter II**, we investigate the evolution of stomatal conductance and water use efficiency (WUE) in the *A. thaliana* population that colonized an island with a montane cloud scrubland ecosystem characterized by seasonal drought and fog-based precipitation. We showed that the colonizing population had evolved an increased stomatal conductance and decreased WUE relative to its closest outgroup population from temperate North Africa. We also identified a nonsynonymous SNP in *MAP KINASE 12* (*MPK12* G53R) that explains a substantial portion of the phenotypic variance in WUE in the island population and provides evidence of positive selection for this allele in harsher regions of the island.

In **Chapter III**, we explore the genetic basis and adaptive evolution of complex traits such as growth rate, leaf color, and stomatal patterning in response to drought in the CVI *A. thaliana* populations. These plants experience extended dry periods and highly variable

rainfall, presenting an opportunity to investigate their adaptation to such challenging conditions. Through genome-wide association mapping analyses, we identified moderately complex genetic architectures with roles for several major effect variants that underlie variation in these traits. Notably, we identified genes with potential pleiotropic functions for complex traits associated with drought stress, emphasizing the intricate nature of plant adaptation in response to drought, but further functional characterization is needed.

In **Chapter IV**, we provide a general discussion of the previous chapters and highlight the major findings of the thesis. The chapter also discusses the implications of the work for understanding plant adaptation to climate change and for predicting future species distributions and adaptive responses to environmental pressures.

1.1   References

Ågren, 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.

Aitken, S.N. and Whitlock, M.C. (2013) Assisted gene flow to facilitate local adaptation to climate change. *Annu. Rev. Ecol. Evol. Syst.*, **44**, 367–388.

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.

Allen, C.D., Macalady, A.K., Chenchouni, H., et al. (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manag.*, **259**, 660–684.

Alonso-Blanco, C., Andrade, J., Becker, C., et al. (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. *Cell*, **166**, 481–491.

Assmann, S.M. (2013) Natural variation in abiotic stress and climate change responses in Arabidopsis: implications for twenty-first-century agriculture. *Int. J. Plant Sci.*, **174**, 3–26.

Beckage, B., Osborne, B., Gavin, D.G., Pucko, C., Siccama, T. and Perkins, T. (2008) A rapid upward shift of a forest ecotone during 40 years of warming in the Green Mountains of Vermont. *Proc. Natl. Acad. Sci.*, **105**, 4197–4202.

Bercovich, M.V., Schubert, N., Saá, A.C.A., Silva, J. and Horta, P.A. (2019) Multi-level phenotypic plasticity and the persistence of seagrasses along environmental gradients in a subtropical lagoon. *Aquat. Bot.*, **157**, 24–32.

Bhargava, S. and Sawant, K. (2013) Drought stress adaptation: metabolic adjustment and regulation of gene expression. *Plant Breed.*, **132**, 21–32.

Bradshaw, A.D. (1965) Evolutionary significance of phenotypic plasticity in plants. *Adv. Genet.*, **13**, 115–155.

Bray, E.A. (1997) Plant responses to water deficit. *Trends Plant Sci.*, **2**, 48–54.

Burger, R. and Lynch, M. (1995) Evolution and extinction in a changing environment: A quantitative-genetic analysis. *Evol. Int. J. Org. Evol.*, **49**, 151–163.

Caujapé-Castells, J., Tye, A., Crawford, D.J., et al. (2010) Conservation of oceanic island floras: present and future global challenges. *Perspect. Plant Ecol. Evol. Syst.*, **12**, 107–129.

Channell, R. and Lomolino, M.V. (2000) Dynamic biogeography and conservation of endangered species. *Nature*, **403**, 84–86.

Chevin, L. and Lande, R. (2010) When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? *Evol. Int. J. Org. Evol.*, **64**, 1143–1150.

Chevin, L.-M. and Hoffmann, A.A. (2017) Evolution of phenotypic plasticity in extreme environments. *Philos. Trans. R. Soc. B Biol. Sci.*, **372**, 20160138.

Chevin, L.-M., Lande, R. and Mace, G.M. (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.*, **8**, e1000357.

Clauw, P., Coppens, F., De Beuf, K., et al. (2015) Leaf responses to mild drought stress in natural variants of Arabidopsis. *Plant Physiol.*, **167**, 800–816.

Colautti, R.I. and Barrett, S.C. (2013) Rapid adaptation to climate facilitates range expansion of an invasive plant. *science*, **342**, 364–366.

Crispo, E. (2007) The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, **61**, 2469–2479.

Cruz de Carvalho, M.H. (2008) Drought stress and reactive oxygen species: production, scavenging and signaling. *Plant Signal. Behav.*, **3**, 156–165.

Davidson, A.M., Jennions, M. and Nicotra, A.B. (2011) Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecol. Lett.*, **14**, 419–431.

Davis, M.B. and Shaw, R.G. (2001) Range Shifts and Adaptive Responses to Quaternary Climate Change. *Science*. Available at: https://www.science.org/doi/abs/10.1126/science.292.5517.673.

Diffenbaugh, N.S. and Field, C.B. (2013) Changes in Ecologically Critical Terrestrial Climate Conditions. *Science*, **341**, 486–492.

Durvasula, A., Fulgione, A., Gutaker, R.M., et al. (2017) African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **114**, 5213–5218.

Excoffier, L., Foll, M. and Petit, R.J. (2009) Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.*, **40**, 481–501.

Excoffier, L. and Ray, N. (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.*, **23**, 347–351.

Exposito-Alonso, M., Becker, C., Schuenemann, V.J., et al. (2018) The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.*, **14**, e1007155.

Exposito-Alonso, M., Burbano, H.A., Bossdorf, O., Nielsen, R. and Weigel, D. (2019) Natural selection on the Arabidopsis thaliana genome in present and future climates. *Nature*, **573**, 126–129.

Exposito-Alonso, M., Vasseur, F., Ding, W., Wang, G., Burbano, H.A. and Weigel, D. (2018) Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nat. Ecol. Evol.*, **2**, 352–358.

Fei, S., Desprez, J.M., Potter, K.M., Jo, I., Knott, J.A. and Oswalt, C.M. (2017) Divergence of species responses to climate change. *Sci. Adv.*, **3**, e1603055.

Foden, W.B., Butchart, S.H.M., Stuart, S.N., et al. (2013) Identifying the World's Most Climate Change Vulnerable Species: A Systematic Trait-Based Assessment of all Birds, Amphibians and Corals. *PLOS ONE*, **8**, e65427.

Fournier-Level, A., Korte, A., Cooper, M.D., Nordborg, M., Schmitt, J. and Wilczek, A.M. (2011) A map of local adaptation in Arabidopsis thaliana. *Science*, **334**, 86–89.

Frankham, R. (1997) Do island populations have less genetic variation than mainland populations? *Heredity*, **78**, 311–327.

Frankham, R. (2005) Genetics and extinction. *Biol. Conserv.*, **126**, 131–140.

Fulgione, A., Neto, C., Elfarargi, A.F., et al. (2022) Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. *Nat. Commun.*, **13**, 1461.

Garant, D., Forde, S.E. and Hendry, A.P. (2007) The Multifarious Effects of Dispersal and Gene Flow on Contemporary Adaptation. *Funct. Ecol.*, **21**, 434–443.

Geng, Y.-P., Pan, X.-Y., Xu, C.-Y., Zhang, W.-J., Li, B., Chen, J.-K., Lu, B.-R. and Song, Z.-P. (2007) Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biol. Invasions*, **9**, 245–256.

Ghalambor, C.K., McKAY, J.K., Carroll, S.P. and Reznick, D.N. (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.*, **21**, 394–407.

Gillespie, J.H. (1984) Molecular evolution over the mutational landscape. *Evolution*, 1116–1129.

Gillespie, J.H. (1983) Some properties of finite populations experiencing strong selection and weak mutation. *Am. Nat.*, **121**, 691–708.

Gurevitch, J., Scheiner, S.M. and Fox, G.A. (2002) *The ecology of plants*, Sinauer Associates Sunderland.

Hancock, A.M., Brachi, B., Faure, N., Horton, M.W., Jarymowycz, L.B., Sperone, F.G., Toomajian, C., Roux, F. and Bergelson, J. (2011) Adaptation to climate across the Arabidopsis thaliana genome. *Science*, **334**, 83–86.

Hausmann, N.J., Juenger, T.E., Sen, S., Stowe, K.A., Dawson, T.E. and Simms, E.L. (2005) Quantitative trait loci affecting δ13C and response to differential water availibility in Arabidopsis thallana. *Evolution*, **59**, 81–96.

He, F., Kang, D., Ren, Y., Qu, L., Zhen, Y. and Gu, H. (2007) Genetic diversity of the natural populations of Arabidopsis thaliana in China. *Heredity*, **99**, 423–431.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hickling, R., Roy, D.B., Hill, J.K., Fox, R. and Thomas, C.D. (2006) The distributions of a wide range of taxonomic groups are expanding polewards. *Glob. Change Biol.*, **12**, 450–455.

Hoffmann, M.H. (2002) Biogeography of Arabidopsis thaliana (l.) heynh.(Brassicaceae). *J. Biogeogr.*, **29**, 125–134.

Hoffmann, M.H. (2005) Evolution of the realized climatic niche in the genus: Arabidopsis (Brassicaceae). *Evolution*, **59**, 1425–1436.

IPCC. (2022) Climate Change 2022: Impacts, Adaptation and Vulnerability. In Working Group II Contribution to the IPCC Sixth Assessment Report. IPCC.

Juenger, T.E. (2013) Natural variation and genetic constraints on drought tolerance. *Curr. Opin. Plant Biol.*, **16**, 274–281.

Jump, A.S. and Peñuelas, J. (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol. Lett.*, **8**, 1010–1020.

Kalladan, R., Lasky, J.R., Chang, T.Z., Sharma, S., Juenger, T.E. and Verslues, P.E. (2017) Natural variation identifies genes affecting drought-induced abscisic acid accumulation in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **114**, 11536–11541.

Kasulin, L., Rowan, B.A., León, R.J., Schuenemann, V.J., Weigel, D. and Botto, J.F. (2017) A single haplotype hyposensitive to light and requiring strong vernalization dominates Arabidopsis thaliana populations in Patagonia, Argentina. *Mol. Ecol.*, **26**, 3389–3404.

Kelly, M. (2019) Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philos. Trans. R. Soc. B*, **374**, 20180176.

Kenney, A.M., McKay, J.K., Richards, J.H. and Juenger, T.E. (2014) Direct and indirect selection on flowering time, water-use efficiency (WUE, δ13C), and WUE plasticity to drought in Arabidopsis thaliana. *Ecol. Evol.*, **4**, 4505–4521.

Kimura, M. (1955) Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. U. S. A.*, **41**, 144.

Lande, R. (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.*, **22**, 1435–1446.

Lande, R. (2015) Evolution of phenotypic plasticity in colonizing species. *Mol. Ecol.*, **24**, 2038–2045.

Lasky, J.R., Des Marais, D.L., McKAY, J.K., Richards, J.H., Juenger, T.E. and Keitt, T.H. (2012) Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol. Ecol.*, **21**, 5512–5529.

Lee, C.-R., Svardal, H., Farlow, A., Exposito-Alonso, M., Ding, W., Novikova, P., Alonso-Blanco, C., Weigel, D. and Nordborg, M. (2017) On the post-glacial spread of human commensal Arabidopsis thaliana. *Nat. Commun.*, **8**, 14458.

Lenoir, J. and Svenning, J. (2015) Climate-related range shifts–a global multidimensional synthesis and new research directions. *Ecography*, **38**, 15–28.

Lynch, M., Conery, J. and Burger, R. (1995) Mutation accumulation and the extinction of small populations. *Am. Nat.*, **146**, 489–518.

Marchadier, E., Hanemian, M., Tisné, S., Bach, L., Bazakos, C., Gilbault, E., Haddadi, P., Virlouvet, L. and Loudet, O. (2019) The complex genetic architecture of shoot growth natural variation in Arabidopsis thaliana. *PLOS Genet.*, **15**, e1007954.

Matesanz, S., Gianoli, E. and Valladares, F. (2010) Global change and the evolution of phenotypic plasticity in plants. *Ann. N. Y. Acad. Sci.*, **1206**, 35–55.

McKay, J.K., Richards, J.H. and Mitchell-Olds, T. (2003) Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.*, **12**, 1137–1151.

McKay, J.K., Richards, J.H., Nemali, K.S., Sen, S., Mitchell-Olds, T., Boles, S., Stahl, E.A., Wayne, T. and Juenger, T.E. (2008) Genetics of Drought Adaptation in Arabidopsis Thaliana Ii. Qtl Analysis of a New Mapping Population, Kas-1 × Tsu-1. *Evolution*, **62**, 3014–3026.

Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D. and Koornneef, M. (1998) Arabidopsis thaliana: a model plant for genome analysis. *Science*, **282**, 662–682.

Merilä, J. and Hendry, A.P. (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.*, **7**, 1–14.

Miner, B.G., Sultan, S.E., Morgan, S.G., Padilla, D.K. and Relyea, R.A. (2005) Ecological consequences of phenotypic plasticity. *Trends Ecol. Evol.*, **20**, 685–692.

Monroe, J.G., Powell, T., Price, N., Mullen, J.L., Howard, A., Evans, K., Lovell, J.T. and McKay, J.K. (2018) Drought adaptation in Arabidopsis thaliana by extensive genetic loss-of-function D. J. Kliebenstein, C. S. Hardtke, and A. Korte, eds. *eLife*, **7**, e41038.

Morueta-Holme, N., Engemann, K., Sandoval-Acuña, P., Jonas, J.D., Segnitz, R.M. and Svenning, J.-C. (2015) Strong upslope shifts in Chimborazo's vegetation over two centuries since Humboldt. *Proc. Natl. Acad. Sci.*, **112**, 12741–12745.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., et al. (2010) Plant phenotypic plasticity in a changing climate. *Trends Plant Sci.*, **15**, 684–692.

Orr, H.A. (1998) The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. *Evolution*, **52**, 935–949.

Orr, H.A. and Unckless, R.L. (2014) The Population Genetics of Evolutionary Rescue. *PLOS Genet.*, **10**, e1004551.

Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annu Rev Ecol Evol Syst*, **37**, 637–669.

Parmesan, C. and Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Peischl, S., Dupanloup, I., Kirkpatrick, M. and Excoffier, L. (2013) On the accumulation of deleterious mutations during range expansions. *Mol. Ecol.*, **22**, 5972–5982.

Perkins, S., Alexander, L. and Nairn, J. (2012) Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophys. Res. Lett.*, **39**.

Pigliucci, M., Murren, C.J. and Schlichting, C.D. (2006) Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.*, **209**, 2362–2367.

Price, T.D., Qvarnström, A. and Irwin, D.E. (2003) The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. Lond. B Biol. Sci.*, **270**, 1433–1440.

Ray, D.K., Mueller, N.D., West, P.C. and Foley, J.A. (2013) Yield trends are insufficient to double global crop production by 2050. *PloS One*, **8**, e66428.

Rieseberg, L.H. and Burke, J.M. (2001) The biological reality of species: gene flow, selection, and collective evolution. *Taxon*, **50**, 47–67.

Ruppert, J.C., Harmoney, K., Henkin, Z., Snyman, H.A., Sternberg, M., Willms, W. and Linstädter, A. (2015) Quantifying drylands' drought resistance and recovery: the importance of drought intensity, dominant life history and grazing regime. *Glob. Change Biol.*, **21**, 1258–1270.

Savolainen, O., Lascoux, M. and Merilä, J. (2013) Ecological genomics of local adaptation. *Nat. Rev. Genet.*, **14**, 807–820.

Segrestin, J., Bernard-Verdier, M., Violle, C., Richarte, J., Navas, M.-L. and Garnier, E. (2018) When is the best time to flower and disperse? A comparative analysis of plant reproductive phenology in the Mediterranean. *Funct. Ecol.*, **32**, 1770–1783.

Sharma, P., Jha, A.B., Dubey, R.S. and Pessarakli, M. (2012) Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J. Bot.*, **2012**.

Shukla, P.R., Skeg, J., Buendia, E.C., et al. (2019) Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.

Stebbins Jr, G.L. (1952) Aridity as a stimulus to plant evolution. *Am. Nat.*, **86**, 33–44.

Stock, A.J., McGoey, B.V. and Stinchcombe, J.R. (2015) Water availability as an agent of selection in introduced populations of Arabidopsis thaliana: impacts on flowering time evolution. *PeerJ*, **3**, e898.

Sturm, M., Racine, C. and Tape, K. (2001) Increasing shrub abundance in the Arctic. *Nature*, **411**, 546–547.

Sultan, S.E. (2000) Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.*, **5**, 537–542.

Tardieu, F. (2012) Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *J. Exp. Bot.*, **63**, 25–31.

Tardieu, F. (2013) Plant response to environmental conditions: assessing potential production, water demand, and negative effects of water deficit. *Front. Physiol.*, **4**, 17.

Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. and Prentice, I.C. (2005) Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci.*, **102**, 8245–8250.

Tisné, S., Serrand, Y., Bach, L., et al. (2013) Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J.*, **74**, 534–544.

Tuberosa, R. (2012) Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.*, **3**, 347.

Van Dooren, T.J.M., Silveira, A.B., Gilbault, E., et al. (2020) Mild drought in the vegetative stage induces phenotypic, gene expression, and DNA methylation plasticity in Arabidopsis but no transgenerational effects. *J. Exp. Bot.*, **71**, 3588–3602.

Verslues, P.E. and Juenger, T.E. (2011) Drought, metabolites, and Arabidopsis natural variation: a promising combination for understanding adaptation to water-limited environments. *Curr. Opin. Plant Biol.*, **14**, 240–245.

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S.M., Schlichting, C.D. and Van Tienderen, P.H. (1995) Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.*, **10**, 212–217.

Waddington, C.H. (1942) Canalization of development and the inheritance of acquired characters. *Nature*, **150**, 563–565.

Waddington, C.H. (1953) Genetic assimilation of an acquired character. *Evolution*, 118–126.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O. and Bairlein, F. (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.

Weigel, D. and Mott, R. (2009) The 1001 genomes project for Arabidopsis thaliana. *Genome Biol.*, **10**, 1–5.

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.*, **111**, 7906–7913.

Williams, S.E., Shoo, L.P., Isaac, J.L., Hoffmann, A.A. and Langham, G. (2008) Towards an Integrated Framework for Assessing the Vulnerability of Species to Climate Change. *PLOS Biol.*, **6**, e325.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97.

Zhang, J., Huang, X. and Jiang, Z. (2014) Physiological responses of the seagrass Thalassia hemprichii (Ehrenb.) Aschers as indicators of nutrient loading. *Mar. Pollut. Bull.*, **83**, 508–515.

<div align="right"># Chapter 2</div>

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

2.1   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*).

2.2   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 (Karger *et al.*, 2021; Stadtmüller, 1987; Bruijnzeel *et al.*, 2011; Walter, 1985; Weathers, 1999; Dawson, 1998; Kerfoot, 1968). 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 (Grant, 1999; Losos *et al.*, 1997). *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 (Fulgione *et al.*, 2022; Brochmann *et al.*, 1997). 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 (Figure 4 and Supplementary Figure 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.

2.3   Results
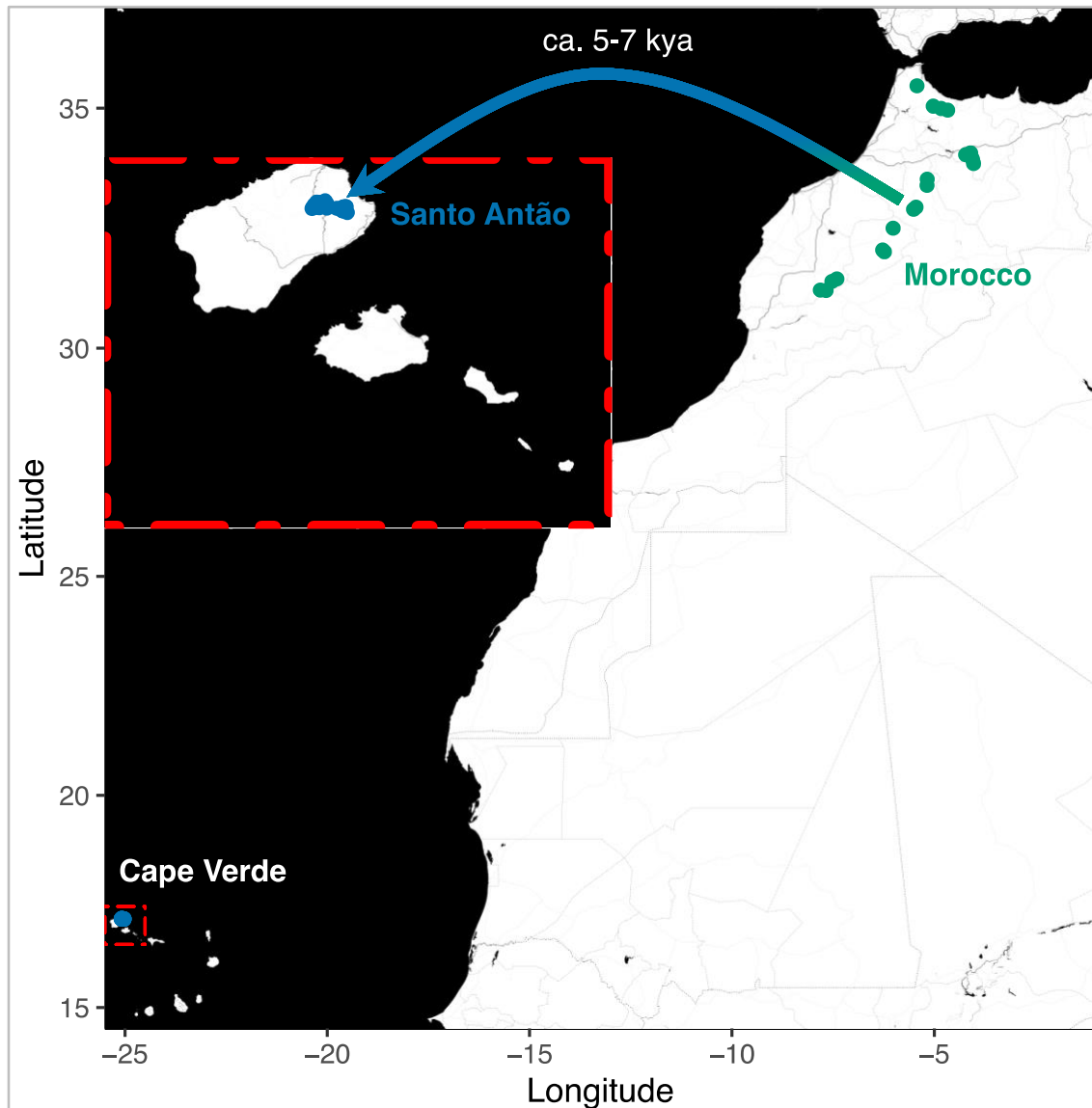
Fig. 1. *Collection locations of Arabidopsis thaliana Santo Antão, CVI (n = 189) and Morocco (n = 61). The arrow indicates the colonization of the CVI from Morocco ∼5–7 kya.*

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

In CVI, rainfall is limited and unpredictable (**Supplementary Figure S1A-B**), and water vapor pressure (specific humidity) is consistently high relative to Moroccan *A. thaliana* sites (**Supplementary Figure 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 Figure 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 - δ^{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**, **Supplementary Material** online). 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**, **Supplementary Material** online), 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 Figure 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 Figure S4**).
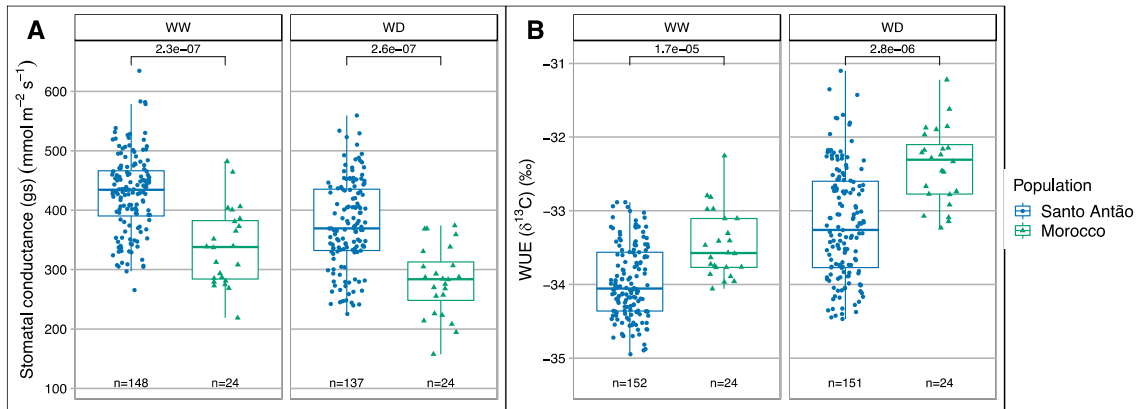
Fig. 2. *Phenotypic variation in (A) stomatal conductance (gs) and (B) WUE for Santo Antão, CVI, and Moroccan Arabidopsis thaliana populations in WW and WD conditions. The line in the center of the boxplots represents the median, the box edges represent the 25th and 75th percentiles (lower and upper bound, respectively), and the whiskers represent the 95% CI. The WUE is measured as carbon isotope discrimination (δ13C), and the carbon isotope ratio is expressed per mil, ‰. The P-values for the Mann–Whitney–Wilcoxon test are shown.*

2.3.2 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**, Supplementary Material online). Furthermore, we found that seven loci were predicted to have effects on WUE in WW and WD conditions (**Supplementary Figure S5A, supplementary table S3-4**, Supplementary Material online). 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 Figure S6A, supplementary table S3-4**, Supplementary Material online). In addition, about 39 and 53 loci were predicted to have major effects in WW and WD conditions, respectively (**Supplementary Figure S7A, supplementary table S3-4**, Supplementary Material online). 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 < 2.2x10^{-16} for the average traits, R^2 = 0.28, P < 2.2x10^{-16} for the drought response of traits, R^2 = 0.16, P < 2.2x10^{-16} for WW, and R^2 = 0.55, P < 2.2x10^{-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.

2.3.3 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 δ^{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 Figure 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 Figure S6B**) as well as for both conditions separately (**Supplementary Figure 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 Figure S8A-B**). Taken together, our results support a central role for the *MPK12* G53R variant in trait variation in the natural CVI population.
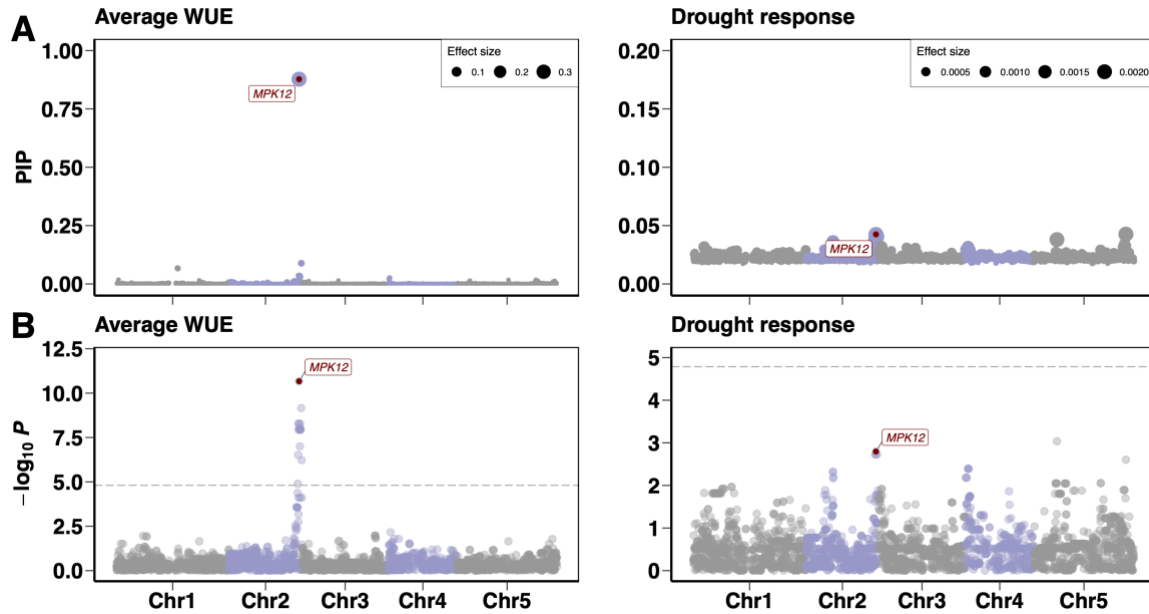
Fig. 3. Genome-wide association (GWA) mapping of WUE. (A) Polygenic modeling of the average WUE across the WW and WD conditions (left) and the drought response of WUE (the difference between both conditions: WW-WD) (right) in the Santo Antão Arabidopsis thaliana population using BSLMM. The y-axis represents the posterior inclusion probability 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 α = 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).

2.3.4 Reconstructing the evolutionary history of variation in water use efficiency

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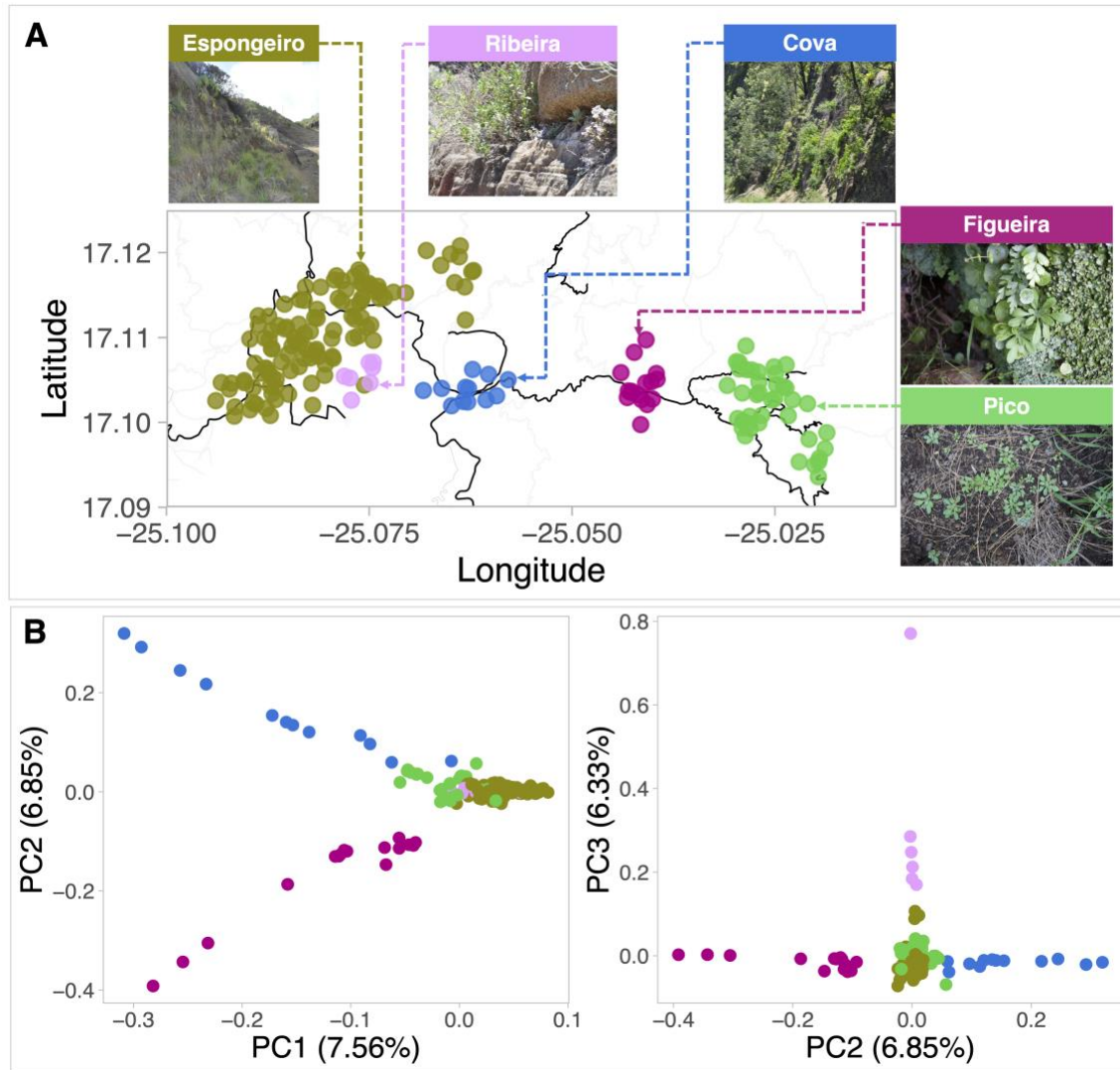
Fig. 4. *Population structure of Arabidopsis thaliana subpopulations in Santo Antão. (A) Geographical distribution of subpopulations across Santo Antão. Images of representative sites for the five subpopulations in Santo Antão show the diversity of habitats that A. thaliana occupies. (B) PCA of genome-wide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

2.3.4.2 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 Figure 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 Figure 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 Figure S9B** and **Supplementary Figure 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. 5D**). 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.
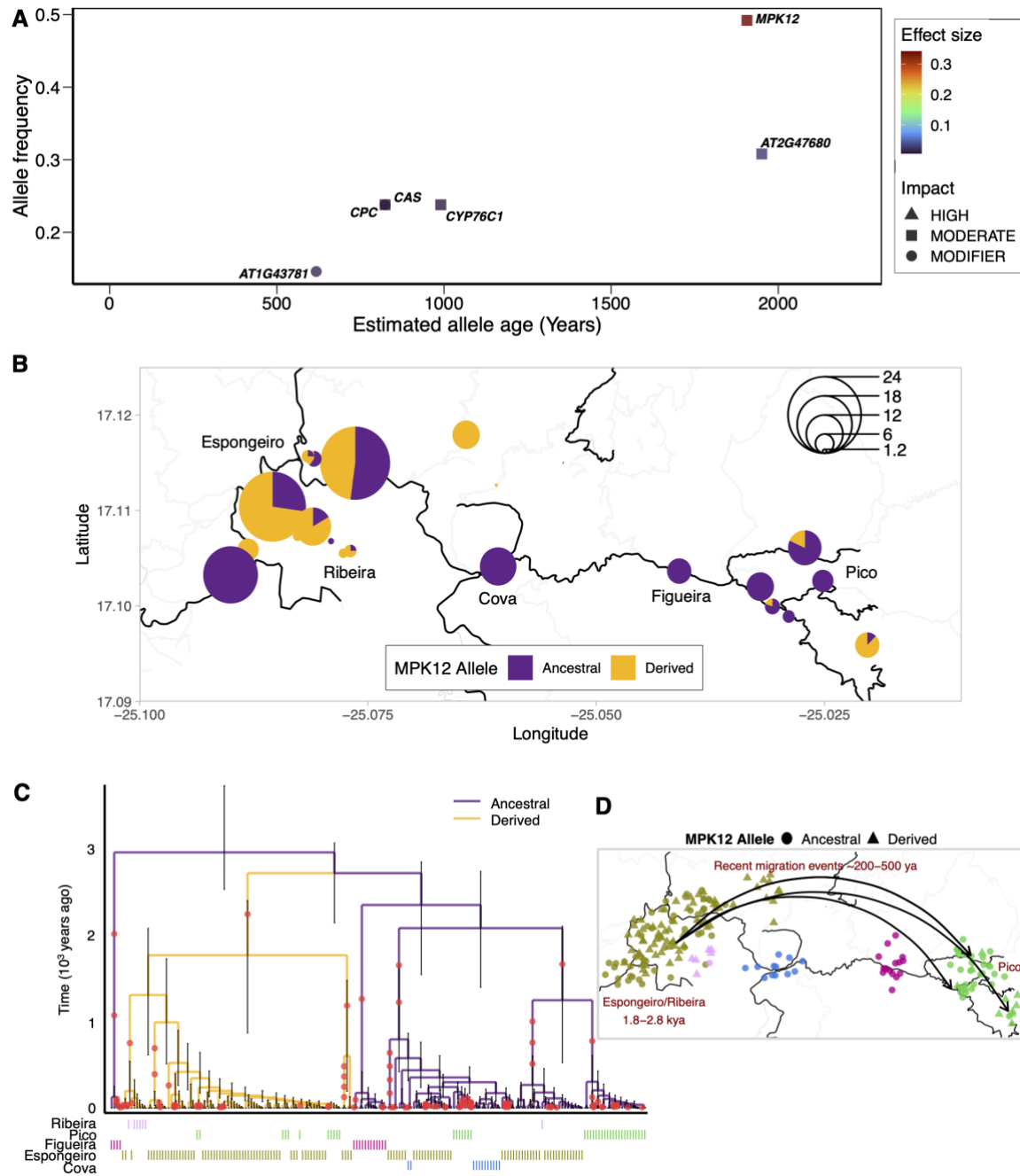
Fig. 5. Evolutionary history of water-use efficiency variation. (A) Estimated allele ages (inferred in RELATE) versus allele frequencies of variants with major effects estimated from GWA mapping of average WUE. Shape denotes predicted impact from gene annotation. (B) Spatial distribution of MPK12 G53R 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. (D) A model of the origin and spread of the MPK12 G53R variant based on the genealogical inference in (C). Figueira: Lombo de Figueira; Cova: Cova de Paúl; Ribeira: Ribeira de Poio; Pico: Pico da Cruz.

