

**Assessment of functional relevance of genes associated with local temperature variables
in *Arabidopsis thaliana***

Yuan Jiang^{a,b,1}, Zhixue Wang^{b,1}, Hui Du^b, Runlong Dong^b, Yaping Yuan^{a,*} and Jian Hua^{b,*}

^a Jilin Engineering Research Center of Crop Biotechnology Breeding, College of Plant Science,
Jilin University, Changchun 130062, P.R. China.

^b Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY
14853, USA.

¹These authors contributed equally to this work.

*Corresponding authors: Jian Hua, Tel: +1-607-255-5554, Email: jh299@cornell.edu

Yaping Yuan, Tel: +86-431-87836266, Email: yuanyp@jlu.edu.cn.

Abstract

Environmental genome wide association study (eGWAS) has been used to identify genetic variations associated with local environmental variables of natural accessions. How likely these variations are functionally relevant to environmental adaptation has not been extensively investigated. Here we used three local temperature variables of 1129 natural accessions of *Arabidopsis thaliana* to assess the functional relevance of associated genes in temperature responses. Candidate genes for top quantitative trait loci (QTLs) associated with each of the variables, minimum, mean or maximum temperatures, were selected for functional analysis. The loss-of-function mutants of these genes were assessed for growth and stress tolerance at 10°C, 16°C, 22°C, 28°C and 35°C. Twenty out of 29 genes analyzed were found to affect growth or tolerance at one or more of these temperatures. In addition, genes associated with maximum temperature more likely have a function at higher temperature, while genes associated with minimum temperature more likely have a function at lower temperature. Examination of geographic distributions of natural variants of five genes revealed that these variations likely contribute to local temperature adaptation. Furthermore, variations in a large proportion of minimum and mean-temperature associated genes exhibited a significant association with growth phenotypes assessed at 10°C for a small set of natural accessions. This study thus shows a functional relevance of genes and variants associated with environmental variables and supports the feasibility of the use of local temperature factors in investigating genetic basis of temperature adaptation.

Keywords

eGWAS; biofactor; temperature; adaptation; *Arabidopsis*

Summary Statement

Whether or not genetic variations associated with local environmental variables are functionally relevant to environmental adaptation is largely unknown. Here we analyzed top associated genes for three temperature variables of 1129 natural accessions of *Arabidopsis thaliana* and showed a functional relevance of these genes in temperature responses, supporting the feasibility of using environment association to investigate environmental adaptation.

Introduction

Local adaptation has been observed in many organisms (Hereford, 2009). Numerous biotic and abiotic environmental factors, such as light, temperature, water availability, CO₂, soil salinity, and surrounding microbes, influence plant growth and development (de Wit et al., 2016, Dusenge et al., 2019, Gupta et al., 2020, Rodriguez et al., 2019, van Zelm et al., 2020). Plants use genome variations to adapt to local environment and to survive a wide range of conditions (Exposito-Alonso et al., 2018, VanWallendael et al., 2019). This is supported by garden plot experiment where natural accessions often had a higher fitness in their local environment than accessions from other locations. For instances, accessions of *Arabidopsis thaliana* (referred as *Arabidopsis* from now on) from colder area in mainland Europe have a higher basal freezing tolerance than accessions from warmer area (Horton et al., 2016), and subgroup of the *Arabidopsis* family from high carbonate soil location showed higher fitness to high carbonate soil than subgroups from the low carbonate soil (Teres et al., 2019). A lagging climate adaptation was observed for banked seeds of *Arabidopsis* where genotypes from historically warmer than the planting site had higher relative fitness than native genotypes (Wilczek et al., 2014). The molecular basis for this adaptation has been studied in many plant species especially in *Arabidopsis* because of the availability of genome sequences of natural accessions and their wide global distributions.

Temperature is an important environmental factor affecting plant growth and development (Quint et al., 2016, Ding et al., 2020). Leaf emergence, flowering and fruiting are some of the traits contributing to temperature climate adaptation (Parmesan and Hanley, 2015). Gene variations responsible for local adaptation to climate temperature variables have been reported in *Arabidopsis* (Prinzenberg et al., 2020); (Wang et al., 2019, Wang et al., 2021); (Lu et al., 2021)). For instance, the *DOG1* (*DELAY OF GERMINATION 1*) gene contributes to seed dormancy variations in *Arabidopsis* (Bentsink et al., 2006), and its functional variants exhibits geographic signature of adaptation to thermal climates (Martinez-Berdeja et al., 2020).

Most of these genetic variations mediating diversity of environmental responses were identified by classical Genome Wide Association Study (GWAS) where phenotypes/traits were assessed for a population of natural accessions and then were associated with genomic variations. This classical GWAS has been widely employed to identify genes important for growth and environmental responses in *Arabidopsis* (Weigel and Mott, 2009, Korte and Farlow, 2013, Xiao et al., 2017) and other plant species (Li et al., 2016, Crowell et al., 2016). However, its use in

capturing natural variations for environmental responses is limited by the labor-intensive measurement of phenotype. Recently, association studies between genetic variations and local climate, named environment GWAS or eGWAS, are being explored in revealing how natural accessions might adapt to their local environment (Hancock et al., 2011, Fournier-Level et al., 2011). This includes seasonal change, temperature, water availability and day length (Lasky et al., 2012, Savolainen et al., 2013).

Despite the promise of eGWAS in identifying gene variants for local adaptation, genomic associations do not always result from functional causal effect. Vigorous testing of functional relevance of associated genes and adaptive value of genetic variants are needed, but it has been carried out only for a few genes (Ferrero-Serrano and Assmann, 2019). It remains unclear to what extent an association of gene variations with environment parameters indicate their roles in adaptation to the local environment. Answering this question is important not only for the use of eGWAS to uncover genetic basis of environment adaptation but also for the selection of DNA variations in generating environmental resilient plants.

Here, we started to address this question by examining the functions and potential adaptation values of candidate genes associated with local temperature variables in *Arabidopsis*. We performed eGWAS analysis on three temperature parameters, including mean temperature, maximum temperature and minimum temperature of their collection sites, for 1129 *Arabidopsis* natural accessions. Thirty most significant QTLs and 29 genes associated with these three temperature parameters were selected for functional analysis in growth and stress tolerance under different temperatures. We show that most of these genes have a temperature-influenced function. Importantly, these associated genes more likely have a measurable function under the temperature range they are associated with in the local environment. This study indicates that eGWAS of local climate factors is a viable approach in identifying environment-related genes and variants for local adaptation.

Materials and methods

Plant materials

The wild-type and mutant plants used in this study are in the Col-0 accession of *Arabidopsis thaliana*. The T-DNA insertion mutants (Alonso et al., 2003) and the 53 natural accessions were obtained from the Arabidopsis Biological Resource Center (ABRC). Genotyping

primers for these mutations were designed using the T-DNA Express tool (<http://signal.salk.edu/cgi-bin/tdnaexpress>).

Plant growth conditions

Seeds were sown on the soil and treated at 4°C for 3 d before moving to growth chambers with light at 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and humidity at 50–70%. Rosette size and fresh weight of seedlings were measured at 14 d after germination for plants grown at 16°C, 22°C or 28°C or at 21 days after germination for plants grown at 10°C. RWC under heat was measured after treating plants (grown for 12d at 22°C) at 35°C for 5 d followed by a recovery at 22°C for 2d. Root length at 35°C were measured on seedlings grown at 22°C for 5 d followed by 35°C growth for 4 days. Rosette size and fresh weight of seedlings were measured at 28 d after germination for the 53 natural accessions grown at 10°C.

