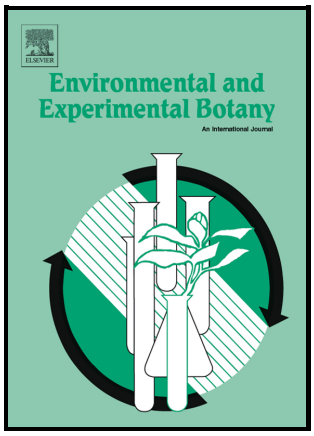


Heterologous expression of coffee *HB12* confers tolerance to water deficit in transgenic plants through an ABA-independent route

Fernanda P Cruz, Roberta KTM Loh, Mariana LC Arcuri, Carlos Dezar, Luis WP Arge, Thais Falcão, Elisson Romanel, Carolina V Morgante, João VA Cerqueira, Thuanne P Ribeiro, Stefanie M Moura, Adriana B Arongaus, Ighor LG Arantes, Bruna P Matta, Regis L Correa, Eduardo Romano, Maria F Grossi-de-Sa, Dorothea Bartels, Raquel L Chan, Márcio Alves-Ferreira

PII: S0098-8472(24)00341-1

DOI: https://doi.org/10.1016/j.envexpbot.2024.105983

Reference: EEB105983

To appear in: *Environmental and Experimental Botany*

Received date: 21 April 2024
Revised date: 28 August 2024

Accepted date: 19 September 2024

Please cite this article as: Fernanda P Cruz, Roberta KTM Loh, Mariana LC Arcuri, Carlos Dezar, Luis WP Arge, Thais Falcão, Elisson Romanel, Carolina V Morgante, João VA Cerqueira, Thuanne P Ribeiro, Stefanie M Moura, Adriana B Arongaus, Ighor LG Arantes, Bruna P Matta, Regis L Correa, Eduardo Romano, Maria F Grossi-de-Sa, Dorothea Bartels, Raquel L Chan and Márcio Alves-Ferreira, Heterologous expression of coffee *HB12* confers tolerance to water deficit in transgenic plants through an ABA-independent route, *Environmental and Experimental Botany,* (2024) doi:https://doi.org/10.1016/j.envexpbot.2024.105983

Heterologous expression of coffee *HB12* confers tolerance to water deficit in transgenic plants through an ABA-independent route

Fernanda P Cruz[1]*; Roberta KTM Loh[1]*; Mariana LC Arcuri[1]; Carlos Dezar[2]; Luis WP Arge[1]; Thais Falcão[3]; Elisson Romanel[3]; Carolina V Morgante[5,6]; João VA Cerqueira[5,6]; Thuanne P Ribeiro[5,6]; Stefanie M Moura[5,6,7]; Adriana B Arongaus[1]; Ighor LG Arantes[8]; Bruna P Matta[4]; Regis L Correa[1]; Eduardo Romano[5,6]; Maria F Grossi-de-Sa[5,6,9]; Dorothea Bartels[10]; Raquel L Chan[2]; Márcio Alves-Ferreira[1,6§].

[1]. Laboratório de Genética Molecular e Biotecnologia Vegetal, Universidade Federal do Rio de Janeiro.

[2]. Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral, CONICET.

[3]. Laboratório de Genômica de Plantas e Bioenergia (PGEMBL), Departamento de Biotecnologia, Escola de Engenharia de Lorena (EEL), Universidade de São Paulo (USP), Lorena, Brazil.

[4]. Laboratório de Evolução de Caracteres Complexos, Universidade Federal do Rio de Janeiro.

[5]. Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, Brazil

[6]. National Institute of Science and Technology, INCT PlantStress Biotech, Embrapa-Brazil

[7]. Embrapa Semiárido, Petrolina-PE, Brazil

[8]. Laboratório do Arbovírus, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, RJ, Brazil.

[9]. Universidade Católica de Brasília, Brasília-DF, Brazil

[10]. Institut für Molekulare Physiologie und Biotechnologie der Pflanzen, Rheinische Friedrich-Wilhelms Universität Bonn.

***** These authors contributed equally to this work.

§ Corresponding author.

E-mail addresses:

FPC: fernandawaltenberg@gmail.com

RKTML: robertakuantchuen@gmail.com

MLCA: mariana.arcuri@unesp.br

CD: carlosdezar@gmail.com

LWPA: l.willianpacheco@gmail.com

TF: thaisfalcao@usp.br

ER: romanel@usp.br

CVM: carolina.morgante@embrapa.br

JVAC: joaovictorufc@gmail.com

TPR: thuanne.florestal@gmail.com

SMM: stefanieshalom@gmail.com

ABA: adriana.arongaus@gmail.com

ILGA: ighorarantes@gmail.com

BPM: bpalma@biologia.ufrj.br

RLCo: regis@biologia.ufrj.br

ER: romano@embrapa.br

MFGS: fatima.grossi@embrapa.br

DB: dbartels@uni-bonn.de

RLC: rchan@fbcb.unl.edu.ar

MAF: marcioaf@ufrj.br

Abstract

Drought is one of the major abiotic stresses affecting plant growth, with serious negative consequences for crop yields worldwide. Among these crops, coffee is severely injured by water deficiency. Despite its economic importance, very little is known about the molecular mechanisms governing coffee responses to water deficit. In the present work, a total of 288 members of the homeobox (HB) gene family were identified in the genome of the *Coffea Arabica* Brazilian Coffee Genome Project database. *In silico* analysis allowed to determine the expression pattern of 33 HD genes. Among them, three genes (*CaZHD4*, *CaHB1-like2* and *CaHB12*) were found to be up-regulated by osmotic stress in the database. Expression analyses revealed that *CaHB12* is highly up-regulated in the leaves and lateral roots of *Coffea arabica* plants under moderate and severe water deficit conditions even after 10 days of drought induction. Functional characterization of transgenic Arabidopsis plants constitutively expressing *CaHB12* resulted in increased tolerance to water deficit at different developmental stages and increased tolerance to salt stress during seed germination. To gain further insights into genes modulated by the ectopic expression of *CaHB12,* a RNA-Seq was performed revealing that classical drought-responsive genes were mostly repressed, suggesting that other mechanisms likely contribute to the tolerant phenotype exhibited by *CaHB12*-expressing plants, such as the pathway signalled by heat shock proteins, reactive oxygen species and heat shock transcription factor signalling pathways. Moreover, to provide further support for the involvement of *CaHB12* in drought stress tolerance, three independent soybean transgenic lines overexpressing *CaHB12* were employed in this study. Accordingly, at a physiological level, the constitutive expression of CaHB12 promotes the regulation of stomatal conductance and antioxidant activity under drought conditions, suggesting that this gene plays a key role in plant responses to water deprivation and can confer tolerance to drought stress. Our data suggest that *CaHB12* is a positive regulator of the stress response in coffee plants and indicate that this gene is a potential candidate for biotechnological approaches.

Keywords:

drought tolerance, osmotic stress, stomatal conductance, HD-Zip, ectopic expression, heat shock proteins

1. Introduction

Brazil is the largest coffee producer worldwide, and coffee ranks among the top 20 most valuable commodities in Brazil (FAOSTAT, 2023). Water deficit caused by drought is considered one of the biggest problems affecting growth efficiency in coffee cultures. The climate change scenarios predict a drastic decrease (50%) in the global area suitable for coffee production by 2050 (Koutouleas et al., 2022). Although a period of moderate water deficit is considered important for flower induction in coffee plants, prolonged limited water conditions can lead to early leaf senescence and reductions in leaf area and grain size (Carr, 2001). Moreover, the fruit development of *Coffea arabica* seems to be quite sensitive to water deficit (Damatta and Ramalho, 2006). The Brazilian Coffee Genome Project initiated in 2002 generated useful tools for the study of coffee genes (Vieira et al., 2006). A total of 37 EST libraries were sequenced, containing samples of three coffee species (*Coffea arabica*, *Coffea canephora* and *Coffea racemosa*) and covering an extended variety of tissues, developmental stages and environmental conditions, including abiotic stresses (Vieira et al., 2006).

Few studies have shed light on coffee molecular response mechanisms to water deficit (Alves et al., 2017; Phillips et al., 2017; Santos et al., 2021; Fernandes et al., 2021; Marques et al., 2022). Expression analysis of carotenoid biosynthetic pathway genes suggested that activation of the xanthophyll cycle in young leaves of coffee trees may be involved in the adaptation process to drought (Simkin et al., 2008). Fernandes *et al.* (2021), after transcriptomic analysis of *C. arabica* cv. Icatu (Icatu) and *C. canephora* cv. Conilon clone CL153 (CL153), revealed a predominance of protective genes (more in CL153) associated with antioxidant activities, embracing including genes involved in water deficit and desiccation,

such as LEA and aspartic proteases (Fernandes et al., 2021). The transcriptional up-regulation elicited by a severe water deficit was proposed to trigger a genotype-specific response involving reticuline oxidase genes in CL153 and heat shock proteins in *C. canephora* (cv. CL53) and *C. arabica* (cv. Icatu), respectively (Fernandes et al., 2021). Both coffee species, *C. arabica* and *C. canephora*, exhibit different mechanisms to induce drought tolerance, and *C. canephora* has been suggested as being the most tolerant to prolonged and severe water deprivation, and this ability reflects directly on their ROS and reactive nitrogen species (RNS) content and the antioxidative apparatus recruited during acclimation (Damatta et al., 2018; Santos et al., 2021; Marques et al., 2022). Another strategy preventing oxidative damage to the photosynthetic machinery is apparently the accumulation of the RuBisCO (RBSC1) protein, as observed in leaves of *C. canephora* var. Conilon subjected to water stress (Marraccini et al., 2011). Although influenced by different factors, some families of transcription factors are considered closely related to responses to drought. A total of 83 transcription factors were found to be involved in the plant response to drought in *Coffea spp.* Most of them were up-regulated (Fernandes et al., 2021). Similarly, a study aimed at the identification of candidate genes in *C. canephora* revealed the expression patterns of two HD-Zip (*CcHDZ1* and *CcHDZ2*) genes in leaves of drought-tolerant (14 and 120) and drought-sensitive (22) cultivars during water deficit acclimation (Marraccini et al., 2012). *CcHDZ1* and *CcHDZ2* are induced under water stress, indicating a possible role for HD-Zip genes in molecular responses to water deficit in coffee. The transcriptional basis of stress memory induced by drought in *C. canephora* was investigated by Guedes and collaborators. Among the transcription factors identified as drought stress memory genes, two HD-Zip, putative homologues of *AtHB7*, were found, genes Cc08_g16780 and Cc02_g01010 (Guedes et al., 2018).

The HD-Zip superfamily of transcription factors (TFs) is unique to plants and is characterized by the presence of a homeobox (HD) associated with a juxtaposed leucine zipper motif (LZ) (Ruberti et al., 1991; Schena and Davis et al., 1992). At present, four subfamilies of HD-Zip proteins can be distinguished (HD-

Zip I, II, III and IV) (Mukherjee et al., 2009). The HD-Zip I subfamily is further subdivided into nine clades (α, β1, β2, γ, δ, ε, ζ, φ1 and φ2) based on sequence similarity and the intron–exon structure of the encoding genes (Agalou et al., 2008; Henriksson et al., 2005). These transcription factors were resolved in a new phylogenetic tree according to the architecture of their carboxy and amino termini (Arce et al., 2011). Transcriptomic analysis revealed that HD-Zip genes are strongly responsive to drought stress in *Coffea spp.* (Fernandes et al., 2021, Guedes et al., 2018). The *Arabidopsis* genes *AtHB7* and *AtHB12* are the best-characterized members within the γ-clade. Strong up-regulation of these two genes was observed during osmotic stress conditions and exogenous treatment with ABA (Olsson et al., 2004). *AtHB12* is preferentially expressed in leaves and stems, whereas *AtHB7* in roots at early developmental stages (Re et al., 2014). During the vegetative stage, *AtHB12* is involved in determining cell size, and at reproductive stage as a negative regulator of inflorescence emission by acting on the pathway mediated by gibberellin (Lee and Chan, 1998; Hur et al., 2015; Son et al., 2010). Another property of *AtHB12* described in the literature is its inducibility by various abiotic stresses (Li et al., 2022). *HaHB4* is a well-characterized divergent sunflower γ-clade member. *Arabidopsis* plants expressing *HaHB4* exhibit a water stress-tolerant phenotype (Dezar et al., 2005). Microarray analysis suggested that *HaHB4* down-regulates genes related to ethylene biosynthesis and ethylene signalling pathways, generating a senescence delay (Manavella et al., 2006). Similarly, the overexpression of *Coffea arabica HB12* (*CaHB12*), the gene characterized in this work, in cotton confers a significant increase in plant tolerance to water deficit, corroborating its potential for plant biotechnology (Basso et al., 2021). The first technology of genetic engineering lines adapted for drought tolerance was a transgenic soybean expressing *HaHB4*, called HB4 and developed by Verdeca, which displays an improved performance of tolerance to drought (Bergau, 2019; Ribichich et al., 2020). Currently, this same technology is present in transgenic wheat, which exhibits similar nutritional composition and an enhanced performance under water deficit conditions (Ayala et al., 2019; Gonzalez et al.,

2019). Several studies have identified and characterized homeobox genes in a wide range of plant species, including nonconventional species, such as sesame, bamboo, poplar, eucalyptus, and physic nut (Wei, 2019; Xu et al., 2019; Hou et al., 2021; Zhang et al., 2020, Tang et al., 2019). The involvement of HD-Zip encoding genes in soybean responses to drought and salinity has been well described, with *Gmdhz32* and *Gmhdz72*, *AtHB12* homologous, emerging as crucial genes at both vegetative and also reproductive developmental stages (Chen et al., 2014; Bhattacharjee et al., 2015).

In this study, we identified and classified 34 coffee HB genes retrieved from the Brazilian Coffee Genome Project database (Vieira et al., 2006). *In silico* analyses allowed the identification of three coffee HB genes, which were predominantly expressed in water deficit and salt stress libraries. Expression profiles of these genes were validated in the leaves and lateral roots of coffee plants during water deficit conditions. *CaHB12*, an HD-Zip I γ-clade, was quickly induced, especially in lateral roots, suggesting that this TF might play a role in the primary coffee water deficit response. We obtained transgenic Arabidopsis and soybean plants expressing *CaHB12* exhibiting increased tolerance to water deprivation compared to their respective controls. RNA-Seq analyses indicated that the mechanism by which *CaHB12* confers drought tolerance to transgenic plants is not related to conventional ABA- dependent routes but to other pathways such as heat shock responses.

2. Results

2.1 Identification, annotation, and phylogenetic analysis of coffee homeobox genes implicated in the water deficit response

To understand the evolutionary relationship of the homeobox gene family among coffee and related species, we used all genome data available from three *Coffea* species (Asterids), including an allotetraploid *Coffea arabica* and its parental diploid *C. canephora* and *C. eugenioides*, tree core Rosids, such as

Arabidopsis thaliana, *Gossypium hirsutum* (cotton), and *Glycine max* (soybean), and two monocots, such as *Oryza sativa* (rice) and *Setaria viridis* (Figure 1). The identification and annotation of the homeobox genes showed 288 in *C. arabica*, 132 in *C. eugenioides*, 81 in *C. canephora*, 111 in *A. thaliana*, 312 in *G. hirsutum*, 281 in *G. max*, 109 in *O. sativa*, and 100 in *S. viridis* (Supplementary Table 1). Arabidopsis and soybean were the only plant species studied here containing members from a set of 14 distinct classes already identified in the plant genome. (Bhattacharjee et al., 2015; Mukherjee et al., 2009). To improve communication concerning gene nomenclature, we adopted the Arabidopsis closest related homeobox gene name to classify coffee homeobox genes, as in the case of *CaHB12* (Ca_66_88.11.1) (Figure 1C).

To gain a better understanding of the evolutionary relationships of HB genes among plant species, 1,414 HB genes were used to construct a phylogenetic tree based on the HD domain. The maximum likelihood phylogenetic analysis among three coffee species and five other angiosperm species revealed that at least one member of each of the 14 classes was identified in coffee-related species (Figure 1A). However, we did not find any member for PINTOX in either *C. eugenioides* or *C. arabica* species or for PHD, NDX, or LD in either *C. canephora* or *C. arabica*, indicating a putative loss of those members in coffee species or an incomplete genome sequence available from the first draft genome. To achieve a better resolution of the coffee HB gene classification, a separate phylogeny was constructed for each family (Supplementary Figures 1, 2, 3, and 4, Supplementary Table 2), and the HD-Zip I proteins were subdivided into nine clades (α, β1, β2, γ, δ, ε, ζ, φ1 and φ2) (Figure 1B), according to the nomenclature of Agalou *et al*. (Agalou et al., 2008). Comparing diploid *Coffea* species, *C. eugenioides* contains many more members for each homeobox class than *C. canephora*, such as for superclasses HD-Zip, BEL, and DDT, which correlates with the coding sequences available thus far (Supplementary Table 1). The tetraploid *C. arabica* also contains many more members (HD- Zip I, HD- Zip II, HD- Zip III, PLINC, WOX, KNOX, and DDT) than the sum of both diploid *Coffea* species, which also correlated with

the total coding sequences available for those species. Exceptions to this rule involve both HD- Zip IV and BEL classes. In cotton, half of the classes (HD- Zip I, HD- Zip II, HD- Zip III, KNOX, BEL, DDT, and PHD) contain up to 3 to 5x genes related to Arabidopsis. Interestingly, *G. hirsutum* species did not contain members of the PINTOX and SAWADEE classes. On the other hand, rice and *S. vidiris* have a similar number of members for each class, showing a highly conservative evolution of the homeobox gene family in monocots.

2.2 Exploring the Diversity and Expression Patterns of Homeobox (HB) Genes in Coffee

A total of 56 coffee ESTs (36 contigs and 20 singletons) were recovered from the Brazilian Coffee Genome Project database (BCGP), and the presence of HD was confirmed in 20 contigs and 13 singletons (33 genes) (Supplemental Table 1). The scarcity of expressed HD genes could be attributed to the restricted dataset of the BCGP, which was the sole data source available at the time of the study. Other domains typically associated with HD were also confirmed in some coffee sequences. The protein sequence alignment of the HD regions identified in coffee allowed the construction of a well-supported phylogenetic tree (data not shown), presenting high bootstrap values among the distinct HB families. The phylogenetic analysis revealed that HB families are present in coffee ESTs, except for DDT, PHD, LD, PINTOX, SAWADEE, and NDX (Supplemental Table 2). Among the recovered coffee sequences, two were classified as WOX, seven as KNOX, six as BEL, five as PLINC, seven as HD- Zip I, three as HD- Zip II, two as HD- Zip III, and one as HD- Zip IV. Interestingly, most of the EST sequences were related to *C. arabica*, while only two ESTs (Contig7385 [HD-Zip I] and CA00_XX_CB1_036_E08_MC_F [PLINC]) were similar to *C. canephora* sequences (Supplementary Table 1).

After the classification of coffee HB sequences, a digital expression analysis was performed, exploiting the information available in the Brazilian Coffee Genome Project database (Figure 2) (Vidal et al., 2010; Vieira et al., 2006). Coffee HB genes showed a range of expression profiles, which were induced in a

wide variety of tissues and experimental conditions. Multiple expression profiles were observed for most of the contigs, except for *CaHAT14* (HD- Zip II - Contig10070), which presented reads exclusively in the embryogenic callus library (EA1), and *CaKNAT3* (KNOX - Contig8161), which presented reads exclusively in seedlings grown with the arachidonic acid library (AR1). For the selection of the target genes, we focused on the analysis of genes expressed in the water deficit (SH) or salt stress (CS1) libraries. Three genes, *CaZHD4*, *CaHB1-like2*, and *CaHB12,* were chosen for further analysis based on their distinctive expression patterns in SH and CS1 according to *in silico* data.