2.3.5   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, $-\log_{10}(p\text{-value})$ = 2.35) (**Fig. 6C**).

We next used gene ontology (GO) 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 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 Figure S12A-B, supplementary table S5, [Supplementary Material](#)** online). 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 Figure S13A**) as well as between Espongeiro and Pico (|XP-EHH|= 2.03, -$\log_{10}$(*p*-value)= 1.37) (**Supplementary Figure 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 Figure S14, supplementary table S6**, <u>Supplementary Material</u> online). 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**, <u>Supplementary Material</u> online) 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 Figure S15**; **supplementary table S8-S9**, <u>Supplementary Material</u> online). 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 Figure 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 Figure S15A**). We next examined the loadings by SNP (**Supplementary Figure S15B**; **supplementary table S10**, [Supplementary Material](#) online) 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 Figure 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 Figure S16**; **supplementary table S11**, [Supplementary Material](#) online). 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**, **Supplementary Material** online), indicating the *MPK12* 53R variant increases fitness independently from *FRI* 232X under CVI (Espongeiro) conditions.
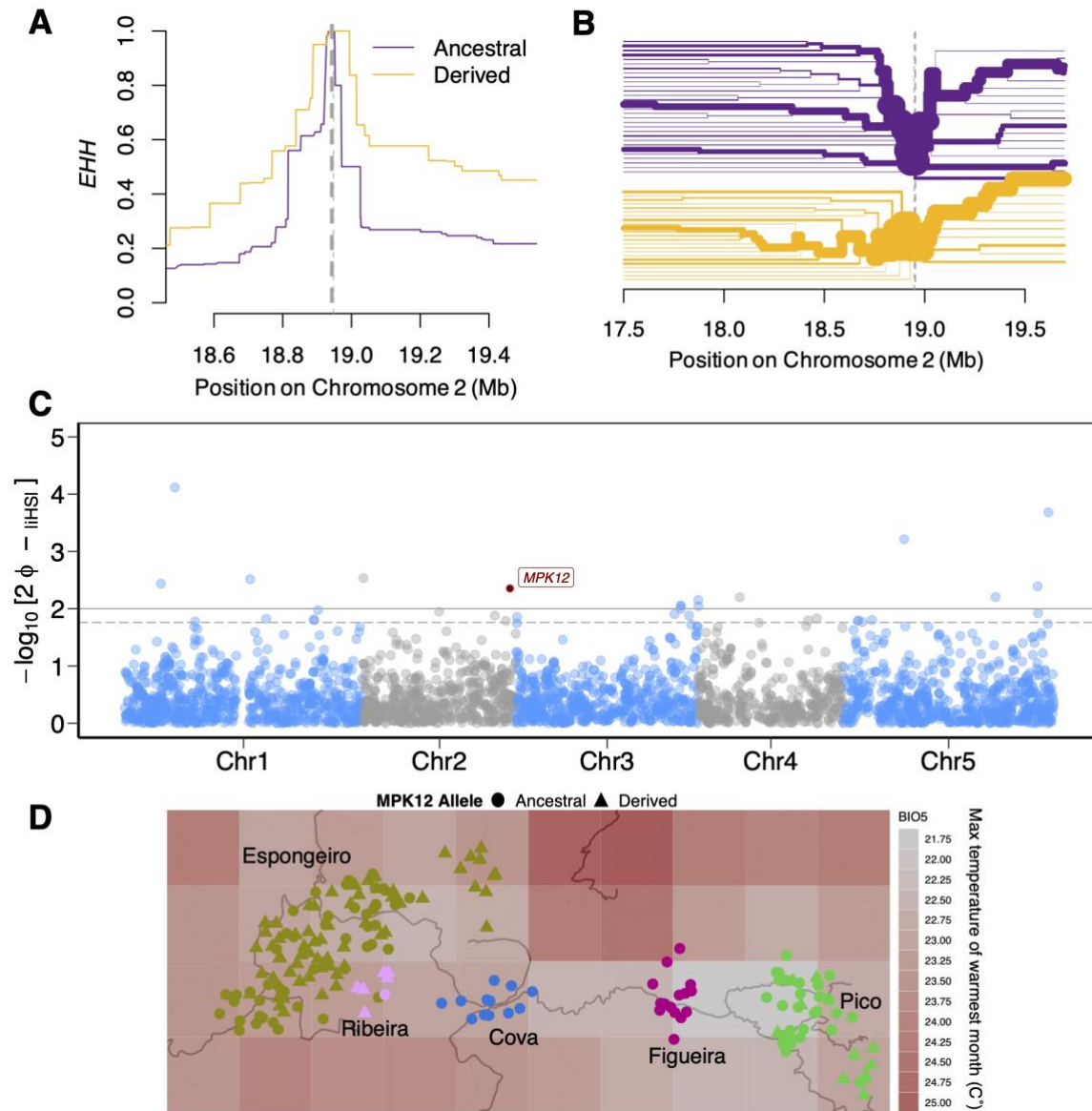
Fig. 6. *Signature of a partial selective sweep at MPK12 G53R. (A) The decay of EHH and (B) bifurcation analysis of the ancestral (upper) and derived (lower) alleles for MPK12 G53R. The dotted line marks the position of the focal MPK12 SNP. The width of the lines in (B) represents the frequency of haplotypes 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 (−log10 (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 point and a box with the text "MPK12". (D) Geographic distribution of Santo Antão Arabidopsis 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.*

2.4 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 Figure 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 Figure 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 (Schemenauer, Bignell, *et al.*, 2016; Schemenauer, Zanetta, *et al.*, 2016; Klemm *et al.*, 2012). 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.

2.5 Materials and methods

2.5.1 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, Supplementary Material** online) (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.

2.5.2 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, Supplementary Material** online) 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.

2.5.3 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.

2.5.4 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.

2.5.5 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).

2.5.6 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 7×10^{-9} derived from (Ossowski *et al.*, 2010) for the percent missing data in 1 Mb sliding windows every 50 kb across the entire genome (2.245×10^{-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.

2.5.7 Climatic variables

We retrieved data for the 19 bioclimatic variables (**supplementary table S7**, **Supplementary Material** online) 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**, **Supplementary Material** online), which we used for a comparison between the climate of collection sites in Santo Antão and Moroccan sites (**Supplementary Figure 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)).

2.5.8 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**, **Supplementary Material** online) 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.

2.5.9 Evolutionary history of *MPK12* 53R

2.5.9.1   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: piHS = $-$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 Figure S12A-B**, **Supplementary Material** online). 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.

2.5.9.2   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**, **Supplementary Material** online).

2.6 Acknowledgements

2.7 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 δ^{13}C analysis. N.K., A.P.M.W and his lab members were responsible for δ^{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.

2.8 Data Availability

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

2.9   References

Allen, J.A. (1877) The influence of physical conditions in the genesis of species. *Radic. Rev.*, **1**, 108–140.

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

Aragón, W., Formey, D., Aviles-Baltazar, N.Y., Torres, M. and Serrano, M. (2021) *Arabidopsis thaliana* Cuticle Composition Contributes to Differential Defense Response to Botrytis cinerea. *Front. Plant Sci.*, **12**:738949.

Assmann, S.M. (2013) Natural variation in abiotic stress and climate change responses in Arabidopsis: implications for twenty-first-century agriculture. *Int. J. Plant Sci.*, **174**, 3–26.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., et al. (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, R.H.B., Singmann, H. and Dai, B. (2014) *Lme4: linear mixed-effects models using eigen and S4 (version 1.1-7)* [computer software]. http://CRAN.R-project.org/package=lme4.

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

Berry, H.L., Waite, T.D., Dear, K.B.G., Capon, A.G. and Murray, V. (2018) The case for systems thinking about climate change and mental health. *Nat. Clim. Change*, **8**, 282–290.

Bhaskara, G.B., Lasky, J.R., Razzaque, S., Zhang, L., Haque, T., Bonnette, J.E., Civelek, G.Z., Verslues, P.E. and Juenger, T.E. (2022) Natural variation identifies new effectors of water-use efficiency in Arabidopsis. *Proc. Natl. Acad. Sci.*, **119**, e2205305119.

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

Bjorkman, A.D., Myers-Smith, I.H., Elmendorf, S.C., et al. (2018) Plant functional trait change across a warming tundra biome. *Nature*, **562**, 57–62.

Blanc, C., Coluccia, F., L'Haridon, F., et al. (2018) The Cuticle Mutant eca2 Modifies Plant Defense Responses to Biotrophic and Necrotrophic Pathogens and Herbivory Insects. *Mol. Plant-Microbe Interactions®*, **31**, 344–355.

Bomblies, K., Yant, L., Laitinen, R.A., Kim, S.-T., Hollister, J.D., Warthmann, N., Fitz, J. and Weigel, D. (2010) Local-Scale Patterns of Genetic Variability, Outcrossing, and

Spatial Structure in Natural Stands of Arabidopsis thaliana. *PLOS Genet.*, **6**, e1000890.

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

Brennan, A.C., Méndez-Vigo, B., Haddioui, A., Martínez-Zapater, J.M., Picó, F.X. and 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 Biol.*, **14**, 17.

Brochmann, C., Rustan, Ø.H., Lobin, W. and Kilian, N. (1997) The endemic vascular plants of the Cape Verde Islands, W Africa. *Sommerfeltia*, **24**, 1–363.

Bruijnzeel, L.A. and Hamilton, L.S. (2017) *Decision Time Cloud Forests*, Available at: http://www.unep.org/resources/report/decision-time-cloud-forests [Accessed May 26, 2022].

Bruijnzeel, L.A., Kappelle, M., Mulligan, M. and Scatena, F.N. (2011) *Tropical montane cloud forests: state of knowledge and sustainability perspectives in a changing world*, Cambridge University Press.

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

Campitelli, B.E., Marais, D.L. Des and Juenger, T.E. (2016) Ecological interactions and the fitness effect of water-use efficiency: Competition and drought alter the impact of natural *MPK12* alleles in *Arabidopsis* B. Enquist, ed. *Ecol. Lett.*, **19**, 424–434.

Chan, E.K., Rowe, H.C., Hansen, B.G. and Kliebenstein, D.J. (2010) The complex genetic architecture of the metabolome. *PLoS Genet.*, **6**, e1001198.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (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 (Austin)*, **6**, 80–92.

Cohen, D. (1970) The expected efficiency of water utilization in plants under different competition and selection regimes. *Isr. J. Bot.*, **19**, 50–4.

Cowan, I. (1986) Economics of carbon fixation in higher plants. *Econ. Carbon Fixat. High. Plants*, 133–170.

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

Davila Olivas, N.H., Kruijer, W., Gort, G., Wijnen, C.L., Loon, J.J. van and Dicke, M. (2017) Genome-wide association analysis reveals distinct genetic architectures for single and combined stress responses in Arabidopsis thaliana. *New Phytol.*, **213**, 838–851.

Dawson, T.E. (1998) Fog in the California redwood forest: ecosystem inputs and use by plants. *Oecologia*, **117**, 476–485.

de Jong G. (1993) Covariances Between Traits Deriving From Successive Allocations of a Resource. *Funct. Ecol.*, **7**, 75–83.

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. *Proc Natl Acad Sci U S A.*, **111**:2836–2841.

Díaz, S., Kattge, J., Cornelissen, J.H.C., et al. (2016) The global spectrum of plant form and function. *Nature*, **529**, 167–171.

Dikmen, S., Cole, J.B., Null, D.J. and Hansen, P.J. (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. and Ehleringer, J. (1992) Contrasting water-use patterns among size and life-history classes of a semi-arid shrub. *Funct. Ecol.*, 482–488.

Donovan, L. and Ehleringer, J. (1994) Water stress and use of summer precipitation in a Great Basin shrub community. *Funct. Ecol.*, 289–297.

Durvasula, A., Fulgione, A., Gutaker, R.M., et al. (2017) African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **114**, 5213–5218.

Exposito-Alonso, M., Vasseur, F., Ding, W., Wang, G., Burbano, H.A. and Weigel, D. (2018) Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nat. Ecol. Evol.*, **2**, 352–358.

Falconer, D.S. and Mackay, T.F. (1983) *Quantitative genetics*, Longman London, UK.

Ferrero-Serrano, Á. and Assmann, S.M. (2019) Phenotypic and genome-wide association with the local environment of Arabidopsis. *Nat. Ecol. Evol.*, **3**, 274–285.

Fick, S.E. and Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.*, **37**, 4302–4315.

Filiault, D.L. and Maloof, J.N. (2012) A genome-wide association study identifies variants underlying the Arabidopsis thaliana shade avoidance response. *PLoS Genet.*, **8**, e1002589.

Fournier-Level, A., Korte, A., Cooper, M.D., Nordborg, M., Schmitt, J. and Wilczek, A.M. (2011) A map of local adaptation in Arabidopsis thaliana. *Science*, **334**, 86–89.

Fulgione, A., Koornneef, M., Roux, F., Hermisson, J. and Hancock, A.M. (2018) Madeiran Arabidopsis thaliana reveals ancient long-range colonization and clarifies demography in Eurasia. *Mol. Biol. Evol.*, **35**, 564–574.

Fulgione, A., Neto, C., Elfarargi, A.F., 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. and Vitalis, R. (2017) rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol. Ecol. Resour.*, **17**, 78–90.

Gazzani, S., Gendall, A.R., Lister, C. and Dean, C. (2003) Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. *Plant Physiol.*, **132**, 1107–1114.

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

Geber, M.A. and Dawson, T.E. (1990) Genetic variation in and covariation between leaf gas exchange, morphology, and development in Polygonum arenastrum, an annual plant. *Oecologia*, **85**, 153–158.

Geber, M.A. and Dawson, T.E. (1997) Genetic variation in stomatal and biochemical limitations to photosynthesis in the annual plant, Polygonum arenastrum. *Oecologia*, **109**, 535–546.

Gloss, A.D., Vergnol, A., Morton, T.C., Laurin, P.J., Roux, F. and Bergelson, J. (2022) Genome-wide association mapping within a local Arabidopsis thaliana population more fully reveals the genetic architecture for defensive metabolite diversity. *Philos. Trans. R. Soc. B*, **377**, 20200512.

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

Grant, P.R. (1999) *Ecology and Evolution of Darwin's Finches*, Princeton University Press.

Hancock, A.M., Brachi, B., Faure, N., Horton, M.W., Jarymowycz, L.B., Sperone, F.G., Toomajian, C., Roux, F. and Bergelson, J. (2011) Adaptation to climate across the Arabidopsis thaliana genome. *Science*, **334**, 83–86.

Heschel, M.S. and Riginos, C. (2005) Mechanisms of selection for drought stress tolerance and avoidance in Impatiens capensis (Balsaminaceae). *Am. J. Bot.*, **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**. https://CRAN.R-project.org/package=raster

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

Jakobson, L., Vaahtera, L., Tõldsepp, K., et al. (2016) Natural Variation in Arabidopsis Cvi-0 Accession Reveals an Important Role of MPK12 in Guard Cell CO2 Signaling. *PLoS Biol.*, **14**, e2000322.

Jammes, F., Song, C., Shin, D., et al. (2009) MAP kinases MPK9 and MPK12 are preferentially expressed in guard cells and positively regulate ROS-mediated ABA signaling. *Proc. Natl. Acad. Sci.*, **106**, 20520–20525.

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

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and 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.

Juenger, T.E., Mckay, J.K., Hausmann, N., Keurentjes, J.J., Sen, S., Stowe, K.A., Dawson, T.E., Simms, E.L. and Richards, J.H. (2005) Identification and characterization of QTL underlying whole-plant physiology in Arabidopsis thaliana: δ13C, stomatal conductance and transpiration efficiency. *Plant Cell Environ.*, **28**, 697–708.

Kahle, D. and Wickham, H. (2013) ggmap: Spatial Visualization with ggplot2. *R J.*, **5**, 144–161.

Kalladan, R., Lasky, J.R., Chang, T.Z., Sharma, S., Juenger, T.E. and Verslues, P.E. (2017) Natural variation identifies genes affecting drought-induced abscisic acid accumulation in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **114**, 11536–11541.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. and Kessler, M. (2017) Climatologies at high resolution for the earth's land surface areas. *Sci. Data*, **4**, 1–20.

Karger, D.N., Kessler, M., Lehnert, M. and 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, A.M., McKay, J.K., Richards, J.H. and Juenger, T.E. (2014) Direct and indirect selection on flowering time, water-use efficiency (WUE, δ ^{13} C), and WUE plasticity to drought in *Arabidopsis thaliana*. *Ecol. Evol.*, **4**, 4505–4521.

Kerfoot, O. (1968) *Mist precipitation on vegetation*. *Forestry Abstracts* **29**, 8–20.

Klein, T. (2014) The variability of stomatal sensitivity to leaf water potential across tree species indicates a continuum between isohydric and anisohydric behaviours. *Funct. Ecol.*, **28**, 1313–1320.

Klemm, O., Schemenauer, R.S., Lummerich, A., et al. (2012) Fog as a Fresh-Water Resource: Overview and Perspectives. *Ambio*, **41**, 221–234.

Koornneef, M., Alonso-Blanco, C. and Vreugdenhil, D. (2004) Naturally occurring genetic variation in Arabidopsis thaliana. *Annu. Rev. Plant Biol.*, **55**, 141–172.

Korves, T.M., Schmid, K.J., Caicedo, A.L., Mays, C., Stinchcombe, J.R., Purugganan, M.D. and Schmitt, J. (2007) Fitness Effects Associated with the Major Flowering Time Gene FRIGIDA in Arabidopsis thaliana in the Field. *Am. Nat.*, **169**, E141–E157.

Kranz, A.R. and Kirchheim, B. (1987) Genetic resources in Arabidopsis. *Arabidopsis Information Service* **24**, 1–167.

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

Lasky, J.R., Des Marais, D.L., McKAY, J.K., Richards, J.H., Juenger, T.E. and Keitt, T.H. (2012) Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol. Ecol.*, **21**, 5512–5529.

Levitt, J. (1972) *Responses of plants to environmental stresses (Physiological Ecology): Chilling, freezing, and high temperature stresses*. New York (NY): Academic Press. 697 pp.

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. and Borevitz, J.O. (2010) Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **107**, 21199–21204.

Liu, Y., Maierhofer, T., Rybak, K., et al. (2019) Anion channel SLAH3 is a regulatory target of chitin receptor-associated kinase PBL27 in microbial stomatal closure J.-M. Zhou, C. S. Hardtke, and S. Hoth, eds. *eLife*, **8**, e44474.

Losos, J.B. and Ricklefs, R.E. (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.

Losos, J.B., Warheitt, K.I. and Schoener, T.W. (1997) Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature*, **387**, 70–73.

Lovell, J.T., Juenger, T.E., Michaels, S.D., et al. (2013) Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Proc. R. Soc. B Biol. Sci.*, **280**, 20131043.

Ludlow, M. (1989) *Strategies of response to water stress*. The Hague, Netherlands: SPB Academic Press.

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

Martin, S.H. and Belleghem, S.M. Van (2017) Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. *Genetics*, **206**, 429–438.

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

McDowell, N., Pockman, W.T., Allen, C.D., et al. (2008) Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New Phytol.*, **178**, 719–739.

McDowell, N.G., Beerling, D.J., Breshears, D.D., Fisher, R.A., Raffa, K.F. and Stitt, M. (2011) The interdependence of mechanisms underlying climate-driven vegetation mortality. *Trends Ecol. Evol.*, **26**, 523–532.

McKay, J.K., Richards, J.H. and Mitchell-Olds, T. (2003) Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.*, **12**, 1137–51.

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

Montillet, J.-L., Leonhardt, N., Mondy, S., et al. (2013) An Abscisic Acid-Independent Oxylipin Pathway Controls Stomatal Closure and Immune Defense in Arabidopsis. *PLOS Biol.*, **11**, e1001513.

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

Moreno-Gutiérrez, C., Dawson, T.E., Nicolás, E. and Querejeta, J.I. (2012) Isotopes reveal contrasting water use strategies among coexisting plant species in a Mediterranean ecosystem. *New Phytol.*, **196**, 489–496.

Morrison, K. and Stacy, E. (2014) Intraspecific divergence and evolution of a life-history trade-off along a successional gradient in Hawaii's Metrosideros polymorpha. *J. Evol. Biol.*, **27**, 1192–1204.

Oksanen, J., Blanchet, F.G., Friendly, M., et al. (2020) vegan: Community Ecology Package. Available from: https://CRAN.R-project.org/package=vegan.

Ossowski, S., Schneeberger, K., Lucas-Lledó, J.I., Warthmann, N., Clark, R.M., Shaw, R.G., Weigel, D. and Lynch, M. (2010) The rate and molecular Spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327**:92–94.

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

Porch, T.G., Tseung, C.-W., Schmelz, E.A. and Mark Settles, A. (2006) The maize Viviparous10/Viviparous13 locus encodes the Cnx1 gene required for molybdenum cofactor biosynthesis. *Plant J.*, **45**, 250–263.

Purcell, S., Neale, B., Todd-Brown, K., et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–75.

Querejeta, J.I., Prieto, I., Torres, P., Campoy, M., Alguacil, M.M. and Roldán, A. (2018) Water-spender strategy is linked to higher leaf nutrient concentrations across plant species colonizing a dry and nutrient-poor epiphytic habitat. *Environ. Exp. Bot.*, **153**, 302–310.

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

Ray, C. (1960) The application of Bergmann's and Allen's rules to the poikilotherms. *J. Morphol.*, **106**, 85–108.

Rheenen, W. van, Peyrot, W.J., Schork, A.J., Lee, S.H. and Wray, N.R. (2019) Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.*, **20**, 567–581.

Roff, D.A. (2002) *Life history evolution*. Sunderland (MA): Sinauer Associates.

Roux, F. and 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, P.C., Reich, D.E., Higgins, J.M., et al. (2002) Detecting recent positive selection in the human genome from haplotype structure. *Nature*, **419**, 832–837.

Sabeti, P.C., Varilly, P., Fry, B., et al. (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature*, **449**, 913–918.

Salam, M.A., Jammes, F., Hossain, M.A., Ye, W., Nakamura, Y., Mori, I.C., Kwak, J.M. and Murata, Y. (2013) Two guard cell-preferential MAPKs, MPK9 and MPK12, regulate YEL signalling in Arabidopsis guard cells. *Plant Biol.*, **15**, 436–442.

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

Schemenauer, R., Bignell, B. and Makepeace, T. (2016) Fog Collection Projects in Nepal: 1997 to 2016. In: *Proceedings of the 7th International Conference on Fog, Fog Collection and Dew, Wroclaw, Poland*, p. 187–190.

Schemenauer, R., Zanetta, N., Rosato, M. and Carter Gamberini, M.V. (2016) *The Tojquia, Guatemala Fog Collection Project 2006 to 2016*. In: Blas M, Sobik M, editors. *Proceedings of the 7th International Conference on Fog, Fog Collection and Dew*. Wroclaw: University of Wroclaw. p. 210–213

Shefferson, R.P., Proper, J., Beissinger, S.R. and Simms, E.L. (2003) Life history trade-offs in a rare orchid: the costs of flowering, dormancy, and sprouting. *Ecology*, **84**, 1199–1206.

Sherrard, M.E. and Maherali, H. (2006) The Adaptive Significance of Drought Escape in Avena barbata, an Annual Grass. *Evolution*, **60**, 2478–2489.

Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M. and 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, M.J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M. and Dean, C. (2005) Role of FRIGIDA and FLOWERING LOCUS C in Determining Variation in Flowering Time of Arabidopsis. *Plant Physiol.*, **138**, 1163–1173.

Skelton, R.P., West, A.G. and Dawson, T.E. (2015) Predicting plant vulnerability to drought in biodiverse regions using functional traits. *Proc. Natl. Acad. Sci.*, **112**, 5744–5749.

Speidel, L., Forest, M., Shi, S. and Myers, S.R. (2019) A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet. 2019 519*, **51**, 1321–1329.

Stadtmüller, T. (1987) *Cloud Forests in the Humid Tropics: A Bibliographic Review*. The United Nations University, Tokyo, and Centro Agronomico Tropical de Investigacion y Ensenanza, Turrialba, Costa Rica.

Stearns, S.C. (1989) Trade-Offs in Life-History Evolution. *Funct. Ecol.*, **3**, 259–268.

Stern, A.J., Wilton, P.R. and Nielsen, R. (2019) An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data R. D. Hernandez, ed. *PLOS Genet.*, **15**, e1008384.

Tabas-Madrid, D., Méndez-Vigo, B., Arteaga, N., Marcer, A., Pascual-Montano, A., Weigel, D., Xavier Pico, F. and 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 Environ.*, **41**, 1806–1820.

Tergemina, E., Elfarargi, A.F., Flis, P., et al. (2022) A two-step adaptive walk rewires nutrient transport in a challenging edaphic environment. *Sci. Adv.*, **8**, eabm9385.

Tisné, S., Serrand, Y., Bach, L., et al. (2013) Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J.*, **74**, 534–544.

Verslues, P.E. and Juenger, T.E. (2011) Drought, metabolites, and Arabidopsis natural variation: a promising combination for understanding adaptation to water-limited environments. *Curr. Opin. Plant Biol.*, **14**, 240–245.

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

Voight, B.F., Kudaravalli, S., Wen, X. and Pritchard, J.K. (2006) A Map of Recent Positive Selection in the Human Genome L. Hurst, ed. *PLoS Biol.*, **4**, e72.

Wallace, A.R. (1855) On the law which has regulated the introduction of new species. *Ann Mag Nat Hist.* **16**:184–196. doi:10.1080/037454809495509.

Walter, H. (1985) *Vegetation of the earth and ecological systems of the geo-biosphere.* 3rd ed. Berlin: Springer.

Weathers, K.C. (1999) The importance of cloud and fog in the maintenance of ecosystems. *Trends Ecol. Evol.*, **14**, 214–215.

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

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

Wieters, B., Steige, K.A., He, F., Koch, E.M., Ramos-Onsins, S.E., Gu, H., Guo, Y.-L., Sunyaev, S. and Meaux, J. de (2021) Polygenic adaptation of rosette growth in Arabidopsis thaliana. *PLoS Genet.*, **17**, e1008748.

Will, R.E., Wilson, S.M., Zou, C.B. and Hennessey, T.C. (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 Phytol.*, **200**, 366–374.

Willi, Y. and Van Buskirk, J. (2022) A review on trade-offs at the warm and cold ends of geographical distributions. *Philos. Trans. R. Soc. B Biol. Sci.*, **377**, 20210022.

Wright, I.J., Reich, P.B., Westoby, M., et al. (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

Zan, Y. and Carlborg, Ö. (2019) A polygenic genetic architecture of flowering time in the worldwide Arabidopsis thaliana population. *Mol. Biol. Evol.*, **36**, 141–154.

Zera, A.J. and Harshman, L.G. (2001) The Physiology of Life History Trade-Offs in Animals. *Annu. Rev. Ecol. Syst.*, **32**, 95–126.

Zhou, X. (2014) Gemma user manual. Chicago (IL): *University of Illinois Chicago*.

Zhou, X., Carbonetto, P. and Stephens, M. (2013) Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *PLOS Genet.*, **9**, e1003264.

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

Zhou, X. and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.*, **44**, 821–824.

Zhou, Z., Li, M., Cheng, H., et al. (2018) An intercross population study reveals genes associated with body size and plumage color in ducks. *Nat. Commun.*, **9**, 1–10.

