The mRNA covalent modification dihydrouridine positively impacts photosynthesis-associated genes in field-grown drought-stressed sorghum

Li'ang Yu[1], Giovanni Melandri[2,36], Anna C. Nelson Dittrich[1], Sebastian Calleja[2], Bruno Rozzi[2], Kyle Palos[1], Aparna Srinivasan[1], Emily K. Brewer[1], Hillary Fischer[4], Toshihiro Obata[5], Hamada Abd Elgawad[6], Gerrit T.S. Beemster[6], Riley Henderson[1], Ciara Denise Garcia[2], Xiaodan Zhang[1], David Stern[1], Andrea Eveland[7], Susan J. Schroeder[8], , Aleksandra Skirycz[4], Eric Lyons[2], A. Elizabeth Arnold[2,3], Brian D. Gregory[9]*, Andrew D. L. Nelson[1]*, Duke Pauli[2,3]*

[1]Boyce Thompson Institute, Cornell University, Ithaca, NY 14850, USA
[2]School of Plant Sciences, University of Arizona, Tucson, AZ, 85721, USA
[3]Center for Agroecosystem Research in the Desert (ARID), University of Arizona, Tucson, AZ 85721, USA
[4]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, 48824, USA
[5]Department of Biochemistry and Center for Plant Science Innovation, University of Nebraska-Lincoln, NE, 68588, USA
[6]Integrated Molecular Plant Physiology Research, University of Antwerp, 2020 Antwerp, Belgium
[7]Donald Danforth Plant Science Center, Olivette, MO, 63131, USA
[8]Department of Chemistry and Biochemistry, The University of Oklahoma, Norman, OK, 73019, USA
[9]Department of Biology, University of Pennsylvania, Philadelphia, PA, US

***Address correspondence to:**
Dr. Duke Pauli
The University of Arizona, School of Plant Sciences
1140 E. South Campus Drive, Tucson, AZ, 85721
dukepauli@arizona.edu

or
Dr. Andrew D. L. Nelson
The Boyce Thompson Institute, Cornell University
533 Tower Road, Ithaca, NY, 14850.
an425@cornell.edu

or
Dr. Brian D. Gregory
University of Pennsylvania, Department of Biology, 103 H Lynch Laboratory
433 South University Avenue, Philadelphia, PA, 19104
bdgregor@sas.upenn.edu

Current Email addresses:
LY: ly76@cornell.edu, GM: gmelandri@arizona.edu, ACND: an426@cornell.edu, SC: scalleja@arizona.edu, BR: brozzi@ufl.edu, KP: krp75@cornell.edu, ApS: as2967@cornell.edu, EKB: ekb78@cornell.edu, HF: fisch333@msu.edu, TO: tobata2@unl.edu, HAE: hamada.abdelgawad@uantwerp.be, GTSB: gerrit.beemster@uantwerp.be, RH: rnh53@cornell.edu, CDG: ciarag@arizona.edu, XZ: xz828@cornell.edu, DS: ds28@cornell.edu, AE: AEveland@danforthcenter.org, SJS: susan.schroeder@ou.edu, AlS: skirycza@msu.edu, EL: ericlyons@arizona.edu, AEA: barnoldaz@gmail.com

Running title: The RNA Modification Dihydrouridine Positively Impacts Plant Stress Responses

Abstract

RNA Covalent Modifications (RCMs) are post-transcriptional changes to RNA's chemical composition. RCMs influence mRNA stability, regulate transcription and translation, and play critical roles in eukaryotic growth and development. However, their role in plants remains poorly understood, particularly in plant field performance under stress conditions. In this study, we grew six diverse *Sorghum bicolor* accessions in the field and examined their physiological and molecular responses to drought and heat stress over time. A detailed systems-level approach unexpectedly uncovered a positive association between the enzyme dihydrouridine synthase (SbDUS2), which synthesizes the deeply conserved RCM dihydrouridine (DHU), photosynthesis, and a core group of photosynthesis-associated transcripts. We then characterized loss-of-function DUS2 mutants in Arabidopsis, revealing that DHU's role in performance and photosynthesis under stress is likely conserved in plants. These results reveal, for the first time in plants, that the DHU modification is a critical and conserved stress-tolerance conferring post-transcriptional regulatory mechanism.

Keywords:

Drought response, RNA covalent modifications, Gene regulatory network, Systems biology

Introduction

Sorghum bicolor (sorghum) is the fifth most important crop worldwide and is grown for grain, brewing, and cellulosic biomass production. As a member of the *Poaceae* and a close relative of maize, sorghum is believed to have been domesticated around 4,000–6,000 years ago in Africa, and to have then diverged into a variety of landraces with diversified morphological variation due to selection across broad geographical locations[1]. Most sorghum cultivars were historically, and are currently, grown in arid and semi-arid areas due to their tolerance to limited irrigation during growth. However, when exposed to prolonged and severe water deficits sorghum experiences reduced in a number of agronomic traits such as biomass and grain yield[2,3]. These traits are controlled by multiple genes and influenced by a number of developmental and environmental cues. Thus, a better understanding of the genetic mechanisms able to interpret and coordinate responses to these environmental cues is essential for maintaining yield in sorghum under stress conditions[4].

Drought is a form of abiotic stress that leads to numerous unfavorable phenotypic and physiological changes such as reduced yield and altered plant growth, and can lead to increased susceptibility to biotic stressors. Plant phenotypic responses to reduced water availability are often predicated physiological and molecular responses, with more tolerant plants usually displaying an increased water use efficiency (WUEi), often achieved by reduced transpiration, that, in turn, lowers down CO_2 net assimilation. Metabolite and enzyme activity levels also change in response to drought, as commonly seen by the increased levels of specific soluble sugars and amino acids able to act as osmoprotectants [5–8]. Combined, these phenotypic, physiological, and metabolic traits can serve as key indicators to interpret the levels of drought tolerance among different accessions within a species.

Drought responses are coordinated at the cellular level through a complex set of regulatory control points, guided by the activation and repression of transcription factors (TFs) that in turn regulate the expression of their target genes[9–11]. These regulatory networks receive and interpret cues from the environment and then govern the transcriptional cascades that ultimately link stress signaling to downstream phenotypes such as yield[12,13]. Systems-level approaches have been quite successful at identifying the elements that influence stress and yield in several crop species[14–16], although much of this work is often done in a single crop accession. Thus, how these findings translate across crop genetic diversity is poorly understood. While large population-level studies of sorghum have identified strong drought tolerance phenotypes useful for breeding and selection[17–19], the crosstalk between many upstream regulatory genes and downstream trait-associated genes remains unclear and largely unexplored. Thus, we are still limited in the genetic and molecular targets at our disposal to develop new cultivars that show improved stress response without the off-target effects that come from expressing a transcriptional regulator outside of its normal context.

Often the act of transcription is used as a direct proxy for translation, and therefore insights on protein function are assumed based on RNA abundance. However, RNAs encounter a number of co- and post-transcriptional events that influence their stability, translation, and even cellular localization, that are not immediately apparent in RNA-sequencing data. One such event is the addition of RNA covalent modifications (RCMs). RCMs, which comprise the so-called epitranscriptome, were initially discovered in the 1970s on tRNAs and rRNAs, but are now known to impact all classes of RNAs (e.g., mRNAs, lncRNAs, and miRNAs), influencing their stability, translation efficiency (mRNAs), structure, and splicing[20–22]. To date, more than 100 RCMs have been discovered, and some, in particular N^6-methyladenosine (m^6A), are known to influence plant development and stress responses[23–26]. Several other RCMs, such as dihydrouridine (DHU), pseudouridine (Ψ), N^1-methyladenosine (m^1A), and 5-methylcytidine (m^5C), are less studied but are also believed to be critical for RNA turnover and translation[20,21]. However, the functional roles of these latter RCMs are poorly understood in plants, predominantly due to difficulty in their detection.

Detection of RCMs generally relies on antibody-based (e.g. m^6A) or modification-specific approaches using chemical probing or conversions (e.g. m^5C and Ψ, respectively)[27–30]. However, RCMs that occur along the Watson-Crick-Franklin base pairing edges can be detected in Illumina RNA-seq data due to nucleotide misincorporation during the reverse transcription step of library preparation. This RT-induced nucleotide misincorporation is non-random, and several statistical algorithms have been developed to infer modification class from the pattern of non-biallelic SNPs present in conventional transcriptome sequencing[31–33]. These algorithms provide us with the ability to monitor both transcript abundance and RCM abundance genome-wide and associate these molecular phenotypes with changes at the whole plant level in response to environmental perturbations.

To identify the species-wide genetic basis by which sorghum effectively responds to drought stress, we leveraged a systems biology approach to study six phylogenetically distinct sorghum accessions with different levels of drought tolerance. These accessions were field-evaluated in the hot, arid environment of the Arizona low desert and exposed to two different watering regimes after flowering. We collected matched transcriptomic, metabolomic, oxidative stress status, physiological, as well as leaf-level and whole plant phenotypic data at multiple points throughout the season to monitor responses to soil water conditions. Gene co-expression networks and TF binding data were integrated with all of the other measured traits to identify the sets of TFs coordinating the phenotypic responses to drought that we observed. Epitranscriptomic data, inferred from Illumina RNA-seq, were also integrated into our analyses, uncovering associations between RCM abundance, phenotype, and plant performance within these accessions. The levels of one such RCM, DHU, and the enzyme responsible for its deposition, Dihydrouridine Synthase (SbDUS2) were positively associated with plant

performance traits (e.g. CO_2 net assimilation) during drought treatment. Despite its deep evolutionary conservation in living organisms, DUS proteins are minimally characterized in plants. Thus, we examined loss-of-function *dus2* mutants in *Arabidopsis thaliana*, where we discovered reduced germination, growth rates, and CO_2 net assimilation under favorable conditions that were exacerbated under abiotic stress. A transcriptomic time course comparison between Arabidopsis wild type and *dus2* mutants under control and stress conditions revealed a decrease in DHU levels and a concomitant reduction in photosynthesis-associated transcript abundance in the mutant plants, similar to what we observed in sorghum in the field. In total, our integrative systems approach uncovered novel protein and RNA-level control points that impact responses to drought stress in sorghum and are strong candidates for targeted breeding approaches.

Methods and Materials

Sorghum growth and phenotypic measurement

Six selected sorghum accessions (PI 533871: M1, PI 533961: DL/59/1530, PI 656041: 80M, PI 656053: N290-B, PI 656076: SC1271, and PI 656096: SC391) were grown at the University of Arizona Maricopa Agricultural Center (MAC). All plants were well-watered (WW) at the same rate (~24% soil volume metric water content, SVWC), corresponding approximately to the soil field capacity) until flag leaf appearance in the whorl (~47 days after planting), after which irrigation was reduced to a subset of plots to mimic drought treatment (Water-limited, WL; 16% SVWC). The destructive plant phenotyping was conducted at the end of the season (maturity stage). Measured traits included above ground biomass, root biomass, and three-leaf weights. For other time-series molecular phenotypes, samples were collected weekly (on Thursdays) for seven weeks (week 1-week 7, from August 13th to September 24th, 2020) on a specific range (Range 28) of a large field experiment conducted at the University of Arizona's Maricopa Agricultural Center (MAC), Maricopa, AZ, US. For transcriptomics, metabolomics, and analysis of other biochemical traits, samples consisting of six leaf disks from five different plants per accession, treatment, and time point were collected (from 10:30 to 11:30) in three different 1.5 mL tubes and immediately snap-frozen in liquid nitrogen. LI-6800 derived photosynthesis-related traits (*N = 29*), (https://www.licor.com, LI-COR Lincoln, NE) were measured in the morning (from 9:00 to 10:00) and afternoon (from 13:00 to 14:00) with the following settings: Flow rate = 600 µmol s^{-1}; Mixing fan speed =12,000 rpm; CO_2 = 420 ppm; Light= 2,000 µmol m^{-2} s^{-1}. The youngest fully expanded leaf on the main stem of each plant was selected for the measurements and data were logged as soon as assimilation and stomatal conductance curves (monitored in real-time on the instrument display) when reaching a plateau (on average 30 seconds after a leaf was clamped).

Sorghum biochemical traits measurement

Biochemical parameters, including levels of highly concentrated metabolites (e.g., total soluble sugars), levels of oxidative stress markers and molecular antioxidants, and antioxidant enzyme and photorespiration enzyme activities were analyzed following established protocols in previous study[34] Chlorophyll and carotenoid contents were determined spectrophotometrically after acetone extraction[35]. Sugar levels were quantified spectrophotometrically using enzyme-coupled assays[36]. Oxidative stress markers, such as malondialdehyde (MDA) and protein carbonyl content, were measured to assess lipid peroxidation and protein oxidation[37,38]. Photorespiration-related enzyme activities, glycolate oxidase (GOX) and hydroxypyruvate reductase (HPR), were assessed based on established methodologies[39,40]. Lipoxygenase (LOX) was extracted in potassium phosphate buffer (pH 7.0), and its activity was measured by changes in conjugated dienes[41]. Total antioxidant capacity (TAC), total polyphenol

content (Poly), and flavonoids were evaluated using standard spectrophotometric assays[42,43]. Enzyme activities of key antioxidants, including ascorbate peroxidase (APX), dehydroascorbate reductase (DHAR), monodehydroascorbate reductase (MDHAR), glutathione reductase (GR), peroxidase (POX), superoxide dismutase (SOD), glutathione peroxidase (GPX), catalase (CAT), ascorbate oxidase (AO), and glutathione *S*-transferase (GST), were determined using microplate-based kinetic assays[44–46].

Sorghum metabolites profiling

Primary sorghum metabolites were quantified using gas chromatography-mass spectrometry (GC-MS) according to published protocol[47]. Briefly, metabolites were extracted by methanol: water: chloroform solution from 30±3 mg fresh weight (FW) of frozen samples. Ribitol was added to the extraction solution as internal standard (IS). An aliquot of the upper aqueous phase was dried by a centrifugal vacuum concentrator. Dried metabolites were derivatized by methoxyamination and trimethylsylilation and analyzed by a 7200 GC-QTOF system (Agilent, St Clara, CA, USA) using a HP-5MS-UI column (30 m 0.25 mm ID 0.25 µm; Agilent) with 100 x split and splitless modes. Metabolite peaks were annotated by MassHunter Unknowns Anaysis (Agilent) using the Fiehn Metabolomics Library (Agilent). The 42 most confidently annotated metabolites were quantified by MassHunter Quantitation (Agilent) and the peak heights were background subtracted, normalized by IS and precise FW of the sample to calculate the relative metabolite contents values.

Arabidopsis growth and phenotypic measurement

Arabidopsis Col-0 and *dus2* mutant lines (Line ID: CS857388, https://abrc.osu.edu) were grown in a greenhouse facility growth chamber at the Boyce Thompson Institute, Ithaca NY for the initial screening of phenotypic variation. The growth conditions were set as follows: temperature: 22°C (control condition) or 37°C (heat treatment), humidity: 60%, and lighting period: 8:00 - 21:00. All seeds were stratified at 4°C for 48~72 hours after planting. For image-based phenotyping, plants were placed under Raspi-controlled phenotyping tools (Pheno-Rigs)[48] to collect time-series plant images (imaging-interval: 30 mins**,** length: 12 days), which were then processed by PlantCV pipelines to extract rosette area represented by RGB images[49], followed by statistical analyses and plots as previously described[48]. Germination assays were performed on ½ MS medium with Col-0 and *dus2* mutant seeds placed side-by-side at 22°C, and the number of germinated seeds was recorded during each 3-hour interval for 27 hours after seed stratification[50]. The photosynthetic efficiency of *dus2* and Col-0 lines was evaluated using a gas exchange approach with the LI-6800 (LI-COR Lincoln, NE). Briefly, plants were maintained in the growth chamber under short-day (SD) growth conditions (photo-period: 8:00-

17:00, temp: 22°C) to ensure a sufficient rosette area to cover the 2cm^2 small chamber of the LI-6800 device[51]. At 4-weeks after germination, half of the plants were kept at 22°C as the control group and the other half were subjected to heat stress conditions (37°C for 24 hours) before LI-6800 measurement. Parameters were set based on previous leaf gas exchange experiments (flow rate = 400 µmol s□¹, relative humidity of air = 55–60□%, reference CO_2 = 400 µmol□¹, mixing fan speed = 10,000□rpm, heat exchanger temperature: 22°C or 37°C, chamber pressure = 0.2 kPa, light source composition 90□% red 10□% blue light, light intensity = 1000 µmol m^{-2} s^{-1}, measurement time = between ZT4 (12:00) and ZT6 (14:00))[51]. To compare the timing of vegetative growth to reproductive growth transition in Col-0 and *dus2* lines, we recorded the rate of bud initiation per Arabidopsis genotypes, as had been done in a previous study[52].

RNA extraction and libraries construction

RNA extraction was performed using methods from previous studies for both Arabidopsis and sorghum[53]. Approximately five flash-frozen ¼" sorghum leaf disks per sample were used as starting material for RNA extractions. RNA extraction was performed using the RNeasy Plant Mini kit (Qiagen #74904) and RNase-free DNase set (Qiagen #79254) according to manufacturer instructions. Depending on total RNA yield for each sample, 2.5-10 ug total RNA was used as starting material for mRNA enrichment using the Dynabeads mRNA purification kit (Invitrogen #61006) according to manufacturer instructions, except with the addition of a second binding and elution step to further enrich mRNA. RNA-seq libraries were then constructed from 10-60 ng mRNA per sample using the Amaryllis Nucleics YourSeq Duet Full Transcript library prep kit with combinatorial dual indexes (now Active Motif #23001 or #23002) according to manufacturer instructions. Prepared libraries were sequenced by Novogene with Illumina NextSeq 500 (PE × 150bp). Arabidopsis RNA was also extracted using Qiagen's RNeasy kit, per manufacturer's recommended protocol. Total Arabidopsis RNA was sent to Novogene for mRNA-enrichment and library preparation, then sequenced on Novogene's NovaSeq X Plus series (PE ×150bp).