Genome-wide association study

The climatic temperature data was obtained as described before (Wang et al., 2019). Three genome-wide association study (GWAS) for mean-temp, max-temp and min-temp of 1129 *Arabidopsis* natural accessions were performed on GWAPP (<https://gwapp.gmi.oeaw.ac.at/index.html>) with a ~10M SNP dataset gained from the Illumina resequencing. The accelerated mixed model method that accounts for population structure was used. Minor allele count (MAC) was set at 52 to remove very minor SNPs for the GWAS result.

Measurement of *Arabidopsis* growth traits

For measurement of seed production, three plants were wrapped together and served as one biological repeat. Eight biological replicates were performed for each genotype. Seeds were harvested once a day when the first batch of seeds started to mature until all seeds were harvested. Student's t-test was performed to assess differences between mutants and Col-0.

For measurement of root and hypocotyl length, plants were grown on ½ MS medium supplemented with 1% sucrose and 0.8% agar at 10°C or 35°C for 4 days. The length of root and hypocotyl was quantified using the ImageJ software (<https://imagej.net/>).

Measurement of relative water content

The relative water content (RWC) was measured as previously described (Morgan, 1984) with some modification. Leaves were weighted immediately after being removed from plants to determine the fresh weight (FW). They were then soaked for 24h at room temperature in deionized water in a tightly sealed tubes before they were weighted to determine the turgid weight (TW).

Finally, leaves were dried at 37°C for 36 h in tubes before their dry weight (DF) was measured to determine. The RWC was calculated by the following equation: $RWC (\%) = [(FW-DF)/(TW-DF)] \times 100\%$.

Results

eGWAS of local temperature variables

Eleven variables or biofactors defined by earlier (Hancock et al., 2011, Andrade-Piedra et al., 2005) are related to local temperature parameters/variables for *Arabidopsis* natural accessions (Fig. S1). These biofactors are not all independent, and most are correlated or anti-correlated with other factors (Fig. S1). We selected three biofactors, annual mean temperature (bio1), maximum temperature of the warmest month (bio5) and minimum temperature of the coldest month (bio6), for further GWAS analysis because they together capture most of the temperature features of the location (Fig. S1). Histogram plotting showed that all three biofactors span a wide range among the 1129 *Arabidopsis* accessions (Fig. 1A-C). The annual mean temperature (referred as mean-temp hereafter) ranges from -8.0°C to 19.0°C; the max temperature of the warmest month (referred as max-temp hereafter) ranges from 9.8°C to 36.0°C, while minimum temperature of the coldest month (referred as min-temp hereafter) ranges from -27.4°C to 16.5°C (Fig. 1A-C). *Arabidopsis* is an annual plant that finishes its cycle in spring/summer or grows over winter. Therefore, these min or max temperatures do not necessarily indicate the temperature *Arabidopsis* plants experience, but is an indicator of the relative value of their growth temperature.

We conducted association analysis of these three temperature-related climate parameters or biofactors (temperature values) with ~10M genome polymorphisms of the 1129 natural accessions using the web tool GWA-Portal (Fig. 1D-F) (Seren et al., 2012, Carlos et al., 2016). The climate data of these 1129 accessions were obtained as described previously (Wang et al., 2019, Wang et al., 2021). Associated single nucleotide polymorphisms (SNPs) that are located within a 10 kb fragment were considered as under the same QTL. The significance of a QTL was defined by the *p* value of the most significant SNP within the QTL. Because the distribution of these parameters or biofactors (using temperature values) was not entirely normal and significance threshold setting for normal distribution might not apply, we used several threshold settings to detect significant SNPs. With a significant threshold set at $p < 10^{-8.5}$ (corresponding to Benjamini–Hochberg correction for multiple testing), few SNPs were found to be associated with any of these

biofactors. At a significance threshold of $p < 10^{-6}$, a total of 1126 SNPs were associated with min-temp, and very few were associated with mean-temp or max-temp (Fig. 1G). At a threshold of $p < 10^{-5}$, 588 SNPs were associated with mean-temp and 144 SNPs with max-temp (Fig. 1G). At a threshold of $p < 10^{-4}$, more associated SNPs were identified but they were under the same QTL with SNPs detected at a higher significance, indicating that this threshold does not detect more QTLs. We therefore used thresholds of $p < 10^{-6}$ for min-temp and $p < 10^{-5}$ for both max-temp and min-temp to capture the associated SNPs. Although the thresholds were below those calculated for a normal distribution population, they were most likely valid because SNPs with a significance value of $p < 10^{-4}$ had been found to be biologically relevant (Tabas-Madrid et al., 2018). In total, 90, 25 and 46 QTLs were identified as associated with mean-temp, max-temp and min-temp, respectively (Fig. 1G). These QTLs were numbered as QTL1xx for mean-temp (bio1), QTL5xx for max-temp (bio5), and QTL6xx for min-temp (bio6) (Table 1). Overlapping QTLs were found for these three factors. Eight QTLs were shared between mean-temp and max-temp, 24 QTLs were shared between mean-temp and min-temp and no QTLs were shared between max-temp and min-temp (Fig. 1H-I).

Functional assessment of associated genes

To start to assess whether or not associated SNPs are relevant to local temperature adaptation, we selected 30 QTLs for functional evaluation. These include 12 most significant QTLs from mean-temp, 9 most significant from max-temp, and 9 most significant from min-temp, with one shared between mean-temp and max-temp (Table 1 and Dataset S1). A linkage disequilibrium (LD) block ($r^2 > 0.5$) was first identified for the SNP with the highest association score, and genes within the LD block that contain significant SNPs were considered as associated genes. This yielded a total of 34 associated genes for the 30 QTLs, with 25 QTLs containing one associated gene and five QTLs containing two genes (Table 1). The number of significant SNPs ranged from 1 to 47 in one associated gene, and they were located in either the coding region only (15 genes), the untranslated region (UTR) or the promoter region (14 genes), or both the coding region and UTR (6 genes) (Dataset S1). For the genes that do not have a published name, we name them as *EMET* (*Environment efactor*) followed by three numbers with the first being the biofactor number and the last two number being the serial number of the isolated QTL (Table 1).

Most of these associated genes do not have a reported function or a fully characterized function, so we set out to determine their roles in growth and stress tolerance over a wide temperature range. A total of 34 genes were characterized by their loss of function (LOF) mutants from the T-DNA insertion mutant library in the Col-0 background (Alonso et al., 2003) (Fig. S2 and Table S1). Five genes were excluded from further study (Table 1), either because of unavailability of two mutant alleles for the gene or being the less likely candidate (e.g. no annotation) among the two genes under the same QTL (for one QTL). Homozygous mutants of the first allele of these 29 genes were isolated by PCR genotyping and propagated for phenotypic analyses. A second allele of these genes (Table S1) were subsequently characterized when the first allele exhibited an altered phenotype compared to the wild type to verify that the defect was indeed due to mutation of the gene. We analyzed mutants of the 29 associated genes at five temperatures 10°C, 16°C, 22°C, 28°C and 35°C by their growth and tolerance phenotypes. A mutant phenotype is defined when both alleles exhibited significant differences ($p < 0.05$, student's *t* test) and at least one allele exhibited high significant difference ($p < 0.01$, student's *t* test) from the wild-type plant. False discovery rate (FDR) adjusted *p* values were corrected for multiple testing with Benjamine Hochberg procedure, and an FDR cutoff of 0.1 was considered (Dataset S2).

At 10°C, mutants of 10 genes had smaller rosettes compared to the wild type Col-0 at three weeks old, and mutants of one gene had a larger rosette compared to the wild type (Fig. S3A-B and Table 2). Mutants of 11 genes and 10 genes had shorter hypocotyl and root, respectively, when grown in the dark for 4 days (Fig. S3C-F and Table 2). Totally, 15 genes impacted growth at 10°C, with five genes affecting growth of all of rosette, hypocotyl and root (Table 2).