2.10 Supplementary figures
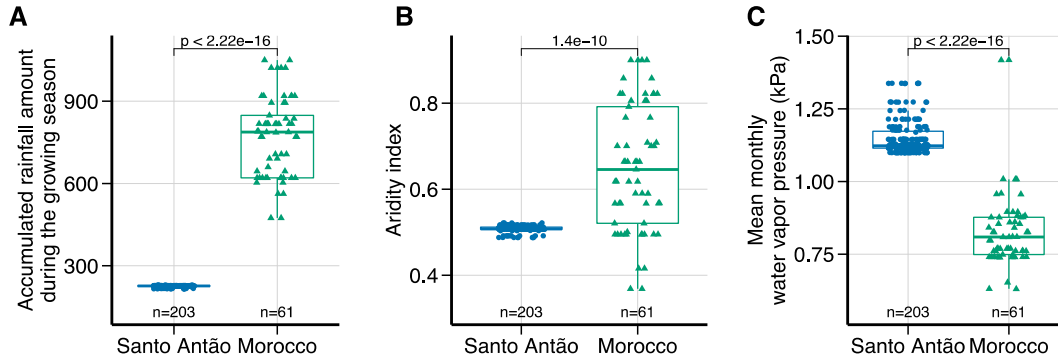
Supplementary Figure S1. *Climate of (A) accumulated rainfall during growing season (kg m^{-2} gsl^{-1}), (B) aridity index, and (C) mean monthly water vapor pressure (KPa) at collection sites in Santo Antão (blue; n=203) relative to Moroccan sites (green; n=61). The line in the center of the boxplots represents the median, the box edges represent the 25th and 75th percentiles (lower and upper bound, respectively), and the whiskers represent 95% CI. The P-values for the Mann-Whitney-Wilcoxon test are shown.*
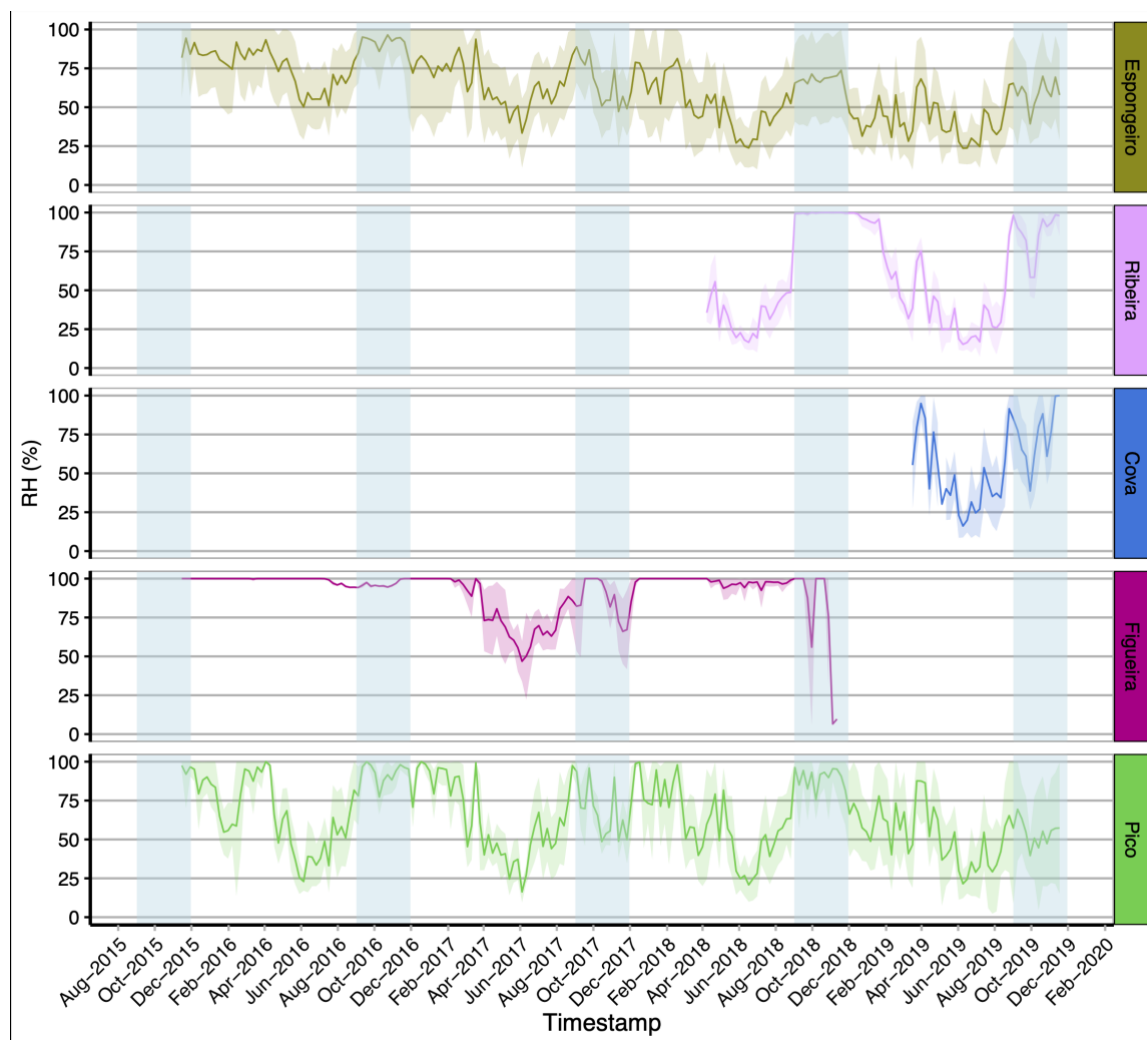
Supplementary Figure S2. *Measurements of the relative humidity (RH) over time using loggers in the field sites of the five Santo Antão sub-populations. The light blue shaded areas over time represent the growth season period (September to November).*
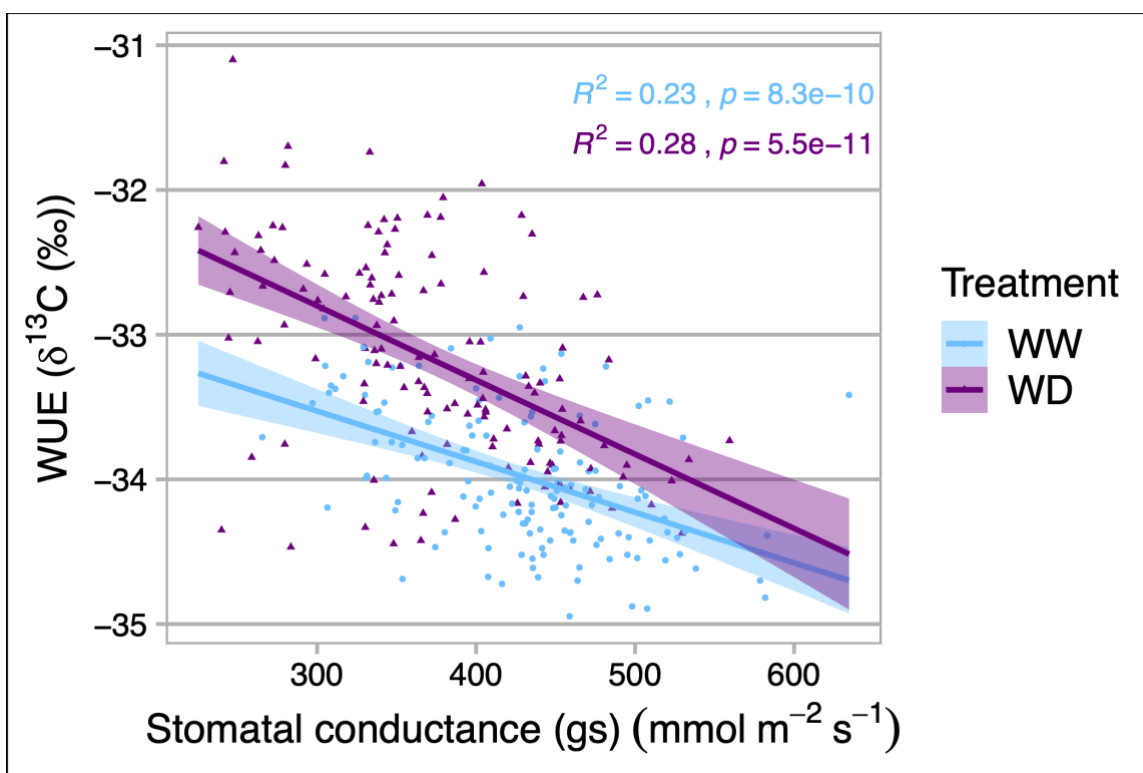
Supplementary Figure S3. *The relationship between water use efficiency (WUE) and stomatal conductance in Santo Antão population. WUE is negatively correlated with stomatal conductance in well-watered (WW) and in response to water deficit (WD). The WUE is measured as carbon isotope discrimination ($\delta^{13}C$), and the carbon isotope ratio is expressed as (per mil, ‰). The shaded areas around the slope represent the confidence interval of the correlation coefficient at 95%. R^2= Pearson's squared correlation coefficient and p= p-value.*
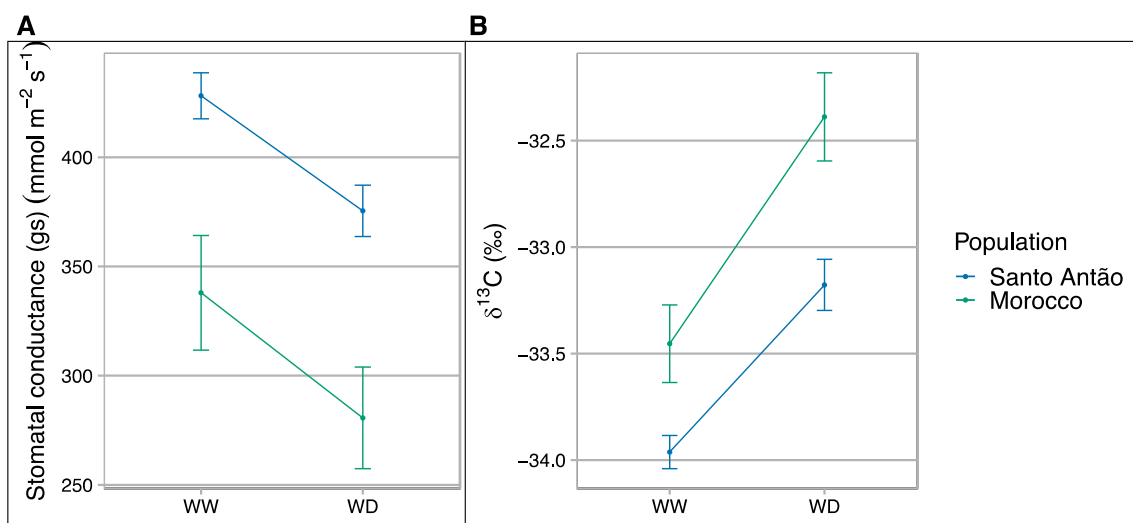
Supplementary Figure S4. *Phenotypic plasticity of (A) water use efficiency (WUE) and (B) stomatal conductance, among Santo Antão A. thaliana population under well-watered (WW) and water deficit (WD) conditions. The WUE is measured as carbon isotope discrimination ($\delta^{13}C$), and the carbon isotope ratio is expressed as (per mil, ‰). The points represent the means and the whiskers represent the 95% CI.*
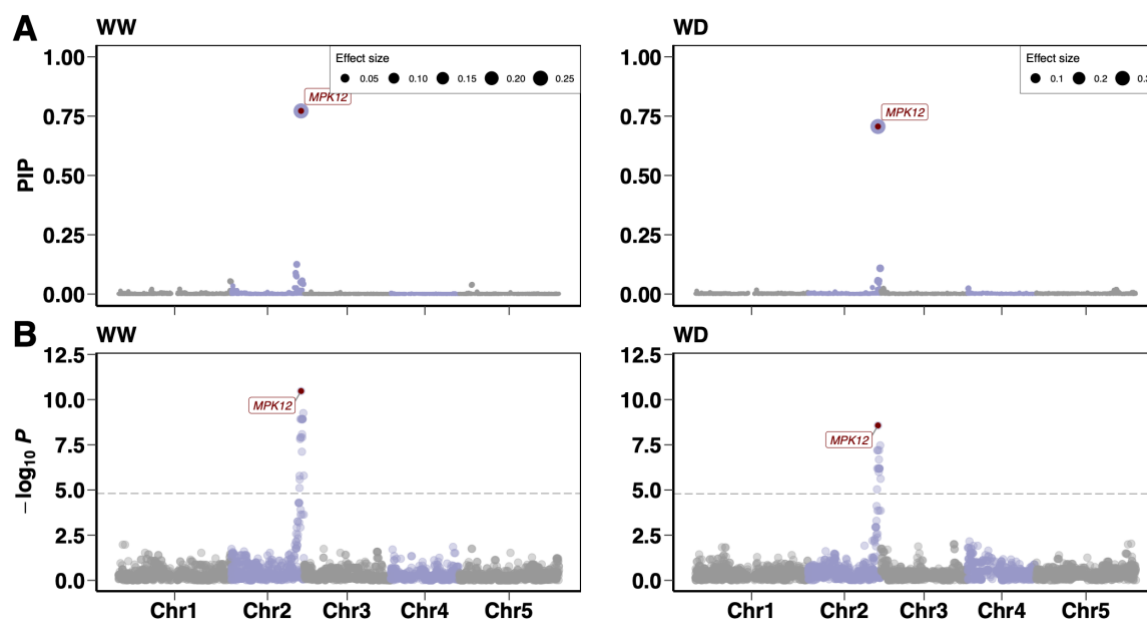
Supplementary Figure S5. Genome-wide association (GWA) mapping. *(A) Polygenic modeling of water use efficiency (WUE) variation in the Santo Antão A. thaliana population in well-watered (WW) and water deficit (WD) conditions using BSLMM. The y-axis represents the posterior inclusion probability (PIP) and the size of symbol denotes the effect size. (B) Genome-wide association mapping of WUE variation in WW and WD conditions using LMM. The horizontal dashed line corresponds to the Bonferroni genome-wide significance. In* ***A*** *and* ***B****, points represent SNPs along the five chromosomes. The red point represents the MPK12 G53R variant (a substitution of arginine for glycine at amino acid position 53).*
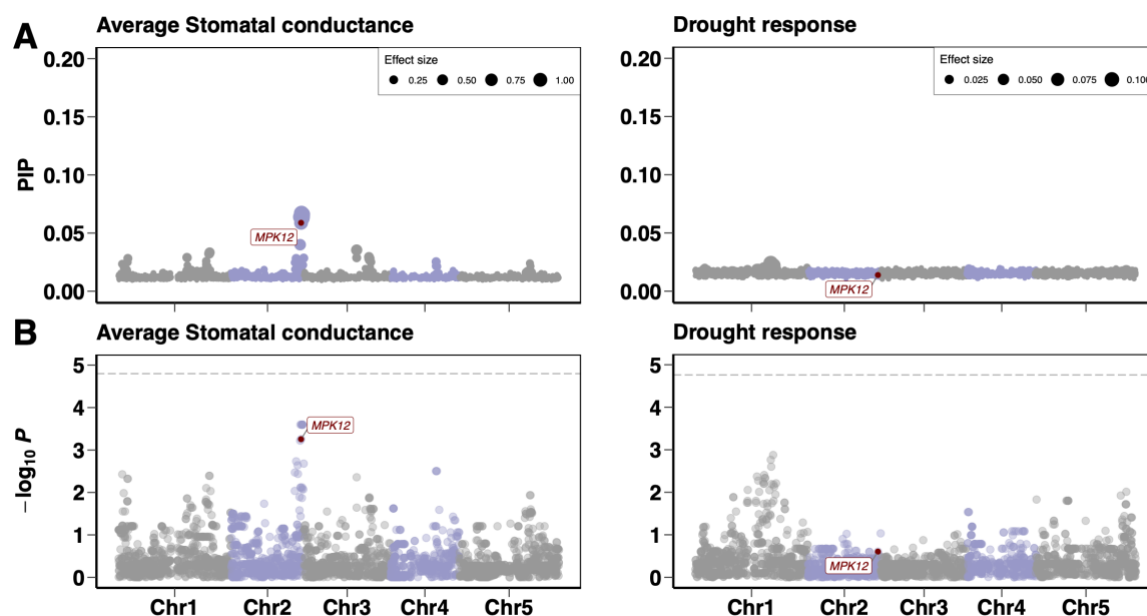
Supplementary Figure S6. Genome-wide association (GWA) mapping of stomatal conductance. *(A) Polygenic modeling of the average stomatal conductance across the well-watered (WW) and water deficit (WD) conditions and the drought response of stomatal conductance (the difference between both conditions: WW-WD) 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 stomatal conductance and its drought response using LMM. The horizontal dashed line corresponds to the Bonferroni significance threshold at α = 0.05. In A and B, points represent SNPs along the five chromosomes. The red point represents the MPK12 G53R variant (a substitution of arginine for glycine at amino acid position 53).*
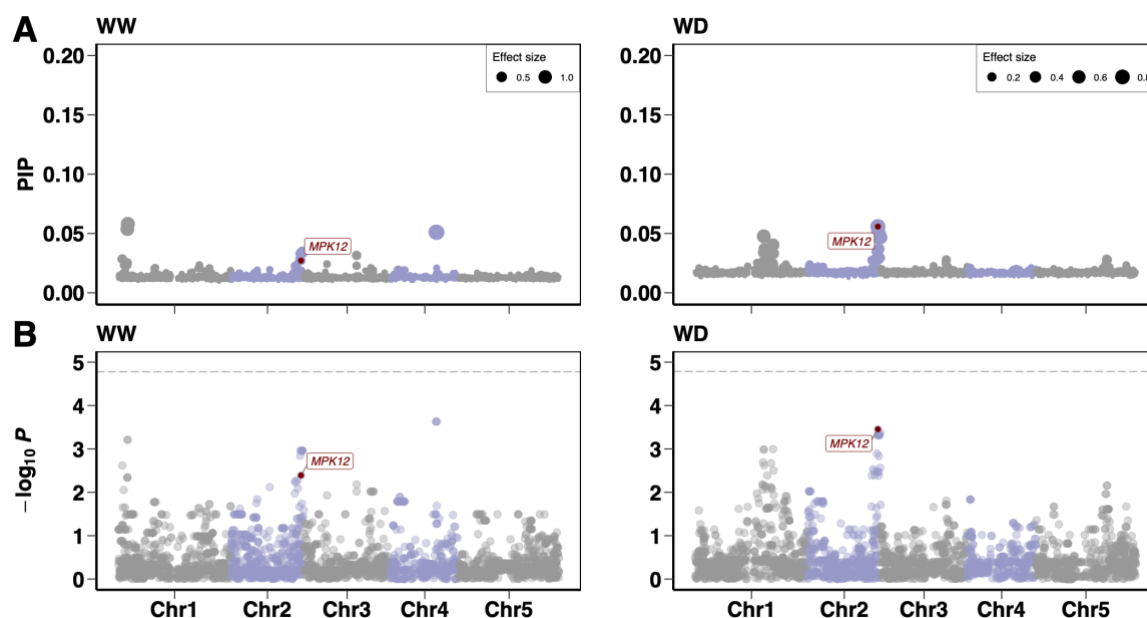
Supplementary Figure S7. (A) Polygenic modeling of stomatal conductance variation in the Santo Antão A. thaliana population in well-watered (WW) and water deficit (WD) conditions using BSLMM. The y-axis represents the posterior inclusion probability (PIP) and the size of symbol denotes the effect size. (B) Genome-wide association mapping of stomatal conductance variation in WW and WD conditions using LMM. The horizontal dashed line corresponds to the Bonferroni significance threshold at α = 0.05. In A and B, points represent SNPs along the five chromosomes. The red point represents the MPK12 G53R variant (a substitution of arginine for glycine at amino acid position 53).

Supplementary Figure S8. *The allelic effects of MPK12 G53R on phenotypic plasticity, (A) water use efficiency (WUE) and (B) stomatal conductance, among Santo Antão A. thaliana population under well-watered (WW) and water deficit (WD) conditions. The WUE is measured as carbon isotope discrimination ($\delta^{13}C$), and the carbon isotope ratio is expressed as (per mil, ‰). The points represent the means and the whiskers represent the 95% CI.*
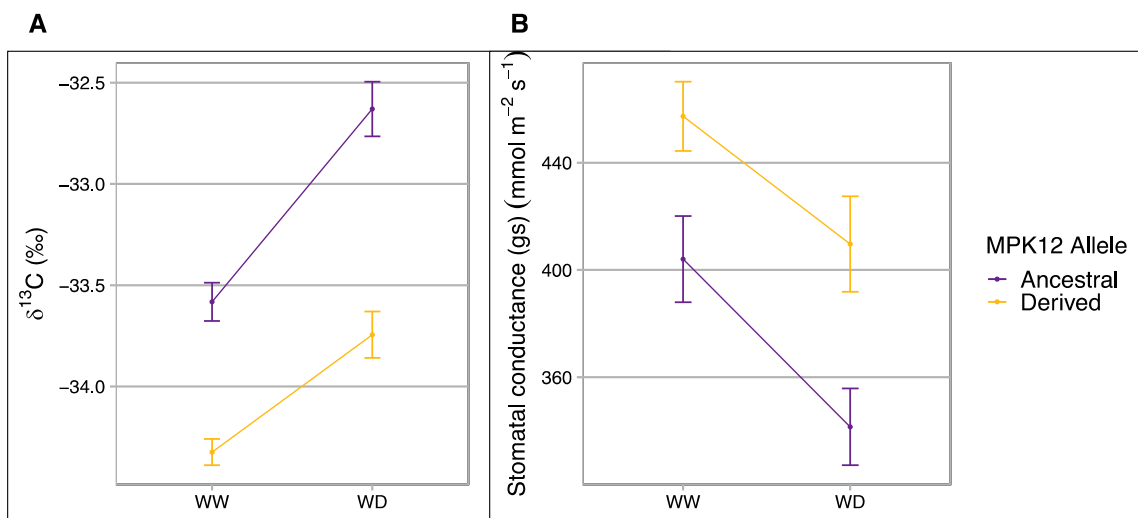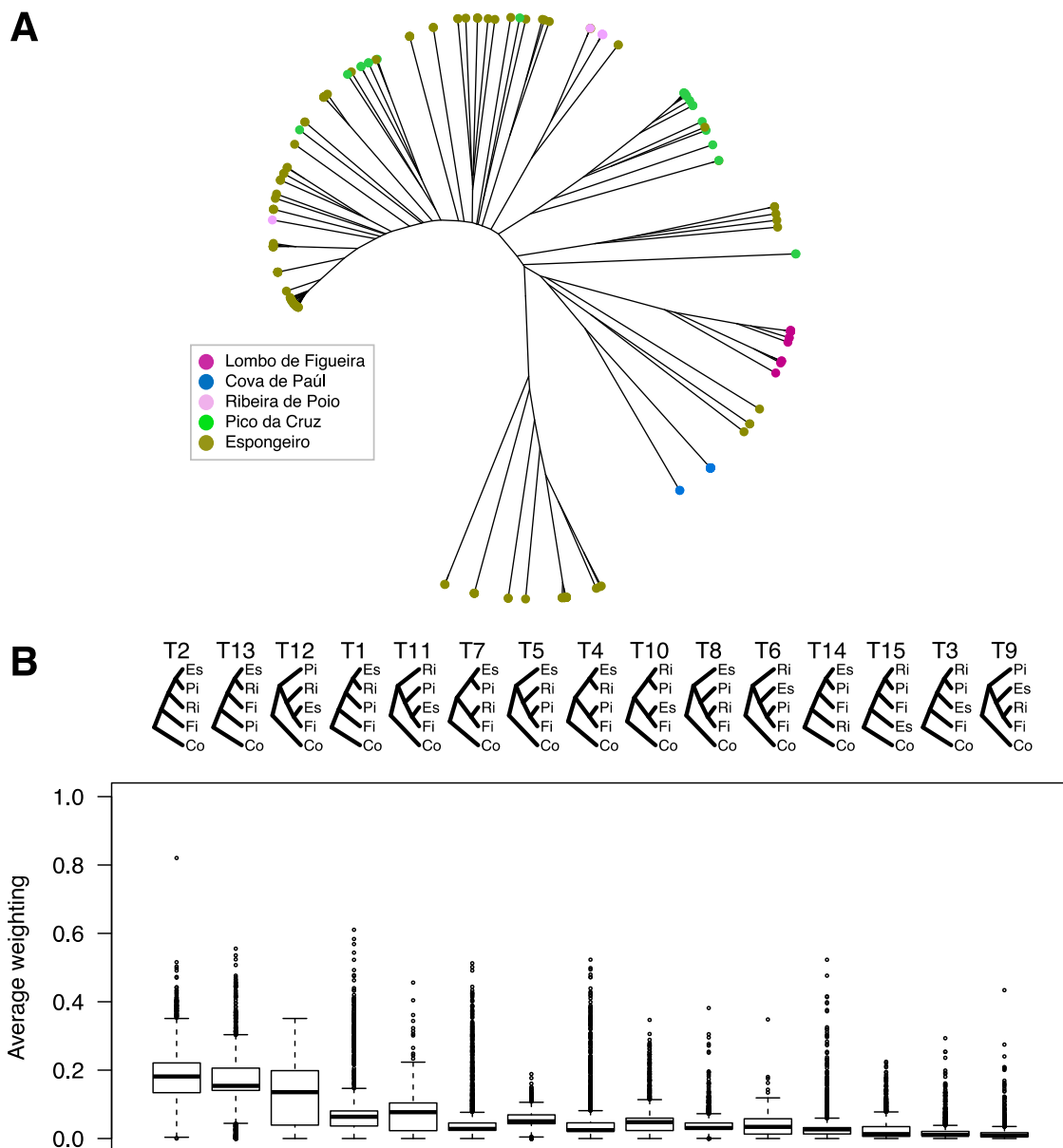
A

B

Supplementary Figure S9. History of the Santo Antão A. thaliana population. *(A) Neighbor-joining tree of Santo Antão population (n=189). Colors represent five sub-populations: Lombo de Figueira (in magenta), Cova de Paúl (in blue), Ribeira de Poio (in amethyst), Pico da Cruz (in green), and Espongeiro (in mustard). (B) Genome-wide Twisst analysis showing the average topology weighting (y-axis) among the fifteen possible phylogenetic topologies using sliding windows containing 50 variable sites. The trees are ordered based on their average topology weightings with the highest topology weighting for topology 2 (T2).*
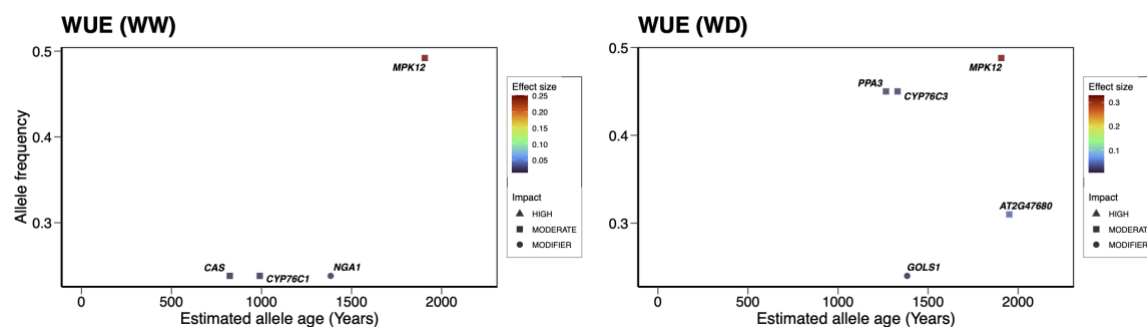
Supplementary Figure S10. *Relationship between estimated allele ages (inferred in RELATE) and allele frequencies of variants with major effects estimated with polygenic GWA mapping for water use efficiency (WUE) in well-watered (WW) and water deficit (WD) conditions in the Santo Antão population. Color denotes variant effect size from the BSLMM and shape denotes predicted impact from gene annotation.*
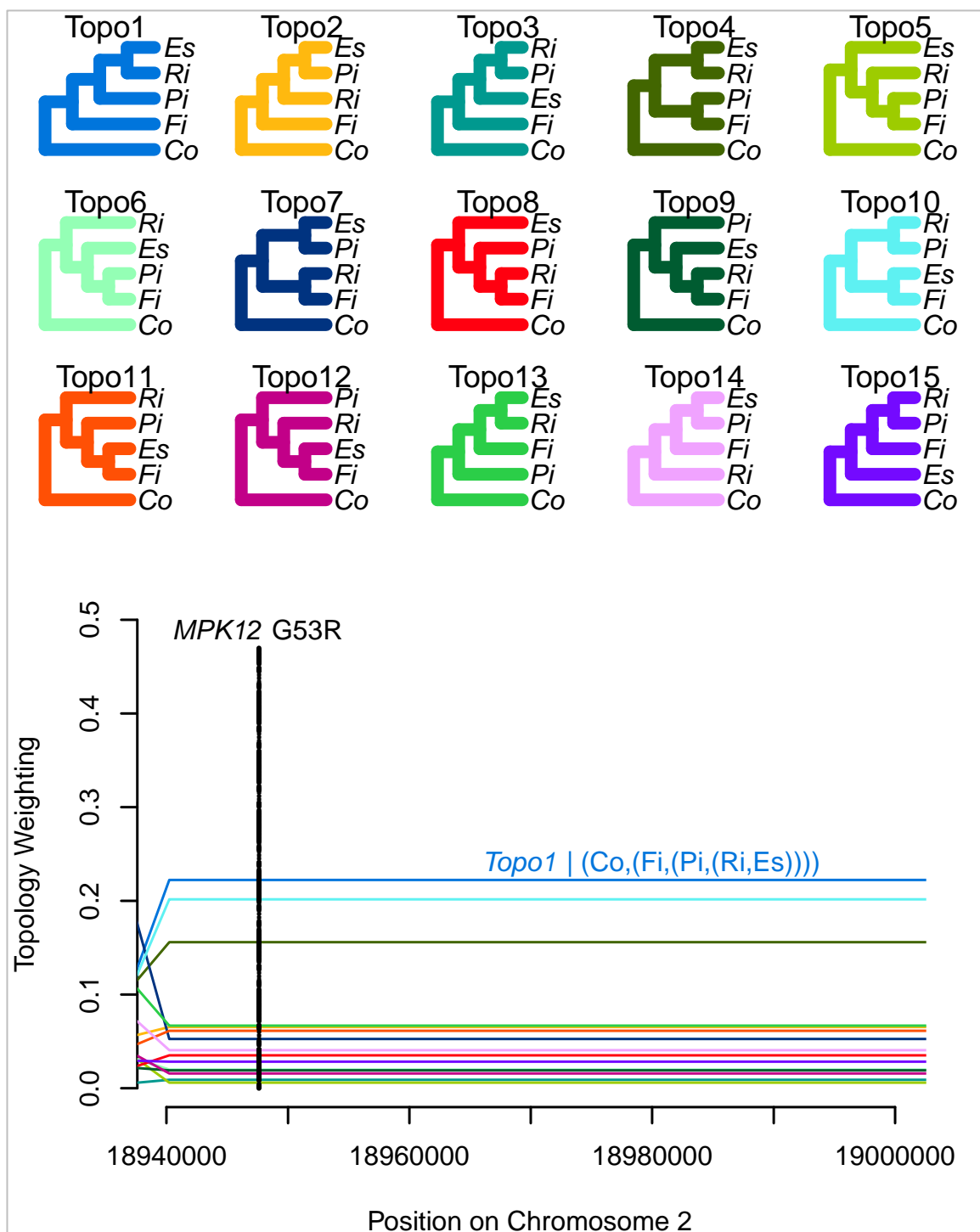
Supplementary Figure S11. *Genome-wide phylogenetic analysis inferred by Twisst, focusing on the MPK12 genomic region. Topology 1 (Topo1) has the highest topology weighting (Bottom panel) for the region surrounding MPK12 G53R among the fifteen possible phylogenetic topologies (Top panel; as previously shown in Supplementary Figure. S9B).*
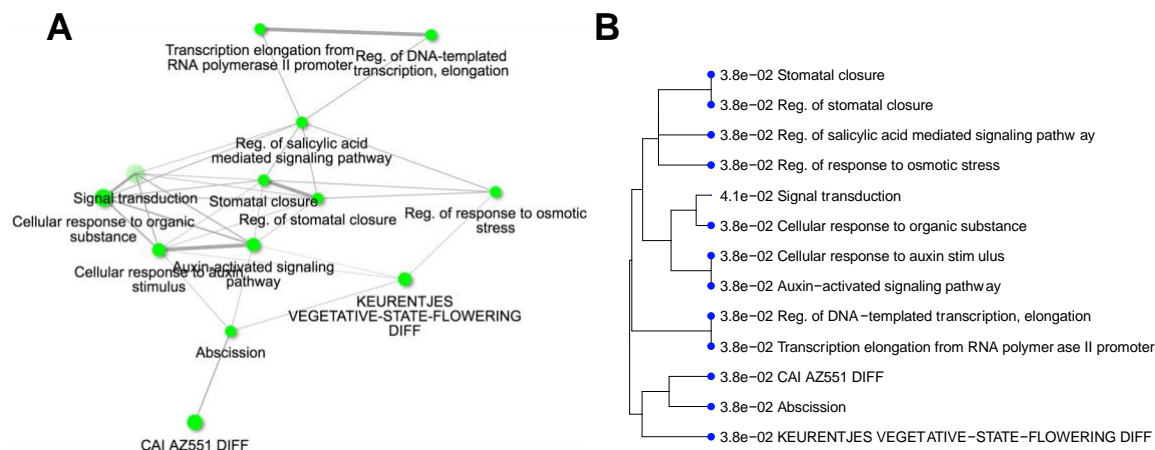
Supplementary Figure S12. *Gene ontology (GO) analysis of the top 1% of genes identified through the genome-wide integrated haplotype score (iHS) across the genomes of the Santo Antão population. (**A**) Significant enriched GO molecular component terms were visualized as a network, in which related GO terms are connected by a line, whose thickness reflects percent of overlapping genes. The size of the node corresponds to number of genes. (B) A hierarchical clustering tree shows the relatedness between GO terms, in which related GO terms are grouped together based on how many genes they share. The size of the solid circle corresponds to the enrichment FDR (see results in supplementary table S5, Supplementary Material online).*
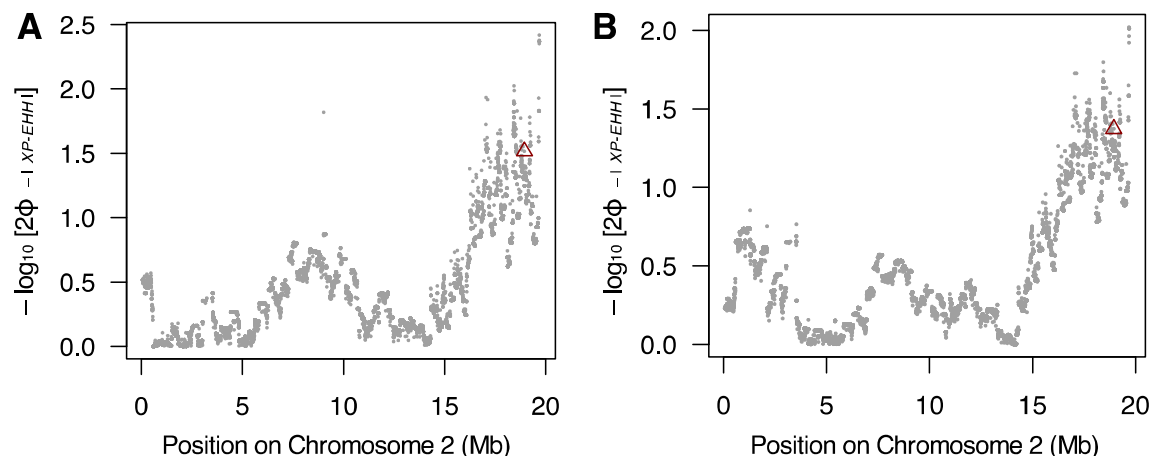
Supplementary Figure S13. *The cross-population extended haplotype homozygosity (XP-EHH) analyses for between Cova de Paúl and Espongeiro (A) and between Pico da Cruz and Espongeiro (B). A dark red triangle marks the derived MPK12 53R variant.*
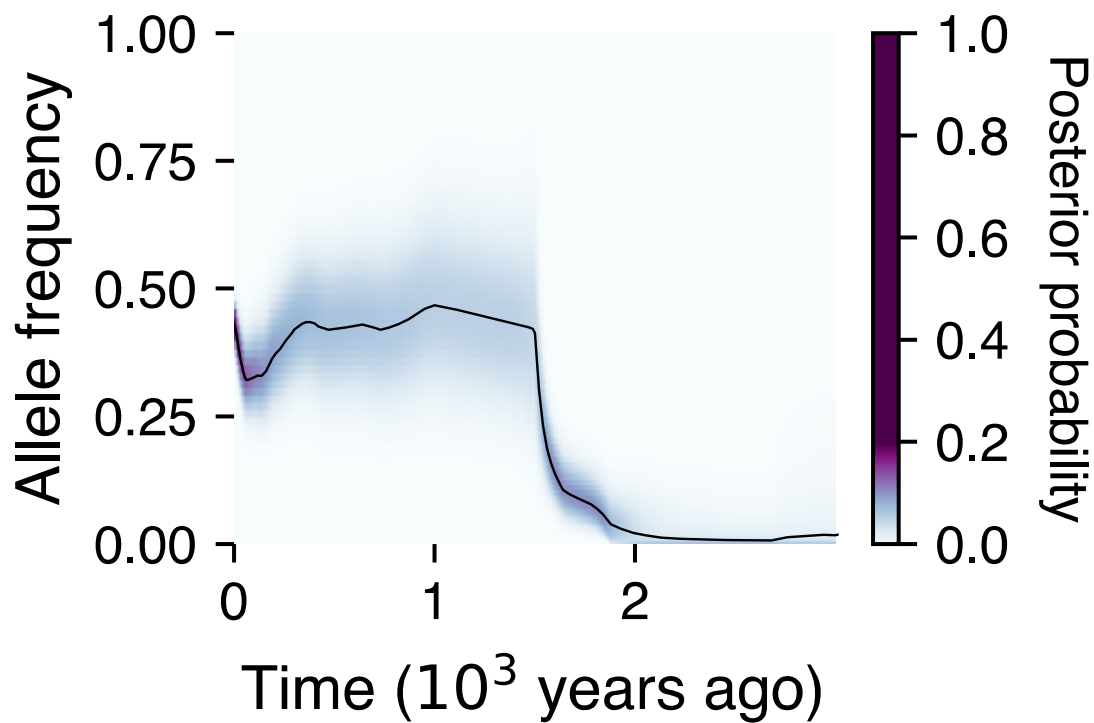
Supplementary Figure S14. *Inferred allele frequency trajectory for the derived MPK12 allele variant in Santo Antão. The black line corresponds to the inferred allele frequency change over time and the surrounding color to the posterior probability.*
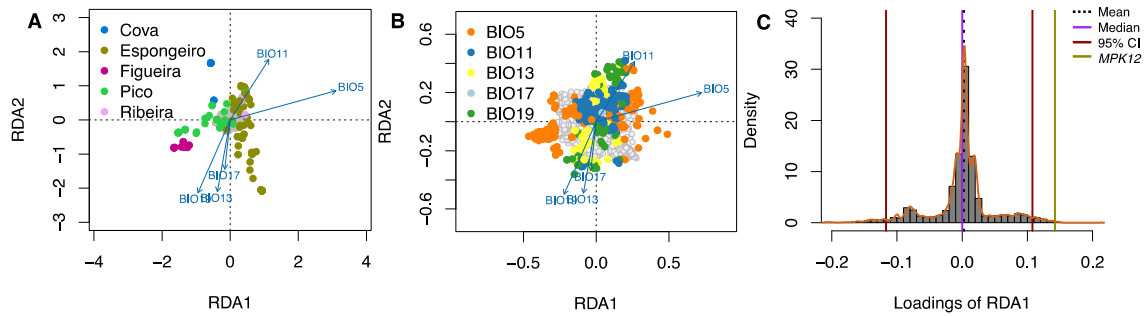
Supplementary Figure S15. Redundancy Analysis (RDA) for Santo Antão population. *(A) A biplot summarizing the result of full RDA on the bioclimatic environmental variables among the five sub-populations in Santo Antão. Individual scores for RDA axes are represented by dots, colored as the same coloring scheme for the previous figures (Lombo de Figueira: magenta, Cova de Paúl: blue, Ribeira de Poio: amethyst, Pico da Cruz: green, and Espongeiro: mustard). Grey dots represent SNPs. (B) A biplot summarizing the result of full RDA among the five bioclimatic variables for the outlier SNPs detected with the RDA outlier detection approach. The coloring scheme is for the SNPs based on the bioclimatic variable that is most strongly correlated with, and grey dots represent the non-candidate SNPs. In A and B, the light blue arrows represent the bioclimatic variables, where the arrow length reflects the amount of variance in SNP genotypes explained by that bioclimatic variable, and the angles represent the correlation between bioclimatic variables. The illustrated axes were statistically significant (P < 0.05). (C) Histogram of the SNP loadings on RDA1. Dashed black, purple, and red lines represent the mean, the median, and 95% of confidence intervals (CI), respectively. The yellow line represents the loading of the MPK12 derived allele variant on the tail of RDA1. Abbreviations:* **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.*
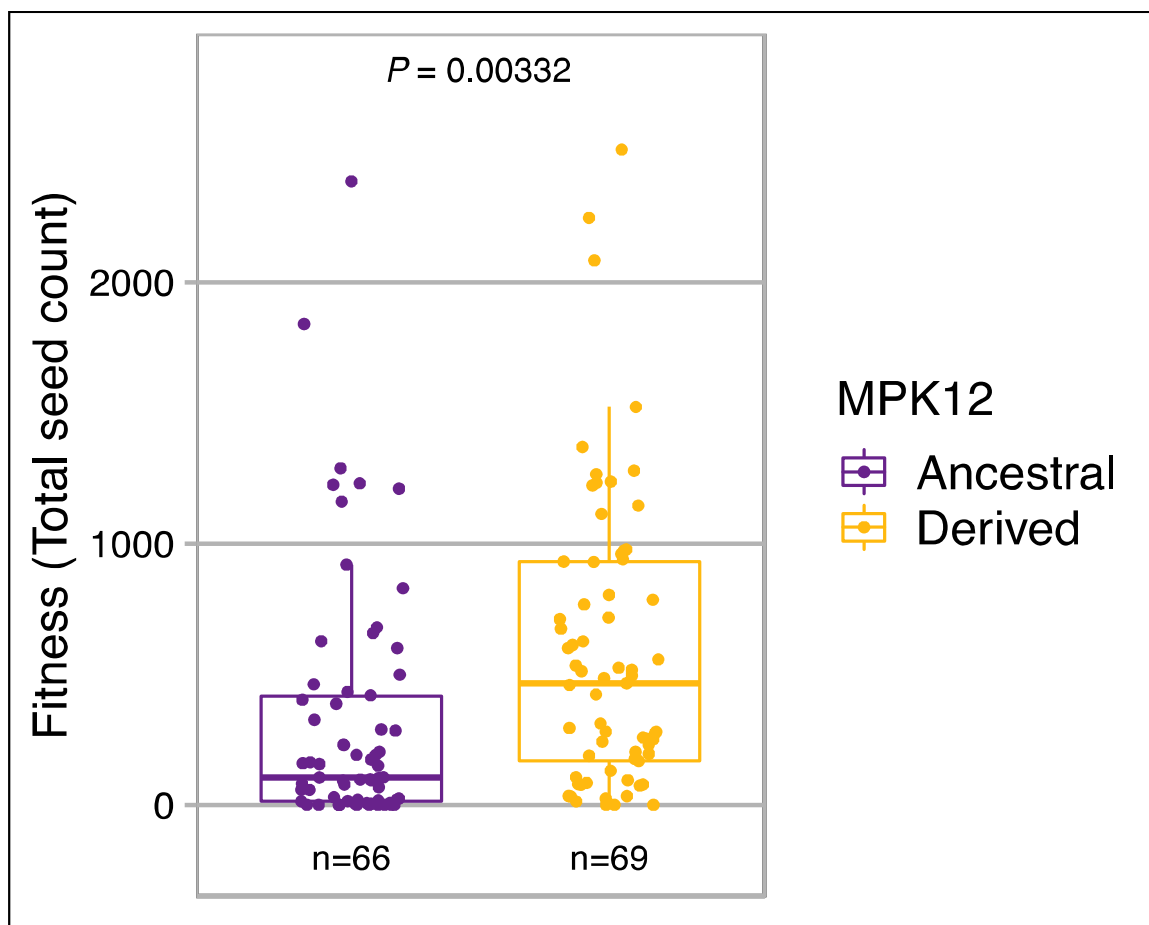
Supplementary Figure S16. *Phenotypic effects of MPK12 variants on fitness (total seed count) among Santo Antão A. thaliana population under simulated Cape Verde growth conditions. The P-value represents the fitness effect of the derived MPK12 allele, using a negative binomial GLM model (supplementary table S11, Supplementary Material online). The line in the center of the boxplots represents the median, the box edges represent the 25th and 75th percentiles (lower and upper bound, respectively), and the whiskers represent 95% CI. Each dot represents the median across four replicates per accession.*