5'-6'-Dihydrouridine quantification in Arabidopsis total RNA

1 µg of total RNA per Arabidopsis leaf sample was digested to the nucleotide level using 0.6 units (U) of Nuclease PI from *Penicillium citrinum* (Sigma-Aldrich), 0.2 U phosphodiesterase I from *Crotalus adamanteus* venom (Worthington), 2 U FastAP Thermosensitive Alkaline Phosphatase (Thermofisher), 10 U benzonase (Sigma-Aldrich) and 200 ng Pentostatin (Fisher Scientific) in 25 mM ammonium acetate buffer at pH 7.5 (methodology based on Sudol et al., 2024). Phosphodiesterase

was reconstituted in 0.11 M Tris salt buffer (0.11M NaCl, 15mM $MgCl_2$, pH 8.9) and Nuclease PI in 20 mM acetate buffer (5 mM ZnCl2, 50 mM NaCl, pH 5.3), both filter sterilized 0.2µm. Samples were incubated overnight at 37°C and stored at -80°C for LC/MS. Before LC-MS/MS, samples digestion enzymes were removed by diluting the samples to 75% MeOH before centrifuging (10min at 21,000g). The supernatant was transferred to a new tube and dried down with a speedvac before reconstitution in HPLC-grade water (Sigma). LC–MS/MS analysis was performed using a Waters Acquity UPLC LC with an HSS-T3 premier column (1.8 µM particle size, 100 × 2.1 mm; Waters) and a Waters Xevo TQ-XS Triple Quadrupole mass spectrometer equipped with an electrospray ion source (ESI). Water with 0.1% formic acid was used as solvent A and LC–MS grade acetonitrile (Sigma) served as solvent B. The elution started with 100% solvent A with a flow rate of 0.3 mL/min for 3 min, followed by a linear gradient to 1% solvent B at 4.65 min, followed by a linear gradient to 6% solvent B at 5.35 min that was held until 6 min, followed by a linear gradient to 50% solvent B at 7 min and then 90% solvent B at 8 min that was held until 9.5 min. 100% solvent A was held for further 2.5 min before starting the next elution. During elution nucleosides were monitored with electrospray in the positive (+) mode with ESI parameters as follows: capillary voltage at 1kV, cone voltage at 40V, desolvation temperature 400°C, and source temperature at 150°C. The mass spectrometer was run in the multiple reaction monitoring (MRM) mode using the parent mass of 247.00 m/z and daughter mass of 114.90 m/z to detect dihydrouridine with the MassLynx software. The quantitative analysis was performed using an external calibration curve with a 5,6-dihydrouridine standard (AdooQ Bioscience).

Identification and characterization of RCMs

Identification of RCMs was performed by intersecting results derived from the High-throughput Annotation of Modified Ribonucleotides (HAMR v1.4) and ModTect (v1.1.5.1) algorithms for field-derived sorghum RNA-seq data[31,33]. Initially, the paired-end stranded RNA reads were trimmed by Trimmonmatic (v0.32, parameter: LEADING:3 TRAILING:3 MINLEN:36, TruSeq3-PE.fa:2:30:10:2) to remove adaptors and low-quality reads[54]. For HAMR, the forward and reverse reads were mapped to the sorghum reference genome (v3.1.1[55]) using Tophat2 (v2.1.1, --read-mismatches 12 --read-edit-dist 12, --max-multihits 10, --transcriptome-max-hits 10, --library-type fr-second strand, and --b2-very-sensitive)[56], followed by the merging of mapped, removal of multiple-mapped reads, and resolving of splicing alignment using Picard (v3.1.1, http://broadinstitute.github.io/picard) and GATK (v3.3.5)[57]. The clean alignment files were processed by HAMR with a stringent parameter setting: --min_read_qual 30, --min_read_coverage 50, --seq_error_rate 0.01, --hypothesis H_4, --max_FDR 0.05, which filtered out potential modified sites with a < 50 mapping depth with high-confidence (FDR < 0.05) to call modified sites[31]. For ModTect prediction, trimmed reads were mapped to the sorghum genome and associated

annotation files using STAR (v 2.7.10b[58]) with the default arguments: --out FilterIntronMotifs RemoveNoncanonical, --alignIntronMax 10000, --out SAMstrandField intronMotif. Raw BAM outputs were then used as input for ModTect with the following arguments: --minBaseQual 30, --readlength 150, and --regionFile: bed file of sorghum (v3.1.1) or *Arabidopsis* (TAIR11) gene annotation[55,59]. To harvest the high-confidence PTMs, only the overlapped sites derived from two replicates and two methods were retained as the modified sites for each sorghum sample. For the public sorghum tissue atlas and large-scale time-series drought experiment RNA-seq datasets[55,60], RCMs were identified using only "ModTect" and were limited to RCMs that were present among samples (i.e. three or more distinct SRAs or between replicates).

To simplify the annotation of RCMs relative to transcripts, only primary transcripts (the longest transcripts per gene) were selected to group the number of modified sites that fell into respective exon regions. The output file of HAMR was cross-referenced to the gene annotation files (primary transcripts only) to assign sites to transcripts. Here, three metrics were applied to characterize the RCMs: the presence and absence of RCMs per transcript (PAM), the number of modified sites per transcript (modification density), and the ratio of modified reads over total mapped reads per transcript (modification frequency).

Genomic analyses of sorghum genome and population

Orthologs of important sorghum candidate genes in other species were confirmed using results of CoGeBlast (https://genomevolution.org/coge/CoGeBlast.pl, parameters E-value: 1e-5, word_size: 3, matrix: BLOSUM62) and the public JGI Phytozome 13 annotation information to identify the best hits from *Arabidopsis* (TAIR 11 and Sbicolor_454.v3 [55,59]). The genomic variants of the six sorghum accessions were identified and annotated using the published Sorghum Association Panel (SAP) resequencing data[61]. Briefly, cleaned DNA reads were mapped to the sorghum reference genome using BWA-mem with default parameters (v0.7.17-r1188)[62]. Duplicated read-pairs and low-quality mapping in each sample were cleaned using Sambamba[63] (v1.28.1, parameters: mapping_quality >= 1 and not (unmapped or secondary_alignment) and not ([XA] != null or [SA] != null) followed by read sorting using samtools[64] Variant calling was conducted using the GATK4 HaplotypeCaller function (v4.2.5.00, parameters: --max-alternate-allele:6, --Ploidy:2)[57]. The single nucleotide polymorphisms (SNPs) were filtered using GATK hard-filtering protocol 'QD < 2.0||FS > 60.0|| MQ< 40.0||MQRankSum < 12.5||ReadPosRankSum < 8.0', and insertions and deletions (INDELs) with parameters 'QD < 2.0||FS > 200.0 || ReadPosRankSum < 20.0'. Those high-quality SNPs and INDELs were further annotated by variants effect predictor (VEP) using the *S. bicolor* Sbicolor_454.v3 gene annotation[55,65]. For

phylogenetic analysis, we retained the bi-allelic SNPs at the four-fold degenerate (4DTV) loci filtered by plink (v1.90b4, --types SNPs -m 2 -M 2) that derived from 365 out of 407 sorghum accessions with complete geographical information[61,66]. Using the SNP filtering with minor allele frequency (MAF > 0.05, missing data < 0.1) and linkage disequilibrium (LD) prune (plink v1.90b4 parameter: --indep-pairwise 1000 500 0.2)[66], a total of 7,520 SNPs were kept to reconstruct the phylogenetic tree using IQtree (v1.6.12, Parameter: -nt 100 -m MFP)[67].

Transcriptome data processing

Transcript abundance was measured using the same sets of RNA reads alignment files used for RCM identification. The FeatureCounts tool was used for quantification using the mRNA (transcripts only) as the meta-level to count reads (v2.0.1, parameter: -s 1 -p -t mRNA -g ID -O). Read counts of 34,130 primary transcripts were normalized using DESeq2 after the removal of transcripts with less than an average of one count across 96 samples[68,69]. We incorporated the time point (week 1, week 3, week 5, and week 7), accession IDs (**A:** M1**, B:** DL/59/1530**, C:** 80M**, D:** N290-B**, E:** SC1271**, and F:** SC391), and watering conditions (control and drought) as the experimental design factors. Condition-wise differentially abundant transcripts (DATs) over each time point per accession were identified using the cutoff of adjusted P-value < 0.05 with FDR. A $\log_2$-transformed fold change (FC) greater than 1 was used to filter up-regulated transcripts and a less than -1 $\log_2$FC filtered the down-regulated transcripts.

Construction of co-expression network

The co-expression network was constructed using the WGCNA package with variance stabilizing transformed (VST) normalized read counts with 34,130 transcripts across 96 samples[70]. Initially, a soft threshold (Beta score) was applied to fit the scale-free topology model assumption which generates a minimum number to make the index curve flatten out reaching a high value (> 0.85)[16]. Further, the adjacency matrix represented by gene expression was transformed into a topological overlap matrix (TOM) to classify genes using a dynamic tree-cut approach. Finally, genes that shared high-confidence correlation (abs(coefficient) > 0.8) were merged into the same module (parameter: deepsplit = 4, tree cut height = 0.2). The pairwise Pearson correlation between each gene from modules and the module-wise eigengene values was tested to characterize each gene's module membership (MM) under the same module. The weight score, a parameter to denote the robustness of connections of each edge, was filtered by taking genes associated with the top 10% high-confidence weight ranking connections for downstream functional analysis. Lastly, the metabolite profiles, oxidative stress-related enzyme activities, and LI-6800-derived traits were integrated into the network using Pearson correlation (PC)

between gene expression and these molecular phenotypes. Trait-correlated modules were selected based on their correlation level between MM scores and phenotypes (cutoff: r > 0.5, or r < -0.5, *p-value* < 0.05). The correlation between gene expression and phenotypes was estimated as gene significance (GS) to interpret the importance of genes related to each molecular phenotype.

Functional and regulatory network analysis

To study the TFs and their downstream target genes (protein-DNA interactions), TFs and protein kinases from each co-expression module were predicted using the iTAK tool based on conserved protein-coding domains[71]. To identify potential upstream regulators of genes in each module, we performed a motif enrichment analysis for the 2kb upstream promoter regions of genes in the same module using AME (--scoring avg --method fisher --hit-lo-fraction 0.25) and the Arabidopsis DAPseq consensus motif database[72,73]. The protein-protein interactions (PPI), PPIs among genes present in the co-expression network were acquired from the STRING database as the third layer of evidence to infer the functions or pathways of modules (parameter: confidence level: > 0.7, source: experiments and database source)[74]. As a result, co-expression patterns, PPIs, and TF-binding interactions were merged into an integrated network.

Functional enrichment analysis

Functional enrichment analysis of pathways (Kyoto Encyclopedia of Genes and Genomes (KEGG) database) and Gene Ontology (GO) analysis was performed to infer the biological function of genes associated with certain RCMs, transcripts assigned in co-expression modules, and differentially abundant transcripts. Here, we used a customized database of GO and KO information derived from sorghum functional annotation (https://phytozome-next.jgi.doe.gov) with the topGO and clusterprofile R packages [75,76]. ShinyGO was used for GO and KEGG enrichment for Arabidopsis [77](FDR < 0.05).

Data visualization and usage of databases

All statistical visualizations were generated in R (v4.3.2)[78] with multiple packages. Briefly, plots include the line, boxplot, PCA, and enrichment graphs generated by the ggplot2 package, and heatmaps using the pheatmap package. R Scripts used for plots were deposited in (https://rpubs.com/LeonYu/Main-Figures and https://rpubs.com/LeonYu/Supp-Figures). The integrated gene networks containing TF-binding, PPIs, and co-expression derived from WGCNA were generated using Cytoscape[79].

Results

Genomic diversity contributes to phenotypic variation among sorghum accessions

Phenotypic diversity under drought conditions is ultimately traced to variation in genomic sequence and rewiring of gene expression[80]. Therefore, to capture the genetic and phenotypic diversity present within sorghum, we initially generated an unrooted phylogeny from publicly available whole genome skim sequencing data for 356 out of 400 accessions with demographic information found within the Sorghum Association Panel (SAP)[61]. (**Figure 1A**). Cross-referencing this phylogeny with performance data from the SAP grown in the field (in Arizona in 2019) under standard conditions, we selected six accessions to comprise a small but diverse panel based on seed viability, genetic distinctiveness, and phenotypic variation (shoot and root biomass, plant height, and yield). Among the selected lines, four accessions (A, B, E, and F) are African accessions obtained from four countries across the continent (**Figure 1A**). The other two accessions (C and D) are U.S. breeding accessions. Variant calling and annotation of SNPs and INDELs were performed to more closely examine the genomic variation between these accessions. In total, 8,725,449 variants were identified among the six accessions relative to the published reference genome, with more than 50% of variants located in intergenic regions, followed by gene regulatory regions (21.00% in the 2Kb upstream region and 19.5% in 2Kb downstream region, 5.61% in the intron regions, and 0.23% affecting splicing in exons, **Figure S1, left panel**). Different proportions of unique SNPs and INDELs were uncovered among the six accessions, with each genotype contributing around 10-20% of genomic variation to the panel (**Figure S1 right panel**). In summary, these data suggest that we were successful in selecting six distinct accessions, representative of the genetic diversity within the species, and likely the wide range of environments in which sorghum can grow.

Based on preliminary field experiments conducted in summer conditions of the Arizona low desert in 2019, the six accessions displayed large phenotypic variation under standard growth conditions but all of them successfully reached reproductive maturity. To assess their performance under drought conditions, we grew the six accessions in replicated plots as part of a large the field trail conducted in the summer 2020 trial and performed a drought stress treatment. All plants were well-watered (WW) until the reproductive stage (appearance of the flag leaf), after which irrigation was reduced to a subset of plots to mimic drought treatment. Phenotypic measurements and molecular sampling (transcriptomics, metabolomics, and biochemical traits) occurred during the 1st, 3rd, 5th, and 7th week after the beginning of the drought treatment (**Figure S2A**)**,** with destructive measurements (e.g., biomass and root microbiome composition) occurring at the end of the season. Treatment effects were readily apparent at the end of the season (**Figures 1B-C**), with the M1 (“A”) accession showing the most pronounced leaf necrosis and alterations to root structure. The “A” (M1) and “D” (N290-B)

accessions had the lowest total biomass under both Control and Drought conditions ($p < 0.05$), potentially indicating poor adaptation to both heat and drought stress (**Figure 1D**). In contrast, the "B" (DL/59/1530) and "C" (80M) accessions displayed a minimal reduction in biomass or leaf weight in response to drought stress. For root phenotypes, "A" exhibited significantly lower length and width compared with other five accession under control conditions (**Figure 1E,** $p < 0.05$), but these differences became less pronounced under drought stress (**Figure 1E**). In addition, the endophytic root microbiome bacterial and fungal components associated with the "A" accession were starkly different under drought stress relative to all other accessions (**Figure S2B and S2C**). Leaf level physiological traits such as water use efficiency (WUE), maximum photosystem II efficiency in light conditions (Fv'/Fm'), and CO_2 net assimilation (A_{net}) exhibited similar results, with the "A" accession displaying hallmarks of water stress, particularly at the later time points (week 5 and week 7, **Figure 1F**), whereas other accessions, such as "B" and "C", showed a robust response. Thus, in addition to genotypic diversity, these data suggest that these six accessions displayed phenotypic diversity to drought and heat stress and may provide information on the molecular mechanisms by which sorghum responds to water limiting conditions.

Sorghum accessions displayed molecular response variation during drought stress

To characterize the molecular responses to drought stress in our panel, we next assessed the change in RNA abundance over the two months of the treatment (week 1, 3, 5, and 7) using RNA-seq. Initial quality control revealed high correlations between replicates (Pearson Correlation; r > 0.95, **Supplementary Table 1**), along with high mapping rates to primary transcripts among samples (**Supplementary Table 2**). Using hierarchical clustering and Pearson correlation analysis of global transcript abundance (N = 22,314, mean TPM > 1), the 96 samples (2 replicates * 4 time points * 2 conditions * 6 accessions) clustered into subgroups that fit their genotypic classification (**Figure 2A**). Separation due to treatment effect was observed for accession "B", whereas accession "A" week 3-week 7 timepoints for control and treatment were separated from week1, suggesting the combined stresses of heat and drought were having a distinct impact on this accession (**Figure 2A**). This global view of the transcriptome suggests that these accessions responded to stress in unique ways that were largely driven by genotype, and not by time or treatment.

To unpack the transcriptomic responses to treatment in the six accessions, we examined the transcripts that were most highly variable among all 96 datasets as well as those that were differentially abundant between treatments (**Supplementary Table 3)**. The top 10% of most variable transcripts were selected, based on the standard deviation of the global transcript abundance for principal component analysis (PCA). This approach uncovered strong genotypic and treatment effects

441 represented by both PC1 and PC2 (PC1 = 19.35%, PC2 = 14.88%, **Figure 2B**). In agreement with the
442 hierarchical clustering, samples from the six accessions clustered separately, with accession "C"
443 exhibiting the least separation due to stress or time, and accession "A" exhibiting the most **(Figure 2B**).
444 Pairwise comparisons of differentially abundant transcripts (DATs; *FDR* < 0.05) between control and
445 treatment for each timepoint and genotype uncovered different transcriptional responses to treatment.
446 Accession "C" showed the lowest response to treatment, whereas accessions "E" and "A" had the most
447 DATs across the four weeks (**Figure 2C**). In addition to vastly different numbers of DATs among
448 accessions, these DATs were often specific to a genotype at any given week during the season (**Figure**
449 **S3**). Gene ontology (GO) term enrichment of these DATs (up and down-regulated, *FDR* < 0.05)
450 reinforced prior observations that each of these genotypes was responding to the treatment by
451 modulating unique pools of transcripts (**Figure 2D**). Interestingly, the "A" and "F" accessions did not
452 appear to be inducing a molecular stress response, as "stress response" GO terms were absent for
453 both of these accessions (**Figure 2D**). Thus, one explanation for the decreased biomass and survival
454 rates for accessions A and F could in part be explained by a lack of initiating an adequate stress
455 response.

456 To gather additional clues as to how these accessions responded to drought treatment at the
457 molecular level, we next examined the stress-induced changes in metabolites and components of the
458 oxidative stress status of the sorghum leaves[81,82] . Metabolites and biochemical traits related to the
459 oxidative stress status were measured on leaf samples simultaneously collected at the same time
460 points as for the RNA-seq analyses, with GC-MS-based quantification and annotation of primary
461 metabolites and spectrophotometric-based quantification of oxidative stress traits. A PCA of global
462 changes in leaf primary metabolome (*N* = 52) across our panel revealed that, as for the transcriptome,
463 the genotype explained most of the sample variation (**Figure 2E**, left panel; PC1: 33.25 %).
464 Interestingly, samples associated with accession "A", weeks 3, 5, and 7 (both conditions), separated
465 from all other samples, and this separation was largely driven by the levels of the three main soluble
466 sugars (sucrose, glucose and fructose) and chlorogenic acid (**Figure 2E**, right panel**; Figures S2D** and
467 **S2E**).