2.2 Expression analysis reveals that *CaHB12* is modulated by water deficit conditions

RT-qPCR analyses were conducted for HD gene expression in leaves and secondary root samples of the six-month-old *C. arabica* cultivar "Catuaí Vermelho" submitted to water deficit (Figure 3). Samples were collected 2, 5 and 10 days after watering was stopped, which corresponds to average water potentials (Ψ_w) of -0.27 MPa, -1.05 MPa, and -4.45 MPa, respectively. *CaZHD4*, *CaHB1-like2*, and *CaHB12* expressions were up-regulated by water deficit. Similar results were observed in the *C. arabica* cultivar "Bourbon Amarelo" (data not shown). The early response to water deficit was stronger in roots than in leaves for the three genes (Figure 3A and B). The *CaHB12* gene showed the most remarkable expression pattern and the strongest induction. Low expression of *CaHB12* was observed in nonstressed leaves and lateral roots, followed by a striking induction after 2 days of stress treatment, when the Ψ_w of stressed plants was slightly different from that of control plants. A marked difference was observed in the expression patterns of *CaHB12* in stressed leaves and lateral roots. While gradual and continuous expression of *CaHB12* was observed during the stress treatment in leaves (Figure 3A), expression peaked during the early stages of water deficit in lateral roots, which was maintained at high levels during the remaining treatment (Figure 3B). Even under severe stress conditions after 10 days of stress treatment

(average of Ψ_w = -4.45 MPa), the expression of *CaHB12* remained high, both in drought-stressed leaves and lateral roots.

2.4 Characterization of *CaHB12* and subcellular localization

The full-length *CaHB12* cDNA sequence is composed of sequences from two singletons, CA00_XX_FR1_002_G03_MC_F (containing the HD and LZ domains) and CA00_XX_SH1_047_H05_EM_F (containing the C-terminal region), of the Brazilian Coffee Genome Project database. The *CaHB12* CDS is 693 bp in length and encodes a CaHB12 protein of 236 amino acids (aa). Analysis of the predicted protein sequence revealed that the HD was located between aa 24 and 94, and the homeodomain-associated leucine zipper was located between aa 89 and 133 (Supplementary Figure 5). The CaHB12 C-terminal domain (CTD) comprises 102 aa after LZ and is similar in size to AtHB12 (106 aa) and MtHB1 (108 aa), in contrast to AtHB7 (127 aa) and HaHB4 (64 aa) (Supplementary Figure 6A). A moderate identity (51% to 59.33%) was observed in the alignment of the HD-LZ domain from CaHB12 and the abiotic stress-related HD-Zip I proteins AtHB7, AtHB12 (*A. thaliana*), MtHB1 (*Medicago truncatula*) and HaHB4 (*Helianthus annus*) (Supplementary Figure 6B). Much lower identity values (11 to 20%) were noticed in the alignment of the CTD region (Supplementary Figure 6C). The C-terminal region (CTR) of CaHB12, which encompasses the final amino acid residues of the protein, contains two Trp residues in conserved positions and is made up of three aromatic, two acidic and five hydrophobic residues juxtaposed to a neutral residue (Ser) described as being an AHA-motif (Arce et al., 2011), which is also longer (nine aa) than the CTRs of AtHB7, AtHB12 and MtHB1 (six aa) and HaHB4 (four aa) (Supplementary Figure 6A).

Transient expression analyses were carried out to confirm the nuclear localization of CaHB12. Coffee leaves were bombarded with a construct carrying the complete *CaHB12* CDS fused to *GFP* driven by the 35S promoter (*CaMV35S:CaHB12:GFP*). The GFP signal was detected in the nuclei (Figure 4C and F), while in cells transformed with the control vector (*CaMV35S:GFP*), the

GFP signal was also located in the cytoplasm (Figure 4B and E). Sequence analysis using PSORT software (Nakai et al., 1999) identified a monopartite nuclear localization signal (NLS) of the KRKR type between aa 24 and 29. To test whether this would be a functional NLS, a construct was prepared that lacked the first 30 aa (*CaMV35S:CaHB12tru:GFP*). Cells transformed with this construct did not show a GFP signal in the nucleus, confirming that the information for nuclear localization is contained within the first 30 aa (Figure 4D and G).

2.5 *A. thaliana* transgenic lines expressing *CaHB12* (CaHB12-OE) are more tolerant to water deficit

Since routine coffee transformation protocols are not available (Mishra and Slater, 2012; Mishra, 2022), a heterologous expression system was chosen to investigate whether the expression of *CaHB12* confers drought tolerance to transgenic plants. Three independent lines of *A. thaliana*, here denominated as CaHB12:OE-A, -B and -C containing a single insert of *CaHB12*, were obtained and selected for analysis. The expression of *CaHB12* in these plants was confirmed through RT-PCR and homozygous T3 generation seeds and plants, which were used in drought and salt tolerance experiments (Figure 5A). Transgenic plants constitutively expressing *CaHB12* apparently showed the same morphology as wild-type plants under physiological conditions. Water deficit tolerance was evaluated during three different developmental stages using three different approaches. In the first approach, malondialdehyde (MDA) values were used to monitor the stress levels of the seedlings. Plants belonging to the transgenic lines expressing *CaHB12* exhibited less membrane lipid peroxidation than wild-type plants when subjected to water deficit conditions in PEG-treated plates (-1.2 MPa) (Figure 5B). The transgenic line C showed the lowest level of MDA (~37 nmol/g), followed by lines B and A, which presented averages of 40 and 41 nmol/g, respectively. In the second approach, transgenic and wild-type plants in transition from vegetative to reproductive growth were subjected to a rapid and abrupt stress, in which watering was withheld until permanent wilting symptoms were observed in wild-type plants. The plants were then rehydrated, and survival rates were

measured after 48h. Transgenic lines showed the same trend during assays, with an average survival rate approximately 29% higher than that of wild-type plants (Figure 5C). The third approach was carried out with plants in the transition from the reproductive to the fruit set phase. After 30 days of regular watering, plants were subjected to slow and continuous dehydration and irrigated with a small quantity of water until obvious wilting symptoms were permanent in wild-type plants. As observed in the case of rapid and abrupt stress, 48 h after rehydration, transgenic plants showed survival rates approximately 87% higher than wild-type plants (Figure 5D).

All transgenic lines displayed a significant increase in germination rates when sown on growth media supplemented with 150 mM NaCl when compared to the wild-type (Supplementary Figure 7A). Moreover, MDA values of transgenic plants under saline stress (100 mM NaCl) revealed lower levels of lipid peroxidation than wild-type plants (Supplementary Figure 7B). The better performance of transgenic lines is shown in Supplementary Figure 7C. Because ABA biosynthesis is stimulated by high salt, stress tolerance was also evaluated using 10-day-old seedlings. No significant differences in salt tolerance between transgenic and wild-type plants were observed when seedlings were transferred from MS medium to the equivalent supplemented with 150 mM NaCl after 7 days (data not shown). These results suggest that the differential performance is limited to the germination stage. Experiments using distinct ABA concentrations (1 and 5 µM) were performed to investigate whether CaHB12-OE lines were ABA insensitive. All transgenic lines were more sensitive than the wild-type to ABA at 1 µM and presented comparable responses at 5 µM (Supplementary Figure 7D,E).

2.6 Global expression analysis of CaHB12-OE plants uncovers the mechanism of tolerance

To better understand the function of *CaHB12* and the downstream genes regulated by its expression, transcriptome analyses were carried out by RNA-Seq with 10-day-old seedlings of the transgenic line CaHB12-OE-B and wild-type lines

grown under normal growth conditions. A total of 96,620,876 cleaned reads were obtained in the four sequenced libraries (two transgenic and two wild-type) (Supplementary Table 3). After mapping reads against the reference genome, 18,598 genes were further used for gene expression analyses. *CaHB12* was detected only in the transgenic libraries. A total of 203 genes were differentially expressed (FDR<0.05) in the transgenic plants compared to the wild-type genotype, with 124 genes induced and 79 genes repressed due to the presence of the transgene (Figure 7, Supplementary Table 4). The expression patterns of nine out of ten genes were confirmed by RT-qPCR (Supplementary Figure 8).

To explore the different annotation classes of genes deregulated by the ectopic expression of *CaHB12* in *A. thaliana*, we identified overrepresented crossed information about transcription factors (TF), GO terms and metabolic pathways. The enrichment analysis of the GO terms identified several terms related to abiotic stress up-regulated in plants overexpressing *CaHB12*, such as the response to heat, response to high light intensity, and responses to reactive oxygen species and hydrogen peroxide. Interestingly, the terms "response to abscisic acid" and "response to salt stress" were enriched, but most of the genes were down-regulated in plants overexpressing *CaHB12* (Figure 7A). The analysis showed that the terms "response to abscisic acid" and "response to salt stress" share several down-regulated genes, such as *COR15A* (AT2G42450), *RD22* (AT5G25610), *ADH* (AT1G77120), and *pEARL1* (AT4G12480), among others (Figure 7B). In addition, several up-regulated genes were also shared among the terms response to hydrogen peroxide, reactive oxygen species and heat, such as *ATHSP17.4* (AT3G46230), *ATHSP17.6A* (AT1G59860) and *HSP17.6II* (AT5G12020) (Figure 7B - Supplementary Table 5). Regarding metabolic pathways, we identified 12 enriched KEGG pathways in the DEG set, being nine identified in down- and 5 in up-regulated genes (Figure 7C and Supplementary Table 6), in which MAPK signalling pathway (map04016) and protein processing in endoplasmic reticulum (map04141) were the pathways with the most significant p value.

Regarding hormone responses affected by *CaHB12* expression, ABA-responsive genes related to seed germination were induced (AT1G75750-*GASA1*, AT3G02885-*GASA5*), while ABA-responsive genes related to drought were repressed (AT5G25610-*RD22*, AT5G59310-*LTP4*) in plants overexpressing *CaHB12*. The transcription factor *WRKY70* (AT3G56400), an activator of jasmonic acid (JA) genes, was also found to be repressed. Regarding ethylene processes, an ERF transcription factor family and *TEM1* (AT1G25560), a known ethylene repressor, were found to be repressed. Taken together, these results suggest crosstalk between JA and ethylene stimuli (e.g., AT4G17490-*ERF6* and AT4G23600-*CORI3*, important for ethylene precursor formation). Other genes, previously, were also repressed (AT3G08770-*LTP6*, AT2G41240-*BHLH100*, and AT5G25610-*RD22*) (Supplementary Table 4).

Additionally, we retrieved from scientific literature genes reported as memory and non-memory (Ding et al., 2012), other ABA related genes (Thonglim et al., 2023) and transcription factors differentially expressed (Supplementary Figure 9 and table 7). Memory genes presented no differential expression and among non-memory genes, just the *COR15A* (AT2G42540) was detected as differentially expressed. Also, we analysed the expression of other ABA related genes and no differential expression was detected. In the set of 10 transcriptional factors detected as differentially expressed, 7 showed down-regulation and only *HD-like* (AT4G03250), *AGL24* (AT4G24540) and *HSFA2* (AT2G26150) were up-regulated. Interestingly, *HD-like* belongs to the same family of the gene studied here and *HSFA2* is a heat shock factor previously reported to drive transcriptional memory after heat stress conditions in Arabidopsis (Friedrich et al., 2021).

Systems biology representation of protein–protein interaction (PPI) of the genes modulated by the overexpression of *CaHB12* is composed of 130 nodes and 324 interactions (Figure 8), a network in which three modules were detected. Module 1 is enriched with genes related to defence responses, oxidative stress and chitin biological processes. Module 2 is enriched with triterpenoid biosynthesis,

root development and oxidation–reduction processes. Finally, module 3 was associated with response to abiotic stresses, reactive oxygen species, hydrogen peroxide, protein complex oligomerization and cellular response to unfolded proteins (Supplementary Table 8).

2.7 Water deficit-responsive microRNAs

Given that microRNAs (miRNAs) are a class of regulatory noncoding RNAs and have been associated with different types of stresses, including ABA-dependent and ABA-independent pathways (Khraiwesh et al., 2012; Xia et al., 2020), the small RNA (sRNA) fraction of *35S:CaHB12* transgenic plants was also analysed in comparison to wild-type plants. A total of 55,405,921 reads were sequenced in the two wild-types and transgenic plants, of which 2,883,452 mapped to 65 known Arabidopsis miRNA genes belonging to 52 families (Supplementary Table 9). Some miRNA isoforms were also detected for several genes, but usually with a lower number of reads than the major reference variant (Supplementary Table 9). The targets of all identified sequences were predicted, and as expected, miRNAs belonging to the same family had highly similar predicted targets (Supplementary Table 10).

A total of 12 miRNA families were significantly differentially expressed between wild-type and transgenic plants (Figure 6C, Supplementary Table 11). miRNA genes known to be involved in several aspects of plant development were directly or indirectly induced by *CaHB12* expression. Such miRNA genes participate in processes, such as vegetative phase change (miR156 and miR159) (Alonso-Peral et al., 2010; Wu et al., 2009), leaf polarity identity (miR166) (Rhoades et al., 2002), regulation of cell proliferation (miR396) (Liu et al., 2009), establishment of stem cell competence (miR394) (Knauer et al., 2013), molecular signalling responses to phytohormones (miR167 and miR171) (Llave et al., 2002; Yang et al., 2006) and homeostasis of the RNA silencing machinery (miR162) (Xie et al., 2003). In addition, all aforementioned miRNAs have been associated with drought, salt or other types of abiotic stresses in *A. thaliana* and other plant species

(Ding et al., 2013; Gentile et al., 2015; Xia et al., 2020). In turn, miR827 and miR395 have been associated with oxidative stresses (Kant et al., 2011; Liang et al., 2010) and were induced and repressed in transgenic plants, respectively. One biotic stress-related miRNA (miR773) (Li et al., 2010) and one recently evolved gene in the Arabidopsis lineage (miR822) (Allen et al., 2004) were also induced in response to *CaHB12* overexpression.

2.8 CaHB12-OE soybean plants limit the decrease in stomatal conductance and CO_2 assimilation under water deprivation

To gain further insights into the functionality of *CaHB12* and substantiate its use as a tool for achieving water deficit-tolerance in crops, we transformed soybean (*Glycine max*) plants with the construct *35:CaHB12* and characterised the overexpressing lines.

In this study, 1,200 soybean embryos underwent bombardment and subsequent in vitro regeneration into plantlets using the Imazapyr herbicide for transformant selection, resulting in the acclimatization of 235 plants in a greenhouse. The confirmation of transgene insertion and expression was conducted through conventional PCR and ELISAs. Ten independent T0 events were obtained, with a transformation efficiency of approximately 0.8%. Among these, three independent events (CaHB12:OE-1, -2 and -3) were selected for further generations (Supplementary Figure 10). Gas exchange parameters were analysed under well-watered and drought conditions for CaHB12-OE T3 lines 1, 2, and 3, as well as non-transformed (NT) plants. Under normal conditions, no significant differences were observed, but under water restriction, CaHB12-OE lines exhibited a minor reduction in net CO_2 assimilation rate (AN) and slightly higher stomatal conductance (gs) compared to NT plants (Figure 9A-C). After rehydration, all genotypes fully recovered gs and transpiration (E), with CaHB12-OE lines maintaining a higher AN compared to NT plants (Figure 10D). AN/Ci curves demonstrated the adjustment of photosynthesis during stress treatment, revealing

that photosynthesis in CaHB12-OE lines was less limited than in NT plants under water restriction (Figure 9E-H). Interestingly, photosynthetic parameters, such as Vcmax, Jmax, and TPU, were less negatively affected by water deficit in CaHB12-OE-2. Additionally, the maximum quantum efficiency of PSII (Fv/Fm) remained unaffected, while the actual quantum yield of photosystem II (φPSII) showed variations across genotypes and treatments (Supplementary Figure 11A-C). The study suggests that the overexpression of CaHB12 enhances drought tolerance in soybean, impacting photosynthetic performance under stress conditions.

2.9 CaHB12-OE soybean plants increase the activity of antioxidant enzymes to cope with water deficit

Principal Component Analysis (PCA) was conducted to comprehensively understand the impact of water deprivation on the physiology of CaHB12-OE and non-transformed (NT) plants. PC1 and PC2 explained significant variations, accounting for 26.1% and 20.4% in the control, 33.2% and 21.7% in water restriction treatment, and 30.6% and 16.7% in rehydration treatment, respectively. Under well-watered conditions, NT plants were characterized by higher chlorophyll, carotenoid, and malondialdehyde (MDA) contents, while CaHB12-OE plants exhibited lower values (Supplementary Figure 11A). CaHB12-OE was distinguished by higher Fv/Fm, ascorbate peroxidase (APX) activity, and non-photochemical quenching (NPQ). During drought treatment, increased antioxidant activity, particularly catalase, in CaHB12-OE, associated with positive variations in transpiration (E) and actual quantum yield of photosystem II (ϕPSII), separated them from NT plants. This distinction was more pronounced in CaHB12-OE-2. Rehydration led to higher NPQ and carotenoid contents in NT plants. Proline and MDA contents, commonly used indicators of drought tolerance, did not directly associate with water deficit tolerance in transgenic plants. The MDA content increase triggered by water deprivation was similar across genotypes, followed by a recovery trend upon rehydration. Water deficiency induced a significant increase in leaf proline content mainly in NT plants, tending to reach well-watered levels

after rehydration (Figure 11B). Curiously, CaHB12-OE lines showed decreased chlorophyll and carotenoid contents in leaves but exhibited no penalties for photosynthetic activity (Supplementary Figure 13A-D).

3 Discussion

There are controversies in the literature about the evolutionary history of HD- Zip genes and their distribution between different kingdoms. However, there is a consensus that they encode a family of transcription factors implicated in various biological processes in plants, especially processes in response to drought and salinity stresses (Ariel et al., 2007; Wei et al., 2019; Hou et al., 2021). HB is composed of 60 amino acids and was originally characterised as a conserved element found in genes encoding TFs. The full picture of plant genomes recently allowed the classification of plant HD-containing proteins into 14 families: HD-Zip (subfamilies HD-Zip I, II, III and IV), PLINC, WOX, NDX, DDT, PHD, LD, PINTOX, SAWADEE, KNOX and BEL (Mukherjee et al., 2009). In the present study, we identified 288, 132, and 81 coffee homeobox genes using the *C. arabica*, *C. eugenioides*, and *C. canephora* databases, respectively, and included a list of diploidy (Arabidopsis), tetraploidy (cotton and soybean), and outgroup (rice and *Setaria*) plant species for comparative and robustness analysis (Figure 1 and Supplementary Table 1).

Similar to other members of the HD- Zip I family, *CaBH12* is stress responsive. The localization of the CaHB12 protein in the nucleus of coffee cells is in accordance with its TF function. The nuclear localization of other HD- Zip I members has already been shown (M. Dai et al., 2008; Deng et al., 2006; Wang et al., 2005; Zhang et al., 2012). Coffee NLS was identified at the beginning of the HD, similar to what has been described for the HD- Zip I protein CPHB7 from *Craterostigma plantagineum* (Deng et al., 2006). Coffee CaHB12 (KRKR) and *Craterostigma* CpHB7 (KKRR) NLS also shared sequence similarities, with a lysine (K) in position one (P1) and basic residues in P2 and P4.

Arce *et al*. (2011) showed that the C-terminal domain (CTD) of HD- Zip I proteins of diverse plant species contains conserved regions in which several motifs were identified, some of which were demonstrated to be functional (Hofer et al., 2009; Sakuma et al., 2010). One of these is the AHA activation motif (aromatic, large hydrophobic, acidic context) first found in HSF (Treuter et al., 1993). Such motifs were shown as functional transactivation domains in HD-Zip I TFs (Capella et al., 2014). On the other hand, sunflower HaHB4 has an atypical AHA motif, presenting a basic aa next to the acidic region, which is possibly related to its inhibitory activity (Arce et al., 2011). The CTR of CaHB12 seems to follow the same trend as AtHB7, AtHB12 and MtHB1, with a high acidic and hydrophobic content juxtaposed to a neutral residue. However, most of the DEGs in the transgenic line were down-regulated, as observed in *HaHB4*-overexpressing Arabidopsis plants (Manavella et al., 2006). CaHB12 has a longer CTR with a higher content of acidic and hydrophobic aa compared to AtHB7, AtHB12, MtHB1 and HaHB4. Further work must be performed to elucidate the role of CaHB12 CTR.