At 35°C, mutants of 11 genes had shorter roots and mutants of two genes had longer roots compared to the wild type (Fig. S4A-B and Table 2). Mutant of 7 genes exhibited decreased relative water content (RWC) compared to wild type Col-0 (Fig. S4C-D and Table 2). Defects in root length and RWC at 35°C were correlated in these mutants (Fig. S5). In total, six genes promote root elongation and maintain RWC at 35°C, indicating their positive regulation of heat tolerance (Table 2).

For growth phenotypes at 16°C, 22°C and 28°C, rosette size and fresh weight as well as seed yield were measured for mutants of these 29 genes. Fresh weight and rosette size had a high positive correlation at all their respective temperatures (Fig. S5), and we define a rosette growth defect for a mutant when either size or weight is different from the wild type.

At 16°C, mutants of three genes had smaller rosettes while mutants of three other genes showed larger rosettes compared to the wild type Col-0 when assayed by both area and fresh weight (Fig. S6A-D and Table 2). Mutants of six genes had reduced seed yields (weight of all seeds produced from one plant) while mutants of one gene had increased seed yield compared to the wild type (Fig. S6E-F and Table 2). In summary, seven genes affect either rosette growth or yield at 16°C, one gene promotes both rosette growth and seed yield, while two genes inhibit rosette growth but promote seed yields (Table 2).

At 22°C, mutants of eight genes and one gene had smaller and larger rosettes than the wild type Col-0, respectively (Fig. S7A-D and Table 2). Mutants of six genes had reduced seed yields compared to wild type Col-0 (Fig. S7E-F). In total, 11 genes affect rosette growth and/or seed yields at 22°C, with four genes promoting both growth and yield (Table 2).

At 28°C, mutants of six genes had smaller rosettes (size or weight) than the Col-0 wild type (Fig. S8A-D and Table 2). Mutants of five genes had reduced seed yields compared to wild type Col-0 (Fig. S8E-F and Table 2). Most of these genes affect either rosette growth or seed yield, and only one gene affect both processes.

The proportion of candidate genes exhibiting an identified mutant phenotype is rather high. A previous large-scale genetic screen identified 54 out of 11000 T-DNA insertion mutants exhibiting a chilling sensitive phenotype (Wang et al, 2016). Although 0.5% is a probably a low estimation (visual inspection may miss subtle phenotypes), the finding of more than 50% with a measurable phenotype is unlikely to be by chance.

We further tested whether these genes have a temperature-dependent function on growth under non-extreme temperatures. Two-way ANOVA analyses were carried out on rosette size, fresh weight, and seed yield of these mutants at 16°C, 22°C and 28°C. Four LOF mutants, *pgyl*, *emet641*, *emet148*, and *cs1b05*, showed strong G (genotype)×E (environment) effects on both rosette size and fresh weight, while *fst1*, *dyrkp*, *ceqorh*, *aln* and *emet517* had G×E effect on the fresh weight (Table 3). Four mutants, *pgyl*, *ski2*, *erf53* and *emet102*, showed strong G×E effects on seeds weight (Table 3). In sum, 12 genes have a temperature-dependent effect on growth measured by size, weight or seed yield. The gene effect analyzed here is based on a LOF mutant allele and is done in the Col-0 background. For those genes that did not exhibit a G×E effect for the knockout allele, it is still possible other alleles from nature (such as a reduced or increased

activity compared to the Col-0 allele) could have a G×E effect or the LOF allele has a G×E effect in a different background.

In addition, function in flowering time was assessed for these genes at 22°C as early flowering was shown to result in reduced rosette growth (Zagotta et al., 1992). Mutants of five genes were found to be early flowering compared to Col-0, while mutants of five genes showed late flowering compared to the Col-0 as measured by bolting time (Fig. S9A). Little correlation was observed between rosette growth and flowering time (Fig. S5), indicating that the rosette growth alteration unlikely resulted from flowering time change. In contrast, anti-correlation was observed between seed yield and flowering time among these mutants, suggesting that seed yield is largely dependent on the length of vegetative growth. Of note is an early flowering phenotype of the *tsf-2* mutant allele (SALK_064104C) which likely results from a higher *TSF* (*TWIN SISTER OF FT*) expression (Fig. S10A-B); while the knockout mutant *tsf-3* allele showed delayed flowering which is consistent with an earlier finding (Kobayashi et al., 1999). Neither allele had an apparent difference in rosette growth and seed yield compared to the wild-type Col-0 at 22°C (Fig. S7A and Table 2). Surprisingly, both mutants have defects in cold and heat tolerance. The function of this gene in cold and heat tolerance needs to be verified by additional alleles or complementation.

Molecular function of associated genes

We examined the molecular characteristics of genes that have a function in different temperature environment. Nine genes have measurable functions at four or all five temperatures as indicated by their mutant phenotypes (Fig. 2A and Table 2), suggesting their general regulation on growth irrespective of temperature. *EMET148* (coding for a calmodulin binding protein), *PGYL* (*PIGGYBACK1*, coding ribosomal protein L10a), and *CEQORH* (*CHLOROPLAST ENVELOPE QUINONE OXIDOREDUCTASE HOMOLOG*, coding for an oxidoreductase) promote growth at the five temperatures analyzed (Table 2). *CSLB05* (*CELLULOSE SYNTHASE LIKE 5*, coding a putative cellulose synthase) and *EMET517* (coding for a GroES-like zinc-binding dehydrogenase) are positive regulators of growth at four temperatures (all except for 16°C). *ALN* (*ALLANTOINASE*), a negative regulator of heat (35°C) tolerance, encodes an allantoinase that converts allantoin to allantoate. This gene also functions in promoting growth at both 22°C and 28°C (Fig. 2A and Table 2) and was found to negatively regulate drought and salt stress tolerance (Irani and Todd, 2016), suggesting that allantoin and allantoate conversion might balance

tolerance to multiple abiotic stresses and plant growth. *EMET641* (coding for a transmembrane protein) has a measurable function at 10°C, 16°C, 22°C and 28°C. *TSF* (*TWIN SISTER OF FT*), a regulator of flowering time, has a measurable function at 10°C, 16°C 28°C and 35°C. In addition, *SUPPERKILLER 2* (*SKI2*, coding an RNA helicase subunit of SKI complex) has a measurable function at four temperatures except for 28°C. SKI proteins are required for 3' to 5' cytoplasmic RNA decay by the exosome, and SKI2 has been implicated in salt and cold stress responses, K⁺ deprivation responses and regulation of wax biosynthesis (Zhang et al., 2015, Luhua et al., 2013, Xu et al., 2011, Zhao and Kunst, 2016). Taken together, the identified genes that regulate growth in general are often involved in essential cellular processes.

Four genes function mainly in growth regulation at low temperatures (10°C and/or 16°C) (Fig. 2A and Table 2). *EMET645*, a positive regulator of growth at 10°C, encodes a chaperone DnaJ-domain superfamily protein. Proteins in this family were reported to confer freezing tolerance, induced during cold acclimation, and were candidate genes for natural variations in freezing tolerance (Horton et al., 2016, Porankiewicz and Clarke, 1997, Chow and Tung, 1998, Teigen et al., 2015). *GT2L* (*GT-2LIKE PROTEIN*), a positive regulator of rosette growth at both 10°C and 16°C, encodes a plant-specific transcription factor that interacts with calcium/calmodulin and is strongly induced by low temperature (Xi et al., 2012). *FST1* (*FLAVONOL SOPHOROSIDE TRANSPORTER 1*), a growth promoter at 10°C, but a growth inhibitor at 16°C, encodes a member of the nitrate/peptide family of transporters which was reported to promote flavanol production in tapetum cells at a low temperature (Bhatia et al., 2018, Grunewald et al., 2020). The basis for its apparently opposite functions at 10°C and 16°C needs further investigation. *EMET142*, a growth promoter at both 10°C and 22°C, encodes a Cysteine/Histidine-rich C1 domain family protein. Therefore, genes that promote low temperature growth encode chaperone proteins, transcription factors, calcium signaling molecules.