Chapter 3

Dissecting the genetic basis of phenotypic adaptation underlying drought in a colonizing *Arabidopsis thaliana* lineage

Ahmed F. Elfarargi,[1] Elodie Gilbault,[2] Nina Döring,[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

3.1   Abstract

Environmental factors exert significant selective pressures on plants, necessitating adaptation to optimize water use efficiency for survival. Modulating the trade-off between the need to open stomata for photosynthesis and close stomata to minimize water loss is crucial for plants to thrive in their specific environments. Complex traits, such as growth rate, leaf color, and stomatal traits, are likely governed by numerous loci distributed across the genome. However, the genetic basis and adaptive evolution of these traits in natural populations in response to environmental stressors remain largely unexplored. Island ecosystems, such as the Cape Verde Islands (CVI), offer valuable opportunities to investigate the genetic basis of adaptation in plants. CVI plants experience extended dry periods with limited and highly variable rainfall, mainly driven by humid trade winds. To determine how plants adapt to this precipitation regime, we examined the genetic architecture of variation in growth rates, leaf color, and stomatal patterning in response to precisely controlled water conditions among CVI *A. thaliana* populations. Genome-wide association mapping analyses revealed that moderately complex genetic architectures with roles for several major effect variants underlie variation in these traits. Furthermore, we found that several identified genes through genetic mapping have pleiotropic functions for complex traits underlying drought stress, highlighting the intricate nature of plant adaptation to these challenging conditions.

Keywords: *Arabidopsis thaliana*, CVI, Phenoscope, growth rate, stomata, drought, GWA, adaptive evolution

3.2 Introduction

Plants are regularly exposed to diverse environmental stresses throughout their life cycle, necessitating the development of specialized mechanisms to adapt to such challenges (Lawlor, 2013; Chaves *et al.*, 2003). Water limitation is a critical environmental factor that limits plant growth and productivity (Boyer, 1982). As a result, water availability plays a crucial role in shaping plant biology, exerting significant selective pressure on the evolution of plant development, morphology, and physiology (Stebbins, 1952; Bohnert *et al.*, 1995). Consequently, plants have evolved three primary strategies to optimize their growth and fitness in environments characterized by limited water availability (Ludlow, 1989). First, drought tolerance is achieved by improving water use efficiency (WUE), which refers to optimizing carbon assimilation through photosynthesis relative to water lost via stomatal transpiration (SCOTT, 2000). Second, drought avoidance involves enhancing water uptake and reducing transpiration rates to conserve water (Levitt, 1972). Lastly, drought escape is facilitated through accelerated flowering, enabling plants to complete their life cycle and secure successful reproduction before the onset of severe drought conditions (McKay *et al.*, 2003; Sherrard and Maherali, 2006). These strategies involve intricate physiological processes and morphological traits governed by numerous genes.

Arabidopsis thaliana, a model organism for higher plants, thrives in diverse habitats with considerable variation in water availability and other environmental factors. Early research on *A. thaliana* emphasized the significance of natural variation within the species as a valuable tool for understanding the connection between molecular perturbations and phenotypic changes (Somerville and Koornneef, 2002). Local adaptation to different environmental conditions, particularly precipitation and humidity, has influenced allele frequency distributions across its extensive range (Hancock *et al.*, 2011). *A. thaliana* has served as a useful model for investigating the molecular genetic basis of complex traits (Mitchell-Olds and Schmitt, 2006). Such studies can enrich our understanding of the molecular changes at many levels associated with plant adaptation in a wide range of natural environments (Borevitz and Nordborg, 2003).

Plant growth is considered a complex trait controlled by many factors, including genetic elements (G), environmental (E) and geographical components, GXE, and developmental stage (Bac-Molenaar *et al.*, 2015; Bouteillé *et al.*, 2012; Zhang *et al.*, 2012; Clauw *et al.*, 2015). Studying plant growth adaptation to drought stress requires a precise experimental setup where soil water content and drought exposure timing are tightly controlled. This kind of experiment can be achieved through an automated phenotyping platform system equipped with the required elements for watering, rotation to avoid the position effect, and an imaging unit for collecting pictures of every plant at a specific time of day throughout the experiment. Several platforms have been developed for controlled drought phenotyping (Bac-Molenaar *et al.*, 2015; Bouteillé *et al.*, 2012; Zhang *et al.*, 2012; Clauw *et al.*, 2015), including the Phenoscope platform (Tisné *et al.*, 2013). Phenoscope has advantages over the other existing platforms, which is the property of precisely rotating more than 700 individual plants with automatic watering adjustments (Tisné *et al.*, 2013).

Here, we examine phenotypic and genetic changes in drought response in an *A. thaliana* lineage from the Cape Verde Islands (CVI) after colonization from North Africa. The Cape Verde Islands are an isolated archipelago that lies 570 km off the west coast of Senegal. A single *A. thaliana* accession, Cvi-0, was collected in CVI 40 years ago (Lobin W, 1983) and has been studied intensively as a member of diverse panels of natural accessions and as a parent in recombinant mapping populations. Cvi-0 is divergent from other well-studied accessions for a range of traits, including transpiration efficiency and WUE (McKay *et al.*, 2003; Juenger *et al.*, 2005; McKay *et al.*, 2008; Monda *et al.*, 2011; Christman *et al.*, 2008), stomatal conductance and stomatal aperture (Juenger *et al.*, 2005; Monda *et al.*, 2011; Aliniaeifard and van Meeteren, 2014; Jakobson *et al.*, 2016), stomatal density (Xu and Zhou, 2008), flooding tolerance (Vashisht *et al.*, 2011), and flowering time (Alonso-Blanco *et al.*, 1998; Kenney *et al.*, 2014; Alonso-Blanco *et al.*, 2016). A recent study elucidated the natural variation in the complex shoot growth architecture in the CviXCol population mapping under mild drought stress response using the Phenoscope platform (Tisné *et al.*, 2013; Marchadier *et al.*, 2019). We recently collected *A. thaliana* from the two Cape Verde

Islands, where it is found: Santo Antão and Fogo. We previously showed that, as a result of strong colonization bottlenecks, CVI island populations are as divergent from the mainland and from each other as any pair of species in the genus (Fulgione *et al.*, 2022) and that the genetic variation in CVI derives almost exclusively (99.9%) from new mutations that occurred after initial colonization (Fulgione *et al.*, 2022). In CVI, the colonizing populations were exposed to an arid climate with limited rainfall and a long dry season, where most of the moisture delivered to the plants comes in the form of high humidity (Brochmann *et al.*, 1997).

Water availability is essential to nearly all aspects of plant biology, development, and physiology (Stebbins Jr, 1952; Bohnert *et al.*, 1995), and variation in water availability is thus expected to exert a strong selective force. Under water-limited conditions, most plants often follow a drought avoidance strategy involving partial stomatal closure to maximize the ratio of carbon gained for photosynthesis to water loss, resulting in WUE optimization (Schulze, 1986; Meinzer, 1993; Price *et al.*, 2002). However, the high humidity in CVI creates a different situation in which moisture is mainly delivered by humid tradewinds. Tuning WUE is crucial for whole-plant fitness in an arid environment (Kenney *et al.*, 2014).

Stomata have pairs of guard cells that control stomatal apertures for gas exchange capacity regulation that take up atmospheric CO_2 and release oxygen and excess water vapor. This gas exchange process between plants and the environment is defined as stomatal conductance (gs) (Tanaka *et al.*, 2014; Horie *et al.*, 2006; Evans *et al.*, 2009). Plants can optimize their stomatal conductance in responding to environmental changes either in the short term by opening and closing stomata or in the long term by epidermal architecture modification through developmental changes in stomatal density (Farquhar and Sharkey, 1982). Several studies have demonstrated that stomatal density and stomatal index (the ratio of a total stomatal number to total cell number in a specific leaf area) increase when plants experience water stress (QUARRIE and JONES, 1977; Spence *et al.*, 1986; Martínez *et al.*, 2007). Depending on the severity of drought, prior research has

indicated that stomatal density increases under moderate drought but decreases under severe drought (Xu and Zhou, 2008). This response to both stress levels could be attributed to a reduction in leaf area leading to increased stomatal density under moderate drought, while the inhibition of guard cell division under severe drought results in decreased stomatal density (Xu and Zhou, 2008). The Cvi-0 accession exhibits more open stomata and higher stomatal conductance than the widely studied Col-0 (Monda et al., 2011).

A previous study on Eurasian accessions of *A. thaliana* revealed a negative correlation between stomatal size and density (Dittberner *et al.*, 2018). Both stomatal size and density are believed to influence stomatal conductance (Franks *et al.*, 2009). Additionally, the distance of the gas molecule diffusion path for a stomatal pore is inversely related to stomatal conductance based on stomatal size for a constant stomatal pore area. Consequently, larger stomata with increased pore depth have a more extended gas exchange diffusion path than smaller stomata (Franks and Beerling, 2009). In response to environments with elevated atmospheric CO_2 levels, plants display a low density of large stomata. Conversely, decreased atmospheric CO_2 levels trigger increased stomatal density and reduced stomatal size to enhance stomatal conductance for optimal carbon gain (Franks and Beerling, 2009).

In this study, we measured several phenotypic traits in response to precisely controlled drought stress in the natural population from Santo Antão, the island where the accession Cvi-0 was collected, as well as populations from Fogo. We compared these two Cape Verde Island populations to the closest outgroup population, which originates from Morocco. We found that the phenotype distributions for the islands are markedly shifted compared to that of Morocco. We further conducted genome-wide association (GWA) mapping to unravel the genetic basis of phenotypic adaptation underlying drought stress in the CVI *A. thaliana* population and to compare it to the Morocco *A. thaliana* population. Phenotypic traits were used as inputs for univariate and multivariate GWAS models to

identify drought stress-specific associations. Overall, these findings advance our knowledge of the role of selection in adaptations to novel environments.

3.3   Results

3.3.1 Population structure based on genome-wide genetic variation

In the present study, we analyzed a collection of *A. thaliana* accessions from two Cape Verde islands, comprising 152 from Santo Antão, 119 from Fogo, and a representative outgroup panel of 64 Moroccan accessions (***Figure 1a***). The Neighbor-joining tree and principal component analysis (PCA) depicted in ***Figure 1b-c*** illustrate the clustering of populations, revealing a distinct grouping of Santo Antão and Fogo *A. thaliana* populations relative to the Moroccan population (***Figure 1b-c***). These findings are in line with previous results, which indicate that the islands diverged without any subsequent admixture between them (Fulgione *et al.*, 2022; Tergemina *et al.*, 2022).
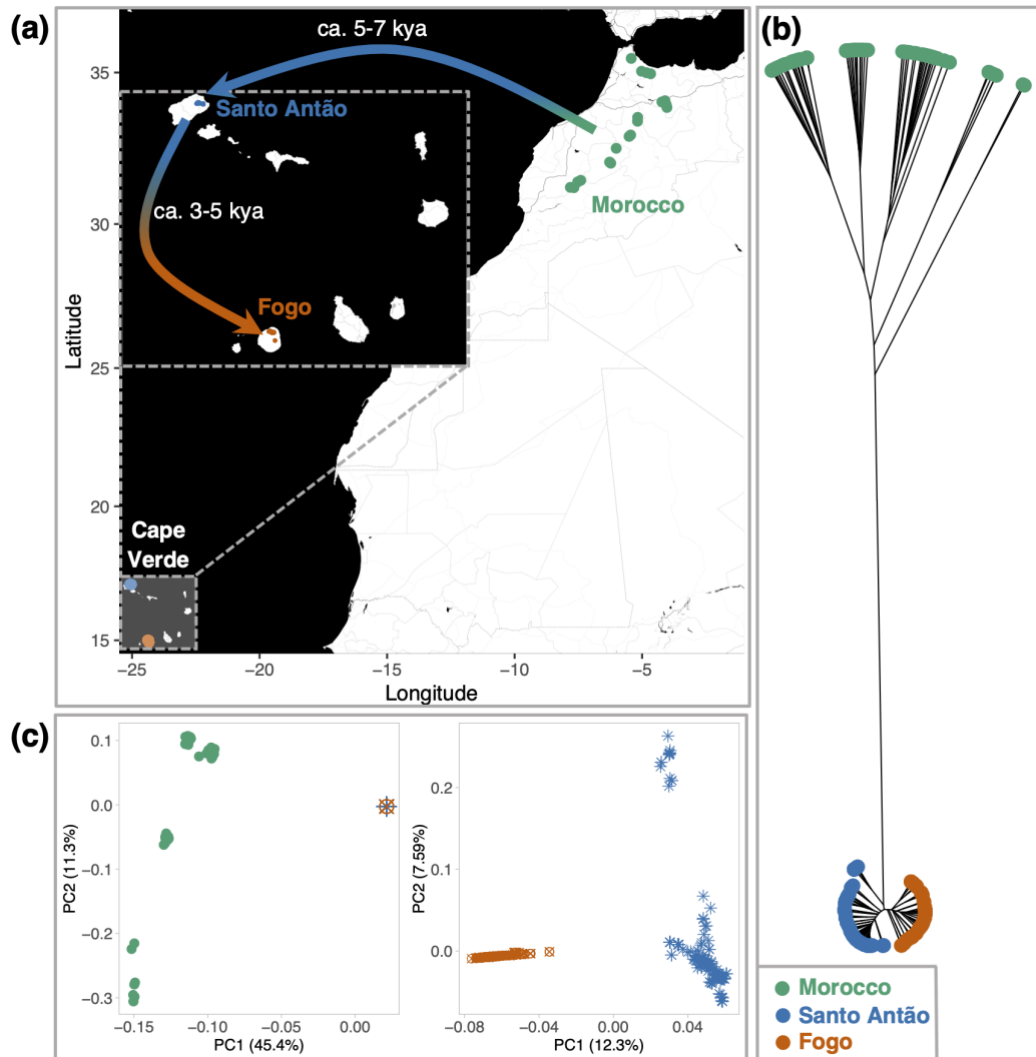
Figure 1. Geographical distribution and genetic structure of the Cape Verde Islands (CVI) and Morocco. (a) Geographic locations populations of A. thaliana in CVI (Santo Antão and Fogo) and the outgroup panel of Moroccan A. thaliana accessions. (b) Phylogenetic tree (genome-wide SNPs) showing the genetic relationship between CVI and Morocco A. thaliana populations. (c) PCA of genome-wide SNPs between CVI and Morocco A. thaliana populations (left) and within the CVI A. thaliana population (right).

3.3.2 Natural variation in drought responses in CVI relative to Morocco

We explored the trait variation in response to drought among CVI islands and the nearest outgroup Moroccan *A. thaliana* populations utilizing the high-throughput phenotyping platform, Phenoscope (Tisné *et al.*, 2013). The implementation of this platform facilitated the automated rotation of pots and adjustment of watering frequencies multiple times daily based on pot weight, enabling experiments that would be impractical through manual methods (Tisné *et al.*, 2013).

3.3.2.1 *Phenotypic variation of rosette-related traits in CVI relative to Morocco*

Our findings revealed that drought impacted plant growth and physiology in Santo Antão and Fogo populations compared to the Moroccan *A. thaliana* population (***Figure 2***). Notably, CVI plants exposed to water deficit (WD) exhibited reduced rosette size, diminished growth rates (***Supplementary Figure 1***), and leaf color-related traits (hue and saturation) relative to those under well-watered (WW) conditions (***Figure 2***). These findings suggest that further investigation into the genetic basis of these differences may offer valuable insights into plant adaptation to drought stress.

3.3.2.2 *A reduction in water use efficiency in CVI relative to Morocco*

In our previous research, we discovered that water use efficiency (WUE) was lower in the Santo Antão *A. thaliana* population compared to the Moroccan population, both under well-watered conditions and in response to drought (Elfarargi *et al.*, 2023). In the present study, we investigated the WUE variation under precisely controlled water conditions on the other island in Cape Verde, Fogo, to study the trait shift compared to Santo Antão and the mainland. Our findings revealed that WUE was higher in the Fogo population compared to Santo Antão but still lower than in the Moroccan population (***Supplementary Figure 1***). This observed variation in WUE across the populations highlights the need to

further explore the underlying genetic factors and environmental influences that shape these phenotypic differences, particularly in the context of the parallelism between Santo Antão and Fogo islands relative to Morocco and the more extreme phenotype observed in Santo Antão.
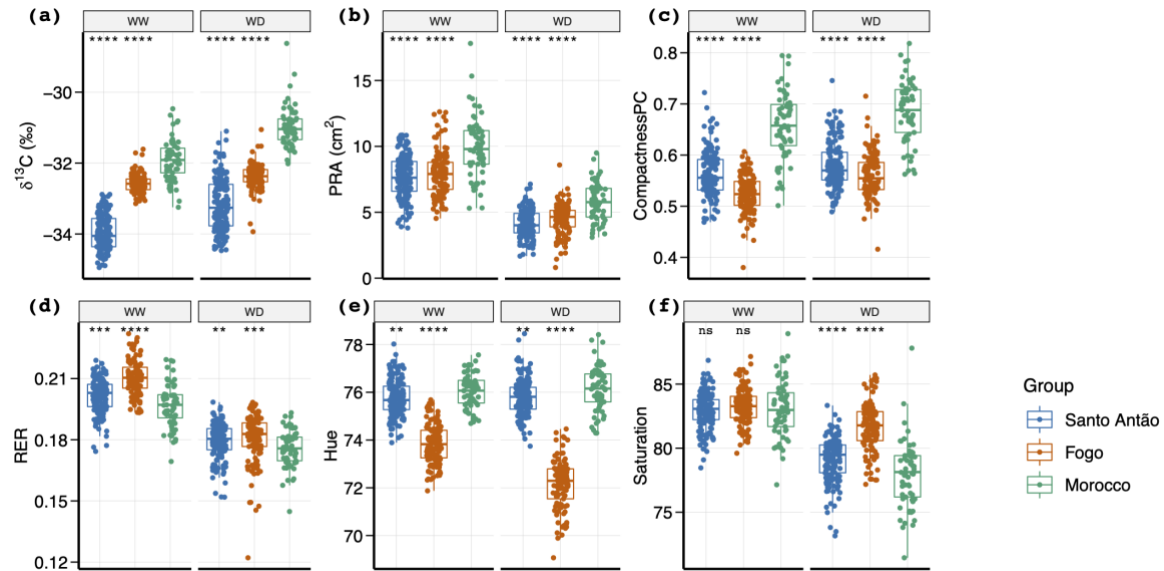
Figure 2. Phenotypic variation in morphological and physiological traits for Santo Antão and Fogo A. thaliana populations relative to Moroccan population in well-watered (WW) and water deficit (WD) conditions. The line in the center of the boxplots represents the median, the box edges represent the 25th and 75th percentiles (lower and upper bound, respectively), and the whiskers represent the 95% CI. (a) Water use efficiency (WUE) is measured as carbon isotope discrimination ($\delta13C$), and the carbon isotope ratio is expressed per mil, ‰. (b) Projected rosette area (PRA). (c) CompactnessPC (ratio of PRA (P) to surface of convex hull area (C)). (d) Relative expansion rate (RER), the growth rate of PRA; integrated over the 18-31 Days' time window. (e-f) Hue is a proxy for chlorophyll content, and saturation represents color intensity. Statistical significance was determined through the Mann-Whitney-Wilcoxon test, using the Moroccan population as a reference. Significance levels are denoted by asterisks, with a single asterisk (*) for $p \leq 0.05$, double asterisks (**) for $p \leq 0.01$, triple asterisks (***) for $p \leq 0.001$, and quadruple asterisks (****) for $p \leq 0.0001$. Non-significant results are indicated by "ns" for $p > 0.05$.

3.3.2.3   Phenotypic variation of stomatal traits in Santo Antão relative to Morocco

To better understand the large difference in WUE between Santo Antão and Morocco, we investigated additional stomatal traits, stomatal density (SD), and stomatal pore width (SPW) in Santo Antão and Morocco populations. We found that SD increased and SPW decreased under drought conditions for both populations (***Figure 3***). Although we did not observe a significant shift in SD between the populations under either condition, we

identified a substantial increase in SPW (Mann-Whitney-Wilcoxon (MWW) test; *p*-value=0.013) in the Santo Antão population compared to the Moroccan population (**Supplementary Figure 1b-c**). Further investigation into the genetic and physiological factors responsible for these differences could provide valuable insights into the underlying mechanisms of WUE variation between populations.
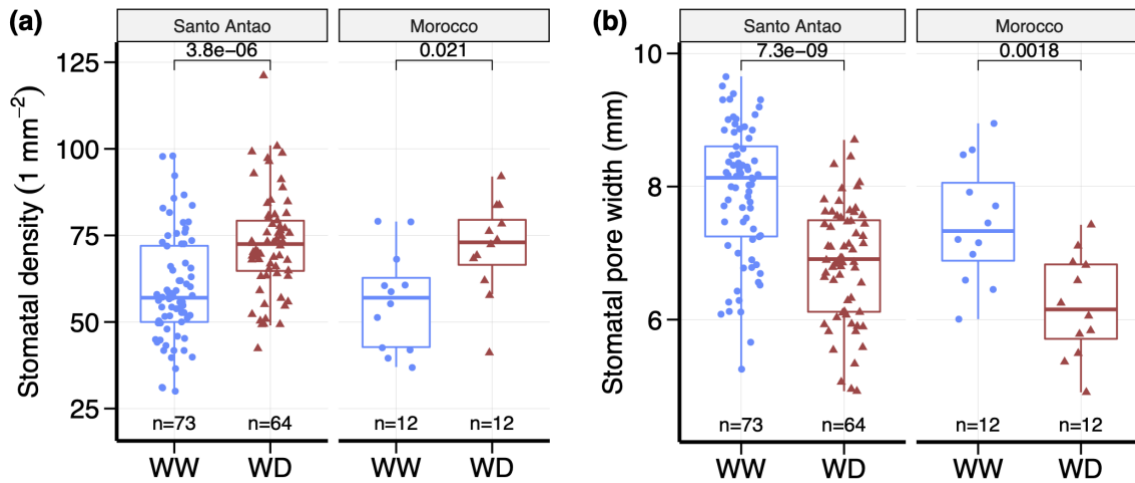
Figure 3. Phenotypic variation in stomatal density (a) and stomatal pore width (b) for Santo Antão and Moroccan A. thaliana populations in well-watered (WW) and water deficit (WD) conditions. The line in the center of the boxplots represents the median, the box edges represent the 25th and 75th percentiles (lower and upper bound, respectively), and the whiskers represent the 95% CI. The P-values for the Mann-Whitney-Wilcoxon test are shown.

3.3.3 Correlations among drought-related traits

We subsequently assessed the genetic correlation between the average traits across treatments in Santo Antão (**Figure 4**), Fogo (**Supplementary Figure 2a**), and Morocco (**Supplementary Figure 2b**). We found that certain drought-related traits exhibit genetic correlations, implying shared genetic factors. For instance, we identified a positive genetic correlation between leaf color intensity (saturation) and rosette growth-related traits, which indicates that similar genetic mechanisms govern these traits. Conversely, we observed a negative genetic correlation between SPW and both rosette growth and other stomatal-related traits, suggesting that selection for increased SPW could result in diminished WUE and growth rate. Moreover, we found that hue (a proxy for chlorophyll content) was negatively correlated with growth rate but positively correlated with WUE and rosette compactness. These findings provide valuable insights into the complex

genetic relationships governing drought-related traits and their potential implications in plant adaptation to drought conditions.
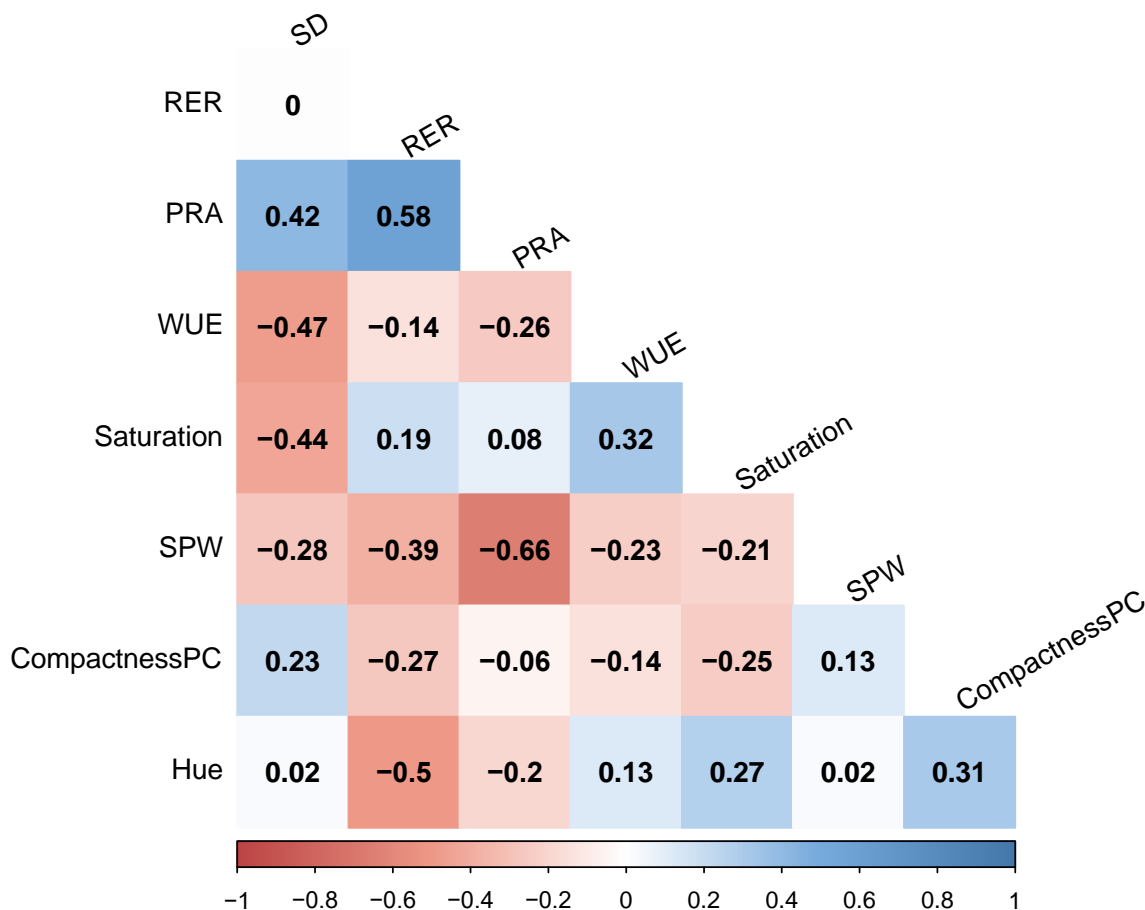
Figure 4. Pairwise Pearson genetic correlation of the average phenotypic traits in Santo Antão. PRA, projected rosette area; WUE, water use efficiency; SD, stomatal density; SPW, stomatal pore width; RER, relative expansion rate. The color spectrum, dark red to dark blue, represents highly negative to highly positive correlations, and the number represents the correlation values.

3.3.4 Heritability of drought-related traits

We estimated trait heritabilities and the number of loci contributing to trait variation (*Table 1*). First, we assessed broad-sense heritability (H^2) for each trait in each population in both WW and WD conditions based on repeatability across replicates. Heritability estimates varied across treatments and ranged from moderate to high for both treatment conditions among the three populations. The highest heritability was observed for SD in Santo Antão under both WW (0.94) and WD (0.93) conditions. The heritability for

compactnessPC ranged between 0.73 and 0.88 among populations under both conditions. Growth rate (RER) exhibited low heritability under both conditions in Santo Antão and under WD condition in Morocco, but higher heritability was observed in Fogo under both conditions and in Morocco under WW conditions. Moderate heritability was observed for PRA, hue, and SPW under both conditions in Santo Antão. Heritability was high for hue (a proxy for chlorophyll content) in Fogo and Morocco under WD conditions but decreased to moderate levels in Morocco under WW conditions and in Santo Antão under both conditions. These findings highlight the diverse heritability patterns of various traits under different conditions, emphasizing the complexity of genetic factors in plant adaptation to changing environments. Next, we estimated heritability (PVE) based on the correlation between traits and a relatedness matrix based on the entire set of genotyped SNPs (a proxy for narrow-sense heritability, h^2_{snp}) in Bayesian Sparse Linear Mixed Model (BSLMM) implemented in GEMMA (Zhou and Stephens, 2012) and observed moderate to high values (***Table 1***).