In addition to the characterization of the Homeobox family in *C. arabica/C. canephora* and the functional characterization of the CaHB12 protein, we also showed that transgenic *A. thaliana* constitutively expressing *CaHB12* exhibited a significant increase in drought tolerance at different developmental stages. These results are consistent with those previously described, corroborating the function of *CaHB12* in the plant response to drought stress and revealing its potential biotechnological application (Basso et al., 2021). When stress was slowly and continuously imposed, similar to field conditions, transgenic CaHB12-OE Arabidopsis showed higher survival rates than in rapid and abrupt water stress experiments. This approach is much more valuable for the selection of genes with biotechnological potential (Bhatnagar-Mathur et al., 2008).

We also demonstrated that CaHB12-OE soybean plants can cope better with drought stress than NT plants through stomatal and photosynthesis regulation without causing irreversible biochemical damage. Transgenic lines exhibited a

decrease in stomatal closure, causing photosynthesis to be less affected under water deprivation, as evidenced by the AN, Vcmax and Jmax values. These responses were accompanied by an increase in antioxidant activity but not by variations in chlorophyll content. After rehydration, CaHB12-OE lines showed a significant better recovery than NT plants, which demonstrated a higher demand for the dissipation of excessive energy due to restrictions in CO_2 assimilation through the xanthophyll cycle (Jahns and Holzwarth, 2012). Interestingly, cotton plants overexpressing *CaHB12* also presented minor stomatal limitations under drought stress, favouring CO_2 fixation capacity, and no alterations in chlorophyll content. These plants also presented lower leaf abscission and enhanced assimilation and water use efficiency (WUE) under stress, and these physiological effects were equally observed in *A. thaliana* CaHB12-OE (Basso et al., 2021; Ré et al., 2014; Alves-Ferreira et al., 2012). Similarly, soybean plants overexpressing another HD-Zip I gene from sunflower showed enhanced transpiration, which may contribute to increased WUE and a delay in leaf senescence as a consequence of a reduction in ethylene biosynthesis (Waltz et al., 2015; Ribichich et al. 2020; Winck et al., 2022). Similar effects were also observed in transgenic wheat expressing *HaHB4* (Gonzalez et al., 2019). This latter characteristic was not observed for CaHB12-OE soybean lines, possibly because drought stress was imposed in the reproductive stage, and the end of the stress coincided with the final stages of soybean development when leaves started to drop. Additionally, the proline content, an osmotic adjustor whose accumulation is positively correlated with drought tolerance in soybean (Nguyen et al., 2020; Pham et al., 2020), was higher in NT plants than in CaHB12-OE lines under drought treatment. A higher proline accumulation in NT plants could mean a response to signals of drought stress that are not present in CaHB12-OE plants due to its protective effect (Szabados and Savouré, 2010).

To further understand the function of *CaHB12* in drought tolerance, we identified and characterized the genes modulated by the overexpression of *CaHB12* in Arabidopsis, which, interestingly, revealed that several members of known water

stress regulatory routes were down-regulated in CaHB12-OE, including traditional drought stress marker genes, such as *RD22* (AT5G25610). These results suggest that enhanced drought tolerance triggered by *CaHB12* overexpression occurs through an ABA-independent pathway.

As mentioned above, among the several repressed genes belonging to the group of water deficit-induced responses, *RD22* was one of them, a gene from the caleosin family known to be a positive regulator of drought tolerance and an ABA-responsive gene that participates in the control of stomatal aperture, production of secondary compounds, regulation of ROS accumulation, ABA-mediated inhibition of germination and ABA stress signalling (Aubert et al., 2011; Phillips et al., 2017, Aubert et al., 2010). Interestingly, *AtHB7* and *AtHB12* expression is inducible by ABA treatment and hydric resource availability in Arabidopsis, and these proteins can regulate themselves and other ABA-dependent pathway elements (Olsson et al., 2004; Ré et al., 2014). In addition, other genes related to seed germination in response to ABA stimulus presented different regulation patterns, with the *GASA1* and *EM1* genes up-regulated and *GASA5* down-regulated (Zhang and Wang, 2011; Tezuka et al., 2012). Although transgenic seedlings expressing *CaHB12* did not seem to be ABA insensitive (Additional file 4), the increased salt tolerance observed during seed germination could still be related to this hormone.

In addition to the heat shock TF (HSF) ATHSFA2, eight heat shock proteins (HSPs) were induced in the presence of *CaHB12*. This HSF was first identified as high-light and heat stress responsive (Vierling, 1991; Nishizawa et al., 2006; Schramm et al., 2006) and as responsive to chilling stress and exogenous treatment with H_2O_2 (Nishizawa et al., 2006). More recently, HSFA2, which, together with HSFA3, composes a complex, was characterized as transcriptional memory after heat stress (Friedrich et al., 2011) which can positively influence H3K4 hyper-methylation, an important biological change involved with epigenetic modulations and priming response to stresses. Some miRNAs associated with heat (miR156, miR166) and H_2O_2 (miR827) stresses were also induced in transgenic lines (Figure

8) (Li et al., 2011; Stief et al., 2014). Transgenic *A. thaliana* and rice plants overexpressing class A2 HSFs presented enhanced tolerance to heat stress and high salinity, indicating a conserved mechanism in distinct plant species (Ogawa et al., 2007; Yokotani et al., 2008). The literature reports that *ATHSFA2* is not only able to regulate HSP expression but also regulates the expression of several antioxidant enzymes, such as GST, GR, POX and APX (Zhuang et al., 2018; Scharf et al., 2012). Therefore, it is likely that HSFs and HSPs could also enhance drought tolerance due to substantial overlapping of osmotic response mechanisms (Knight and Knight, 2001). Accordingly, the physiological impacts on transgenic soybeans harbouring the *35S:CaHB12* construct, such as its unaltered chlorophyll content and effects on stomatal aperture, suggest that the tolerance to drought is not ABA-mediated but rather an antioxidative response mediated by heat shock proteins, a hypothesis supported by the enhanced antioxidant enzyme activities observed in CaHB12-OE soybean events and confirmed by transcriptomic analysis in transgenic Arabidopsis, which exhibited increased expression of, for example, a chloroplastic superoxide dismutase (ATSOD2) as well as its activator ATCCS (Figure 9). The overexpression of *ATHSP17.6A*, for instance, a potential target of *ATHSFA2*, was able to induce enhanced tolerance to salt and osmotic stress (Sun et al., 2001). The literature provides evidence that the response mechanism activated by HSPs and HSFs in *Arabidopsis* by the ectopic expression of *CaHB12* is probably conserved in coffee (dos Santos et al., 2011; Marraccini et al., 2012). *CcHSP1* transcripts, for example, accumulated in the leaves of *Coffea canephora* during water deficit acclimation (Marraccini et al., 2012). In parallel, the up-regulation of HSPs is related to tolerance to drought and heat, as reported in transgenic tobacco and rice lines, where overexpression of *NtHSP70* and *OssHP17.7* enhanced tolerance to heat as well as drought in transgenic plants (Udvardi et al. 2007; Sato and Yokoya, 2008). In agreement, the overexpression of *OsTF1L*, a member of HD-Zip in rice, promotes an increase in survival rates and lignin accumulation and a strong induction in drought-related gene expression, such as *HSP70*, *CYP450* and *AP2*, supporting our transcriptomic data (Bang et al.,

2018). Our data suggest that the expression of *CaHB12* evokes protecting proteins, including HSPs and antioxidant enzymes, which reduces intracellular oxidative stress. Moreover, this inhibits the downstream ABA signalling cascade corroborated by the down-regulation of RCAR9, a regulatory component of the ABA receptor (Figure 9). In addition, several genes reported to be deregulated by drought and salt stresses and even ABA treatment exhibit clear alterations in this expression pattern in *CaHB12* transgenic lines, such as *COR15A, RCAR9, PCC*1 and the *HSP* gene family (Kilian et al., 2007; Winter et al., 2007).

The two most significant overrepresented pathways, MAPK signalling pathway and protein processing in endoplasmic reticulum are well-recognized pathways to be involved with plant drought stress. MAPK signalling cascade works as an early signaling network through the regulation of ROS in response to drought stress (Majeed et al., 2023). Regarding protein processing in endoplasmic reticulum, this pathway is important for storing misfolding proteins, which this misfolding can result in non-functional proteins and the accumulation of these non-functional proteins in other cellular components can cause intracellular damage (Manghwar and Li, 2022). Both pathways reinforce the phenotype observed by the up-regulation of CaHB12.

The responses to salicylic (SA) and jasmonic (JA) acid stimuli, as well as to ethylene stimuli, were generally repressed in CaHB12-OE *Arabidopsis*. The relationship of HD- Zip genes with jasmonate has already been reported in tobacco; for example, *NtHD9* and *NtHD12* regulate jasmonate signalling (Zhang et al., 2023). Although JA-related genes were induced in *HaHB4* plants (Manavella et al., 2008), ethylene-related genes were down-regulated (Manavella et al., 2006). The group of genes modulated in CaHB12-OE and HaHB4-OE Arabidopsis presented a small overlap (Supplementary Figure 6A), for instance, only three ethylene-related genes are repressed in CaHB12-OE and HaHB4-OE plants (*AP2*, *ERF-3* and *ERF-2*). However, several genes with similar functions were found to be down-regulated in CaHB12-OE and HaHB4-OE, indicating a similar mode of action of CaHB12 and HaHB4 to trigger drought tolerance (Supplementary Figure 6B).

The involvement of *DREB* and *SnRK* genes in the crosstalk between ABA-dependent and ABA-independent pathways has been extensively characterized. In addition, the literature reports the relationship of other TFs implicated in this convergence step (Agarwal et al., 2010; Yoshida et al., 2014; Hu and Yu., 2013; Ma et al., 2019). Interestingly, the overexpression of *CaHB12* in transgenic cotton promotes a significant increase in drought tolerance mediated by a fine modulation in ABA-related gene expression, although no significant differences in ABA content were observed in transgenic lines when compared to NT (Basso et al., 2021). Similarly, the heterologous expression of *AtHB6* in maize also enhances drought tolerance, reduces MDA levels after rewatering, promotes the activities of ROS scavenging enzymes and induces the gene expression of members of the ABA-dependent pathway (Jiao et al., 2022). Curiously, *AtHB12* plays a crucial role in ABA signalling, acting specifically in the coordination of the ABA-dependent pathway through the regulation of transcription of *AtPP2Cs* and *AtSnRK2.3s* (up-regulated) and the *AtPYR/PYL* gene family (down-regulated) (Olsson et al., 2004; Valdes et al., 2012). Comprehensive analysis of mutant lines, such as transcriptome and proteome analyses, should provide further insights into transcriptional networks under stress conditions and highlight this crosstalk during stress signalling.

The results presented here, together with the reported literature, indicate that delayed senescence could contribute to increased drought and salt tolerance conferred by *CaHB12*. Among the ethylene-related set of genes, *RAV1* is a TF that positively regulates leaf senescence in *Arabidopsis* (Woo et al., 2010) and leads to retardation of lateral root development (Hu et al., 2004). The observed repression of *RAV1* is, therefore, consistent with a more exuberant root system, which has been related to increased drought tolerance in coffee plants (Pinheiro et al., 2005). Other genes associated with root growth and development were induced in CaHB12-OE plants, e.g., *THAS1*, *CYP708A2* and *CYP705A5*. Although no differences were observed in root growth in CaHB12-OE plants, it cannot be excluded that *CaHB12* leads to increased root growth under appropriate

physiological conditions.

4. Conclusions

The coffee HD-Zip I γ-clade gene, *CaHB12*, encodes a nuclear-localized protein, significantly up-regulated in the leaves and lateral roots of *Coffea arabica* under water deficit conditions. RNA-Seq analysis of CaHB12-overexpressing Arabidopsis indicates that the HSF and HSP signalling pathways play a major role in the enhanced drought tolerance conferred by CaHB12. Furthermore, overexpression of CaHB12 also confers increased drought tolerance in both soybean and cotton, suggesting conservation of these pathways across distantly related species. Notably, the ectopic expression of CaHB12 did not result in any discernible morphological changes in Arabidopsis or soybean, underscoring the gene's potential as a target for enhancing stress tolerance in crops.

5. Methods

5.1 Identification of coffee homeobox genes

A local *A. thaliana* HB database was created as a result of repeated surveys of The Arabidopsis Information Resource (http://www.arabidopsis.org/index.jsp), including searches against the TAIR10_cdna_20101214 file using the local BLASTN tool. A total of 110 sequences of *Arabidopsis* were used to conduct local BLASTN against the rice genome database (http://rice.plantbiology.msu.edu/) and the Brazilian Coffee Genome Project database (http://www.lge.ibi.unicamp.br/cafe/) using the search parameter matrix BLOSUM62 and an e-value of 0.111 as the cut-off.

5.2 Bioinformatics analysis of homeobox genes in *Coffea* and related plant species

The homeobox sequences from three coffea plant species (*Coffea arabica*, *C. canephora*, and *C. eugenioides*), *Arabidopsis thaliana*, *Gossypium hirsutum*, *Glycine max*, *Oryza sativa*, and *Setaria viridis* were obtained from the public databases World Coffee Research (https://worldcoffeeresearch.org; Scalabrin et al., 2020), Coffee Genome Hub (https://coffee-genome.org; Denoeud et al., 2014), NCBI GenBank (https://www.ncbi.nlm.nih.gov/genome/73741?genome_assembly_id=418179), TAIR (https://www.arabidopsis.org), and Phytozome v12.0 (https://phytozome.jgi.doe.gov/pz/portal.html).

All *A. thaliana* 110 and *O. sativa* 110 Homeobox genes previously identified (Mukherjee et al., 2009) were downloaded from TAIR and *Oryza sativa* v7_JGI from Phytozome v12.0 and used as bait to search against three coffea plant species, Arabidopsis, cotton, soybean, rice, and the *Setaria* genome and proteome database. All protein recovery sequences from these plant species were analysed in the PFAM v33.1 (https://pfam.xfam.org; Mistry et al., 2021) and SMART (http://smart.embl-heidelberg.de; Letunic et al., 2021) databases. All those sequences containing the Homeodomain (PF00046 and SM000389) typical of HD-Zip I, HD- Zip II, HD- Zip III, HD- Zip VI, WOX, PHD, DDT, SAWADEE, LD, PINTOX, POX, and NDX classes, containing the ZF-HD_dimer (PF04770) typical of the PLINC class, and containing KNOX1 (PF03790 and SM001255), KNOX2 (PF03791 and SM001256), and/or Homeobox_KN (PF05920) typical of KNOX I and KNOX II classes, were kept for the following analysis.

All multiple sequence alignments were conducted using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (https://www.ebi.ac.uk/Tools/msa/muscle/; Madeira et al., 2019) using default settings. Phylogenetic analysis was estimated using the Maximum Likelihood (ML) method at PhyML 3.0 (Guindon et al., 2010) and using Smart Model Selection (SMS) (Lefort et al., 2017) to select the best model, such as the VT+G+F model of sequence evolution and aLRT branch support testing (Anisimova and

Gascuel, 2006). The phylogenetic tree was visualized using iTOL software (https://itol.embl.de; Letunic and Bork, 2019).

5.3 *In silico* expression analysis

For *in silico* expression analysis, the number of sequence reads computed for each gene in several *C. arabica* libraries from the Brazilian Coffee Genome Project database was normalized by the total number of reads in each library. The expression values were used as input for Cluster and Tree View software (Eisen et al., 1998; Page, 1996). Only libraries presenting reads for at least one of the putative coffee HB genes were considered for the analysis (Additional file 8).

5.4 Isolation of RNA and cDNA synthesis

Total RNA isolation from coffee leaves or lateral root tissue samples, as well as *DNase* I (Invitrogen by Life Technologies, California, USA) treatments and cDNA synthesis, were carried out as previously described (Cruz et al., 2009).

5.5 Quantitative RT-PCR (RT-qPCR)

The RT-qPCRs were performed in triplicate and carried out on the 7500 Fast Real-Time PCR platform (Life Technologies, California, USA), according to Cruz *et al.* (2009), using primers C11116F, C11116R, C11933F, C11933R, 7747F and 7747R for *CaHB12*, *CaHB1-like2* and *CaZHD4* expression analyses, respectively (Supplementary Table 11). Primer amplification efficiency and cycle threshold (Ct) values were calculated using Miner software (Zhao and Fernald, 2005). Ct values were converted to normalized relative quantities using qBase v1.3.5 software (Hellemans et al., 2007). Specific reference genes were used for data normalization of drought-stressed leaves and roots as previously described (Cruz et al., 2009) (Supplementary table 11).

5.7 Subcellular localization

Complete and truncated coffee CaHB12 CDS were amplified using the HB12NcoF and HB12NcoR and the HB12Nco-truF and HB12NcoR primer pairs,

respectively, which contained NcoI restriction sites (Supplementary Table 11). Fragments were fused in frame to green fluorescent protein (GFP) and cloned under the control of a double cauliflower mosaic virus 35S promoter (CaMV35S:CaHB12:GFP or CaMV35S:CaHB12tru:GFP) into the pGJ280 vector (Willige et al., 2009). Fully expanded leaves of *C. arabica* were bombarded with Biolistic® PDS-1000/He (BIO-RAD) equipment with 3 mg of 1 µm gold particles (BIORAD) coated with 15 µg of plasmid DNA obtained from a large-scale extraction with the NucleoBond® Xtra Midi/Maxi kit (Macherey-Nagel, Düren, Germany). Coffee leaves were subsequently immersed in water and transferred to a growth chamber (18 h light/6 h dark) for 16 h. GFP fluorescence was analysed after 16 h using a Nikon Eclipse TE2000-U/D-Eclipse C1 confocal microscope. Images were captured with EZ-C1 v 3.20 software (Nikon, Düsseldorf, Germany).

5.8 Generation of transgenic plants

Arabidopsis thaliana

The full-length *CaHB12* coding sequence (CDS) was amplified using primers GW11116/11116mon (Supplementary Table 11) and cloned into the entry vector pENTR®D-TOPO. Then the insert was recombined with the binary vector pB2GW7. Electrocompetent *Agrobacterium tumefaciens* (strain GV3101) cells were transformed with the recombinant binary vector pB2GW7 (Karimi et al., 2002) containing the construction CaMV35S:CaHB12. Twelve *A. thaliana* plants (Col-0) were transformed using the floral dip procedure (Desfeux et al., 2000). The T2 plants presenting 100% herbicide resistance (glufosinate ammonium salt - 10 $\mu g.mL^{-1}$) were selected, and the T3 seeds were used in stress tolerance assays. The expression of *CaHB12* was verified by RT–PCR, and *actinII* was used as a reference, using primers GW11116 and 1116mon for *CaHB12* and Ath_ActinII-F and Ath_ActinII-R for *ActinII* (Supplementary Table 11).

Glycine max

Seeds from the *G. max* Conquista Embrapa (Brazil) variety were soaked in

70% ethanol for 2 min, decontaminated with 2.5% sodium hypochlorite for 20 min, and incubated in sterile water for 16 h at room temperature under dark condition. Then, embryonic axes were excised and used for biolistic transformation, as described by Rech *et* al. (2008).The cassette used for soybean transformation was engineered for the overexpression of the coding sequence of HB12 gene from *Coffea arabica* under the control of the *G. max* strong constitutive ubiquitin-conjugating enzyme gene promoter (pUceS8.3), including its 5′-untranslated region (UTR) (Fragoso et al., 2022) The screening of transgenic plants was performed across generations by PCR analysis (primers *HB12.Gm.F* and HB12.Gm.R, Supplementary table 11) and semiquantitative indirect Enzyme Linked Immunosorbent assays (ELISA), as described by Basso *et al.* (2021) (Supplementary Table 11).

