Positive Selection and Heat-Response Transcriptomes Reveal Adaptive Features of the *Arabidopsis* Desert Relative, *Anastatica hierochuntica*

Gil Eshel[a,1,], Nick Duppen[a,1], Guannan Wang[b], Dong-Ha Oh[b], Yana Kazachkova[a,2], Pawel Herzyk[c], Anna Amtmann[c], Michal Gordon[d,], Vered Chalifa-Caspi[d], Michelle Arland Oscar[e], Shirli Bar-David[f], Amy Marshall-Colon[g], Maheshi Dassanayake[b,3] and Simon Barak[h,3]

[a]Albert Katz International School for Desert Studies, Ben-Gurion University of the Negev, Sde Boqer Campus, 8499000, Israel

[b]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA

[c]Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

[d]Bioinformatics Core Facility, The National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel

[e]Blaustein Center for Scientific Cooperation, Ben-Gurion University of the Negev, Sde Boqer Campus, 8499000, Israel

[f]Mitrani Department of Desert Ecology, Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion, 8499000, Israel

[g]Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[h]French Associates' Institute for Agriculture and Biotechnology of Drylands, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sde Boqer Campus, 8499000, Israel

[1]These authors contributed equally to the article.

[2]Current address: Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel

[3]Address correspondence to: simon@bgu.ac.il or maheshid@lsu.edu

Short title: Adaptive Transcriptome Features of a Desert Plant

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) are: Simon Barak (simon@bgu.ac.il) and Maheshi Dassanayake (maheshid@lsu.edu)

Abstract

Extremophytes have evolved genetic adaptations for tolerance to abiotic stresses characteristic of their extreme environments. Comparative molecular analyses of *Arabidopsis thaliana* with its halophytic extremophyte relatives have revealed that the halophytes exist in a pre-adapted, stress-ready state. We generated a reference transcriptome of the heat-tolerant *A. thaliana* desert relative, *Anastatica hierochuntica* (True Rose of Jericho) and used two approaches to identify adaptations that could facilitate an extremophyte lifestyle: (i) We identified common positively selected extremophyte genes that target stomatal opening, nutrient acquisition, and UV-B induced DNA repair. In *A. hierochuntica*, we identified genes consistent with a photoperiod-insensitive, early-flowering phenotype that could maximize fitness in the desert environment; (ii) Using RNA-seq analysis, we demonstrate that *A. thaliana* and *A. hierochuntica* transcriptomes exhibit similar transcriptional adjustment in response to heat, and that the *A. hierochuntica* transcriptome does not exist in a heat stress-ready state, unlike its halophytic relatives. Furthermore, the *A. hierochuntica* global transcriptome as well as orthologs belonging to specific functional groups, display a lower basal expression but higher heat-induced expression than in *A. thaliana*. We suggest that the increased reactiveness of the *A. hierochuntica* transcriptome in response to heat stress is related to specific conditions native to a desert environment.

Introduction

Plant species inhabiting extreme environments - so-called extremophytes - are able to thrive in some of the most inhospitable environments on Earth that are characterized by severe abiotic stresses. These stresses include drought and temperature extremes in deserts, intense cold in the Antarctic, and habitats both on land and in the sea that are typified by acute salinity (John and Spangenberg, 2005; Amtmann, 2009; Dassanayake et al., 2010; Oh et al., 2012; Lawson et al., 2014; Farrant et al., 2015; Kazachkova et al., 2018; Oscar et al., 2018). Understanding how extremophytes adapt to their stressful environments could aid in identifying targets for molecular breeding efforts to improve crop stress tolerance, as well as facilitate the development of extremophyte-based agriculture (Bressan et al., 2013; Shabala, 2013; Cheeseman et al., 2015; Ventura et al., 2015).

To gain insight into genetic adaptations that facilitate an extremophyte lifestyle, comparative physiological and molecular analyses of stress-sensitive *Arabidopsis thaliana* and its extremophyte relatives have proven to be a powerful approach (Kraemer, 2010; Koenig and Weigel, 2015; Kazachkova et al., 2018). Indeed, these extremophyte relatives are becoming premier models for understanding plant adaptation to extreme environments with the development of a number of genetic resources including chromosome-level genome assemblies, natural accession collections, transformation protocols, and web resources (http://extremeplants.org/) (Zhu et al., 2015; Kazachkova et al., 2018; Wang et al., 2019). For example, the halophyte models *Eutrema salsugineum* and *Schrenkiella parvula*, have revealed that differences between salt-sensitive and salt-tolerant relatives are associated with altered regulation of basic physiological and molecular processes including: (i) global pre-adaptation to stress, which is manifest as reduced adjustment of the transcriptome, proteome, and metabolome in response to a stress challenge compared to *A. thaliana*, and constitutively high or low expressed genes in the halophytes that are induced or repressed, respectively, in stress-sensitive plants. In other words, these halophytes appear to exist in a stress-ready state even under stress-neutral conditions (Taji et al., 2004; Kant et al., 2006; Kant et al., 2008a; Lugan et al., 2010; Pang et al., 2010; Kazachkova et al., 2013; Oh et al., 2014; Wang et al., 2021); (ii) *E. salsugineum* is able to maintain energy supply under saline conditions via alternative pathways acting as sinks for excess electrons and via protection of chloroplasts by a highly active ROS-scavenging system (Stepien and Johnson, 2009; Wiciarz et al., 2015; Pilarska et al., 2016); (iii) at the genome level, structural changes have led to the selective expansion (e.g. tandem duplication) of genes with subfuctionalization and/or neofuctionalization leading to changes in expression and/or function (Sun et al., 2010; Dassanayake et al., 2011a; Dassnayake et al., 2011b; Ali et al., 2012; Oh et al., 2014; Ali et al., 2016; Ali et al., 2018;

Kazachkova et al., 2018). Many of the genes with copy number variation possess functions that could be important for stress tolerance. In addition, the model halophyte genomes retain a substantial proportion of taxonomically restricted genes with unknown functions that could denote unique stress tolerance determinants (Dassanayake et al., 2011a; Wu et al., 2012; Yang et al., 2013; Oh et al., 2019).

Extremophyte species are tolerant to the multiple abiotic stresses that are characteristic of their native environments. Yet, an extremophyte Brassicaceae model that represents desert species has not hitherto been developed. Such a model could leverage the functional genomics knowledge that exists for *A. thaliana* thereby facilitating comparative analyses to understand plant adaptations to the extreme desert environment. We have therefore been studying the *A. thaliana* relative, *Anastatica hierochuntica*, also known as the 'True Rose of Jericho', a Saharo-Arabian desert species (Fig. 1A) which also occupies the uppermost, driest zones of wadies or runnels of the Israeli Negev desert (Friedman and Stein, 1980; Friedman et al., 1981; Fig. 1B). This arid region has temperatures varying between -3.6 and 46 °C, an annual rainfall between 25 and 200 mm, and soil nitrate levels ranging from 0.4 to 4 mM (Gutterman, 2002; Ward, 2009; Eshel et al., 2017). We have demonstrated that *A. hierochuntica* is highly tolerant to heat, low soil nitrogen and oxidative stresses, and moderately tolerant to salt stress (Eshel et al., 2017). We further showed that in response to salt stress, *A. hierochuntica* shares common salt tolerance mechanisms with its halophytic relatives including tight control of salt uptake into shoots, and resilient photochemistry (Eppel et al., 2014; Eshel et al., 2017). Furthermore, metabolic profiling indicates constitutive upregulation under control and saline conditions of metabolites that have a role in scavenging of reactive oxygen species. These data suggest that *A. hierochuntica* might also exist in a stress-ready state similar to its halophytic relatives.

In the current study, we furthered our molecular investigations by *de novo* sequencing and assembly of an *A. hierochuntica* reference transcriptome. We then used coding sequences of *A. hierochuntica* and other available Brassicaceae to identify positively selected genes that may contribute to adaptation to extreme conditions in general, and desert conditions, in particular. We also tested our hypothesis that *A. hierochuntica* exists in a stress-ready state by comparing the response of the *A. thaliana* and *A. hierochuntica* transcriptomes to heat stress.

Results

De novo assembly and annotation of the *Anastatica hierochuntica* reference transcriptome

To generate a high-quality *A. hierochuntica* reference transcriptome that maximizes coverage of genes contained in the genome, we sequenced and assembled transcripts using RNA pooled from

multiple plant organs (root, shoot, flower, seeds), at different developmental stages (early seedling stage, and mature plants before and after flower initiation), and under control, heat, drought and salinity stress conditions (Supplemental Fig. S1; Supplemental Methods). We identified 30,670 putative protein coding genes out of the high-confidence 36,871 assembled transcripts (Fig. 1C; Supplemental Methods), and the distribution of transcript lengths was similar to that of *A. thaliana* cDNA length distribution (Fig. 1C). We assessed the completeness of the reference transcriptome based on the expected presence of core genes in land plants as identified by the BUSCO database (Simão et al., 2015) and by mapping sequenced reads back to the assembled reference. We detected 93.6% BUSCOs (Fig. 1D), comparable to *de novo* assembled transcriptomes from other Brassicaceae (Lopez et al., 2017) and obtained 88% mapped reads. These data indicated that we have generated a high-quality reference transcriptome appropriate for our downstream analyses.

We annotated the reference transcriptome using sequence similarity to protein databases including NCBI, InterPro, and KEGG databases (Supplemental Table S1). This resulted in an annotation based on a previously known sequences for 96% of our assembled transcripts (Supplemental Methods)

Positively selected genes in extremophyte Brassicaceae

As an initial approach to identifying adaptations to an extremophyte lifestyle, in general, and to desert conditions, in particular, we pinpointed positively selected genes that might be indicative of adaptive evolution of stress tolerance. We first used phylogenomics to infer evolutionary relationships between 16 Brassicaceae species including *A. hierochuntica* and representing all major lineages in this family (Supplemental Table S2). *Tarenaya hassleriana* (Cleomaceae) was used as an outgroup. This led to a selection of 13,806 ortholog groups found in 17 taxa. The generated phylogenomic tree partitioned the species in concordance with their previously assigned lineages (I, II, and III), where *Aethionema arabicum* is considered to belong to a basal clade within the Brassicaceae (Fig. 2A; Franzke et al., 2011; Kiefer et al., 2014). *A. hierochuntica* (*Anastaticeae*) was assigned to LIII. It is important to note that *A. hierochuntica* is the single representative species used for LIII due to this lineage being sparsely represented in publicly available genomic databases compared to transcriptomes available for LI and LII. LIII sequences are mostly limited to plastid genomic data created for systematics studies. Thus, to the best of our knowledge our study provides the first substantial genetic resource that enables exploration into adaptive traits that have evolved in a representative Lineage III species.

The tree contains five extremophyte species considered to be naturally tolerant to various abiotic stresses (Fig 2A, red asterisks): the halophytes *Eutrema salsugineum* and *Schrenkiella parvula* (tolerant to high salinity and multiple other stresses; Kazachkova et al., 2018), *Thlaspi arvense* (freezing-tolerant; Sharma et al., 2007; Zhou et al., 2007), *A. hierochuntica* (heat-, salt-, low N-tolerant; Eshel et al., 2017) and *Arabidopsis halleri* (heavy metal hyperaccumulator, semi-alpine conditions; Hanikenne et al., 2008; Honjo and Kudoh, 2019). Therefore, to identify genes under common positive selective pressure in the extremophytes, we used the branch-site model (Yang, 1997; Yang, 2007) to test the external branches (foreground) of the five extremophyte species against all the other branches (background). We then repeated this procedure to test for positively selected genes in three specific extremophytes - the well-studied halophyte models, *E. salsugineum* and *S. parvula*, and *A hierochuntica* - by labelling each species' external branch as the foreground. Overall, we identified 194, 120, 131 and 99 genes with a positive selection signal in the "all extremophyte species", *A. hierochuntica*, *E. salsugineum*, and *S. parvula* runs, respectively (Supplemental Tables S3-S6). We also tested *A. thaliana*, as an abiotic stress-sensitive control, and identified 112 positively selected genes (Supplemental Table S7).

We next examined whether any common genes are under positive selection in the extremophytes or between the extremophytes and stress-sensitive *A. thaliana*. While we could not detect a clear convergence in the use of same positively selected orthologs in the extremophytes (Fig. 2B), the functional attributes shared by those positively selected orthologs in each extremophyte exhibited convergence (Fig. 2C; Supplemental Fig. S2; Supplemental Tables S8 to S12). Notably, orthologs associated with the GO-term "response to stress" [GO:0006950] were highly enriched in the extremophytes suggesting major selective pressure for stress tolerance imposed on the plants by their extreme environments.

Inspection of the positively selected genes from the "all extremeophyte species" run, supported association of these genes with adaptations to harsh environments. For instance, *AKS2*, *MYB52*, *WRKY75*, *ASF1B* and *PHR1*/*UVR2* that have known functions in ABA responses, phosphate starvation, heat stress, and UV-B radiation stress (Table 1 and refs. therein), were among the positively selected group of genes in the extremophytes. Interestingly, *bZIP1* (salt/drought tolerance, and nitrogen signaling) and *APX6* (ROS-scavenging) showed signatures of positive selection unique to *A. hierochuntica* (Table 1). These genes are particularly noteworthy because *A. hierochuntica* is highly tolerant to low N and oxidative stresses, and moderately tolerant to salt stress (Eshel et al., 2017). Positively selected genes unique to *S. parvula* included *CAX11*/*CCX5* and *RAB28* that are involved in

high-affinity K^+ uptake and Na^+ transport, and lithium toxicity, respectively (Table 1; Borrell et al., 2002; Zhang et al., 2011). The pinpointing of these two genes added validity to our positive selection analysis because the native soils of *S. parvula* contain levels of Li^+ and K^+ that are highly toxic to most plants (Helvaci et al., 2004; Ozfidan-Konakci et al., 2016), and this species displays extreme tolerance to both Li^+ and K^+ toxicity (Oh et al., 2014). In contrast to the extremophyte species, positively selected genes in *A. thaliana* were related to biotic stress responses (Table 1).

We further examined whether any of the "all extremophyte species" or *A. hierochuntica*-specific genes were differentially regulated between *A. thaliana* and *A. hierochuntica* exposed to heat stress (see below for RNA-seq details). Figure 3A shows that *AKS2*, *bZIP1* and *PHR1/UVR1* expression displayed clear and significantly higher transcript levels in *A. hierochuntica* compared to *A. thaliana* whereas the expression of *APX6* exhibited lower transcript levels in *A. hierochuntica*.

Exclusively in *A. hierochuntica*, we identified, *CYP71*, *FAS1*, *FBH2*, *SBI1/LCMT1*, and *VIP5* that are involved in photoperiodic flowering, regulation of meristems, and control of morphology including shoot branching (Table 1). Furthermore, two of these *A. hierochuntica* genes exhibited differential regulation compared to *A. thaliana* in response to heat stress (Fig. 3B). *AhFAS1* expression was highly upregulated by heat stress while *AtFAS1* expression was downregulated. *AhSBI1/LCMT1* expression was unaffected by heat stress but transcript levels were lower than *AtSBI1/LCMT1* over all time points whereas *AtSB1/LCMT1* expression was highly upregulated by heat. Considering that *A. hierochuntica* ontogeny is very different from *A. thaliana*, *E. salsugineum* and *S. parvula* – it exhibits a multi-branched sympodial shoot structure supporting multiple axillary inflorescences that flower independent of day length (Fig 3C; Gutterman, 1998; Eshel et al., 2017) - positive selection of these genes could indicate an important adaptation to the desert environment.

Comparative global analysis of the *A. thaliana* and *A. hierochuntica* heat response transcriptomes

Our previous work demonstrated that certain stress-associated metabolites are constitutively maintained at high levels in *A. hierochuntica* even under stress-neutral control conditions (Eshel et al., 2017). Therefore, we hypothesized that the *A. hierochuntica* transcriptome will show similar traits of constitutive expression at stress-ready levels in response to heat stress. Thus, as a second approach to identify extremophyte adaptations to a desert lifestyle, we performed a comparative analysis of the *A. thaliana* and *A. hierochuntica* transcriptome response to heat stress in young plants at similar developmental stages before anthesis. To simulate a realistic scenario, temperature data near *A. hierochuntica* populations during their growing season were obtained from the Israel Meteorological

Service and used as a guide for heat treatments (Supplemental Fig. S3). Thus, plants were exposed to either control (23 °C) or three consecutive daily heat waves covering the early heat response and acquired heat tolerance phases (Lindquist, 1986; Hong and Vierling, 2000), with day/night temperatures of 40 °C/25 °C followed by 2 d recovery at 23 °C (Fig. 4A). Plants were well-watered throughout the entire experiment to avoid any dehydration effects that could arise due to the heat treatment.

Heat stress had no significant effect on *A. hierochuntica* leaf area in contrast to *A. thaliana* where growth in leaf area was significantly retarded by heat stress although it had almost recovered to control levels 2 d after the end of the heat treatment (Fig. 4B). *A. thaliana* shoot fresh weight was also significantly reduced by heat stress but did not recuperate after 2 d recovery under control conditions while *A. hierochuntica* fresh weight was not affected by heat stress (Fig. 4C). These results illustrate that *A. hierochuntica* is highly tolerant to heat stress and confirm our previous *in vitro* experiments (Eshel et al., 2017).

A PCA analysis based on genome-wide transcript levels showed that the transcriptomes of both species under elevated temperature were clearly distinct from those in control conditions (Fig. 4D). The control and heat-stressed samples harvested in the morning were positioned separately from the samples harvested in the afternoon, which could be due to differences in early vs. late heat-mediated gene expression or/and diurnal changes in gene expression (Kant et al., 2008b; Li et al. 2019; Mody et al., 2020). Transcriptomes of plants recovering from heat stress clustered near control samples suggesting that the overall transcriptomes return to pre-stress conditions. Because these results indicate that both species undergo conditional transcriptional adjustment in response to heat stress, we examined the median expression level across the whole transcriptome for each condition. Figure 4E shows that the heat treatments conferred a similar effect upon transcript abundance between the two species. Compared to their respective controls, the median transcript abundance (and total abundance as depicted by the distribution) decreased under heat stress in the morning samples, increased in response to heat treatments in the noon samples, and decreased during recovery. Furthermore, the percentage of differentially expressed genes (out of the total number of protein-coding genes) was similar for both species under all heat conditions (Fig. 4F; Supplemental Table S19). Thus, these data support the proposition that the *A. thaliana* and *A. hierochuntica* global transcriptomes adjust to heat stress with a similar magnitude and trend.

The similar heat-mediated transcriptome adjustment exhibited by both species suggests that the *A. hierochuntica* transcriptome is not globally pre-adapted to stress - i.e. it does not exist in a stress-

ready state. To further test this hypothesis, we used Weighted Gene Co-expression Network Analysis (WGCNA; Langfelder and Horvath, 2008) to identify five types of transcriptional response mode among the expression patterns of 17,962 orthologous pairs for each species (Wang et al., 2021; Supplemental Fig. S5; Supplemental Table S13): (a) "Stress-ready" where transcript level under control conditions in one species is equal to the ortholog transcript level under heat in the other species; (b) "Shared response" where expression of both orthologs exhibit a similar response to heat; (c) "Unique response" where expression of an ortholog exhibits a heat response specifically in one species but not in the other; (d) "Opposite response" where expression of the ortholog in one species shows the opposite response in the other species; (e) "No response" where the expression of both orthologs does not respond to heat. Analysis of the total number of orthologs that could be categorized within the five response modes revealed that only 4.4% of the orthologs belonged to the "No response" mode (Fig. 5A). The vast majority (82%) of orthologs displayed a shared response mode while about 5% exhibited a unique response and 2.1% showed an opposite response. Importantly, while 535 genes did exhibit a stress-ready mode in *A. hierochuntica*, we also detected 221 *A. thaliana* genes displaying a stress-ready mode. Similar results were obtained when we examined only orthologs associated with GO terms for abiotic stress responses (Fig. 5A; Supplementary Methods). Overall, our data show that: (i) *A. thaliana* and *A. hierochuntica* exhibit similar transcriptional adjustment; (ii) a large majority of orthologs display a shared response mode unlike the extremophyte *S. parvula* response to boron toxicity where almost no genes showed a shared response (Wang et al., 2020); (iii) only a few hundred orthologs in both species show a stress-ready mode of expression. These findings suggest that *A. hierochuntica* does not display the classic features of a stress-ready transcriptome (Kazachkova et al., 2018; Wang et al., 2021).