Additionally, we used BSLMM to evaluate the genetic contributions of sparse (PGE; the proportion of the PVE explained by SNPs with a non-zero effect) and polygenic components (PVE; the proportion of phenotypic variance explained by the polygenic term) (***Table 1***). The PGE denotes the proportion of PVE that is accounted for by genetic variants with large effects, providing insight into the genetic architecture of the traits. We observed PVE and PGE values exceeding 30% for all studied traits. Finally, we inferred trait architecture using a polygenic GWA BSLMM to estimate the number of loci contributing to the trait. BSLMM models the polygenic architecture as a mixture of large and small effects. The heritability estimation results suggest that the genetic architecture of drought-related traits is complex, with a few major loci accounting for a large proportion of the phenotypic variance.

Table 1. Summary statistics, estimation of broad-sense heritability (H2), and genetic architecture for the average traits across treatments. Heritabilities above 0.4 are marked with (*).

Population	Trait	N	H²		PVE	PGE	BSLMM
			WW	WD			n_gamma
Santo Antão	WUE	152	NA	NA	0.77	0.72	7
	PRA	152	0.50*	0.44*	0.42	0.37	15
	CompactnessPC	152	0.75*	0.75*	0.76	0.40	5
	RER	152	0.27	0.24	0.30	0.37	17
	Hue	152	0.45*	0.48*	0.65	0.39	38
	Saturation	152	0.28	0.65*	0.72	0.51	21
	SD	64	0.94*	0.93*	0.58	0.45	13
	SPW	64	0.45*	0.44*	0.50	0.51	18
Fogo	WUE	146	NA	NA	0.74	0.68	6
	PRA	146	0.84*	0.69*	0.40	0.42	16
	CompactnessPC	146	0.74*	0.73*	0.57	0.40	44
	RER	146	0.63*	0.64*	0.37	0.41	27
	Hue	146	0.89*	0.82*	0.69	0.44	33
	Saturation	146	0.74*	0.81*	0.81	0.44	48
Morocco	WUE	61	NA	NA	0.43	0.40	16
	PRA	61	0.68*	0.38*	0.46	0.43	22
	CompactnessPC	61	0.86*	0.88*	0.84	0.31	26
	RER	61	0.52*	0.16	0.35	0.37	18
	Hue	61	0.59*	0.75*	0.70	0.43	27
	Saturation	61	0.79*	0.72*	0.49	0.48	10
	SD	12	0.91*	0.91*	NA	NA	NA
	SPW	12	0.43*	0.43*	NA	NA	NA

Variables are studied populations (Population), drought-related traits (Trait), number of individuals (N), and broad-sense heritability (H^2) under well-watered (WW) and water deficit (WD) conditions. Examined traits comprise water use efficiency (WUE), projected rosette area (PRA), compactnessPC (the ratio of PRA to convex hull area), relative expansion rate (RER), hue (a proxy for chlorophyll content), leaf color intensity (saturation), stomatal density (SD), and stomatal pore width (SPW). Results from the Bayesian Sparse Linear Mixed Model (BSLMM) implemented in GEMMA are also included, featuring the proportion of variance explained (PVE), the proportion of variance explained by sparse effects (PGE), and the posterior estimate of the number of SNPs with major effect (n_gamma (median)).

3.3.5 Polygenic basis of drought adaptation in natural populations

We conducted genome-wide association studies (GWAS) using linear mixed models implemented in GEMMA (Zhou and Stephens, 2012) to adjust for population structure by incorporating a kinship matrix as a random effect. We performed GWAS for each trait in

each of the two Cape Verde Island and the Moroccan populations. We report results for the average for each trait across both conditions and the drought response, i.e., the difference between traits in the two conditions. Afterward, we used the local score method (Bonhomme *et al.*, 2019) to improve detection power by integrating scores over regions accounting for the LD between SNPs in a short genomic region containing a causal variant. The local score method is calculated by combining association signals and finding the maximum of a Lindley process over the sequence of these scores (Bonhomme *et al.*, 2019).

Given that several of the traits studied here (water-use efficiency (WUE), stomatal density (SD), and stomatal pore width (SPW)) are related and may potentially have a shared genetic basis, we also applied a multivariate LMM (mvLMM) approach (Zhou and Stephens, 2014) to improve the statistical power of GWAS. The mvLMM GWAS approach incorporates data from multiple traits to identify additional associated loci. By employing these strategies, we could identify significant associations across the genome.

To facilitate the presentation, we first report the genetic basis of WUE in natural *A. thaliana* populations from CVI and Morocco. Subsequently, we categorize the remaining drought-related traits, including WUE, into three groups based on their correlations (***Figure 4***). The first group encompasses stomatal-related traits, such as WUE, SD, and SPW. The second group contains rosette growth-related traits, including PRA, RER, and compactnessPC. Lastly, the third group includes leaf color-related traits like hue and saturation.

3.3.5.1 Genes underlying the reduced WUE in response to drought in CVI relative to Morocco

In Santo Antão, the top peak is one that we previously reported, G53R in *Mitogen-Activated Protein Kinase12 (MPK12;* AT2G46070) (***Figure 5a***) (Elfarargi *et al.*, 2023). Here, we identified additional associations with WUE using *MPK12* G53R as a covariate in the model to account for its major effect on WUE. We found three additional genomic regions

to be significantly associated with average WUE and two additional regions significantly associated with the drought response of WUE (***Figure 5b-c***). A peak in average WUE on chromosome 2 contains the gene *VQ18* (AT2G44340) (***Figure 5b***). VQ18 is responsive to abscisic acid (ABA), negatively modulating ABA signaling through interaction with ABI5 (Pan *et al.*, 2018). The second associated region is on chromosome 3 (***Figure 5b***) and contains a variant at *DI19-3* (AT3G05700), which encodes a drought-responsive family protein and participates in plant response to drought and ABA stress (Qin *et al.*, 2014). The third associated region on chromosome 3 (***Figure 5b***) contains the *PYL8* locus (AT5G53160), which encodes RCAR3, a regulatory component of the ABA receptor, which interacts with protein phosphatase 2Cs ABI1 and ABI2 (Belda-Palazon *et al.*, 2018).

Regarding the drought response of WUE, we found a significant region at the beginning of chromosome 4 (***Figure 5c***), which contains the well-known *FRIGIDA* (*FRI*), a major determinant of flowering time in *A. thaliana* (Johanson *et al.*, 2000; Gazzani *et al.*, 2003; Shindo *et al.*, 2005; Fulgione *et al.*, 2022). The other highest peak on chromosome 5 (***Figure 5c***) contained the *NHL26* gene, a member of the *NHL* (*NDR1/HIN1-like*) family of genes involved in response to abiotic stresses in plants. *NHL26* was induced by ABA, a plant hormone that plays an important role in stress response and was highly expressed in senescent leaves, which mimic dehydration (Bao *et al.*, 2016).

In Fogo, we found that six loci explained 74% (95% CI: 52%-92%) of the genetic variance for average WUE (***Figure 5d, Table 1***), and 16 loci explained 23% (95% CI: 8%-41%) of the genetic variance for drought response of WUE (***Figure 5e***). Loci associated with the average WUE include *NRAMP2* (AT1G47240), AT5G42490, *TOC75-I* (AT1G35860), *MTP1* (AT2G46800). *NRAMP2* and *MTP1* are involved in ion homeostasis. *NRAMP2* belongs to a gene family of metal ion transporters and is essential for root growth under low Mn conditions. On the other hand, *MTP1* encodes a zinc transporter and plays a crucial role in mediating Zn ion homeostasis. These genes are essential for plant growth and development, particularly under stress conditions, as maintaining proper ion homeostasis is vital for plant survival in harsh environments. Among the associated candidates with the

drought response of WUE, *AP2C1* (AT2G30020), a PP2C-type protein phosphatase, is a particularly good candidate since it plays an important role in regulating plant stress responses by negatively controlling MAPK activation (Ayatollahi *et al.*, 2022).

In Morocco, GWAS did not yield any Bonferroni significant findings for average WUE (see ***Figure 5d***). Nevertheless, three genomic regions were significantly associated with the drought response of WUE (***Figure 5e***), and one region was significantly associated with well-watered (WW) conditions. Among the identified genes, we identified *MYB23* (AT5G40330), which is a transcription factor implicated in ABA signaling. Previous research has demonstrated that *MYB23* functions as a transcriptional activator for ABA-responsive genes, emphasizing its crucial role in modulating plant reactions to stress and other environmental factors (Abe *et al.*, 2003). Regarding WUE under WW conditions, a single Bonferroni significant peak was identified at the end of chromosome 3 (see ***Supplementary Figure 3***). This peak includes the *homeobox 1* (*ATHB-1* or *HAT5*; AT3G01470) gene, encoding a homeodomain leucine zipper class I (HD-Zip I) transcriptional activator involved in leaf and hypocotyl development. Intriguingly, *HAT5* has been found within transcription factors that exhibit differential expression at the beginning and end of guard cell differentiation (Lopez-Anido *et al.*, 2021). These findings highlight the importance of genes associated with ion homeostasis, ABA signaling, and mitogen-activated protein kinase (MAPK) activation, which are essential for plant growth and development under stress conditions.
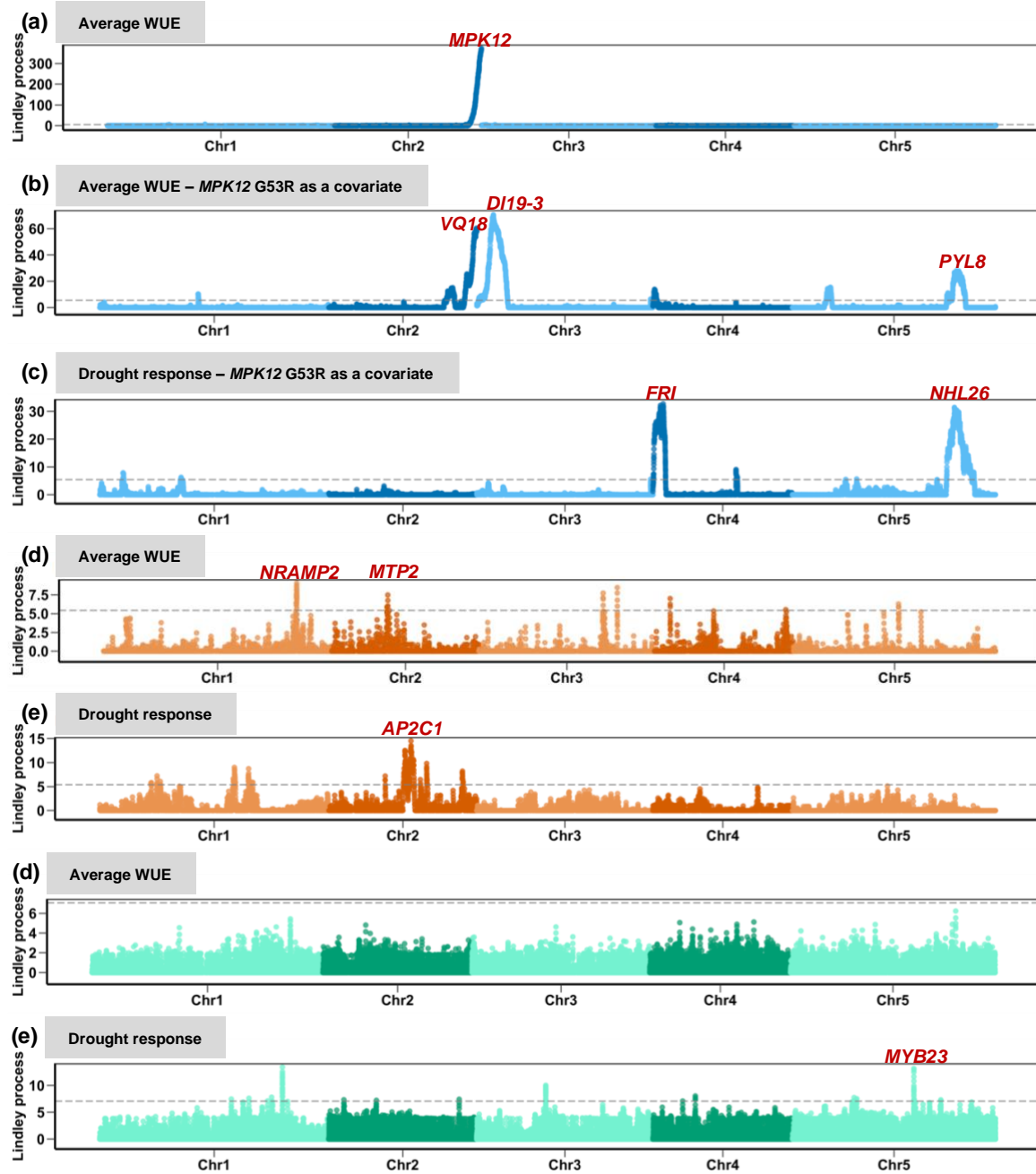
Figure 5. Genomic associations of water use efficiency (WUE) in response to drought in Santo Antão (blue), Fogo (orange), and Morocco (green) populations using LMM in GEMMA followed by the local score approach. The y-axis represents the Lindley process score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.

3.3.5.2 The genetic basis of stomatal trait variation in Santo Antão

The stomatal traits WUE, SD, and SPW were combined to form a multi-trait set because these traits are highly intercorrelated (***Figure 4***) and represent the leaf stomata as a

structural unit. Multivariate GWAS analysis, conducted on the averaged data for each of the three traits across both conditions, revealed several signals across the genome (***Figure 6***). In addition to the previously identified variant (G53R) in *MPK12* (Elfarargi *et al.*, 2023), we identified several potentially interesting associations across the genome. One of the highest peaks on chromosome 1 (***Figure 6***) contains Gibberellin (GA) 3-oxidase 2 (*GA3OX2*; AT1G80340), which is a member of the GA3OX class of enzymes. It plays a crucial role in regulating the levels of bioactive GAs in plants, which is essential for the modulation of plant vegetative growth and development (Mitchum *et al.*, 2006). Further, the *GA3OX2* gene codes for the GA3 activation enzyme, which regulates the steady-state level of DELLA proteins, and bioactive GA3 signaling depends on the proteolysis of DELLA proteins (Davière and Achard, 2013). It has been shown that the expression of GA3OX2 is down-regulated in dormant *Cvi* seeds and thermo-inhibited *Col* seeds (Cadman *et al.*, 2006). Moreover, studies have found that the *ga3ox1 ga3ox2* double mutant has a severe dwarf phenotype with strongly decreased expression levels of lignin deposition transcription factors in the stem and cell wall thickening (Mitchum *et al.*, 2006). Another peak on chromosome 3 contained *NPF8.1* (AT3G54140) (***Figure 6***), which plays a critical role in the stomatal movement in *A. thaliana* by mediating the uptake of ABA into guard cells, which in turn regulates stomatal closure in response to environmental cues such as drought (Tal *et al.*, 2016; Shimizu *et al.*, 2021). Another genomic region on chromosome 5 contains the *SUVH4* gene (***Figure 6***), which encodes a histone methyltransferase that catalyzes H3K9 methylation along with SUVH5 and SUVH6 (Du *et al.*, 2015). The methylation status of H3K9 can be modulated in response to drought stress, with decreased H3K9me2 levels observed around drought stress-inducible genes (Q., Wang *et al.*, 2021). Although a direct role of SUVH4 in drought stress response has not been explicitly shown, its involvement in H3K9 methylation suggests that SUVH4, along with other H3K9 methyltransferases and demethylases like JMJ27, may contribute to regulating drought stress responses by affecting the chromatin landscape around stress-responsive genes (Q., Wang *et al.*, 2021). In conclusion, our findings from the multivariate GWAS analysis on stomatal-related traits in Santo Antão highlight the potential involvement of variation in several key genes, such as *GA3OX2*, *NPF8.1*, and *SUVH4*, in modulating stomatal responses to drought.
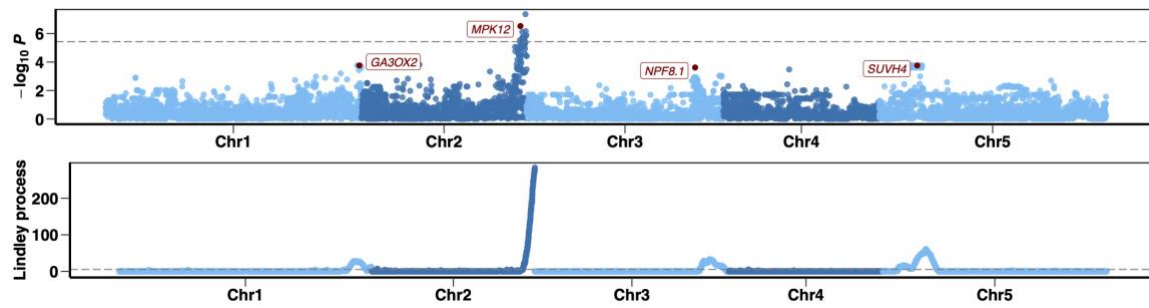
Figure 6. Genomic associations in Santo Antão for stomatal-related traits (WUE, SD, and SPW) using the multivariate model in GEMMA (upper) followed by the local score approach (bottom). The dashed line in plots corresponds to a genome-wide Bonferroni significance.

3.3.5.3 The genetic basis of rosette growth-related traits in response to drought in CVI relative to Morocco

In Santo Antão, we investigated the genetic basis of projected rosette area (PRA). Our findings revealed that 15 loci explained 42% (95% CI: 27%-59%) of the genetic variance for average PRA (***Figure 7a, Table 1***), while 15 loci explained 15% (95% CI: 3%-32%) of the genetic variance for drought response of PRA (***Figure 7b***). We identified three significantly associated regions for average PRA (***Figure 7a***). The first associated region is located on chromosome 1, including *TPST* (AT1G08030), which encodes a tyrosyl protein sulfotransferase involved in the auxin pathway. The second significant association is found in the *disproportionating enzyme 2* (*DPE2*; AT2G40840) on chromosome 2. This gene encodes a cytosolic protein with amylomaltase and transglucosidase activity and plays a crucial role in converting starch to sucrose in leaves at night via its signaling pathway (Chia *et al.*, 2004). Lastly, on chromosome 5, we identified the same variant in *SUVH4* that was obtained in the mvLMM analysis of stomatal-related traits (***Figure 6***), suggesting a potential pleiotropic effect of this gene on multiple drought-related phenotypes.

Furthermore, we found five significantly associated regions for the drought response of PRA (***Figure 7b***). Three out of five were located on chromosome 2. The first associated region contains *ABCG5* (AT2G13610) gene, which encodes the ABC-2 type transporter family protein. The second associated region contains the *AT2G38995* gene, which encodes O-acyltransferase (WSD1-like) family protein. The third associated region

contains the *MPK12* gene, which we obtained before (***Figure 5a, Figure 6***), suggesting a potential pleiotropic effect of *MPK12* on multiple drought-related phenotypes. On chromosome 3, we identified two associated regions containing *DRE/CRT-binding protein 2B* (*DREB2B*; AT3G11020) and *DNAJ homolog 3* (*ATJ3*; AT3G44110) (***Figure 7b***). On the other associated region on chromosome 3 (***Figure 7b***), *ATJ3* plays a crucial role in regulating flowering time and response to drought in *A. thaliana* by integrating multiple flowering signals and promoting the expression of *SOC1* and *FT* through direct interaction with *SVP* in the nucleus, which in turn prevents *SVP* binding to their regulatory sequences and facilitates the floral transition (Shen and Yu, 2011; Shen *et al.*, 2011).

Subsequently, we investigated the genetic underpinnings of compactness (CompactnessPC) in relation to drought in Santo Antão. Our findings revealed that five loci explained 76% (95% CI: 63%-88%) of the genetic variance for average compactness (***Figure 7c, Table 1***), while eight loci contributed to 11% (95% CI: 6%-28%) of the genetic variance in drought response of compactness (***Figure 7d***). Among the loci identified for both average and drought-responsive compactness, a significant association region was found on chromosome 3, which includes *AtDWARF4* (*AtDWF4*; AT3G50660) (***Figure 7d***). Next, we examined the genetic basis of RER underlying drought in Santo Antão. Our analysis indicated that 17 loci explained 30% (95% CI: 14%-49%) of the genetic variance for average RER (***Figure 7e, Table 1***), while five loci were responsible for 18% (95% CI: 2%-38%) of the genetic variance in drought response of RER (***Figure 7f***). We detected a significantly associated region on chromosome 1, encompassing *AtAHG2* (AT1G55870), which encodes a poly(A)-specific ribonuclease (*AtPARN*). Additionally, we identified multiple significant associations for the drought response of RER (***Figure 7f***). Notably, the associated region on chromosome 2 contains the *MPK12* gene we previously observed (***Figure 5a, Figure 6, Figure 7b***). This finding implies a potential pleiotropic influence of *MPK12* on various drought-related phenotypes.

In Fogo, we examined the genetic determinants of PRA in response to drought. Our analysis identified 16 loci that explained 41% (95% CI: 20%-63%) of the genetic variance in

average PRA (**Supplementary Figure 4a, Table 1**), while there were 11 loci contributing to 48% (95% CI: 30%-67%) of the genetic variance in the drought response of PRA (**Supplementary Figure 4b**). Among these associated loci, the most significant region on chromosome 5 includes *AtDOF5.8* (AT5G66940) (**Supplementary Figure 4b**). Subsequently, we examined the genetic factors contributing to compactness under drought conditions. Our study revealed that 44 loci accounted for 57% (95% CI: 29%-82%) of the genetic variance in average compactness (**Supplementary Figure 4c, Table 1**), while 12 loci were responsible for 26% (95% CI: 10%-46%) of the genetic variance in drought response of compactness (**Supplementary Figure 4d, Table 1**). For average compactness, we identified a significant region on chromosome 4 containing *AtUGT79B2* (AT4G27560) (**Supplementary Figure 4c**). In the drought response of compactness, we identified a significant region on chromosome 2 containing *PHYTOCHROME INTERACTING FACTOR1* (*AtPIF1*; AT2G20180) (**Supplementary Figure 4d**).

Next, we investigated the genetic basis of relative expansion rate (RER) under drought conditions. Our study identified 27 loci accounting for 37% (95% CI: 13%-65%) of the genetic variance in average RER (**Supplementary Figure 4e, Table 1**) and eight loci responsible for 30% (95% CI: 12%-53%) of the genetic variance in the drought response of RER (**Supplementary Figure 4f**). We discovered a significant region on chromosome 4 harboring the *FOREVER YOUNG* (*FEY*; AT4G27760) gene for average RER (**Supplementary Figure 4e**). *FEY* is implicated in leaf positioning and meristem maintenance in *A. thaliana*, with the fey mutant displaying disruptions in these developmental processes (Callos *et al.*, 1994). The predicted FEY protein exhibits similarity to nodulins and specific reductases, indicating a potential function in shoot apex communication through modifying a regulatory factor in meristem development (Callos *et al.*, 1994). For the drought response of RER, we detected a significant region on chromosome 1 encompassing *AtPHO1-H10* (AT1G69480) (**Supplementary Figure 4f**).

In the average PRA analysis for Morocco (**Supplementary Figure 5a**), a significant association was identified on chromosome 4, encompassing the AT4G03570 gene, the

function of which remains unknown. Conversely, no Bonferroni significant findings were observed in the drought response of PRA (**Supplementary Figure 5b**). Regarding compactness, although no Bonferroni significant regions were detected in the average trait (**Supplementary Figure 5c**), several significant associations were found in the drought response (**Supplementary Figure 5d**), including a notable region on chromosome 4 containing the *JMJ28* gene (AT4G21430). *JMJ28*, an *A. thaliana* KDM3 group H3K9 demethylase, interacts with *FLOWERING BHLH1* (*FBH*) transcription factors to facilitate the activation of the core flowering regulator *CONSTANS* (*CO*) by eliminating the repressive H3K9me2 mark (Hung *et al.*, 2021). This interaction uncovers the role of *JMJ28* and *FBH* transcription factors in the epigenetic regulation of *CO*, thus promoting flowering in *A. thaliana* (Hung *et al.*, 2021). For RER, no Bonferroni significant regions were found in either the average trait (**Supplementary Figure 5e**) or the drought response (**Supplementary Figure 5f**).
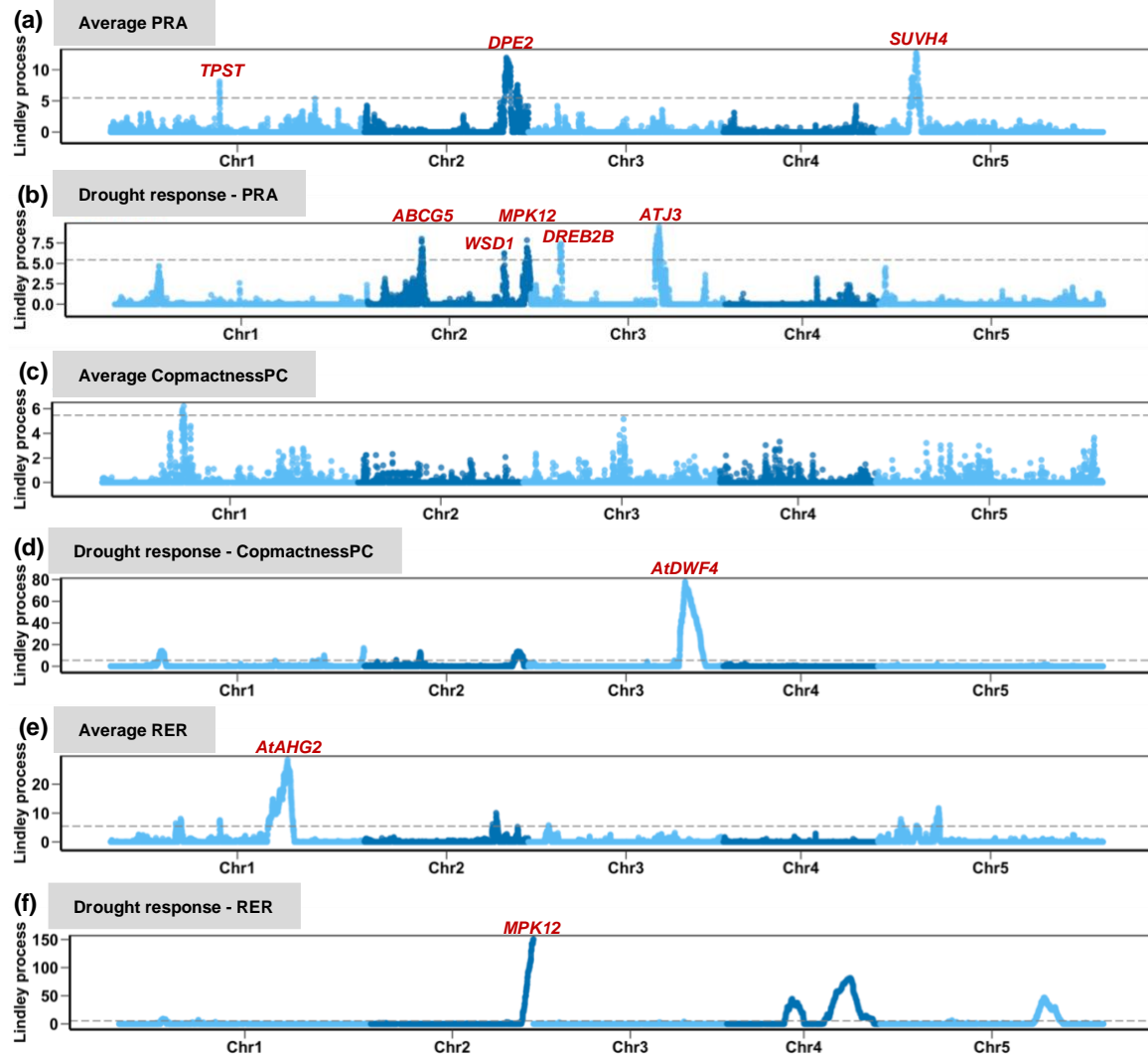
Figure 7. Genomic associations of rosette growth-related traits (PRA, compactnessPC, and RER) in Santo Antão using LMM in GEMMA followed by the local score approach. The x-axis and y-axis represent the chromosomes and Lindley score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.

3.3.5.4 The genetic basis of leaf color-related traits (hue and saturation) in response to drought in CVI relative to Morocco

In Santo Antão, we investigated the genetic basis of hue (a proxy for chlorophyll content) underlying drought. Our findings revealed that 38 loci explained 65% (95% CI: 47%-81%) of the genetic variance for average hue (***Figure 8a, Table 1***), while 30 loci explained 44% (95% CI: 27%-63%) of the genetic variance for drought response of hue (***Figure 8b***). We discovered two significant regions on chromosome 3 harboring *TMAC2* (AT3G02140) and *ATG4* (AT3G51820) genes for average hue (***Figure 8a***). *TMAC2* is an ABA-responsive gene

in *A. thaliana* that is highly induced by ABA and NaCl and is targeted to the nucleus (Huang and Wu, 2007). Constitutive overexpression of *TMAC2* results in insensitivity to ABA and salt stress, suggesting that it negatively regulates these responses (Huang and Wu, 2007). Additionally, *TMAC2* plays a critical role in controlling root elongation, floral initiation, and starch degradation, which is evidenced by the altered phenotypes of *TMAC2*-overexpressing plants, including short roots, late flowering, and starch excess (Huang and Wu, 2007). CHLG, or chlorophyll synthase, plays a crucial role in a salvage pathway for chlorophyll biosynthesis by re-esterifying chlorophyllide produced during chlorophyll turnover (Lin *et al.*, 2014). This pathway is essential for recycling chlorophyll in *A. thaliana* under heat stress, as demonstrated by the CLD1 (chlorophyll dephytylase 1) and CHLG enzymes in *chlg* and *cld1* mutants (Lin *et al.*, 2022). CHLG also interacts with CLD1 and the light-harvesting complex-like proteins OHP1 and LIL3, suggesting that auxiliary factors are required for the chlorophyll recycling process (Lin *et al.*, 2022). Mutations in CHLG can lead to the accumulation of chlorophyllide a and the generation of phototoxic singlet oxygen, causing light-dependent, heat-induced cotyledon bleaching in *A. thaliana* seedlings (Lin *et al.*, 2014). Additionally, we identified multiple significant associations for the drought response of hue (***Figure 8b***). The significantly associated region on chromosome 5 contains *the Clade E Growth-Regulating2* (*EGR2*, AT5G27930) gene. The EGR Type 2C protein phosphatases, including EGR2, act as negative growth regulators to restrain plant growth during drought conditions (Bhaskara *et al.*, 2017). EGRs target cytoskeleton and plasma membrane-associated proteins, such as Microtubule-Associated Stress Protein 1 (MASP1), which binds, bundles, and stabilizes microtubules (Bhaskara *et al.*, 2017). The EGR-MASP1 system selectively regulates microtubule recovery and stability, allowing plants to adjust their growth and cell expansion in response to drought stress (Bhaskara *et al.*, 2017). The significantly associated region on chromosome 1 contains *the AtbZIP62* (AT1G19490) gene. The transcription factor *AtbZIP62* plays a role in oxidative, nitro-oxidative, and drought stress conditions (Rolly *et al.*, 2020). It appears to be involved in the transcriptional regulation of *AtPYD1*, a gene involved in pyrimidine biosynthesis and modulates cellular redox state and photosynthetic processes (Rolly *et al.*, 2020). *AtbZIP62* is suggested to positively regulate drought stress response in *A. thaliana*,

with the knock-out plants showing sensitivity to drought stress and an inability to recover after re-watering (Rolly *et al.*, 2020). The associated region at the end of chromosome 2 contains the *MPK12* gene, which we previously observed in other associations.

In Fogo, we explored the genetic basis of hue underlying drought. We identified 33 loci that explained 69% (95% CI: 54%-83%) of the genetic variance for average hue (***Figure 8c, Table 1***), while seven loci explained 46% (95% CI: 25%-68%) of the genetic variance for drought response of hue (***Figure 8d***). We found a region on chromosome 1 containing the *AGP31* (AT1G28290) gene was associated with average hue (***Figure 8a***). *AGP31*, a nonclassical arabinogalactan protein in *A. thaliana*, has its mRNA levels decreased in response to wounding, methyl jasmonate, and ABA, primarily through transcription repression (Liu and Mehdy, 2007). The AGP31 protein is localized to the cell wall and is a galactose-rich AGP, with expression found in the vascular bundle throughout the plant, except in the flower (Liu and Mehdy, 2007). The preferential expression in vascular tissues suggests that *AGP31* may play a role in vascular tissue function during defense responses and development (Liu and Mehdy, 2007). Regarding the drought response of hue, we identified a significant region on chromosome 5 containing the *AtU2AF35b* (AT5G42820). *AtU2AF35b* functions as a negative regulator of flowering time in *A. thaliana* and is involved in ABA-mediated flowering (Xiong *et al.*, 2019). The *atu2af35b* mutants exhibit early-flowering phenotypes under both long-day and short-day conditions, with reduced transcript accumulation of the flowering repressor gene *FLOWERING LOCUS C* (*FLC*) (Xiong *et al.*, 2019). *AtU2AF35b* is implicated in the pre-mRNA splicing of *FLC* and *ABI5* in the shoot apex, thereby playing a role in ABA-mediated flowering transition in *A. thaliana* (Xiong *et al.*, 2019).