468 In contrast to transcriptomic and metabolomic profiles, the levels of biochemical traits (*N* = 24)
469 related to the leaf oxidative stress status did not exhibit strong genotypic effects (**Figure S2F**). Thus,
470 we analyzed these data by using pairwise correlation among each entry of the four classes of
471 chemicals, including antioxidants, oxidative markers, antioxidant enzymes/proteins, and others (**Figure**
472 **2F**). As expected, we observed strong correlation among similar classes of antioxidant markers (e.g.
473 polyphenols and flavonoids). We then used one representative marker from each of the four main
474 clusters to examine accession-specific changes over time and between treatments (**Figure 2G**).

Interestingly, although each of these chemicals displayed a similar profile between accessions over time in response to treatment, the degree of response varies. The stronger performers (based on biomass), “C” and “F”, for the most part displayed reduced levels of oxidative stress markers relative to the other accessions under drought conditions at week 7, whereas levels were elevated for the weaker performers, “A” and “D”. Overall, we identified strong genotypic effects on both visible and molecular phenotypes. As expected for this species, most accessions displayed resilience to combined heat and drought stress, whereas others (e.g. accession “A”) performed poorly in this environment. Thus, this panel presents a suitable comparative framework to identify the components responsible for drought and heat tolerance in sorghum under field conditions.

Certain RCMs display a genotype- and treatment-associated enrichment on photosynthesis and stress-related transcripts

As changes in RNA covalent modification (RCM) levels have been associated with plant stress responses[21,26,83,84], we next examined changes at the epitranscriptomic level in our panel. A subset of RCMs interferes with RNA-seq library preparation (specifically the reverse transcription step), leaving patterns of mis-incorporated nucleotides that can be assigned to specific modification classes bioinformatically[30,31,33]. We processed our RNA-seq data through two independent RCM predictors (HAMR and Modtect) and took the intersection of called sites, resulting in 20,096 high-confidence RCMs associated with 3,715 transcripts (**Supplementary Table 4 and 5**). These RCMs are mainly composed of six modification classes based on the resolution of the RCM caller, including 3-methylcytidine (m^3C, *N* = 1688 transcripts), N1-methyladenosine / 1-methylinosine / 2-methylthio-N6-isopentenyladenosine (m^1A / m^1I / ms^2i^6A, *N* = 2324 transcripts), N1-methylguanosine (m^1G. *N* = 1250 transcripts), pseudouridine (Y, *N* = 1435 transcripts), D (pihydrouridine, *N* = 1101 transcripts), and N^6-isopentenyl adenosine or N^6-threonylcarbamoyladenosine (i^6A / t^6A, *N* = 1236 transcripts).

In order to associate epitranscriptomic changes with other molecular and physiological traits, we developed three metrics based on our previous work to explain variation in the modification status for each transcript across our data[85]: **1**) Presence or absence of a modification (PAM) within a transcript, scored as 1 (present/modified) or 0 (absent/not modified), which simply denotes if a transcript was modified or not in a particular context, ; **2**) RCM density, which describes the actual number of modification sites within a transcript within a dataset, ; and **3**) RCM frequency, which describes the ratio of total modified reads for a transcript relative to the total mapped reads for that transcript (**Figure S4A and S4B, Supplementary Tables 6 and 7;** see Methods). As with the other molecular markers, genotype, not treatment, explained most of the variation in RCM PAM and density in our samples regardless of RCM class (PC1: 16.93%; **Figure S5A**. When the density of each RCM class was

examined individually, separation based on treatment was more apparent only for the “A” accession (**Figures S5B**). By examining all RCMs per transcript, regardless of class (**Supplementary Tables 6** and **7**), a stronger genotypic separation was observed for PAM, density, and frequency, again predominantly driven by the “A” accession (**Figure S6A**). An examination of average density and frequency changes over time revealed that accessions “A” displayed the most variation, and predominantly under drought conditions (**Figure S6B**). To further characterize the RCM variation driven by genotype and treatment, we performed pairwise comparisons of RCM features under control and drought using the numbers of modified transcripts, mean modification density, and mean modification frequency for each accession (**Figures S7A-D**). While accession “A” had a higher number of modified transcripts when combining treatments and timepoints (**Figure S7A**), accessions “A” displayed slight reduction in number of modified transcripts (**Figure S7B**), frequency (**Figure S7C**), and density (**Figure S7D**) under drought conditions relative to control.

To better understand the biological implications of a transcript being modified, we selected a core set of 342 transcripts that had a non-zero PAM across all time points, treatments, and genotypes. A hierarchical clustering split these 342 transcripts into three groups based on RNA abundance changes (Z score-normalized, **Figure 3A**, upper panel, **Supplementary Table 8**). Gene Ontology (GO) enrichment analysis revealed an overrepresentation of stress-response-related terms in the Group I genes, translation and metabolic processes for Group II, and photosynthetic processes for Group III (**Figure 3B,** top 20 *FDR* GO terms shown). Within these groups, we highlighted time points and accessions where these modified transcripts were most or least abundant (C1-C6, **Figure 3A**). Interestingly, samples associated with accession “A” showed the highest abundance of stress response transcripts (C2; **Figure 3A**) and the least abundance of transcripts associated with photosynthesis (C6; **Figure 3A**). An examination of the relationship between RNA abundance and RCM frequency for the transcripts found within each of these groups revealed an inverse relationship for most clusters (**Figure 3C**). In contrast, RCM density is largely correlated with RNA abundance (**Figure 3D**). Directly comparing the scaled transcript abundance vs RCM frequency and density values revealed that transcripts with high RCM density and low RCM frequency are associated with high transcript abundance (**Figure 3E**). This result is consistent with the global transcript-wide correlation between transcript abundance and RCM status (**Figure S4C**). We also observed this pattern in transcripts that are only modified in a single genotype (genotype-specific RCMs; **Figure S7E**). Despite these genotype-specific modified transcripts being transcribed across all genotypes, their RNA abundance was lower in genotypes in which they were not modified compared with genotypes in which they were modified (**Figure S7F**), indicating RCM density is positively associated with the transcript abundance. In

summary, these data suggest that RCMs are enriched in a transcript- and genotype-specific manner that may influence how each accession responds to its environment.

An integrated co-expression network analysis reveals two modules with strong RCM and trait associations

With each of these molecular and physiological traits examined, we next took a comparative systems-level approach to connect these traits to their molecular pathways and putative regulators. To do so, we constructed a co-expression network starting with the top 10% of most variable transcripts, as measured by median absolute deviation (MAD) score. We then recursively explored parameter space, increasing the number of included transcripts until we achieved an optimized network with the fewest number of unassigned transcripts and the highest scale-free topology model (**Supplemental Table 9**). The resulting network contained 22,314 transcripts classified into 44 modules with a low portion of unclassified transcripts (*N* = 2,646) and high scale-free model correlation (R^2 = 0.91, slope = -1.93, **Supplemental Table 10**). Using this framework, we incorporated phenotype data into modules using trait-module correlations. Briefly, the first principal component value (eigenvalue) per module was calculated using their transcript abundance PCA. Leaf level physiological measurements (LI-6800), biochemical traits related to the oxidative stress status, and metabolite abundances were then correlated with the eigenvalues of each module (**Supplemental Tables 11, 12,** and **13**), leading to 13 modules with significant trait-module correlations ($P < 0.05$, **Figure 4A**, left panel). Several modules (e.g. “Lightyellow”, “Salmon”, “Dark-turquoise”, and “Darkmagenta”) were strongly correlated with multiple traits. To further prioritize modules, we examined pairwise correlation values between each gene within a module and the associated traits (referred to as gene significance, GS), to uncover the modules harboring genes with the strongest correlations to traits of interest. gene-trait correlations. This approach identified a number of transcripts with high GS (denoted in **Figure 4A**; *N* = 1,167). In particular, we observed a suite of transcripts with strong positive (*N* = 213; Darkmagenta) and negative (*N* = 207; Salmon) with leaf-level physiological measurements. These genes, and their respective modules, are therefore of interest for identifying potential drivers of responses to heat and drought stress in these accessions.

To gain insights into the biological processes associated with each of these modules, GO and KEGG pathway enrichment analyses were performed (modules with *FDR* < 0.05 are displayed, **Figure 4A**, middle panel). This approach revealed modules with biological processes of interest, such as stress response, translation, metabolic pathways, and RNA modification terms. In addition, we performed a motif enrichment analysis using the Arabidopsis DAP-seq motif database to identify TF binding motifs over-represented in the promoters of genes in the 13 modules (*FDR* < 0.05)[72]. We then

determined which of the enriched motifs matched to TFs found within the corresponding module, suggesting those TFs may target co-expressed genes (**Figure 4A**, right panel). For example, in the "Salmon" module, we identified 22 TFs with enriched motifs (**Supplemental Table 14**). Of these TFs, 12 are annotated as regulators of abiotic stress (e.g. CRF, ERF, and MYB family TFs; **Supplemental Table 14**), agreeing with our assertion that this represents a module of stress-associated genes.

Lastly, we overlaid RCM features onto each module by calculating the portion of modified transcripts with high GS value, and proportions of modified transcripts/unmodified transcripts associated with the enriched GO/KEGG terms in a module (**Figure 4B**). These data uncovered several modules with strong associations between each of our molecular and physiological datasets that are associated with stress responses or photosynthesis. Of these, the "Darkmagenta" module stands out for its strong positive association with leaf-level physiological parameters and enrichment of stress-response GO terms, whereas the "Salmon" module shows a negative correlation with leaf-level physiology and RCM status but a positive correlation with photosynthetic terms.

Based on the information above, we further characterized co-expression modules to determine if any exhibit distinct patterns that can be attributed to genotype, condition, or time point. To do so, we plotted the eigenvalue of each accession under each treatment per module to select modules with strong genotype or treatment effects over time. Using this approach, we identified two modules with strong genotypic effects (**Figure 4C,** "Darkmagenta" and "Salmon" modules) where accession "A" performed quite differently at weeks 3-7 relative to the other accessions. Interestingly, of the 342 constantly modified transcripts (**Figure 3A**), a large number of the photosynthesis-associated transcripts (Group III) appear in the "Salmon" and "Darkmagenta" modules (**Figure 3B** and **Figure 4D**), further reinforcing a connection between RCM status and leaf-level responses to the environment. Overall, this network-based approach overlaid multiple lines of evidence to uncover a subset of transcripts whose abundance is strongly associated with physiological and metabolic traits. In addition, this approach uncovered putative transcriptional and post-transcriptional regulators, TFs and RCMs, respectively, that might lead to the accession-specific stress resilience we observe in this panel.

Transcripts associated with drought resilience are regulated by SbCDF3 and RCMs

To further explore the sets of transcripts that have the strongest trait-transcription associations, we initially focused on identifying TFs within the "Darkmagenta" and "Salmon" modules that likely exert the largest effect on transcription within their respective modules (hub TFs). Using the Analysis of Motif Enrichment (AME)[73], we found the binding motifs of 10 TFs that were significantly enriched in the 2 kb upstream promoter regions of genes in the "Darkmagenta" module (**Supplemental Table 14**). Among

these putative hub TFs, one in particular, CDF3 (Arabidopsis DAP-seq database ID: C2C2dof_tnt.CDF3, TAIR ID: AT3G47500, sorghum ID: Sobic.002G421900), stood out based on AME ranking, numbers of edges (connection between two transcripts), and number of connected target genes within the module. Importantly, CDF3 has previously been reported as a transcriptional activator that enhances drought and oxidative stress tolerance in Arabidopsis and tomatoes[86–88]. CDF3 is encoded by a single copy gene in sorghum, and importantly, appears to be connected to similar classes of stress response genes within the "Darkmagenta" module. Out of the 1,093 genes within the "Darkmagenta" module, 681 (62% in total) were predicted to be bound by CDF3 in their promoters (**Figure 5A**). Careful examination of the binding motifs of those 681 putative CDF3-bound and co-expressed genes revealed that the consensus motif contains the core binding motif (5'-T/AAAAG-3'; **Figure 5A**) reported in previous molecular studies[89], suggesting conservation in both CDF3 as a regulator of environmental responses and in its preferred binding motif. Interestingly, an examination of CDF3's transcript abundance pattern across the 96 sorghum samples revealed a dramatic decrease in the "A" accession in weeks 3-7 under both control and drought conditions (**Figure 5B**). Additionally, we examined the correlation between CDF3 and each CDF3 target (N = 412**,** selected by random extraction) in the "Darkmagenta" module, a similar number of non-CDF3 targeted transcripts within the module, as well as randomly selected transcripts from outside of the "Darkmagenta" module (N = 412). A significantly higher correlation was observed between CDF3 and within-module targets compared with the other two groups (p = 0.0018 and p < 2.2e-16, **Figure 5C**). Lastly, we compared the $\log_2$FC score (Control/Drought) of CDF3 targets and non-CDF3 targets across all accessions and time points. This comparison revealed a significantly higher $\log_2$FC (p = 0.001) of CDF3-targets (**Figure 5D)**, suggesting these target genes are induced under drought conditions in a CDF3-dependent manner.

The above CDF3-target data, along with the traits correlated with the "Darkmagenta" module, suggest that CDF3 is acting as a potential central regulator connecting the drought response in these accessions to downstream photosynthetic traits. To further investigate this hypothesis, we more closely examined the relationship between CDF3 target genes and the physiological and molecular traits (e.g. RCMs and metabolites) associated with the "Darkmagenta" module. To do so, we selected the transcripts (***N* = 82**) with the highest trait-RNA abundance correlation (cut-off: |r| > 0.70) and then integrated GO terms, TF-binding, and protein-protein interaction data (**Figure 5E; Supplementary Table 15**). We further filtered these CDF-interacting transcripts based on RCM status (boxes = modified, circles = unmodified; **Figure 5E, Supplementary Table 15**: column: RCMs), resulting in a sub-network of 230 out of the 1,093 transcripts in the original "Darkmagenta" module. Of the 93 modified transcripts found within this sub-network, 67 transcripts displayed a strong correlation with CDF3 and were predicted to be CDF3 targets based on AME (blue dashed lines, **Figure 5E**). Of these,

ten genes were strongly correlated with CO_2 net assimilation (Blue boxes, **Figure 5E**), eleven with flavonoid content (grey boxes, **Figure 5E**), four with leaf temperature (orange boxes, **Figure 5E**), and six with total antioxidant capacity (peach boxes, **Figure 5E**). Based on GO enrichment, ~ 200 genes in the module were associated with GO terms for abiotic stress response, RNA modification (predominantly spliceosomal RNAs), and metabolic processes. In addition, several genes with strong trait-transcript abundance correlation in this module coincide with genes previously identified through QTL mapping (marked with red stars, **Figure 5E, Supplementary Table 15** in the "QTL" column) and therefore could be ideal candidates for downstream functional analyses.

Interestingly, ten of these CDF3 targets were not only strongly trait-correlated, but were also part of the previously identified constantly modified transcripts (**Figures 3A** and **5F**). In addition to being strongly correlated with CDF3 transcript abundance ($R > 0.6$, $P < 0.05$; **Figure 5F**) and RCM status (**Figure 5G,** $P < 0.05$ with Pearson correlation coefficient (PCC) marked), these ten transcripts were also positively correlated with CO_2 net assimilation, and negatively associated with flavonoid content and total antioxidant activity (**Figure 5F**). Indeed, the poorly performing accession "A" had reduced transcript abundance for each of these genes (and CDF3) and showed reduced CO_2 net assimilation in the later time points (weeks 3-7; **Figure 5F**). Some of these trait and CDF3-associated genes have been functionally characterized in previous studies, including the orthologs of *SbPPC2, SbCP5,* and *SbPSAL* in Arabidopsis[90–92]. In agreement with the data we describe here, these genes are integral to photosynthesis-related processes and confer stress tolerance. Thus, we uncovered a suite of photosynthesis and stress-associated genes that are regulated at the transcriptional level by CDF3 and at the post-transcriptional level by a suite of RCMs in sorghum.

Network analysis uncovers a link between the RCM dihydrouridine and photosynthesis under drought stress

Using a similar approach, we examined transcripts within the "Salmon" module that were also associated with photosynthesis and metabolism. We first prioritized hub TFs based on AME motif enrichment, identifying two potential upstream TF regulators, ARID-HMG DNA-binding PROTEIN15 (HMGDB15) and BASIC PENTACYSTEINE6 (BPC6, **Supplementary Table 14**). Both of these TFs were reported to be involved in abiotic stress response in Arabidopsis (AtHMGDB15: AT5G21430; AtBPC6: AT5G42520)[93–95]. While these TFs may be of interest, this module contained a more interesting target gene. When examining genes with high module membership, we identified a transcript that encodes a dihydrouridine synthase (hereafter referred to as SbDUS2: *Sobic.006G158100.1*), the enzyme responsible for synthesizing the RCM dihydrouridine (DHU). There

are multiple DUS enzymes in eukaryotes; a phylogenetic analysis including all potential *Dus* genes from multiple plant species and budding yeast suggests that this particular enzyme falls into the *Dus2* clade (**Figure S8**). SbDUS2 shares strong sequence similarity at the amino acid and functional domain level with other plant DUS2s, including in Arabidopsis (CDD; ID: AT3G49640**,** CDD number**:** cd02801-FMN-binding domain) and non-plant DUS2s (e.g., ScDUS2). DUS2 is an ancient enzyme shared between eukaryotic, prokaryotic, and some archaeal lineages, and DHU has been predominantly reported on functional non-coding RNAs (e.g. tRNAs and rRNAs) and more recently on mRNAs in vertebrates and yeast[113,115,116]. Intrigued by its inclusion in the "Salmon" module, we examined DUS2 transcript abundance over time and across treatments in each of the accessions (**Figure 6A**). Accessions "A" and "F" generally showed decreased *SbDUS2* transcript abundance over time compared with the other accessions, regardless of treatment (**Figure 6A**). In addition, the PAM of DHU among 1101 DHU-modified transcripts showed a strong genotype effect, with most "A" and "F" samples separating from all other accessions and time points in a PCA (**Figure 6B**). Interestingly, the "A" and "F" accessions showed an elevated number (*N* = 119) of homozygous non-reference alleles within the *SbDUS2* coding sequencing relative to the other four accessions (**Figure 6C**), in agreement with their phylogenetic relationship (**Figure 1A**). Of these, eight SNPs were predicted to impact splicing and potentially affect transcript abundance (**Figure 6C**; VEP annotation results, see methods). The true impact of these SNPs is unclear, but altogether, DUS2 expression and the pattern of DHU deposition suggest that dihydrouridine is used differentially across these six accessions.