Five genes have roles in growth at normal and high temperatures (Fig. 2A and Table 2). *EMET102*, a positive regulator of root elongation at 35°C, encodes a DnaJ domain containing heat shock protein. *RAVI* (*RELATED TO ABI3/VPI 1*), a positive regulator of seed production at 28°C, encodes an ERF (ETHYLENE RESPONSIVE FACTOR) transcriptional factor that has been implicated in ABA stress response (Shin and Nam, 2018). *CYP702A6* (*CYTOCHROME P450, FAMILY 702, SUBFAMILY A, POLYPEPTIDE 6*), a positive regulator of heat tolerance at 35°C and growth at 28°C, encodes one of the cytochrome P450 family enzymes, which are a superfamily

of hemethiolate enzymes that catalyze diverse oxidative reactions and have been implicated in regulation of plant stress responses (Pandian et al., 2020, Hansen et al., 2021). In addition, *CYP51G1*, a positive regulator of heat tolerance at 35°C, encodes a putative demethylase which involved in sterol biosynthesis (Kim et al., 2005). *TCF1 (TOLERANT TO CHILLING AND FREEZING1)*, a positive regulator of heat tolerance at 35°C (Fig. S4), encodes a chromatin associated protein which was reported to function in freezing tolerance by mediating expression of lignin biosynthesis genes (Ji et al., 2015). Therefore, genes that positively regulate growth at normal and high temperatures are involved in heat shock response, transcriptional regulation, metabolism and signaling. Another gene, *DYRKP-2B (PLANT-SPECIFIC DUAL-SPECIFICITY TYROSINE PHOSPHORYLATION-REGULATED KINASE 2B)*, a positive regulator of heat tolerance at 35°C and growth at 10°C and 22°C, encodes a putative protein kinase.

Only one gene, *ERF53*, was identified as having functions in both low and high temperatures but not normal temperatures (Fig. 2A and Table 2). It is a target gene regulated by *HSF* and *REV* genes during heat response (Li et al., 2019). The identification of this transcription factor as a positive growth regulator at 10°C and 35°C suggests that it may mediate responses to both low and high temperature stresses.

Link between temperature range of gene function and temperature association of genetic variants

We further asked if a gene more likely has a measurable growth-regulation function in the temperature range that its variation is associated with. In another word, do mutants of genes associated with max-temp or min-temp more likely show altered phenotypes under a high or low temperature respectively? To this end, we calculated the proportion of mutants that exhibited a growth defect at five temperatures for each of the three temperature association groups. At high temperature of 35°C, the percentages of mutants exhibiting a measurable defect in RWC/root length were 63%/63% for max-temp associated genes, 23%/54% for mean-temp associated genes, and 0%/22% for min-temp associated genes. Similarly, at low temperature of 10°C, the percentages of mutants showed different performance on rosette size/hypocotyl length/root length were 67%/56%/56% for min-temp associated genes, 31%/31%/23% for mean-temp associated genes and 25%/38%/25% for max-temp associated genes (Fig. 2A-B and Table 2). In addition, mutants of genes associated with min-temp more frequently had altered phenotypes at 16°C (33% in rosette size, 33% in seed production) compared to those associated with mean-temp (23% in

rosette size, 23% in seed production) and those with max-temp (0% in rosette size, 13% in seed production) (Fig. 2A-B and Table 2). Similarly, mutants of genes associated with max-temp more likely had an altered seed yield at 28°C (38%) than those associated with mean-temp (15%) and min-temp (0%) (Fig. 2A-B and Table 2).

To determine if the difference in percentage for the three group of genes is statistically significant, we performed the chi-square test on the proportions of genes showing altered phenotype for three gene groups (min-temp, mean-temp and max-temp) at each temperature. All phenotypes measured were combined for each temperature in order to have sufficient sample size for this test. The three group of genes have significant proportion differences in proportions at 10°C, 16°C and 35°C, but not at 22°C and 28°C (Fig. 2C). This further supports that genes from max-temperature more likely have function under high temperature while genes from min-temperature likely have function under low temperature condition.

Assessment of local adaptation of gene variants

To determine whether or not variations of the associated genes were relevant for local adaptation, we asked if a gene variant likely confers better growth or tolerance in the local environment. We first assessed the functional impact of the most significant SNPs for each of these 20 associated genes with a detected function in growth and stress adaptation, in order to identify gene variants with a predicted reduced or enhanced activity for the testing. Functional impact was assessed by either gene expression level or protein function. Association of gene variants with their own gene expression level was assessed by using the RNA-seq data of 665 natural accessions (Kawakatsu et al., 2016). Three genes, *TCF1* (SNP: Chr3_20616707, $p = 0.000536$, Welch's t test), *ALN* (SNP: Chr4_2525988, $p = 0.0390419$, Welch's t test), and *ERF53* (SNP: Chr2_8985825, $p = 0.0268618$, Welch's t test), had their SNPs associated with their respective gene expressions (Fig. S11A), suggesting that these gene variants have different gene expression levels which likely lead to different gene activity levels. The impact of the most significant SNP of each gene on its protein activity was determined using the web-based tool *POLYMOPH 1001* (Carlos Alonso-Blanco. Jorge Andrade. Claude Becker, 2016). None of the SNPs were predicted to have a high impact, which is defined as frame shift, stop codon gain, stop codon loss or start loss. SNPs from *SKI2* and *CSLB05* were predicted to cause splicing variations and thus could alter protein activities (Fig. S11B). Combining these two analyses, we identified 5 genes (out of 20 genes) whose most significant SNPs likely have a significant impact on their respective gene activity. More genes

would likely have variations that impact their activities when all associated SNPs (not only the most significant one) are examined for their functional impact (Fig. S12). We assessed whether or not a gene variant likely confers a more robust growth at its local environment than another variant on these 5 genes by the global distribution of natural variants.

ERF53, identified from min-temp association, is a positive regulator of plant growth under low temperature (Table 2). Its minor allele was associated with a higher gene expression (Fig. S11A) and was distributed in regions with a lower min-temp (Fig. 3A). This suggests an adaptation as a higher expression of *ERF53* might have been preserved for adaptation to lower min-temp.

TCF1, identified from the mean-temp association, is a positive regulator of seed yield at 16°C. Its minor allele was associated with a high gene expression and was distributed in regions with a lower mean-temp (Fig. 3B), suggesting an adaptation to lower temperature for higher seed yield.

SKI2, identified from mean-temp association, is a negative regulator of plant growth at lower temperature (Table 2). Its minor allele, with two most significant SNPs (Chr3_17297514 and Chr3_17295089) predicted to cause splicing variations leading to reduced protein activity (Table S1), was distributed at the locals with lower mean-temp (Fig. 3C). This allele likely confers a more robust growth at low temperature and was likely selected for low temperature environment.

ALN, isolated from max-temp association, is a negative regulator of root growth (not water control) at 35°C but a positive regulator of growth at 10°C, 22°C and 28°C (Table 2). Its minor allele was associated with a relative higher *ALN* expression and was distributed more in locals with a higher max-temp (Fig. 3D). This minor allele is expected to have a less growth at 35°C but better growth at 28°C, 22°C, and 10°C. Further examination revealed that the minor allele of *ALN* is also more distributed at a lower temperature region for both mean-temp and min-temp and this distribution pattern was not observed for other genes associated with max-temp (Fig. 3E). Therefore, the minor allele of *ALN* (with higher expression) might enable a more robust growth at a lower temperature and this advantage might have overridden the disadvantage at a very high temperature.