In Morocco, we explored the genetic basis of hue underlying drought. We identified 59 loci explained 70% (95% CI: 29%-93%) of the genetic variance for average hue (***Figure 8c, Table 1***), while 36 loci explained 66% (95% CI: 34%-88%) of the genetic variance for drought response of hue (***Figure 8d***). A notable region at the end of chromosome 2 containing the *ATHB-4* (AT2G44910) gene was discovered to be associated with average

hue. *ATHB-4*, a member of the class II homeodomain-leucine zipper (HD-ZIPII) transcription factor family, plays a crucial role in establishing the dorsoventral axis in cotyledons and developing leaves, as well as in responding to shade (Bou-Torrent *et al.*, 2012). Loss-of-function mutations in *ATHB-4* and another HD-ZIPII gene, *HAT3*, result in severely abaxialized and radialized leaves, while overexpression of *HAT3* leads to adaxialized leaf development (Bou-Torrent *et al.*, 2012). The involvement of the HD-ZIPII/HD-ZIPIII module in both the dorsoventral patterning of leaves and shade avoidance highlights its significance in the intersection of plant patterning and growth promotion (Bou-Torrent *et al.*, 2012). Regarding the drought response of hue, no Bonferroni significant regions were detected.
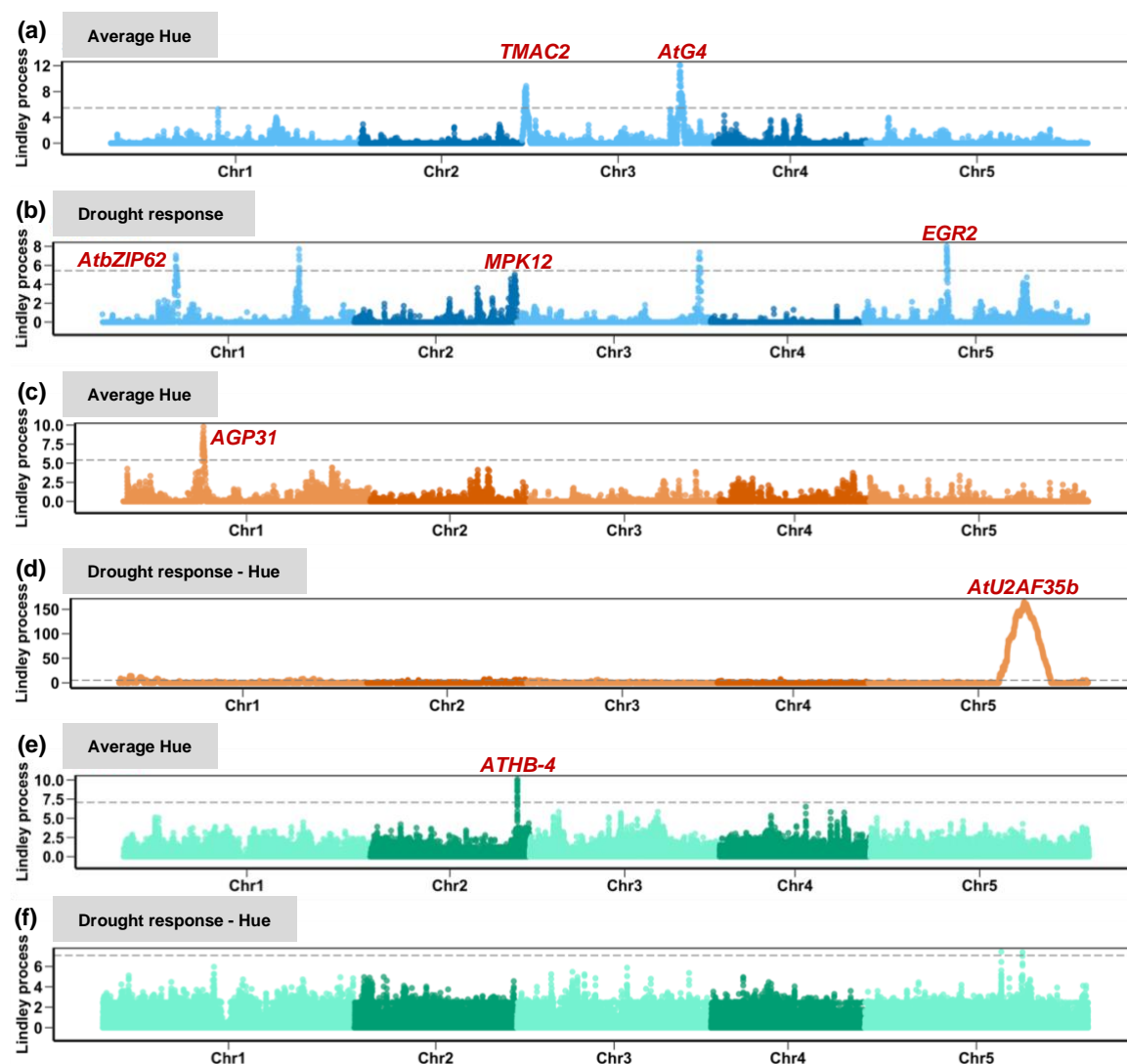

Figure 8. Genomic associations of hue (a proxy for chlorophyll content) in response to drought in Santo Antão (blue), Fogo (orange), and Morocco (green) populations using LMM in GEMMA followed by the local score approach. The y-axis

represents the Lindley process score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.

We next investigated the genetic basis of saturation (color intensity) under drought conditions in Santo Antão. We found that 21 loci explained 72% (95% CI: 59%-84%) of the genetic variance for average saturation (*Figure 9a, Table 1*), while 16 loci contributed to 36% (95% CI: 16%-60%) of the genetic variance for drought-induced saturation response (*Figure 9b*). A significant region at the beginning of chromosome 4 (*Figure 9a*) encompassed the well-known *FRIGIDA* (*FRI*), a key regulator in determining flowering time in *A. thaliana* (Johanson *et al.*, 2000; Gazzani *et al.*, 2003; Shindo *et al.*, 2005; Fulgione *et al.*, 2022). Another region on chromosome 5 (*Figure 9a*) harbored the *NHL26* gene. Interestingly, both *FRI* and *NHL26* were also detected in the genetic basis of the drought response of WUE when *MPK12* was included as a covariate in the model (*Figure 5c*), highlighting the shared genetic components across different traits. For the drought response of saturation, we discovered a significant region on chromosome 5 containing the *CONSTANS-LIKE1* (*COL1*, AT5G15850) gene, homologous to the *A. thaliana* flowering-time gene *CONSTANS*. Recent research has revealed that a soybean homolog of *COL1*, *GmCOL1a*, plays a crucial role in enhancing soybean stress tolerance (Xu *et al.*, 2023). Under salt and drought stress, mRNA levels of *GmCOL1a* increased, and the protein was stabilized, resulting in enhanced stress tolerance (Xu *et al.*, 2023). *GmCOL1a* promotes the expression of genes related to salt tolerance, ultimately reducing the Na+/K+ ratio in soybean plants and contributing to their survival under harsh conditions (Xu *et al.*, 2023).

In Fogo, we investigated the genetic basis of saturation underlying drought. We identified 48 loci explained 81% (95% CI: 69%-90%) of the genetic variance for average saturation (*Figure 9c, Table 1*), while 23 loci explained 42% (95% CI: 25%-60%) of the genetic variance for drought response of saturation (*Figure 9d*). A notable region at chromosome 3 containing the *MIR169F* (AT3G14385) gene was discovered to be associated with average saturation (*Figure 9c*). The *MIR169* gene plays a crucial role in drought resistance by regulating the nuclear factor Y (NF-Y) transcription factor in *A. thaliana*. Specifically, *NFYA5*, a subunit of NF-Y, is induced by drought stress in an ABA-dependent manner,

while *MIR169* is downregulated under the same conditions (Li *et al.*, 2008). Similarly, in tomato plants, the accumulation of *Sly-miR169* is induced by drought stress, leading to the downregulation of its target genes, including three nuclear factor Y subunit genes (*SlNF-YA1/2/3*) and one multidrug resistance-associated protein gene (*SlMRP1*) (Zhang *et al.*, 2011). Overexpression of *Sly-miR169c* in tomato plants results in reduced stomatal opening, decreased transpiration rate, lower leaf water loss, and enhanced drought tolerance (Zhang *et al.*, 2011). With the drought response of saturation, our analysis revealed a significant region on chromosome 3 containing the *CYSTEINE-RICH REPEAT SECRETORY PROTEIN 38* (*CRRSP38*, AT3G22060) gene. Previous research has demonstrated that the *CRRSP38* gene may serve as a novel regulator of leaf growth (Nikonorova *et al.*, 2018).
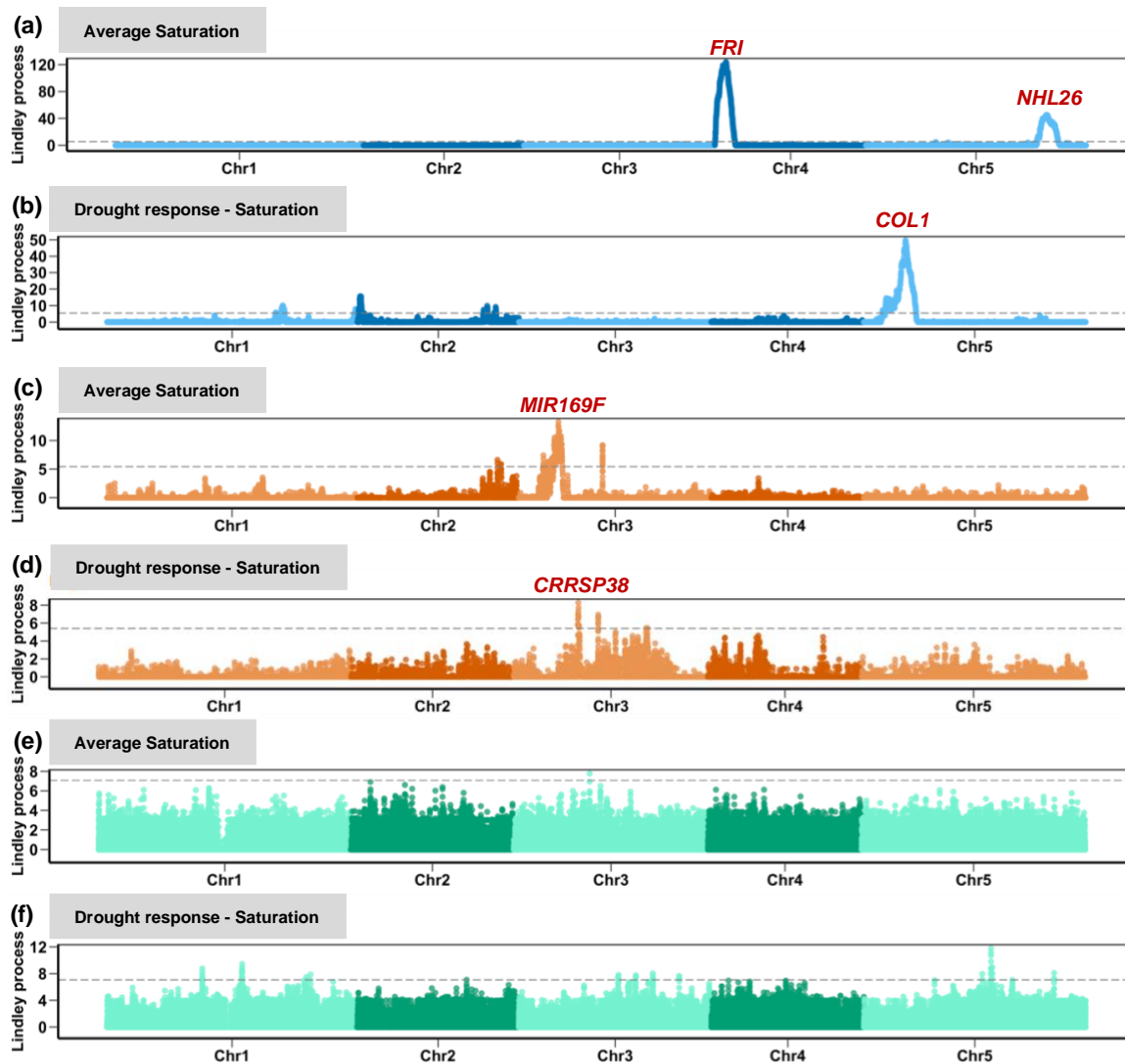
Figure 9. Genomic associations of saturation (color intensity) in response to drought in Santo Antão (blue), Fogo (orange), and Morocco (green) populations using LMM in GEMMA followed by the local score approach. The y-axis represents the Lindley process score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.

3.4   Discussion

We utilized the Phenoscope (high-throughput phenotyping robot) platform to investigate the genetic basis of rosette growth, leaf color, and stomatal traits under precisely controlled drought conditions in Santo Antão, where the accession *Cvi-0* was collected, and Fogo populations, comparing these to the closest outgroup population originating from Morocco. We focused on several traits describing the dynamic rosette growth, including projected rosette area (PRA), relative growth rate (RER), and compactness. Rosette growth-related traits exhibit different genetic architectures throughout the plant lifecycle, with individual loci contributing variably to cumulative or age-specific traits. Previous research has demonstrated that heritabilities for growth-related traits change over time, potentially due to QTLs (quantitative trait loci) acting more or less likely at specific time points (Zhang *et al.*, 2012; Flood *et al.*, 2016). Studying growth dynamics at different time resolutions has allowed for identifying genotype-phenotype associations not detectable by considering only cumulative traits at a single time point (Moore *et al.*, 2013; Muraya *et al.*, 2017; Zhang *et al.*, 2017).

While the number of identified genes responsible for the natural variation of individual traits continues to grow (e.g., (Alonso-Blanco *et al.*, 2005; Loudet *et al.*, 2007; Baxter *et al.*, 2010; Kronholm *et al.*, 2012; Brachi *et al.*, 2015; Vidigal *et al.*, 2016)), these studies primarily focus on specific types of adaptation with relatively simple genetic architectures in the investigated populations. Our study aims to address the knowledge gap in understanding genetic constraints affecting organism adaptation. By dissecting the genetic basis of complex and/or correlated traits related to drought stress in a colonizing *A. thaliana* lineage, we provide insights into the genetic underpinnings of drought adaptation in more intricate contexts.

Here, we investigated the genetic basis of water use efficiency (WUE) in response to drought across three distinct regions: Santo Antão, Fogo, and Morocco, utilizing several different GWAS models. Several GWAS candidate genes were associated with the ABA signaling pathway, which plays a crucial role in plant responses to drought stress. ABA

concentration in leaf tissues influences transpiration via stomatal closure and modulates water uptake by regulating the root-to-shoot biomass ratio (Tuberosa, 2012). Moreover, ABA contributes to proline accumulation in plants under abiotic stress conditions (Sharma *et al.*, 2011). Proline and antioxidants such as ascorbic acid, anthocyanins, and tocopherol serve as a scavenger for Reactive Oxygen Species (ROS). These compounds, in conjunction with enzymes like superoxide dismutase and catalases that degrade ROS, help mitigate ROS-induced damage to the cellular structures of stressed plants (Bhatt *et al.*, 2011; Setter, 2012; Bhargava and Sawant, 2013; Liang *et al.*, 2013).

We previously showed that a phenotypic shift towards higher stomatal conductance and lower WUE on Santo Antão island was due largely to a nonsynonymous variant in the *MPK12* gene (Elfarargi *et al.*, 2023). In this study, by including the *MPK12* G53R variant as a covariate in the model, we uncovered additional significant associations, all of which are implicated in the abscisic acid (ABA) signaling pathway. Subsequently, we integrated stomatal density and stomatal pore width with WUE to create a multitrait set, as these traits are highly intercorrelated (*Figure 4*) and represent the leaf stomatal as a structural unit. Multivariate GWAS (Zhou and Stephens, 2014) analysis of the average data for these three traits uncovered several signals across the genome (*Figure 6*). Among the associated loci, we found the *NPF8.1* gene, which plays a crucial role in the stomatal movement in *A. thaliana* by facilitating ABA uptake into guard cells and, consequently, modulating stomatal closure in response to drought (Tal *et al.*, 2016; Shimizu *et al.*, 2021). Additionally, we found an association with *SUVH4* (*Figure 6*), which has also been shown to be influenced by ABA modulation (Zheng *et al.*, 2012). Interestingly, we identified *SUVH4* in association with the average PRA (*Figure 7a*), suggesting that *SUVH4* has potential pleiotropic effects on several traits underlying drought.

Furthermore, we identified several candidate genes associated with rosette growth-related traits underlying drought stress in Santo Antão. Our findings revealed genes involved in various pathways, such as auxin signaling (*TPST*), starch metabolism (*DPE2*), cuticle formation (*ABCG5*), wax ester accumulation (*WSD1-like*), and stress-responsive

signaling (*MPK12*, *DREB2B*, *ATJ3*). Loss-of-function *tpst* mutants exhibit a pronounced dwarf phenotype, early senescence, and a reduced number of flowers and siliques (Komori *et al.*, 2009). For *DPE2*, a previous study reported a significant decrease in both leaf and rosette area for *dpe2* mutants under drought stress compared to control conditions (Massonnet *et al.*, 2015). A recent study has shown that ABCG5 plays a crucial role in promoting proper seedling development under drought conditions by forming a dense cuticle layer, which helps maintain air spaces within and on the plant, thereby preventing excessive water absorption and ensuring normal growth (Lee *et al.*, 2021). It has been shown that *WSD1* plays a key role in drought tolerance by contributing to the accumulation of wax esters on *A. thaliana* leaves and stems, thereby reducing water loss and maintaining plant growth and relative water content under water deficit conditions (Patwari *et al.*, 2019). The expression of *DREB2B* is induced by drought stress and hence regulates the transcription of the drought-responsive genes via recognizing and binding to the DRE/CRT sequence (Liu *et al.*, 2019). It has been known that the dehydration-responsive element (DRE) has an essential trans-acting function for the regulation of abiotic responsive genes induced in the ABA-independent downstream signaling pathway in response to drought, salinity, and cold (Yamaguchi-Shinozaki and Shinozaki, 1994). All DRE binding proteins, including DREB2B, showed a significant sequence similarity in the ERF/AP2 DNA binding region (77 – 134 amino acid). In *A. thaliana*, *DREB2B* has seven homologs, and it has a higher expression level with *DREB2A* compared to the low expression level of the remaining six DREB2-type proteins under drought stress (Nakashima *et al.*, 2000). A previous study showed that the overexpressing *DREB2A*, the homolog to *DREB2B*, has an impact on growth under normal conditions, and this constitutive active form improved tolerance to drought stress via regulation of drought stress-responsive gene expression (Sakuma *et al.*, 2006). *ATJ3* is a member of the *A. thaliana* cytosolic *HSP40* family, which plays a vital role in protecting plants against prolonged heat stress by interacting with HSP70-4 and participating in heat stress granules, with its chaperone function being regulated by protein farnesylation, essential for heat tolerance in plants (T.,-Y., Wang *et al.*, 2021).

Furthermore, we detected associations with genes involved in brassinosteroid biosynthesis (*AtDWF4*) and ABA hypersensitivity and stress responses (*AtAHG2/AtPARN*). Overexpression of *AtDWF4* in *Brassica napus* has demonstrated enhanced tolerance to abiotic and biotic stress, as well as increased plant productivity (Sahni *et al.*, 2016). Similarly, transgenic *A. thaliana* plants overexpressing *BdDWF4*, a homolog from *Brachypodium distachyon*, displayed elongated and slender phenotypes (Corvalán and Choe, 2017). The *ahg2-1* mutant plants exhibit higher ABA accumulation compared to wild-type plants, resulting in increased ABA sensitivity during germination and later developmental stages, as well as modified responses to various stressors (Nishimura *et al.*, 2005). The rosette growth rate in *ahg2-1* plants is reduced by over half, although the number of rosette leaves is not significantly impacted (Nishimura *et al.*, 2005). Lastly, *AtPARN* expression is induced by ABA, high salinity, and osmotic stress, implying that the mRNA-destabilizing activity of *AtPARN* is essential for regulating ABA, salicylic acid, and stress responses in plants (Nishimura *et al.*, 2005).

In Fogo, we identified several genes associated with rosette growth-related traits underlying drought. Our findings revealed genes involved in abiotic stress responses, such as *AtDOF5.8*, *AtUGT79B2*, *AtPIF1*, *FEY*, and *AtPHO1-H10*. The *AtDOF5.8* transcription factor has been implicated in the regulation of the *ANAC069* gene, which plays a role in integrating auxin and salt signals to modulate *A. thaliana* seed germination (He *et al.*, 2015). Both *AtDOF5.8* and *ANAC069* share similar expression profiles in response to salt, drought, and ABA exposure, indicating their participation in a regulatory network that mediates abiotic stress responses in plants (He *et al.*, 2015). The *AtUGT79B2* has been shown to be strongly induced by various abiotic stresses, including cold, salt, and drought (Li *et al.*, 2017). Overexpression of *AtUGT79B2* enhances plant tolerance to these stresses, while double mutants with *Atugt79b3* show increased susceptibility, suggesting that these anthocyanin rhamnosyltransferase genes play a role in abiotic stress tolerance through modulating anthocyanin accumulation (Li *et al.*, 2017). *AtPIF1* influences ABA biosynthesis and signaling component genes (Oh *et al.*, 2007), while *FEY* is implicated in leaf positioning and meristem maintenance (Callos *et al.*, 1994). Lastly, *AtPHO1-H10* is involved in

phosphate homeostasis and is upregulated during Pi deficiency (Ribot *et al.*, 2008). Additionally, its expression is modulated by abscisic acid, which increases *AtPHO1-H10* expression in both Pi-sufficient and Pi-deficient situations, implying a potential role in drought stress response due to the involvement of abscisic acid in plant stress responses (Ribot *et al.*, 2008).

Hue and saturation are an alternative representation of the RGB (Red Green Blue) color model extracted from visual images of plant rosettes (Sass *et al.*, 2012). The Hue value represents the color range and saturation the color intensity (Walter *et al.*, 2007). It was reported that hue values have a highly positive correlation with chlorophyll content and were suitable for photochemical yield estimation of photosystem II (Faragó *et al.*, 2018). Therefore, since the hue values represent the chlorophyll content, the saturation values represent the intensities of chlorophyll content. Although previous research has indicated that the saturation parameter, a proxy for chlorophyll intensity, is less important than hue values (De Vylder *et al.*, 2012), our research discovered an association between average saturation and the *FRIGIDA* (*FRI*) gene located on chromosome 4 (**Figure 9a**). The *FRI* gene is known to influence a range of traits, such as flowering, water use efficiency, photosynthesis, and plant growth (Lovell *et al.*, 2013). Consistently, we identified the *FRI* gene for the drought response of WUE through incorporating the *MPK12* G53R variant as a covariate in the model (**Figure 5c**), suggesting potential pleiotropic effects of *FRI* in leaf color together with WUE underlying drought stress.

A connection between flowering time and vegetative growth dynamics in *A. thaliana* has been previously established (Bac-Molenaar *et al.*, 2016). The role of pleiotropy in drought adaptation has also been explored by examining correlated physiological traits in nature, like WUE and flowering time (McKay *et al.*, 2003). Additionally, hue and saturation traits were used in a prior study examining natural variation in an *A. thaliana* population derived from a Can-0 x Col-0 cross (Hanemian *et al.*, 2020). This study identified a QTL involving the *FLM* (*FLOWERING LOCUS M*) gene, which regulates pleiotropic aspects of plant growth and leaf color (Hanemian *et al.*, 2020). Considering these findings, we propose that leaf

color traits (hue and saturation) may be linked to the transition from the vegetative to the flowering stage, a critical adaptive process that significantly contributes to reproductive success.

Upon a sudden shift to an extreme climate in CVI, adaptation must be fast enough to modulate alleles with major effects for simple traits or alleles with pleiotropic functions for complex traits. An optimal balance exists between loci with targeted function and loci with pleiotropic effects for accelerated adaptation to manage complex traits. Here, we showed that *MPK12* is associated with WUE, stomatal traits, plant growth, and leaf color underlying drought, suggesting pleiotropic effects of the *MPK12* gene in a much wider role than initially described.

This study sheds light on the genetic underpinnings of drought-related traits and their potential involvement in plant adaptation to drought environments. Moreover, identifying genomic associations sets the stage for future investigations to decipher the molecular mechanisms underlying plant responses to drought stress. While we have uncovered promising links between genes and drought-related traits, further validation of these candidate genes as drivers of observed phenotypic variation require additional studies employing knockout lines, reciprocal transformations, and analyses of differential gene expression and co-expression networks. Once validated, this knowledge could deepen our understanding of the roles these genes play in *A. thaliana's* drought response and might potentially be leveraged to improve drought tolerance in crop plants.

3.5   Materials and methods

3.5.1 Whole-Genome Sequencing

We sequenced 15 new Santo Antão accessions using Illumina HiSeq3000 machines. Genomic DNA was extracted using the DNeasy Plant Mini kits (Qiagen), fragmented using sonication (Covaris S2), and libraries were prepared with Illumina TruSeq DNA sample prep kits (Illumina), NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs) and NEBNext Ultra II DNA Library Prep Kit (New England Biolabs). Libraries were sequenced with 2x 100-150 bp paired-end reads. We assessed DNA quality and quantity via capillary electrophoresis (TapeStation, Agilent Technologies) and fluorometry (Qubit and Nanodrop, Thermo Fisher Scientific).

3.5.2 Variant identification and genotyping

Here, we called variants from the newly sequenced 15 Santo Antão accessions together with the previously released whole-genome short-read data for 189 Santo Antão accessions (ENA: PRJEB39079 (ERP122550)) (Fulgione *et al.*, 2022). We aligned the raw Illumina sequence for the whole data to the *Arabidopsis* TAIR10 reference genome. Further, we used SHORE (Ossowski *et al.*, 2008) to identify SNP variants following a pipeline described here (https://github.com/HancockLab/SNP_calling_Arabidopsis). Furthermore, the obtained variant call format (VCF) file was filtered to minimize variant calling bias and to retain only high-quality variants: (1) retain only bi-allelic variants; (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.

3.5.3 Study populations

In this study, we used 350 lines from CVI, including 189 accessions (Fulgione *et al.*, 2022) together with 15 new accessions collected from 26 different locations in Santo Antão (*Figure 1*), 146 accessions collected from 18 different locations in Fogo (Tergemina *et al.*, 2022; Fulgione *et al.*, 2022), and 61 Moroccan accessions (Durvasula *et al.*, 2017) for genetic analysis.

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.

3.5.4 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 (n=152), Fogo (n=146) (Tergemina *et al.*, 2022; Fulgione *et al.*, 2022) and Moroccan (n=61) *A. thaliana* accessions (Brennan *et al.*, 2014) 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). Phenoscope experimental setup and sample collection for WUE and its measurements can be found here (Elfarargi *et al.*, 2023). Different phenotypic traits were extracted from images taken daily at the same time throughout 23 days on Phenoscope. The image-based traits were defined as follows: projected rosette area (PRA), hue and saturation (converted from RGB color component), compactnessPC: (ratio of PRA (P) to surface of convex hull area (C)), and relative expansion rate (RER) over time window from day 8 to day 32.

In this study, we measured further leaf stomatal density and stomatal pore width in the Santo Antão population. We adapted a previously published protocol, "Microscopy-Based Stomata Analyses" (Eisele *et al.*, 2016), using a small piece of Scotch tape. The tape was applied firmly onto the abaxial surface of a detached fully developed leaf per accession and per treatment of each genotype. Then, we gently peeled off the tape from the leaf surface, not tearing it. We transferred the tape immediately to a petri dish containing 4% formaldehyde solution for 10-20 minutes at room temperature to fix the status of stomatal guard cells and prevent their responses to subsequent signals. Next, we put the samples (tapes with epidermal peels) into a clean tissue and added 50 µl of 1 mg/ml propidium iodide solution to each sample and incubated them for 5-10 minutes in the dark. Then we placed the tape on a microscope slide and gently pressed it to remove any air bubbles. As a final step, we examined for stomatal densities and stomatal apertures

using confocal images of epidermal peels taken with a confocal microscope (LSM 700, ZEISS). The images were captured and analyzed using ImageJ software (https://imagej.nih.gov/ij/). The measurements were performed across several days (29-32 DAS) around mid-day.

3.5.5 Phenotype data analysis

To assess differences in phenotype distributions, we employed both parametric and non-parametric tests. Wilcoxon rank sum tests were conducted using the wilcox.test function in the '*ggpubr*' package (Kassambara, 2020). Linear models were also used to determine the fixed effects of treatment, geographic region, and their interaction on the measured phenotypic traits. The '*lme4*' package in R (Bates *et al.*, 2014) was utilized to run a model for each phenotype with the following equation:

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

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

Pearson correlations were calculated in R using the *cor.test* function to assess the correlations between phenotypes in both treatments. The significance of the correlations was evaluated with the t-test implemented in the *cor.test* function.

3.5.6 Genome-wide association study

3.5.6.1   Univariate and Multivariate GWAS

To perform the association analysis, we used the SNP data. We first filtered out non-biallelic variants from the VCF. We considered only variants 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 studied traits using the univariate linear mixed model implemented in GEMMA (Zhou and Stephens, 2012), separately for well-watered and drought conditions as well as the average for each trait across both conditions and the drought response (difference between conditions: WW-WD). For stomatal traits, we utilized a multivariate linear mixed

model (mvLMM) in GEMMA (Zhou and Stephens, 2014), which jointly models the relationships between the traits.

According to Shim *et al.* 2015 (Shim *et al.*, 2015) and based on the GEMMA outputs, we calculated the proportion of variance in each trait explained by a given variant (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 variant, MAF is the minor allele frequency for the variant, and N is the sample size.

Next, we applied a local score approach to the output of *p*-values provided by LMM and mvLMM in GEMMA. This local score approach, which takes LD into account when estimating associations, increases the power of detecting significant genomic regions associated with trait variation and is subsequently converted into scores (Bonhomme *et al.*, 2019). Scripts used to calculate the Lindley scores are available here (https://forge-dga.jouy.inra.fr/documents/940).

3.5.6.2 Inference of genetic architecture

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 variants by inferring locus effect sizes (β) while controlling for other variants included in the model. Using this approach, we modeled two effect hyperparameters: a basal effect (β), which captures small-effect loci contributing 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 variants, the sparse effect size for each locus was calculated by multiplying (β) by (γ). We used the 1% variants (> 99% quantile) distribution of locus effects to identify which genes with the highest sparse effects on the studied trait.

3.5.7 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).

3.6 Acknowledgments

3.7 Author 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 δ13C analysis. N.K., A.P.M.W and his lab members were responsible for δ13C 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. and A.M.H. wrote the manuscript with input from all authors.

3.8 References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2003) Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. *Plant Cell*, **15**, 63–78.

Aliniaeifard, S. and Meeteren, U. van (2014) Natural variation in stomatal response to closing stimuli among Arabidopsis thaliana accessions after exposure to low VPD as a tool to recognize the mechanism of disturbed stomatal functioning. *J. Exp. Bot.*, **65**, 6529–6542.

Alonso-Blanco, C., Andrade, J., Becker, C., et al. (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. *Cell*, **166**, 481–491.

Alonso-Blanco, C., El-Assal, S.E.-D., Coupland, G. and Koornneef, M. (1998) Analysis of Natural Allelic Variation at Flowering Time Loci in the Landsberg erecta and Cape Verde Islands Ecotypes of Arabidopsis thaliana. *Genetics*, **149**, 749–764.

Alonso-Blanco, C., Gomez-Mena, C., Llorente, F., Koornneef, M., Salinas, J. and Martínez-Zapater, J.M. (2005) Genetic and Molecular Analyses of Natural Variation Indicate CBF2 as a Candidate Gene for Underlying a Freezing Tolerance Quantitative Trait Locus in Arabidopsis. *Plant Physiol.*, **139**, 1304–1312.

Ayatollahi, Z., Kazanaviciute, V., Shubchynskyy, V., et al. (2022) Dual control of MAPK activities by AP2C1 and MKP1 MAPK phosphatases regulates defence responses in Arabidopsis. *J. Exp. Bot.*, **73**, 2369–2384.

Bac-Molenaar, J.A., Fradin, E.F., Becker, F.F.M., Rienstra, J.A., Schoot, J. van der, Vreugdenhil, D. and Keurentjes, J.J.B. (2015) Genome-Wide Association Mapping of Fertility Reduction upon Heat Stress Reveals Developmental Stage-Specific QTLs in Arabidopsis thaliana. *Plant Cell*, **27**, 1857–1874.

Bac-Molenaar, J.A., Granier, C., Keurentjes, J.J.B. and Vreugdenhil, D. (2016) Genome-wide association mapping of time-dependent growth responses to moderate drought stress in Arabidopsis. *Plant Cell Environ.*, **39**, 88–102.

Bao, Y., Song, W.-M. and Zhang, H.-X. (2016) Role of Arabidopsis NHL family in ABA and stress response. *Plant Signal. Behav.*, **11**, e1180493.

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

Baxter, I., Brazelton, J.N., Yu, D., et al. (2010) A coastal cline in sodium accumulation in Arabidopsis thaliana is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genet.*, **6**, e1001193.

Belda-Palazon, B., Gonzalez-Garcia, M.-P., Lozano-Juste, J., et al. (2018) PYL8 mediates ABA perception in the root through non-cell-autonomous and ligand-stabilization–based mechanisms. *Proc. Natl. Acad. Sci.*, **115**, E11857–E11863.

Bhargava, S. and Sawant, K. (2013) Drought stress adaptation: metabolic adjustment and regulation of gene expression. *Plant Breed.*, **132**, 21–32.

Bhaskara, G.B., Wen, T.-N., Nguyen, T.T. and Verslues, P.E. (2017) Protein Phosphatase 2Cs and Microtubule-Associated Stress Protein 1 Control Microtubule Stability, Plant Growth, and Drought Response. *Plant Cell*, **29**, 169–191.

Bhatt, D., Negi, M., Sharma, P., Saxena, S.C., Dobriyal, A.K. and Arora, S. (2011) Responses to drought induced oxidative stress in five finger millet varieties differing in their geographical distribution. *Physiol. Mol. Biol. Plants Int. J. Funct. Plant Biol.*, **17**, 347–353.

Bohnert, H.J., Donald, ', Nelson, E. and Jensenayb, R.G. (1995) *Adaptations to Environmental Stresses*, American Society of Plant Physiologists.

Bonhomme, M., Fariello, M.I., Navier, H., et al. (2019) A local score approach improves GWAS resolution and detects minor QTL: application to Medicago truncatula quantitative disease resistance to multiple Aphanomyces euteiches isolates. *Heredity*, **123**, 517–531.

Borevitz, J.O. and Nordborg, M. (2003) The impact of genomics on the study of natural variation in Arabidopsis. *Plant Physiol.*, **132**, 718–25.

Bouteillé, M., Rolland, G., Balsera, C., Loudet, O. and Muller, B. (2012) Disentangling the Intertwined Genetic Bases of Root and Shoot Growth in Arabidopsis. *PLoS ONE*, **7**, e32319.

Bou-Torrent, J., Salla-Martret, M., Brandt, R., Musielak, T., Palauqui, J.-C., Martínez-García, J.F. and Wenkel, S. (2012) ATHB4 and HAT3, two class II HD-ZIP transcription factors, control leaf development in Arabidopsis. *Plant Signal. Behav.*, **7**, 1382–1387.

Boyer, J.S. (1982) Plant productivity and environment. *Science*, **218**, 443–8.

Brachi, B., Meyer, C.G., Villoutreix, R., Platt, A., Morton, T.C., Roux, F. and Bergelson, J. (2015) Coselected genes determine adaptive variation in herbivore resistance throughout the native range of Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 4032–4037.

Brennan, A.C., Méndez-Vigo, B., Haddioui, A., Martínez-Zapater, J.M., Picó, F.X. and 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 Biol.*, **14**, 17.

Brochmann, C., Rustan, Ø.H., Lobin, W. and Kilian, N. (1997) The endemic vascular plants of the Cape Verde Islands, W Africa. *Sommerfeltia*, **24**, 1–363.

Cadman, C.S.C., Toorop, P.E., Hilhorst, H.W.M. and Finch-Savage, W.E. (2006) Gene expression profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. *Plant J.*, **46**, 805–822.

Callos, J.D., DiRado, M., Xu, B., Behringer, F.J., Link, B.M. and Medford, J.I. (1994) The forever young gene encodes an oxidoreductase required for proper development of the Arabidopsis vegetative shoot apex. *Plant J.*, **6**, 835–847.