5.9 Plant growth conditions and stress treatments

Coffea arabica

The *C. arabica* cultivar "Catuaí Vermelho IAC44" was grown in a greenhouse located at the Universidade Federal do Rio de Janeiro (22° 54' 10" S/43° 12' 27" W – Rio de Janeiro – Brazil) in large trays containing 20 plants each (temperature: 21± 2 °C and natural photoperiod). Control plants received 500 mL of water that was poured directly into the tray at 1-day intervals. The water potential (ψ_w) of each plant was measured at predawn with a Scholander-type pressure chamber. Water deficit was induced by withholding water for 10 days. Samples of fully expanded leaves (third pair) and lateral roots were collected after 2, 5 and 10 days. Experiments were performed in duplicate (June/2006 and June/2007, 12 ±1 h/light and 12 ±1 h/dark), and each sample was composed of leaf or root material from five plants. All samples were immediately frozen in liquid nitrogen and kept at -80 °C for RNA isolation.

Arabidopsis thaliana

Wild-type and transgenic T3 lines of *A. thaliana* expressing coffee

homeobox *CaHB12* were grown in a growth chamber at 22 °C (± 2 °C) under a photoperiod of 18 h light/6 h dark.

To mimic water deficit conditions, PEG8000 (polyethylene glycol)-infused plates were used to reduce the water potential of the culture medium to -1.2 MPa (Verslues and Bray, 2004). Plants were sown on half-strength MS (Murashige and Skoog, 1962) supplemented with 6 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer (pH 5.7) for 20 days. Thirty plants of each genotype were transferred to PEG-treated plates and plates without PEG treatment. Plant material was collected seven days after transfer, frozen in liquid nitrogen, and stored at -80 °C for malondialdehyde (MDA) quantification. Two biological replicates were evaluated for each genotype.

For water deficit assays conducted in soil, four seeds of a given genotype were sown directly in 100 g of soil in 8x7 cm pots (Dezar et al., 2005). Rapid water deficit was imposed by stopping watering after the first pair of leaves was apparent until permanent wilting symptoms were observed in wild-type plants. All experiments were conducted in duplicate using 98 wild-type plants and 64, 32 and 64 plants of transgenic lines CaHB12-OE 1-3 respectively. A continuous water deficit was imposed 30 days after the first pair of leaves was apparent. One millilitre of water per day was then added to each pot until wilting symptoms were observed in wild-type plants. All experiments were conducted in duplicate in 64 wild-type plants and 56 plants of each transgenic line. In both experiments, the number of survivors was counted 48 h after rehydration.

For salt stress assays, plants were sown on MS medium supplemented with 100 mM (four plates with 50 seeds each) and 150 mM (two plates with 50 seeds each) NaCl (Liu et al., 2009). Germination rates were determined seven days after transferring to the plates. Only individuals with roots completely inserted in the culture medium were considered positive for germination. In each experiment, three sets of 10 plants were frozen in liquid nitrogen and stored at -80 °C for MDA measurements. Control plants were grown in MS medium without NaCl.

Glycine max

Seeds representing the T3 self-pollinated offspring of three independent transgenic lines and seeds from non-transformed (NT) plants were germinated in 5-L pots filled with a mixture of clay soil, sand, and Vivatto commercial substrate (Technes, Sao Paulo, Brazil) (3:1:1), supplemented with NPK 4-14-8 fertiliser. Plants were grown in a greenhouse at 28°C ± 7 and 75% humidity, under natural light regime. The experiment was carried out in a completely randomised design with six replicates. Plants were watered to soil field capacity (FC) every other day until reaching the reproductive (R3) stage, when water was withheld. The analyses were performed in well-watered plants before water suspension, considered as control condition, in drought-stressed plants at 15 % FC for 15 days, and in plants rehydrated for five days. Soil relative humidity was monitored daily with the HS2 HydroSense II soil moisture sensor (Campbell Scientific, Shepshed, UK). Lateral leaflets from the first fully expanded young trifoliate leaf from the top of the plant were collected at each analysis point, immediately frozen in liquid nitrogen, and stored at − 80 °C for further assays.

5.10 Evaluation of physiological and biochemical variables

The first fully expanded young trifoliate leaf from the top of the soybean plant was used for phytochemical analyses. Gas exchange and fluorescence measurements were performed simultaneously from 8:00 am to 12:30 am on the central leaflet with the portable open-flow gas exchange system (LI-6400XT, LI-COR, Nebraska, USA) attached with an integrated fluorescence chamber head (LI-6400-40, LI-COR Inc.), using saturating photosynthetic photon flux density of 1,000 μmol photons m−2 s−1 and CO_2 concentration of 400 ppm. The following parameters were determined: net CO_2 assimilation rate (AN), stomatal conductance (gs), intercellular CO_2 concentration (Ci), transpiration (E), steady-state fluorescence (Fs), maximum fluorescence during a light-saturating pulse (Fm'), and minimal fluorescence during a light-saturating pulse (Fo'). At 8:00 pm of the same day, the maximum (Fm) and minimal (Fo) fluorescence were determined in dark-adapted leaflets. All measurements followed procedures described by Genty *et al.* (1989). Chlorophyll a fluorescence variables were calculated as Maxwell and

Johnson (2000). AN/Ci curves were measured between 2 and 4 pm using saturating photosynthetic photon flux density of 1,000 µmol photons m−2 s−1 and CO_2 concentration varying from 50 to 1,600 µmol mol-1. Electron transport rate (JMAX), maximum ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) carboxylation rate (VCMAX), and triose phosphate utilization rate (TPU) were calculated from AN/Ci curves, as described by Farquhar *et al*. (1980).

Pigments were extracted from 100 mg of leaf tissue in 1 mL of ethanol. The chlorophylls a, b and carotenoids contents were calculated as described by Porra *et al*. (1989). The activity of the antioxidant enzymes SOD, APX and CAT were determined in six biological samples and three technical replicates following previously described protocols (Giannopolitis and Ries, 1977; Nakano and Asada 1981; Havir and McHale 1987). Membrane damage was measured for soybean and Arabidopsis by determination of malonyl dialdehyde (MDA) content, which is a marker of lipid peroxidation, from 100 mg of grounded leaf, according to Cakmak and Horst (1991). Lipid peroxidation in Arabidopsis transgenic lines was determined by measuring the MDA content in seedlings using the thiobarbituric acid (TBA) test (Kotchoni et al., 2006).

5.11 Transcriptome sequencing and analysis (RNA-Seq and sRNA-Seq)

Whole transcriptome deep sequencing on the Illumina Hi-Seq platform was performed for two biological replicates of wild-type (Col-0) and transgenic (CaHB12-OE) genotypes.

Total RNA was extracted from pools of 50 seedlings harvested 10 days after germination with a *RNeasy® Plant Mini Kit* (Qiagen). Total RNA (4 and 10 µg) was used for mRNA and sRNA library construction, respectively. The sequencing reaction was performed by Fasteris SA (Switzerland). Regarding sRNA-seq, six libraries were produced, since an additional technical replicate was obtained for one wild-type and one transgenic library. Such technical replicates were analysed as one by averaging their reads for each condition (wild-type or transgenic libraries).The read quality in all libraries was assessed using FastQC ver. 0.11.5

(Andrews, 2010), followed by a cleaning step in which Trimmomatic ver. 0.38 (Bolger et al., 2014) software using the following parameters: sensitive detection of Illumina adapters by 2:30:10, sliding windows of four bases to eliminate bases with a quality (Q) less than 15, and reads with a length less than 50 were discarded. As a final library quality control step, we ran FastQC software over the cleaned libraries for comparison with the raw libraries.

Each cleaned library was mapped against the *A. thaliana* genome (ARAPORT 11) using STAR v. 2.7 (Dobin et al., 2013) software with default parameters. The count step of reads mapped to the features (genes) was performed with HTSeq ver. 0.11.2 (Anders et al., 2015) software. Only the reads mapped uniquely over the genes were used for the differential expression analysis step.

For the differential expression analysis, we discarded genes with low expression (CPM row sum < 1). EdgeR ver. 3.36 (Robinson et al., 2010) statistical software package for R was used to carry out the differential expression analysis. Leading analysis was used for quality control (Supplementary Figure 14) and a Fisher-like exact test based on a negative binomial common dispersion was estimated, and then, a moderate dispersion estimate was obtained for each tag (e.g., each gene for the RNA-Seq data and each miRNA family for the sRNA-Seq data) through its weighted individual dispersion (prior.n=10, trend=FALSE) (Robinson and Smyth, 2007; Robinson and Smyth, 2008). Next, Fisher-like exact tests based on negative binomial distribution were conducted to test whether group means (wild type versus transgenic) were significantly different from one another. *P* values were adjusted for multiple tests with Benjamini–Hochberg correction (Benjamini and Hochberg, 1995) to control the false discovery rate (FDR). Genes were considered differentially expressed with FDR<0.05.

To identify which gene ontology (GO) terms and KEGG pathways which were enriched in the set of differentially expressed genes, up- and down-regulated genes separately, a set-based enrichment was performed using a custom R script based on Fisher's exact test. GO terms were considered overrepresented based on an FDR threshold of less than 0.05, and KEGG pathways based on p value with the

same value. For visualization of enriched features, we used the ggplot2 ver. 3.3.5 (Wickham, 2016) R package to build bubble charts and Cytoscape ver. 3.7.2 (Shannon et al., 2003) to build a network of enriched GO terms and their associated DEGs.

String database ver. 11 (Szklarczyk et al., 2019) was used to retrieve the interactions of DEGs from *A. thaliana*. The graphical representation of the network was built with Cytoscape.

Lastly, genes of interest previously reported as drought responsive (memory and non-memory, and ABA) and differentially expressed transcriptional factor genes had their expression analysed. For memory and non-memory genes we retrieved them from Ding *et al.* (2012), ABA related genes from Thonglim *et al*. (2023) and transcription factors from PlantTFDB (Jin et al., 2017).

The transcriptional profile of selected genes was determined by RT-qPCR as aforementioned (see 5.4 and 5.5 items). Linear data were normalized to the mean cycle threshold (C_t) of AT1G05850 and AT4G34270 (Czechowski et al., 2005).

miRNA identification

miRCat software (plant version) was initially used to align all 18- to 24-nt-long sRNAs sequenced to the *A. thaliana* TAIR 9 genome using default parameters (Moxon et al., 2008). All putative miRNA precursor sequences extracted from the genome by the miRCat software were mapped against the redundant set of 18- to 24-nt-long sRNAs using the Bowtie software (Langmead et al., 2009). The Bowtie output was filtered with an *in-house* filter tool (FilterPrecursor) to identify candidate sequences as miRNA precursors. The software scans the data and identifies patterns of one or two blocks of perfect matching sRNAs that correspond to where mature and/or star miRNA sequences align (Kulcheski et al., 2011). The filtering was performed with the following default parameters. FilterPrecursor`s source code can be downloaded at http://code.google.com/p/filter-precursors/downloads/list. The resulting files, containing the 18- to 24-nt sRNAs mapped to the putative miRNA precursors, were then visually inspected using

Tablet software (Milne et al., 2010). Sequences having the expected pattern of sRNA mapping to their respective precursors were checked by the RNA folding tool from the UEA sRNA toolkit (Moxon et al., 2008) using default parameters. The following criteria were used to define a good miRNA candidate: no more than four unpaired nucleotides between the putative mature and star sequences, of which no more than three nucleotides were consecutive and no more than three nucleotides were without a corresponding unpaired nucleotide in the near complementary sequence within the hairpin structure (Meyers et al., 2008). miRNAs were then classified according to BLAST searches against the miRBase database (Griffiths-Jones, 2004).

The targets of the predicted miRNAs were identified in the *A. thaliana* genome with psRNATarget software (Dai et al., 2018) using the following parameters: maximum expectation value: 3; multiplicity of target sites: 2; and nucleotide range of central mismatch for translational inhibition: 9–11. Candidate sequences were annotated based on BLASTN and PFam searches (Punta et al., 2012).

Acknowledgements

We thank the researchers Eduardo Arcoverde Ph.D. for helping with a Scholander-type pressure chamber, Alan Carvalho Ph.D. and Felipe Rodrigues da Silva Ph.D. from Embrapa Genetic Resources and Biotechnology for allowing and helping with the initial searches of the Brazilian Coffee Genome Project database, Guilherme Loss Morais Ph.D. for helping with bioinformatics for sRNA discovery, David da Cunha Valença Ph.D for helping with statistical analysis and Fátima Barbosa, Luiz Frade and Durvalina Felix for technical assistance. This work was part of Fernanda P Cruz Ph.D.'s research in Genetics at the Department of Genetics of the Universidade Federal do Rio de Janeiro (UFRJ),and was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (INCT Biotec Seca-Pragas, 465480/2014-4), by awards from the National Council for Scientific and Technological Development (CNPq; Grant 316821/2021-7) by the Fundação

de Amparo à Pesquisa do Rio de Janeiro (FAPERJ CNE, E-26/200.464/2023) to MA-F. ; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); Centro Brasileiro-Argentino de Biotecnologia (CBABIO); and Deutscher Akademischer Austauschdienst (DAAD), DAAD-BMBF- Scholarship within the program "Studies and Research in Sustainability". We declare no conflicts of interest for any work cited in this study.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

FPC carried out the water deficit experiments and qPCR analysis of the coffee samples. Genetic constructs, subcellular localization studies and the generation of transgenic *Arabidopsis thaliana* plants were performed by FPC. CD performed the water deficit experiments using transgenic *Arabidopsis*. FPC and RKTML carried out the salt, ABA and MDA experiments performed *in vitro* using transgenic *Arabidopsis*. TF and E Romanel performed the identification of coffee homeobox genes, the phylogenetic analysis, digital expression analysis, interpretation, and discussion. RNA-Seq analysis, including statistical analysis, was performed by RKTML, BPM and LWPA. Digital expression, enrichment analyses and systems biology approaches were performed by LWPA, together with results and discussion writing. RT-qPCR validations were performed by RKTML. Small RNA identification was performed by ILGA, and statistics were performed by BPM and MLCA. CVM and TPR contributed to the production and characterization of soybean transgenic lines. CVM, JVAC, and SMM performed soybean water deficit experiments and statistical analysis. RLC, DB, MFGS and MAF contributed to the experimental design, data interpretation and discussion. MLCA, E Romano, ABA and RLCo also contributed to data interpretation. FPC, RKTML, RLCo,, E. Romanel, MLCA and LWPA wrote the manuscript. All authors read and approved the final version of the manuscript.

References

Agalou, A., Purwantomo, S., Övernäs, E., Johannesson, H., Zhu, X., Estiati, A., De Kam, R. J., Engström, P., Slamet-Loedin, I. H., Zhu, Z., Wang, M., Xiong, L., Meijer, A. H., Ouwerkerk, P. B. F. (2008). A genome-wide survey of HD-Zip genes in rice and analysis of drought-responsive family members. *Plant Molecular Biology*, *66*(1–2), 87–103. https://doi.org/10.1007/s11103-007-9255-7

Agarwal P. K., Shukla P. S., Gupta K., Jha B. (2013). Bioengineering for salinity tolerance in plants: state of the art. *Mol. Biotechnol*. 54 102–123. https://doi.org/10.1007/s12033-012-9538-3

Allen, E., Xie, Z., Gustafson, A. M., Sung, G.-H., Spatafora, J. W., Carrington, J. C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nature Genetics*, *36*, 1282–1290. https://doi.org/10.1038/ng1478.

Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., White, R. G., Millar, A. A. (2010). The MicroRNA159-Regulated GAMYB-like genes inhibit growth and promote programmed Cell Death in Arabidopsis. *Plant Physiology*, *154*(2), 757–771. https://doi.org/10.1104/pp.110.160630

Alves, G. S. C., Torres, L. T., Déchamp, E., Breitler, J.C., Joët, T., Gatineau, F., Andrade, A. C., Bertrand, B., Marraccini, P., Etienne, H. (2017). Differential fine-

tuning of gene expression regulation in coffee leaves by CcDREB1D promoter haplotypes under water deficit, *Journal of Experimental Botany*, Volume 68, Issue 11, Pages 3017–3031, https://doi.org/10.1093/jxb/erx166

Alves-Ferreira, M., da Cruz Waltenberg, F.P., de Campos Pinto, E.R., Grossi-de-Sa, M.F. (2012). Use of the coffee homeobox gene CaHB12 to produce transgenic plants with greater tolerance to water scarcity and salt stress. Patent. US20130340113A1.

Anders, S., Pyl, P. T., Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics, 31(2), 166-169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Anisimova, M., Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology*, *55*(4), 539–552. https://doi.org/10.1080/10635150600755453

Arce, A. L., Raineri, J., Capella, M., Cabello, J. V., Chan, R. L. (2011). Uncharacterized conserved motifs outside the HD-Zip domain in HD-Zip subfamily I transcription factors; a potential source of functional diversity. *BMC Plant Biology*, *11*(42), 1–19. https://doi.org/10.1186/1471-2229-11-42

Ariel, F., Diet, A., Verdenaud, M., Gruber, V., Frugier, F., Chan, R., Crespi, M.

(2010). Environmental regulation of lateral root emergence in *Medicago truncatula* requires the HD-zip i transcription factor HB1. *Plant Cell*, *22*(7), 2171–2183. https://doi.org/10.1105/tpc.110.074823

Ariel, F., Manavella, P. A., Dezar, C. A., Chan, R. L. (2007). The true story of the HD-Zip family. *Trends in Plant Science*, *12*(9), 419–426. https://doi.org/10.1016/j.tplants.2007.08.003

Aubert, Y., Leba, L. J., Cheval, C., Ranty, B., Vavasseur, A., Aldon, D., Galaud, J. P. (2011). Involvement of RD20, a member of caleosin family, in ABA-mediated regulation of germination in *Arabidopsis thaliana*. *Plant Signaling and Behavior*, *6*(4), 538–540. https://doi.org/10.4161/psb.6.4.14836

Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T., Vavasseur, A., Galaud, J. P. (2010). RD20, a Stress-Inducible Caleosin, Participates in Stomatal Control, Transpiration and Drought Tolerance in *Arabidopsis thaliana*, *Plant and Cell Physiology*, Volume 51, Issue 12, December 2010, Pages 1975–1987. https://doi.org/10.1093/pcp/pcq155

Ayala, F., Fedrigo, G. V., Burachik, M., Miranda, P. V. (2019). Compositional equivalence of event IND-ØØ412-7 to non-transgenic wheat. *Transgenic Res. 28* (2), 165–176. https://doi.org/10.1007/s11248-019-00111-y

Bang, S.W., Lee, D.-K., Jung, H., Chung, P.J., Kim, Y.S., Choi, Y.D., Suh, J.-W. and Kim, J.-K. (2019). Overexpression of OsTF1L, a rice HD-Zip transcription factor, promotes lignin biosynthesis and stomatal closure that improves drought

tolerance. *Plant Biotechnol J*, 17: 118-131. https://doi.org/10.1111/pbi.12951

Basso, M. F., Costa, J. A., Ribeiro, T. P., Arraes, F. B. M., Lourenço-Tessutti, I. T., Macedo, A. F., Lisei-de-As, M.E., Santos Amorim, R.M., Pinto, E.R.C., Silva, M.C.M., Morgante, C.V., Floh, E.I.S., Alves-Ferreira, M., Grossi-de-Sa, M.F. (2021). Overexpression of the CaHB12 transcription factor in cotton (*Gossypium hirsutum*) improves drought tolerance. *Plant Physiol. Biochem*. 165, 80–93. https://doi.org/10.1016/j.plaphy.2021.05.009

Benjamini, Y., Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bergau, J. (2019). Verdeca Introduces HB4® Drought Tolerant Soybeans to Growers at Argentina's Expoagro [Online]. https://www.businesswire.com https://www.businesswire.com/news/home/20190326005300/en/VerdecaIntroduces-HB4%C2%AE-Drought-Tolerant-Soybeans-Growers. *Arcadia Biosciences.*

Bhatnagar-Mathur, P., Vadez, V., Sharma, K. K. (2008). Transgenic approaches for abiotic stress tolerance in plants: Retrospect and prospects. *Plant Cell Reports*, *27*(3), 411–424. https://doi.org/10.1007/s00299-007-0474-9

Bhattacharjee, A., Ghangal, R., Garg, R., Jain, M. (2015). Genome-wide analysis of homeobox gene family in legumes: Identification, gene duplication and expression profiling. *PLoS ONE*, *10*(3), 1–22.