To further explore the relationship between DHU and the traits associated with the "Salmon" module, we developed a sub-network that integrated phenotypic and transcriptional associations. We built up the sub-network (*N* = 77) of transcripts with high GS and module membership and then overlaid RCM status, TF-binding, trait association, and GO term data (**Figure 6D**). Within this module, there was a set of nine transcripts that consistently showed a non-zero dihydrouridine status (blue boxes, **Figure 6D**). The abundance of these constantly DHU-modified transcripts was positively correlated with *DUS2* transcription (Average PCC > 0.38, **Figure 6E** values to the left of the heatmap) as well as correlated with CO_2 net assimilation rate (PCC > 0.65, **Figure 6E** values to the right of gene names). We also observed that two of the nine DUS2-co-expressed genes are found within QTL intervals that were significantly associated with leaf carbon assimilation rate derived from a previous GWAS study of sorghum (*p* = 2.9e-4 and *p* = 4.5e-5; **Figure 6E,** dashed box**,** Pyruvate phosphate dikinase (PPDK): Sobic.009G132900 and Chloro-respiratory reduction L (CRRL)*:* Sobic.008G102000)[96,97]. Furthermore, we found that one of them, *SbPPDK*, is a putative target of the TF HMGD15. The Arabidopsis ortholog of each of these genes has been characterized as stress-associated (ID: *AtPPDK:* AT4G15530 and *AtCRRL:* AT5G21430). AtCRRL is essential for NAD(P)H dehydrogenase activity (NDH) that serves as

the machinery of stress resistance to promote stress tolerance[98,99].AtPPDK is known to enhance photosynthesis while improving stress tolerance and is also important for seed development in Arabidopsis[100,101]. In maize, the C4 grass ortholog of sorghum PPDK1 (*Zm00001eb287770.1p: ZmPPDK-C4*) is also responsible for CO_2 net assimilation, with loss-of-function mutants displaying reduced expression of other C4-related genes[102]. Thus, these genes are of particular interest for understanding plant performance under drought stress.

To assess the robustness of the observed connection between *SbPPDK, SbCRRL*, and DHU modifications, we reprocessed two publicly available RNA-seq datasets for sorghum encompassing a field drought-stress experiment and a transcriptome tissue atlas[55,60]. The large drought time-series dataset generated by Varoquaux and colleagues[38] examined the transcriptional profiles of two sorghum lines under both pre- and post-flowering drought treatment. The two accessions in question, neither of which were included in our experimental panel, displayed either pre-flowering drought tolerance (RTx430) or post-flowering drought tolerance (BTX642). We focused on the analysis of post-flowering imposed drought treatment data, as this matched our field experiment. Analysis of these RNA-seq datasets revealed a decrease in *SbPPDK1* and *SbCRRL* levels under drought in the susceptible background (RTX430; **Figure 6F**) but remained unperturbed in the tolerant background (BTX642, **Figure 6F**), suggesting that *SbPPDK* and *SbCRRL* may positively impact the sorghum performance under water-limited conditions, as observed in our field trial (**Figure 6E**). In addition, *DUS2* transcript abundance displayed a strong positive Pearson correlation with both *SbPPDK1* and *SbCRRL* (**Figure S8A, left:** r = 0.88, p < 2.2e-16; r = 0.7, p = 2.2e-16, N = 395), suggesting a potential post-transcriptional regulatory role for DUS2 on *SbPPDK* and *SbCRRL*. In the tissue atlas dataset, we again observed a strong correlation between *SbPPDK-SbDUS2* and *SbCRRL-SbDUS2* (**Figure S8A, middle**: r = 0.63, p = 3.1e-16; r = 0.47, p = 7.9e-19; N = 137). Notably, *SbPPDK* exhibited higher transcript abundance in upper leaves, tissue that we sampled in the field (**Figure 6G**). Furthermore, we analyzed the DHU modification status in *SbPPDK and SbCRRL* transcripts relative to their abundance in the tissue atlas data (**Figure 6H**). The results revealed a strong correlation between abundance and both modification density (*SbPPDK:* r = 0.93, p = 8.6e-09 and *SbCRRL:* r = 0.85, p = 2e-04) and frequency (*SbPPDK:* r = -0.58, p = 0.0094 and *SbCRRL:* r = -0.7, p = 0.008). These findings revealed that core photosynthesis-related mRNAs are targeted for modification by DHU across a number of environmental and developmental contexts, and the presence of this modification is associated with increased transcript abundance and CO_2 net assimilation levels under drought stress.

DUS2-mediated DHU addition affects transcript abundance and net CO_2 assimilation in Arabidopsis

The data presented thus far suggests a role for DHU in promoting drought tolerance in sorghum. To determine if DUS2 has a direct and conserved role in influencing plant performance under stress, we focused on DUS2 loss-of-function mutants in the genetic model system, *Arabidopsis thaliana*. The uncharacterized Arabidopsis DUS2 ortholog (AT3G49640) is 79.81% identical to the SbDUS2 protein and also contains the dihydrouridine synthase domain (**Supplemental Figure S8** and **Figure 7A**). We obtained a T-DNA insertion line that disrupted the third exon within the coding region of *AtDus2* (**Figure 7A**; CS857388, https://abrc.osu.edu/). Consistent with expectations, LC-MS analysis of DHU levels in mature rosette leaves revealed a significant reduction in the *atdus2* background relative to the wild type control (**Figure 7B**).

It was previously observed that *AtDUS2* transcript abundance peaks in germinating seeds, leaves, and flowers (https://www.arabidopsis.org/locus?key=126568)[103]. In agreement with this expression pattern, we observed reduced germination (**Figure 7C**; $p < 0.05$, $N > 50$) and growth rates (**Figure 7D**), as well as shorter stature and increased time to flower (**Figure 7E**), in *atdus2* mutant plants relative to wild type controls (Col-0). We also examined the ability of the *atdus2* mutant to withstand heat stress by growing ~2-week seedlings at 37°C (heat) and 22°C (control) for 12 days. While both Col-0 Arabidopsis and *atdus2* mutants exhibited reduced growth rates at 37°C, the survival rate at the end of the 12 days of treatment was significantly reduced in mutant as compared to Col-0 plants (**Figure 7D**, right and inset, $P < 0.05$). As *SbDUS2* was positively associated with CO_2 net assimilation under heat conditions in our field data, we examined the leaf-level responses to heat stress in *atdus2* mutant plants. We observed slightly reduced ($p = 0.12$ and $p = 0.33$, $N = 10$; **Figure 7F**) CO_2 net assimilation and stomatal conductance in *dus2* mutant lines under control conditions (22°C). However, photosynthetic parameters were significantly reduced under heat stress (24h at 37°C) in the *atdus2* background relative to Col-0 controls ($p = 0.00016$ and $p = 0.0024$, $N = 10$, **Figure 7F**), suggesting that DUS2 is critical for a proper heat stress response in Arabidopsis. Interestingly, although Arabidopsis does not utilize *PPDK1* to perform C4 photosynthesis, we still observed a positive correlation between *AtDUS2* and *AtPPDK1* across the DUS2 and Col-0 comparative heat stress data and the public abiotic stress data (heat, wounding, and cold treatment, **Figure S9A** right panel, r = 0.6, p = 4e-8, r = 0.48, $p = 0.05$)[103], highlighting a potential deeply conserved post-transcriptional regulatory connection between these two genes across two plant species.

We then asked how extensive the changes to the transcriptome would be in DUS2 loss-of-function mutants as compared to wild type Col-0 in Arabidopsis. First, we examined the transcription

levels of *AtDUS2* from the two backgrounds. As expected, the *atdus2* mutant lines displayed significantly lower *DUS2* transcript levels due to the T-DNA insertion occurring in the 5' portion of the coding region (**Figure S9B** and **S9C**). Furthermore, we found lower DHU modification frequency in the transcriptome of *atdus2* mutants under both control and heat stress conditions (**Figure 7F**, p = 0.072 and p = 0.0042) and a slightly reduced overall number of modified sites (**Figure S9B)**. These data align with the overall reduction in 5′,6-dihydrouridine concentration in total RNA observed by LC-MS (p = 0.0099, **Figure 7B**). Finally, we examined the transcriptomic responses to heat stress in control (Col-0) and *atdus2* backgrounds. We uncovered strong treatment and genotype effects based on a PCA of the top 25% most abundant transcripts (**Figure 7H**), suggesting that, in agreement with our morphological data, loss of DHU has a pronounced impact on plant molecular responses. A GO term enrichment comparison of genes upregulated by heat in WT vs *atdus2* mutants on the last day of stress (Day 4) revealed a stronger response in the WT plants, where stress response terms are either significantly more enriched in WT as compared to the mutants (**Figure 7I**, left inset; $-\log_{10}(FDR) > 10$), or are completely lacking in the mutant plants (**Figure 7I**, left). Interestingly, among the transcripts suppressed by stress in the *atdus2* mutant, we noticed an enrichment in multiple stress-associated terms, whereas the suppressed transcripts in the WT background were predominantly associated with cell division and metabolism (**Figure 7I**, right). These findings suggest that the loss of DUS2 alters the transcriptome globally in response to heat. Whether this alteration is due to loss of global DHU levels in the transcriptome, or because of reduced DHU levels on a few key specific mRNAs (e.g., PPDK) is still unclear. However, these genetic experiments in Arabidopsis, in combination with our sorghum field work, point to a novel and conserved role for dihydrouridine in modulating plant performance under abiotic stress conditions.

Discussion

Combined heat and drought stress typically induce complex phenotypic and molecular responses that often lead to reduced biomass, seed set, and photosynthesis, thereby adversely affecting the overall productivity of agricultural crops[104]. Sorghum is generally well adapted to hot and dry environments, although there is variation in how different sorghum accessions perform, variation that can be leveraged to shed light on the exact regulatory mechanisms by which sorghum achieves maximal growth under these conditions. Given the complex nature of plant responses to water deprivation, a variety of stress-responsive genes have been identified in sorghum and other plant species[60,105–108]. However, induction of a transcriptional regulator does not imply that its downstream response is beneficial to plant growth and yield. Thus, we still lack an understanding of which of these regulatory mechanisms and their targets contribute to improved agricultural traits in a field setting and therefore could serve as breeding markers or targets for manipulation.

We used a representative subset of the genetic and phenotypic diversity found within sorghum to identify transcriptional and post-transcriptional regulatory mechanisms associated with improved responses to hot and dry environments. The benefit of the approach applied here is that systemically induced stress-response genes are ignored in favor of those that track with phenotypic and genetic variation. Our panel of six distantly related sorghum accessions revealed distinct physiological and molecular responses to hot and arid conditions, with some accessions exhibiting minimal reductions in biomass and photosynthetic efficiency, whereas others (e.g. “A”) performed quite poorly. Leaf-level physiological traits tracked closely with oxidative stress status, metabolomic, transcriptomic, and epitranscriptomic features across the six accessions, allowing us to make inferences as to which molecular mechanisms were contributing to phenotypic success. Using a comparative systems-level approach, we identified cohorts of transcripts whose expression patterns were most closely linked to molecular and physiological responses to drought stress (**Figure 4**). Two co-expression modules in particular exhibited variation due to both treatment and genotype, allowing us to identify regulators that were associated with stress tolerance. At the transcriptional level, we identified three stress-induced TFs that may serve as key positive regulators of photosynthetic traits and plant performance (*SbCDF3;* **Figure 5***, SbHMGDB15*, and *SbBPC6;* **Figure 6**). As expected based on their expression in leaf tissue, these TFs, and their AME-predicted targets, such as *SbPPDK, SbCRRL, SbPPC2, SbCP5,* and *SbPSAL,* were highly correlated with photosynthetic traits. Several of these genes, such as *Cdf3* and *Ppdk1*, have been characterized in other plant species[86,93,95], and have also been shown to be part of a large suite of genes induced under drought stress across development in sorghum[60]. In addition, association studies in sorghum under abiotic stresses uncovered a strong association between CO_2 assimilation and QTLs harboring these genes[96]. As these genes in particular show strong trait-linked

transcriptional variation between sorghum accessions and are conserved across Arabidopsis and sorghum, they are ideal candidates for manipulation or selective breeding in sorghum or other grasses.

While RCMs have been shown to influence or be associated with plant stress responses [24–26,83,109–111],, we were still surprised to see that an enzyme associated with the 5'-6'-dihydrouridine, DUS2, was clustered into one of the two photosynthesis and stress-response modules. The dihydrouridine RCM is most commonly associated with tRNAs, and indeed the DUS protein family is conserved across the tree of life, including in eukaryotes, prokaryotes, and some archaea[112–114]. Due to advances in detection capabilities, DHU has recently been described on mRNAs in humans and yeast, where it has been reported to influence splicing, translation, and fundamental biological processes such as meiosis[113,115,116]. However, no reports exist on DHU in plants, and thus the findings presented here offer insights into a deeply conserved post-transcriptional regulatory mechanism.

Interestingly, not only was *DUS2* strongly co-expressed with a group of transcripts related to net CO_2 assimilation, the presence of DHU itself was also linked to the abundance of these transcripts. One gene in particular, *Ppdk1*, is integral to both photosynthesis and gluconeogenesis in plants. Both *PPDK1*'s transcript abundance and DHU levels on this transcript were positively associated with photosynthetic parameters under drought stress and thus we hypothesize that this RCM may be positively influencing translation or RNA stability. We observed similar connections between DHU levels and transcript abundance for a number of other genes during stress conditions. Overall, our transcriptomic and epitranscriptomic approach uncovered both the transcriptional and post-transcriptional regulators of key enzymes associated with central metabolism and photosynthesis.

Given that this was the first report of DHU in plants, and in particular how DHU might contribute to plant stress responses, we sought to develop a clearer picture of what this RCM might be doing by examining Arabidopsis lines with insertions in the orthologous *Dus2* locus. These *atdus2* mutant lines exhibited both decreased germination, slower growth, and reduced biomass at the time of flowering (**Figure 7**). Given the overall reduced DHU levels in these plants, each of these traits could be linked to overall reduced translation rates due to inadequate DHU on tRNAs [112,115]. However, we also observed a heat-stress-dependent reduction in photosynthetic traits and reduced survival after treatment in Arabidopsis *dus2* mutants, suggesting a conserved role for DHU in mounting an effective response to this stress in plants (**Figure 7**). We also observed reduced DHU levels on mRNA transcripts in this background and a concomitant decrease in transcript abundance for stress-response genes in the mutant background during heat stress. In line with a model whereby DHU is impacting transcripts more directly, a time-series transcriptomic analysis of Col-0 and *dus2* mutants during heat stress revealed a similar enrichment of stress and photosynthesis differentially abundant transcripts as observed in the "A" sorghum accession (**Figure 7H**). Interestingly, despite the fact that Arabidopsis performs C3-

photosynthesis, we still observe a positive correlation between *DUS2* and *PPDK1* transcripts in Arabidopsis, both in our data and in the Arabidopsis abiotic stress atlas[103], suggesting these two transcripts are co-regulated across the two species. Indeed, DUS2 and the RCM DHU are positively associated with many nuclear-encoded genes linked to photosynthesis and metabolism in Arabidopsis and sorghum. While the mechanism by which DUS2 targets specific transcripts is unclear, the deep conservation of many of these genes, and *Dus2* itself, suggests that this modification is an important addition to the repertoire of critical mRNA RCMs.

In summary, our comparative systems-level approach has allowed us to better understand the transcriptional and post-transcriptional regulatory mechanisms underlying heat and drought tolerance in sorghum. Using a small panel that represents high genetic diversity within sorghum, alongside intensive phenotypic and molecular phenotypic analyses, we were able to identify regulatory factors, such as CDF3, associated with improved biomass and photosynthetic parameters. Among the many differentially abundant transcripts, we were able to identify a subset of strong candidates to be used as breeding targets, similar to our past efforts in other agronomic species[16]. Lastly, our integrative approach uncovered an unexpected role for the RCM DHU in plant stress responses. Indeed, this deeply conserved RCM influences abiotic stress responses in both Arabidopsis and sorghum by acting on many of the same photosynthetic and metabolic mRNA transcripts. Thus, these data highlight the complex transcriptional and post-transcriptional regulatory landscape influencing plant stress responses but offer promising insights as to how they might be used to select more stress-resilient plants in the future.

Data availability

Transcriptome sequencing reads have been deposited in the National Center for Biotechnology Information (NCBI) BioProject database under the accession number: PRJNA1119650**.** Chromatograms of 5,6-dihydrouridine Mass spectrometry data was deposited under massIVE online repository (ftp://massive.ucsd.edu/v09/MSV000097205/). Computational pipelines used for data analysis were deposited under the GitHub page (https://github.com/Leon-Yu0320/Sorghum_Range28). The R markdown files (R code) for data visualization in main figures and supplementary figures were deposited in Rpubs workspace (https://rpubs.com/LeonYu/Main-Figures and https://rpubs.com/LeonYu/Supp-Figures).

Acknowledgments

We would like to thank the members of the epitranscriptomic working group within the Nelson, Pauli, Schroeder, and Gregory labs for their insightful comments. We would like to thank Dr. Cliff Weil for insightful comments about the SAP, Dr. Tony Schilmiller and the MSU Mass Spectrometry and Metabolomics Core for assistance in metabolomics measurements and analysis. Many thanks to Dr. Alyssa Kearly for her assistance in technical proofing of the manuscript. This work was supported by NSF IOS PGRP 2023310 (to ADLN, EHL, DP, and BDG), NSF IOS PGRP 1849708 (to BDG), NSF MCB 2427729 (to BDG), NSF MCB 2120131 (to ADLN and DP), Department of Energy (DOE) DE-AR0001101 (to DP), DE-SC0023305 (to DP), DE-SC0020401 (to DP), Cotton Incorporated 18-384, 20720, and 21-830 (to DP), NSF DBI-2019674 (to DP and ADLN), and NIH R35GM153298 (to AIS).