CSLB05, identified from both mean-temp and max-temp associations, is a positive regulator of plant growth under all five temperatures except for 16°C. Its minor allele of most significant SNP at Chr4_8722723 is predicted to cause splicing variation and a truncated protein product (Fig. S11B) and therefore should confer a reduced growth under most of the temperature

conditions. This minor allele is distributed more at the locals with higher max-temp and higher mean temp than the major allele (Fig. 3F). It is not obvious that the minor allele conferring a reduced growth offers an advantage over the major allele at high temperature.

In sum, variants of four genes out of the five genes could offer advantage in their local regions and therefore likely have resulted from adaptation to local temperature environment. A variant of the fifth gene did not seem to offer advantage in the local region in terms of temperature factors. Further study should reveal whether or not it is selected by its associated genes or by other environmental factors.

Variations in biofactor associated genes are associated with variations of plant growth in natural accessions at 10°C

To further examine whether or not genes identified from eGWAS contribute to phenotypic variations of natural accessions, we assessed rosette size and rosette fresh weight of a small set of 53 accessions grown at 10°C (Dataset S3). Ten out of 13 mean-temp associated genes and five out of nine min-temp associated genes were significantly associated with rosette growth at 10°C (Fig. S13). Among these fifteen genes, the loss of function mutants of six genes had measurable rosette growth differences compared to Col-0 at 10°C (Table 2). Because the impact on gene function from these variations cannot be predicted solely from the sequence change, we were not able to test the correlation of the phenotypic impact of these alleles with the phenotypic impact of the LOF alleles. Nevertheless, these associations suggest that even with a small set of accessions where the power of association detection is weak, a large proportion of candidate genes identified from eGWAS exhibited association with phenotypic variations among natural accessions.

Discussion

In this study, we assessed the viability of using eGWAS to identify adaptive gene variants for local environment. We performed eGWAS on three temperature factors (max-temp, mean-temp, and min-temp) on 1129 Arabidopsis natural accessions and identified 129 unique QTLs. Functional analyses of the top 29 associated genes revealed that 20 had measurable roles in growth and temperature stress tolerance and that many have temperature-specific functions. Interestingly, genes associated with min-temp more likely have a function under low temperature while genes associated with max-temp more likely have a function at high temperature. Further analysis of variants of five genes revealed that genetic variants of four genes are distributed more in locales

where they may confer better growth, suggesting an adaptive value of these genetic variants. In addition, we assessed the 10°C growth phenotypes of a small set of natural accessions and found an association of the growth phenotype with a large proportion of genes associated with min-temp or mean-temp but not high-temp. Together, these data indicate that eGWAS is a viable approach for identifying adaptive gene variants for local temperature environment.

Nine out of 29 genes associated genes studied did not exhibit any growth regulating function under the tested conditions. Besides the possibilities of genetic redundancy of analyzed gene with its homologous gene or only a hyperactive allele has a detectable phenotype, the absence of a LOF mutant phenotype of the candidate gene might happen when the causal gene for a QTL does not have the highest association and was therefore overlooked in this study. Alternatively, the association may result from functional adaptation to other environmental factor that is correlated with temperature factor. Additional environmental conditions especially those assimilating natural environment, could be used in the future for further functional assessment. We also did not find clear adaptive values for one of the five genes analyzed. This gene has growth regulating function over a wide range of temperatures, which likely make it difficult to assess the advantage of having a variant under a certain temperature environment. Alternatively, its variants may be adaptive for the factors other than temperature. Future studies on additional associated genes and variants will further reveal the effectiveness of the eGWAS in identifying adaptive genome variations.

This study also identifies genes that are important for growth and development at low or high temperatures. Although the number of genes studied is small, these genes appear to preferentially fall into different functional categories depending on the temperature association. The five genes (*CYP51G1*, *CYP702A6*, *ALN*, *EMET517* and *CSLB05*) that are associated with max-temp and have function at high temperature are involved in metabolism and biosynthesis. In contrast, the six genes that are associated with min-temp and regulate low temperature growth encode transcription factors (*GT2L* and *ERF53*), chaperone (*EMET645*), ribosomal subunit (*PGYI*) transmembrane protein (*EMET64I*) and transporter (*FSTI*). This suggest that plants may use different molecular strategies to adapt to local temperature. Variations in transcription reprogramming may help plants to more effectively adapt to low temperature environment, while variations in metabolic processes may enable effective adaptation to a high temperature environment. It is also interesting to note that *ERF53* affects tolerance to both low and high temperatures, suggesting some shared mechanisms in temperature stress tolerance.

This study identifies 10 new genes for flowering time control. Molecularly, they are likely involved in metabolism (*DYRKP-2B*, *CSLB05*, *ALN* and *CEQORH*), transcription (*ERF53* and *RAVI*), signaling (*EMET148*, *EMET641*), translation (*PYGI*), and transporting (*FSTI*). Together with the known flowering time gene *TSF*, they comprise over 1/3 of 29 temperature factor associated genes analyzed and half of the 20 associated genes with a measurable growth function. A relatively high frequency in identifying flowering time genes from eGWAS has been observed before (Tabas-Madrid et al., 2018), suggesting that changing flowering time is an important strategy for plants to avoid reproduction at unfavorable environment. However, identifying these genes from temperature eGWAS is unlikely due to their flowering time regulation. All these 11 flowering time genes affect rosette growth under at least one temperature tested. While mutants of three genes are early flowering and have smaller rosette, mutants of eight other genes do not have that correlation. Therefore, the majority of these (at least 8/11) have a flowering time-independent function in growth regulation. In addition, mutants of all these genes except for *TSF* had a small difference in flowering time (within 2 to 5 days) compared to the wild type, which unlikely will have a large impact on plant growth.

It is worth noting that the 1129 accessions used in this study contain some ambiguous and problematic accessions that may not have true climate factors (Horton et al., 2012) (Pisupati et al., 2017). We later carried out eGWAS using the 897 accessions from the native range removing accessions from North America and the British Isles and excluding contaminants (Ferrero-Serrano and Assmann, 2019). Under the same significant thresholds, majority of QTLs (74% mean-temp, 91% min-temp and 60% max-temp) from eGWAS-1129 are also found in eGWAS-897. Among the top 30 QTLs that were further analyzed in this study, 24 were identified from both eGWAS-1129 and eGWAS-897, while six additional QTLs (three mean- and three max-) from eGWAS-1129 were also identified from eGWAS-897 under a lower significant threshold. Therefore, the use of the 1129 accessions, though not ideal, did not significantly impact the QTLs identified.

In sum, this is the first study to our knowledge that tests the function of a relatively large number of genes associated with local environment parameters and shows the viability of eGWAS in revealing potential molecular mechanisms of local adaptation. Additional QTLs identified from these three eGWAS can be investigated to reveal more genes involved in temperature adaptation. With a similar design, eGWAS will likely be effective in identifying gene variants for other environmental factors such as salinity, precipitation, and seasonality in Arabidopsis. It may also

be used to reveal local adaptation in other plant species. Harnessing the rich genetic diversity in plants will greatly enhance our understanding of plant adaptation to environment and provide means to cope with global climate change.