We next searched for other characteristics that differentiate the response of the *A. thaliana* and *A. hierochuntica* transcriptomes to heat stress. By examining global transcript abundance under control conditions, we observed that the median basal expression level of the *A. hierochuntica* transcriptome was significantly lower than in *A. thaliana* (Fig. 5B). Additionally, extremophyte differentially expressed genes (DEGs) displayed a greater heat-mediated fold-change expression than *A. thaliana* DEGs (Figs. 5C and 5D). To support these findings, we increased the resolution of our analysis by comparing the basal and fold-change expression of specific functional groups of *A. thaliana* and *A. hierochuntica* orthologs that exhibited either a shared or unique response mode to heat stress (Supplemental Table S19). Orthologs associated with GO-terms for abiotic stress whose expression displayed shared upregulation by heat exhibited an average lower basal expression in *A. hierochuntica*

compared to *A. thaliana* and no significant difference in average % change in expression (Fig. 6A). Abiotic stress-associated orthologs showing shared heat-mediated downregulated expression displayed both a lower basal and higher % change in expression in *A. hierochuntica* compared to *A. thaliana* (Fig. 6B). Similarly, both heat-mediated upregulated and downregulated abiotic stress-associated, unique-expressed orthologs showed lower basal and higher % change in expression in the extremophyte (Figs. 6A and 6B).

Plants actively and early on reduce their growth in response to stress independent of photosynthesis, and in *A. thaliana* and *E. salsugineum* leaves this is apparent in a reduction in both cell size and cell elongation (Aguirrezabal et al., 2006; Skirycz et al., 2010; Kazachkova et al., 2013). Stress-mediated reduction in cell division can be linked to downregulation of cell cycle-associated genes and arrest of the cell cycle (Skirycz et al., 2011). Subsequently, expression of genes involved in photosynthesis is downregulated under stress (Rizhsky et al., 2002; Zhang et al., 2018a; Huang et al., 2019). We observed that the majority of shared- and unique-expressed orthologs associated with photosynthesis or the cell cycle were downregulated by heat stress in both species. However, for both shared- and unique-expressed photosynthesis orthologs, *A. hierochuntica* exhibited a similar basal, and higher % change in expression than *A. thaliana* (Fig. 6C). Cell cycle orthologs with shared heat-meditated downregulated expression showed a lower basal and higher % change in expression in the extremophyte while unique-expressed cell-cycle orthologs exhibited lower basal expression in *A. hierochuntica* (Fig. 6D).

To ensure that the lower basal expression of genes in *A. hierochuntica* was not due to an overall lower metabolic rate in the extremophyte compared to *A. thaliana*, we examined the basal expression of 15 orthologous housekeeping genes from both species. We found that the average ratio of basal expression of *A. thaliana* to *A. hierochuntica* housekeeping genes was 1.0 ± 0.34 (Supplemental Fig. S4), supporting the contention that the average lower basal gene expression observed in *A. hierochuntica* compared to *A. thaliana* is not due to lower metabolic activity in the extremophyte.

Comparative analysis of the *A. thaliana* and *A. hierochuntica* early heat response-specific transcriptomes

Our findings suggest that although the *A. hierochuntica* transcriptome does not exist in a stress-ready state, it is more reactive to imposition of stressful conditions than that of *A. thaliana*. To strengthen this assertion, we used WGCNA to cluster genes from all conditions into modules with similar expression profiles and detected 22 *A. thaliana* and 21 *A. hierochuntica* co-expression modules

(Supplemental Fig. S5). In both species, two modules were observed that clearly covered early heat-induced genes (1.5 h [morning] and 7 h [afternoon] after onset of heat stress (Fig. 7A; Supplemental Tables S14-S18). The morning heat response modules of both species were enriched in GO biological terms such as "response to heat", "response to high light intensity" and "response to reactive oxygen species", (Fig. 7B; Supplemental Methods) while the *A. thaliana* and *A. hierochuntica* afternoon heat response modules were not enriched in any GO-terms. Importantly, both shared and unique expressed *A. hierochuntica* genes associated with GO terms for abiotic stress exhibited the same or lower basal expression, and greater heat-mediated % change in expression than their *A. thaliana* orthologs (Fig. 7C). Thus, the early heat response module analysis provides further support for the notion that the *A. hierochuntica* transcriptome is more reactive to a heat stress challenge than *A. thaliana*.

Validation of "between species" RNA-seq analyses

The above comparisons of gene expression between the two species utilized DeSeq2 (Love et al., 2014) as a normalized measure of gene expression and to identify differentially expressed genes. DeSeq2 normalizes read counts for different sequencing depths between samples. However, when dealing with two different species, several other factors can affect direct comparison of expression levels between orthologs including whether a few highly expressed genes constitute a large proportion of the sequenced transcripts, as well as differences in gene numbers and orthologous transcript length (Zhou et al., 2019; Zhao et al., 2020). Therefore, we performed a number of further analyses to validate our results. Figure 8A shows that there was no significant difference between the species in the proportion of the top 10 most highly expressed genes out of the total transcripts sequenced across all treatments (*A. thaliana*, ~7% to 12%, and *A. hierochuntica*, ~10% to 15% of the total sequenced transcripts). We further re-normalized our raw read count data (normalized for transcript length) using a new between-species method that applies Scale-Based Normalization (SCBN) to the most conserved orthologs, thereby obtaining a scaling factor that minimizes the false discovery rate of differentially expressed genes (Zhou et al., 2019). Applying SCBN to the 109 most conserved orthologs between *A. thaliana* and *A. hierochuntica* (Supplemental Table S20) and using the scaling factor to correct normalized gene counts, we obtained similar comparative basal expression results as observed with DeSeq2 (Fig. 8B). Finally, QPCR analysis confirmed the RNA-seq fold-change gene expression patterns of four selected *A. thaliana* and *A. hierochuntica* genes (Fig. 8C). These genes included *AKS2*, a gene found to be positively selected in the "all extremophyte species" run (Table 1), two genes involved in abiotic stress responses (*ELIP1* and *HSP17.6II*; Sun et al., 2001; Rizza et al., 2011), and an

A. thaliana-specific and an *A. hierochuntica*-specific Taxonomically Restricted Gene (*A. hierochuntica* ID: TRINITY_DN7044_c0_g2_i1; AGI: At2g07719; Supplemental Methods). Additionally, we selected an ethylene signaling gene *CPR5* (Wang et al., 2017) that exhibited lower basal expression in *A. hierochuntica* than in *A. thaliana* in the RNA-seq analysis and confirmed this result via absolute QPCR (Fig. 8D).

Discussion

Positively selected *A. hierochuntica* genes suggest adaptive evolution for an opportunistic desert lifestyle

In this study, we sequenced and assembled a high-quality reference transcriptome of the *Brassica* lineage III desert plant *A. hierochuntica* and then constructed a phylogenomic tree from 16 Brassicaceae. Using this phylogeny, we pinpointed a number of positively selected genes specifically in *A. hierochuntica* indicating adaptation to the desert environment (Table 1; Fig. 3B). Intriguingly, these genes function in *A. thaliana* in the transition from vegetative to reproductive growth and meristem development: (i) *VERNALIZATION INDEPENDENCE VIP5* encodes a protein that is part of the RNA polymerase II-associated factor 1 (PAF1c) complex, which is involved in enhancing transcription of the floral repressor *FLOWERING LOCUS C* (*FLC*) gene and other *MADS AFFECTING FLOWERING* (*MAF*) gene family members, via histone modifications (Oh et al., 2004; Yu and Michaels, 2010; Crevillen and Dean, 2011; Lu et al., 2017). *A. thaliana vip5* mutants exhibit early flowering and floral abnormalities such as incomplete sepal closure of the developing flower bud, and diminished petal development; (ii) FLOWERING BHLH 2 (FBH2) is a transcriptional activator of the *CONSTANS* gene, a central regulator of photoperiodic flowering that is essential for proper measurement of day-length. Overexpression of *FBH2* causes early flowering regardless of photoperiod (Ito et al., 2012); (iii) *FASCIATA1* (*FAS1*) encodes a subunit of the CHROMATIN ASSEMBLY FACTOR-1 (CAF-1) that functions as a histone chaperone to deposit histones H3/H4 onto DNA at the replication fork and appears to function in the organization of shoot and root apical meristems, and in cellular differentiation (Kaya et al., 2001; Exner et al., 2006). Mutations in *fas1* cause stem fasciation, abnormal leaf and flower morphology, and defects in the organization of apical meristems (Leyser and Furner, 1992; Kaya et al., 2001). Interestingly, the Caf1 complex interacts with ASF1 in humans, *Drosophila* and yeast, and they have overlapping functions in *A. thaliana* (Sharp et al., 2001; Tyler et al., 2001; Mello et al., 2002; Zhu et al., 2011). *ASF1b* was identified as a gene undergoing positive

selection in the extremophytes and both ASF1b and FAS1 also function in DNA repair (Hisanaga et al., 2013; Lario et al., 2013); (iv) *CYP71* encodes a cyclophilin, that plays a critical role in epigenetic silencing of homeotic genes that regulate meristem development, including the floral meristem (Li et al., 2007). Furthermore, CYP71 physically interacts with FAS1 and is required for targeting FAS1 to the *KNAT1* locus (Li and Luan, 2011). KNAT1 is a KNOX family homeodomain protein that is essential for maintenance of apical meristems (Hake et al., 2004), and FAS1 is probably involved in silencing the *KNAT1* gene. In addition, CYP71 interacts with LIKE HETEROCHROMATIN PROTEIN 1 (LHP1), which is associated with repressive H3K27 methylation deposited by Polycomb repressor complex, PRC2 (Li and Luan, 2011; Feng and Lu, 2017). Mutations in the *LHP1* gene cause strong early flowering and LHP1 is involved in repressing expression of flowering time and floral identity genes (Gaudin et al., 2001; Kotake et al., 2003); (v) *SUPPRESSOR OF BRI1 (SBI1)/LEUCINE CARBOXYLMETHYLTRANSFERASE (LCMT1)* encodes a protein that methylates protein phosphatase A (PP2A), which then dephosphorylates components of the brassinosteroid signaling pathway (Di Rubbo et al., 2011; Wu et al., 2011). The *sbi1/lcmt* mutant is early flowering in both long and short days and exhibits downregulation of *FLC* expression. Furthermore, genes that are significantly differentially regulated between WT and the *sbi1/lcmt* mutant are enriched in stress response genes (Creighton et al., 2017), consistent with the role of brassinosteroids in flowering and stress responses (Li and He, 2010; Nolan et al., 2020).

The finding that genes involved in regulating flowering and meristem development were only under positive selection in *A. hierochuntica* is consistent with the very different developmental program in the desert relative compared to many other Brassicaceae including the additional four extremophyte plants included in our analysis. *A. hierochuntica* does not display the distinctive transition from the vegetative rosette leaf stage to the reproductive bolting stage, which is accelerated in long-day conditions (Pouteau and Albertini, 2009; Song et al., 2013). Instead, regardless of photoperiod, the shoot repeatedly bifurcates from the four true-leaf stage onwards, developing an axillary inflorescence at each branch point thereby leading to a multi-branched shoot morphology (Fig. 1B, panel (i); Eshel et al., 2017). Most interestingly, mutation in the *A. thaliana FAS1* gene (whose *A. hierochuntica* ortholog is under positive selection) can induce stem bifurcation and enlargement (Leyser and Furner, 1992). The shoot bifurcation, multi-branch, photoperiod-insensitive, early flowering traits could maximize fitness in the unpredictable desert environment where plants need to ensure development of seeds but might not survive until a critical day length induces flowering. This idea is supported by our observations of *A. hierochuntica* populations in the Dead Sea valley of Israel,

where tiny dead plants that have still managed to produce a few seeds can be seen alongside much larger plants presumably from a year with higher rainfall (Fig. 1B).

Brassicaceae extremophytes possess common positively selected genes that are indicative of adaptation to harsh environments

Extremophytes are present in all three Brassicaceae lineages (Fig. 2A; Franzke et al., 2011) illustrating that adaptation to stressful habitats has occurred independently, multiple times within the Brassicaceae and, in agreement with previous reports, is indicative of convergent evolution (Kazachkova et al., 2018; Birkeland et al., 2020). Using the external branches of the five extremophytes in our phylogeny as the foreground, we identified common positively selected genes that could contribute to plant adaptation to extreme environments. Other studies have detected signatures of positive selection in genes that function in stress tolerance (Xia et al., 2010; Fischer et al, 2011; Mboup et al., 2012; Vigeland et al., 2013; Yang et al., 2014; Bondel et al., 2018; Birkeland et al., 2020; Liu et al., 2020), and with respect to Brassicaceae extremophytes, three reports are noteworthy. Birkeland et al. (2020) found evidence for positive selection of stress-associated genes in three Arctic species. Most interestingly, they discovered little overlap of positively selected genes between the extremophytes but considerable overlap in functional pathways, similar to our findings with the extremophyte plants in the current study (Figs. 2B and 2C; Supplemental Fig. S2). They suggested that these data do not support adaptive molecular convergence but rather indicate evolution of similar adaptations via distinct evolutionary pathways. In the second report, positive selection was detected in the tandem duplicates of the *COR15* gene in species of the highly freezing-tolerant alpine/arctic *Draba* genus (Azocar et al., 1988; Koch and Al-Shehbaz, 2002) suggesting neo-functionalization (Zhou et al., 2009). *COR15* genes function in freezing tolerance by stabilizing cell membranes (Thalhammer and Hincha, 2014). In the third study, the gene encoding the *E. salsugineum* Na^+/H^+ antiporter gene, *SALT OVERLY SENSITIVE 1* (*SOS1*), was found to have evolved under positive selection (Jarvis et al., 2014). EsSOS1 is an essential component of *E. salsugineum* halophytism (Oh et al., 2007; Oh et al., 2009), and confers greater tolerance to a salt-sensitive yeast strain than its orthologs from *A. thaliana* or *S. parvula* partly due to a single amino acid change in the putative autoinhibitory domain. This finding coupled with differential regulation of the *AtSOS1* and *EsSOS1* genes (Taji et al., 2004; Kant et al., 2006; Oh et al., 2010; Taji et al., 2010; Dassanayake et al., 2011b) suggests that *EsSOS1* has undergone positive selection to enhance its ancestral functional role in Na^+ exclusion thereby facilitating adaptation of the extremophyte to saline habitats.

In our current study, we identified a set of 194 genes under positive selection across the five extremophyte species. Among these genes, we identified two genes encoding ABA-responsive transcription factors (TFs), AKS2 and MYB52 (Table 1), which, along with other studies showing positive selection in ABA biosynthesis and signaling genes in wild tomato populations from dry environments, illustrates the importance of the ABA response networks in adaptive evolution of stress tolerance (Xia et al., 2010; Fischer et al, 2011; Bondel et al., 2018). Of particular interest was the gene encoding the bHLH TF, ABA-RESPONSIVE KINASE SUBSTRATE (AKS) 2. This gene is a member of a family of three TFs in *A. thaliana* that activate transcription of K^+ channels in stomata guard cells thereby enhancing stomatal opening (Takahashi et al., 2013). ABA-dependent phosphorylation of AKS2 represses its activity contributing to stomata closure (Takahashi et al., 2013; Bu et al., 2016). It is intriguing that a regulator of stomatal aperture has undergone positive selection in *A. hierochuntica* because alterations in stomatal aperture is a crucial early response to multiple abiotic stresses, although specific stresses lead to distinct stomatal responses. For instance, in most plant species including *A. thaliana*, drought and salt stresses cause stomatal closure in order to prevent water loss albeit at the expense of CO_2 availability for carbon fixation (Brugnoli and Lauteri, 1991; Lawlor, 1995; Chaves et al., 2009; Stepien and Johnson, 2009). On the other hand, heat shock can cause stomatal opening presumably for evaporative cooling while prolonged, non-lethal warming causes stomatal closure to prevent water loss (Devireddy et al., 2020; Wang et al., 2020a). *A. thaliana* and *E. salsugineum* exhibit differential regulation of stomatal aperture whereby salt stress causes stomatal closure and inhibition of net CO_2 assimilation in the glycophyte but only marginal reduction in stomatal aperture and net CO_2 assimilation in the halophyte, even under extreme salt concentrations (Stepien and Johnson, 2009). Thus, positive selection of non-synonymous amino acid changes in the coding region of *AKS2*, plus differences in heat-mediated regulation of *A. thaliana* and *A. hierochuntica AKS2* expression (Fig. 3A) suggest that this gene may have been naturally selected for survival in extreme environments.

We also identified the *WRKY75* as a positively selected gene in the extremophytes (Table 1). In *A. thaliana*, the WRKY75 TF is involved in the response to inorganic phosphate (Pi) starvation by regulating the expression of several key phosphate starvation-induced genes (Devaiah et al., 2007). Extremophytes often exist on soils with low Pi availability due to patchy distribution, low soil concentrations or high alkalinity, which reduces phosphorus solubility (Thompson et al., 2006; Holzapfel, 2008; Guevara et al., 2012). For instance, the *E. salsugineum* Yukon ecotype grows in the semi-arid, sub-arctic, saline and highly alkaline soils of the Yukon region in Canada (Guevara et al., 2012). Plants of this ecotype are highly tolerant to Pi deficiency compared to *A. thaliana*, and this

tolerance is associated with higher basal expression of several genes involved in the response to Pi starvation including *WRKY75* (Velasco et al., 2016). We previously showed (Eshel et al., 2017), that both *E. salsugineum* and *A. hierochuntica* exhibit significantly higher basal levels of Pi than *A. thaliana* consistent with several reports on *E. salsugineum* Shandong and Yukon ecotypes (Gong et al., 2005; Kazachkova et al., 2013; Velasco et al., 2016). The higher Pi accumulation in these extremophytes, positive selection across the five extremophyte plants, and differential expression of *A. thaliana* and *E. salsugineum WRKY75* suggests that selection for more efficient extraction of soil Pi is a common feature of evolution of stress tolerance particularly in extreme environments.

In addition to low soil nutrients, plants can be exposed to other abiotic stresses including temperature extremes and UV-B radiation that cause damage to DNA, proteins and lipids both directly, and via ROS production. UV-B exposure causes direct DNA damage by inducing the formation of cyclobutane pyrimidine dimers (CPDs) (Kimura and Sakaguchi, 2006). If CPDs accumulate they can block processes such as transcription and replication, and can eventually cause cell death. *PHOTOLYASE1*/*UV-RESISTANCE2* (*PHR1*/*UVR2*) specifically repairs UV-B-induced CPDs in a process known as photoreactivation (Ahmad et al., 1997; Landry et al., 1997; Jiang et al., 1997). Furthermore, CPD repair by *PHR1*/*UVR2* is the major mechanism maintaining transgenerational genome stability in *A. thaliana* exposed to continuous UV-B (Willing et al., 2016). We found that *PHR1*/*UVR2* has undergone positive selection in the extremophyte plants, as has *ANTI-SILENCING FUNCTION 1b* (*ASF1B*), an ASF1 histone H3/H4 chaperone that is also important for the repair of UV-B-induced DNA damage, and functions in the transcription of genes involved in basal and acquired thermotolerance (Table 1; Lario et al., 2013; Nie et al., 2014; Weng et al., 2014). *PHR1*/*UVR2* also displays higher basal expression in *A. hierochuntica* compared to *A. thaliana*, and while heat leads to downregulation of the gene in both species, expression is reduced to a lesser extent in the extremophyte. (Fig. 3A). *ASF1B*, on the other hand exhibits similar expression in both species.