Author contributions

LY, GM, AEA, EHL, BDG, ADLN, and DP developed the experiment. GM, SC, BR, and DP designed the field experiment and collected field-related data. ACND and EKB extracted RNA and performed other sorghum-related molecular experiments. AE provided accessions from the SAP for this study. LY, ApS, XZ, and EKB worked with the Arabidopsis DUS2 mutants. CDG and AEA worked on the microbiome-related experiments. GM and TO generated the metabolomics data. HA and GB generated the biochemical data related to oxidative stress status. LY, RH, and DS helped and provided expertise in acquiring physiological measurements for the Arabidopsis DUS2 mutants. HF and AIS performed mass-spec experiments on the DUS2 mutants. LY analyzed all RNA-seq and developed co-expression networks and data integration. LY, KP, and SJS analyzed epitranscriptomic data. LY, BDG, ADLN, and DP wrote the manuscript.

Figure Legends

Main Figures

Figure 1. Genomic and phenotypic diversity of selected sorghum genotypes
(A) Phylogenetic tree of 356 domesticated sorghum accessions: Reconstruction of an unrooted population-level phylogeny of the Sorghum Association Panel (SAP) was performed using 7,520 bi-allelic SNPs recovered following low-quality loci removal and LD-pruning. The six selected sorghum accessions are marked along with the geographical locations colored to represent African (**green**), Asian(**purple**), and North American (**blue**) regions. The placement of the six accessions used in this study are denoted with different colored stars that match the color scheme throughout this work. **(B, C) Morphological responses of sorghum accessions under drought:** Leaf and root tissues collected at the terminal stage of sampling in the field. **(D, E) Measurement of whole plant, leaf, and root characteristics:** These terminal traits were derived from plants grown under the two watering conditions (***N*** = 5 per accession). Asterisks indicate significant differences of parameters derived from comparison of each accession relative to "A" ($P < 0.05$). **(F) Photosynthesis-related status of sorghum under drought:** Line graphs highlighting the mean WUEi, internal CO_2 concentration, and Fv'/Fm' values measured during the morning (top) and afternoon (bottom) at each time point for drought treated plants (*N* = 5 per accession). Asterisks indicate significant differences of parameters derived from comparison of each accession relative to "A" ($P < 0.05$).

Figure 2. Molecular response to drought among sorghum accessions
(A) Clustering of samples based on similarities in global transcriptome: A hierarchical clustering (Pearson correlation) of 34,130 transcripts normalized by DESeq2 was used to group transcriptomic features among samples derived from six accessions under two watering conditions over four time points. The scale indicates the Pearson correlation coefficient (PCC) score. **(B) Transcriptome profiling of the six sorghum accessions:** Two principal component analyses (PCAs), based on the top 10% most abundant transcripts, highlight the separation caused by genotypic and treatment effects (left) and genotypic and time effects (right). **(C) Numbers of differentially abundant transcripts (DATs; Drought vs Control) among six sorghum accessions:** The line graph represents number of up (circle) or down (triangle) regulated transcripts over time. **(D) GO enrichment of DATs (Drought vs Control) in the six sorghum accessions:** The enrichment levels of GO terms that were shared between DATs from at least two accessions are shown. GO terms must have been enriched at least at *FDR* < 0.05 to be shown. The scale of enrichment is shown to the left. **(E) Profiles of metabolites under two watering conditions:** The left PCA displays separation based on treatment (triangle = drought; circle = control) and genotype (colors similar as in **C**). The right PCA loading plot highlights major contributors to the left PCA, with length of the line denoting the contribution of a particular molecule to the PCA. **(F) Correlation of oxidative stress markers:** Pairwise Pearson correlation coefficient (PCC) of 24 chemicals was displayed by heatmap with the four chemical classes denoted. (**G) Profiles of representative oxidative stress chemicals under drought:** Representative markers from (**F**) were selected to display activity across time for each of the six genotypes under both control and drought conditions.

Figure 3. RCM landscape of a core set of constantly modified transcripts

(A) Transcript abundance of constantly modified transcripts: A clustered heatmap of Z-scored variance stabilized abundance values for the 342 constantly modified transcripts across the six accessions and time points. Transcripts were clustered into three groups (I, II, and III; y-axis) based on their correlation to each other, and six subgroups (C1-C6) based on similarities between accessions and timepoints (x-axis). **(B) GO enrichment of modified transcripts per group:** Numbers of transcripts per term (Counts) and level of the enrichment (-log10(*FDR*)) were plotted to represent different sets of transcripts being modified among the three groups. **(C, D) Relationship between transcript abundance and modification frequency (C) or density (D):** The scaled (Z-score) transcript and modification values (either frequency or density) were compared across the accessions and time points. Colors along the x-axis, as well as subgroups (C1-C6), correspond to those shown in (**A**). **(E) Relationship between transcript abundance and modification frequency (left) or density (right):** The scaled (Z-score) transcript and modification levels for all transcripts within a particular subgroup (C1-C6) are plotted.

Figure 4. Selection of co-expression modules contributing to drought tolerance

(A) Description of features associated with the transcripts within each module: The number of transcripts per module, correlation between traits and the module (**blue and red scale**), levels of KEGG and GO enrichment levels (purple and brown scale), and the number of TFs with enriched motifs for each module is displayed. Scales for module-trait correlation values, as well as GO-enrichment levels, are shown to the bottom right. Only those module-trait correlations with a P-value < 0.05 are shown. **(B) Description of features associated with the modified transcripts within each module:** The ratio of modified transcripts with module-trait correlations (green to yellow scale) and the enriched KEGG and GO terms for these modified transcripts (dark blue to yellow scale) are shown. Scales indicated to the right, traits and terms correspond to those shown above in (**A**). **(C) Eigene value of the Salmon (left) and Darkmagenta (right) modules:** The eigenvalue of each module, divided by accession, treatment, and time, was plotted. **(D) Percent of constantly modified transcripts within modules:** Percent of the 342 constantly modified transcripts, separated by major group as described in **Figure 3A**, assigned to the “Salmon”, “Darkmagenta”, or all other modules.

Figure 5. Examination of the Darkmagenta (photosynthesis and stress-related) network

(A) Number of genes within the Darkmagenta module that are targets of CDF3: Proximal (2 kb upstream) promoter elements for the 1,093 genes in the salmon module were examined for motif enrichment by AME. The consensus motif for CDF3 in these promoters (bottom left), and the number of CDF3 targets within the Darkmagenta module, are shown. The core motif for the Arabidopsis CDF3 is shown to the right. **(B) Transcript abundance of CDF3:** Relative normalized abundance (TPM) of CDF3 among genotypes, time points, and treatments. **(C) Pairwise correlation between CDF3 and putative targets within the module:** The correlation between CDF3 and each module was compared among three groups, including CDF3 targets, CDF3 non-targets, and randomly selected genes outside of the module ($N = 412$ for each). The number in each group was subset to match the maximum number of non-CDF3 targets within the Darkmagenta module. Significance between groups is shown (t-test, $P < 0.05$). **(D) Change in abundance in response to drought for transcripts within the Darkmagenta module:** The $\log_2$FC value of all samples (Drought/Control) for the CDF3 targets and non-targets from (C) is shown. Significance (shown) calculated with a student’s t-test. **(E) Integration of regulatory information onto a subnetwork of transcripts within the Darkmagenta module:** A

child network of 230 genes with high module membership was visualized using Cytoscape. Solid grey lines indicate co-expression only, dashed red lines indicate protein-protein interactions derived from the STRING database, and the blue dashed and arrowed lines highlight DAP-seq derived targets of CDF3. Nodes (transcripts) are colored based on trait-association, and grouped based on GO-term enrichment. Node shape (circle or square) indicate transcript modification status. Red stars indicate genes overlapping QTLs from previous GWAS (Ortiz et al., 2017). **(F) Examination of constantly modified trait-associated genes targeted by CDF3:** (Upper panel) Transcript abundance of 10 selected genes targeted by CDF3 is shown in a clustered heatmap, with the respective CDF3 correlation displayed on the left, with the correlated traits indicated on the right with gene names. (Lower panel) Measurements for correlated traits shown in (**F**) with the correlation values shown in the top left of each plot. **(G) Correlation between transcript abundance and RCM frequency and density:** The sample-wise Pearson correlation coefficient between transcript abundance and either RCM frequency (left side) or RCM density (right side) was plotted. Colors correspond to the highlighted traits in (**F**).

Figure 6. Exploring the connection between DHU and photosynthesis-associated genes

(A) Transcript abundance of DUS2: Relative normalized abundance (TPM) of *SbDUS2* among genotypes, time points, and treatments is shown. **(B) Clustering of samples based on PAM status of DHU-modified transcripts:** The DHU-modified transcripts were examined using a PCA (N = 1101 transcripts). As in (**A**), circles indicate control, triangles indicate drought stressed samples. Timepoints are indicated next to highly separated samples (e.g. W3). **(C) Genomic variation within *SbDUS2* among the six genotypes:** Schematic showing the *SbDus2* locus (gene structure shown below), with the 119 SNPs identified within the gene shown for each accession (colors defined to the right). SNPs that could impact mRNA splicing (N = 8) are boxed. **(D) Integration of regulatory information onto a subnetwork of transcripts within the Salmon module:** A child network of 77 transcripts with high module membership was visualized using Cytoscape. Solid grey lines indicate co-expression only, dashed red lines indicate protein-protein interactions derived from the STRING database, and the blue dashed and arrowed lines highlight DAP-seq derived targets of HMGDB15 or BPC6. Nodes (transcripts) are colored based on trait-association, and grouped based on GO-term enrichment. Node shape (circle or square) indicate transcript modification status. Red stars indicate genes overlapping QTLs from previous GWAS[96]. **(E) Heatmap of DHU-modified transcripts:** Transcript abundance of 9 constantly DHU- modified transcripts is shown in a clustered heatmap, with the correlation between each transcript and *DUS2* listed on the left ($P < 0.05$) and transcript-trait correlations shown to the right. **(F) Transcript abundance of DUS2 and select DHU-modified transcripts in an independent stress experiment:** Normalized (Z-score) transcript abundances are shown for *SbDUS2*, *SbCRRL*, and *SbPPDK* based on data derived from a drought study performed by Varoquaux et al., 2019[60]. Shape of each point corresponds to accession, color corresponds to gene (see key). **(G) Transcript abundance of selected transcripts in a sorghum tissue atlas dataset:** A heatmap of Z-score normalized transcript abundances are shown, divided by tissue. Developmental stage is shown to the left, specific tissue profile is shown to the right. Data originally from McCormick et al., 2018[55]. **(H) Correlation between *SbPPDK* and *SbCRRL* abundance and DHU density (top) and frequency (bottom):** Transcript abundance (derived from the tissue atlas) was Z-score normalized and plotted against normalized DHU density or frequency. Line color corresponds to the two transcripts.

Figure 7. Functional characterization of *atdus2* mutants

(A) Examination of the *AtDUS2* locus: Schematic of the *AtDUS2* locus (top) with the T-DNA insertion site in the third exon denoted. A protein alignment of the Arabidopsis and sorghum DUS2 proteins is depicted below, with the DUS-FMN-like domain outlined in orange. Alignment coverage and amino acid identity is shown below. **(B) Germination percentage of *atdus2* mutants compared to Col-0 (control) plants:** Following stratification, MS plates containing both *atdus2* mutants and WT seeds were imaged every three hours, with germination calculated as the emergence of cotyledons. Nine plates were included for analysis ($P < 0.05$). **(C) Vegetative growth monitoring of Col-0 and *dus2* plants under control and heat conditions:** (Top panel) Image-based phenotyping was performed over 12 days of vegetative growth under control and heat conditions, beginning after the emergence of the second set of true leaves. Images were captured every 30 minutes, with growth rates calculated using PlantCV. The right panel depicts survival rates over time. (Bottom panel) Selected plant images from the 10_{th} day of image collection. **(D) Reproductive growth features of Col-0 and *atdus2* plants:** Growth habit of the two backgrounds at 36 days (top) and the percent of plants with flower buds visible plotted over time (bottom). The number of plants examined is shown. **(E) Photosynthetic efficiency of Col-0 and *atdus2* plants:** Photosynthetic efficiency was measured on the largest rosette leaf using a LI-6800 system. CO_2 net assimilation rate and stomatal conductance under heat (24h at 37°C) and control are shown. **(F) DHU modification levels between Col-0 and *atdus2* mutants:** Concentration of 5,6-dihydrouridine between two genetic background using LC–MS/MS analysis (left) and DHU modification frequency is shown for each background and treatment (right). This modification level was compared between the two genotypes under the two conditions using a Student's t-test. **(G) PCA of top 25% most variable transcripts in *atdus2* and Col-0:** A PCA of the transcriptomic profiles for the two backgrounds (indicated by color) over time (size of point) and between heat and control conditions (shape of point) is shown ($N = 5207$ transcripts). **(H) Comparison of enriched GO terms associated with heat responsive transcripts in Col-0 and *dus2* plants:** Differentially abundant transcripts associated with the last day of treatment were used for GO enrichment calculations. GO terms are divided by genotype in which they were enriched, and whether they were associated with up (left) or down-regulated transcripts (right; cut-off, $FDR < 0.05$).

Supplementary Figures

Figure S1. Genomic composition of variants of six sorghum accessions: 8,725,449 variants (SNPs and INDELs) derived from variants calling (GATK4, reference genome: **McCormick et al., 2018)** were annotated based on genomic classifications using VEP tool, the unique variants of each accession were calculated across upstream, downstream, intron, and splicing regions.

Figure S2. Phenotypic variation of six sorghum accessions

(A) Schematic graph of experiment: Three factors (genotype, condition, and timepoint) of the experiment design are displayed. The "blue" and "and "red" box highlight the samples been collected for metabolites data profile and microbiome data collection. **(B, C) Metabolites contents:** The content of sucrose and chlorogenic acids measured across six genotypes, four timepoints, and two conditions highlight the distinct response of accession "A". **(D, E) Features of microbiome derivatives:** The PCA and PCA loading of bacteria ($N = 515$) and fungi ($N = 46$) uncovered the lower level of response of accession "A" under drought conditions.

Figure S3. Upset plots of DATs between two watering conditions

(A, B) Comparisons of the up-regulated (left) and down-regulated (right) transcripts among samples per genotype across each time point, identified in a pair-wise manner. Numbers of DATs were displayed using the upset plots. The top bar charts denote the shared (or unique) DATs corresponding to the inclusion of samples below, and bottom-left bars denote the total numbers of DATs derived from each pair-wise comparison.

Figure S4. Global characteristics of RNA covalent modifications (RCMs)

(A) Modification density and modification frequency: The RCM density is characterized by the number of modified sites per transcript **(numbers of red triangles)**, and the RCM frequency is estimated by the ratio between the total reads containing modifications **(red lines)** and the total mapped reads per transcript. **(B) PCC of modification frequency between replicates among samples:** The correlation levels of modification frequency between replicates per sample derived from HAMR and Modtect were compared using a boxplot where the HAMR has a more stable estimation of modification frequency for each sample. **(C) Relationship between transcript abundance and RCM frequency or density:** The transcripts-wise correlation between RCM frequency and transcript abundance was estimated using the $\log_2$(TPM) of each transcript and the $\log_2$(modification frequency). The correlation between RCM density and transcript abundance was measured using the $\log_2$(TPM) score and the $\log_2$(transcript length normalized density).

Figure S5. Effects of modification type, genotype, time points, and treatment on global RCMs

(A) RCMs landscape of modified transcripts: The effects of six modification types, six genotypes, four-time points, and two watering conditions were estimated using PCA based on the presence and absence of particular modification types (denoted by a matrix comprised of “0” and “1” to represent the status of each RCM type). **(B) RCMs density of modified transcripts for each modification type:** The effects of six genotypes, four-time points, and two watering conditions were estimated using PCA based on the modification density of particular modification types (denoted by a matrix comprised of modification density score of each RCM type).

Figure S6. Characterization of global RCM status

(A) RCMs landscape of modified transcripts: Four sets of inputs were used to uncover the status of modified transcripts, including the presence and absence of RMCs (PAM), modification density, modification frequency, and transcript abundance. **(B) Characteristics of modification density or frequency relative to treatment, time, and genotypes:** The mean of density and frequency score uncovered the accession “A” with lower density under drought and higher frequency under control relative to other accessions.

Figure S7. Comparison of RCMs status among genotypes

(A, B) Number of modified transcripts: Numbers of modified transcripts per genotype were compared using a boxplot (t-test, $P < 0.05$). The accessions “A” and “F” revealed a reduced number of modified transcripts under drought conditions while other accessions exhibited a reverse pattern. **(C) Mean modification frequency under two conditions per genotype:** The mean modification frequency scores were compared for each genotype under drought and control conditions. **(D) Mean modification density under two conditions per genotype:** The mean modification density scores were compared for each genotype under drought and control conditions. **(E) Status of transcription of**

modified transcripts unique to each genotype: Transcript abundance of genotypic-specific modified transcripts were displayed using a heatmap where the results suggest the higher levels of transcription of genotypic-specific modified transcripts. **(F) Comparison of transcript abundance of genotype-specific modified transcripts relative to background:** Abundance levels of genotype-specific modified transcripts in the genotype in which they're modified compared to all other genotypes. The boxplots highlight the higher transcription level in certain genotypes compared with the background (t-test, $P < 0.05$).

Figure S8. Maximum likelihood phylogeny of plant and budding yeast DUS proteins

A maximum likelihood phylogeny was inferred from a protein alignment of DUS genes identified in Arabidopsis (AT), Sorghum (Sobic), Oryza sativa (rice, Os), Physcomitreum patens (Pp), and Saccharomyces cerevisiae (Sc) through reciprocal best BLASTp. Values at each node represent bootstrap support (out of 100 replicates).

Figure S9. Transcriptome features of *DUS2 in* sorghum and Arabidopsis

(A) Correlation of DUS2 transcript abundance relative to PPDK and CRRL: The transcript abundance was quantified using the TPM score derived from the salmon pipeline from independent RNA-seq samples, including two sorghum datasets (**McCormick et al., 2018 and Varoquaux et al., 2019**), the heat stress data developed in this study and three types of abiotic stress in Klepikova et al 2016 in Arabidopsis. Pearson correlation was calculated to estimate the relation of PPDK and CRRL relative to DUS2. **(B) Comparison of epitranscriptome features between *atdus2* mutants and Col-0:** Normalized reads been mapped to *AtDUS2* transcripts and number of dihydrouridine-modified sites detected using Modtect were compared in two genetic backgrounds under control and heat conditions. **(C) Confirmation of T-DNA insertion within the 5'-end third exon of AtDUS using RNA-seq reads mapping:** Mapped reads from DUS2 mutants and Col-0 background were visualized using the IGV tool to confirm mutant samples contain homozygous T-DNA insertion that disrupts the expression of full-length *DUS2*.