Chaves, M.M., Maroco, J.P. and Pereira, J.S. (2003) Understanding plant responses to drought — from genes to the whole plant. *Funct. Plant Biol.*, **30**, 239–264.

Chia, T., Thorneycroft, D., Chapple, A., Messerli, G., Chen, J., Zeeman, S.C., Smith, S.M. and Smith, A.M. (2004) A cytosolic glucosyltransferase is required for conversion of starch to sucrose in Arabidopsis leaves at night. *Plant J.*, **37**, 853–863.

Christman, M.A., Richards, J.H., Mckay, J.K., Stahl, E.A., Juenger, T.E. and Donovan, L.A. (2008) Genetic variation in Arabidopsis thaliana for night-time leaf conductance. *Plant Cell Environ.*, **31**, 1170–1178.

Clauw, P., Coppens, F., De Beuf, K., et al. (2015) Leaf responses to mild drought stress in natural variants of Arabidopsis. *Plant Physiol.*, **167**, 800–816.

Corvalán, C. and Choe, S. (2017) Identification of brassinosteroid genes in Brachypodium distachyon. *BMC Plant Biol.*, **17**, 5.

Davière, J.-M. and Achard, P. (2013) Gibberellin signaling in plants. *Development*, **140**, 1147–1151.

De Vylder, J., Vandenbussche, F., Hu, Y., Philips, W. and Van Der Straeten, D. (2012) Rosette Tracker: An Open Source Image Analysis Tool for Automatic Quantification of Genotype Effects. *Plant Physiol.*, **160**, 1149–1159.

Dittberner, H., Korte, A., Mettler-Altmann, T., Weber, A.P.M., Monroe, G. and Meaux, J. de (2018) Natural variation in stomata size contributes to the local adaptation of water-use efficiency in *Arabidopsis thaliana*. *Mol Ecol., **20**, 4052–4065.*

Du, J., Johnson, L.M., Jacobsen, S.E. and Patel, D.J. (2015) DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.*, **16**, 519–532.

Durvasula, A., Fulgione, A., Gutaker, R.M., et al. (2017) African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.*, **114**, 5213–5218.

Eisele, J.F., Fäßler, F., Bürgel, P.F. and Chaban, C. (2016) A Rapid and Simple Method for Microscopy-Based Stomata Analyses. *PLoS ONE*, **11**, e0164576.

Elfarargi, A.F., Gilbault, E., Döring, N., Neto, C., Fulgione, A., Weber, A.P.M., Loudet, O. and Hancock, A.M. (2023) Genomic Basis of Adaptation to a Novel Precipitation Regime. *Mol. Biol. Evol.*, **40**, msad031.

Evans, J.R., Kaldenhoff, R., Genty, B. and Terashima, I. (2009) Resistances along the CO2 diffusion pathway inside leaves. *J. Exp. Bot.*, **60**, 2235–2248.

Faragó, D., Sass, L., Valkai, I., Andrási, N. and Szabados, L. (2018) PlantSize Offers an Affordable, Non-destructive Method to Measure Plant Size and Color in Vitro. *Front. Plant Sci.*, **9**.

Farquhar, G.D. and Sharkey, T.D. (1982) Stomatal Conductance and Photosynthesis. *Annu. Rev. Plant Physiol.*, **33**, 317–345.

Flood, P.J., Kruijer, W., Schnabel, S.K., Schoor, R. van der, Jalink, H., Snel, J.F.H., Harbinson, J. and Aarts, M.G.M. (2016) Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. *Plant Methods*, **12**, 14.

Franks, P.J. and Beerling, D.J. (2009) Maximum leaf conductance driven by CO2 effects on stomatal size and density over geologic time. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 10343–7.

Franks, P.J., Drake, P.L. and Beerling, D.J. (2009) Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: an analysis using Eucalyptus globulus. *Plant Cell Environ.*, **32**, 1737–1748.

Fulgione, A., Neto, C., Elfarargi, A.F., et al. (2022) Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. *Nat. Commun.*, **13**, 1461.

Gazzani, S., Gendall, A.R., Lister, C. and Dean, C. (2003) Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. *Plant Physiol.*, **132**, 1107–1114.

Hancock, A.M., Brachi, B., Faure, N., Horton, M.W., Jarymowycz, L.B., Sperone, F.G., Toomajian, C., Roux, F. and Bergelson, J. (2011) Adaptation to climate across the Arabidopsis thaliana genome. *Science*, **334**, 83–6.

Hanemian, M., Vasseur, F., Marchadier, E., Gilbault, E., Bresson, J., Gy, I., Violle, C. and Loudet, O. (2020) Natural variation at FLM splicing has pleiotropic effects modulating ecological strategies in Arabidopsis thaliana. *Nat. Commun.*, **11**, 4140.

He, L., Su, C., Wang, Y. and Wei, Z. (2015) ATDOF5.8 protein is the upstream regulator of ANAC069 and is responsive to abiotic stress. *Biochimie*, **110**, 17–24.

Horie T., Matsuura S., Takai T., Kuwasaki K., Ohsumi A., and Shiraiwa T. (2006) Genotypic difference in canopy diffusive conductance measured by a new remote-sensing method and its association with the difference in rice yield potential. *Plant Cell Environ.*, **29**, 653–660.

Huang, M.-D. and Wu, W.-L. (2007) Overexpression of TMAC2, a novel negative regulator of abscisic acid and salinity responses, has pleiotropic effects in Arabidopsis thaliana. *Plant Mol. Biol.*, **63**, 557–569.

Hung, F.-Y., Lai, Y.-C., Wang, J., et al. (2021) The Arabidopsis histone demethylase JMJ28 regulates CONSTANS by interacting with FBH transcription factors. *Plant Cell*, **33**, 1196–1211.

Jakobson, L., Vaahtera, L., Tõldsepp, K., et al. (2016) Natural Variation in Arabidopsis Cvi-0 Accession Reveals an Important Role of MPK12 in Guard Cell CO2 Signaling. *PLOS Biol.*, **14**, e2000322.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and 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.

Juenger, T.E., Mckay, J.K., Hausmann, N., Keurentjes, J.J., Sen, S., Stowe, K.A., Dawson, T.E., Simms, E.L. and Richards, J.H. (2005) Identification and characterization of QTL underlying whole-plant physiology in Arabidopsis thaliana: δ13C, stomatal conductance and transpiration efficiency. *Plant Cell Environ.*, **28**, 697–708.

Kahle, D. and Wickham, H. (2013) ggmap: Spatial Visualization with ggplot2. *R J.*, **5**, 144–161.

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

Kenney, A.M., McKay, J.K., Richards, J.H. and Juenger, T.E. (2014) Direct and indirect selection on flowering time, water-use efficiency (WUE, δ13C), and WUE plasticity to drought in Arabidopsis thaliana. *Ecol. Evol.*, **4**, 4505–4521.

Komori, R., Amano, Y., Ogawa-Ohnishi, M. and Matsubayashi, Y. (2009) Identification of tyrosylprotein sulfotransferase in Arabidopsis. *Proc. Natl. Acad. Sci.*, **106**, 15067–15072.

Kronholm, I., Picó, F.X., Alonso-Blanco, C., Goudet, J. and Meaux, J. de (2012) Genetic basis of adaptation in Arabidopsis thaliana: local adaptation at the seed dormancy QTL DOG1. *Evol. Int. J. Org. Evol.*, **66**, 2287–2302.

Lawlor, D.W. (2013) Genetic engineering to improve plant performance under drought: physiological evaluation of achievements, limitations, and possibilities. *J. Exp. Bot.*, **64**, 83–108.

Lee, E.-J., Kim, K.Y., Zhang, J., et al. (2021) Arabidopsis seedling establishment under waterlogging requires ABCG5-mediated formation of a dense cuticle layer. *New Phytol.*, **229**, 156–172.

Levitt, J. (1972) *Responses of plants to environmental stresses (Physiological Ecology): Chilling, freezing, and high temperature stresses*. New York (NY): Academic Press. 697 pp.

Li, P., Li, Y.-J., Zhang, F.-J., Zhang, G.-Z., Jiang, X.-Y., Yu, H.-M. and Hou, B.-K. (2017) The Arabidopsis UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. *Plant J.*, **89**, 85–103.

Li, W.-X., Oono, Y., Zhu, J., et al. (2008) The Arabidopsis NFYA5 Transcription Factor Is Regulated Transcriptionally and Posttranscriptionally to Promote Drought Resistance. *Plant Cell*, **20**, 2238–2251.

Liang, X., Zhang, L., Natarajan, S.K. and Becker, D.F. (2013) Proline mechanisms of stress survival. *Antioxid. Redox Signal.*, **19**, 998–1011.

Lin, Y.-P., Lee, T., Tanaka, A. and Charng, Y. (2014) Analysis of an Arabidopsis heat-sensitive mutant reveals that chlorophyll synthase is involved in reutilization of chlorophyllide during chlorophyll turnover. *Plant J.*, **80**, 14–26.

Lin, Y.-P., Shen, Y.-Y., Shiu, Y.-B., Charng, Y. and Grimm, B. (2022) Chlorophyll dephytylase 1 and chlorophyll synthase: a chlorophyll salvage pathway for the turnover of photosystems I and II. *Plant J.*, **111**, 979–994.

Liu, C. and Mehdy, M.C. (2007) A Nonclassical Arabinogalactan Protein Gene Highly Expressed in Vascular Tissues, AGP31, Is Transcriptionally Repressed by Methyl Jasmonic Acid in Arabidopsis. *Plant Physiol.*, **145**, 863–874.

Liu, Y., Maierhofer, T., Rybak, K., et al. (2019) Anion channel SLAH3 is a regulatory target of chitin receptor-associated kinase PBL27 in microbial stomatal closure J.-M. Zhou, C. S. Hardtke, and S. Hoth, eds. *eLife*, **8**, e44474.

Lobin W (1983) The occurrence of Arabidopsis thaliana in the Cape Verde Islands. *Arab Info Ser*, **20**, 119–123.

Lopez-Anido, C.B., Vatén, A., Smoot, N.K., Sharma, N., Guo, V., Gong, Y., Anleu Gil, M.X., Weimer, A.K. and Bergmann, D.C. (2021) Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf. *Dev. Cell*, **56**, 1043-1055.e4.

Loudet, O., Saliba-Colombani, V., Camilleri, C., Calenge, F., Gaudon, V., Koprivova, A., North, K.A., Kopriva, S. and Daniel-Vedele, F. (2007) Natural variation for sulfate content in Arabidopsis thaliana is highly controlled by APR2. *Nat. Genet.*, **39**, 896–900.

Lovell, J.T., Juenger, T.E., Michaels, S.D., et al. (2013) Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Proc. R. Soc. B Biol. Sci.*, **280**, 20131043.

Ludlow, M. (1989) *Strategies of response to water stress*. The Hague, Netherlands: SPB Academic Press.

Marchadier, E., Hanemian, M., Tisné, S., Bach, L., Bazakos, C., Gilbault, E., Haddadi, P., Virlouvet, L. and Loudet, O. (2019) The complex genetic architecture of shoot growth natural variation in Arabidopsis thaliana. *PLOS Genet.*, **15**, e1007954.

Martínez, J.P., Silva, H., Ledent, J.F. and Pinto, M. (2007) Effect of drought stress on the osmotic adjustment, cell wall elasticity and cell volume of six cultivars of common beans (Phaseolus vulgaris L.). *Eur. J. Agron.*, **26**, 30–38.

Massonnet, C., Dauzat, M., Bédiée, A., Vile, D. and Granier, C. (2015) Individual Leaf Area of Early Flowering Arabidopsis Genotypes Is More Affected by Drought than Late Flowering Ones: A Multi-Scale Analysis in 35 Genetically Modified Lines. *Am. J. Plant Sci.*, **6**, 955–971.

McKay, J.K., Richards, J.H. and Mitchell-Olds, T. (2003) Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.*, **12**, 1137–1151.

McKay, J.K., Richards, J.H., Nemali, K.S., Sen, S., Mitchell-Olds, T., Boles, S., Stahl, E.A., Wayne, T. and Juenger, T.E. (2008) Genetics of Drought Adaptation in Arabidopsis Thaliana Ii. Qtl Analysis of a New Mapping Population, Kas-1 × Tsu-1. *Evolution*, **62**, 3014–3026.

Meinzer, F.C. (1993) Stomatal control of transpiration. *Trends Ecol. Evol.*, **8**, 289–294.

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

Mitchum, M.G., Yamaguchi, S., Hanada, A., Kuwahara, A., Yoshioka, Y., Kato, T., Tabata, S., Kamiya, Y. and Sun, T. (2006) Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. *Plant J.*, **45**, 804–818.

Monda, K., Negi, J., Iio, A., Kusumi, K., Kojima, M., Hashimoto, M., Sakakibara, H. and Iba, K. (2011) Environmental regulation of stomatal response in the Arabidopsis Cvi-0 ecotype. *Planta*, **234**, 555–563.

Moore, C.R., Johnson, L.S., Kwak, I.-Y., Livny, M., Broman, K.W. and Spalding, E.P. (2013) High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics*, **195**, 1077–1086.

Muraya, M.M., Chu, J., Zhao, Y., Junker, A., Klukas, C., Reif, J.C. and Altmann, T. (2017) Genetic variation of growth dynamics in maize (Zea mays L.) revealed through automated non-invasive phenotyping. *Plant J. Cell Mol. Biol.*, **89**, 366–380.

Nakashima, K., Shinwari, Z.K., Sakuma, Y., Seki, M., Miura, S., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) Organization and expression of two Arabidopsis DREB2 genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression. *Plant Mol. Biol.*, **42**, 657–665.

Nikonorova, N., Broeck, L.V. den, Zhu, S., Cotte, B. van de, Dubois, M., Gevaert, K., Inzé, D. and Smet, I.D. (2018) Early mannitol-triggered changes in the Arabidopsis leaf (phospho)proteome reveal growth regulators. *J. Exp. Bot.*, **69**, 4591.

Nishimura, N., Kitahata, N., Seki, M., Narusaka, Y., Narusaka, M., Kuromori, T., Asami, T., Shinozaki, K. and Hirayama, T. (2005) Analysis of ABA Hypersensitive Germination2 revealed the pivotal functions of PARN in stress response in Arabidopsis. *Plant J.*, **44**, 972–984.

Oh, E., Yamaguchi, S., Hu, J., et al. (2007) PIL5, a Phytochrome-Interacting bHLH Protein, Regulates Gibberellin Responsiveness by Binding Directly to the GAI and RGA Promoters in Arabidopsis Seeds. *Plant Cell*, **19**, 1192–1208.

Ossowski, S., Schneeberger, K., Clark, R.M., Lanz, C., Warthmann, N. and Weigel, D. (2008) Sequencing of natural strains of Arabidopsis thaliana with short reads. *Genome Res.*, **18**, 2024–2033.

Pan, J., Wang, H., Hu, Y. and Yu, D. (2018) Arabidopsis VQ18 and VQ26 proteins interact with ABI5 transcription factor to negatively modulate ABA response during seed germination. *Plant J.*, **95**, 529–544.

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

Patwari, P., Salewski, V., Gutbrod, K., et al. (2019) Surface wax esters contribute to drought tolerance in Arabidopsis. *Plant J.*, **98**, 727–744.

Price, A.H., Cairns, J.E., Horton, P., Jones, H.G. and Griffiths, H. (2002) Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach:

progress and new opportunities to integrate stomatal and mesophyll responses. *J. Exp. Bot.*, **53**, 989–1004.

Purcell, S., Neale, B., Todd-Brown, K., et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–75.

Qin, L.-X., Li, Y., Li, D.-D., Xu, W.-L., Zheng, Y. and Li, X.-B. (2014) Arabidopsis drought-induced protein Di19-3 participates in plant response to drought and high salinity stresses. *Plant Mol. Biol.*, **86**, 609–625.

QUARRIE, S.A. and JONES, H.G. (1977) Effects of Abscisic Acid and Water Stress on Development and Morphology of Wheat. *J. Exp. Bot.*, **28**, 192–203.

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

Ribot, C., Wang, Y. and Poirier, Y. (2008) Expression analyses of three members of the AtPHO1 family reveal differential interactions between signaling pathways involved in phosphate deficiency and the responses to auxin, cytokinin, and abscisic acid. *Planta*, **227**, 1025–1036.

Rolly, N.K., Imran, Q.M., Shahid, M., Imran, M., Khan, M., Lee, S.-U., Hussain, A., Lee, I.-J. and Yun, B.-W. (2020) Drought-induced AtbZIP62 transcription factor regulates drought stress response in Arabidopsis. *Plant Physiol. Biochem.*, **156**, 384–395.

Sahni, S., Prasad, B.D., Liu, Q., Grbic, V., Sharpe, A., Singh, S.P. and Krishna, P. (2016) Overexpression of the brassinosteroid biosynthetic gene DWF4 in Brassica napus simultaneously increases seed yield and stress tolerance. *Sci. Rep.*, **6**, 28298.

Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2006) Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression. *Plant Cell*, **18**, 1292–1309.

Sass, L., Majer, P. and Hideg, É. (2012) Leaf Hue Measurements: A High-Throughput Screening of Chlorophyll Content. In J. Normanly, ed. *High-Throughput Phenotyping in Plants: Methods and Protocols*. Methods in Molecular Biology. Totowa, NJ: Humana Press, pp. 61–69. Available at: https://doi.org/10.1007/978-1-61779-995-2_6.

Schulze, E.D. (1986) Carbon Dioxide and Water Vapor Exchange in Response to Drought in the Atmosphere and in the Soil. *Annu. Rev. Plant Physiol.*, **37**, 247–274.

SCOTT, P. (2000) Resurrection Plants and the Secrets of Eternal Leaf. *Ann. Bot.*, **85**, 159–166.

Setter, T.L. (2012) Analysis of Constituents for Phenotyping Drought Tolerance in Crop Improvement. *Front. Physiol.*, **3**, 180.

Sharma, S., Villamor, J.G. and Verslues, P.E. (2011) Essential role of tissue-specific proline synthesis and catabolism in growth and redox balance at low water potential. *Plant Physiol.*, **157**, 292–304.

Shen, L., Kang, Y.G.G., Liu, L. and Yu, H. (2011) The J-Domain Protein J3 Mediates the Integration of Flowering Signals in Arabidopsis. *Plant Cell*, **23**, 499–514.

Shen, L. and Yu, H. (2011) J3 regulation of flowering time is mainly contributed by its activity in leaves. *Plant Signal. Behav.*, **6**, 601–603.

Sherrard, M.E. and Maherali, H. (2006) The Adaptive Significance of Drought Escape in Avena barbata, an Annual Grass. *Evolution*, **60**, 2478–2489.

Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M. and 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.

Shimizu, T., Kanno, Y., Suzuki, H., Watanabe, S. and Seo, M. (2021) Arabidopsis NPF4.6 and NPF5.1 Control Leaf Stomatal Aperture by Regulating Abscisic Acid Transport. *Genes*, **12**, 885.

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

Somerville, C. and Koornneef, M. (2002) A fortunate choice: the history of Arabidopsis as a model plant. *Nat. Rev. Genet.*, **3**, 883–889.

Spence, R.D., Wu, H., Sharpe, P.J.H. and Clark, K.G. (1986) Water stress effects on guard cell anatomy and the mechanical advantage of the epidermal cells. *Plant Cell Environ.*, **9**, 197–202.

Stebbins Jr, G.L. (1952) Aridity as a stimulus to plant evolution. *Am. Nat.*, **86**, 33–44.

Tal, I., Zhang, Y., Jørgensen, M.E., et al. (2016) The Arabidopsis NPF3 protein is a GA transporter. *Nat. Commun.*, **7**, 11486.

Tanaka, Y., Kumagai, E., Tazoe, Y., Adachi, S. and Homma, K. (2014) Leaf Photosynthesis and Its Genetic Improvement from the Perspective of Energy Flow and CO_2 Diffusion. *Plant Prod. Sci.*, **17**, 111–123.

Tergemina, E., Elfarargi, A.F., Flis, P., et al. (2022) A two-step adaptive walk rewires nutrient transport in a challenging edaphic environment. *Sci. Adv.*, **8**, eabm9385.

Tisné, S., Serrand, Y., Bach, L., et al. (2013) Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J.,* **74**, 534–544.

Tuberosa, R. (2012) Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.,* **3**, 347.

Vashisht, D., Hesselink, A., Pierik, R., et al. (2011) Natural variation of submergence tolerance among Arabidopsis thaliana accessions. *New Phytol.,* **190**, 299–310.

Vidigal, D.S., Marques, A.C.S.S., Willems, L.A.J., Buijs, G., Méndez-Vigo, B., Hilhorst, H.W.M., Bentsink, L., Picó, F.X. and Alonso-Blanco, C. (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.

Walter, A., Scharr, H., Gilmer, F., et al. (2007) Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. *New Phytol.,* **174**, 447–455.

Wang, Q., Liu, P., Jing, H., Zhou, X.F., Zhao, B., Li, Y. and Jin, J.B. (2021) JMJ27-mediated histone H3K9 demethylation positively regulates drought-stress responses in Arabidopsis. *New Phytol.,* **232**, 221–236.

Wang, T.-Y., Wu, J.-R., Duong, N.K.T., Lu, C.-A., Yeh, C.-H. and Wu, S.-J. (2021) HSP70-4 and farnesylated AtJ3 constitute a specific HSP70/HSP40-based chaperone machinery essential for prolonged heat stress tolerance in Arabidopsis. *J. Plant Physiol.,* **261**, 153430.

Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis. *Springer-Verl. N. Y.*

Xiong, F., Ren, J.-J., Yu, Q., Wang, Y.-Y., Lu, C.-C., Kong, L.-J., Otegui, M.S. and Wang, X.-L. (2019) AtU2AF65b functions in abscisic acid mediated flowering via regulating the precursor messenger RNA splicing of ABI5 and FLC in Arabidopsis. *New Phytol.,* **223**, 277–292.

Xu, C., Shan, J., Liu, T., Wang, Q., Ji, Y., Zhang, Y., Wang, M., Xia, N. and Zhao, L. (2023) CONSTANS-LIKE 1a positively regulates salt and drought tolerance in soybean. *Plant Physiol.,* **191**, 2427–2446.

Xu, Z. and Zhou, G. (2008) Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass. *J. Exp. Bot.,* **59**, 3317–25.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994) A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell*, **6**, 251–264.

Zhang, X., Hause, R.J. and Borevitz, J.O. (2012) Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in Arabidopsis thaliana. *G3 Bethesda Md*, **2**, 29–34.

Zhang, X., Huang, C., Wu, D., et al. (2017) High-Throughput Phenotyping and QTL Mapping Reveals the Genetic Architecture of Maize Plant Growth. *Plant Physiol.*, **173**, 1554–1564.

Zhang, X., Zou, Z., Gong, P., Zhang, J., Ziaf, K., Li, H., Xiao, F. and Ye, Z. (2011) Over-expression of microRNA169 confers enhanced drought tolerance to tomato. *Biotechnol. Lett.*, **33**, 403–409.

Zheng, J., Chen, F., Wang, Z., Cao, H., Li, X., Deng, X., Soppe, W.J.J., Li, Y. and Liu, Y. (2012) A novel role for histone methyltransferase KYP/SUVH4 in the control of Arabidopsis primary seed dormancy. *New Phytol.*, **193**, 605–616.

Zhou, X. and Stephens, M. (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods*, **11**, 407–409.

Zhou, X. and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.*, **44**, 821–824.