Taken together, our positive selection analysis suggests common selective pressures in extremophyte plants that target key components in stomatal opening, in nutrient acquisition, and in UV-B induced DNA repair. On the other hand, we found that *A. thaliana* positively selected genes were principally involved in defense against pathogens (Table 1). This finding is consistent with the discovery that *A. thaliana* orthologs exhibiting higher basal expression compared to *S. parvula* are enriched in pathogen-defense genes whereas *S. parvula* orthologs displaying higher basal expression are enriched in genes involved in ion transport (Oh et al., 2014). Thus, our positive selection analysis coupled with the findings of Oh et al (2014) support the hypothesis that because *A. thaliana* evolved in

temperate regions where pathogen density is relatively high compared to extremophyte habitats such as salt flats or deserts, it encountered greater evolutionary pressures for adaptation to biotic stresses.

The *A. hierochuntica* transcriptome does not exist in a heat stress-ready state and is more reactive to heat stress than *A. thaliana*

The distinction between *A. hierochuntica* and the other extremophyte relatives was further illustrated by our finding that the *A. hierochuntica* transcriptome behaves very differently in response to stress. The extent of transcriptomic, proteomic and metabolic adjustment of the halophytes *E. salsugineum* and *S. parvula* in response to ionic stress is much lower than that of *A. thaliana* (Taji et al., 2004; Gong et al., 2005; Lugan et al., 2010; Pang et al., 2010; Oh et al., 2014; Vera-Estrella et al., 2014; Kazachkova et al., 2018; Wang et al., 2021). In contrast, *A. thaliana* and *A. hierochuntica* exhibit similar transcriptome adjustment in response to heat stress and during recovery (Figs. 4E and 4F). The lower transcriptomic and metabolic adjustment of the halophytic relatives is reflected in their stress-ready state in stress-neutral conditions whereby transcript and metabolite accumulation that is induced or repressed in *A. thaliana* in response to ionic stress, exhibits constitutively high or low accumulation, respectively, in the halophytes (Gong et al., 2005; Lugan et al., 2010; Kazachkova et al., 2013; Oh et al., 2014; Kazachkova et al., 2018; Wang et al., 2021). For instance, *S. parvula* possesses a pre-boron stress-adapted genome with over 1,000 stress-ready orthologs whose basal expression levels match the post-stress expression levels observed in *A. thaliana* (Wang et al., 2021). No boron stress-ready *A. thaliana* orthologs were detected and there were almost no orthologs showing a shared response to excess boron.

In contrast to the halophytic relatives, the great majority of *A. hierochuntica* genes exhibited a shared response mode with their *A. thaliana* orthologs (Fig 5A). Moreover, while we could detect some heat stress-ready *A. hierochuntica* genes, we also observed stress-ready *A. thaliana* genes. On the other hand, we discerned almost universal lower basal and higher fold-change gene expression in *A. hierochuntica* compared to *A. thaliana*. This was observed whether we compared global transcriptomes (Figs. 5B to 5D), shared/unique-expressed specific functional categories of orthologs (Fig. 6) or early heat response module-specific genes (Fig. 7).

Taking all our results together, the similar transcriptome adjustment to heat stress and recovery between heat stress-sensitive *A. thaliana* and the heat-tolerant extremophyte, as well as the analyses of transcriptome expression modes, support our contention that the *A. hierochuntica* transcriptome does not exist in a heat stress-ready state in contrast to its halophytic relatives. However, the *A.*

hierochuntica transcriptome is more reactive (shows greater fold-change in expression) to heat stress than *A. thaliana*. One possible reason for this difference in the transcriptome responses of *A. hierochuntica* vs. its halophytic relatives may be the native environment of each species. *A. hierochuntica* survives and thrives in an environment with seasonal temperatures that range from -3.6 to 46.8 °C (Eshel et al., 2017), with diurnal maximum variations exceeding 18 °C (Israeli Meteorological Services). On the other hand, *E. salsugineum* can be found in locations such as China's Shandong peninsula where temperatures range from -5 °C in the winter to 32 °C in the summer and with diurnal temperature differences rarely exceeding 10 °C (Guedes et al., 2015). Furthermore, diurnal temperature differences in the cold spring growing period of *E. salsugineum* are likely to be much more moderate than those experienced by *A. hierochuntica* during the warm Negev desert spring. Under such potentially extreme diurnal temperature variations, *A. hierochuntica* may possess a lower basal gene expression to allow for a greater relative transcriptome response during diurnal temperature shifts. Moreover, lower basal expression levels would require less energy to be expended in the low nutrient desert environment. Another possible difference between *A. hierochuntica* and its halophytic relatives relates to the type of stress challenge each species faces. *E. salsugineum* and *S. parvula* habitats possess levels of ions such as Na^+ and B^{3+} that are toxic to the majority of plant species and the two halophytes are constantly exposed to ionic stress throughout their life cycle. This situation might have led to the evolution of a stress-associated transcriptome that is continuously "switched-on". Conversely, *A. hierochuntica* is generally exposed to heat stress later in its life cycle and on a diurnal basis (unless there is a hamsin). Thus, evolution of a flexible, energy-conserving transcriptome that confers the ability to quickly react to the extreme daytime heat could be advantageous for adaptation to a desert environment.

Materials and methods

Plant material and growth conditions

F_4 generation *A. hierochuntica* seeds descended from a single seed from plants originally collected in the Negev Desert (Nahal Hayun, 30.191424N and 35.009926E), Israel, were used in this study.

Seeds for plants used to extract RNA for sequencing of the *de novo* reference transcriptome were germinated and grown on nutrient agar plates for 5 d in the growth room (16 h light (150 µmol photons m^{-2} s^{-1})/8 h dark; 22 °C), as described in Eshel et al. (2017). Plant material was prepared for sequencing on two platforms: (i) For Illumina sequencing, plate-grown seedlings were harvested and

snap-frozen in liquid nitrogen; (ii) For Roche 454 sequencing, plate-grown 5 day-old seedlings were transferred to pots containing autoclaved *A. thaliana* soil growth medium (Weizmann Institute of Science), irrigated to field capacity with 1 g l^{-1} 20-20-20 NPK + micronutrients solution (Haifa Chemicals), and kept in the growth room until plants developed four true fully-expanded leaves. These plants were then treated with the following conditions: (a) Control (field-capacity, 22 °C); (b) Drought stress (25% field capacity for 1 week); (c) Salt shock (200 mM NaCl in the fertilizer solution), harvested after 1, 3 and 6 h; (d) Heat shock (45 °C), harvested after 0.5, 1, and 2 h. Roots, shoots and flowers (where available) from these soil-grown plants, were harvested separately and snap-frozen in liquid nitrogen. In addition to these soil-grown samples, mature seeds, from the same F_4 generation seed stock, were imbibed in H_2O for 8.5 h, and then snap-frozen in liquid nitrogen.

For RNA-seq heat stress experiments, *A. thaliana* and *A. hierochuntica* were germinated and grown on nutrient agar plates according to Eshel et al. (2017). Seedlings were grown on plates until cotyledons were fully expanded before transfer to 7 cm x 7 cm x 8 cm pots containing *Arabidopsis* nitrogen-less soil (Weizmann Institute of Science; 70% fine peat [1–10mm], 30% perlite 4) irrigated to field capacity with a custom-made fertilizer solution (5 mM KNO_3, 2 mM $MgSO_4$, 1 mM $CaCl_2$ x $2H_2O$, 10 mM KH_2PO_4 [pH 6.0, adjusted with KOH] plus MS micronutrients (Murashige and Skoog, 1962). Flats containing pots were placed in the growth room under the same conditions as for the plate experiments. Flats were covered with plastic domes for 1 to 2 d, which were then gradually removed to allow seedlings to harden. Each day, pots were shuffled so that all plants received equal illumination and to remove shelf position effects. Plants were irrigated alternatively every 3 d with either fertilizer solution or water in order to maintain constant nutrient concentrations. After 6 d in the growth room, uniform plants were transferred to two growth chambers (KBWF 720, BINDER GmbH, Tuttlingen, Germany) (16 h light/8 h dark; 23 °C; 60% relative humidity) for heat treatments. The light/dark transitions at the beginning and end of the day comprised 0.5 h at 100 μmol photons m^{-2} s^{-1} and 0.5 h at 150 μmol photons m^{-2} s^{-1} to mimic sunrise and sunset. Light intensity for the remaining 14 h was 250 μmol m^{-2} s^{-1}. Plants were allowed to acclimate for 4 d and were moved randomly every day between the chambers, to avoid chamber effects. At day 10 after transfer to soil, (*A. hierochuntica* plants had two true leaves and *A. thaliana* had six true leaves), heat treatment was initiated in one chamber, keeping the other chamber as the control (23 °C). The heat treatment included 3 d at 40/25 °C, day/night temperatures, followed by two days of recovery at control conditions (Fig. 4A). Similar to the 1 h light transition, the temperature was also gradually increased/decreased for 1 h, between the light/dark states, to reach the appropriate temperatures. Plants from both chambers were harvested at

eight time points (Fig. 4A), either in the morning (1.5 h after the onset of the light/heat period) or at midday (7 h after the onset of the light/heat period). For each condition, three biological replicates comprising 6 pooled plants per replicate (27 samples per species) were used for downstream analyses.

Phylogenomics and positive selection analysis

To identify positively selected genes, which are unique to *A. hierochuntica* or common to extremophyte Brassicaceae species, we used coding sequences of the *A. hierochuntica* reference transcriptome and 16 sequenced Brassicaceae species (Supplementary Table S2) into the automated Agalma phylogenomics pipeline (Dunn et al., 2013).

The Agalma pipeline identified orthologous genes among these species by: (i) Identifying homologous genes among all input sequences from all the species using an all-by-all TBLASTX search followed by a Markov Clustering Algorithm (MCL) tool (Enright et al., 2002); (ii) For each homolog group, a peptide multiple sequence alignment (MSA) is produced using MAFFT with the E-INS-i algorithm (Katoh et al., 2005); (iii) The MSA is further used to build a maximum likelihood (ML) phylogenetic tree with RAxML v8.2.3 (Stamatakis, 2014). (iv) The homolog group tree is further pruned into maximally inclusive subtrees to define ortholog groups. MSAs of 13,806 ortholog groups with a sequence representation in at least 4 taxa were concatenated into a supermatrix for ML species tree search, using RAxML (with the PROTGAMMAWAG model of evolution, and 100 rapid bootstrap searches) under the WAG rate matrix (Whelan and Goldman, 2001), with gamma-distributed among-site rate variation.

To detect positive selection in the five extremophyte species, ortholog groups with sequence representation in at least 2 extremophytes, were selected to ensure sufficient statistical power (Anisimova et al., 2001). For each ortholog group, the peptide MSA was converted into the corresponding codon alignment using the pal2nal.pl program (Suyama et al., 2006), and the ML species tree was pruned using PHAST tree_doctor (Hubisz et al., 2011), to keep only sequence-represented taxa. Codon alignments together with pruned trees were further analyzed with the PAML v4.8, CODEML program (Yang, 1997; Yang, 2007), using the Branch-Site model. To test for positive selection, the tested branch(s) were labeled (foreground), and the log likelihood of two models (M1a and M2a), were calculated for each ortholog group. The difference between the two models is that in the M1a (null) model, the non-synonymous to synonymous rate ratio (dN/dS) is fixed to 1 (fix_omega = 1 and omega = 1), indicative of neutral selection, while in the M2a (alternative) model, the initial dN/dS ratio is set to 1, and is further estimated by the model (fix_omega = 0 and omega = 1). A

Likelihood Ratio Test was performed (with X^2 distribution), to identify genes with log likelihood values significantly different between the two models, indicative of deviation from neutral selection. Ortholog groups with portion of sites in the foreground branches, that had an estimated dN/dS ratio greater than 1, were considered under positive selection. To account for multiplicity, a Benjamini–Yekutieli false discovery rate (FDR) correction (Benjamini and Yekutieli, 2001) was applied using the "qvalue" R package, with a *q*-value < 0.05 cutoff for a gene to be considered as positively selected. Sites under positive selection were identified using the empirical Bayes approach with a posterior probability $p > 0.95$.

The above procedure was repeated to identify positive selected genes in *A. hierochuntica*, and in the other extremophyte species. For each analysis, different branches on the tree were tested (labeled as foreground) compared with all other branches (background): (i) labeling the external branches of all five extremophyte species as the foreground (4,723 ortholog groups); (ii) labeling the *A. hierochuntica* external branch as the foreground (3,093 ortholog groups); (iii) labeling the *E. salsugineum* external branch as the foreground (4,457 ortholog groups); (iv) labeling the *S. parvula* external branch as the foreground (4,369 ortholog groups); and (v) labeling the *A. thaliana* external branch as the foreground (5,513 ortholog groups). *A. thaliana* was considered as a control/comparator species sensitive to abiotic stresses (Kazachkova et al., 2018). The Venn diagram comparing positive selected genes (Fig. 2B) was generated using an online tool: http://bioinformatics.psb.ugent.be/webtools/Venn/).

To further assess the functionality of the positively selected genes, Gene Ontology (GO) terms were assigned to each ortholog group based on *A. thaliana* GO annotation. In cases where an ortholog group did not contain an *A. thaliana* ortholog, the closest *A. thaliana* homolog (best BLASTP hit) was used. Significant positively selected genes were further tested for enriched GO terms (Fisher's exact test, with a *q*-value < 0.05 cutoff) using the online AgriGO tool (Du et al., 2010; http://bioinfo.cau.edu.cn/agriGO/analysis.php), where the *A. thaliana* genome served as the background. Enriched GO terms with more than 2,000 genes in the *A. thaliana* genome were excluded, as these are broad and less informative terms. In addition, due to the hierarchical structure of the Gene Ontology, enriched GO terms often tend to be highly redundant. Therefore, GO terms that share the majority (>50%) of genes among them, were further clustered using the MCL algorithm (Wang et al., 2020b), with an inflation value of 3, and the GO term with the lowest enrichment *q*-value per cluster was chosen as a representative.

Reference transcriptome and RNA-seq

For sequencing the *A. hierochuntica* reference transcriptome, total RNA was extracted from 100 mg of ground tissue using the RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA was further treated with DNase I digestion. For RNA-seq, total RNA was extracted from 50 mg of ground tissue from *A. thaliana* or *A. hierochuntica* plants using the Plant/Fungi Total RNA Purification Kit (Norgen Biotek Corp. Ontario, Canada). RNA integrity, purity and concentrations were determined by NanoDrop spectrophotometer (NanoDrop ND-1000) and by denaturing formaldehyde agarose gel electrophoresis. Equal amounts of RNA from all samples designated for 454 sequencing of a normalized cDNA library were pooled into one sample, and dried on GenTegra™ RNA transport tubes (Integen X, Pleasanton, CA, USA). Samples for reference transcriptome Illumina sequencing were sent to the Glasgow Polyomics Facility at the University of Glasgow, UK, while the 454 samples were sent to GenePool genomics facility at the University of Edinburgh, UK. The reference transcriptome was assembled using a hybrid assembly approach that utilized both Illumina and 454 reads, and was annotated based on public databases (Supplemental Methods).

Samples for RNA-seq were delivered to the Roy J. Carver Biotechnology Center at the University of Illinois, Urbana-Champaign, USA. Libraries were prepared with the Illumina TruSeq Stranded mRNA Sample Prep Kit (Illumina) and 100 bp HiSeq2500 Illumina single-end reads were *de novo* assembled using the Trinity assembler (v 2.1.1, Grabherr et al., 2011). DEGs were identified using DESeq2 (Love et al., 2014). For full details of sequencing, read yields, quality control, transcriptome assembly, annotation, normalized gene expression estimations, and differential expression analysis, see Supplemental Methods. For raw read counts and DEGs identified in each species and for various functional groups, see Supplemental Table S19. Reference transcriptome and RNA-seq reads as well as the full assembled transcriptome are available via the NCBI SRA and TSA databases under BioProject PRJNA731383.

To assign ortholog pairs to the five idealized modes of expression in response to heat stress (Fig. 5A), Weighted Gene Co-expression Network Analysis (WGCNA) was employed to cluster normalized and quantified expression data into modules containing genes with similar expression profiles (Supplemental Fig. S5; Supplemental Table S13). DESeq2 was then used to call statistically significant DEGs. For direct comparison of absolute orthologous transcript levels, TPM values of minimum or maximum expression were analyzed for statistically significant difference ($p \leq 0.05$) with a Student *t*-test.

Validation of RNA-seq data

Scale-based normalization (SCBN)

To validate the differential gene expression analysis performed with DeSeq2, we re-normalized our RNA-seq read count data (normalized to transcript length) using the Scale-Based Normalization (SCBN) method that aids in removing systematic variation between different species (Zhou et al. 2019), For the first step - pinpointing highly conserved genes - we utilized the 17,962 *A. thaliana* and *A. hierochuntica* orthologous genes identified using the Agalma phylogenomics pipeline (see below "Phylogenomics and positive selection analysis"). Applying BLASTn (ver, 2.10.1) software revealed 109 highly conserved orthologs with an E-value $\leq$ 1e-^{100}, query coverage of $\geq$ 98%, and identical matches of $\geq$ 99%. The SCBN R package (http://www.bioconductor.org/packages/devel/bioc/html/SCBN.html) was then used on the 109 highly conserved genes to obtain a scaling factor of 0.9223461, which it then applied to the 17,962 orthologs to call 12,808 common differentially expressed genes ($p \leq$ 1e-^{05}) between the two species. To generate an approximate corrected gene count, the scaling factor was applied to each individual gene and the corrected average and median basal (control) expression is depicted in Fig. 9B.

Real-time QPCR analysis

Total RNA was extracted from whole shoots with TRIzol (38% Phenol (w/v), 0.8 M guanidine thiocyanate, 0.4 M ammonium thiocyanate, 0.1 M sodium acetate pH 5, 5% glycerol (v/v)) according to Rio et al. (2010). To remove residual genomic DNA, 7 µg of total RNA was treated with RNase-free PerfeCta DNase I (Quanta Biosciences, Inc., Gaithersburg, MD, USA) according to the manufacturer's instructions, and cDNA was synthesized from 1 µg of total RNA using the qScript™ cDNA Synthesis Kit (Quanta Biosciences). For amplification of PCR products, primers were designed using the NCBI Primer-BLAST tool (Ye et al., 2012), and analyzed for any secondary structure with the IDT OligoAnalyzer™ Tool (Supplementary Table S21). QPCR was performed with the ABI PRISM 7500 Sequence Detection System (Applied Biosystems). Each reaction contained 5 µl Applied Biosystems™ Power SYBR® Green PCR Master Mix (Thermo Fisher Scientific Inc., Waltham, MA, USA), 40 ng cDNA, and 300 nM of each gene-specific primer. The QPCR amplification protocol was: 95 °C for 60 s, 40 cycles of 95 °C for 5 s (denaturation) and 60 °C for 30 s (annealing/extension). Data were analyzed using the SDS 2.3 software (Applied Biosystems). To check the specificity of annealing of the primers, dissociation kinetics was performed at the end of each PCR run. All reactions were performed in triplicates. Relative quantification of target genes was calculated using the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001), using *A. thaliana eIF4A1* and the *A. hierochuntica eIF4A1* ortholog as

internal references. To ensure the validity of the $2^{-\Delta\Delta C}T$ method, standard curves of 2-fold serial dilutions of cDNA were created and amplification of efficiencies of target and reference gene were shown to be approximately equal. For absolute quantification of basal expression, QPCR products were gel-purified (Gel/PCR DNA Fragments Extraction Kit [Geneaid Biotech Ltd, New Tapei City, Taiwan]), and quantified with a Nanodrop spectrophotometer. Fivefold serial dilutions of each PCR product were used to create a standard curve for determination of transcript copy number. As a loading control, the absolute transcript copy number of *eIF4A1*was also calculated and normalized to the highest *eIF4A1*level, which was assigned a value of 1. The target gene transcript copy number was then adjusted for loading differences by dividing by the normalized *eIF4A1* level.