Acknowledgement

This study was designed by J Hua, Z Wang, and Y Jiang. Y Jiang, Z Wang, H Du, and R Dong conducted the experiments. Y Jiang, Z Wang, and J Hua analyzed the data and wrote the manuscript. J Hua and Y Yuan supervised the project. Research in J Hua's lab is supported by NSF IOS-1946174. Y Jiang and H Du were supported by a fellowship from China Scholarship Council.

Reference

- ALONSO, J. M., STEPANOVA, A. N., LEISSE, T. J., KIM, C. J., CHEN, H., SHINN, P., STEVENSON, D. K., ZIMMERMAN, J., BARAJAS, P., CHEUK, R., GADRINAB, C., HELLER, C., JESKE, A., KOESEMA, E., MEYERS, C. C., PARKER, H., PREDNIS, L., ANSARI, Y., CHOY, N., DEEN, H., GERALT, M., HAZARI, N., HOM, E., KARNES, M., MULHOLLAND, C., NDUBAKU, R., SCHMIDT, I., GUZMAN, P., AGUILAR-HENONIN, L., SCHMID, M., WEIGEL, D., CARTER, D. E., MARCHAND, T., RISSEEUW, E., BROGDEN, D., ZEKO, A., CROSBY, W. L., BERRY, C. C. & ECKER, J. R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 301, 653-7.
- ANDRADE-PIEDRA, J. L., FORBES, G. A., SHTIENBERG, D., GRUNWALD, N. J., CHACON, M. G., TAIPE, M. V., HIJMANS, R. J. & FRY, W. E. 2005. Qualification of a Plant Disease Simulation Model: Performance of the LATEBLIGHT Model Across a Broad Range of Environments. *Phytopathology*, 95, 1412-22.
- BENTSINK, L., JOWETT, J., HANHART, C. J. & KOORNNEEF, M. 2006. Cloning of DOG1, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proc Natl Acad Sci U S A*, 103, 17042-7.
- BHATIA, C., PANDEY, A., GADDAM, S. R., HOECKER, U. & TRIVEDI, P. K. 2018. Low Temperature-Enhanced Flavonol Synthesis Requires Light-Associated Regulatory Components in *Arabidopsis thaliana*. *Plant Cell Physiol*, 59, 2099-2112.

CARLOS, A.-B., JORGE, A., CLAUDE, B., FELIX, B., JOY, B., KARSTEN, M. B., JUN, C.,
EUNYOUNG, C., TODD, M. D., WEI, D., JOSEPH, R. E., MOISES, E.-A., ASHLEY, F.,
JOFFREY, F., XIANGCHAO, G., DOMINIK, G. G., ANGELA, M. H., STEFAN, R. H.,
SVANTE, H., MATTHEW, H., MIKE, J., RANDALL, A. K., ARTHUR, K., PAMELA,
K., CHRISTA, L., CHENG-RUEI, L., DAZHE, M., TODD, P. M., RICHARD, M., NI
WAYAN, M., THOMAS, N., MATTHIAS, N., VIKTORIA, N., MAGNUS, N., POLINA
YU, N., PICÓ, F. X., ALEXANDER, P., FERNANDO, A. R., ALEX, R., BETH, A. R.,
PATRICE, A. S., KARL, J. S., ROBERT, J. S., ÜMIT, S., FELICE GIANLUCA, S.,
MITCHELL, S., HANNES, S., MATT, M. T., DONALD, T., SAMUEL, L. V.,
CONGMAO, W., GEORGE, W., XI, W., WOLFRAM, W., DETLEF, W. & XUEFENG,
Z. 2016. 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis*
thaliana. *Cell*, 166, 481-491.

CARLOS ALONSO-BLANCO. JORGE ANDRADE. CLAUDE BECKER, F. B., JOY
BERGELSON, KARSTEN M. BORGWARDT, JUN CAO, EUNYOUNG CHAE, TODD
M. DEZWAAN, WEI DING, JOSEPH R. ECKER, MOISES EXPOSITO-ALONSO,
ASHLEY FARLOW, JOFFREY FITZ, XIANGCHAO GAN, DOMINIK G. GRIMM,
ANGELA M. HANCOCK, STEFAN R. HENZ, SVANTE HOLM, MATTHEW
HORTON, MIKE JARSULIC, RANDALL A. KERSTETTER, ARTHUR KORTE,
PAMELA KORTE, CHRISTA LANZ, CHENG-RUEI LEE, DAZHE MENG, TODD P.
MICHAEL, RICHARD MOTT, NI WAYAN MULIYATI, THOMAS NÄGELE,
MATTHIAS NAGLER, VIKTORIA NIZHYNKA, MAGNUS NORDBORG, POLINA
YU. NOVIKOVA, F. XAVIER PICÓ, ALEXANDER PLATZER, FERNANDO A.
RABANAL, ALEX RODRIGUEZ, BETH A. ROWAN, PATRICE A. SALOMÉ, KARL
J. SCHMID, ROBERT J. SCHMITZ, ÜMIT SEREN, FELICE GIANLUCA SPERONE,
MITCHELL SUDKAMP, HANNES SVARDAL, MATT M. TANZER, DONALD TODD,
SAMUEL L. VOLCHENBOUM, CONGMAO WANG, GEORGE WANG, XI WANG,
WOLFRAM WECKWERTH, DETLEF WEIGEL, XUEFENG ZHOU, 2016. 1,135
Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell*, 166,
481-491.

CHOW, K. C. & TUNG, W. L. 1998. Overexpression of *dnaK/dnaJ* and *groEL* confers freeze
tolerance to *Escherichia coli*. *Biochem Biophys Res Commun*, 253, 502-5.

CROWELL, S., KORNILIEV, P., FALCAO, A., ISMAIL, A., GREGORIO, G., MEZEY, J. & MCCOUCH, S. 2016. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat Commun*, 7, 10527.

DING, Y., SHI, Y. & YANG, S. 2020. Molecular Regulation of Plant Responses to Environmental Temperatures. *Mol Plant*, 13, 544-564.

FERRERO-SERRANO, A. & ASSMANN, S. M. 2019. Phenotypic and genome-wide association with the local environment of *Arabidopsis*. *Nat Ecol Evol*, 3, 274-285.

FOURNIER-LEVEL, A., KORTE, A., COOPER, M. D., NORDBORG, M., SCHMITT, J. & WILCZEK, A. M. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science*, 334, 86-9.

GRUNEWALD, S., MARILLONNET, S., HAUSE, G., HAFERKAMP, I., NEUHAUS, H. E., VESS, A., HOLLEMANN, T. & VOGT, T. 2020. The Tapetal Major Facilitator NPF2.8 Is Required for Accumulation of Flavonol Glycosides on the Pollen Surface in *Arabidopsis thaliana*. *Plant Cell*, 32, 1727-1748.

HANCOCK, A. M., BRACHI, B., FAURE, N., HORTON, M. W., JARYMOWYCZ, L. B., SPERONE, F. G., TOOMAJIAN, C., ROUX, F. & BERGELSON, J. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, 334, 83-6.

HANSEN, C. C., NELSON, D. R., MOLLER, B. L. & WERCK-REICHHART, D. 2021. Plant cytochrome P450 plasticity and evolution. *Mol Plant*, 14, 1772.

HEREFORD, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *Am Nat*, 173, 579-88.

HORTON, M. W., HANCOCK, A. M., HUANG, Y. S., TOOMAJIAN, C., ATWELL, S., AUTON, A., MULIYATI, N. W., PLATT, A., SPERONE, F. G., VILHJALMSSON, B. J., NORDBORG, M., BOREVITZ, J. O. & BERGELSON, J. 2012. Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nat Genet*, 44, 212-6.

HORTON, M. W., WILLEMS, G., SASAKI, E., KOORNNEEF, M. & NORDBORG, M. 2016. The genetic architecture of freezing tolerance varies across the range of *Arabidopsis thaliana*. *Plant Cell Environ*, 39, 2570-2579.