Refences

1. Morris, G. P. *et al.* Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 453–458 (2013).
2. Abreha, K. B. *et al.* Sorghum in dryland: morphological, physiological, and molecular responses of sorghum under drought stress. *Planta* **255**, 20 (2021).
3. Sehgal, A. *et al.* Drought or/and heat-stress effects on seed filling in food crops: Impacts on functional biochemistry, seed yields, and nutritional quality. *Front. Plant Sci.* **9**, 1705 (2018).
4. Alizadeh, M. R. *et al.* A century of observations reveals increasing likelihood of continental-scale compound dry-hot extremes. *Sci. Adv.* **6**, eaaz4571 (2020).
5. Fàbregas, N. & Fernie, A. R. The metabolic response to drought. *J. Exp. Bot.* **70**, 1077–1085 (2019).
6. Zargar, S. M. *et al.* Impact of drought on photosynthesis: Molecular perspective. *Plant Gene* **11**, 154–159 (2017).
7. Hasanuzzaman, M., Nahar, K., Gill, S. S. & Fujita, M. Drought stress responses in plants, oxidative stress, and antioxidant defense. *Climate change and plant abiotic stress tolerance* 209–250 (2013).
8. Melandri, G. *et al.* Assessing drought and heat stress-induced changes in the cotton leaf metabolome and their relationship with hyperspectral reflectance. *Front. Plant Sci.* **12**, 751868 (2021).
9. Zhu, J.-K. Abiotic stress signaling and responses in plants. *Cell* **167**, 313–324 (2016).
10. Kim, J.-S., Kidokoro, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. Regulatory networks in plant responses to drought and cold stress. *Plant Physiol.* **195**, 170–189 (2024).
11. Mahmood, T. *et al.* Insights into drought stress signaling in plants and the molecular genetic basis of cotton drought tolerance. *Cells* **9**, 105 (2019).
12. Sato, H., Mizoi, J., Shinozaki, K. & Yamaguchi-Shinozaki, K. Complex plant responses to drought and heat stress under climate change. *Plant J.* **117**, 1873–1892 (2024).
13. Joshi, R. *et al.* Transcription factors and plants response to drought stress: Current understanding and future directions. *Front. Plant Sci.* **7**, 1029 (2016).
14. Manna, M. *et al.* Transcription factors as key molecular target to strengthen the drought stress tolerance in plants. *Physiol. Plant.* **172**, 847–868 (2021).
15. Bang, S. W. *et al.* Overexpression of OsTF1L, a rice HD-Zip transcription factor, promotes lignin biosynthesis and stomatal closure that improves drought tolerance. *Plant Biotechnol. J.* **17**, 118–131 (2019).
16. Yu, L. *et al.* Regulation of a single inositol 1-phosphate synthase homeologue by HSFA6B contributes to fibre yield maintenance under drought conditions in upland cotton. *Plant Biotechnol. J.* **22**, 2756–2772 (2024).
17. Ortiz, D. & Salas-Fernandez, M. G. Dissecting the genetic control of natural variation in sorghum photosynthetic response to drought stress. *J. Exp. Bot.* **73**, 3251–3267 (2022).
18. Enyew, M. *et al.* Novel sources of drought tolerance in sorghum landraces revealed via the analyses of genotype-by-environment interactions. *Front. Plant Sci.* **13**, 1062984 (2022).
19. Tsehaye, Y., Menamo, T. M., Abay, F., Tadesse, T. & Bantte, K. Multi-locus genome-wide association study for grain yield and drought tolerance indices in sorghum accessions. *Plant Genome* **17**, e20505 (2024).
20. Sharma, B., Prall, W., Bhatia, G. & Gregory, B. D. The diversity and functions of plant RNA modifications: What we know and where we go from here. *Annu. Rev. Plant Biol.* **74**, 53–85 (2023).
21. Prall, W., Ganguly, D. R. & Gregory, B. D. The covalent nucleotide modifications within plant mRNAs: What we know, how we find them, and what should be done in the future. *Plant Cell* **35**, 1801–1816 (2023).

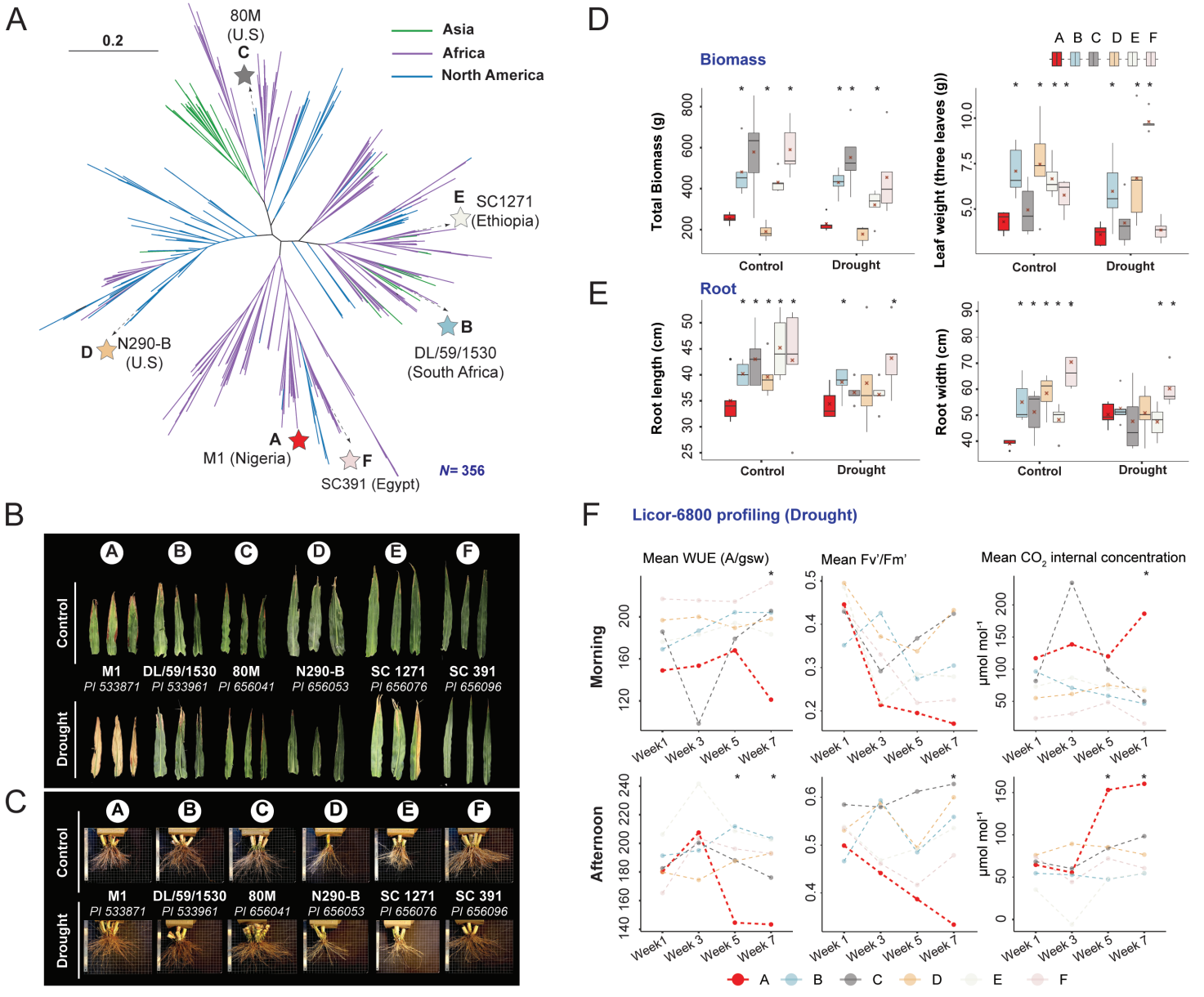
22. Gilbert, W. V. & Nachtergaele, S. MRNA regulation by RNA modifications. *Annu. Rev. Biochem.* **92**, 175–198 (2023).
23. Arribas-Hernández, L. *et al.* An m6A-YTH module controls developmental timing and morphogenesis in Arabidopsis. *Plant Cell* **30**, 952–967 (2018).
24. Luo, J.-H. *et al.* Natural variation in RNA m6A methylation and its relationship with translational status. *Plant Physiol.* **182**, 332–344 (2020).
25. Cheng, Q. *et al.* Coordination of m6A mRNA methylation and gene transcriptome in rice response to cadmium stress. *Rice (N. Y.)* **14**, 62 (2021).
26. Anderson, S. J. *et al.* N6-methyladenosine inhibits local ribonucleolytic cleavage to stabilize mRNAs in Arabidopsis. *Cell Rep.* **25**, 1146-1157.e3 (2018).
27. Schaefer, M., Pollex, T., Hanna, K. & Lyko, F. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res.* **37**, e12 (2009).
28. Sun, H., Li, K., Liu, C. & Yi, C. Regulation and functions of non-m6A mRNA modifications. *Nat. Rev. Mol. Cell Biol.* **24**, 714–731 (2023).
29. Zhang, Y., Lu, L. & Li, X. Detection technologies for RNA modifications. *Exp. Mol. Med.* **54**, 1601–1616 (2022).
30. Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat. Rev. Genet.* **18**, 275–291 (2017).
31. Ryvkin, P. *et al.* HAMR: high-throughput annotation of modified ribonucleotides. *RNA* **19**, 1684–1692 (2013).
32. Vandivier, L. E. *et al.* Chemical modifications Mark alternatively spliced and uncapped messenger RNAs in Arabidopsis. *Plant Cell* **27**, 3024–3037 (2015).
33. Tan, K.-T., Ding, L.-W., Wu, C.-S., Tenen, D. G. & Yang, H. Repurposing RNA sequencing for discovery of RNA modifications in clinical cohorts. *Sci. Adv.* **7**, eabd2605 (2021).
34. Melandri, G. *et al.* Biomarkers for grain yield stability in rice under drought stress. *J. Exp. Bot.* **71**, 669–683 (2020).
35. Porra, R. J., Thompson, W. A. & Kriedemann, P. E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta Bioenerg.* **975**, 384–394 (1989).
36. Riewe, D. *et al.* A tyrosine aminotransferase involved in tocopherol synthesis in Arabidopsis: Tyrosine aminotransferase affecting tocopherols. *Plant J.* **71**, 850–859 (2012).
37. Hodges, D. M., DeLong, J. M., Forney, C. F. & Prange, R. K. Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds. *Planta* **207**, 604–611 (1999).
38. Levine, R., Williams, J. A., Stadtman, E. P. & Shacter, E. Carbonyl assays for determination of oxidatively modified proteins. *Methods Enzymol.* **233**, 346–357 (1994).
39. Feierabend, J. & Beevers, H. Developmental studies on microbodies in wheat leaves : I. Conditions influencing enzyme development: I. Conditions influencing enzyme development. *Plant Physiol.* **49**, 28–32 (1972).
40. Schwitzguebel, J. P. & Siegenthaler, P. A. Purification of peroxisomes and mitochondria from spinach leaf by percoll gradient centrifugation. *Plant Physiol.* **75**, 670–674 (1984).
41. Steczko, J., Donoho, G. A., Dixon, J. E., Sugimoto, T. & Axelrod, B. Effect of ethanol and low-temperature culture on expression of soybean lipoxygenase L-1 in Escherichia coli. *Protein Expr. Purif.* **2**, 221–227 (1991).
42. Benzie, I. & Strain, J. Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods Enzymol.* **299**, 15–27 (1999).
43. Zhang, Q. *et al.* A simple 96-well microplate method for estimation of total polyphenol content in seaweeds. *J. Appl. Phycol.* **18**, 445–450 (2006).

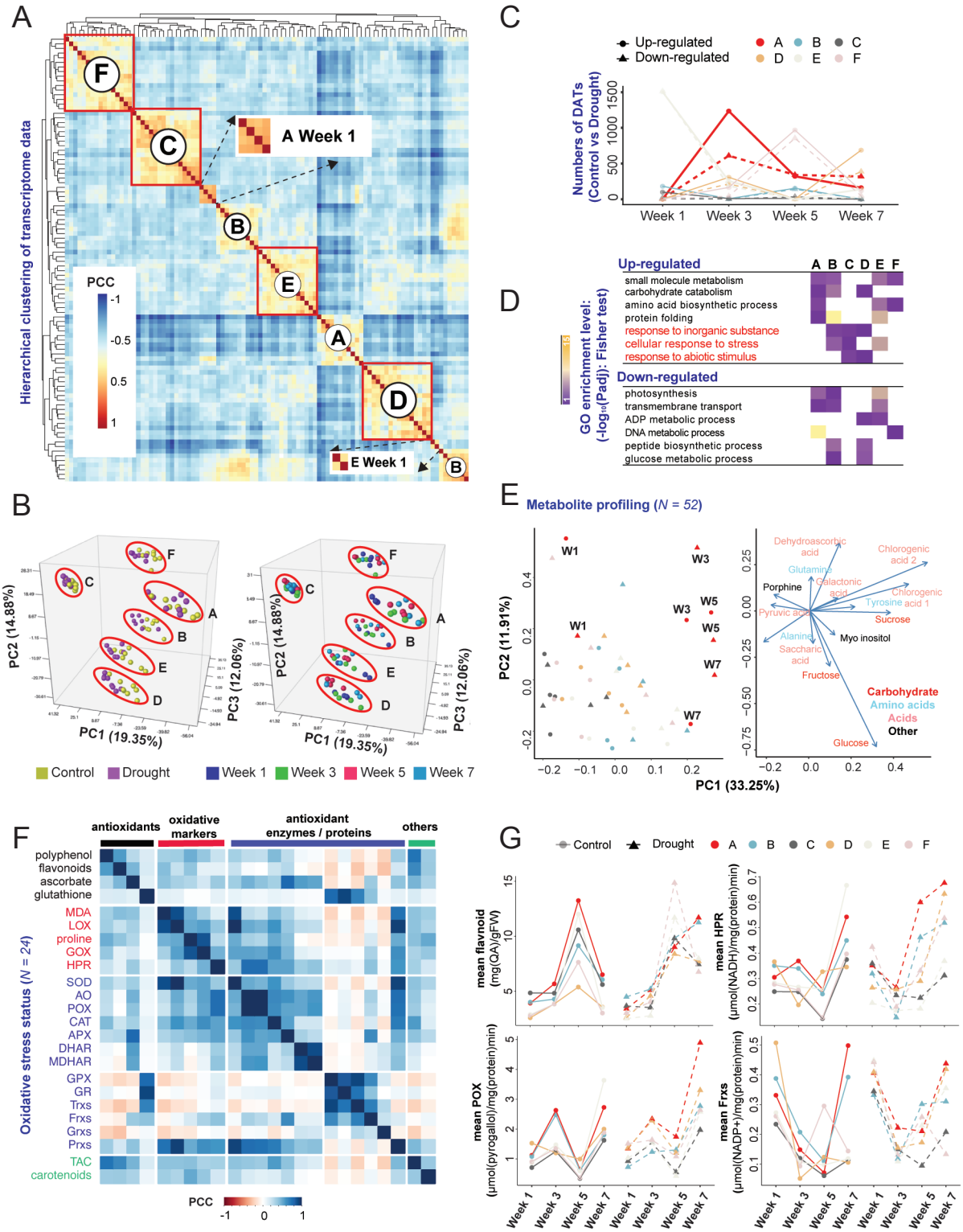
44. Murshed, R., Lopez-Lauri, F. & Sallanon, H. Microplate quantification of enzymes of the plant ascorbate-glutathione cycle. *Anal. Biochem.* **383**, 320–322 (2008).
45. Dhindsa, R. S., Plumb-Dhindsa, P. L. & Reid, D. M. Leaf senescence and lipid peroxidation: Effects of some phytohormones, and scavengers of free radicals and singlet oxygen. *Physiol. Plant.* **56**, 453–457 (1982).
46. Aebi, H. [13] Catalase in vitro. *Methods Enzymol.* (1984).
47. Wase, N., Abshire, N. & Obata, T. High-throughput profiling of metabolic phenotypes using high-resolution GC-MS. *Methods Mol. Biol.* **2539**, 235–260 (2022).
48. Yu, L. *et al.* Development of a mobile, high-throughput, and low-cost image-based plant growth phenotyping system. *Plant Physiol.* (2024) doi:10.1093/plphys/kiae237.
49. Fahlgren, N. *et al.* A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in Setaria. *Mol. Plant* **8**, 1520–1535 (2015).
50. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497 (1962).
51. Brazel, A. J., Manoj, N. S., Turck, F. & Ó'Maoiléidigh, D. S. Measuring CO2 assimilation of Arabidopsis thaliana whole plants and seedlings. *Plant Sci.* **350**, 112295 (2025).
52. Boyes, D. C. *et al.* Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants: A model for high throughput functional genomics in plants. *Plant Cell* **13**, 1499–1510 (2001).
53. Burgess, S. J. QIAGEN RNeasy Plant RNA Extraction Protocol (Modified). (2023).
54. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
55. McCormick, R. F. *et al.* The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* **93**, 338–354 (2018).
56. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
57. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. Dobin, A. & Gingeras, T. R. Mapping RNA-seq reads with STAR. *Curr. Protoc. Bioinformatics* **51**, 11.14.1-11.14.19 (2015).
59. Cheng, C.-Y. *et al.* Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J.* **89**, 789–804 (2017).
60. Varoquaux, N. *et al.* Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 27124–27132 (2019).
61. Boatwright, J. L. *et al.* Sorghum Association Panel whole-genome sequencing establishes cornerstone resource for dissecting genomic diversity. *Plant J.* **111**, 888–904 (2022).
62. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).
63. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
64. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *BIOINFORMATICS APPLICATIONS NOTE* **25**, (2009).
65. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
66. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
67. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
68. Liao, Y., Smyth, G. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2013).