Supplemental Data

Supplemental Figure S1. Transcriptome sequencing and hybrid assembly workflow.

Supplemental Figure S2. GO term enrichment analysis of positively selected genes.

Supplemental Figure S3. An example of a three-day heat wave event in the Dead Sea valley during April 2008.

Supplemental Figure S4. Basal (control) expression of 15 orthologous *A. thaliana* and *A. hierochuntica* housekeeping genes.

Supplemental Figure S5. Clustering dendrogram of module eigenvalues for *A. thaliana* and *A. hierochuntica* transcriptome profiles under and heat stress conditions.

Supplemental Table S1. *A. hierochuntica* transcriptome functional annotation.

Supplemental Table S2. Species included in the custom *Brassicaceae* CDS database in the *Brassicaceae* phylogenomic analysis.

Supplemental Table S3. CODEML positive selected genes (q<0.05) in the all extremophyte species run with the branch-site model.

Supplemental Table S4. CODEML positive selected genes (q<0.05) in *A. hierochuntica* run with the branch-site model.

Supplemental Table S5. CODEML positive selected genes (q<0.05) in *E. salsugineum* run with the branch-site model.

Supplemental Table S6. CODEML positive selected genes (q<0.05) in *S. parvula* run with the branch-site model.

Supplemental Table S7. CODEML positive selected genes (q<0.05) in *A. thaliana* run with the branch-site model.

Supplemental Table S8. GO terms over represented in the all extremophyte species 194 positive selected genes.

Supplemental Table S9. GO terms over represented in *A. hierochuntica* 120 positive selected genes.

Supplemental Table S10. GO terms over represented in *E. salsugineum* 131 positive selected genes.

Supplemental Table S11. GO terms over represented in *S. parvula* 99 positive selected genes.

Supplemental Table S12. GO terms over represented in *A. thaliana* 112 positive selected genes.

Supplemental Table S13. Assignment of ortholog pairs modes of expression ($\log_2$ fold-change data and WGCNA modules).

Supplemental Table S14. Early heat response modules (rlog expression data).

Supplemental Table S15. *A. thaliana* early heat module (Thistle1) gene list.

Supplemental Table S16. *A. thaliana* late heat module (Ivory) gene list.

Supplemental Table S17. *A. hierochuntica* early heat module (Black) gene list.

Supplemental Table S18. *A. hierochuntica* early heat module (Floralwhite) gene list.

Supplemental Table S19. *A. thaliana* and *A. hierochuntica* raw read data plus GO-terms, and DEGs.

Supplemental Table S20. 109 most conserved orthologs (≥ 98% query coverage, ≥ 99% nucleotide identity, ≥ BLAST value of e-100) between *A. thaliana* and *A. hierochuntica***.**

Supplemental Table S21. PCR primers used in this study.

Supplemental Methods.

Acknowledgements

We would like to dedicate this paper to Dr. Dirk Hincha, who passed away in 2020. A great plant stress scientist, colleague and friend. He will be sorely missed. We thank Ruth Shaked and Beery Yaakov for their dedicated technical support. We are also grateful to Patrick Barko from the University of Illinois Urbana-Champaign for excellent aid with WGCNA. Our appreciation to Alvaro Hernandez and all the team at the Roy J. Carver Biotechnology Center, University of Illinois Urbana-Champaign for superb sequencing services. This work was supported by the Goldinger Trust Jewish Fund for the Future, the Koshland Foundation for Support of Interdisciplinary Research In Combatting Desertification, and the I-CORE Program of the Planning and Budgeting Committee to S.B., National Science Foundation award MCB 1616827, Next-Generation BioGreen21 Program of Republic of Korea (PJ01317301).to M.D., and NSF-BSF-IOS-EDGE 1923589/2019610 to M.D. and S.B. G.E. was

supported by an Israel President Fellowship for Excellence and Scientific Innovation award, N.D was supported by a Ben-Gurion University, Kreitman School for Advanced Research Studies High-tech, Bio-tech and Chemo-tech award, and G.W. was supported by an Economic Development Assistantship award from Louisiana State University, The authors also acknowledge the LSU High Performance Computing Services, and the BGU Bioinformatics Core Facility for providing computational resources needed for data analyses.

References

Aguirrezabal L, Bouchier-Combaud S, Radziejwoski A, Dauzat M, Cookson SJ, Granier C (2006) Plasticity to soil water deficit in *Arabidopsis thaliana*: dissection of leaf development into underlying growth dynamic and cellular variables reveals invisible phenotypes. Plant Cell Environ 29: 2216-2227

Ahmad M, Jarillo JA, Klimczak LJ, Landry LG, Peng T, Last RL, Cashmore AR (1997) An enzyme similar to animal type II photolyases mediates photoreactivation in *Arabidopsis*. Plant Cell 9: 199-207

Ali A, Raddatz N, Aman R, Kim S, Park HC, Jan M, Baek D, Khan IU, Oh DH, Lee SY, (2016) A single amino-acid substitution in the sodium transporter HKT1 associated with plant salt tolerance. Plant Physiol 171: 2112-2126

Ali A, Khan IU, Jan M, Khan HA, Hussain S, Nisar M, Chung WS, Yun DJ (2018) The high-affinity potassium transporter EpHKT1;2 from the extremophile *Eutrema parvula* mediates salt tolerance. Front Plant Sci 9: 1108

Ali Z, Park HC, Ali A, Oh DH, Aman R, Kropornicka A, Hong H, Choi W, Chung WS, Kim WY et al. (2012) TsHKT1;2, a HKT1 homolog from the extremophile *Arabidopsis* relative *Thellungiella salsuginea*, shows K^+ specificity in the presence of NaCl. Plant Physiol 158: 1463-1474

Amtmann A (2009) Learning from evolution: *Thellungiella* generates new knowledge on essential and critical components of abiotic stress tolerance in plants. Mol Plant 2: 3-12

Anisimova M, Bielawski JP, Yang Z (2001) Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol Biol Evol 18: 1585-1592

Azocar A, Rada F, Goldstein G (1988) Freezing tolerance in *Draba chionophila*, a 'miniature' caulescent rosette species. Oecologia 75: 56-60

Bao F, Azhakanandam S Franks RG (2010) SEUSS and SEUSS-LIKE transcriptional adaptors regulate floral and embryonic development in *Arabidopsis*. Plant Physiol 152; 821-836

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under depencency. Ann of Stat 29: 1165-1188

Birkeland S, Gustafsson ALS, Brysting AK, Brochmann C, Nowak MD (2020) Multiple genetic trajectories to extreme abiotic stress adaptation in Arctic Brassicaceae. Mol Biol Evol 37: 2052-2068

Bondel KB, Nosenko T, Stephan W (2018) Signatures of natural selection in abiotic stress-responsive genes of *Solanum chilense*. R. Soc open sci 5: 171198

Borrell A, Cutanda MC, Lumbreras V, Pujal J, Goday A, Culiáñez-Macià FA, Pagès M (2002) *Arabidopsis thaliana* Atrab28: A nuclear targeted protein related to germination and toxic cation tolerance. Plant Mol Biol 50: 249-259

Bressan R, Park H, Orsini F, Oh DH, Dassanayake M, Inan G, Yun J-Y, Bohnert H.J, Maggio A (2013) Biotechnology for mechanisms that counteract salt stress in extremophile species: a genome-based view. Plant Biotechnol Rep. 7: 27-37

Brugnoli E, Lauteri M (1991) Effects of salinity on stomatal conductance, photosynthetic capacity, and carbon isotope discrimination of salt-tolerant (*Gossypium hirsutum* L.) and salt-sensitive (*Phaseolus vulgaris* L.) C_3 Non-Halophytes. Plant Physiol 95: 628-635

Bu S-L, Liu C, Liu N, Zhao J-L, Ai L-F, Chi H, Li KL, Chien C-W, Burlingame AL, Zhang S-W (2016) Immunopurification and mass spectrometry identifies protein phosphatase 2A (PP2A) and BIN2/GSK3 as regulators of AKS transcription factors in *Arabidopsis*. Mol Plant 10: 345-348

Cassan-Wang H, Goué N, Saidi MN, Legay S, Sivadon P, Goffner D, Grima-Pettenati J (2013) Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. Front Plant Sci 4: 189

Chaves MM, Flexas J, Pinheiro C (2009) Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. Ann Bot 103: 551-560

Cheeseman JM (2015) The evolution of halophytes, glycophytes and crops, and its implications for food security under saline conditions. New Phytol 206: 557-570

Chen X, Goodwin SM, Liu X, Chen X, Bressan RA, Jenks MA (2005) Mutation of the RESURRECTION1 locus of *Arabidopsis* reveals an association of cuticular wax with embryo development. Plant Physiol 139: 909-919

Chen C, Letnik I, Hacham Y, Dobrev P, Ben-Daniel B-H, Vanková R, Amir R, Miller G (2014) ASCORBATE PEROXIDASE6 protects *Arabidopsis* desiccating and germinating seeds from

stress and mediates cross talk between reactive oxygen species, abscisic acid, and auxin. Plant Physiol 166: 370-383

Chen H, Ruan J, Chu P, Fu W, Liang Z, Li Y, Tong J, Xiao L, Liu J, Li C (2020) AtPER1 enhances primary seed dormancy and reduces seed germination by suppressing the ABA catabolism and GA biosynthesis in *Arabidopsis* seeds. Plant J 101: 310-323

Creighton MT, Kolton A, Kataya ARA, Maple-Grødem J, Averkina IO, Behzad Heidari B, Lillo C (2017) Methylation of protein phosphatase 2A-Influence of regulators and environmental stress factors. Plant Cell Environ 40: 2347-2358

Crevillen P, Dean C (2011) Regulation of the floral repressor gene FLC: the complexity of transcription in a chromatin context. Curr Opin Plant Biol 14: 38-44

Dassanayake M, Haas JS, Bohnert HJ, Cheeseman JM (2010) Comparative transcriptomics for mangrove species: an expanding resource. Funct Integr Genomics 10: 523-532

Dassanayake M, Oh DH, Haas JS, Hernandez A, Hong H, Ali S, Yun DJ, Bressan RA, Zhu JK, Bohnert HJ, (2011a) The genome of the extremophile crucifer *Thellungiella parvula*. Nat Genet 43: 913-918

Dassanayake M, Oh DH, Hong H, Bohnert HJ, Cheeseman JM (2011b) Transcription strength and halophytic lifestyle. Trends Plant Sci 16: 1-3

Devaiah BN, Karthikeyan AS, Raghothama KG (2007) WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *Arabidopsis*. Plant Physiol 143: 1789-1801

Devireddy AR, Arbogast J, Mittler R (2020) Coordinated and rapid whole plant systemic stomatal response. New Phytol 225: 21-25

Di Rubbo S, Irani NG, Russinova E (2011) PP2A Phosphatases: The “on-off” regulatory switches of brassinosteroid signaling. Sci Signal 4: pe25

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: A GO analysis toolkit for the agricultural community. Nucleic Acids Res 38: 64-70

Dunn CW, Howison M, Zapata F (2013) Agalma: an automated phylogenomics workflow. BMC Bioinformatics 14: 330

Enright A J, Von Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. Nuclei Acids Res 30: 1575-1584

Eppel A, Shaked R, Eshel G, Barak S, Rachmilevitch S (2014) Low induction of non-photochemical quenching and high photochemical efficiency in the annual desert plant *Anastatica hierochuntica*. Physiol Plant 151: 544-558

Eshel G, Shaked R, Kazachkova Y, Khan A, Eppel A, Cisneros A, Acuna T, Gutterman Y, Tel-Zur N, Rachmilevitch S et al. (2017) *Anastatica hierochuntica*, an *Arabidopsis* desert relative, is tolerant to multiple abiotic stresses and exhibits species-specific and common stress tolerance strategies with its halophytic relative, *Eutrema* (*Thellungiella*) *salsugineum*. Front Plant Sci 7: 1992

Exner V, Taranto P, Schonrock N, Gruissem W, Hennig L (2006) Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development. Development 133: 4163-4172

Fan Y, van den Dool H (2008) A global monthly land surface air temperature analysis for 1948-present. J Geophys Res 113: D01103

Farrant JM, Cooper K, Hilgart A, Abdalla KO, Bentley J, Thomson JA, Dace HJW, Peton N, Mundree SG, Rafudeen MS (2015) A molecular physiological review of vegetative desiccation tolerance in the resurrection plant *Xerophyta viscosa* (Baker). Planta 242: 407-426

Feng J, Lu J (2017) LHP1 could act as an activator and a repressor of transcription in plants. Front Plant Sci 8: 2041

Fischer I, Camus-Kulandaivelu L, Allal F, Stephan W (2011) Adaptation to drought in two wild tomato species: the evolution of the *Asr* gene family. New Phytol 190: 1032-1044

Franzke A, Lysak MA, Al-Shehbaz IA, Koch MA, Mummenhoff K (2011) Cabbage family affairs: The evolutionary history of Brassicaceae. Trends Plant Sci 16: 108-116

Friedman J, Stein Z (1980) The influence of seed dispersal mechanisms on the distribution and survival of *Anastatica hierochuntica* (Cruciferae) in the Negev Desert. J Ecol 68: 43-50

Friedman J, Stein Z, Rushkin E (1981) Drought tolerance of germinating seeds and young seedlings of *Anastatica hierochuntica* L. Oecologia 51: 400-403

Gaudin V, Libault M, Pouteau S, Juul T, Zhao G, Lefebvre D, Grandjean O (2001) Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in *Arabidopsis*. Development 128: 4847-4858

Gong Q, Li P, Ma S, Indu Rupassara S, Bohnert HJ (2005) Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. Plant J 44: 826-839

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644-52

Guedes JDA, Jin G, Bocinsky RK (2015) The impact of climate on the spread of rice to north-eastern China: A new look at the data from Shandong Province. PLoS One 10: 1-19

Guevara DR, Champigny MJ, Tattersall A, Dedrick J, Wong CE, Li Y, Labbe A, Ping CL, Wang Y, Nuin P, et al. (2012) Transcriptomic and metabolomics analysis of Yukon *Thellungiella* plants grown in cabinets and their natural habitat show phenotypic plasticity. BMC Plant Biol 12: 175

Guo P, Li Z, Huang P, Li B, Fang S, Chu J, Guo H (2017) A tripartite amplification loop involving the transcription factor WRKY75, salicylic acid, and reactive oxygen species accelerates leaf senescence. Plant Cell 29: 2854-2870

Gutterman Y. (1998) Ecophysiological genotypic and phenotypic strategies affecting seed ‘‘readiness to germinate’’ in plants occurring in deserts. In: Taylor, AG., Huang, X.-L. (Eds.), Progress in Seed Research: Proceedings of Second International Conference on Seed Science and Technology. Communication services of the New York State Agricultural Experiment Station, Geneva, NY, pp. 348

Gutterman Y (2002) Survival Strategies of Annual Desert Plants (Adaptations of Desert Organisms). Heidelberg: Springer

Hake S, Smith HM, Holtan H, Magnani E, Mele G, Ramirez J (2004) The role of *knox* genes in plant development. Annu Rev Cell Dev Biol 20: 125-151

Hanikenne, M, Talke IN, Haydon M., Lanz C, Nolte A, Motte P et al. (2008) Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature 453: 391-396

Helvaci C, Mordogan H, Çolak M, Gündogan I. (2004) Presence and Distribution of Lithium in Borate Deposits and Some Recent Lake Waters of West-Central Turkey. *International Geology Review* **46**: 177-190.

Hisanaga T, Ferjani A, Horiguchi G, Ishikawa N, Fujikura U, Kubo M, Demura T, Fukuda H, Ishida T, et al. (2013) The ATM-dependent DNA damage response acts as an upstream trigger for compensation in the *fas1* mutation during *Arabidopsis* leaf development. Plant Physiol 162: 831-841

Holzapfel, C. (2008) Deserts. In S.E. Jørgensen & B.D. Fath (eds.), Vol.2 of Encyclopedia of Ecology, pp. 879-898. Oxford: Elsevier.

Hong SW, Vierling E (2000) Mutants of *Arabidopsis thaliana* defective in the acquisition of tolerance to high temperature stress. Proc Natl Acad Sci USA 97: 4392-4397

Honjo MN, Kudoh H (2019) *Arabidopsis halleri*: a perennial model system for studying population differentiation and local adaptation. AoB Plants 11: plz076

Hu Y, Wu Q, Sprague SA, Park J, Oh M, Rajashekar CB, Koiwa H, Nakata PA, Cheng N, Hirschi KD, et al. (2015) Tomato expressing *Arabidopsis* glutaredoxin gene *AtGRXS17* confers tolerance to chilling stress via modulating cold responsive components. Hortic Res 2: 15051

Huang J, Zhao X, Chory J (2019) The *Arabidopsis* transcriptome responds specifically and dynamically to high light stress. Cell Rep 29: 4186-4199

Hubisz MJ, Pollard KS, Siepel A (2011) Phastand Rphast: Phylogenetic analysis with space/time models. Brief Bioinformatics 12: 41-51

Ito S, Song YH, Josephson-Day AR, Miller RJ, Breton G, Richard G. Olmstead RG, Imaizumi T (2012) FLOWERING BHLH transcriptional activators control expression of the CONSTANS in *Arabidopsis*. Proc Natl Acad Sci USA 109: 3583-3587

Jarvis DE, Ryu C-H, Beilstein MA, Schumaker KS (2014) Distinct roles for SOS1 in the convergent evolution of salt tolerance in *Eutrema salsugineum* and *Schrenkiella parvula*. Mol Biol Evol 31: 2094-2107

Jiang C-Z, Yee J, Mitchell DL, Britt AB (1997) Photorepair mutants of *Arabidopsis*. Proc Natl Acad Sci USA 94: 7441-7445

John UP. Spangenberg GC. (2005) Xenogenomics: genomic bioprospecting in indigenous and exotic plants through EST discovery, cDNA microarray-based expression profiling and functional genomics. Comp. Funct. Genomics 6: 230-235

Kant S, Kant P, Raveh E, Barak S (2006) Evidence that differential gene expression between the halophyte, *Thellungiella halophila*, and *Arabidopsis thaliana* is responsible for higher levels of the compatible osmolyte proline and tight control of Na^+ uptake in *T. halophila*. Plant, Cell Environ 29: 1220–1234

Kant S, Bi YM, Weretilnyk E, Barak S, Rothstein SJ (2008a) The *Arabidopsis* halophytic relative *Thellungiella halophila* tolerates nitrogen-limiting conditions by maintaining growth, nitrogen uptake, and assimilation. Plant Physiol 147: 1168–1180

Kant P, Gordon M, Kant S, Zolla G, Davydov O, Heimer YM, Chalifa-Caspi V, Shaked R, Barak S (2008b) Functional genomics-based identification of genes that regulate *Arabidopsis* responses to multiple abiotic stresses. Plant Cell Environ 31: 697-714

Katoh K, Kuma KI, Toh H, Miyata T (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33: 511-518

Kaya H, Shibahara K, Taoka K, Iwabuchi M, Stillman B, Araki T (2001) *FASCIATA* genes for Chromatin Assembly Factor-1 in *Arabidopsis* maintain the cellular organization of apical meristems. Cell 104: 131-142

Kazachkova Y, Batushansky A, Cisneros A, Tel-Zur N, Fait A, Barak S (2013) Growth platform-dependent and -independent phenotypic and metabolic responses of *Arabidopsis* and its halophytic relative, *Eutrema salsugineum*, to salt stress. Plant Physiol 162: 1583-1598

Kazachkova Y, Eshel G, Pantha P, Cheeseman JM, Dassanayake M, Barak S (2018) Halophytism: What have we learnt from *Arabidopsis thaliana* relative model systems? Plant Physiol 178: 972-988

Kiefer M, Schmickl R, German DA, Mandakova T, Lysak MA, Al-Shehbaz IA, Franzke A, Mummenhoff K, Stamatakis A, Koch MA (2014) BrassiBase: Introduction to a novel knowledge database on brassicaceae evolution. Plant Cell Physiol 55: 1-9

Kimura S, Sakaguchi K (2006) DNA Repair in Plants. Chem Rev 106: 753-766

Koch M, Al-Shehbaz IA (2002) Molecular data indicate complex intra- and intercontinental differentiation of American *Draba* (Brassicaceae). Ann Missouri Bot Gard 89: 88-109

Koenig D, Weigel D (2015) Beyond the thale: comparative genomics and genetics of *Arabidopsis* relatives. *Nat Rev Genet* 16**,** 285–298.

Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K (2003) *Arabidopsis* TERMINAL FLOWER 2 gene encodes a Heterochromatin Protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes.

Kramer, U (2010) Metal hyperaccumulation in plants. Annu Rev Plant Biol 61: 517-534

Landry G, Stapleton AE, Lim J, Hoffman P, Hays JB, Walbot V, Last RL (1997) An *Arabidopsis* photolyase mutant is hypersensitive to ultraviolet-B radiation. Proc Natl Acad Sci USA 94: 328-332

Lange H, Ndecky SYA, Gomez-Diaz C. Pflieger D, Butel N, Zumsteg J, Kuhn L, Piermaria C, Chicher J, Christie M, et al. (2019) RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in *Arabidopsis*. Nat Commun **10:** 3871

Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 9: 559

Laporte D, Olate E, Salinas P, Salazar M, Jordana X, Holuigue L (2012) Glutaredoxin GRXS13 plays a key role in protection against photooxidative stress in *Arabidopsis*. J Exp Bot 63: 503-515

Lario LD, Ramirez-Parra E, Gutierrez C, Spampinato CP, Casati P (2013) ANTI-SILENCING FUNCTION1 proteins are involved in ultraviolet-induced DNA damage repair and are cell cycle regulated by E2F transcription factors in *Arabidopsis*. Plant Physiol 162: 1164-1177

Lawlor D. (1995) The effects of water deficit on photosynthesis. In *Environment and Plant Metabolism* (ed. N. Smirnoff), pp. 129-160. Bios Scientific Publishers, Oxford, UK

Lawson T, Davey PA, Yates SA, Bechtold U, Baeshen M, Baeshen N, Mutwakil MZ, Sabir J, Baker NR, Mullineaux PM (2014) C_3 photosynthesis in the desert plant *Rhazya stricta* is fully functional at high temperatures and light intensities. New Phytol 201: 862-873

Lee JE, Lampugnani ER, Bacic A, Golz JF (2014) SEUSS and SEUSS-LIKE 2 coordinate auxin distribution and KNOXI activity during embryogenesis. Plant J 80: 122-135

Leyser HMO, Furner IJ (1992) Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana*. Development 116: 397-403

Li H, He Z, Lu G, Lee SC, Alonso J, Ecker JR, Sheng Luan (2007) A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in *Arabidopsis*. Plant Cell 19: 2403-2416

Li Z, He Y (2010) Roles of brassinosteroids in plant reproduction. Int J Mol Sci 21: 872

Li H, Luan S (2011) The cyclophilin AtCYP71 interacts with CAF-1 and LHP1 and functions in multiple chromatin remodeling processes. Mol Plant 4: 748-758

Li J. Liu J, Wang G, Cha J-Y, Li G, Chen S., Li Z, Guo J, Zhang C, Yang Y et al. (2015). A chaperone function of NO CATALASE ACTIVITY1 is required to maintain catalase activity and for multiple stress responses in *Arabidopsis*. Plant Cell 27: 908-925

Li B, Gao Z, Liu X, Sun D, Tang W (2019) Transcriptional profiling reveals a time-of-day-specific role of REVEILLE 4/8 in regulating the first wave of heat shock-induced gene expression in *Arabidopsis*. Plant Cell 31: 2353-2369

Liang X, Ding P, Lian K, Wang J, Ma M, Li L, Li L, Li M, Zhang X, She Chen S, et al. (2016) *Arabidopsis* heterotrimeric G proteins regulate immunity by directly coupling to the FLS2 receptor. eLife 5: e13568

Lindquist S (1986) The heat shock response. Annu Rev Biochem 55: 1151-1191

Liu Y, Liu N, Deng X, Liu D, Li M, Cui D, Hu Y, Yan Y (2020) Genome-wide analysis of wheat DNA-binding with one finger (Dof) transcription factor genes: evolutionary characteristics and diverse abiotic stress responses. BMC Genomics 21: 276

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods 25*:* 402-408.

Lopez L, Wolf EM, Pires JC, Edger PP, Koch MA (2017) Resources from transcriptomes in the Brassicaceae family. Front Plant Sci 8: 1488

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550

Lu C, Tian Y, Wang S, Su Y, Mao T, Huang T, Chen Q, Xu Z, Ding Y (2017) Phosphorylation of SPT5 by CDKD;2 is required for VIP5 recruitment and normal flowering in *Arabidopsis thaliana*. Plant Cell 29: 277-291

Lugan R, Niogret MF, Leport L, Guegan J-P, Larher FR, Savouré A, Kopka J, Bouchereau A (2010) Metabolome and water homeostasis analysis of *Thellungiella salsuginea* suggests that dehydration tolerance is a key response to osmotic stress in this halophyte. Plant J 64: 215-229

Mang HG, Laluk KA, Parsons EP, Kosma DK, Cooper BR, Park HC, AbuQamar S, Boccongelli C, Miyazaki S, Consiglio F, et al. (2009) The *Arabidopsis* RESURRECTION1 gene regulates a novel antagonistic interaction in plant defense to biotrophs and necrotrophs. Plant Physiology, 151: 290-305.

Mboup M, Fischer I, Lainer H, Stephan W (2012) Trans-species polymorphism and allele-specific expression in the *CBF* gene family of wild tomatoes Mol Biol Evol 29: 3641-3652

Mello JA, Sillje HH, Roche DM, Kirschner DB, Nigg EA, Almouzni G (2002) Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. EMBO Rep 3: 329-334

Mody T, Titouan Bonnot T, Nagel DH (2020) Interaction between the circadian clock and regulators of heat stress responses in plants. Genes 11: 156

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15: 473-497

Nekrasov V, Li J, Batoux M, Roux M, Chu ZH, Lacombe S, Rougon A, Bittel P, Kiss-Papp M, Chinchilla D, et al. (2009) Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. EMBO J 28: 3428-3438

Nie X, Wang H, Li J, Holec S, Berger F (2014) The HIRA complex that deposits the histone H3.3 is conserved in *Arabidopsis* and facilitates transcriptional dynamics. Biology Open 3: 794-802

Nolan TM, Vukasinovic N, Liu D, Russinova E, Yina Y (2020) Brassinosteroids: Multidimensional regulators of plant growth, development, and stress responses. Plant Cell 32: 295-318

Obertello M, Krouk G, Katari MS, Runko SJ, Coruzzi GM (2010) Modeling the global effect of the basic-leucine zipper transcription factor 1 (bZIP1) on nitrogen and light regulation in *Arabidopsis*. BMC Systems Biol 4: 111

Oh S, Zhang H, Ludwig P, van Nocker S (2004) A mechanism related to the yeast transcriptional regulator Paf1c is required for expression of the *Arabidopsis* FLC/MAF MADS box gene family. Plant Cell 16: 2940-2953

Oh DH, Gong Q, Ulanov A, Zhang Q, Li Y, Ma W, Yun DJ, Bressan RA, Bohnert HJ (2007) Sodium stress in the halophyte *Thellungiella halophile* and transcriptional changes in a *thsos1*-RNA interference line. J Integr Plant Biol 49: 1484-1496

Oh DH, Leidi E, Zhang Q, Hwang SM, Li Y, Quintero FJ, Jiang X, D'Urzo MP, Lee SY, Zhao Y, (2009) Loss of halophytism by interference with *SOS1* expression. Plant Physiol 151: 210-222

Oh DH, Dassanayake M, Haas JS, Kropornika A, Wright C, D'Urzo MP, Hong H, Ali S, Hernandez A, Lambert GM, et al (2010) Genome structures and halophyte-specific gene expression of the extremophile *Thellungiella parvula* in comparison with *Thellungiella salsuginea* (*Thellungiella halophila*) and *Arabidopsis*. Plant Physiol 154: 1040−1052

Oh DH, Dassanayake M, Bohnert HJ, Cheeseman JM (2012) Life at the extreme: lessons from the genome. Genome Biol 13: 241

Oh DH, Hong H, Lee SY, Yun DJ, Bohnert HJ, Dassanayake M (2014) Genome structures and transcriptomes signify niche adaptation for the multiple-ion-tolerant extremophyte *Schrenkiella parvula*. Plant Physiol 164: 2123-2138

Oh DH, Dassanayake M (2019) Landscape of gene transposition-duplication within the Brassicaceae family. DNA Res 26: 21-36

Oscar MA, Barak S, Winters G (2018) The tropical invasive seagrass, *Halophila stipulacea*, has a superior ability to tolerate dynamic changes in salinity levels compared to its freshwater relative, *Vallisneria americana*. Front Plant Sci 9: 950

Ozfidan-Konakci C, Uzilday B, Ozgur R, Yildiztugay E, Sekmen AH, Turkan I (2016) Halophytes as a source of salt tolerance genes and mechanisms: a case study for the Salt Lake area, Turkey. Funct Plant Biol 43: 575-589

Pang Q, Chen S, Dai S, Chen Y, Wang Y, Yan X (2010) Comparative proteomics of salt tolerance in *Arabidopsis thaliana* and *Thellungiella halophila*. J Proteome Res 9: 2584-2599

Para A, Li Y, Marshall-Colon A, Varala K, Francoeur NJ, Moran TM, Edwards MB, Hackley C, Bargmann BOR, Birnbaum KD (2014) Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in *Arabidopsis*. Proc Natl Acad Sci 111: 10371-10376

Park MY, Kang JY, Kim SY (2011). Overexpression of *AtMYB52* confers ABA hypersensitivity and drought tolerance. Molecules and Cells 31: 447-454

Patel S, Dinesh-Kumar SP (2008) *Arabidopsis* ATG6 is required to limit the pathogen-associated cell death response. Autophagy 4: 20-27

Pilarska M, Wiciarz M, Jajić I, Kozieradzka-Kiszkurno M, Dobrev P, Vankova R, Niewiadomska E (2016) A different pattern of production and scavenging of reactive oxygen species in halophytic *Eutrema salsugineum* (*Thellungiella salsuginea*) plants in comparison to *Arabidopsis thaliana* and its relation to salt stress signaling. Front Plant Sci 7: 1179

Pouteau S, Albertini C (2009) The significance of bolting and floral transitions as indicators of reproductive phase change in *Arabidopsis*. J Exp Bot 60: 3367-3377

Rio DC, Ares Jr M, Hannon GJ, Nilsen TW (2010) Purification of RNA Using TRIzol (TRI Reagent). Cold Spring Harb Protoc; doi:10.1101/pdb.prot5439

Rizhsky L, Liang H, Mittler R (2002) The combined effect of drought stress and heat shock on gene expression in tobacco. Plant Physiol 130: 1143-1151

Rizza A, Boccaccini A, Lopez-Vidriero I, Costantino P, Vittorioso P (2011) Inactivation of the *ELIP1* and *ELIP2* genes affects *Arabidopsis* seed germination. New Phytol 190: 896–905

Serrano M, Guzmán P (2004) Isolation and gene expression analysis of *Arabidopsis thaliana* mutants with constitutive expression of ATL2, an early elicitor-response RING-H2 zinc-finger gene. Genetics, 167: 919-929

Shabala S (2013) Learning from halophytes: physiological basis and strategies to improve abiotic stress tolerance in crops. Ann Bot 112: 1209-1221

Sharma N, Cram D, Huebert T, Zhou N Parkin IAP (2007) Exploiting the wild crucifer *Thlaspi arvense* to identify conserved and novel genes expressed during a plant's response to cold stress. Plant Mol Biol 63: 171-184

Sharp JA, Fouts ET, Krawitz DC, Kaufman D (2001) Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. Curr Biol 11: 463-473

Shrestha B, Guragain B, Sridhar VV (2014) Involvement of co-repressor LUH and the adapter proteins SLK1 and SLK2 in the regulation of abiotic stress response genes in *Arabidopsis*. BMC Plant Biol 14: 54

Shi D, Ren A, Tang X, Qi G, Xu Z, Chai G, Hu R, Zhou G, Kong Y (2018) MYB52 negatively regulates pectin demethylesterification in seed coat mucilage. Plant Physiol 176: 2737-2749

Simão FA Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210-3212

Skirycz A, De Bodt S, Obata T, De Clercq I, Claeys H, De Rycke R, Andriankaja M, Van Aken O, Van Breusegem F, Fernie AR, et al (2010) Developmental stage specificity and the role of mitochondrial metabolism in the response of *Arabidopsis* leaves to prolonged mild osmotic stress. Plant Physiol 152: 226-244

Skirycz A, Claeys H, De Bodt S, Oikawa A, Shinoda S, Andriankaja M, Maleux K, Eloy NB, Coppens F, Yoo S-D, et al (2011) Pause-and-stop: the effects of osmotic stress on cell proliferation during early leaf development in *Arabidopsis* and a role for ethylene signaling in cell cycle arrest. Plant Cell 23: 1876-1888

Song YH, Kubota A, Kwon MS, Covington MF, Lee N, Taagen ER, Cintron DL, Hwang DY, Akiyama R, Hodge SK, et al. (2013) Molecular basis of flowering under natural long-day conditions in *Arabidopsis*. Nat Plants 4: 824-835

Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312-1313

Stepien P, Johnson GN (2009) Contrasting responses of photosynthesis to salt stress in the glycophyte *Arabidopsis* and the halophyte *Thellungiella*: role of the plastid terminal oxidase as an alternative electron sink. Plant Physiol 149: 1154-1165

Sun W, Bernard C, van de Cotte B, Van Montagu M, Verbruggen N (2001) *At-HSP17.6A*, encoding a small heat-shock protein in *Arabidopsis*, can enhance osmotolerance upon overexpression. Plant J 27: 407-415

Sun Q, Gao F, Zhao L, Li K, Zhang J (2010) Identification of a new 130 bp *cis*-acting element in the *TsVP1* promoter involved in the salt stress response from *Thellungiella halophila*. BMC Plant Biol 10: 90

Sun X, Li Y, Cai H, Bai X, Ji W, Ding X, Zhu Y.(2012) The *Arabidopsis* AtbZIP1 transcription factor is a positive regulator of plant tolerance to salt, osmotic and drought stresses. J Plant Res: 125, 429-438

Suyama M, Torrents D, Bork P (2006) PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res, 34(Web Server Issue) W609-612

Taji T, Seki M, Satou M, Sakurai T, Kobayashi M, Ishiyama K, Narusaka Y, Narusaka M, Zhu JK, Shinozaki K (2004) Comparative genomics in salt tolerance between *Arabidopsis* and *Arabidopsis*-related halophyte salt cress using *Arabidopsis* microarray. Plant Physiol 135: 1697-1709

Taji T, Komatsu K, Katori T, Kawasaki Y, Sakata Y, Tanaka S, Kobayashi M, Toyoda A, Seki M, Shinozaki K (2010) Comparative genomic analysis of 1047 completely sequenced cDNAs from an *Arabidopsis*-related model halophyte, *Thellungiella halophila*. BMC Plant Biol 10: 261

Takahashi Y, Ebisu Y, Kinoshita T, Doi M, Okuma E, Murata Y, Shimazaki K (2013) bHLH transcription factors that facilitate K^+ uptake during stomatal opening are repressed by abscisic acid through phosphorylation. Sci Signal 6: ra48

Thalhammer A, Hincha DK (2014) A mechanistic model of COR15 protein function in plant freezing tolerance: integration of structural and functional characteristics. Plant Signal Behav, 9: 12, e977722

Thompson TL, Zaady E, Huancheng P, Wilson TB, Martens DA (2006) Soil C and N pools in patchy shrublands of the Negev and Chihuahuan Deserts. Soil Biol Biochem 38: 1943-1955

Tyler JK, Collins KA, Prasad-Sinha J, Amiott E, Bulger M, Harte PJ, Kobayashi R, Kadonaga, JT (2001) Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors. Mol Cell Biol 21: 6574-6584

Velasco VME, Mansbridge J, Bremner S, Carruthers K, Summers PS, Sung WWL, Champigny MJ, Weretilnyk EA (2016) Acclimation of the crucifer *Eutrema salsugineum* to phosphate limitation is associated with constitutively high expression of phosphate-starvation genes. Plant Cell Environ 39: 1818-1834

Ventura Y, Eshel A, Pasternak D, Sagi M (2015) The development of halophyte-based agriculture: past and present. Ann Bot 115: 529-540

Vera-Estrella R, Barkla BJ, Pantoja **O** (2014) Comparative 2D-DIGE analysis of salinity responsive microsomal proteins from leaves of salt-sensitive *Arabidopsis thaliana* and salt-tolerant *Thellungiella salsuginea*. J Proteomics 111: 113-127

Vigeland MD, Spannagl M, Asp T, Paina C, Rudi H, Rognli O-A, Fjellheim S, Sandve SR (2013) Evidence for adaptive evolution of low-temperature stress response genes in a Pooideae grass ancestor. New Phytol 199: 1060-1068

Wang F, Wang L, Qiao L, Chen J, Pappa MB, Pei H, Zhang T, Chang C, Dong C-H (2017) *Arabidopsis* CPR5 regulates ethylene signaling via molecular association with the ETR1 receptor. J. Intgr Plant Biol 59810-824.

Wang G, Pantha P, Tran KN, Oh DH, Dassanayake M (2019) Plant growth and Agrobacterium-mediated floral-dip transformation of the extremophyte *Schrenkiella parvula*. J Vis Exp 2019: 1-8

Wang L, Ma K-B, Lu Z-G, Ren S-X, Jiang H-R, Cui J-W, Chen G, Teng N-J, Lam H-M, Jin B (2020a) Differential physiological, transcriptomic and metabolomic responses of *Arabidopsis* leaves under prolonged warming and heat shock. BMC Plant Biol 20: 86

Wang G, Oh DH Dassanayake M (2020b) GOMCL: a toolkit to cluster, evaluate, and extract non-redundant associations of Gene Ontology-based functions. BMC Bioinformatics 21: 139.

Wang G, DiTusa SF, Oh DH, Herrmann AD, Mendoza-Cozatl DG, O'Neill MA, Smith AP, Dassanayake M (2021) Cross species multi-omics reveals cell wall sequestration and elevated global transcript abundance as mechanisms of boron tolerance in plants. New Phytol. In press.