https://doi.org/10.1371/journal.pone.0119198

Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170, https://doi.org/10.1093/bioinformatics/btu170

Cakmak, I. and Horst, W.J. (1991). Effect of aluminium on lipid peroxidation, superoxide dismutase, catalase, and peroxidase activities in root tips of soybean (*Glycine max*). *Physiologia Plantarum*, 83: 463-468. https://doi.org/10.1111/j.1399-3054.1991.tb00121.x

Capella, M., Ré, D.A., Arce, A.L., Chan, R.L. (2014). Plant homeodomain-leucine zipper I transcription factors exhibit different functional AHA motifs that selectively interact with TBP or/and TFIIB. *Plant Cell Reports* 33: 955–967. https://doi.org/10.1007/s00299-014-1576-9

Carr, M. K. V. (2001). The water relations and irrigation requirements of coffee. *Experimental Agriculture*, *37*, 1–36. https://doi.org/10.1017/S0014479701001090

Chai, W., Si, W., Ji, W., Qin, Q., Zhao, M., Jiang, H. (2018). Genome-Wide Investigation and Expression Profiling of HD-Zip Transcription Factors in Foxtail Millet (*Setaria italica* L.). *BioMed Research International*, *2018*, 1–18. https://doi.org/10.1155/2018/8457614

Chen X., Chen Z., Zhao H., Zhao Y., Cheng B., Xiang Y. (2014). Genome-Wide Analysis of Soybean HD-Zip Gene Family and Expression Profiling under Salinity

and Drought Treatments. *PLoS ONE*, *9*(2): e87156. https://doi.org/10.1371/journal.pone.0087156

Cruz, F., Kalaoun, S., Nobile, P., Colombo, C., Almeida, J., Barros, L. M. G., Romano, E., Grossi-De-Sá, M. F., Vaslin, M., Alves-Ferreira, M. (2009). Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. *Molecular Breeding*, *23*(4), 607–616. https://doi.org/10.1007/s11032-009-9259-x

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., Scheible, W.-R. (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. *Plant Physiology*, *139*, 5–17. https://doi.org/10.1104/pp.105.063743

Dai, M., Hu, Y., Ma, Q., Zhao, Y., Zhou, D. X. (2008). Functional analysis of rice HOMEOBOX4 (*Oshox4*) gene reveals a negative function in gibberellin responses. *Plant Molecular Biology*, *66*(3), 289–301. https://doi.org/10.1007/s11103-007-9270-8

Dai, X., Zhuang, Z., Zhao, P. X. (2018). psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, *46*(W1), W49–W54. https://doi.org/10.1093/nar/gky316

Damatta, F. M., Ramalho, J. D. C. (2006). Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology*, *18*(1), 55–81. https://doi.org/10.1590/S1677-04202006000100006

Damatta, F.M., Avila, R.T., Cardoso, A.A., Martins, S.C.V, Ramalho, J. C. (2018). Physiological and Agronomic Performance of the Coffee Crop in the Context of Climate Change and Global Warming: A Review. *J Agric Food Chem.* 30;66(21):5264-5274. https://doi.org/10.1021/acs.jafc.7b04537

Deng, X., Phillips, J., Bräutigam, A., Engström, P., Johannesson, H., Ouwerkerk, P. B. F., Ruberti, I., Salinas, J., Vera, P., Iannacone, R., Meijer, A. H., Bartels, D. (2006). A homeodomain leucine zipper gene from *Craterostigma plantagineum* regulates abscisic acid responsive gene expression and physiological responses. *Plant Molecular Biology*, *61*(3), 469–489. https://doi.org/10.1007/s11103-006-0023-x

Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., Zheng, C., Alberti, A., Anthony, F., Aprea, G., Aury, J.M., Bento, P., Bernard, M., Bocs, S., Campa, C., Cenci, A., Combes, M,C., Crouzillat, D., Da Silva, C., Daddiego, L., De Bellis, F., Dussert, S., Garsmeur, O., Gayraud, T., Guignon, V., Jahn, K., Jamilloux, V., Joët, T., Labadie, K., Lan, T., Leclercq, J., Lepelley, M., Leroy, T., Li, L.T., Librado, P., Lopez, L., Muñoz, A., Noel, B., Pallavicini, A., Perrotta, G., Poncet, V., Pot, D., Priyono, Rigoreau, M., Rouard, M., Rozas, J., Tranchant-Dubreuil, C., VanBuren, R., Zhang, Q., Andrade, A.C., Argout, X., Bertrand, B., de Kochko, A., Graziosi, G., Henry, R.J., Jayarama, Ming, R., Nagai, C., Rounsley, S., Sankoff, D., Giuliano, G., Albert, V.A., Wincker, P., Lashermes, P. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*. *345*(6201):1181-4.

https://doi.org/10.1126/science.1255274.

Desfeux, C., Clough, S. J., Bent, A. F. (2000). Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. *Plant Physiology*, *123*(3), 895–904. https://doi.org/10.1104/pp.123.3.895

Dezar, C. A., Gago, G. M., González, D. H., Chan, R. L. (2005). *Hahb-4*, a sunflower homeobox-leucine zipper gene, is a developmental regulator and confers drought tolerance to *Arabidopsis thaliana* plants. *Transgenic Research*, *14*(4), 429–440. https://doi.org/10.1007/s11248-005-5076-0

Ding, Y., Fromm, M., Avramova, Z. (2012). Multiple exposures to drought 'train' transcriptional responses in Arabidopsis. *Nature Communications*, *3*(740). https://doi.org/10.1038/ncomms1732

Ding, Y., Tao, Y., Zhu, C. (2013). Emerging roles of microRNAs in the mediation of drought stress response in plants. *Journal of Experimental Botany*, *64*(11), 3077–3086. https://doi.org/10.1093/jxb/ert164

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21. https://doi.org/10.1093/bioinformatics/bts635

dos Santos, T. B., Budzinski, I. G. F., Marur, C. J., Petkowicz, C. L. O., Pereira, L. F. P., Vieira, L. G. E. (2011). Expression of three galactinol synthase isoforms in

Coffea arabica L. and accumulation of raffinose and stachyose in response to abiotic stresses. *Plant Physiology and Biochemistry*, *49*(4), 441–448. https://doi.org/10.1016/J.PLAPHY.2011.01.023

Eisen, M. B., Spellman, P. T., Brown, P. O., Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA*, *95*(25), 14863–14868. https://doi.org/10.1073/pnas.95.25.14863

FAOSTAT. (2023). https://www.fao.org/faostat/en/#data

Farquhar GD, von Caemmerer S, Berry JA. (1980) A biochemical model of photosynthetic CO2 assimilation in leaves of C 3 species. *Planta*. *149*(1):78-90. https://doi.org/10.1007/BF00386231

Fernandes, I., Marques, I., Paulo, O. S., Batista, D., Partelli, F. L., Lidon, F. C., DaMatta, F. M., Ramalho, J. C., Ribeiro-Barros, A. I. (2021). Understanding the Impact of Drought in Coffea Genotypes: Transcriptomic Analysis Supports a Common High Resilience to Moderate Water Deficit but a Genotype Dependent Sensitivity to Severe Water Deficit. *Agronomy* *11*(2255). https://doi.org/10.3390/agronomy11112255

Friedrich, T., Oberkofler, V., Trindade, I., Altmann, S., Brzezink, K., Lämke, J., Gorka, M., Kappel, C., Sokolowska, E., Skirycz, A., Graf, A, Bäurle, I. (2021). Heteromeric HSFA2/HSFA3 complexes drive transcriptional memory after heat stress in Arabidopsis. *Nature Communications*, *12*(1), 3426. https://doi.org/10.1038/s41467-021-23786-6

Gentile, A., Dias, L.I., Mattos, R.S., Ferreira, T.H., Menossi, M. (2015) MicroRNAs and drought responses in sugarcane. *Front Plant Sci.* 23; 6:58. https://doi.org/10.3389/fpls.2015.00058

Genty, B., Briantais, J.M., Baker, N.R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA)*, *990*(1) 87-92. https://doi.org/10.1016/S0304-4165(89)80016-9

Giannopolitis, C.N., Ries, S.K. (1977) Superoxide dismutases: I. Occurrence in higher plants. *Plant Physiol.* 59(2):309-14. https://doi.org/ 10.1104/pp.59.2.309

Gonzalez, F. G., Capella, M., Ribichich, K. F., Curin, F., Giacomelli, J. I., Ayala, F., Watson, G., Otegui, M,E., Chan, R.L. (2019). Field-grown transgenic wheat expressing the sunflower gene *HaHB4* significantly outyields the wild type. *J. Exp. Bot*. 70 (5), 1669–1681. https://doi.org/10.1093/jxb/erz037

Griffiths-Jones, S. (2004). The microRNA registry. *Nucleic Acids Research*, *32*, 109–111. https://doi.org/10.1093/nar/gkh023

Guedes, F. A. F., Nobres, P., Ferreira, D. C. R., Menezes-Silva, P. E., Ribeiro-Alves, M., Correa, R. L., DaMatta, F. M., Alves-Ferreira, M. (2018). Transcriptional memory contributes to drought tolerance in coffee (*Coffea canephora*) plants *Environmental and Experimental Botany*, 147:220-233. https://doi.org/10.1016/j.envexpbot.2017.12.004

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O.

(2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Havir, E.A, McHale, N.A. (1987) Biochemical and developmental characterization of multiple forms of catalase in tobacco leaves. *Plant Physiol.* 84(2):450-5. doi: 10.1104/pp.84.2.450.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, *8*(2). https://doi.org/10.1186/gb-2007-8-2-r19

Henriksson, E., Olsson, A. S. B., Johannesson, H., Johansson, H., Hanson, J., Engström, P., Sö, E. (2005). Homeodomain Leucine Zipper Class I Genes in Arabidopsis. Expression Patterns and Phylogenetic Relationships. *Plant Physiology*, *139*(1), 509–518. https://doi.org/10.1104/pp.105.063461

Hofer, J., Turner, L., Moreau, C., Ambrose, M., Isaac, P., Butcher, S., Weller, J., Dupin, A., Dalmais, M., Signor, C. Le, Bendahmane, A., Ellis, N. (2009). Tendril-less regulates tendril formation in pea leaves. *Plant Cell*, *21*(2), 420–428. https://doi.org/1105/tpc.108.064071

Hou, J., Sun, Y., Wang, L., Jiang, Y., Chen, N., Tong, S. (2021). Genome-Wide Analysis of the Homeobox Gene Family and Identification of Drought-Responsive Members in *Populus trichocarpa*. *Plants* (Basel). 2021 Oct 25;10(11):2284.

https://doi.org/10.3390/plants10112284.

Hu, Y. X., Wang, Y. H., Liu, X. F., Li, J. Y. (2004). *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. *Cell Research, 14*(1), 8–15. https://doi.org/10.1038/sj.cr.7290197

Hu, Y., Yu, D. (2014). BRASSINOSTEROID INSENSITIVE2 interacts with ABSCISIC ACID INSENSITIVE5 to mediate the antagonism of brassinosteroids to abscisic acid during seed germination in Arabidopsis. *Plant Cell*. 26(11):4394-408. https://doi.org/10.1105/tpc.114.130849.

Hur, Y.-S., Um, J.-H., Kim, S., Kim, K., Park, H.-J., Lim, J.-S., Kim, W.-Y., Jun, S.E., Yoon, E.K., Lim, J., Ohme-Takagi, M., Kim, D., Park, J., Kim, G.-T. Cheon, C.-I. (2015), *Arabidopsis thaliana homeobox 12* (*ATHB12*), a homeodomain-leucine zipper protein, regulates leaf growth by promoting cell expansion and endoreduplication. *New Phytol, 205*: 316-328. https://doi.org/10.1111/nph.12998

Jahns, P., Holzwarth, A.R. (2012). The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *Biochim Biophys Acta*. *1817*(1):182-93. https://doi.org/10.1016/j.bbabio.2011.04.012.

Jiao, P., Jiang, Z., Wei, X., Liu, S., Qu, J. Guan, S., Ma, Y. (2022). Overexpression of the homeobox-leucine zipper protein *ATHB-6* improves the drought tolerance of maize (*Zea mays* L.), *Plant Science*, *316*, 111159. https://doi.org/10.1016/j.plantsci.2021.111159.

Johannesson, H., Wang, Y., Engström, P. (2001). DNA-binding and dimerization

preferences of Arabidopsis homeodomain-leucine zipper transcription factors *in vitro*. *Plant Molecular Biology*, *45*(1), 63–73. https://doi.org/10.1023/A:1006423324025

Kant, S., Peng, M., Rothstein, S. J. (2011). Genetic regulation by *NLA* and *microRNA827* for maintaining nitrate-dependent phosphate homeostasis in Arabidopsis. *PLoS Genetics*, *7*(3). https://doi.org/10.1371/journal.pgen.1002021

Karimi, M., Inzé, D., Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*, *7*(5), 193–195. https://doi.org/10.1016/s1360-1385(02)02251-3

Khraiwesh, B., Zhu, J.-K., Zhuc, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta.*, *1819*(2), 137– https://doi.org/148. 10.1016/j.bbagrm.2011.05.001

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J., Harter, K. (2007), The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal*, 50: 347-363. https://doi.org/10.1111/j.1365-313X.2007.03052.x

Knauer, S., Holt, A. L., Rubio-Somoza, I., Tucker, E. J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M. C., Tucker, M. R., Laux, T. (2013). A Protodermal *miR394* Signal Defines a Region of Stem Cell Competence in the *Arabidopsis*

Shoot Meristem. *Developmental Cell*, *24*(2), 125–132. https://doi.org/10.1016/J.DEVCEL.2012.12.009

Knight, H., Knight, M. R. (2001). Abiotic stress signalling pathways: Specificity and cross-talk. *Trends in Plant Science*, *6*(6), 262–267. https://doi.org/10.1016/S1360-1385(01)01946-X

Kotchoni, S. O., Kuhns, C., Ditzer, A., Kirch, H. H., Bartels, D. (2006). Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant, Cell and Environment*, *29*(6), 1033–1048. https://doi.org/10.1111/j.1365-3040.2005.01458.x

Koutouleas, A., Sarzynski, T., Bordeaux, M., Bosselmann, A.S., Campa, C., Etienne, H., Turreira-García, N., Rigal, C., Vaast, P., Ramalho, J.C., Marraccini, P., Ræbild, A. (2022) Shaded-coffee: a nature-based strategy for coffee production under climate change? A review. *Front. Sustain. Food Syst., 6*, 877476. https://doi.org/10.3389/fsufs.2022.877476

Kulcheski, F. R., de Oliveira, L. F. V., Molina, L. G., Almerão, M. P., Rodrigues, F. A., Marcolino, J., Barbosa, J. F., Stolf-Moreira, R., Nepomuceno, A. L., Marcelino-Guimarães, F. C., Abdelnoor, R. V., Nascimento, L. C., Carazzolle, M. F., Pereira, G. A. G., Margis, R. (2011). Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *BMC Genomics*, *12*(307), 1–17. https://doi.org/10.1186/1471-2164-12-307

Langmead, B., Trapnell, C., Pop, M., Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3). https://doi.org/1186/gb-2009-10-3-r25

Lefort, V., Longueville, J. E., Gascuel, O. (2017). SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution*, *34*(9), 2422–2424. https://doi.org/10.1093/molbev/msx149

Lee YH, Chun JY. (1998). A new homeodomain-leucine zipper gene from *Arabidopsis thaliana* induced by water stress and abscisic acid treatment. *Plant Molecular Biology 37:* 377– 384. https://doi.org/10.1023/a:1006084305012

Letunic, I., Bork, P. (2019). Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), 256–259. https://doi.org/10.1093/nar/gkz239

Letunic, I., Khedkar, S., Bork, P. (2021). SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Research*, *49*(D1), D458–D460. https://doi.org/10.1093/nar/gkaa937

Li, T., Li, H., Zhang, Y.-X., Liu, J.-Y. (2011). Identification and analysis of seven H_2O_2-responsive miRNAs and 32 new miRNAs in the seedlings of rice (*Oryza sativa* L. *ssp. indica*). *Nucleic Acids Research*, *39*(7), 2821–2833. https://doi.org/10.1093/nar/gkq1047

Li, Y., Zhang, Q. Q., Zhang, J., Wu, L., Qi, Y., Zhou, J. M. (2010). Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant

innate immunity. *Plant Physiology*, *152*(4), 2222–2231. https://doi.org/10.1104/pp.109.151803

Li, Y., Yang, Z., Zhang, Y., Guo, J., Liu, L., Wang, C., Wang, B., Han, G. (2022). The roles of HD-ZIP proteins in plant abiotic stress tolerance. *Front Plant Sci. 12*;13:1027071. https://doi.org/10.3389/fpls.2022.1027071.

Liang, G., Yang, F., Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant Journal*, *62*(6), 1046–1057. https://doi.org/10.1111/j.1365-313X.2010.04216.x

Liu, D., Song, Y., Chen, Z., Yu, D. (2009). Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. *Physiologia Plantarum*, *136*(2), 223–236. https://doi.org/10.1111/j.1399-3054.2009.01229.x

Liu, H.-H., Tian, X., Li, Y.-J., Wu, C.-A., Zheng, C.-C. (2008). Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA*, *14*(5), 836–843. https://doi.org/10.1261/rna.895308

Liu, L., Hu, X., Song, J., Zong, X., Li, D., Li, D. (2009). Over-expression of a *Zea mays* L. protein phosphatase 2C gene (*ZmPP2C*) in *Arabidopsis thaliana* decreases tolerance to salt and drought. *Journal of Plant Physiology*, *166*(5), 531–542. https://doi.org/10.1016/J.JPLPH.2008.07.008

Llave, C., Xie, Z., Kasschau, K. D., Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA Targets Directed by a Class of Arabidopsis miRNA.