IRANI, S. & TODD, C. D. 2016. Ureide metabolism under abiotic stress in *Arabidopsis thaliana*. *J Plant Physiol*, 199, 87-95.

JI, H., WANG, Y., CLOIX, C., LI, K., JENKINS, G. I., WANG, S., SHANG, Z., SHI, Y., YANG, S. & LI, X. 2015. The Arabidopsis RCC1 Family Protein TCF1 Regulates Freezing Tolerance and Cold Acclimation through Modulating Lignin Biosynthesis. *PLoS Genet*, 11, e1005471.

KAWAKATSU, T., HUANG, S. C., JUPE, F., SASAKI, E., SCHMITZ, R. J., URICH, M. A., CASTANON, R., NERY, J. R., BARRAGAN, C., HE, Y., CHEN, H., DUBIN, M., LEE, C. R., WANG, C., BEMM, F., BECKER, C., O'NEIL, R., O'MALLEY, R. C., QUARLESS, D. X., GENOMES, C., SCHORK, N. J., WEIGEL, D., NORDBORG, M. & ECKER, J. R. 2016. Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell*, 166, 492-505.

KIM, H. B., SCHALLER, H., GOH, C. H., KWON, M., CHOE, S., AN, C. S., DURST, F., FELDMANN, K. A. & FEYEREISEN, R. 2005. Arabidopsis cyp51 mutant shows postembryonic seedling lethality associated with lack of membrane integrity. *Plant Physiology*, 138, 2033-2047.

KOBAYASHI, Y., KAYA, H., GOTO, K., IWABUCHI, M. & ARAKI, T. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, 286, 1960-2.

KORTE, A. & FARLOW, A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9, 29.

LASKY, J. R., DES MARAIS, D. L., MCKAY, J. K., RICHARDS, J. H., JUENGER, T. E. & KEITT, T. H. 2012. Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol Ecol*, 21, 5512-29.

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.

LI, C., SUN, B., LI, Y., LIU, C., WU, X., ZHANG, D., SHI, Y., SONG, Y., BUCKLER, E. S., ZHANG, Z., WANG, T. & LI, Y. 2016. Numerous genetic loci identified for drought tolerance in the maize nested association mapping populations. *BMC Genomics*, 17, 894.

LU, S., ZHU, T., WANG, Z., LUO, L., WANG, S., LU, M., CUI, Y., ZOU, B. & HUA, J. 2021. Arabidopsis immune-associated nucleotide-binding genes repress heat tolerance at the reproductive stage by inhibiting the unfolded protein response and promoting cell death. *Mol Plant*, 14, 267-284.

LUHUA, S., HEGIE, A., SUZUKI, N., SHULAEV, E., LUO, X., CENARIU, D., MA, V., KAO,
 S., LIM, J., GUNAY, M. B., OOSUMI, T., LEE, S. C., HARPER, J., CUSHMAN, J.,
 GOLLERY, M., GIRKE, T., BAILEY-SERRES, J., STEVENSON, R. A., ZHU, J. K. &
 MITTLER, R. 2013. Linking genes of unknown function with abiotic stress responses by
 high-throughput phenotype screening. *Physiol Plant*, 148, 322-33.

MARTINEZ-BERDEJA, A., STITZER, M. C., TAYLOR, M. A., OKADA, M., EZCURRA, E.,
 RUNCIE, D. E. & SCHMITT, J. 2020. Functional variants of DOG1 control seed chilling
 responses and variation in seasonal life-history strategies in *Arabidopsis thaliana*. *Proc
 Natl Acad Sci U S A*, 117, 2526-2534.

MORGAN, J. A. 1984. Interaction of water supply and N in wheat. *Plant Physiol*, 76, 112-7.

PANDIAN, B. A., SATHISHRAJ, R., DJANAGUIRAMAN, M., PRASAD, P. V. V. &
 JUGULAM, M. 2020. Role of Cytochrome P450 Enzymes in Plant Stress Response.
Antioxidants (Basel), 9.

PARMESAN, C. & HANLEY, M. E. 2015. Plants and climate change: complexities and surprises.
Ann Bot, 116, 849-64.

PISUPATI, R., REICHARDT, I., SEREN, U., KORTE, P., NIZHYNSKA, V., KERDAFFREC,
 E., UZUNOVA, K., RABANAL, F. A., FILIAULT, D. L. & NORDBORG, M. 2017.
 Verification of *Arabidopsis* stock collections using SNPmatch, a tool for genotyping high-
 plexed samples. *Sci Data*, 4, 170184.

PORANKIEWICZ, J. & CLARKE, A. K. 1997. Induction of the heat shock protein ClpB affects
 cold acclimation in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J Bacteriol*,
 179, 5111-7.

PRINZENBERG, A. E., CAMPOS-DOMINGUEZ, L., KRUIJER, W., HARBINSON, J. &
 AARTS, M. G. M. 2020. Natural variation of photosynthetic efficiency in *Arabidopsis*
thaliana accessions under low temperature conditions. *Plant Cell Environ*, 43, 2000-2013.

QUINT, M., DELKER, C., FRANKLIN, K. A., WIGGE, P. A., HALLIDAY, K. J. & VAN
 ZANTEN, M. 2016. Molecular and genetic control of plant thermomorphogenesis. *Nat
 Plants*, 2, 15190.

SAVOLAINEN, O., LASCOUX, M. & MERILA, J. 2013. Ecological genomics of local
 adaptation. *Nat Rev Genet*, 14, 807-20.

665 SEREN, U., VILHJALMSSON, B. J., HORTON, M. W., MENG, D., FORAI, P., HUANG, Y. S.,
 666 LONG, Q., SEGURA, V. & NORDBORG, M. 2012. GWAPP: a web application for
 667 genome-wide association mapping in Arabidopsis. *Plant Cell*, 24, 4793-805.

668 SHIN, H. Y. & NAM, K. H. 2018. RAV1 Negatively Regulates Seed Development by Directly
 669 Repressing MINI3 and IKU2 in Arabidopsis. *Mol Cells*, 41, 1072-1080.

670 TABAS-MADRID, D., MENDEZ-VIGO, B., ARTEAGA, N., MARCER, A., PASCUAL-
 671 MONTANO, A., WEIGEL, D., XAVIER PICO, F. & ALONSO-BLANCO, C. 2018.
 672 Genome-wide signatures of flowering adaptation to climate temperature: Regional
 673 analyses in a highly diverse native range of Arabidopsis thaliana. *Plant Cell Environ*, 41,
 674 1806-1820.

675 TEIGEN, L. E., ORCZEWSKA, J. I., MCLAUGHLIN, J. & O'BRIEN, K. M. 2015. Cold
 676 acclimation increases levels of some heat shock protein and sirtuin isoforms in threespine
 677 stickleback. *Comp Biochem Physiol A Mol Integr Physiol*, 188, 139-47.

678 WANG, S., BAI, G., WANG, S., YANG, L., YANG, F., WANG, Y., ZHU, J. K. & HUA, J. 2016.
 679 Chloroplast RNA-Binding Protein RBD1 Promotes Chilling Tolerance through 23S rRNA
 680 Processing in Arabidopsis. *PLoS Genet*, 12, e1006027.

681 WANG, Z., YANG, L., LIU, Z., LU, M., WANG, M., SUN, Q., LAN, Y., SHI, T., WU, D. &
 682 HUA, J. 2019. Natural variations of growth thermo-responsiveness determined by
 683 SAUR26/27/28 proteins in Arabidopsis thaliana. *New Phytol*, 224, 291-305.