Ward D (2009) The biology of deserts. Oxford: Oxford University Press

Weng M, Yang Y, Feng H, Pan Z, Shen W-H, Zhu Y, Dong A (2014) Histone chaperone ASF1 is involved in gene transcription activation in response to heat stress in *Arabidopsis thaliana*. Plant Cell Environ 37: 2128-2138

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Molecular Biol Evol 18: 691-699

Wiciarz M, Gubernator B, Kruk J, Niewiadomska E (2015) Enhanced chloroplastic generation of H_2O_2 in stress-resistant *Thellungiella salsuginea* in comparison to *Arabidopsis thaliana*. Physiol Plant 153: 467-476

Widiez T, Kafafia ESE, Girina T, Berr A, Ruffel S, Krouk G, Vayssièresa A, Shen W-H, Coruzzi GM, Gojon A et al. (2011) HIGH NITROGEN INSENSITIVE 9 (HNI9)-mediated systemic repression of root NO_3^- uptake is associated with changes in histone methylation. Proc Natl Acad Sci USA 108: 13329-13334

Willing E-M, Piofczyk T, Albert A, Winkler JB, Schneeberger K, Pecinka A (2016) UVR2 ensures transgenerational genome stability under simulated natural UV-B in *Arabidopsis thaliana*. Nat Commun | 7: 13522

Wu G, Wang X, Li X, Kamiya Y, Otegui MS, Chory J (2011) Methylation of a phosphatase specifies dephosphorylation and degradation of activated brassinosteroid receptors. Sci Signal 4: ra29

Wu HJ, Zhang Z, Wang JY, Oh DH, Dassanayake M, Liu B, Huang Q, Sun HX, Xia R, Wu Y et al. (2012) Insights into salt tolerance from the genome of *Thellungiella salsuginea*. Proc Natl Acad Sci USA 109: 12219-12224

Wu JX, Li J, Liu Z, Yin J, Chang Z-Y, Rong C, Wu J-L, Bi F-C, Yao N (2015) The *Arabidopsis* ceramidase AtACER functions in disease resistance and salt tolerance. Plant J 81: 767-780

Xia H, Camus-Kulandaivelu L, Stephan W, Tellier A, Zhang Z (2010) Nucleotide diversity patterns of local adaptation at drought-related candidate genes in wild tomatoes. Mol Ecol 19: 4144-4154

Xu T Lee K Gu L Kim J-I Kang H (2013) Functional characterization of a plastid-specific ribosomal protein PSRP2 in *Arabidopsis thaliana* under abiotic stress conditions. Plant Physiol Biochem 73: 405-411

Yamamoto M, Uji S, Sugiyama T, Sakamoto T, Kimura S, Endo T, Nishikawa S (2020) ERdj3B-mediated quality control maintains anther development at high temperatures. Plant Physiol 182: 1979-1990

Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13: 555-556

Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24: 1586-1591

Yang R, Jarvis DE, Chen H, Beilstein MA, Grimwood J, Jenkins J, Shu S, Prochnik S, Xin M, Ma C, (2013) The reference genome of the halophytic plant *Eutrema salsugineum*. Front Plant Sci 4: 46

Yang Z, Wang Y, Gao Y, Zhou Y, Zhang E, Hu Y, Yuan Y, Liang G, Xu C (2014) Adaptive evolution and divergent expression of heat stress transcription factors in grasses. BMC Evol Biol 14: 147

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13: 134

Yu X, Michaels SD (2010) The *Arabidopsis* Paf1c complex component CDC73 participates in the modification of FLOWERING LOCUS C chromatin. Plant Physiol 153: 1074 -1084

Zhang X, Zhang M, Takano T, Liu, S (2011) Characterization of an *AtCCX5* gene from *Arabidopsis thaliana* that involves in high-affinity K^+ uptake and Na^+ transport in yeast. Biochemi Biophys Res Commun 414: 96-100

Zhang X, Lei L, Lai J, Zhao H, Song W (2018a) Effects of drought stress and water recovery on physiological responses and gene expression in maize seedlings. BMC Plant Biol 18: 68

Zhang L, Chen L, Yu D (2018b) Transcription factor WRKY75 interacts with DELLA proteins to affect flowering. Plant Physiol 176: 790-803

Zhao Q, Shen J, Gao C, Cui Y, Wang Y, Cui J, Cheng L, Cao W, Zhu Y, Huang S, et al. (2019) RST1 is a FREE1 suppressor that negatively regulates vacuolar trafficking in *Arabidopsis*. Plant Cell 31: 2152-2168

Zhao S, Ye Z, Stanton R (2020) Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. RNA 26: 903-909

Zhou N, Robinson SJ, Huebert T, Bate NJ and Parkin IA (2007) Comparative genome organization reveals a single copy of CBF in the freezing tolerant crucifer *Thlaspi arvense*. Plant Mol Biol 65: 693-705

Zhou D, Zhou J, Meng L, Wang Q, Xie H, Guan Y, Ma Z, Zhong Y, Chen F, Liu J (2009) Duplication and adaptive evolution of the *COR15* genes within the highly cold-tolerant *Draba* lineage (Brassicaceae). Gene 441: 36-44

Zhou Y, Zhu J, Tong T, Wang J, Lin B and Zhang J (2019) A statistical normalization method and differential expression analysis for RNA-seq data between different species. BMC Bioinformatics 20: 163

Zhu Y, Weng M, Yang Y, Zhang C, Li Z, Wen-Hui Shen W, Dong A (2011) *Arabidopsis* homologues of the histone chaperone ASF1 are crucial for chromatin replication and cell proliferation in plant development. Plant J 66: 443-455

Zhu JK, Whited J, Seluanov A, Gorbunova V, Kasahara M, Amdam GV, Ulanovsky N, Feng G, Brunet A, Margoliash D (2015) The next top models. Cell 163: 18-20

Figure legends

Figure 1. Geographic distribution of *A. hierochuntica* and *de-novo* reference transcriptome. (A), Geographic distribution data are based on *Anastatica* L. in the Global Biodiversity Information Facility database (GBIF Secretariat (2019). GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2021-03-12). Average temperature data for this region are from 1948 to Feb 2021 acquired by the Physical Sciences Laboratory (Fan and van den Dool, 2008); (B), Lab-grown and wild *A. hierochuntica* plants. Panels: (i) 40 d-old lab-grown plant. Note the axillary inflorescence at each branch point; (ii) large mature plant from the Ovda valley population in the Negev desert. Ruler length = 30 cm; (iii) young seedlings growing near the dead mother plant from a Neot Smadar population in the Negev desert; (iv) a large population of *A. hierochuntica* in the Ovda valley with high variation in plant size due to spatial and temporal variations in water availability; (v) *A. hierochuntica* seedling already beginning to flower after producing four

true leaves (Neot Smadar); (vi) two tiny dead plants (white arrows) from a population near the Dead Sea valley, having already dispersed their few seeds; (C), Transcript length distribution and *A. hierochuntica* assembly descriptive statistics; (D), Assessment of reference transcriptome assembly completeness using the Benchmarking Universal Single-Copy Orthologs (BUSCO) tool (Simão et al., 2015). The percentages of 1,375 single-copy genes, conserved among land plants, identified in the *A. hierochuntica* transcriptome are shown.

Figure 2. Phylogenomic and comparative positive selection analyses of *A. hierochuntica* and other representative *Brassicaceae* genomes. (A), Maximum-likelihood tree topology based on supermatrix analysis of 13,806 ortholog groups that contain an amino acid sequence from at least four taxa. All nodes are 100% supported by 100 rapid bootstrapping repeats. Red asterisks, extremophyte species; (B), Comparison of the number of positively selected genes among species. Positively selected genes in each species were identified using the AGALMA-PAML pipeline; (C), Comparative GO-term enrichment analysis of positively selected genes. The red color intensity corresponds to the number of positively selected genes assigned with that GO term (the numbers are indicated within the cells). Cells with a white color correspond to GO terms that were not significantly enriched. The *A. thaliana* genome was used as the background gene set and significance (q-value < 0.05) of enrichment was assessed via the Fisher's exact test. For the full list of enriched GO-terms see Supplemental Fig. S2 and Supplemental Tables S8 to S12).

Figure 3. Expression of positively selected genes in response to heat stress. Gene expression was assessed by RNA-seq transcriptome analysis of *A. thaliana* and *A. hierochuntica* plants grown under control conditions or exposed to heat stress (see Fig. 4A for experimental design). Expression is expressed as transcripts per kilobase million (TPM) normalized gene expression. (A), Positively selected genes from the "all extremophyte species" and *A. hierochuntica* analyses that are associated with abiotic stress responses (Table 1); (B), Positively selected *A. hierochuntica* genes that function in photoperiodic flowering, regulation of meristems, and control of morphology (Table 1). Data are mean ± S.D. (n = 3) Asterisks indicate significant difference at $p < 0.05$ between *A. thaliana* and *A. hierochuntica* at the same time point and condition (Student's t-test). CM, control morning; CN, control afternoon; HW1M, heat wave 1 morning; HW1N, heat wave 1 afternoon; HW2N, heat wave 2 afternoon; HW3N, heat wave 3 afternoon; R1N, day 1 recovery from heat stress afternoon; CR2N,

control plants parallel to the R2N time point afternoon; R2N, day 2 recovery from heat stress afternoon; Blue shading, control conditions; Pink shading, heat conditions.

Figure 4. *A. thaliana* and *A. hierochuntica* exhibit similar transcriptome adjustment to heat stress. (A), Experimental design for *A. thaliana* and *A. hierochuntica* control and heat stress conditions. Control plants were harvested the day before the initiation of heat stress and on the last day of the experiment (indicated by arrows) from a parallel 23 °C control chamber. Red and blue circles represent samples harvested 1.5 h (morning) or 7 h (afternoon) respectively, after onset of light/heat. Each circle represents 3 independent experiments, each comprising 6 pooled plants. (B and C), Effect of heat stress on *A. thaliana* and *A. hierochuntica* leaf area (B) and fresh weight (C). Data are mean ± S.D. (n = 5) and are representative of two independent experiments. Letters above bars indicate significant difference at $p < 0.05$ (Tukey HSD test). Blue shading, control conditions; Pink shading, heat conditions. (D), Principal component analysis (PCA) of *A. thaliana* and *A. hierochuntica* transcript levels. Each point represents one biological replicate and the three replicates for each condition are depicted with the same symbol. Symbols are explained in the legend box and refer to the experimental design shown in (A); (E), Comparison of the abundance of 27,416 and 30,670 protein-coding *A. thaliana* and *A. hierochuntica* transcripts, respectively. Asterisks represent significant difference at $p < 0.05$ (Wilcoxon rank sum test) between the treatment compared to its respective control; (F), Percent of *A. thaliana* and *A. hierochuntica* differentially expressed genes (DEGs) in response to heat stress. In total, 17,989 *A. thaliana* and 19,443 *A. hierochuntica* genes were differentially expressed in response to heat stress in at least one condition (Supplemental Table S19), and % DEGs was calculated based on 27,416 and 30,670 protein-coding genes for *A. thaliana* and *A. hierochuntica*, respectively. CM, control morning; CN, control afternoon; HW1M, heat wave 1 morning; HW1N, heat wave 1 afternoon; HW2N, heat wave 2 afternoon; HW3N, heat wave 3 afternoon; R1N, day 1 recovery from heat stress afternoon; CR2N, control plants parallel to the R2N time point afternoon; R2N, day 2 recovery from heat stress afternoon; Union, DEGs identified under either HW1M or HW1N or HW2N or HW3N.

Figure 5. The *A. hierochuntica* transcriptome does not exist in a stress-ready state but exhibits a lower basal expression and higher fold-change expression than *A. thaliana* in response to heat stress. (A), Modes of expression of ortholog pairs between *A. thaliana* and *A. hierochuntica* in response to heat stress. WGCNA followed by DESeq2 was used to assign orthologs to response modes (Supplemental Fig. 5; Supplemental Table S13). Differences in absolute transcripts levels were

identified by comparing TPM minimum or maximum expression values (Student's *t*-test, $p \leq 0.05$). The green (*A. thaliana*) and orange (*A. hierochuntica*) lines indicate idealized expression patterns of the ortholog pairs in each species under control and heat conditions, compared to the *A. thaliana* control (dashed line). Ctrl, control; + heat, heat stress treatment; SA, genes associated with GO-terms for abiotic stress (see Supplemental Methods for GO-terms); (B), Transcript abundance of 17,962 *A. thaliana* and *A. hierochuntica* ortholog pairs under control conditions. (C), Combined violin and box plots showing absolute $\log_2$ fold-changes of *A. thaliana* and *A. hierochuntica* DEGs in response to heat stress (Supplemental Table S19). The median $\log_2$ fold-change is shown as a black square inside each box plot; (D), Number of DEGs, median $\log_2$ fold-change values and *p*-values for panel C. CM, control morning; CN, control afternoon; CR2N, control plants parallel to the R2N (day 2 recovery from heat stress afternoon) time point; HW1M, heat wave 1 morning; HW1N, heat wave 1 afternoon; HW2N, heat wave 2 afternoon; HW3N, heat wave 3 afternoon; Union, DEGs identified under either HW1M or HW1N or HW2N or HW3N. Asterisks represent significant difference at $p < 0.05$ (Wilcoxon rank sum test) between *A. thaliana* and *A. hierochuntica*.

Figure 6. *A. hierochuntica* shared- and unique-expressed orthologs in specific functional groups display lower basal and greater heat-mediated % change in expression than in *A. thaliana*. All genes used in this analysis possess a unique AGI code (putative *A. hierochuntica* orthologs were assigned *A. thaliana* AGI codes). Genes were chosen based on their association with GO terms for their respective categories (Supplemental Table S19; Supplemental Methods). Basal expression levels were based on CM conditions. % change in expression from basal level was calculated based on the minimum expression levels of downregulated genes (cell cycle, photosynthesis) or maximum expression levels of upregulated genes (abiotic stress) in response to heat stress over the three heat waves. Basal and % change in expression values for all genes in each category are in Supplemental Table S19. For box and whisker plots, the median (thick black line), the mean (cross below the median line) and interquartile range (IQR) of the observed differences are shown. Whiskers indicate the maximum/minimum range. Open circles correspond to extreme observations with values >1.5 times the IQR. Underlined numbers above the circles indicate the number of shared or unique expressed genes. Letters above the circles indicate significant differences at $p < 0.05$ (Student's *t*-test). Numbers next to boxes are median values. *A.t*, *Arabidopsis thaliana*; *A.h*, *Anastatica hierochuntica*.

Figure 7. The *A. hierochuntica* early heat response transcriptome displays lower basal and greater heat-mediated % change in expression than in *A. thaliana*. (A), Expression profiles of *A. thaliana* (left two panels) and *A. hierochuntica* (right two panels) morning and afternoon early heat response modules. These modules were assigned standard color-based names by WGCNA (e.g. Thistle, Ivory etc.; Supplemental Fig. S5; Supplemental Tables S14-S18). Transcript levels were scaled to visualize patterns of expression. The relative intensity of gene expression (red, upregulated; green, downregulated) is shown in the scale bar. Gene expression in each condition represents the average of three biological replicates. The number of genes in each module is shown under the respective module. CM, control morning; CN, control afternoon; HW1M, heat wave 1 morning; HW1N, heat wave 1 afternoon; HW2N, heat wave 2 afternoon; HW3N, heat wave 3 afternoon; R1N, day 1 recovery from heat stress afternoon; CR2N, control plants parallel to the R2N time point afternoon; R2N, day 2 recovery from heat stress afternoon; Blue lines above heat map, control conditions; Pink lines, heat conditions. (B), Non-redundant enriched GO-terms in the early heat response morning modules. The afternoon modules were not enriched in any GO-terms; (C), Expression of orthologs associated with abiotic stress GO terms (Supplemental Table S19; Supplemental Methods) Underlined numbers above the circles indicate the number of shared- or unique-expressed genes. Letters above the circles indicate significant differences at $p < 0.05$ (Student's T-test). Numbers next to boxes are median values. *A.t*, *Arabidopsis thaliana*; *A.h*, *Anastatica hierochuntica*.

Figure 8. Validation of "between species" RNA-seq analysis. (A), Top 10 expressed transcripts as a percentage of all expressed transcripts; (B), Comparison of basal expression of control samples using DeSeq2 rlog normalization or the between species Scale-Based Normalization (SCBN) method (Zhou et al., 2019). Numbers next to boxes are median values. Letters above the circles indicate significant differences at $p < 0.05$ (Wilcoxon rank sum test); (C), Relative QPCR expression of selected *A. thaliana* and *A. hierochuntica* genes. Gene expression was determined according to the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001) using *eIF4A1* from each species as a reference gene. Expression was normalized to the expression level in the control morning sample, which was assigned a value of 1. Data are mean ± SD (n = 3 to 4) and are representative of two independent experiments; (D), Comparison of the basal (control) expression levels of *CPR5* estimated by RNA-seq or absolute QPCR quantification of transcript copy number. Absolute quantification was performed using a fivefold serial dilution of gel-purified *CPR5* and *eIF4A1* (reference gene) QPCR products to create a standard curve. CM, control morning; CN, control afternoon; HW1M, heat wave 1 morning; HW1N, heat wave 1

afternoon; HW2N, heat wave 2 afternoon; HW3N, heat wave 3 afternoon; R1N, day 1 recovery from heat stress afternoon; CR2N, control plants parallel to the R2N time point afternoon; R2N, day 2 recovery from heat stress afternoon. Blue shading, control conditions; Pink shading, heat conditions. *A.t*, *Arabidopsis thaliana*; *A.h*, *Anastatica hierochuntica*; TPM, transcripts per kilobase million.

Table 1. Positively selected genes[a], with a potential role in adaptation to extreme environments. Selected genes from five CODEML branch-site model analyses are indicated based on their *A. thaliana* ortholog identifier. TF, Transcription Factor.

Positively selected gene	Function	References
All extremophyte species (*A. hierochuntica*, *E. salsugineum*, *S. parvula*, *T. arvense* and *A. halleri*)		
AKS2 (At1g05805)	TF; facilitates stomatal opening, ABA response	Takahashi et. al. 2013
ASF1B (At5g38110)	Histone H3/H4 chaperone; repair of UV-B-induced DNA damage, basal and acquired thermotolerance	Lario et al., 2013; Nie et al., 2014; Weng et al., 2014
MYB52 (At1g17950)	TF; ABA response, drought tolerance, involved in the regulation of secondary wall formation, seed mucilage	Park et al., 2011; Cassan-Wang et al., 2013; Shi et al., 2018
PHR1/UVR2 (At1g12370)	Photolyase enzyme; repair of UV-B-induced DNA damage	Ahmad et al., 1997; Landry et al., 1997; Jiang et al., 1997
WRKY75 (At5g13080)	TF; Pi starvation, root development, GA-mediated flowering, defense response	Devaiah et al., 2007; Velasco et al., 2016; Guo et al., 2017; Zhang et al., 2018b
A. hierochuntica		
APX6 (At4g32320)	Hydrogen peroxide-scavenging enzyme; alleviation of ROS damage	Chen et al., 2014
bZIP1 (At5g49450)	TF, light and nitrogen sensing, salt and drought tolerance	Obertello et al., 2010; Sun et al., 2012; Para et al., 2014
CYP71 (At3g44600)	Cyclophilin; silencing of homeotic genes; meristem development, interacts with FAS1 and the floral repressor LHP1	Li et al., 2007; Li and Luan, 2011
FAS1 (At1g65470)	Subunit of CaF-1; organization of apical meristems, cellular differentiation, DNA repair	Leyser & Furner, 1992; Kaya et al., 2001; Hisanaga et al., 2013
FBH2 (At4g09180)	TF; photoperiodic flowering	Ito et al., 2012
SBI1/LCMT1 (At1g02100)	Leucine carboxylmethyltransferase; brassinosteroid signaling; flowering, stress responses	Di Rubbo et al., 2011; Wu et al., 2011; Creighton et al., 2017
VIP5 (At1g61040)	PAF1c component; activates floral repressors and photoperiodic pathway regulators. regulation of N uptake	Oh et al., 2004; Yu and Michaels, 2010; Crevillen & Dean, 2011; Widiez et al., 2011; Lu et al., 2017

E. salsugineum		
ATCES1/ACER (At4g22330)	Alkaline ceramidase; sphingolipid homeostasis, disease resistance, salt tolerance	Wu et al., 2015
GRXS13 (At1g03850)	Glutaredoxin; Chilling and photooxidative stress tolerance	Laporte et al., 2012; Hu et al., 2015
NCA1 (At3g54360)	Chaperone; regulates catalase 2 (ROS-scavenging enzyme) activity, salt, cold, high pH stresses	Li et al., 2015
PSRP2 (At3g52150)	Plastid-specific ribosomal protein; RNA chaperone activity, negative regulator of seed germination under abiotic stress	Xu et al., 2013
SLK2 (At5g62090)	Transcriptional adaptor; embryogenesis, organ development, repression of stress-responsive gene transcription	Bao et al., 2010; Lee et al., 2014; Shrestha et al., 2014
S. parvula		
Atrab28 (At1g03120)	LEA protein; Li^+ tolerance	Borrell et al., 2002
CAX11/CCX5 (At1g08960)	Cation calcium exchanger; K^+ uptake, Na^+ transport in yeast	Zhang et al., 2011
PER1 (At1g48130)	Peroxiredoxin; ROS scavenging, enhances primary seed dormancy	Chen et al., 2020
A. thaliana		
ATG6 (At3g61710)	AuTophGy-related protein; autophagy, pathogen defense	Patel & Dinesh-Kumar, 2008
ATL2 (At3g16720)	RING-H2 zinc-finger protein; pathogen defense	Serrano & Guzmán, 2004
ERDJ3B (At3g62600)	ER-localized DNAJ chaperone; anther development under heat stress, pathogen defense	Nekrasov et al., 2009; Yamamoto et al., 2020
RST1 (At3g27670)	ARM-repeat protein; RNA exosome cofactor, vacuolar trafficking, cuticular wax production, embryo development, pathogen defense	Chen et al., 2005; Mang et al., 2009; Lange et al., 2019; Zhao et al., 2019;
XLG2 (At4g34390)	Heterotrimeric G protein; pathogen defense	Liang et al., 2016

[a]For log-likelihood values of the alternative and null models, log-likelihood ratio tests and *p*-values, see Supplemental Tables S3 to S7.