3.9 Supplementary figures
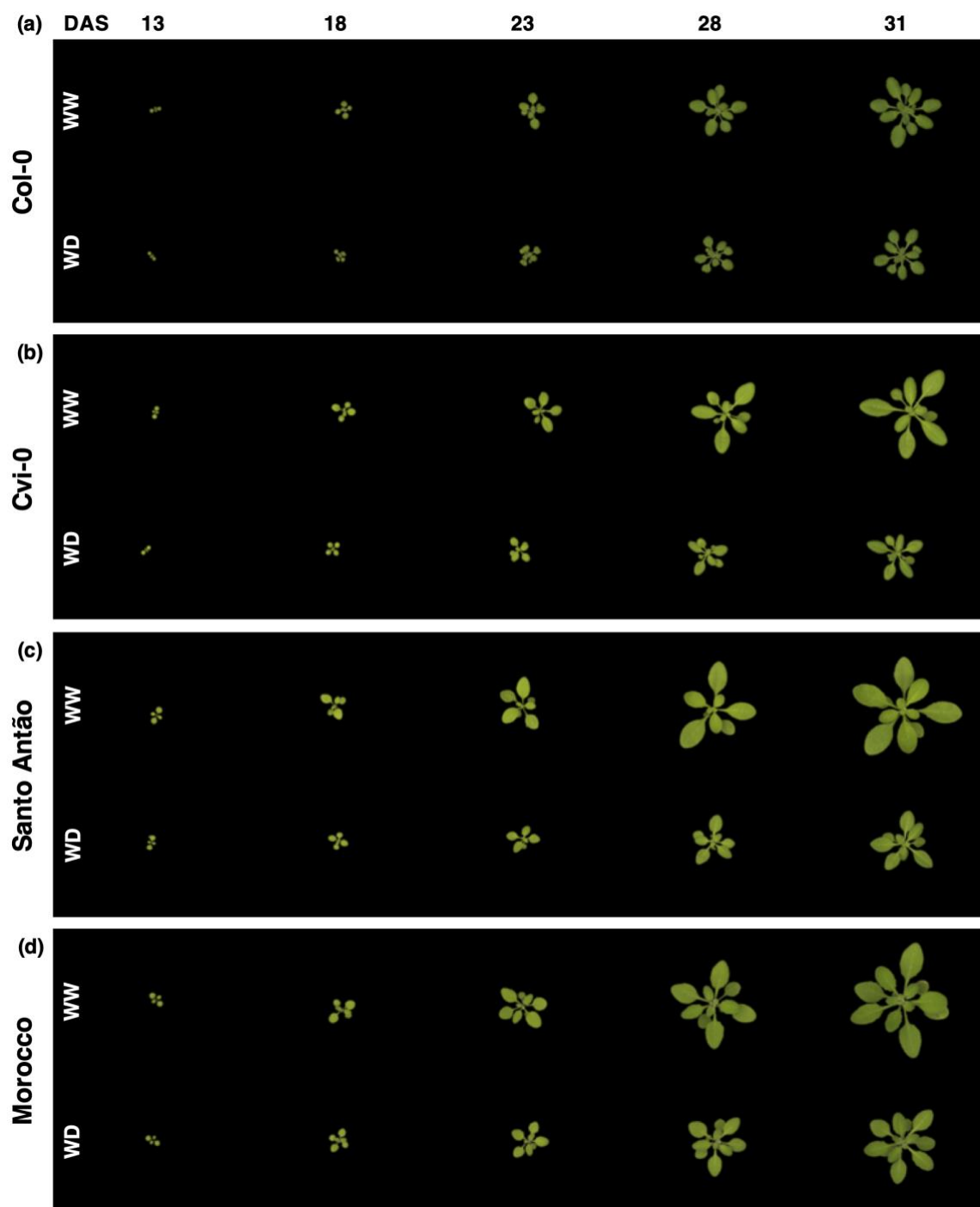

Supplementary Figure 1. Images of representative accessions on time-points (13, 18, 23, 28, and 31 Days After Sowing (DAS)) of Phenoscope drought experiment in well-watered (WW) and water deficit (WD) conditions. (a) Col-0, (b) Cvi-0, (c) S15-3 as a representative accession from Santo Antão, (b) Cvi-0 from Santo Antão, and (d) Elh-2 as a representative accession from Morocco.
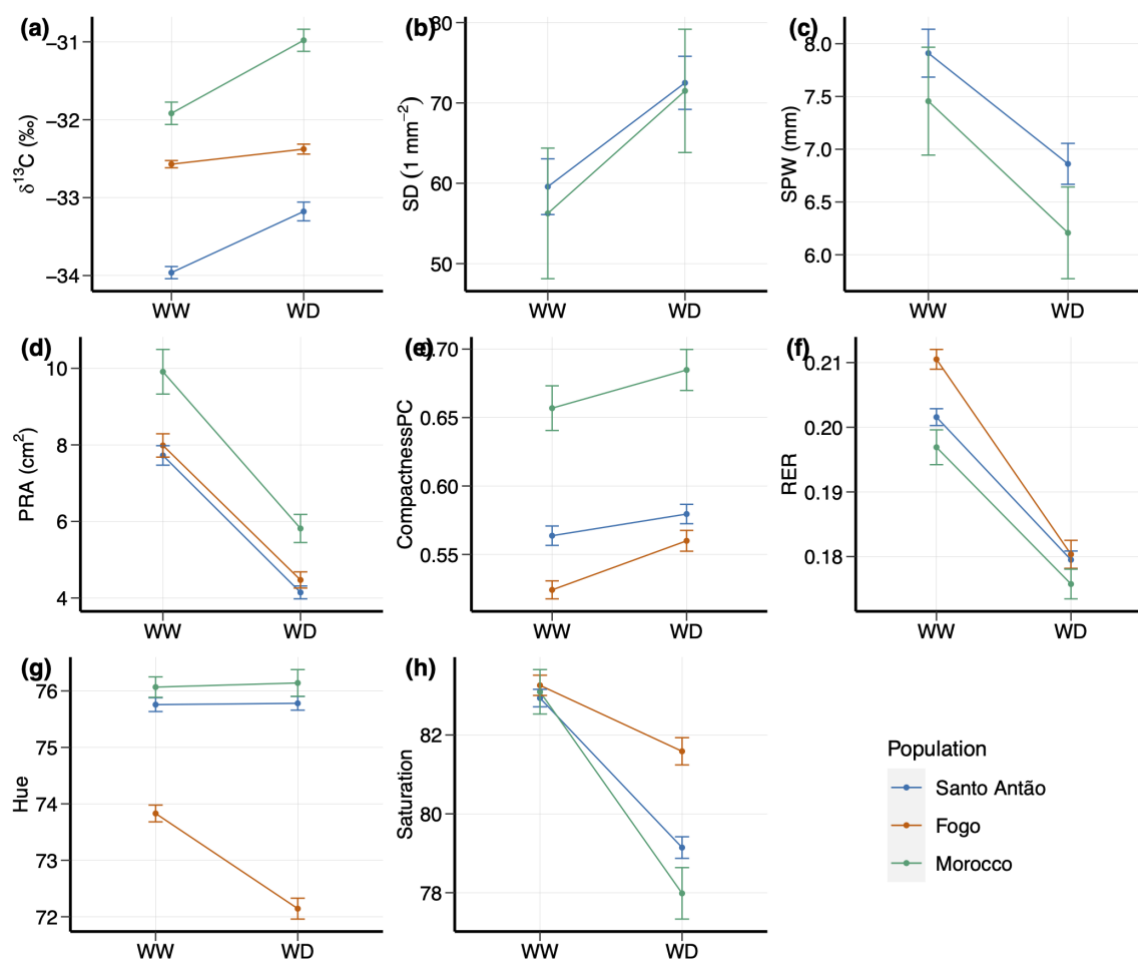
Supplementary Figure 1. Phenotypic plasticity among Santo Antão, Fogo, and Morocco A. thaliana population under well-watered (WW) and water deficit (WD) conditions. (a) Water Use Efficiency (WUE) is measured as carbon isotope discrimination (δ13C), and the carbon isotope ratio is expressed per mil, ‰. (b) Stomatal density (SD). (c) Stomatal pore width (SPW). (d) projected rosette area (PRA). (e) compactnessPC (ratio of PRA (P) to surface of convex hull area (C)). (f) relative expansion rate (RER), the growth rate of PRA; integrated over the 18-31 Days' time window. (g-h) Hue is a proxy for the chlorophyll content and saturation its color intensity. The points represent the means, and the whiskers represent the 95% CI.
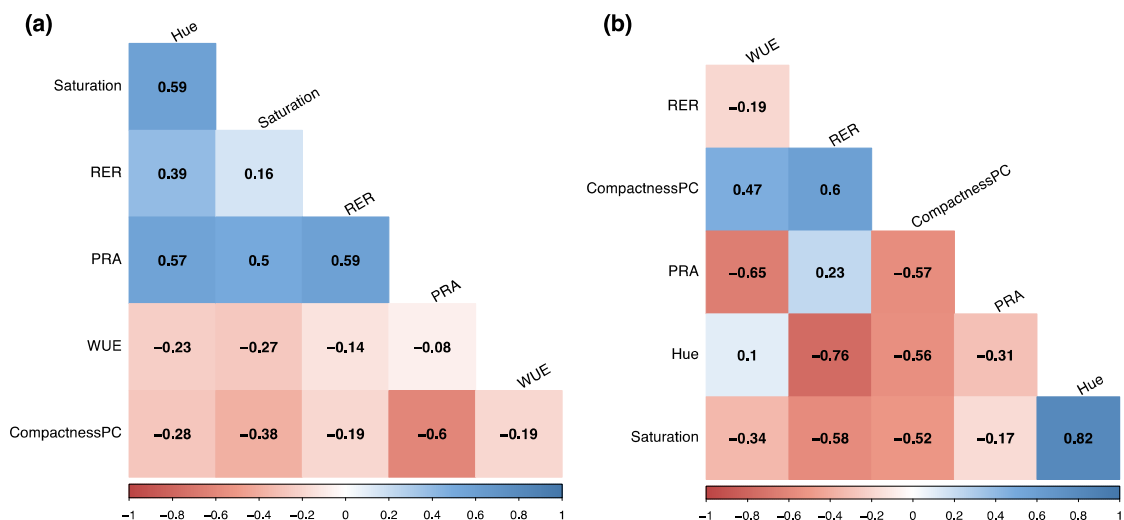
Supplementary Figure 2. Pairwise Pearson genetic correlation of the average phenotypic traits in (a) Fogo and (b) Morocco. PRA, projected rosette area; WUE, water use efficiency; SD, stomatal density; SPW, stomatal pore width; RER, relative expansion rate. The color spectrum, dark red to dark blue, represents highly negative to highly positive correlations, and the number represents the correlation values.
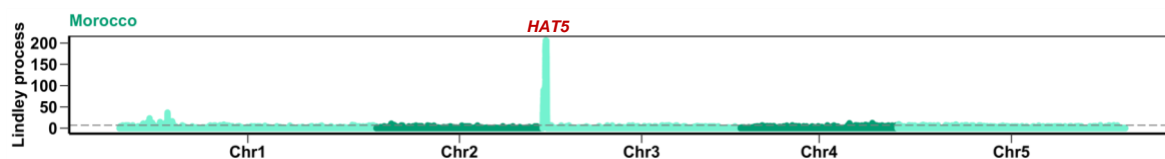
Supplementary Figure 3. Genetic basis of WUE in Morocco in well-watered condition. Manhattan plot for WUE in Morocco in well-watered condition. The x-axis and y-axis represent the chromosomes and Lindley score from the local score approach. The dashed line corresponds to a genome-wide Bonferroni significance.
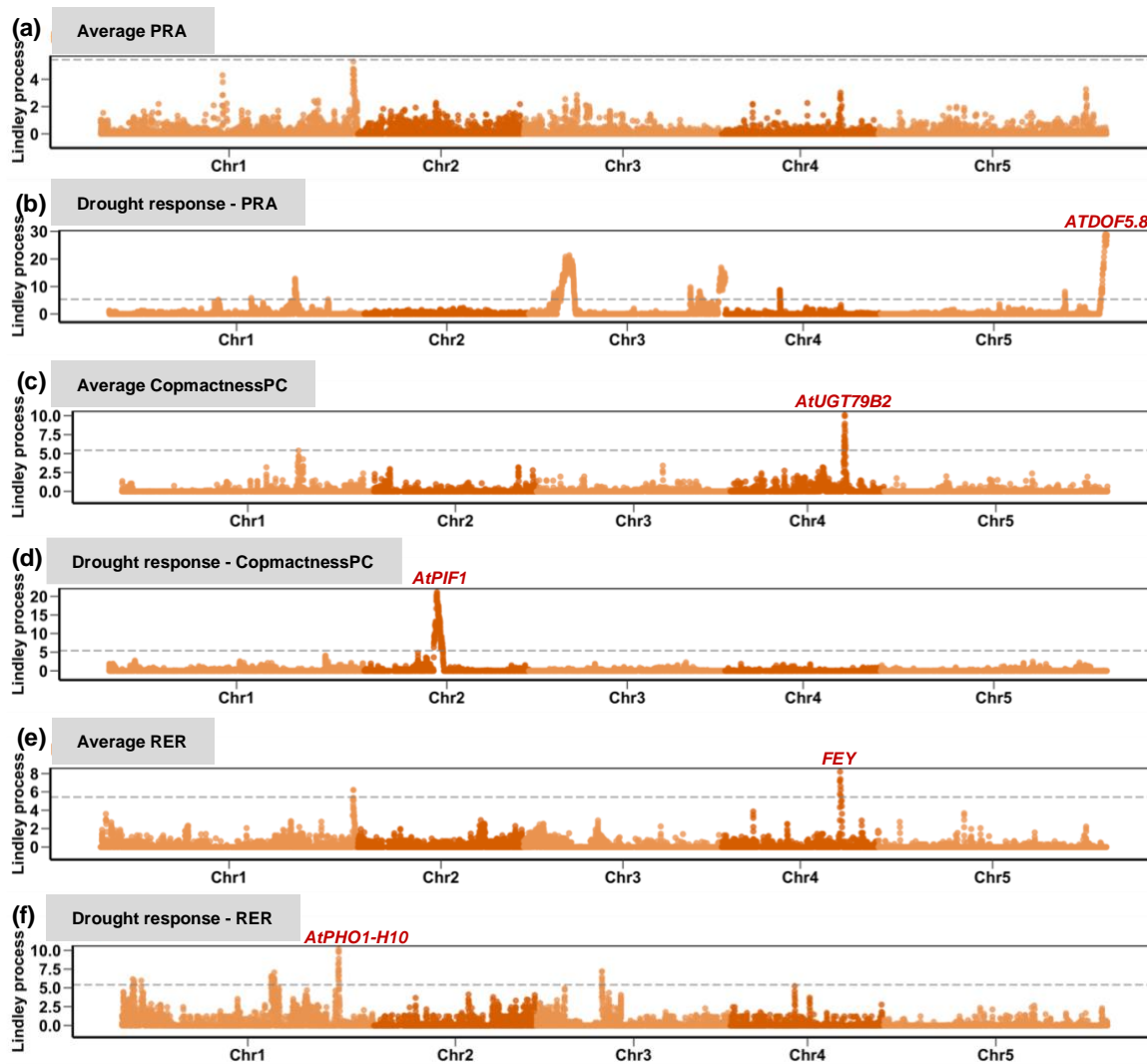
Supplementary Figure 4. Genomic associations of rosette growth-related traits (PRA, compactnessPC, and RER) in Fogo Antão using LMM in GEMMA followed by the local score approach. The x-axis and y-axis represent the chromosomes and Lindley score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.
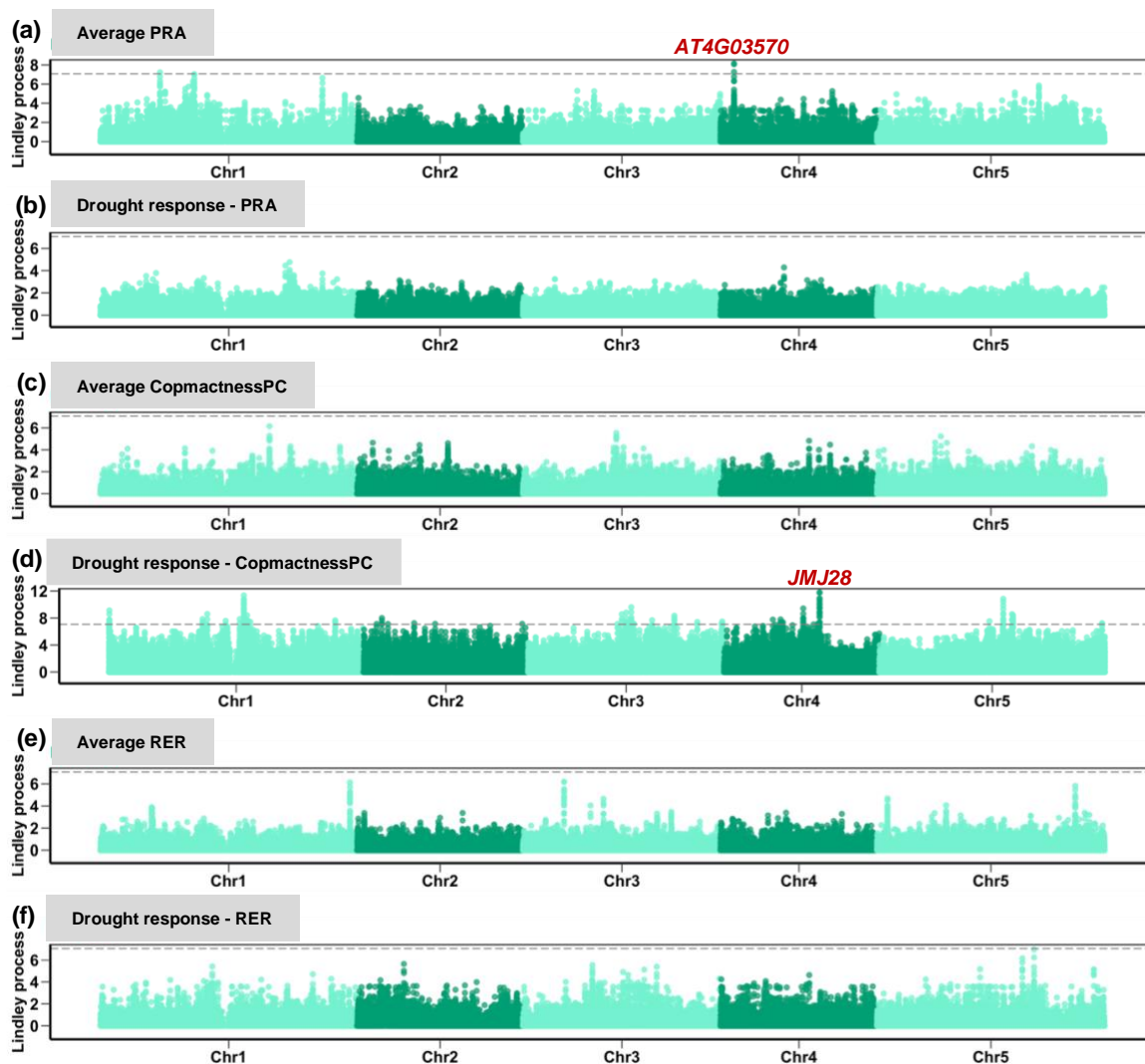
Supplementary Figure 5. Genomic associations of rosette growth-related traits (PRA, compactnessPC, and RER) in Morocco using LMM in GEMMA followed by the local score approach. The x-axis and y-axis represent the chromosomes and Lindley score from the local score approach. The dashed line in plots corresponds to a genome-wide Bonferroni significance.

Chapter 4

General discussion

In this thesis, the aim was to investigate the genetic basis and adaptive relevance of drought response in Cape Verdean *Arabidopsis thaliana* populations by using a genome-wide association (GWA) mapping approach, which utilizes natural variation to identify genes that account for natural variation in plant performance (morphological and physiological traits) under drought conditions. In this general discussion, I will present the key findings of the study and compare them to previous work in the fields of plant phenotyping, quantitative genetics, and evolutionary biology. Further, I will focus on the strong associations detected by GWA mapping and the candidate genes identified and understand their ecological roles.

4.1 Decoding adaptation strategies in Cape Verde *Arabidopsis* populations

The physiological and molecular mechanisms underlying the adaptation of natural populations to novel environments are poorly understood (Scheffers *et al.*, 2016). One possibility for natural populations to adapt to new environments is by relying on the emergence of new mutations (Hermisson and Pennings, 2005; Przeworski *et al.*, 2005). Natural island populations provide robust systems to explore the biological significance of genetic variations in adaptation. A single *A. thaliana* line was collected from the Cape Verde Islands (Cvi-0) 40 years ago (Lobin W, 1983) and has been an enigma to the *A. thaliana* community due to its phenotypic and climatic divergence as well as its surprising location far from Eurasia. A recently discovered *A. thaliana* population from CVI is confined to the islands of Santo Antão and Fogo (Fulgione *et al.*, 2022). Previous evolutionary reconstructions, combined with environmental data, suggest that *A. thaliana* colonized Santo Antão approximately 5-7 thousand years ago (kya) before adapting to Fogo, an active stratovolcano with a challenging edaphic environment, around 3-5 kya (Fulgione *et al.*, 2022). Upon colonizing both islands, *A. thaliana* plants encountered a sudden shift to extreme climate conditions characterized by a long dry season where moisture is delivered primarily from humid tradewinds, reduced photoperiod and growing season, and altered edaphic factors (Fulgione *et al.*, 2022).

4.2   Polygenic basis of drought-related traits

In this thesis, GWA analyses revealed that drought-related traits possess a polygenic nature ((Elfarargi *et al.*, 2023), Chapter 3). Similar findings have been observed in other GWA studies focusing on various traits in plants (Rasamivelona *et al.*, 1995; Marwede *et al.*, 2004; Atwell *et al.*, 2010; Alonso-Blanco *et al.*, 2016; Bartoli and Roux, 2017; Mayer *et al.*, 2020; Sun *et al.*, 2020; Wang *et al.*, 2020; Saini *et al.*, 2022; Xiong *et al.*, 2023), as well as in animals, including humans (Hirschhorn *et al.*, 2002; Goddard and Hayes, 2009; Vink *et al.*, 2014; Escott-Price *et al.*, 2015; Loh *et al.*, 2015; Field *et al.*, 2016; Boyle *et al.*, 2017; Bosse *et al.*, 2017; Martin *et al.*, 2017; Zeng *et al.*, 2021). These studies indicate that multiple genes contribute to the majority of traits. Considering the highly polygenic nature of many traits, adaptive responses to changing selective pressures often necessitate shifts in these traits, facilitated by alterations in allele frequencies at multiple underlying loci. As a result, polygenic adaptation in complex traits is expected to be widespread. This hypothesis is supported by over a century of research on the response to artificial selection for various traits in plants and animals.

Polygenic adaptation at the edges in populations facing extreme drought is a critical mechanism for species to persist in challenging environments. These adaptations involve multiple genes working together to enable organisms to cope with environmental stressors. In *A. thaliana*, populations at the edge of their range, such as those in the CVI islands, have shown evidence of rapid adaptation and evolutionary rescue in response to extreme drought conditions (Fulgione *et al.*, 2022). The adaptation in these populations is likely to be polygenic, involving multiple genes that contribute to physiological and developmental traits associated with drought response ((Elfarargi *et al.*, 2023), Chapter 3). Such adaptations may include adjustments in water-use strategies, root architecture optimization, and alterations in reproductive timing (Franks *et al.*, 2009; Nicotra *et al.*, 2010; Yeh and Price, 2004).

Overall, polygenic adaptation at the edges in populations facing extreme drought is an essential mechanism for plant species to cope with environmental stressors. Studying

these adaptations in *A. thaliana* and other plant species can provide valuable insights into the genetic and phenotypic factors contributing to resilience and persistence in changing environments.

4.3   Dissecting complex traits

We previously showed that a phenotypic shift towards higher stomatal conductance and lower WUE on Santo Antão island was due largely to a nonsynonymous variant in the *MPK12* gene (Elfarargi *et al.*, 2023). In this study, by including the *MPK12* G53R variant as a covariate in the model, we uncovered additional significant associations with *VQ18*, *DI19-3*, and *PYL8* genes, all of which are implicated in the abscisic acid (ABA) signaling pathway. Previous research has established that *MPK12* exhibits pleiotropic effects on guard cell size and ABA response in *A. thaliana* (Des Marais *et al.*, 2014). Moreover, a single amino acid substitution in *MPK12* has been shown to result in constitutively larger stomata and altered responses to hormonal inhibition of stomatal opening, as well as short-term changes in vapor pressure deficit, ultimately leading to lower WUE and decreased fitness under water-limited conditions (Des Marais *et al.*, 2014). The expression of ABA receptors *PYL12*, a close relative to *PYL8*, and *RCAR6* have been reported to enhance WUE and growth rate in *A. thaliana* (Yang *et al.*, 2016). In both CVI islands and Morocco, we found several genes associated with ion homeostasis, ABA signaling, and mitogen-activated protein kinase (MAPK) activation, which are essential for plant growth and development under stress conditions.

Furthermore, several candidate genes associated with rosette growth-related traits underlying drought stress were identified in Santo Antão. These genes are involved in auxin signaling (*TPST*), starch metabolism (*DPE2*), cuticle formation (*ABCG5*), wax ester accumulation (*WSD1-like*), and stress-responsive signaling (*MPK12*, *DREB2B*, *ATJ3*). Additionally, associations were detected with genes involved in brassinosteroid biosynthesis (*AtDWF4*) and ABA hypersensitivity and stress responses (*AtAHG2/AtPARN*). In Fogo, we found several genes associated with rosette growth-related traits underlying drought. These genes include *AtDOF5.8*, *AtUGT79B2*, *AtPIF1*, *FEY*, and *AtPHO1-H10*, which are involved in abiotic stress responses. *AtDOF5.8* and *ANAC069* participate in a regulatory

network mediating abiotic stress responses, *AtUGT79B2* enhances plant tolerance to various stresses, *AtPIF1* influences ABA biosynthesis and signaling component genes, *FEY* plays a role in leaf positioning and meristem maintenance, and *AtPHO1-H10* is involved in phosphate homeostasis and ABA modulation.

4.4   Genetic complexity and pleiotropy

Pleiotropy refers to a single gene affecting multiple traits, which can lead to complex genetic architectures and relationships among traits (Paaby and Rockman, 2013). In this thesis, we showed that some genes such as *MPK12*, *SUVH4*, *FRI*, and *NHL26* have potential pleiotropic effects on water use efficiency, stomatal patterning, leaf color, and rosette growth underlying drought in CVI populations ((Elfarargi *et al.*, 2023), Chapter 3). Pleiotropy has been observed in various studies, and it plays a significant role in the adaptation to environmental changes, including drought stress. In *A. thaliana*, pleiotropy has been observed in multiple studies investigating the genetic basis of various traits, such as flowering time, disease resistance, and response to abiotic stresses, including drought (Juenger *et al.*, 2005; Atwell *et al.*, 2010). For example, genes involved in drought tolerance have also been found to affect other traits like water use efficiency, flowering time, plant growth, and photosynthesis (McKay *et al.*, 2003; Lovell *et al.*, 2013; Hanemian *et al.*, 2020). The presence of pleiotropy in drought-related traits implies that adaptation to drought may involve complex genetic networks and trade-offs among various traits.

In other plant species, pleiotropy has been reported in the context of drought stress. For instance, in rice, a single gene called *DRO1* has been shown to affect both root growth angle and drought avoidance (Uga *et al.*, 2013). In maize, pleiotropy has been observed in genes that contribute to both flowering time and drought tolerance (Buckler *et al.*, 2009; Wallace *et al.*, 2014). In animals, pleiotropy is also common, with genes influencing multiple traits like body size, fecundity, and life history traits (Hau, 2007; Wright *et al.*, 2010; Bolormaa *et al.*, 2014). In humans, pleiotropy has been observed in genes associated with complex traits and diseases, such as obesity, type 2 diabetes, and cardiovascular disease (Sivakumaran *et al.*, 2011; Solovieff *et al.*, 2013).

In conclusion, this work uncovered several significant associations between genetic loci and traits related to drought stress adaptation in *A. thaliana* populations on the Cape Verde Islands. Considering the previously identified *MPK12* G53R variant, we have identified additional associations with genes involved in the ABA signaling pathway, ion homeostasis, and MAPK activation. Additionally, our findings revealed candidate genes associated with rosette growth-related traits and leaf color traits in Santo Antão and Fogo, which play crucial roles in auxin signaling, starch metabolism, cuticle formation, wax ester accumulation, stress-responsive signaling, brassinosteroid biosynthesis, and ABA hypersensitivity. These results shed light on the complex genetic architecture underlying plant adaptation to drought stress and highlight the potential pleiotropic functions of some identified genes, which need further investigation to elucidate their precise roles in plant adaptation to challenging environmental conditions.

4.5   References

Alonso-Blanco, C., Andrade, J., Becker, C., et al. (2016) 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. *Cell*, **166**, 481–491.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., et al. (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, **465**, 627–631.

Bartoli, C. and Roux, F. (2017) Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. *Front. Plant Sci.*, **8**, 763.

Bolormaa, S., Pryce, J.E., Reverter, A., et al. (2014) A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for Stature, Fatness and Reproduction in Beef Cattle. *PLOS Genet.*, **10**, e1004198.

Bosse, M., Spurgin, L.G., Laine, V.N., et al. (2017) Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, **358**, 365–368.

Boyle, E.A., Li, Y.I. and Pritchard, J.K. (2017) An expanded view of complex traits: from polygenic to omnigenic. *Cell*, **169**, 1177–1186.

Buckler, E.S., Holland, J.B., Bradbury, P.J., et al. (2009) The Genetic Architecture of Maize Flowering Time. *Science*, **325**, 714–718.

Elfarargi, A.F., Gilbault, E., Döring, N., Neto, C., Fulgione, A., Weber, A.P.M., Loudet, O. and Hancock, A.M. (2023) Genomic Basis of Adaptation to a Novel Precipitation Regime. *Mol. Biol. Evol.*, **40**, msad031.

Escott-Price, V., Sims, R., Bannister, C., et al. (2015) Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain*, **138**, 3673–3684.

Field, Y., Boyle, E.A., Telis, N., et al. (2016) Detection of human adaptation during the past 2000 years. *Science*, **354**, 760–764.

Franks, P.J., Drake, P.L. and Beerling, D.J. (2009) Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: an analysis using Eucalyptus globulus. *Plant Cell Environ.*, **32**, 1737–1748.

Fulgione, A., Neto, C., Elfarargi, A.F., et al. (2022) Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. *Nat. Commun.*, **13**, 1461.

Goddard, M.E. and Hayes, B.J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet.*, **10**, 381–391.

Granier, C., Aguirrezabal, L., Chenu, K., et al. (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis

thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol.*, **169**, 623–635.

Hanemian, M., Vasseur, F., Marchadier, E., Gilbault, E., Bresson, J., Gy, I., Violle, C. and Loudet, O. (2020) Natural variation at FLM splicing has pleiotropic effects modulating ecological strategies in Arabidopsis thaliana. *Nat. Commun.*, **11**, 4140.

Harbinson, J., Prinzenberg, A.E., Kruijer, W. and Aarts, M.G.M. (2012) High throughput screening with chlorophyll fluorescence imaging and its use in crop improvement. *Curr. Opin. Biotechnol.*, **23**, 221–226.

Hau, M. (2007) Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. *BioEssays*, **29**, 133–144.

Hermisson, J. and Pennings, P.S. (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335–2352.

Hirschhorn, J.N., Lohmueller, K., Byrne, E. and Hirschhorn, K. (2002) A comprehensive review of genetic association studies. *Genet. Med.*, **4**, 45–61.

Juenger, T.E., Mckay, J.K., Hausmann, N., Keurentjes, J.J., Sen, S., Stowe, K.A., Dawson, T.E., Simms, E.L. and Richards, J.H. (2005) Identification and characterization of QTL underlying whole-plant physiology in Arabidopsis thaliana: δ13C, stomatal conductance and transpiration efficiency. *Plant Cell Environ.*, **28**, 697–708.

Lobin W (1983) The occurrence of Arabidopsis thaliana in the Cape Verde Islands. *Arab Info Ser*, **20**, 119–123.

Loh, P.-R., Bhatia, G., Gusev, A., et al. (2015) Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat. Genet.*, **47**, 1385–1392.

Lovell, J.T., Juenger, T.E., Michaels, S.D., et al. (2013) Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Proc. R. Soc. B Biol. Sci.*, **280**, 20131043.

Martin, A.R., Lin, M., Granka, J.M., et al. (2017) An unexpectedly complex architecture for skin pigmentation in Africans. *Cell*, **171**, 1340–1353.

Marwede, V., Schierholt, A., Möllers, C. and Becker, H.C. (2004) Genotype× environment interactions and heritability of tocopherol contents in canola. *Crop Sci.*, **44**, 728–731.

Mayer, M., Hölker, A.C., González-Segovia, E., Bauer, E., Presterl, T., Ouzunova, M., Melchinger, A.E. and Schön, C.-C. (2020) Discovery of beneficial haplotypes for complex traits in maize landraces. *Nat. Commun.*, **11**, 4954.

McKay, J.K., Richards, J.H. and Mitchell-Olds, T. (2003) Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.*, **12**, 1137–1151.

Moore, C.R., Johnson, L.S., Kwak, I.-Y., Livny, M., Broman, K.W. and Spalding, E.P. (2013) High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics*, **195**, 1077–1086.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., et al. (2010) Plant phenotypic plasticity in a changing climate. *Trends Plant Sci.*, **15**, 684–692.

Paaby, A.B. and Rockman, M.V. (2013) The many faces of pleiotropy. *Trends Genet. TIG*, **29**, 66–73.

Przeworski, M., Coop, G. and Wall, J.D. (2005) The Signature of Positive Selection on Standing Genetic Variation. *Evolution*, **59**, 2312–2323.

Rasamivelona, A., Gravois, K.A. and Dilday, R.H. (1995) Heritability and genotype× environment interactions for straighthead in rice. *Crop Sci.*, **35**, 1365–1368.

Richard, C.A., Hickey, L.T., Fletcher, S., Jennings, R., Chenu, K. and Christopher, J.T. (2015) High-throughput phenotyping of seminal root traits in wheat. *Plant Methods*, **11**, 13.

Saini, D.K., Chopra, Y., Singh, J., Sandhu, K.S., Kumar, A., Bazzer, S. and Srivastava, P. (2022) Comprehensive evaluation of mapping complex traits in wheat using genome-wide association studies. *Mol. Breed.*, **42**, 1–52.

Scheffers, B.R., Meester, L.D., Bridge, T.C.L., et al. (2016) The broad footprint of climate change from genes to biomes to people. *Science*., **354** (6313):aaf7671.

Sivakumaran, S., Agakov, F., Theodoratou, E., et al. (2011) Abundant Pleiotropy in Human Complex Diseases and Traits. *Am. J. Hum. Genet.*, **89**, 607–618.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. and Smoller, J.W. (2013) Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.*, **14**, 483–495.

Sun, S., Wang, X., Wang, K. and Cui, X. (2020) Dissection of complex traits of tomato in the post-genome era. *Theor. Appl. Genet.*, **133**, 1763–1776.

Tisné, S., Serrand, Y., Bach, L., et al. (2013) Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J.*, **74**, 534–544.

Uga, Y., Sugimoto, K., Ogawa, S., et al. (2013) Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nat. Genet.*, **45**, 1097–1102.

Vink, J.M., Hottenga, J.J., Geus, E.J. de, Willemsen, G., Neale, M.C., Furberg, H. and Boomsma, D.I. (2014) Polygenic risk scores for smoking: predictors for alcohol and cannabis use? *Addiction*, **109**, 1141–1151.

Wallace, J.G., Larsson, S.J. and Buckler, E.S. (2014) Entering the second century of maize quantitative genetics. *Heredity*, **112**, 30–38.

Wang, Q., Tang, J., Han, B. and Huang, X. (2020) Advances in genome-wide association studies of complex traits in rice. *Theor. Appl. Genet.*, **133**, 1415–1425.

Wright, D., Rubin, C.-J., Martinez Barrio, A., Schütz, K., Kerje, S., Brändström, H., Kindmark, A., Jensen, P. and Andersson, L. (2010) The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. *Mol. Ecol.*, **19**, 5140–5156.

Xiong, J., Chen, D., Chen, Y., Wu, D. and Zhang, G. (2023) Genome-wide association mapping and transcriptomic analysis reveal key drought-responding genes in barley seedlings. *Curr. Plant Biol.*, **33**, 100277.

Yazdanbakhsh, N. and Fisahn, J. (2012) High-throughput phenotyping of root growth dynamics. *Methods Mol. Biol. Clifton NJ*, **918**, 21–40.

Yeh, P.J. and Price, T.D. (2004) Adaptive phenotypic plasticity and the successful colonization of a novel environment. *Am. Nat.*, **164**, 531–542.

Zeng, J., Xue, A., Jiang, L., et al. (2021) Widespread signatures of natural selection across human complex traits and functional genomic categories. *Nat. Commun.*, **12**, 1164.

Zhang, X., Hause, R.J. and Borevitz, J.O. (2012) Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in Arabidopsis thaliana. *G3 Bethesda Md*, **2**, 29–34.

Acknowledgments

First and foremost, I would like to dedicate this work to **my late father**, whose memory, love, and unwavering belief in my potential continue to guide and inspire me. His wisdom and guidance will always remain close to my heart. My deepest appreciation goes to **my loving mother**, whose steadfast support and encouragement have been the bedrock of my accomplishments. Your unwavering faith in my abilities has been a constant source of strength and motivation throughout my academic journey.

I would also like to express my gratitude to my supervisor, **Dr. Angela Hancock**, for giving me the opportunity to work on such an interesting research project. Words cannot express how grateful I am to work under your supervision; you have been an exceptional mentor to me. Your ability to identify research questions and devise ingenious solutions has inspired and motivated me every single day. Beyond your scientific prowess, you are a remarkable person with a genuine heart. Despite overseeing multiple projects and numerous team members, your unwavering dedication to organization and ensuring everyone's well-being is a rare and admirable quality. You were always going to the point during our discussions and trusted me for the details. I felt I had total freedom in my research, and I hope to continue to work this way in the future. Thank you for your support, input, and valuable discussions that helped me complete my thesis.

My sincere appreciation goes to **Prof. Dr. Ute Höcker** and **Prof. Dr. Christine Heim** for their time and effort in evaluating my dissertation and being part of my committee defense. I am thrilled and honored to have you as part of the committee.

I am also grateful to my TAC co-advisors, **Prof. Dr. Juliette de Meaux** and **Prof. Dr. Maria von Korff**, for their invaluable feedback and comments on my project. Their expertise and knowledge have enriched my understanding of my research topic, and their guidance has

significantly contributed to the development of my work. Thank you for always being available for me.

My heartfelt thanks go to the entire MPIPZ community for providing a friendly and stimulating scientific environment, as well as the institute for offering all the necessary facilities. The opportunity to be part of such a vibrant and supportive research community has been a privilege that I will always treasure. I would like to extend my gratitude to the PhD Coordinators, **Dr. Johanna Spandl** and **Dr. Stephan Wagner**, for their unwavering support, guidance, and encouragement throughout my PhD journey. Their dedication to ensuring a smooth and well-organized experience has been instrumental in my academic success.

I would like to extend my heartfelt appreciation to my fellow colleagues in the Hancock group: **Manu**, **Yulia**, **Sophie-Asako**, **Sofia**, **Shifa**, and **Yasin**, as well as past members: **Célia**, **Mehmet**, **Nina**, **Andrea**, **Pádraic**, and **Johan**. Their camaraderie and intellectual stimulation have been a constant source of inspiration and motivation. I am thankful to all of you for your support and for creating such a friendly group.

I am incredibly grateful to **my wife**. Her love, understanding, and unwavering faith in my abilities have been a beacon of light during the most challenging moments of my PhD. Her presence has made all the difference, and I cannot thank her enough for her support. My heartfelt thanks go to **my siblings**, who have been a constant source of encouragement and camaraderie throughout my life. Their love and support have played an integral role in my determination to persevere and reach my goals.

Lastly, I want to extend my deepest gratitude to **my friends**, who have stood by me through thick and thin. Their companionship, laughter, and shared experiences have enriched my life and made this journey all the more rewarding. Their unwavering support and belief in my abilities have been instrumental in my academic success.

Erklärung zur Dissertation

gemäß der Promotionsordnung vom 12. März 2020

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen:

Ahmed F Elfarargi, Elodie Gilbault, Nina Döring, Célia Neto, Andrea Fulgione, Andreas P M Weber, Olivier Loudet, Angela M. Hancock, *Genomic basis of adaptation to a novel precipitation regime*, *Molecular Biology and Evolution*, Volume 40, Issue3, March 2023, msad031, https://doi.org/10.1093/molbev/msad031

13th April, 2023

Ahmed Elfarargi

Curriculum Vitae

Ahmed Fathi Elfarargi

Carl-von-Linne-Weg 10, 50829 Köln, Germany

E-mail: elfarargi@mpipz.mpg.de | ahmedelfarargi@gmail.com

Phone: +49 015736942640

EDUCATION & ACADEMIC

10/2017 – 07/2023	**Ph.D.** at University of Cologne and Max Planck Institute for Plant Breeding Research (MPIPZ) in Cologne. **Project |** "Genetic basis and adaptive relevance of drought response in Cape Verde *Arabidopsis*" **Advisor |** Dr. Angela M. Hancock
09/2015 – 08/2017	**M.Sc. Pant Biotechnology** (Spec. Functional Plant Genomics) Wageningen University (WUR), Wageningen, the Netherlands **M.Sc. Thesis |** "Genome-wide identification, characterization and expression analysis of resistance (*R*) genes in *Parasponia*, *Trema*, and hybrids" **Advisors |** Dr. René Geurts and Dr. Robin Van Velzen
01/2017 – 06/2017	Research **MSc intern** at Parker lab, MPIPZ **Project |** "Toward understanding the role of SUVH-family proteins in EDS1-dependent plant immunity" **Advisors |** Prof. Dr. Jane Parker and Dr. Dmitry Lapin
10/2013 – 06/2014	**Diploma** in Horticulture Genetics and Biotechnology, CIHEAM-MAICh, Greece.
09/2006 – 07/2011	**B.Sc. in Biotechnology**, Biotechnology Department, Faculty of Agriculture, Al-Azhar University, Cairo, Egypt.

PUBLICATIONS

- **Elfarargi, Ahmed F**., Elodie Gilbault, Nina Döring, Célia Neto, Andrea Fulgione, Andreas PM Weber, Olivier Loudet, and Angela M. Hancock. "***Genomic basis of adaptation to a novel precipitation regime.***" *Molecular Biology and Evolution* 40, no. **3** (2023): msad031.

- Tergemina, Emmanuel, **Ahmed F. Elfarargi**, Paulina Flis, Andrea Fulgione, Mehmet Göktay, Célia Neto, Marleen Scholle et al. "***A two-step adaptive walk rewires nutrient transport in a challenging edaphic environment.***" *Science advances* 8, no. **20** (2022): eabm9385.

- Fulgione, Andrea, Célia Neto, **Ahmed F. Elfarargi**, Emmanuel Tergemina, Shifa Ansari, Mehmet Göktay, Herculano Dinis et al. "***Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages.***" *Nature Communications* 13, no. **1** (2022): 1461.

CONFERENCES

- **ESEB Conference**, Prague Congress Centre, Prague, Czech Republic, August 14th – 19th, 2022. (Participated with a talk)

- **GRC meeting (Salt and Water Stress in Plants)**, Les Diablerets Conference Center in Les Diablerets, Vaud (fr), Switzerland, May 22nd – 27th, 2022. (Participated with a poster)

- **Society for Molecular Biology and Evolution (SMBE)** 2021, Online Meeting, July 3rd – 8th, 2021. (Participated with a poster)

- **Population Genetics Group (PopGroup54)**, Online Meeting, January 4th – 6th 2021.

- **ESEB Conference**, Turku, Finland, August 19th – 24th, 2019. (Participated with a poster)

- **International Plant Molecular Biology (IPMB) Conference**, Montpellier, France, August 5th – 10th, 2018.

WORKSHOPS & TRAINING COURSES

- **MaxBi Basics in Python**, online course provided by MPI-GWDG, Göttingen, Germany, April 18th – 20th, 2021.

- **24th Summer Institute in Statistical Genetics (SISG)**, University of Washington, Seattle, Washington, United States, July 8th – 26th, 2019. Participated in four training modules on Population and Quantitative Genetics (two awarded as a scholarship).

- **Data Analysis with R**, MPIPZ, Cologne, Germany, April 24th – 26th, 2019.

- **Statistical Literacy**, MPIPZ, Cologne, Germany, October 8th – 10th, 2018.

- **RNA Sequencing**, provided by CEPLAS, Düsseldorf, Germany, September 17th – 19th, 2018.

- **Cytoscape Automation**, Max Planck Institute for Biology of Ageing, Cologne, Germany, November 30th, 2017.

- **Protein-Protein Interaction and Introduction to Cytoscape**, Max Planck Institute for Biology of Ageing, Cologne, Germany, November 27th – 28th, 2017.

- **Self- and Time Management**, MPIPZ, Cologne, Germany, November 20th – 21st, 2017.

AWARDS & SCHOLARSHIPS

- **SISG2019 Award**, July 2019

  Awarded two modules to attend the 24th Summer Institute in Statistical Genetics at the University of Washington.

- **NUFFIC-NFP Scholarship**, September 2015 – August 2017

  Full scholarship awarded for pursuing a Master's Degree in Plant Biotechnology at Wageningen University, the Netherlands.

- **CIHEAM-MAICh Scholarship**, October 2013 – June 2014

  Scholarship for completing a Diploma in Horticulture Genetics and Biotechnology at CIHEAM-MAICh, Greece.