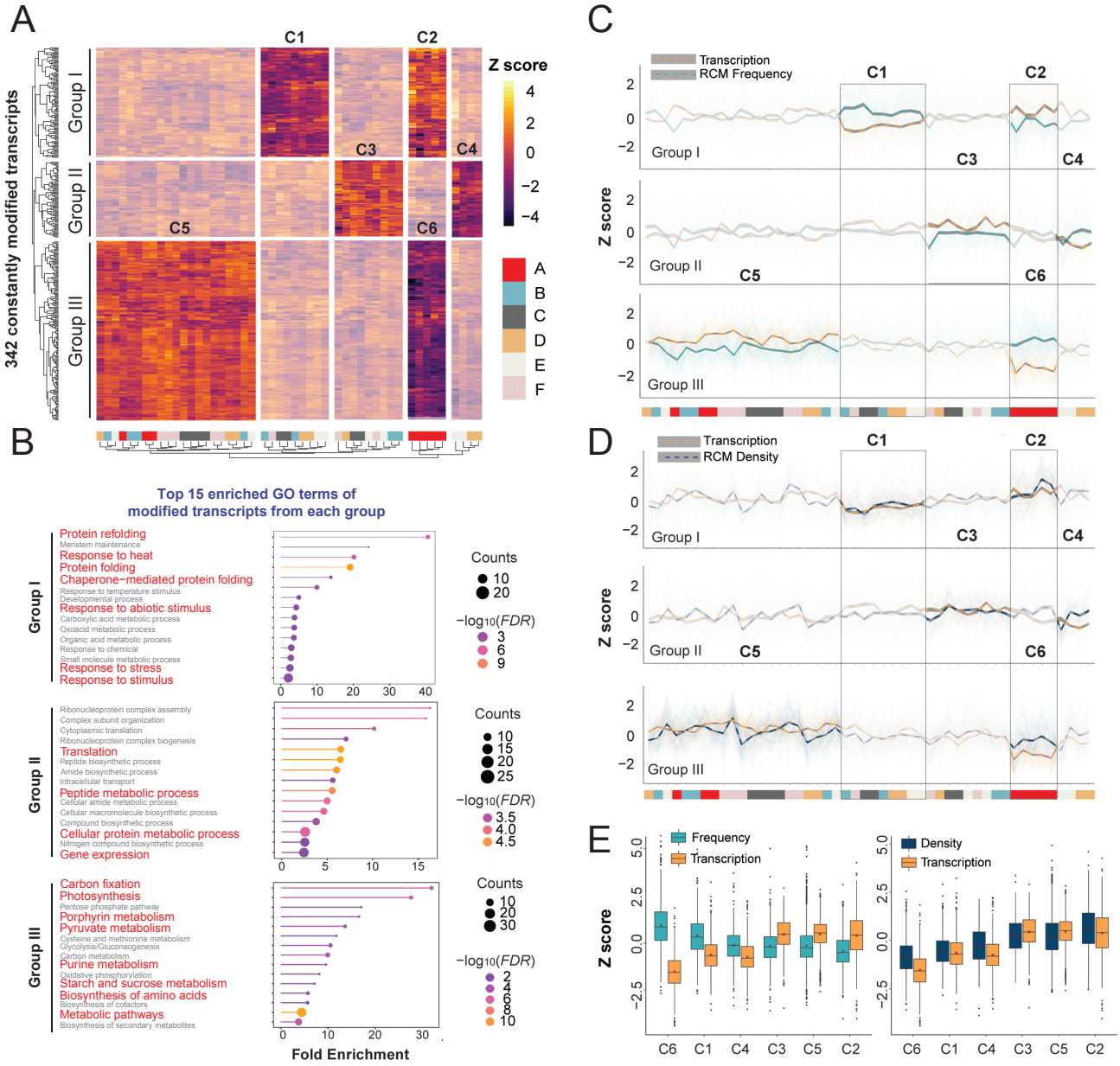
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
70. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
71. Zheng, Y. *et al.* ITAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* **9**, 1667–1670 (2016).
72. O'Malley, R. C. *et al.* Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**, 1280–1292 (2016).
73. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**, W39-49 (2015).
74. von Mering, C. *et al.* STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* **31**, 258–261 (2003).
75. Alexa, A., Rahnenführer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).
76. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
77. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
78. Team, R. C. RA language and environment for statistical computing, R Foundation for Statistical. *Computing* (2020).
79. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
80. Wang, M. *et al.* Genomic innovation and regulatory rewiring during evolution of the cotton genus Gossypium. *Nat. Genet.* **54**, 1959–1971 (2022).
81. Obata, T. *et al.* Metabolite profiles of maize leaves in drought, heat, and combined stress field trials reveal the relationship between metabolism and grain yield. *Plant Physiol.* **169**, 2665–2683 (2015).
82. Das, K. & Roychoudhury, A. Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Front. Environ. Sci.* **2**, (2014).
83. Cai, Z. *et al.* The m6A reader ECT8 is an abiotic stress sensor that accelerates mRNA decay in Arabidopsis. *Plant Cell* **36**, 2908–2926 (2024).
84. Lee, K. P. *et al.* The m6A reader ECT1 drives mRNA sequestration to dampen salicylic acid-dependent stress responses in Arabidopsis. *Plant Cell* **36**, 746–763 (2024).
85. Palos, K., Nelson Dittrich, A. C., Lyons, E. H., Gregory, B. D. & Nelson, A. D. L. Comparative analyses suggest a link between mRNA splicing, stability, and RNA covalent modifications in flowering plants. *BMC Plant Biol.* **24**, 768 (2024).
86. Corrales, A.-R. *et al.* Multifaceted role of cycling DOF factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in Arabidopsis. *Plant Cell Environ.* **40**, 748–764 (2017).
87. Corrales, A.-R. *et al.* Characterization of tomato Cycling Dof Factors reveals conserved and new functions in the control of flowering time and abiotic stress responses. *J. Exp. Bot.* **65**, 995–1012 (2014).
88. Noguero, M., Atif, R. M., Ochatt, S. & Thompson, R. D. The role of the DNA-binding One Zinc Finger (DOF) transcription factor family in plants. *Plant Sci.* **209**, 32–45 (2013).
89. Yanagisawa, S. & Schmidt, R. J. Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J.* **17**, 209–214 (1999).
90. Ascencio-Ibáñez, J. T. *et al.* Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiol.* **148**, 436–454 (2008).
91. Leonetti, P. *et al.* Endogenous activated small interfering RNAs in virus-infected Brassicaceae crops show a common host gene-silencing pattern affecting photosynthesis and stress response. *New Phytol.* **229**, 1650–1664 (2021).

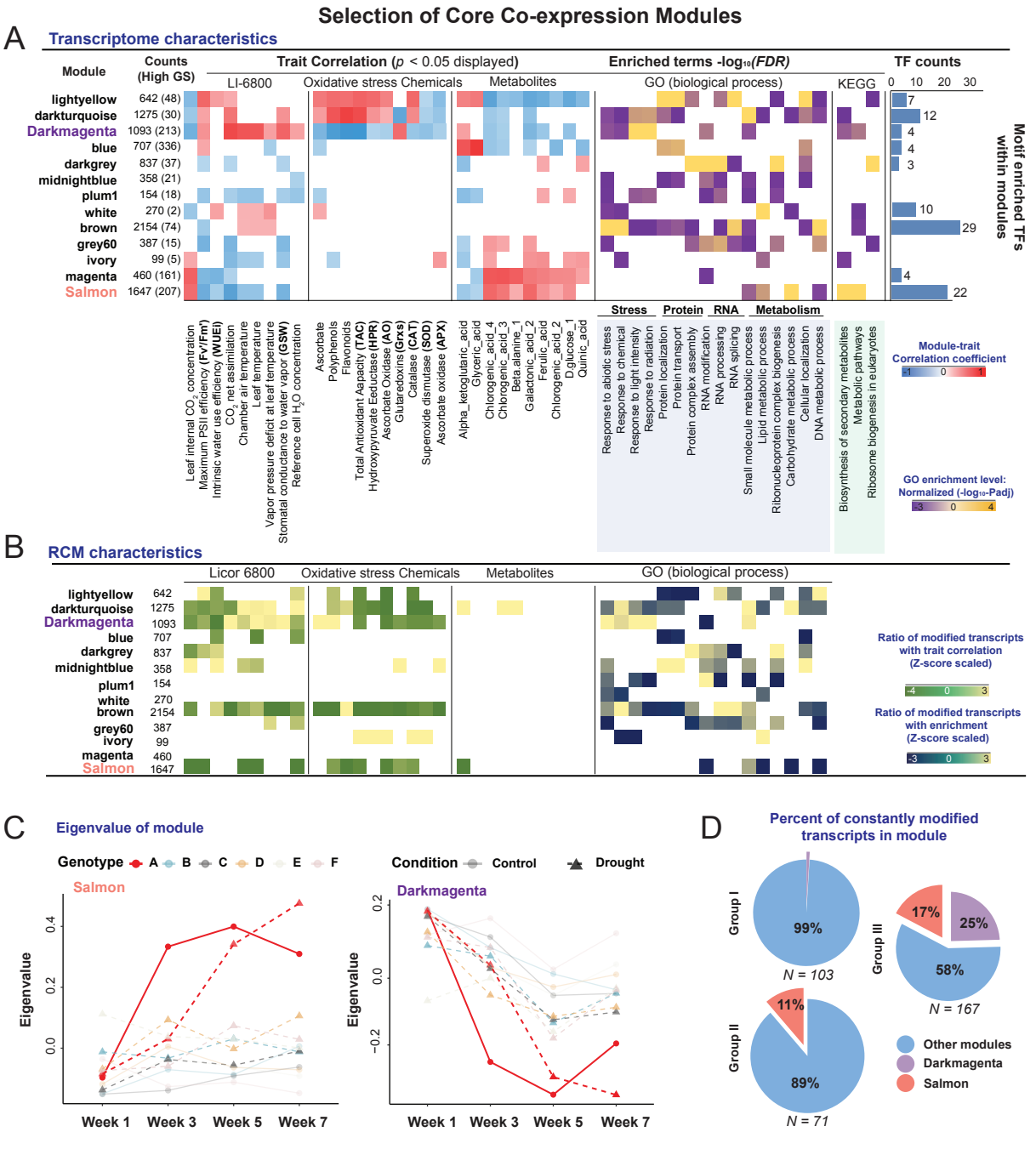
92. Feria, A. B. *et al.* Phosphoenolpyruvate carboxylase (PEPC) and PEPC-kinase (PEPC-k) isoenzymes in Arabidopsis thaliana: role in control and abiotic stress conditions. *Planta* **244**, 901–913 (2016).
93. Antosch, M., Mortensen, S. A. & Grasser, K. D. Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes. *Plant Physiol.* **159**, 875–883 (2012).
94. Zhang, Z., Zhang, T. & Ma, L. Analysis of basic pentacysteine6 transcription factor involved in abiotic stress response in Arabidopsis thaliana. *Front. Genet.* **14**, 1097381 (2023).
95. Mallik, R. *et al.* Identification of genome-wide targets and DNA recognition sequence of the Arabidopsis HMG-box protein AtHMGB15 during cold stress response. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194644 (2020).
96. Ortiz, D., Hu, J. & Salas Fernandez, M. G. Genetic architecture of photosynthesis in Sorghum bicolor under non-stress and cold stress conditions. *J. Exp. Bot.* **68**, 4545–4557 (2017).
97. Mace, E. *et al.* The Sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. *Züchter Genet. Breed. Res.* **132**, 751–766 (2019).
98. Yamamoto, H., Peng, L., Fukao, Y. & Shikanai, T. An Src homology 3 domain-like fold protein forms a ferredoxin binding site for the chloroplast NADH dehydrogenase-like complex in Arabidopsis. *Plant Cell* **23**, 1480–1493 (2011).
99. Wang, P. *et al.* Chloroplastic NAD(P)H dehydrogenase in tobacco leaves functions in alleviation of oxidative damage caused by temperature stress. *Plant Physiol.* **141**, 465–474 (2006).
100. Eastmond, P. J. *et al.* Arabidopsis uses two gluconeogenic gateways for organic acids to fuel seedling establishment. *Nat. Commun.* **6**, 6659 (2015).
101. Yadav, S. & Mishra, A. Ectopic expression of C4 photosynthetic pathway genes improves carbon assimilation and alleviate stress tolerance for future climate change. *Physiol. Mol. Biol. Plants* **26**, 195–209 (2020).
102. Zhang, Y. *et al.* Characterization of maize leaf pyruvate orthophosphate dikinase using high throughput sequencing. *J. Integr. Plant Biol.* **60**, 670–690 (2018).
103. Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D. & Penin, A. A. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant J.* **88**, 1058–1070 (2016).
104. Lesk, C. *et al.* Compound heat and moisture extreme impacts on global crop yields under climate change. *Nat. Rev. Earth Environ.* **3**, 872–889 (2022).
105. Pardo, J. *et al.* Cross-species predictive modeling reveals conserved drought responses between maize and sorghum. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2216894120 (2023).
106. Johnson, S. M., Cummins, I., Lim, F. L., Slabas, A. R. & Knight, M. R. Transcriptomic analysis comparing stay-green and senescent Sorghum bicolor lines identifies a role for proline biosynthesis in the stay-green trait. *J. Exp. Bot.* **66**, 7061–7073 (2015).
107. Pardo, J. & VanBuren, R. Evolutionary innovations driving abiotic stress tolerance in C4 grasses and cereals. *Plant Cell* **33**, 3391–3401 (2021).
108. Marks, R. A., Van Der Pas, L., Schuster, J., Gilman, I. S. & VanBuren, R. Convergent evolution of desiccation tolerance in grasses. *Nat. Plants* **10**, 1112–1125 (2024).
109. Tang, Y. *et al.* OsNSUN2-mediated 5-methylcytosine mRNA modification enhances rice adaptation to high temperature. *Dev. Cell* **53**, 272-286.e7 (2020).
110. Yang, W. *et al.* The N1-methyladenosine methylome of petunia mRNA. *Plant Physiol.* **183**, 1710–1724 (2020).
111. Prall, W. *et al.* Pathogen-induced m6A dynamics affect plant immunity. *Plant Cell* **35**, 4155–4172 (2023).
112. Yu, F. *et al.* Molecular basis of dihydrouridine formation on tRNA. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19593–19598 (2011).
113. Brégeon, D. *et al.* Dihydrouridine in the transcriptome: New Life for this ancient RNA chemical modification. *ACS Chem. Biol.* **17**, 1638–1657 (2022).

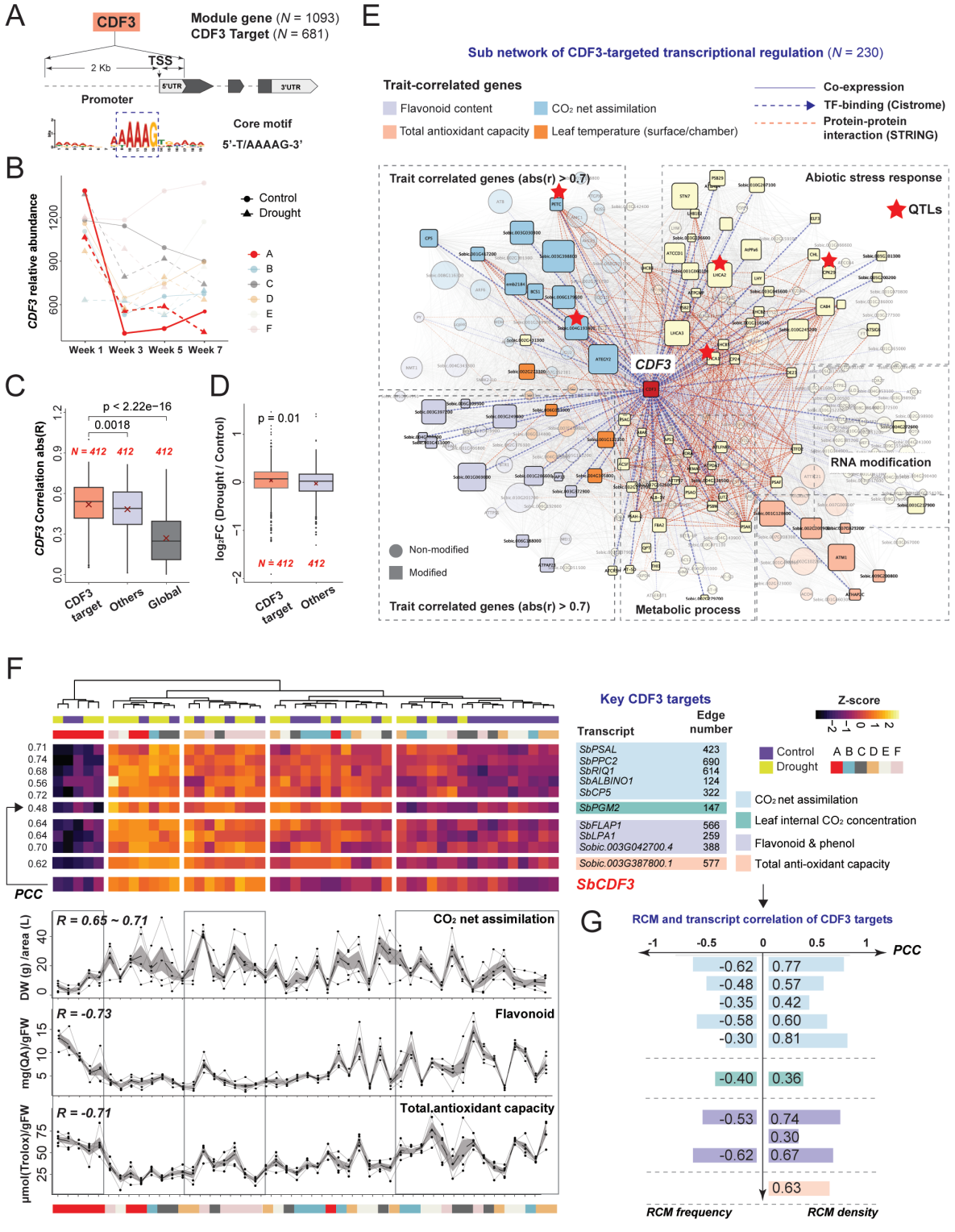
114. Lombard, M. *et al.* Evolutionary diversity of Dus2 enzymes reveals novel structural and functional features among members of the RNA dihydrouridine synthases family. *Biomolecules* **12**, 1760 (2022).
115. Finet, O., Yague-Sanz, C., Marchand, F. & Hermand, D. The Dihydrouridine landscape from tRNA to mRNA: a perspective on synthesis, structural impact and function. *RNA Biol.* **19**, 735–750 (2022).
116. Rodell, R., Robalin, N. & Martinez, N. M. Why U matters: detection and functions of pseudouridine modifications in mRNAs. *Trends Biochem. Sci.* **49**, 12–27 (2024).