Figure 1

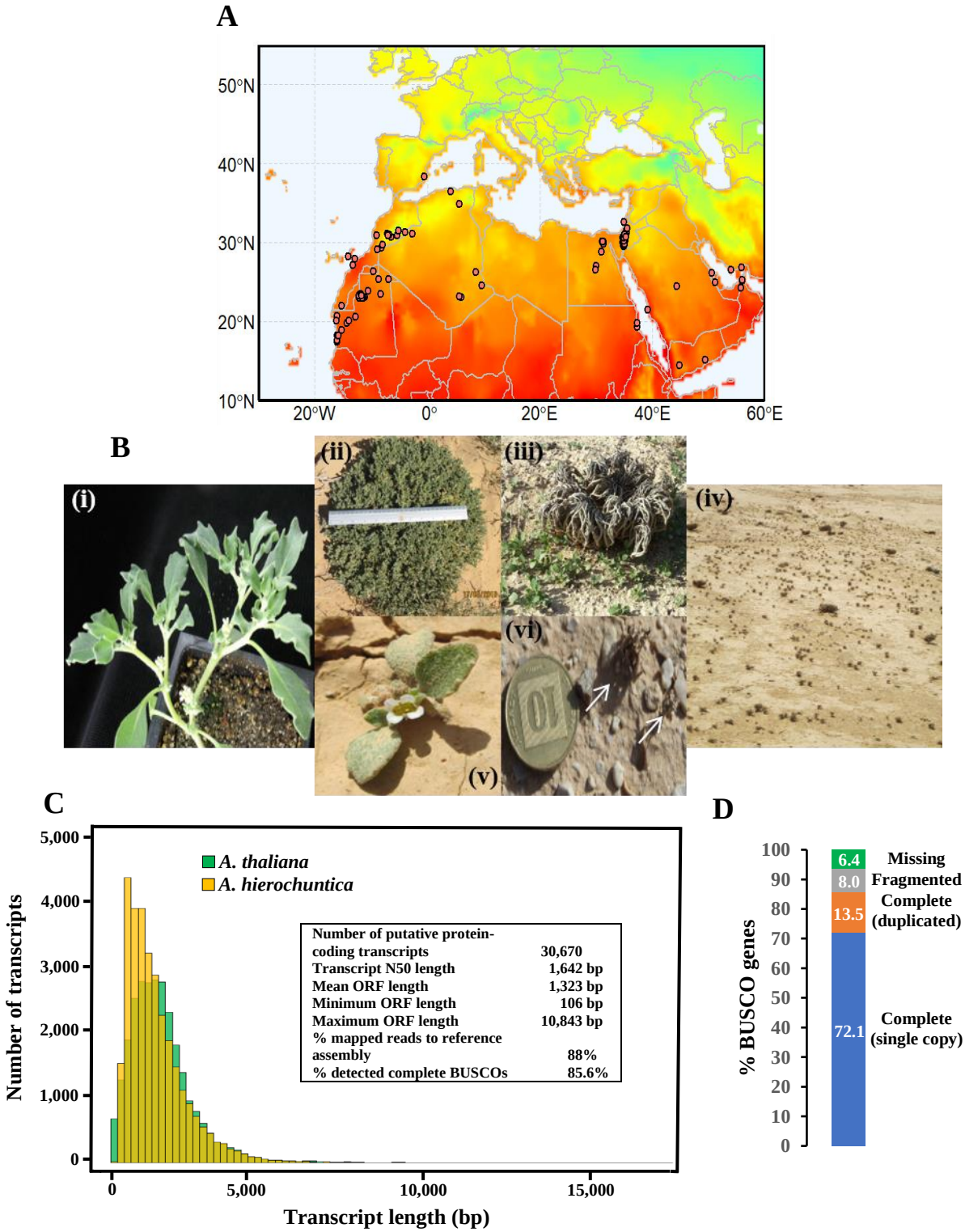

A
50°N
40°N
30°N
20°N
10°N
20°W
0°
20°E
40°E
60°E
B
(i)
(ii)
(iii)
(iv)
(v)
(vi)
C
A. thaliana
A. hierochuntica
Number of putative protein-coding transcripts 30,670
Transcript N50 length 1,642 bp
Mean ORF length 1,323 bp
Minimum ORF length 106 bp
Maximum ORF length 10,843 bp
% mapped reads to reference assembly 88%
% detected complete BUSCOs 85.6%
Number of transcripts
5,000
4,000
3,000
2,000
1,000
0
0
5,000
10,000
15,000
Transcript length (bp)
D
% BUSCO genes
100
90
80
70
60
50
40
30
20
10
0
6.4
Missing
8.0
Fragmented
13.5
Complete (duplicated)
72.1
Complete (single copy)

Figure 2

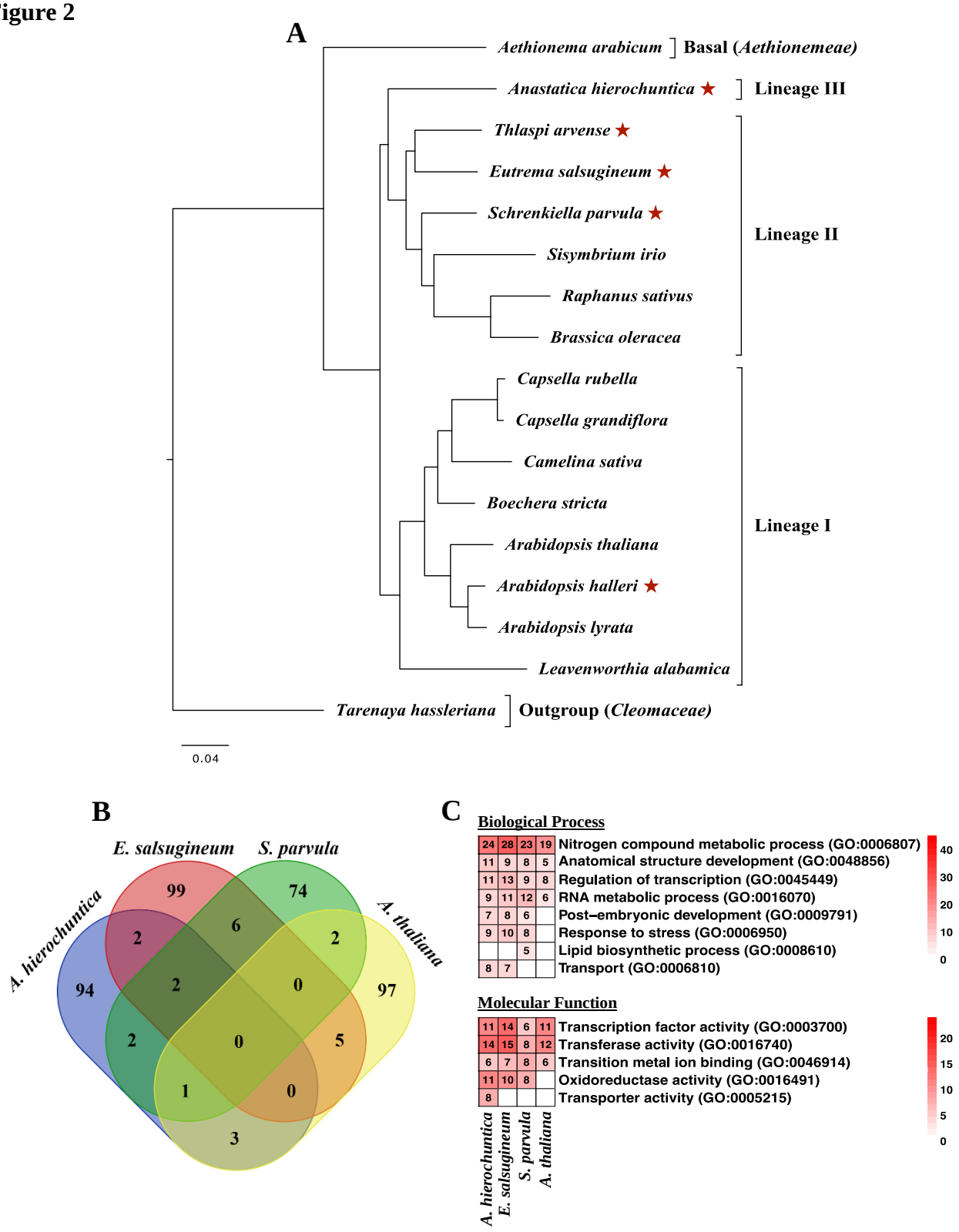

A
Aethionema arabicum
Basal (Aethionemeae)
Anastatica hierochuntica ★
Lineage III
Thlaspi arvense ★
Eutrema salsugineum ★
Schrenkiella parvula ★
Sisymbrium irio
Raphanus sativus
Brassica oleracea
Lineage II
Capsella rubella
Capsella grandiflora
Camelina sativa
Boechera stricta
Arabidopsis thaliana
Arabidopsis halleri ★
Arabidopsis lyrata
Leavenworthia alabamica
Lineage I
Tarenaya hassleriana
Outgroup (Cleomaceae)
0.04
B
E. salsugineum
S. parvula
A. hierochuntica
A. thaliana
99
74
2
6
2
94
2
0
97
2
0
5
1
0
3
C
Biological Process
24 28 23 19 Nitrogen compound metabolic process (GO:0006807)
11 9 8 5 Anatomical structure development (GO:0048856)
11 13 9 8 Regulation of transcription (GO:0045449)
9 11 12 6 RNA metabolic process (GO:0016070)
7 8 6 Post–embryonic development (GO:0009791)
9 10 8 Response to stress (GO:0006950)
5 Lipid biosynthetic process (GO:0008610)
8 7 Transport (GO:0006810)
40
30
20
10
0
Molecular Function
11 14 6 11 Transcription factor activity (GO:0003700)
14 15 8 12 Transferase activity (GO:0016740)
6 7 8 6 Transition metal ion binding (GO:0046914)
11 10 8 Oxidoreductase activity (GO:0016491)
8 Transporter activity (GO:0005215)
20
15
10
5
0
A. hierochuntica
E. salsugineum
S. parvula
A. thaliana

Figure 3

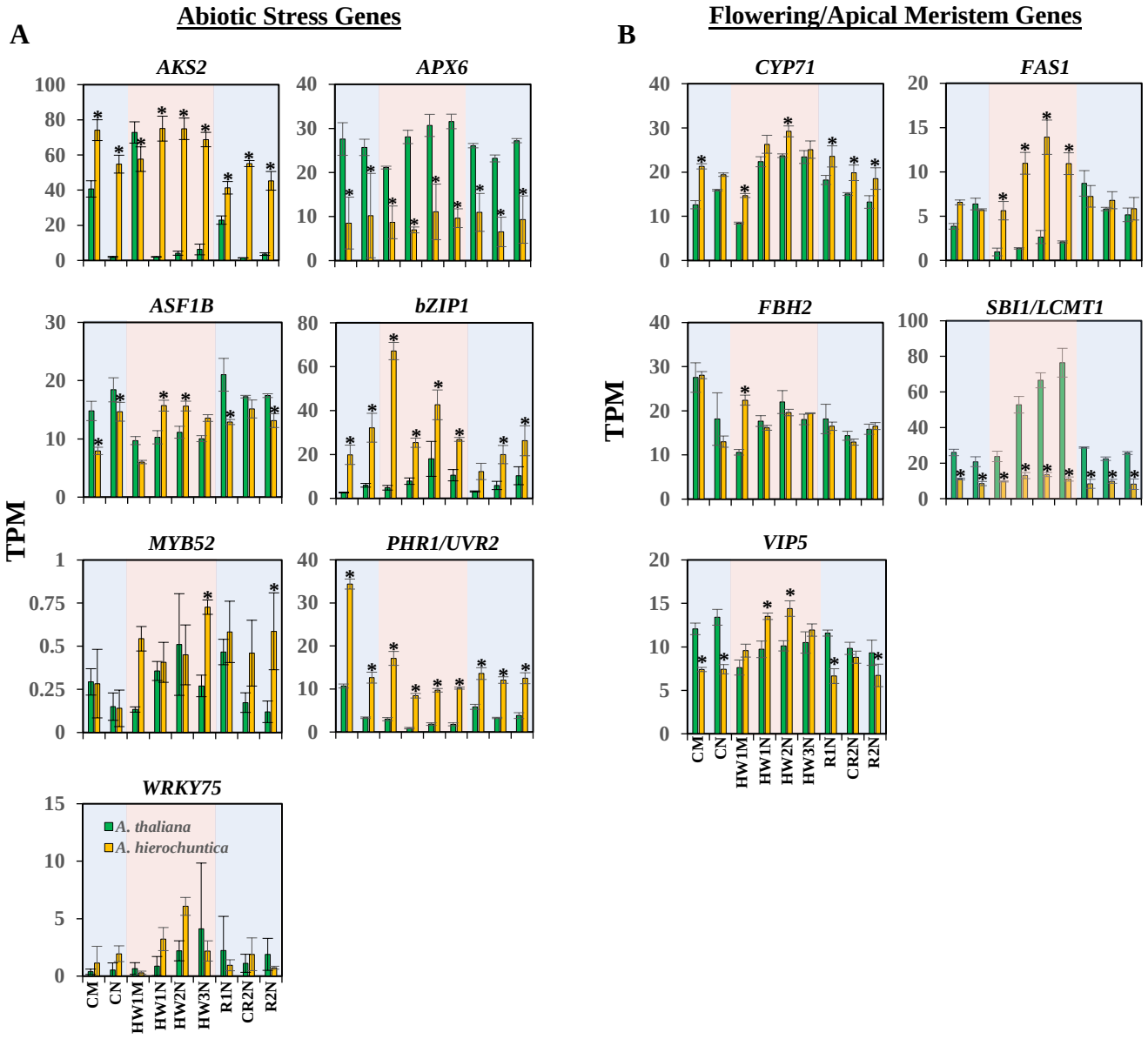
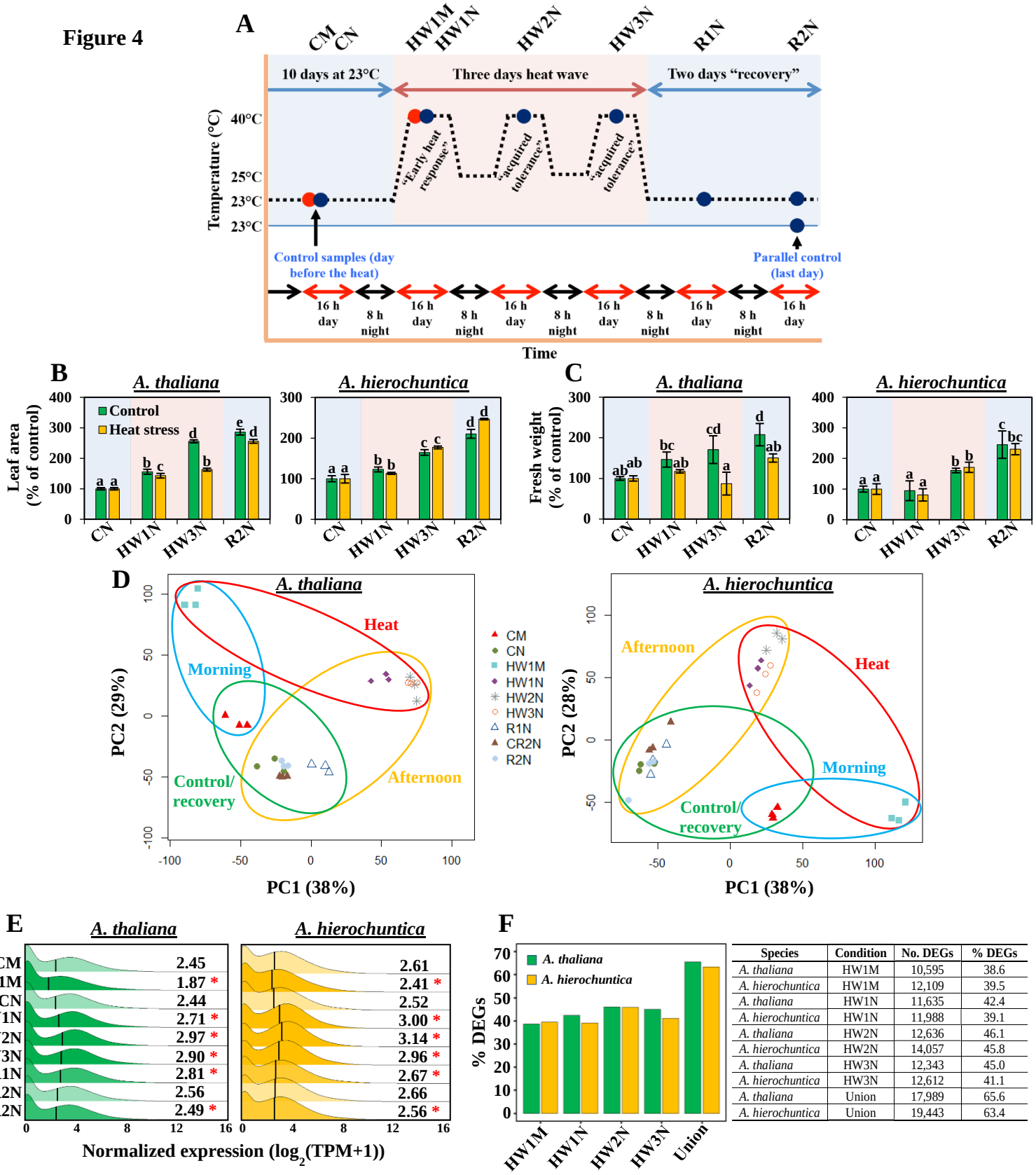
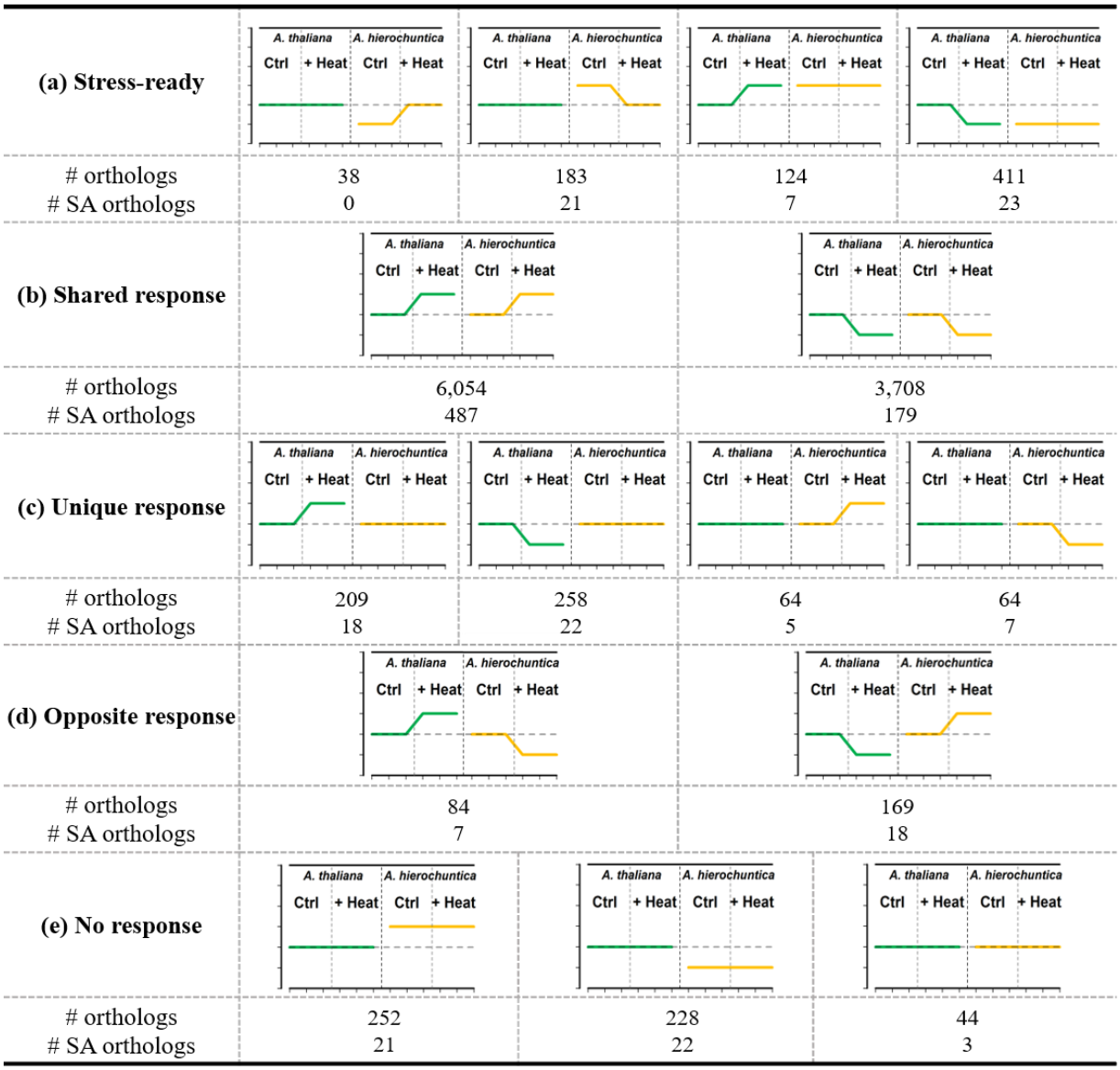