684 WANG, Z., YANG, L., WU, D., ZHANG, N. & HUA, J. 2021. Polymorphisms in cis-elements
 685 confer SAUR26 gene expression difference for thermo-response natural variation in
 686 Arabidopsis. *New Phytol*, 229, 2751-2764.

687 WEIGEL, D. & MOTT, R. 2009. The 1001 genomes project for Arabidopsis thaliana. *Genome*
 688 *Biol*, 10, 107.

689 XI, J., QIU, Y., DU, L. & POOVAIAH, B. W. 2012. Plant-specific trihelix transcription factor
 690 AtGT2L interacts with calcium/calmodulin and responds to cold and salt stresses. *Plant*
 691 *Sci*, 185-186, 274-80.

692 XIAO, Y., LIU, H., WU, L., WARBURTON, M. & YAN, J. 2017. Genome-wide Association
 693 Studies in Maize: Praise and Stargaze. *Mol Plant*, 10, 359-374.

- XU, R. R., QI, S. D., LU, L. T., CHEN, C. T., WU, C. A. & ZHENG, C. C. 2011. A DExD/H box RNA helicase is important for K⁺ deprivation responses and tolerance in *Arabidopsis thaliana*. *Febs Journal*, 278, 2296-2306.
- ZAGOTTA, M., SHANNON, S., JACOBS, C. & MEEKS-WAGNER, D. 1992. Early-Flowering Mutants of *Arabidopsis thaliana*. *Functional Plant Biology*, 19, 411-418.
- ZHANG, X., ZHU, Y., LIU, X., HONG, X., XU, Y., ZHU, P., SHEN, Y., WU, H., JI, Y., WEN, X., ZHANG, C., ZHAO, Q., WANG, Y., LU, J. & GUO, H. 2015. Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in *Arabidopsis*. *Science*, 348, 120-3.
- ZHAO, L. & KUNST, L. 2016. SUPERKILLER Complex Components Are Required for the RNA Exosome-Mediated Control of Cuticular Wax Biosynthesis in *Arabidopsis* Inflorescence Stems. *Plant Physiol*, 171, 960-73.

Tables

Table 1 Summary of associated genes identified from environmental GWAS analyses.

Listed are eGWAS, QTL name, Gene ID, annotation, and gene name for each associated gene.

The gene annotation was obtained from TAIR. ‘Δ’ indicates the shared gene (*CSLB05*) by mean-temp and max-temp eGWAS. ‘\’ indicates the genes not included in further analysis.

Table 2 Summary of phenotypes of mutants grown at different temperatures.

Listed are phenotypes of two mutant alleles of the associated genes in 15 measurements at 5 temperatures. Phenotypes are compared to the wild-type Col-0: larger (L) or smaller (S) in growth, delayed (D) or early (E) in flowering time, reduced (R) in relative water content. Significant difference between mutant and the wild type was assessed by Student’s *t* test. Upper case indicates $p < 0.01$, lower case indicates $0.05 < p < 0.01$, and ‘ns’ indicates no significant difference. Mutant phenotype is highlighted by green (small, early or reduced) or pink (large or delayed) when both alleles exhibited significant differences ($p < 0.05$) and at least one allele exhibited high significant difference ($p < 0.01$) from the wild type plant. ‘Δ’ indicates the mutants of the shared gene (*CSLB05*) detected by mean-temp and max-temp eGWAS.

Table 3 Summary of the effects of temperature and genotype on rosette area, rosette weight and seed production.

Two-way ANOVA analyses were performed on rosette area, fresh weight and seed production of plants grown at 16°C, 22°C and 28°C. The *p* value and FDR were shown in the table. False discovery rate adjusted *p* values (FDR) were calculated with Benjamini Hochberg procedure. Significances are indicated by asterisk: ***FDR < .001; **FDR < .01; *FDR < .05. G: genotype; GxE: genotype x environment. Light blue: significant impact from genotype; Yellow: significant impact from interaction of genotype and environment. Temperature significantly impacts all phenotypes analyzed. The first allele of mutants was assessed by the analyses. ‘Δ’ indicates the shared gene (*CSLB05*) by mean-temp and max-temp eGWAS.

Figure legends

Figure 1 Genome wide association study (GWAS) of local climatic temperature variables in *Arabidopsis thaliana* natural accessions.

(A-C) Distributions of mean-temp (annual mean temperature) (A), max-temp (max temperature of warmest month) (B), and min-temp (minimum temperature of coldest month) (C) of 1129 *Arabidopsis* natural accessions. (D-F) Manhattan plots of GWAS of mean-temp (D), max-temp (E) and min-temp (F). Red dashed horizontal line indicates significant threshold. (G) Number of SNPs and QTLs identified from each GWAS analysis. (H-I) Venn diagram of QTLs (H) and SNPs (I) identified from GWAS analyses of mean-temp, max-temp and min-temp.

Figure 2 Summary of biofactor associated genes that are identified with a function in growth under at least one of the five temperatures.

(A) Venn diagram of genes whose loss-of-function mutants exhibited growth differences compared to the wild-type Col-0 at five temperatures: 10°C (purple), 16°C (blue), 22°C (green), 28°C (yellow), and 35°C (red). Genes in blue, green, and brown were identified from their association with min-temp, mean-temp and max-temp, respectively. (B) Percentage of genes that have a detected function for each group of the associated genes. Color scheme for associated genes is the same as in (A). (C) *Chi*-square test on the phenotype proportions for each group.

Figure 3 Geographic and climatic distribution of polymorphisms in five associated genes identified from environmental GWAS.

Allele distributions of single nucleotide polymorphisms (SNPs) in *ERF53* (A), *TCF* (B), *SKI2* (C), *ALN* (D-E), *CSLB05* (F). Insertion shows the density plot of the major (in blue) allele and the minor (in red) allele along the min-temp (A, E), mean-temp (B, C) and max-temp (D, F) of accessions locales. The *p* value is from Welch's *t*-test assessing differences of the relevant local temperature variable of minor allele to major allele. The predicted activity and effect on phenotype of each allele was shown.

Supplemental information

Supplementary Figure 1 Correlation coefficient of the 11 local temperature variables.

Supplementary Figure 2 Schematic diagrams of T-DNA insertion mutations in the 29 associated genes.

Supplementary Figure 3 Quantification of rosette area, hypocotyl length and root length of mutants grown at 10°C.

Supplementary Figure 4 Quantification of root length and relative water content of mutants grown at 35°C.

Supplementary Figure 5 Correlation coefficient of 15 phenotypes analyzed.

Supplementary Figure 6 Quantification of rosette area, rosette weight and seed production of mutants grown at 16°C.

Supplementary Figure 7 Quantification of rosette area, rosette weight and seed production of mutants grown at 22°C.

Supplementary Figure 8 Quantification of rosette area, rosette weight and seed production of mutants grown at 28°C.

Supplementary Figure 9 Flowering times of mutants grown at 22°C.

Supplementary Figure 10 Flowering time of *tsf-2* mutant and Col-0 grown at 22°C.

Supplementary Figure 11 Polymorphisms at five associated genes and their impacts on gene function.

Supplementary Figure 12 Geographic and climatic distribution of polymorphisms in 16 genes identified from environmental GWAS.

Supplementary Figure 13 Polymorphisms of candidate genes from eGWAS are associated with plant growth variations at 10 °C in Arabidopsis natural accessions.

Supplementary Table 1 Summary of the information of T-DNA insertion mutants.

Supplementary Dataset 1: Summary of the SNPs information in the 30 QTLs evaluated.

Supplementary Dataset 2: Summary of the results of statistical tests on phenotyping data.

Supplementary Dataset 3: Quantification of rosette size and weight of the 53 natural accessions at 10°C.