Science, *297*(5589), 2053–2056. https://doi.org/10.1126/science.1076311

Ma, Q., Xia, Z., Cai, Z., Li, L., Cheng, Y., Liu, J., Nian, H. (2019). GmWRKY16 Enhances Drought and Salt Tolerance Through an ABA-Mediated Pathway in *Arabidopsis thaliana*. *Front. Plant Sci.* 9:1979. https://doi.org/10.3389/fpls.2018.01979

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Majeed, Y.; Zhu, X.; Zhang, N.; ul-Ain, N.; Raza, A.; Haider, F. U.; Si, H. (2023). Harnessing the role of mitogen-activated protein kinases against abiotic stress in plants. *Frontiers in Plant Science*, *14*(932923). https://doi.org/10.3389/fpls.2023.932923

Manavella, P. A., Arce, A. L., Dezar, C. A., Bitton, F., Renou, J. P., Crespi, M., Chan, R. L. (2006). Cross-talk between ethylene and drought signalling pathways is mediated by the sunflower Hahb-4 transcription factor. *Plant Journal*, *48*(1), 125–137. https://doi.org/10.1111/j.1365-313X.2006.02865.x

Manavella, P. A., Dezar, C. A., Bonaventure, G., Baldwin, I. T., Chan, R. L. (2008). HAHB4, a sunflower HD-Zip protein, integrates signals from the jasmonic acid and ethylene pathways during wounding and biotic stress responses. *Plant Journal*, *56*(3), 376–388. https://doi.org/10.1111/j.1365-313X.2008.03604.x

Manghwar, H., Li, J. (2022). Endoplasmic reticulum stress and unfolded protein response signaling in plants. *International Journal of Molecular Sciences*, 23(828). https://doi.org/10.3390/ijms23020828

Marques, I., Rodrigues, A.P., Gouveia, D., Lidon, F.C., Martins, S., Semedo, M.C., Gaillard, J.C., Pais, I.P., Semedo, J.N., Scotti-Campos, P., Reboredo, F.H., Partelli, F.L., DaMatta, F.M., Armengaud, J., Ribeiro-Barros, A.I., Ramalho, J.C. (2022). High-resolution shotgun proteomics reveals that increased air [CO_2] amplifies the acclimation response of coffea species to drought regarding antioxidative, energy, sugar, and lipid dynamics. *J Plant Physiol.* 276:153788. https://doi.org/10.1016/j.jplph.2022.153788

Marraccini, P., Freire, L. P., Alves, G. S. C., Vieira, N. G., Vinecky, F., Elbelt, S., Ramos, H. J. O., Montagnon, C., Vieira, L. G. E., Leroy, T., Pot, D., Silva, V. A., Rodrigues, G. C., Andrade, A. C. (2011). *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. *BMC Plant Biology*, *11*(85), 1–23. https://doi.org/10.1186/1471-2229-11-85

Marraccini, P., Vinecky, F., Alves, G. S. C., Ramos, H. J. O., Elbelt, S., Vieira, N. G., Carneiro, F. A., Sujii, P. S., Alekcevetch, J. C., Silva, V. A., DaMatta, F. M., Ferrão, M. A. G., Leroy, T., Pot, D., Vieira, L. G. E., Silva, F. R. da, Andrade, A. C. (2012). Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. *Journal of Experimental Botany*, *63*(11), 4191–4212. https://doi.org/10.1093/jxb/err313

Maxwell K, Johnson GN. (2000) Chlorophyll fluorescence--a practical guide. *J Exp Bot.* 51(345):659-68. https://doi.org/10.1093/jxb/51.345.659

Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., Cao, X., Carrington, J. C., Chen, X., Green, P. J., Griffiths-Jones, S., Jacobsen, S. E., Mallory, A. C., Martienssen, R. A., Poethig, R. S., Qi, Y., Vaucheret, H., Voinnet, O., Watanabe, Y., Zhui, J. K. (2008). Criteria for annotation of plant microRNAs. *Plant Cell*, *20*(12), 3186–3190. https://doi.org/10.1105/tpc.108.064311

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., Marshall, D. (2010). Tablet-next generation sequence assembly visualization. *Bioinformatics*, *26*(3), 401–402. https://doi.org/10.1093/bioinformatics/btp666

Mishra, M. (2022) - Current progress in tissue culture and transgenic research in coffee (*Coffea spp*.) https://doi.org/10.1201/9781003043133-22

Mishra, M. K., Slater, A. (2012). Recent Advances in the Genetic Transformation of Coffee. *Biotechnology Research International*, *2012*, 1–17. https://doi.org/10.1155/2012/580857

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/nar/gkaa913

Moxon, S., Schwach, F., Dalmay, T., MacLean, D., Studholme, D. J., Moulton, V. (2008). A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics*, *24*(19), 2252–2253. doi:10.1093/bioinformatics/btn428

Mukherjee, K., Brocchieri, L., Bürglin, T. R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Molecular Biology and Evolution*, *26*(12), 2775–2794. https://doi.org/10.1093/molbev/msp201

Murashige, T., Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, *15*(3), 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Nakai, K., Horton, P., Nakai, K., Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Computer Corner*, *24*(1), 34–35. https://doi.org/10.1016/s0968-0004(98)01336-x

Nakano Y, Asada K (1981) Hydrogen Peroxide is Scavenged by Ascorbate-specific Peroxidase in Spinach Chloroplasts. *Plant and Cell Physiology*, *22*(5), 867–880. https://doi.org/10.1093/oxfordjournals.pcp.a076232

Nguyen, T. T.Q., Trinh, L. T. H., Pham, H. B. V., Le, T. V., Phung, T. K. H., Lee, S. H. Cheong, J. J. (2020). Evaluation of proline, soluble sugar and ABA content in soybean *Glycine max* (L.) under drought stress memory[J]. *AIMS Bioengineering*, *7*(3): 114-123. https://doi.org/10.3934/bioeng.2020011

Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K. Shigeoka, S. (2006), Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *The Plant Journal*, *48*: 535-547. https://doi.org/10.1111/j.1365-313X.2006.02889.x

Ogawa, D., Yamaguchi, K., Nishiuchi, T. (2007). High-level overexpression of the *Arabidopsis HsfA2* gene confers not only increased themotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *Journal of Experimental Botany*, *58*(12), 3373–3383. https://doi.org/10.1093/jxb/erm184

Olsson, A. S. B., Engström, P., Söderman, E. (2004). The homeobox genes *ATHB12* and *ATHB7* encode potential regulators of growth in response to water deficit in *Arabidopsis*. *Plant Molecular Biology*, *55*(5), 663–677. https://doi.org/10.1007/s11103-004-1581-4

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*, *12*(4), 357–358. https://doi.org/10.1093/bioinformatics/12.4.357

Paquette, J., Tokuyasu, T. (2010). EGAN: Exploratory gene association networks. *Bioinformatics*, *26*(2), 285–286. https://doi.org/10.1093/bioinformatics/btp656

Pham, T. T. N., Nguyen, H. Q., Nguyen, T. N. L., Dao, X. T., Sy, D. T., Van Son, L., Hoang Mau, C. (2020). Overexpression of the "*GmDREB2*" gene increases proline accumulation and tolerance to drought stress in soybean plants. *Australian Journal of Crop Science*, *14*(3), 495–503.

https://doi.org/10.3316/informit.123331655384478

Phillips, K., Ludidi, N. (2017). Drought and exogenous abscisic acid alter hydrogen peroxide accumulation and differentially regulate the expression of two maize RD22-like genes. *Sci Rep* 7, 882. https://doi.org/10.1038/s41598-017-08976-x

Pinheiro, H. A., DaMatta, F. M., Chaves, A. R. M., Loureiro, M. E., Ducatti, C. (2005). Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. *Annals of Botany*, *96*(1), 101–108. https://doi.org/10.1093/aob/mci154

Porra, R.J., Thompson, W.A., Kriedemann, P.E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *975*(3),384-394. https://doi.org/10.1016/S0005-2728(89)80347-0

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., Bateman, A., Finn, R. D. (2012). The Pfam protein families database. *Nucleic Acids Research*, *40*(D1), 290–301. https://doi.org/10.1093/nar/gkr1065

Ré, D.A., Capella, M., Bonaventure, G., Chan, R. L. (2014). Arabidopsis *AtHB7* and *AtHB12* evolved divergently to fine tune processes associated with growth and

responses to water stress. *BMC Plant Biol* 14, 150. https://doi.org/10.1186/1471-2229-14-150

Rech, E.L., Vianna, G.R., Aragão, F.J. (2008). High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nat Protoc.* *3*(3):410-8. https://doi.org/10.1038/nprot.2008.9

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., Bartel, D. P. (2002). Prediction of Plant MicroRNA Targets. *Cell*, *110*(4), 513–520. https://doi.org/10.1016/S0092-8674(02)00863-2

Ribichich, K.F., Chiozza, M., Ávalos-Britez, S., Cabello, J,V., Arce, A.L., Watson, G., Arias, C., Portapila, M., Trucco, F., Otegui, M.E., Chan, R.L. (2020). Successful field performance in warm and dry environments of soybean expressing the sunflower transcription factor HB4. *J Exp Bot*. *71*(10):3142-3156. https://doi.org/10.1093/jxb/eraa064

Robinson, M. D., Smyth, G. K. (2007). Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*, *23*(21), 2881–2887. https://doi.org/10.1093/bioinformatics/btm453

Robinson, M. D., Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, *9*(2), 321–332. https://doi.org/10.1093/biostatistics/kxm030

Robinson, M. D., McCarthy, D. J., Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.

Bioinformatics, *26*(1), 139-140. https://doi.org/10.1093/bioinformatics/btp616

Ruberti, I., Sessa, G., Lucchetti, S., Morelli, G. (1991). A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. *EMBO Journal*, *10*(7), 1787–1791. https://doi.org/10.1002/j.1460-2075.1991.tb07703.x

Sakuma, S., Pourkheirandish, M., Matsumoto, T., Koba, T., Komatsuda, T. (2010). Duplication of a well-conserved homeodomain-leucine zipper transcription factor gene in barley generates a copy with more specific functions. *Functional and Integrative Genomics*, *10*(1), 123–133. https://doi.org/10.1007/s10142-009-0134-y

Santos, M. O., Coelho, L. S., Carvalho, G. R., Botelho, C. E., Torres, L. F., Vilela, D. J. M., Andrade, A. C., Silva, V. A. (2021). Photochemical efficiency correlated with candidate gene expression promote coffee drought tolerance. *Sci Rep* 1;11(1):7436. https://doi.org/10.1038/s41598-021-86689-y

Sato, Y., Yokoya, S. (2008). Enhanced tolerance to drought stress in transgenic rice plants overexpressing a small heat-shock protein, sHSP17.7. *Plant* Cell Rep. 27, 329– 334. https://doi.org/10.1007/s00299-007-0470-0

Scalabrin, S., Toniutti, L., Di Gaspero, G., Scaglione, D., Magris, G., Vidotto, M., Pinosio, S., Cattonaro, F., Magni, F., Jurman, I., Cerutti, M., Suggi Liverani, F., Navarini, L., Del Terra, L., Pellegrino, G., Ruosi, M. R., Vitulo, N., Valle, G., Pallavicini, A., Graziosi, G., Klein, P. E., Bentley, N., Murray, S., Solano, W, Al-

Hakimi, A., Schilling, T., Montagnon, C., Morgante, M., Bertrand, B. (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Scientific Reports, 10*(1), 1–13. https://doi.org/10.1038/s41598-020-61216-7

Scharf, K.D., Berberich, T., Ebersberger, I., Nover, L. (2012). The plant heat stress transcription factor (Hsf) family: structure, function and evolution. *Biochim Biophys Acta*. 1819(2):104-19. https://doi.org/10.1016/j.bbagrm.2011.10.002. Erratum in: *Biochim Biophys Acta.* 2017 Dec 13

Schena, M., Davis, R.W. (1992) HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. *Proc. Natl. Acad. Sci. USA* 89, 3894–3898. https://doi.org/10.1073/pnas.89.9.3894.

Schramm, F., Ganguli, A., Kiehlmann, E., Walch, D., von Koskull-Döring, P. (2006). The Heat Stress Transcription Factor HsfA2 Serves as a Regulatory Amplifier of a Subset of Genes in the Heat Stress Response in Arabidopsis . *Plant Mol Biol 60*, 759–772 https://doi.org/10.1007/s11103-005-5750-x

Sessa, G., Morelli, G., Ruberti, I. (1993). The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO Journal, 12*(9), 3507–3517. https://doi.org/10.1002/j.1460-2075.1993.tb06025.x

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T. (2003). Cytoscape: a software environment for

integrated models of biomolecular interaction networks. *Genome Resource*, 13(11), 2498-504. https://doi.org/10.1101/gr.1239303

Simkin, A. J., Moreau, H., Kuntz, M., Pagny, G., Lin, C., Tanksley, S., McCarthy, J. (2008). An investigation of carotenoid biosynthesis in *Coffea canephora* and *Coffea arabica*. *Journal of Plant Physiology*, *165*(10), 1087–1106. https://doi.org/10.1016/J.JPLPH.2007.06.016

Son, O., Hur, Y.S., Kim, Y.K., Lee, H.J., Kim, S., Kim, M.R., Nam, K.H., Lee, M.S., Kim, B.Y., Park, J., Lee, S.C., Hanada, A., Yamaguchi, S., Lee, I.J., Kim, S.K., Yun, D.J., Söderman, E., Cheon, C.I. (2010). ATHB12, an ABA-inducible homeodomain-leucine zipper (HD-Zip) protein of Arabidopsis, negatively regulates the growth of the inflorescence stem by decreasing the expression of a gibberellin 20-oxidase gene. *Plant Cell Physiology* 51: 1537– 1547. https://doi.org/10.1093/pcp/pcq108

Stief, A., Altmann, S., Hoffmann, K., Pant, B. D., Scheible, W. R., Bäurle, I. (2014). Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell*, *26*(4), 1792–1807. https://doi.org/10.1105/tpc.114.123851

Sun, W., Bernard, C., Van Cotte, B. De, Van Montagu, M., Verbruggen, N. (2001). *At-HSP17.6A*, encoding a small heat-shock protein in *Arabidopsis*, can enhance osmotolerance upon overexpression. *Plant Journal*, *27*(5), 407–415. https://doi.org/10.1046/j.1365-313X.2001.01107.x

Szabados, L., Savouré A. (2010). Proline: a multifunctional amino acid. *Trends Plant Sci.* 15(2):89-97. https://doi.org/10.1016/j.tplants.2009.11.009

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., Mering, C. V. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607-D613. https://doi.org/10.1093/nar/gky1131

Tang, Y., Wang, J., Bao, X., Liang, M., Lou,H., Zhao, J., Sun, M., Liang, J., Jin, L., Li, G.. Qiu, Y., Liu, K. (2019). Genome-wide identification and expression profile of HD-ZIP genes in physic nut and functional analysis of the JcHDZ16 gene in transgenic rice. *BMC Plant Biol* 19, 298. https://doi.org/10.1186/s12870-019-1920-x

Tezuka, K., Taji, T., Hayashi, T., Sakata, Y. (2013) A novel *abi5* allele reveals the importance of the conserved Ala in the C3 domain for regulation of downstream genes and salt tolerance during germination in *Arabidopsis*, *Plant Signaling & Behavior*, 8:3. https://doi.org/10.4161/psb.23455

Thonglim, A., Bortolami, G., Delzon, S., Larter, M., Offringa, R., Keurentjes, J. J. B., Smets, E., Balazadeh, S., Lens, F. (2023). Drought response in Arabidopsis displays synergistic coordination between stems and leaves. *Journal of Exp. Botany*, 74, 1004-1021. https://doi.org/10.1093/jxb/erac446

Treuter, E., Nover, L., Ohme, K., Scharf, K-D (1993), Promoter specificity and deletion analysis of three heat stress transcription factors of tomato. *Mol Gen Genet 240*:113–125. https://doi.org/10.1007/BF00276890

Udvardi, M.K., Kakar, K., Wandrey, M., Montanari, O., Murray, J., Andriankaja, A., Zhang, J.Y., Benedito, V., Hofer, J.M., Chueng, F., Town, C.D. (2007) Legume transcription factors: global regulators of plant development and response to the environment. *Plant Physiol, 144*(2):538-49. https://doi.org/10.1104/pp.107.09806

Valdés, A. E., Övernäs, E., Johansson, H., Rada-Iglesias, A., Engström, P. (2012). The homeodomain-leucine zipper (HD-Zip) class I transcription factors ATHB7 and ATHB12 modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities. *Plant Molecular Biology*, *80*(4–5), 405–418. https://doi.org/10.1007/s11103-012-9956-4

Verslues, P. E., Bray, E. A. (2004). *LWR1* and *LWR2* are required for osmoregulation and osmotic adjustment in Arabidopsis. *Plant Physiology*, *136*(1), 2831–2842. https://doi.org/10.1104/pp.104.045856

Vidal, R. O., Mondego, J. M. C., Pot, D., Ambrósio, A. B., Andrade, A. C., Pereira, L. F. P., Colombo, C. A., Vieira, L. G. E., Carazzolle, M. F., Pereira, G. A. G. (2010). A high-throughput data mining of single nucleotide polymorphisms in Coffea species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. *Plant Physiology*, *154*(3), 1053–1066. https://doi.org/10.1104/pp.110.162438

Vieira, L. G. E., Andrade, A. C., Colombo, C. A., Araújo, A. H., Metha, Â., Oliveira, A. C., Labate, C. A., Marino, C. L., Monteiro-vitorello, D. B., Monte, D. D. C., Giglioti, É., Kimura, E. T., Kuramae, E. E., Gertrudes, E., Lemos, M., Rita, E., Almeida, P. De, Dorry, H. F. A., Carrer, H., … Oliveira, M. C. (2006). Brazilian coffee genome project an EST-based genomic resource. *Brazilian Journal of Plant Physiology*, *18*(1), 95–108. https://doi.org/10.1590/S1677-04202006000100008

Vierling, E. (1991). The Roles of Heat Shock Proteins in Plants. *Annu. Rev. Plant Biol.*, 42, 579–620. https://doi.org/10.1146/annurev.pp.42.060191.003051

Waltz, E. (2015). First stress-tolerant soybean gets go-ahead in Argentina. *Nat. Biotechnol.* 33, 682. https://doi.org/10.1038/nbt0715-682

Wang, Y. J., Li, Y. D., Luo, G. Z., Tian, A. G., Wang, H. W., Zhang, J. S., Chen, S. Y. (2005). Cloning and characterization of an HDZip I gene *GmHZ1* from soybean. *Planta*, *221*(6), 831–843. https://doi.org/10.1007/s00425-005-1496-6

Wei, M., Liu, A., Zhang, Y., Zhou, Y., Li, D., Dossa, K., Zhou, R,, Zhang, X., You, J. (2019). Genome-wide characterization and expression analysis of the HD-Zip gene family in response to drought and salinity stresses in sesame. *BMC Genomics* 20, 748. https://doi.org/10.1186/s12864-019-6091-5

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York.* Retrieved from https://ggplot2.tidyverse.org

Willige, B. C., Kutzer, M., Tebartz, F., Bartels, D. (2009). Subcellular localization and enzymatic properties of differentially expressed transketolase genes isolated

from the desiccation tolerant resurrection plant *Craterostigma plantagineum*. *Planta*, *229*(3), 659–666. https://doi.org/10.1007/s00425-008-0863-5

Winck, J.E.M., Sarmento, L.F.V., Zanon, A.J., Librelon, S.S., Garcia, A. Streck, N.A. (2022). Growth and transpiration of soybean genotypes with HaHB4® transcription factor for drought tolerance. *Physiologia Plantarum*, 174(1), e13557. https://doi.org/10.1111/ppl.13557

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., Provart, N.J. (2007) An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLOS ONE* 2(8): e718. https://doi.org/10.1371/journal.pone.0000718

Woo, H. R., Kim, J. H., Kim, J., Kim, J., Lee, U., Song, I. J., Kim, J. H., Lee, H. Y., Nam, H. G., Lim, P. O. (2010). The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*. *Journal of Experimental Botany*, *61*(14), 3947–3957. https://doi.org/10.1093/jxb/erq206

Wu, G., Park, M. Y., Conway, S. R., Wang, J.-W., Weigel, D., Poething, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell*, *138*(4), 750–759. https://doi.org/10.1016/j.cell.2009.06.031

Xia, H., Yu, S., Kong, D., Xiong, J., Ma, X., Chen, L., Luo, L. (2020). Temporal responses of conserved miRNAs to drought and their associations with drought tolerance and productivity in rice. *BMC Genomics* 21, 232.

https://doi.org/10.1186/s12864-020-6646-5

Xie, Z., Kasschau, K. D. Carrington, J. C. (2003). Negative feedback regulation of *Dicer-Like1* in *Arabidopsis* by microRNA-guided mRNA degradation. *Current Biology*, *13*, 784–789. https://doi.org/10.1016/S0960-9822(03)00281-1

Xu, X., Lou, Y., Yang, K., Shan, X., Zhu, C., Gao, Z. (2019). Identification of Homeobox Genes Associated with Lignification and Their Expression Patterns in Bamboo Shoots. *Biomolecules*, *9*(12):862. https://doi.org/10.3390/biom9120862

Yang, J. H., Han, S. J., Yoon, E. K., Lee, W. S. (2006). "Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells." *Nucleic Acids Research*, *34*(6), 1892–1899. https://doi.org/10.1093/nar/gkl118

Yokotani, N., Ichikawa, T., Kondou, Y., Matsui, M., Hirochika, H., Iwabuchi, M., Oda, K. (2008). Expression of rice heat stress transcription factor OsHsfA2e enhances tolerance to environmental stresses in transgenic Arabidopsis. *Planta*, *227*(5), 957–967. https://doi.org/10.1007/s00425-007-0670-4

Yoshida, T., Mogami, J., Yamaguchi-Shinozaki, K. (2014). ABA-dependent and ABA-independent signalling in response to osmotic stress in plants. *Curr Opin Plant Biol*, 21:133-139. https://doi.org/10.1016/j.pbi.2014.07.009

Zhang, H., Xu, H., Xu, M., Yan, X., Wang, Z., Yu, J., Lei, B., Cui, H. (2023). Transcription factors *NtHD9* and *NtHD12* control long glandular trichome formation via jasmonate signaling, *Plant Physiology*, 191(4), 2385–2399.