A
Abiotic Stress Genes
B
Flowering/Apical Meristem Genes
AKS2
APX6
ASF1B
bZIP1
MYB52
PHR1/UVR2
WRKY75
CYP71
FAS1
FBH2
SBI1/LCMT1
VIP5
TPM
A. thaliana
A. hierochuntica
CM
CN
HW1M
HW1N
HW2N
HW3N
R1N
CR2N
R2N

Figure 4

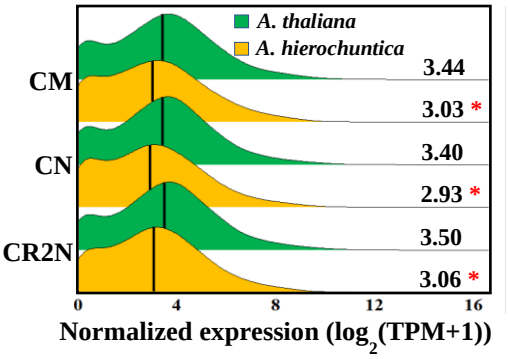


Species	Condition	No. DEGs	% DEGs
A. thaliana	HW1M	10,595	38.6
A. hierochuntica	HW1M	12,109	39.5
A. thaliana	HW1N	11,635	42.4
A. hierochuntica	HW1N	11,988	39.1
A. thaliana	HW2N	12,636	46.1
A. hierochuntica	HW2N	14,057	45.8
A. thaliana	HW3N	12,343	45.0
A. hierochuntica	HW3N	12,612	41.1
A. thaliana	Union	17,989	65.6
A. hierochuntica	Union	19,443	63.4

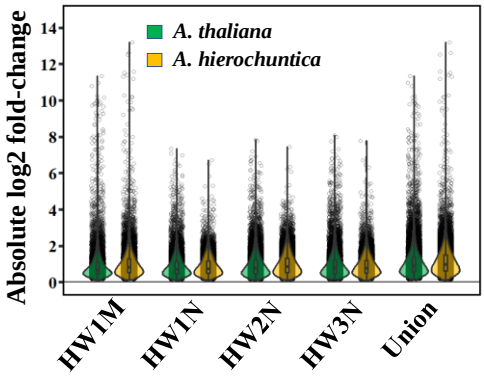
Figure 5



B



C



D

Condition	No. DEGs		Median $\log_2$ fold-change		*p*-value
	A. thaliana	*A. hierochuntica*	*A. thaliana*	*A. hierochuntica*	
HW1M	10,595	12,109	0.68	0.82	4.35E-63
HW1N	11,635	11,988	0.72	0.75	1.26E-05
HW2N	12,636	14,057	0.76	0.85	5.12E-33
HW3N	12,343	12,612	0.76	0.79	3.38E-02
Union	17,989	19,433	0.89	0.98	2.37E-34

Figure 6

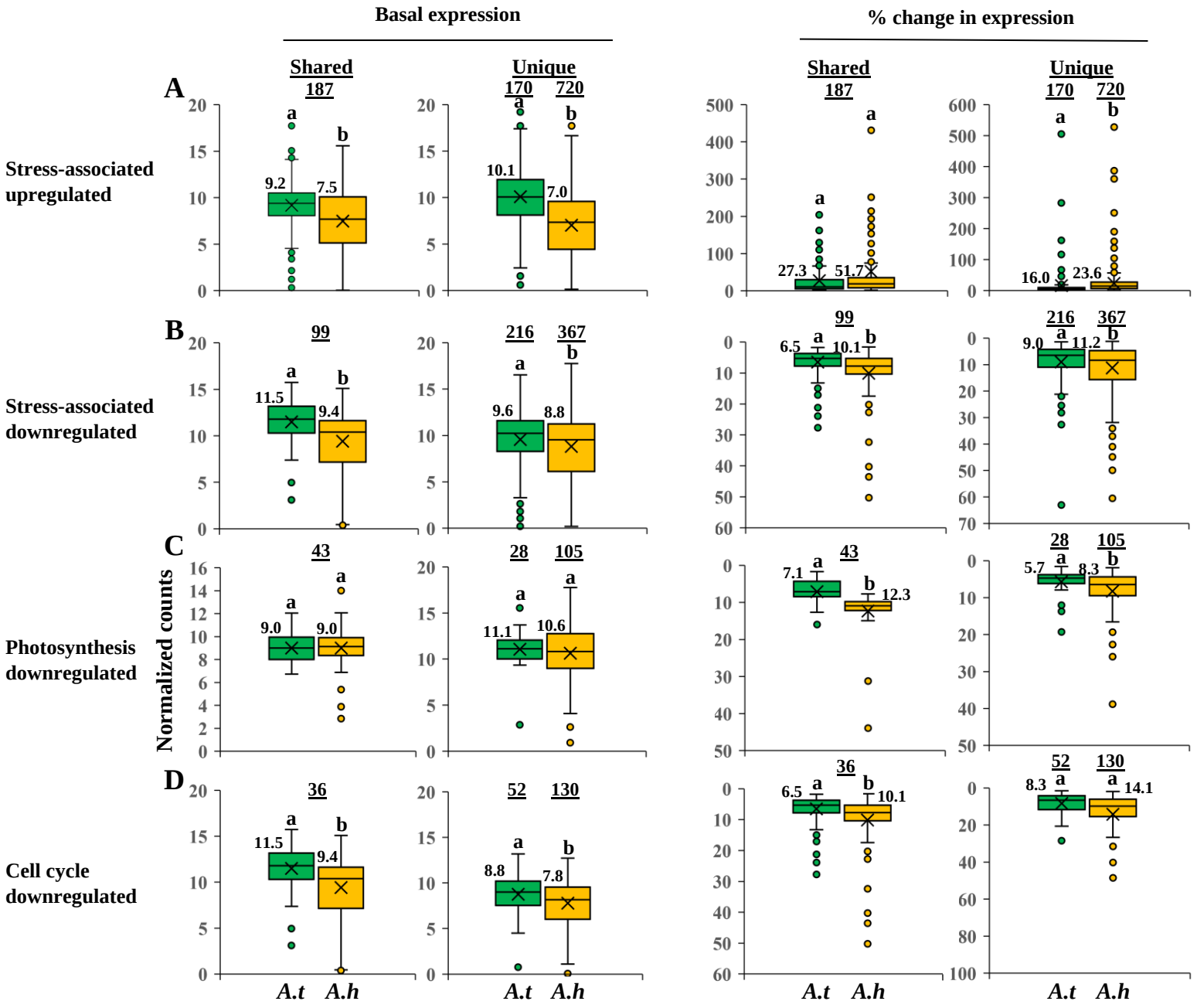
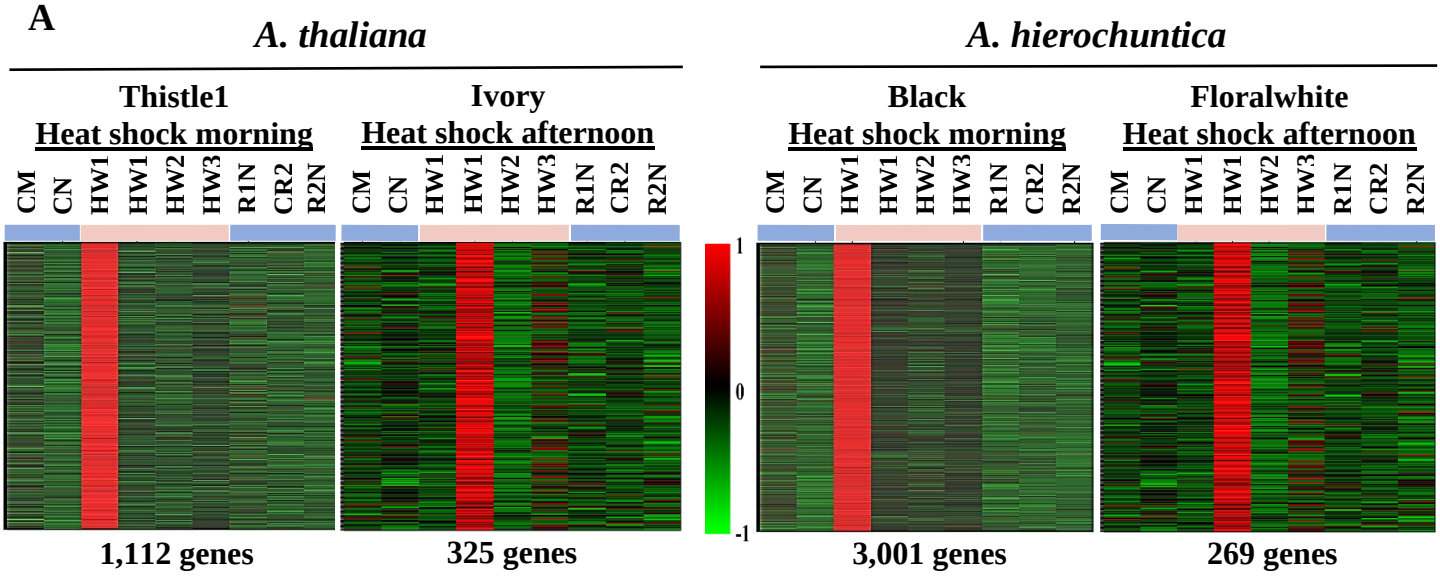

Basal expression
% change in expression
Shared
Unique
A
B
C
D
Stress-associated upregulated
Stress-associated downregulated
Photosynthesis downregulated
Cell cycle downregulated
Normalized counts
187
170
720
99
216
367
43
28
105
36
52
130
9.2
7.5
10.1
7.0
27.3
51.7
16.0
23.6
11.5
9.4
9.6
8.8
6.5
10.1
9.0
11.2
9.0
9.0
11.1
10.6
7.1
12.3
5.7
8.3
11.5
9.4
8.8
7.8
6.5
10.1
8.3
14.1
A.t
A.h

Figure 7

A



B

Enriched GO-terms in early heat response morning modules			
***A. thaliana* Module Thistle1**		***A. hierochuntica* Module Black**	
GO-term	***p*-value**	**GO-term**	***p*-value**
response to heat (GO:0009408)	8.92E-14	response to heat (GO:0009408)	4.27E-19
response to light intensity (GO:0009642)	2.26E-06	response to stress (GO:0009642)	2.46E-18
response to stress (GO:0006950)	4.64E-04	response to chitin (GO:0010200)	6.68E-09
response to hydrogen peroxide (GO:0042542)	1.61E-03	regulation of nitrogen compound metabolic process (GO:0051171)	9.08E-07
response to water deprivation (GO:0009414)	1.76E-02	response to high light intensity (GO:0009644)	9.08E-07
response to reactive oxygen species (GO:0000302)	3.44E-02	response to reactive oxygen species (GO:0000302)	5.26E-06

C

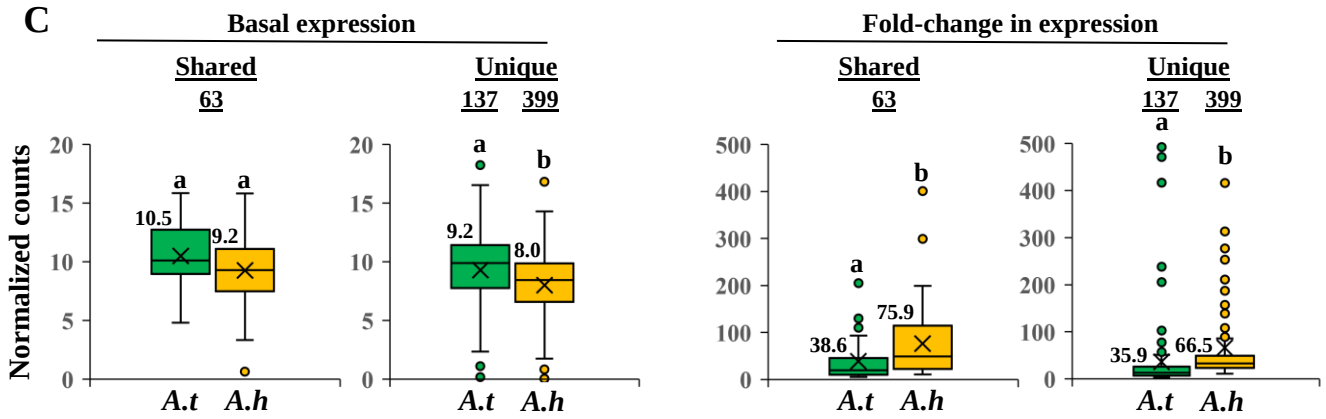
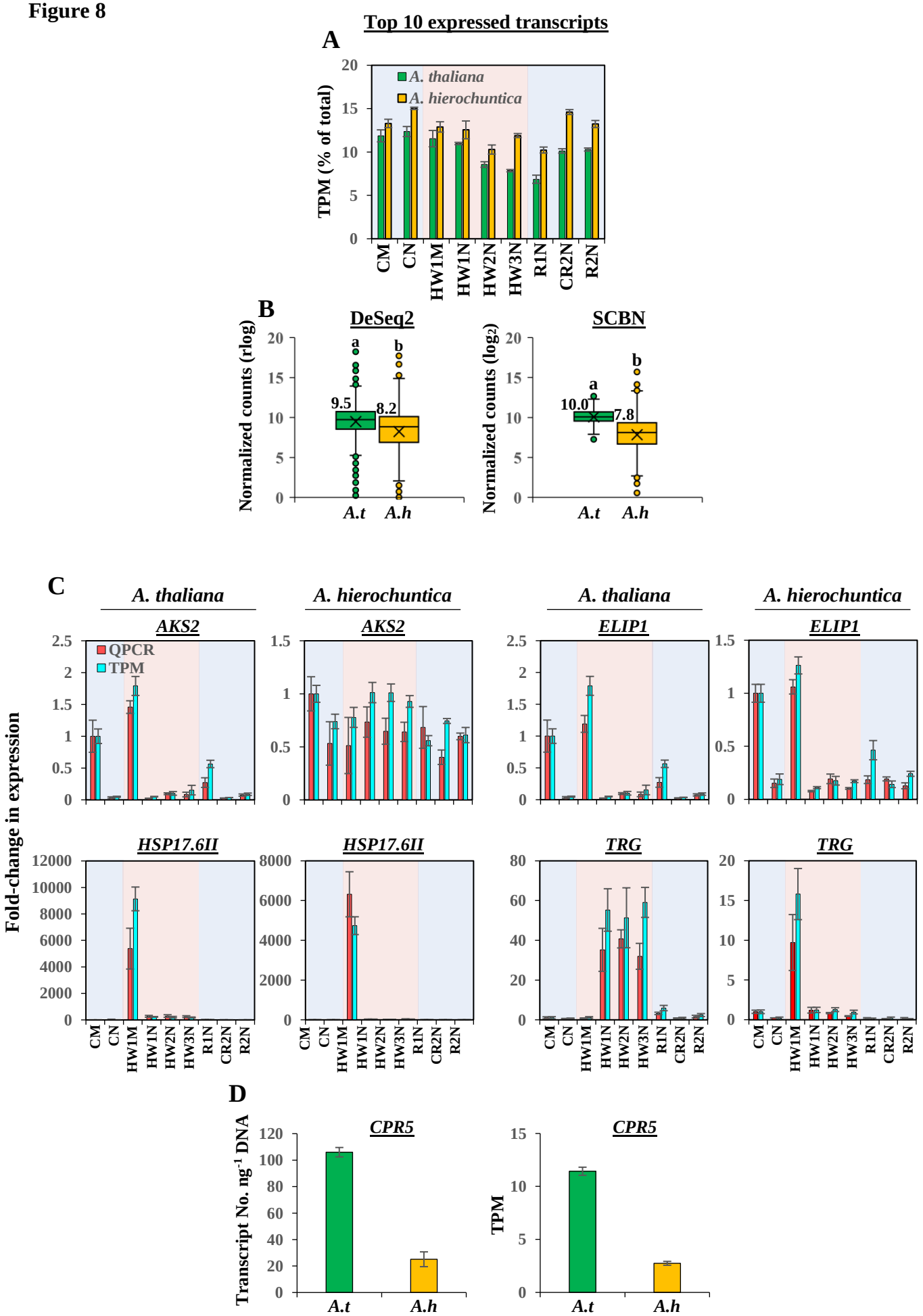


Figure 8


Top 10 expressed transcripts
A
A. thaliana
A. hierochuntica
TPM (% of total)
CM
CN
HW1M
HW1N
HW2N
HW3N
R1N
CR2N
R2N
B
DeSeq2
SCBN
Normalized counts (rlog)
Normalized counts (log2)
a
b
9.5
8.2
10.0
7.8
A.t
A.h
C
A. thaliana
A. hierochuntica
AKS2
ELIP1
HSP17.6II
TRG
QPCR
TPM
Fold-change in expression
D
CPR5
Transcript No. ng-1 DNA
TPM