https://doi.org/10.1093/plphys/kiad003

Zhang, J., Wu, J., Guo, M., Aslam, M., Wang, Q., Ma, H., Li, S., Zhang, X., Cao, S. (2020). Genome-wide characterization and expression profiling of *Eucalyptus grandis* HD-Zip gene family in response to salt and temperature stress. *BMC Plant Biol* 20, 451. https://doi.org/10.1186/s12870-020-02677-w

Zhang, S., Haider, I., Kohlen, W., Jiang, L., Bouwmeester, H., Meijer, A. H., Schluepmann, H., Liu, C. M., Ouwerkerk, P. B. F. (2012). Function of the HD-Zip I gene *Oshox22* in ABA-mediated drought and salt tolerances in rice. *Plant Molecular Biology*, *80*(6), 571–585. https://doi.org/10.1007/s11103-012-9967-1

Zhang, Z., Wang, X. (2011) Overexpression of *GASA5* increases the sensitivity of *Arabidopsis* to heat stress, *Journal of Plant Physiology*, 168(17), 2093-2101. https://doi.org/10.1016/j.jplph.2011.06.010

Zhao, S., Fernald, R. D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology*, *12*(8), 1047–1064. https://doi.org/10.1089/cmb.2005.12.1047

Zhuang, L., Cao, W., Wang, J., Yu, J., Yang, Z., Huang, B. (2018). Characterization and Functional Analysis of *FaHsfC1b* from *Festuca arundinacea* Conferring Heat Tolerance in Arabidopsis. *Int J Mol Sci.* 11;19(9):2702. https://doi.org/10.3390/ijms19092702

Figures

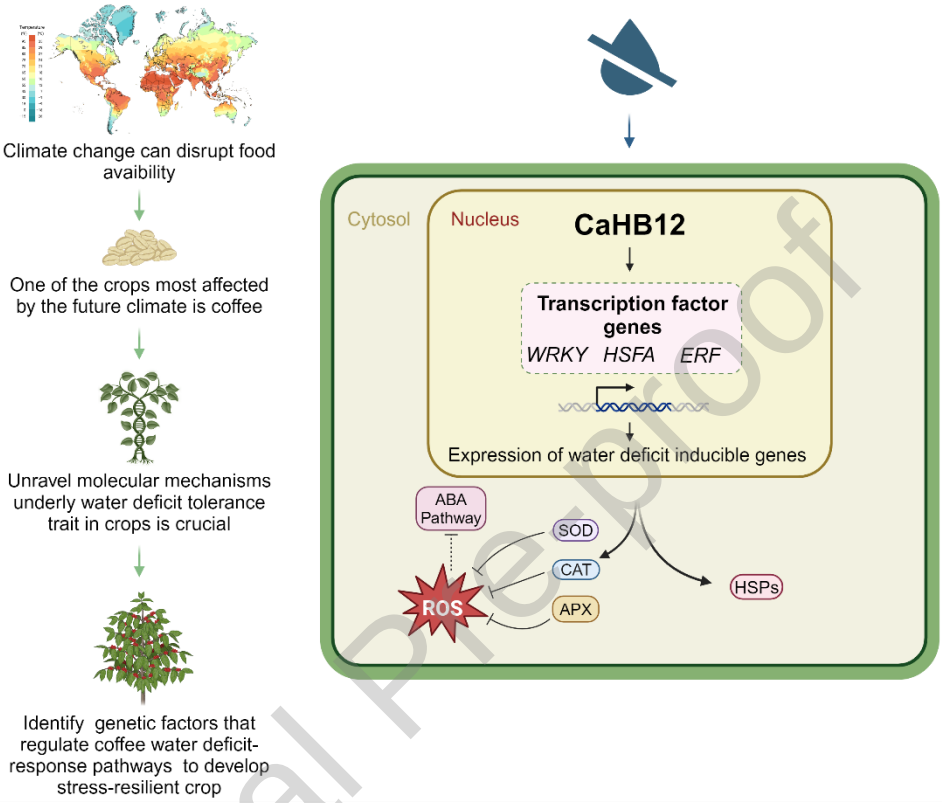


Figure 1 - The phylogenetic trees were generated using full-length amino acid sequences from three coffee species (*Coffea arabica, C. canephora,* and *C. eugenioides*) and five angiosperms (*Arabidopsis thaliana, Gossypium hirsutum, Glycine max, Oryza sativa,* and *Setaria viridis*), and statistical confidence was assessed by an aLRT branch. (A) Phylogenetic tree showing 14 classes labelled of homeobox gene family, as described based on Mukherjee et al., (2009). (B) HD-Zip I phylogenetic tree showing nine clades (α, β1, β2, δ, ε, γ, ϕ1, ϕ2, and ζ), as described in Agalou *et al.* (2008), highlighting the γ clade in blue. (C) Phylogenetic tree of the γ clade from HD-Zip I from eight plant species studied here and *HaHB4* from *Helianthus annuus* (sunflower) showing the gene code and gene name, as described for Arabidopsis and rice (Mukherjee et al., 2009), for cotton (Zhang et al., 2016), for soybean (Chen et al., 2014), and for *S. italica* (Chai et al., 2018). The

scale bars indicate the number of substitutions per site. For simplicity, the original code gene for each species (Supplemental Table 1) was substituted as follows: the underline present in *C. arabica* and *C. canephora* code was removed. Detailed information about the code and classification of genes is available in Supplemental Table 1.

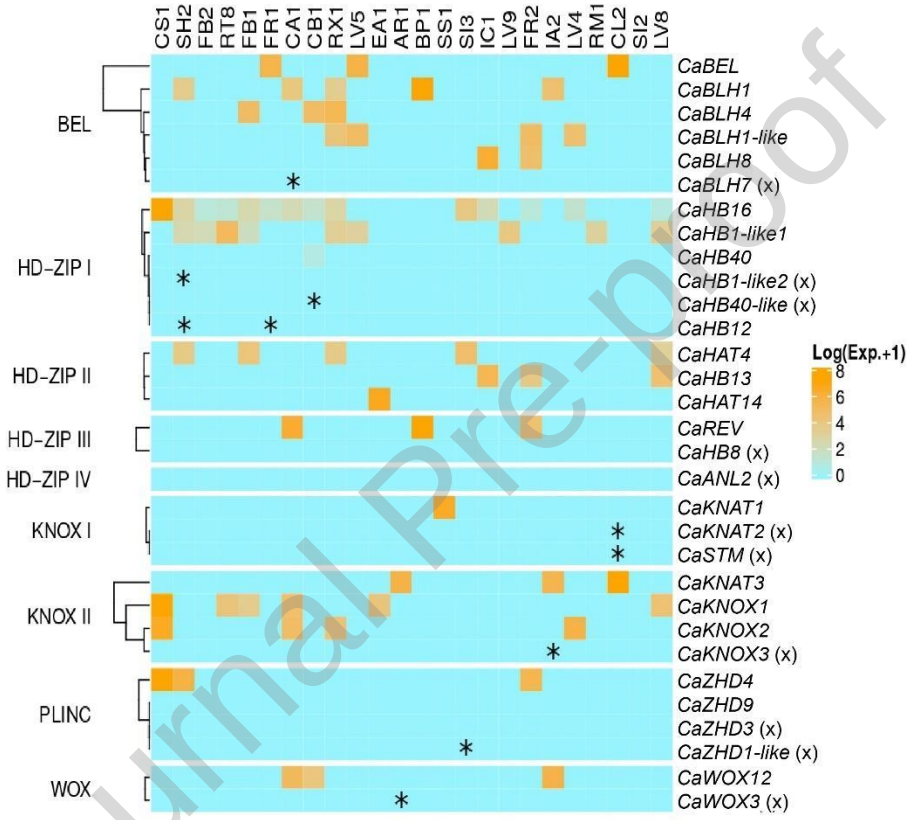


Figure 2 – *In silico* expression profile and phylogenetic relationships of coffee homeobox genes. Columns indicate the included libraries, and squares indicate the expression of each homeobox sequence. CS1: individual cells resuspended in mannose, NaCl and KCl solution; SH2: field-grown plants under hydric stress; FB2: flower buds (stages 1 and 2) - short; RT8: Roots and individual cells suspended in aluminium presence; FB1: flower buds (stages 1 and 2) - long; FR1: green unripe fruits- long ; CB1: individual cells treated with benzothiadozole and brassinosteroids; RX1: infected branches (*Xylella spp*.); CA1: Non-embryogenic callus; LV5: orthotropic branches shoot tips (young plants); EA1: embryogenic

callus; AR1: Leaves treated with arachidonic acid; BP1: root and leave's cells suspension; SS1: tissue of field-grown plants; SI3: germinating seeds; IC1: non-embryogenic leaves; LV9: plagiotropic branches leaves – short (mature plants) ; FR2: green unripe fruits- short; embryogenic leaves; LV4: orthotropic branches leaves – long (young plants) ; RM1: Leaves of plants infested with *Leucoptera spp*. and *Hemileia spp*.; CL2: hypocotyls treated with benzothiadozole; SI2: 'not described'; LV8: plagiotropic branches leaves – long (mature plants) The colour scale indicates the level of expression of each sequence according to the software package ComplexHeatmap, which was used for data visualisation. Singletons expressions are given by the library from which they were isolated as represented by an asterisk (*).

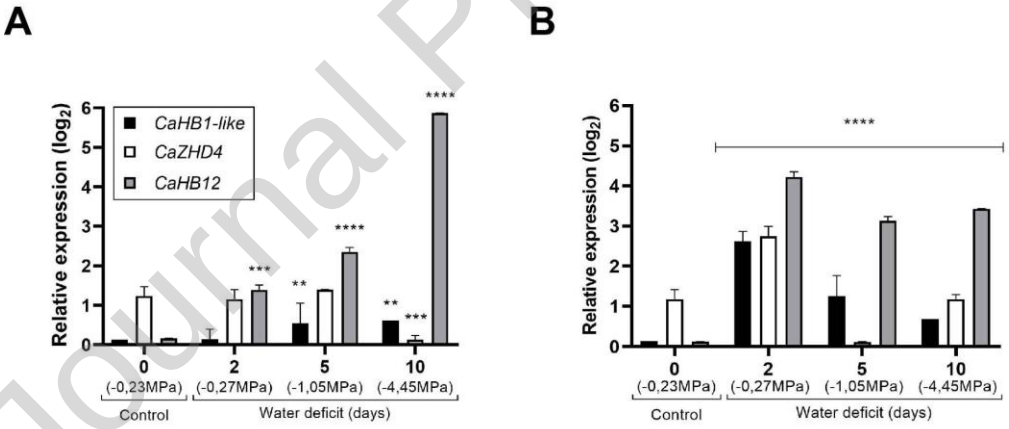


Figure 3 - Modulation of *CaHB12, CaHB1-like* 2 and *CaZHD1* expressions under water stress conditions in the leaves (A) and roots (B) of coffee plants. The Y-axis represents the relative levels of gene expression obtained by qPCR on a logarithmic scale. The X-axis represents the number of days of water suspension. The expression levels of *GAPDH, S24* and *UBQ10* were used for data normalization in leaves, while *UBI9, AP47* and S24 were used for data normalization in the roots. The asterisk denotes a statistically significant difference

between the expression levels of the genes in water stress conditions compared to control ($p<0,05$).

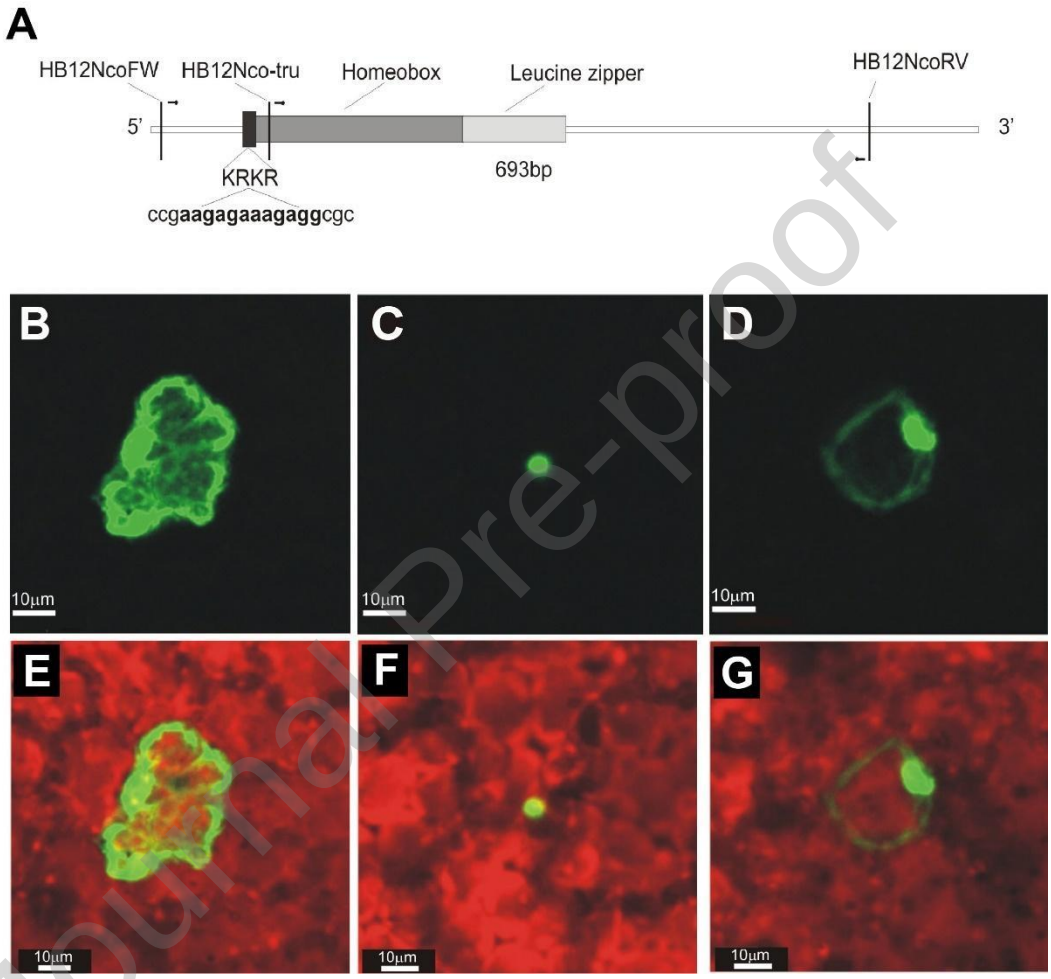


Figure 4 - Subcellular localization of CaHB12. (A) Graphic scheme of the *CaHB12* CDS. The nuclear location signal (NLS) KRKR is represented by the black box, and the homeodomain and leucine zipper are represented by grey boxes. Primer annealing sites used for amplification of the complete and truncated *CaHB12* CDS (lacking the NLS) are represented by black arrows. Visualization of CaHB12 fused to green fluorescent protein (GFP) in coffee (*C. arabica*) leaves is shown in the lower panels (B-G). **(B)** Fluorescent pattern resulting from *35S:GFP* expression. **(C)** and **(D)** Fluorescent pattern of expression from *35S:CaHB12:GFP*

with complete and truncated *CaHB12* CDS (lacking the NLS), respectively. B-D: green filter. E-G: Green plus red filters. Transient expression was observed 24 h after bombardment. Scale bars = 10 µm.

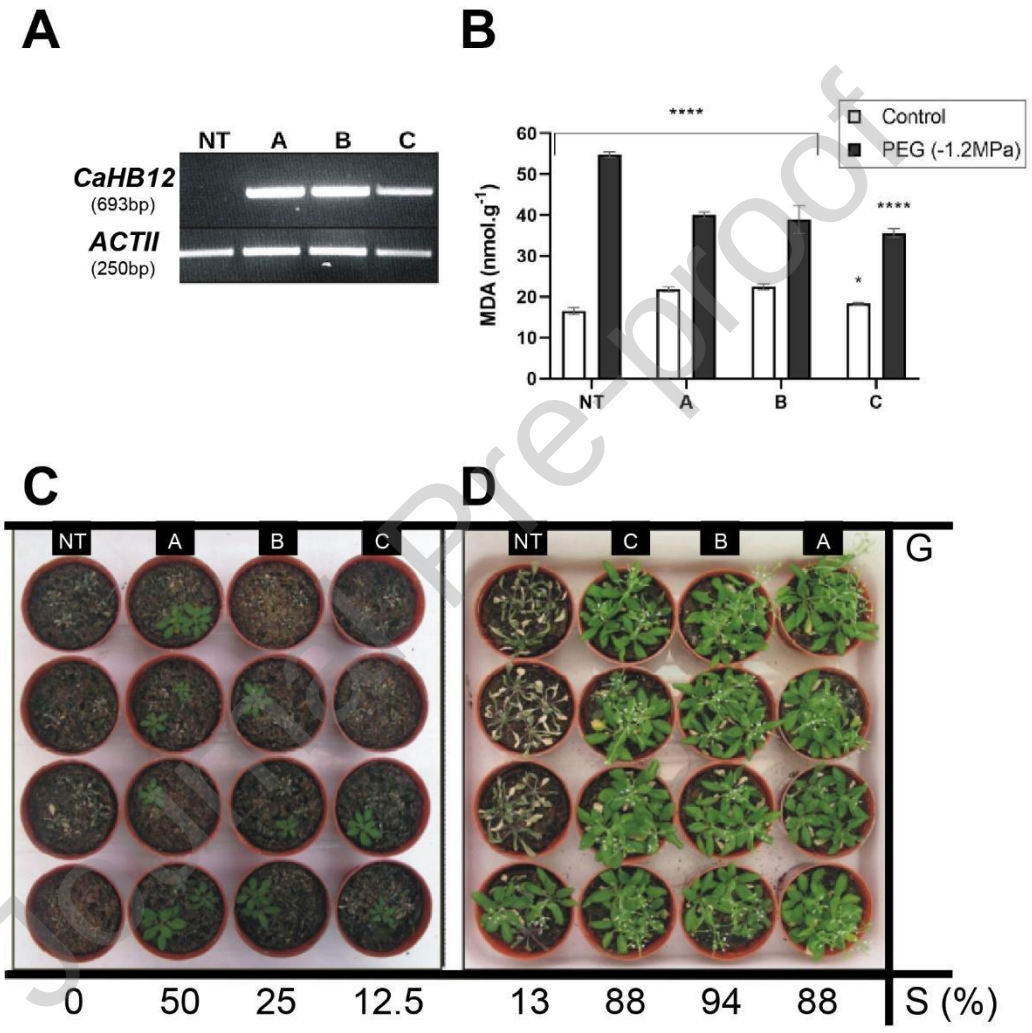


Figure 5 - Water deficit experiments. (A) PCR detection of the transgene *CaHB12* in arabidopsis events. **(B)** MDA levels after PEG (-1.2 MPa) treatment. **(C)** Survival rates of Arabidopsis plants submitted to rapid and abrupt water stress treatment. (**D)** Survival rates after slow and continuous water stress treatment. Survival rates (S), given as percentages (%), were measured two days after rewatering. G – genotypes; NT – nontransgenic; A, B and C – CaHB12-OE

transgenic lines. The asterisk denotes a statistically significant difference between the plant lines ($p<0,05$).

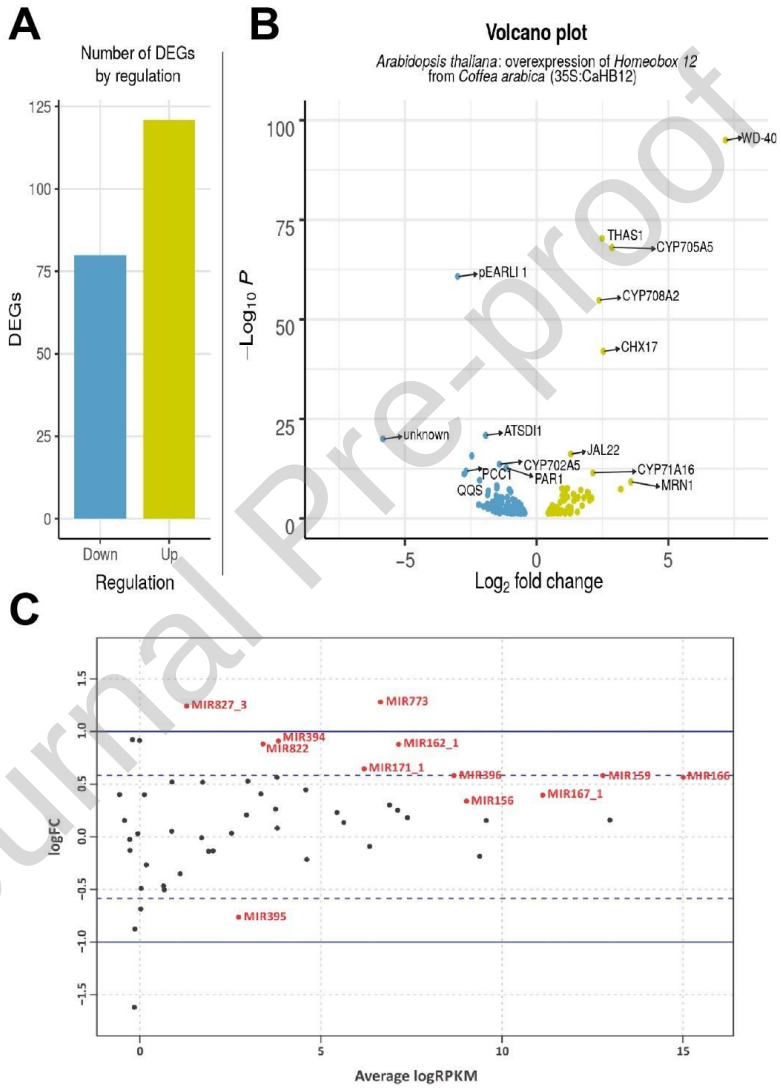


Figure 6 - Summary of differentially expressed genes of *Arabidopsis thaliana* expressing CaHB12-OE. (A) Number of DEGs separated by down- and up-regulated genes. **(B)** Volcano plot of all genes identified as DEGs (FDR < 0.05) with symbol annotation. **(C)** Fold change (in logarithmic scale) of all identified

miRNA families plotted against the logarithmic scale of average reads per million. miRNA families with significant differential expression between wild-type and transgenic lines are highlighted in red.

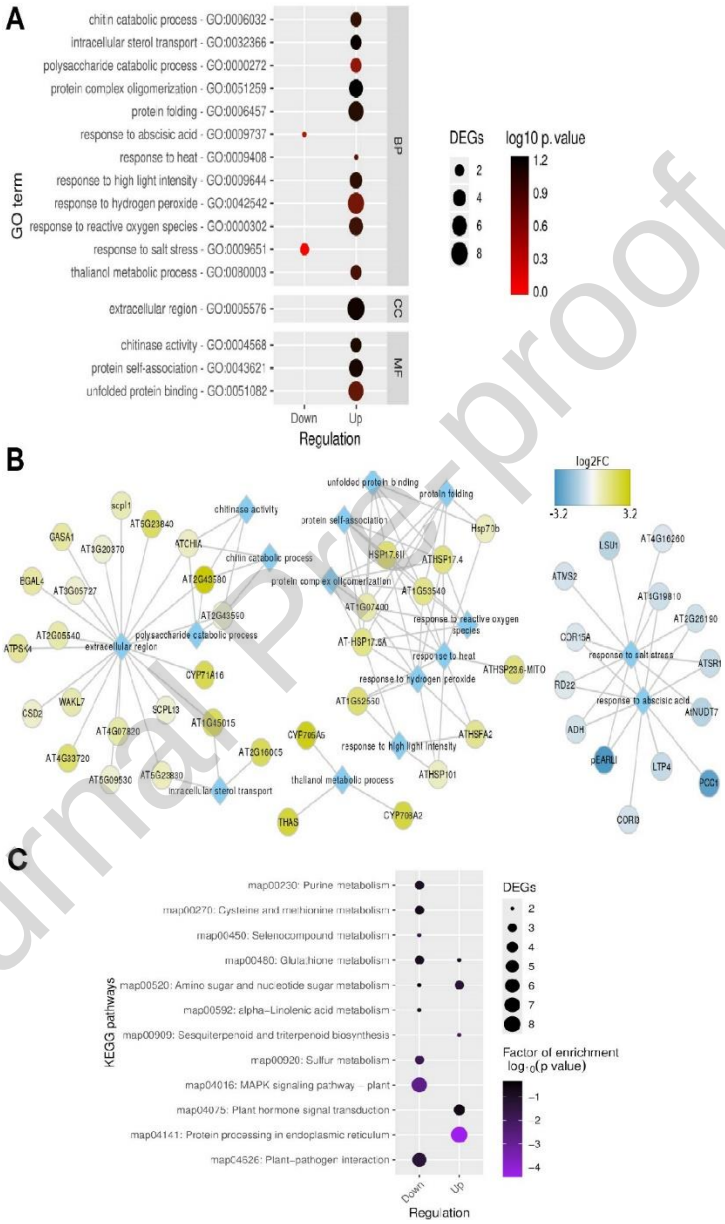


Figure 7 - Enriched GO categories within the set of up- and down-regulated differentially expressed genes. (A) Functional enrichment indicating the overrepresented GO categories after a hypergeometric test using Fisher's exact test [FDR < 0.05]. Darker tones indicate a higher fold change between transgenic and

wild-type lines. **(B)** Graph representation of genes annotated with the enriched GO terms. Green diamonds represent the GO terms, and the circle represents the genes coloured yellow (induced) or blue (repressed). **(C)** Graphical representation of overrepresented KEGG pathways identified in up- and down-regulated genes datasets in *Arabidopsis thaliana* overexpressing CaHB12.

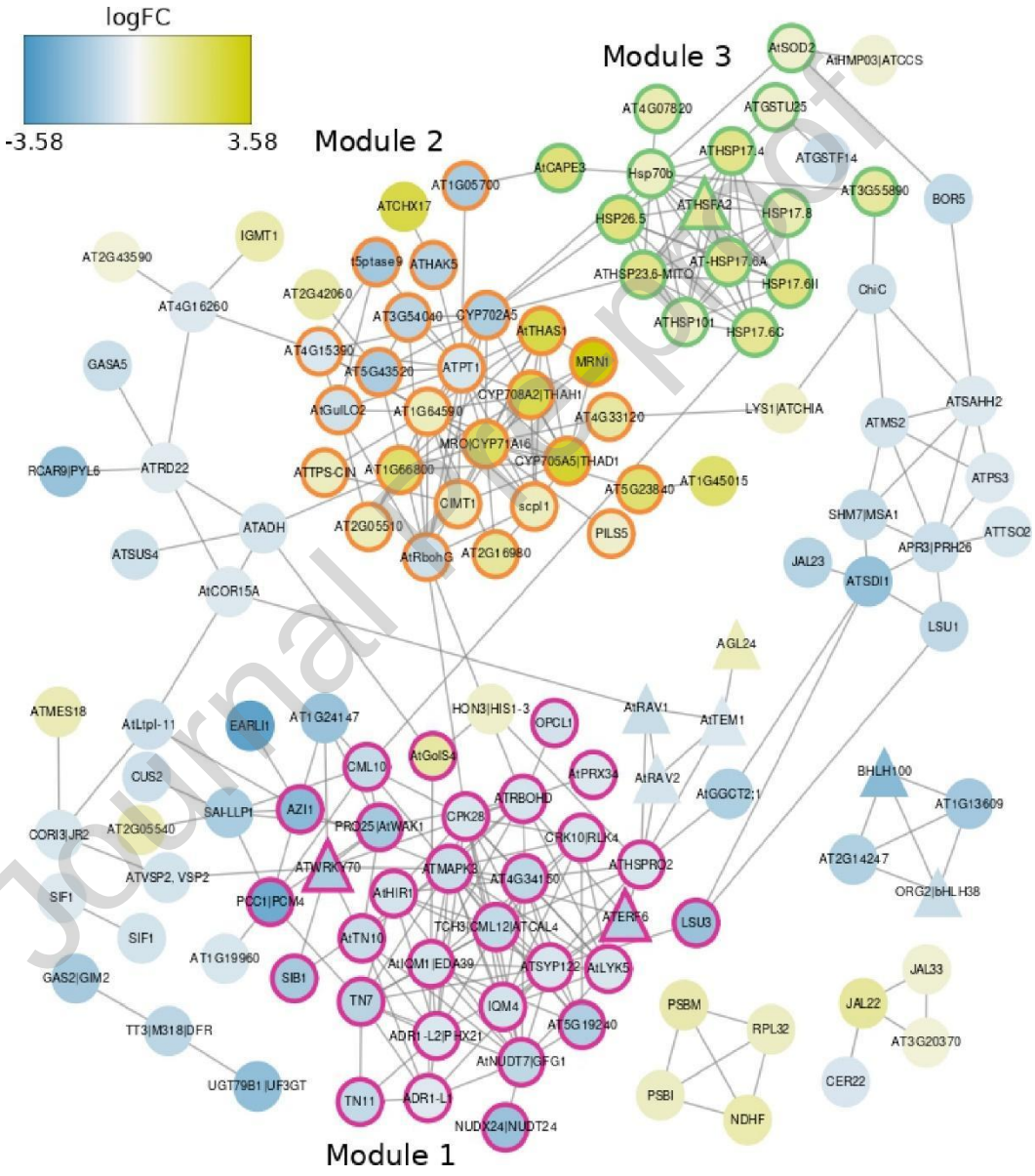


Figure 8 - Systems biology representation of protein–protein interaction (PPI) for the DEGs identified in *Arabidopsis thaliana* mutants expressing CaHB12-OE. Nodes represent the genes in which yellow indicates induced expression and

blue indicates repressed expression. Border colours are module notations from network topological analysis, and each module was submitted to GO term enrichment analysis. Module 1: enriched with defence response and oxidative stress biological processes; module 2 with triterpenoid biosynthesis, root development and oxidation–reduction processes; and module 3 with response to abiotic stresses, reactive oxygen species, hydrogen peroxide, protein complex oligomerization and cellular response to unfolded protein.

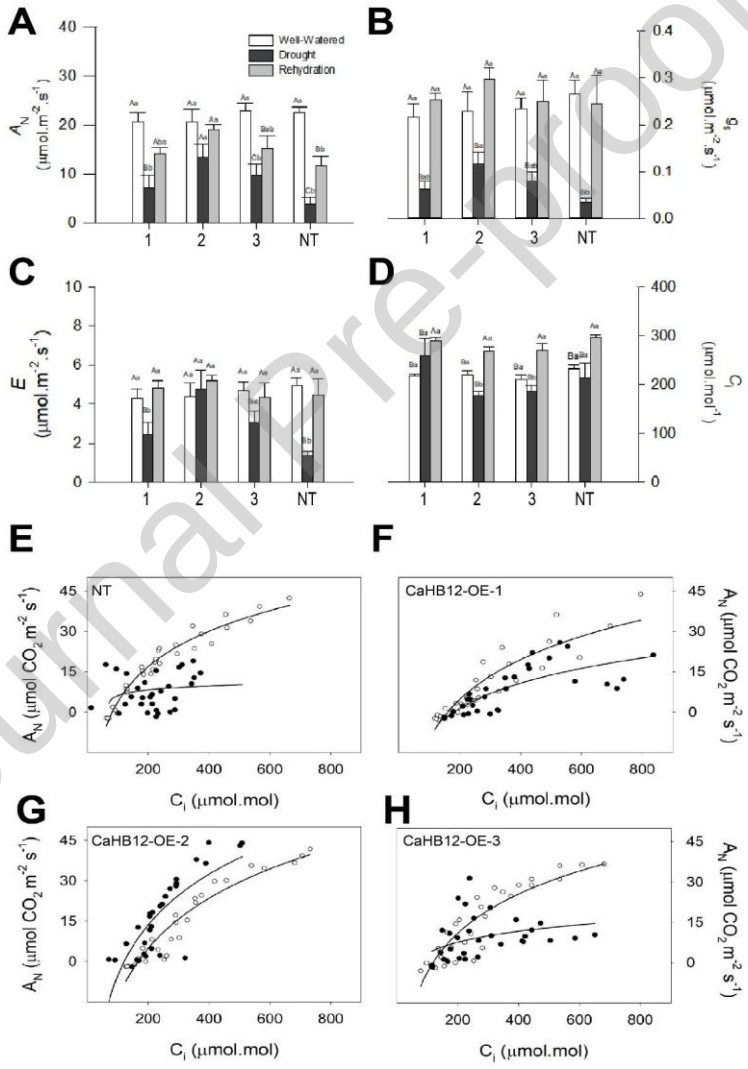


Figure 9 – Gas exchange and photosynthetic variables of CaHB12-OE soybean events (1,2 and) and non-transformed (NT) plants under drought-rehydration treatments. (A) Net CO_2 assimilation rate per unit leaf area (A_N). **(B)**

Stomatal conductance (g_s). **(C)** Transpiration (E). **(D)** Intercellular CO_2 concentration (C_i). Different capital letters indicate significant differences among treatments (well-watered, drought and rehydration) of a same genotype (transgenic lines and NT plants), while lower-case letters indicate significant differences among genotypes in the same treatment by Tukey-Kramer HSD ($p \leq 0.05$). Error bars represent standard errors of the mean (N=6). **(E-H)** N/C_i curves of (E) NT plant and CaHB12-OE lines **(F)** 1 **(G)** 2 and **(H)** 3. White and black circles indicate data of well-watered and drought-stressed plants, respectively.

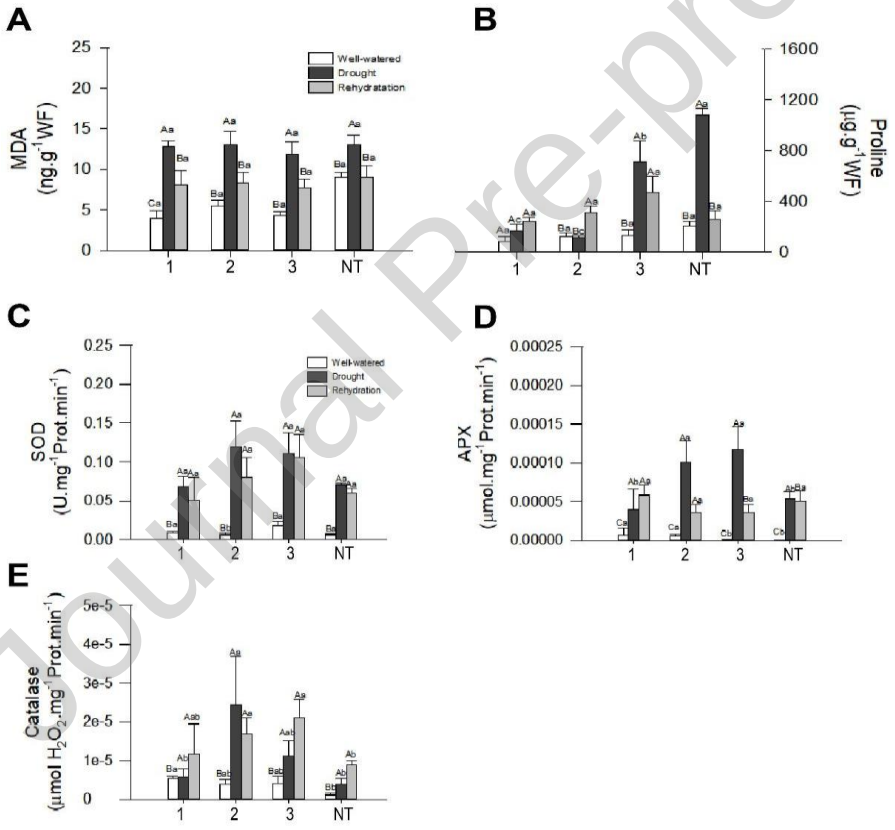
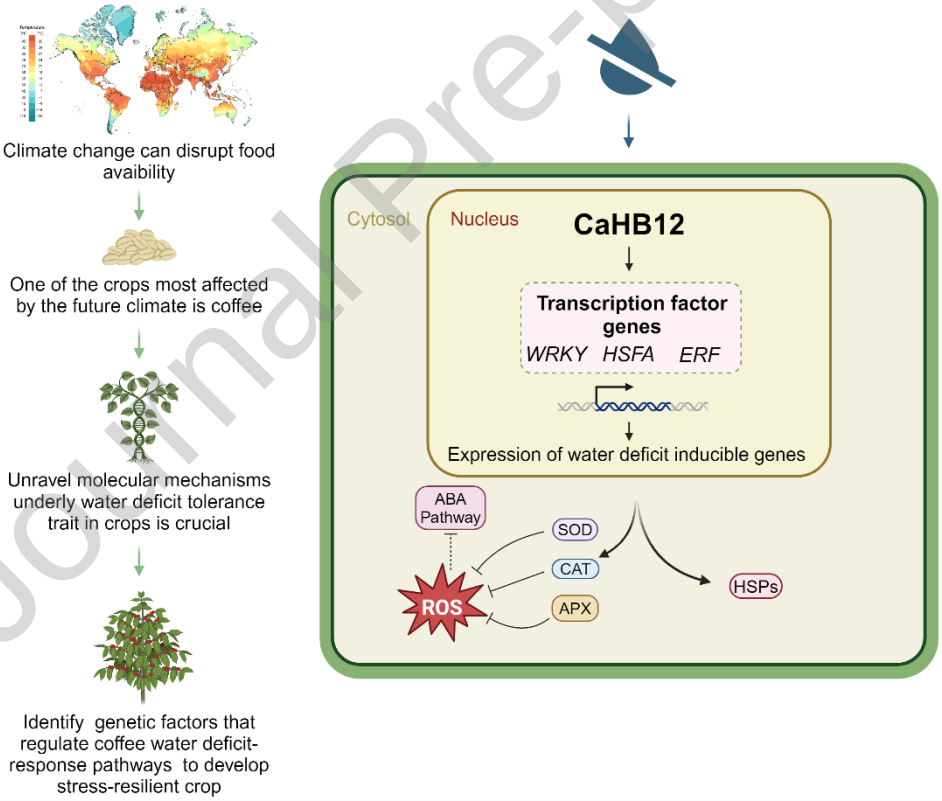


Figure 10 – Functional characterization of CaHB12-OE soybean events (1, 2 and 3) and non-transformed (NT) plants under drought-rehydration treatments. (A) Malondialdehyde (MDA) and **(B)** proline content in CaHB12-OE lines (1, 2 and 3) and non-transformed (NT) plants under different watering

treatments**.** **(C-E)** Antioxidant enzyme activity in CaHB12-OE lines and non-transformed plants under different watering treatments. **(C)** Superoxide dismutase activity (SOD). **(D)** Ascorbate peroxidase activity (APX). **(E)** Catalase (CAT). Different capital letters indicate significant differences among genotypes under different watering treatments (well-watered, drought and rehydration), while lower-case letters indicate significant differences among different genotype in the same treatment by Tukey-Kramer HSD ($P \leq 0.05$). Error bars represent standard errors of the mean (N=6).

Graphical abstract



Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Marcio Alves-Ferreira reports financial support was provided by National Council for Scientific and Technological Development. Marcio Alves-Ferreira has patent #UTILIZAÇÃO DO GENE HOMEOBOX DE CAFÉ CAHB12 NA PRODUÇÃO DE PLANTAS TRANSGÊNICAS MAIS TOLERANTES AO DÉFICIT HÍDRICO E ESTRESSE SALINO issued to PI 1015903-7. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Highlights

- Many transcription factors are implicated in regulation of gene expression in plants responses to drought, one of them is the homeobox domain (HD).
- *CaHB12* gene, a HD of *Coffea arabica*, increases the survival rate and regulates the drought tolerance.
- *CaHB12* revealed clear effects on stomatal conductance and antioxidant activity.
- CaHB12 plays a crucial role in the acquisition of drought tolerance and has biotechnological potential.