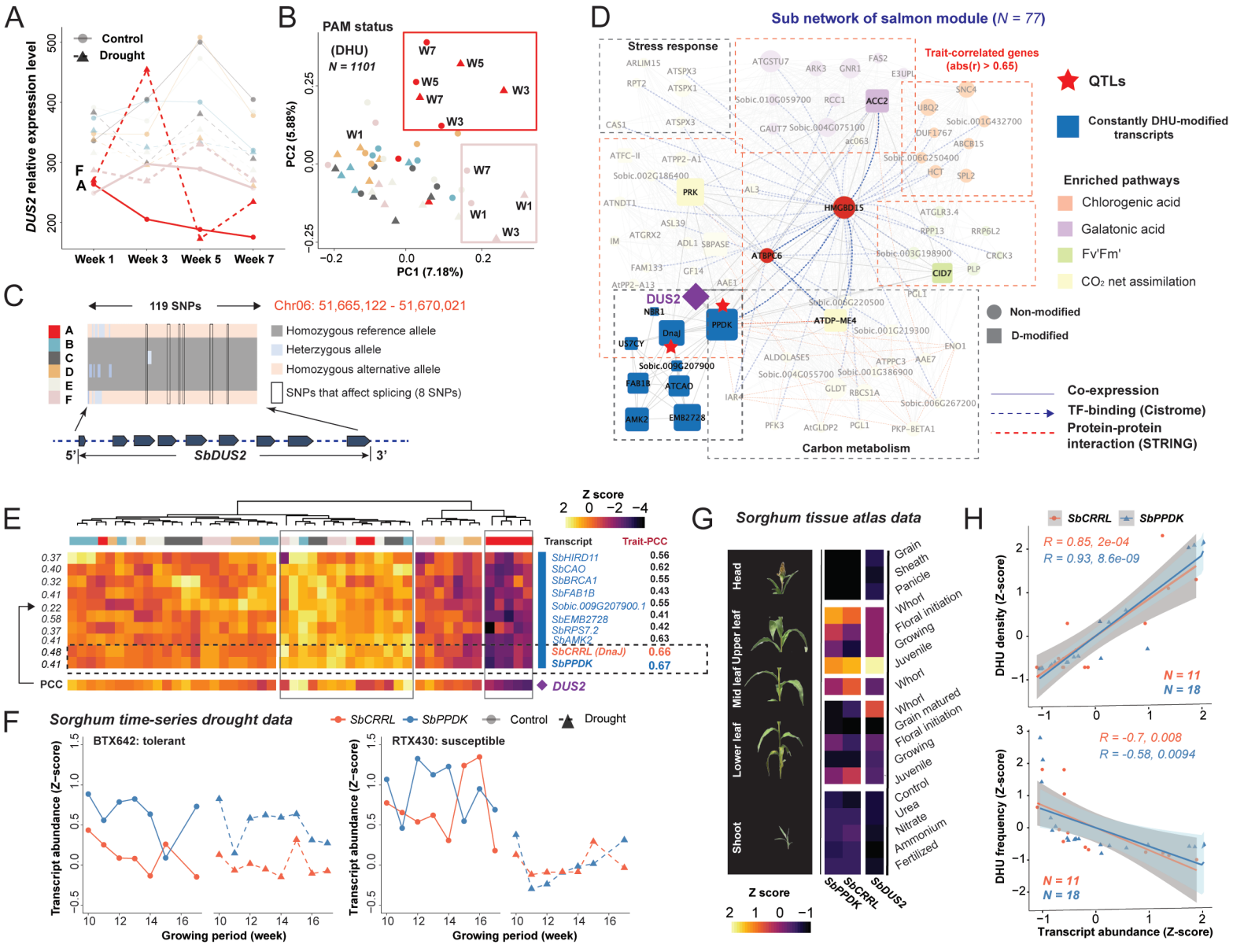

A
0.2
80M
(U.S)
C
Asia
Africa
North America
E SC1271
(Ethiopia)
B
DL/59/1530
(South Africa)
D N290-B
(U.S)
A
M1 (Nigeria)
F
SC391 (Egypt)
N= 356
B
A B C D E F
Control
Drought
M1 PI 533871
DL/59/1530 PI 533961
80M PI 656041
N290-B PI 656053
SC 1271 PI 656076
SC 391 PI 656096
C
A B C D E F
Control
Drought
M1 PI 533871
DL/59/1530 PI 533961
80M PI 656041
N290-B PI 656053
SC 1271 PI 656076
SC 391 PI 656096
D
Biomass
A B C D E F
Total Biomass (g)
Control
Drought
Leaf weight (three leaves (g))
Control
Drought
E
Root
Root length (cm)
Control
Drought
Root width (cm)
Control
Drought
F
Licor-6800 profiling (Drought)
Mean WUE (A/gsw)
Mean Fv'/Fm'
Mean CO2 internal concentration
Morning
Afternoon
µmol mol-1
Week1
Week 3
Week 5
Week 7
A B C D E F

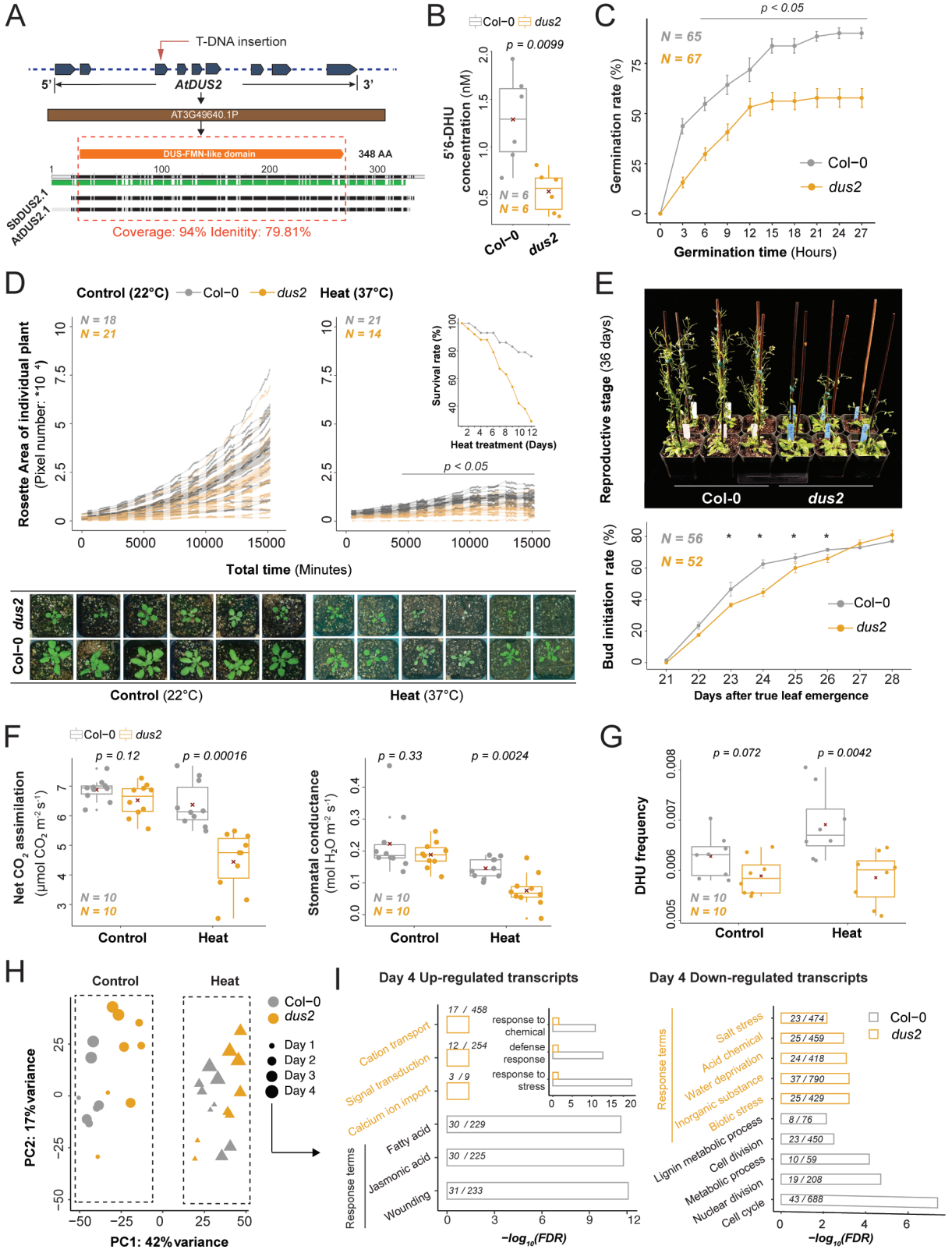

A
Hierarchical clustering of transcriptome data
PCC
-1
-0.5
0.5
1
A Week 1
E Week 1
B
PC1 (19.35%)
PC2 (14.88%)
PC3 (12.06%)
Control
Drought
Week 1
Week 3
Week 5
Week 7
C
Up-regulated
Down-regulated
Numbers of DATs (Control vs Drought)
Week 1
Week 3
Week 5
Week 7
D
Up-regulated
small molecule metabolism
carbohydrate catabolism
amino acid biosynthetic process
protein folding
response to inorganic substance
cellular response to stress
response to abiotic stimulus
Down-regulated
photosynthesis
transmembrane transport
ADP metabolic process
DNA metabolic process
peptide biosynthetic process
glucose metabolic process
GO enrichment level: (-log10(Padj): Fisher test)
E
Metabolite profiling (N = 52)
PC1 (33.25%)
PC2 (11.91%)
Dehydroascorbic acid
Chlorogenic acid 2
Glutamine
Galactonic acid
Porphine
Chlorogenic acid 1
Tyrosine
Pyruvic acid
Sucrose
Alanine
Myo inositol
Saccharic acid
Fructose
Glucose
Carbohydrate
Amino acids
Acids
Other
F
antioxidants
oxidative markers
antioxidant enzymes / proteins
others
polyphenol
flavonoids
ascorbate
glutathione
MDA
LOX
proline
GOX
HPR
SOD
AO
POX
CAT
APX
DHAR
MDHAR
GPX
GR
Trxs
Frxs
Grxs
Prxs
TAC
carotenoids
Oxidative stress status (N = 24)
PCC
-1
0
1
G
Control
Drought
mean flavnoid (mg(QA)/gFW)
mean HPR (µmol(NADH)/mg(protein)min)
mean POX (µmol(pyrogallol)/mg(protein)min)
mean Frxs (µmol(NADP+)/mg(protein)min)
Week 1
Week 3
Week 5
Week 7


A
C1
C2
C3
C4
C5
C6
Z score
4
2
0
−2
−4
342 constantly modified transcripts
Group I
Group II
Group III
A
B
C
D
E
F
B
Top 15 enriched GO terms of
modified transcripts from each group
Protein refolding
Meristem maintenance
Response to heat
Protein folding
Chaperone−mediated protein folding
Response to temperature stimulus
Developmental process
Response to abiotic stimulus
Carboxylic acid metabolic process
Oxoacid metabolic process
Organic acid metabolic process
Response to chemical
Small molecule metabolic process
Response to stress
Response to stimulus
Ribonucleoprotein complex assembly
Complex subunit organization
Cytoplasmic translation
Ribonucleoprotein complex biogenesis
Translation
Peptide biosynthetic process
Amide biosynthetic process
Intracellular transport
Peptide metabolic process
Cellular amide metabolic process
Cellular macromolecule biosynthetic process
Compound biosynthetic process
Cellular protein metabolic process
Nitrogen compound biosynthetic process
Gene expression
Carbon fixation
Photosynthesis
Pentose phosphate pathway
Porphyrin metabolism
Pyruvate metabolism
Cysteine and methionine metabolism
Glycolysis/Gluconeogenesis
Carbon metabolism
Purine metabolism
Oxidative phosphorylation
Starch and sucrose metabolism
Biosynthesis of amino acids
Biosynthesis of cofactors
Metabolic pathways
Biosynthesis of secondary metabolites
Counts
10
20
−log10(FDR)
3
6
9
Counts
10
15
20
25
−log10(FDR)
3.5
4.0
4.5
Counts
10
20
30
−log10(FDR)
2
4
6
8
10
Fold Enrichment
C
Transcription
RCM Frequency
Z score
Group I
Group II
Group III
D
Transcription
RCM Density
E
Frequency
Transcription
Density
Transcription
C6
C1
C4
C3
C5
C2


Selection of Core Co-expression Modules
A
Transcriptome characteristics
Module
Counts (High GS)
Trait Correlation (p < 0.05 displayed)
Enriched terms -log10(FDR)
TF counts
LI-6800
Oxidative stress Chemicals
Metabolites
GO (biological process)
KEGG
lightyellow 642 (48)
darkturquoise 1275 (30)
Darkmagenta 1093 (213)
blue 707 (336)
darkgrey 837 (37)
midnightblue 358 (21)
plum1 154 (18)
white 270 (2)
brown 2154 (74)
grey60 387 (15)
ivory 99 (5)
magenta 460 (161)
Salmon 1647 (207)
Motif enriched TFs within modules
Leaf internal CO2 concentration
Maximum PSII efficiency (Fv/Fm)
Intrinsic water use efficiency (WUEi)
CO2 net assimilation
Chamber air temperature
Leaf temperature
Vapor pressure deficit at leaf temperature
Stomatal conductance to water vapor (GSW)
Reference cell H2O concentration
Ascorbate
Polyphenols
Flavonoids
Total Antioxidant Aapacity (TAC)
Hydroxypyruvate Eeductase (HPR)
Ascorbate Oxidase (AO)
Glutaredoxins (Grxs)
Catalase (CAT)
Superoxide dismutase (SOD)
Ascorbate oxidase (APX)
Alpha_ketoglutaric_acid
Glyceric_acid
Chlorogenic_acid_4
Chlorogenic_acid_3
Beta.alanine_1
Galactonic_acid_2
Ferulic_acid
Chlorogenic_acid_2
D.glucose_1
Quinic_acid
Stress
Protein
RNA
Metabolism
Response to abiotic stress
Response to chemical
Response to light intensity
Response to radiation
Protein localization
Protein transport
Protein complex assembly
RNA modification
RNA processing
RNA splicing
Small molecule metabolic process
Lipid metabolic process
Ribonucleoprotein complex biogenesis
Carbohydrate metabolic process
Cellular localization
DNA metabolic process
Biosynthesis of secondary metabolites
Metabolic pathways
Ribosome biogenesis in eukaryotes
Module-trait Correlation coefficient
GO enrichment level: Normalized (-log10-Padj)
B
RCM characteristics
Licor 6800
Oxidative stress Chemicals
Metabolites
GO (biological process)
Ratio of modified transcripts with trait correlation (Z-score scaled)
Ratio of modified transcripts with enrichment (Z-score scaled)
C
Eigenvalue of module
Genotype A B C D E F
Condition Control Drought
Salmon
Darkmagenta
Eigenvalue
Week 1 Week 3 Week 5 Week 7
D
Percent of constantly modified transcripts in module
Group I
Group II
Group III
99%
11%
89%
17%
25%
58%
N = 103
N = 71
N = 167
Other modules
Darkmagenta
Salmon


A
CDF3
Module gene (N = 1093)
CDF3 Target (N = 681)
2 Kb
TSS
5'UTR
3'UTR
Promoter
Core motif
5'-T/AAAAG-3'
B
CDF3 relative abundance
Control
Drought
A
B
C
D
E
F
Week 1
Week 3
Week 5
Week 7
C
CDF3 Correlation abs(R)
p < 2.22e−16
0.0018
N = 412
412
412
CDF3 target
Others
Global
D
log2FC (Drought / Control)
p = 0.01
N = 412
412
CDF3 target
Others
E
Sub network of CDF3-targeted transcriptional regulation (N = 230)
Trait-correlated genes
Flavonoid content
CO2 net assimilation
Total antioxidant capacity
Leaf temperature (surface/chamber)
Co-expression
TF-binding (Cistrome)
Protein-protein interaction (STRING)
Trait correlated genes (abs(r) > 0.7)
Abiotic stress response
QTLs
CDF3
RNA modification
Non-modified
Modified
Trait correlated genes (abs(r) > 0.7)
Metabolic process
F
Key CDF3 targets
Transcript
Edge number
SbPSAL 423
SbPPC2 690
SbRIQ1 614
SbALBINO1 124
SbCP5 322
SbPGM2 147
SbFLAP1 566
SbLPA1 259
Sobic.003G042700.4 388
Sobic.003G387800.1 577
SbCDF3
Z-score
Control
Drought
A B C D E F
CO2 net assimilation
Leaf internal CO2 concentration
Flavonoid & phenol
Total anti-oxidant capacity
PCC
R = 0.65 ~ 0.71
CO2 net assimilation
DW (g) /area (L)
R = -0.73
Flavonoid
mg(QA)/gFW
R = -0.71
Total antioxidant capacity
μmol(Trolox)/gFW
G
RCM and transcript correlation of CDF3 targets
PCC
-0.62 0.77
-0.48 0.57
-0.35 0.42
-0.58 0.60
-0.30 0.81
-0.40 0.36
-0.53 0.74
0.30
-0.62 0.67
0.63
RCM frequency
RCM density


A
Control
Drought
DUS2 relative expression level
Week 1
Week 3
Week 5
Week 7
B
PAM status
(DHU)
N = 1101
PC2 (5.88%)
PC1 (7.18%)
C
119 SNPs
Chr06: 51,665,122 - 51,670,021
Homozygous reference allele
Heterzygous allele
Homozygous alternative allele
SNPs that affect splicing (8 SNPs)
5'
SbDUS2
3'
D
Sub network of salmon module (N = 77)
Stress response
Trait-correlated genes (abs(r) > 0.65)
QTLs
Constantly DHU-modified transcripts
Enriched pathways
Chlorogenic acid
Galatonic acid
Fv'Fm'
CO2 net assimilation
Non-modified
D-modified
Co-expression
TF-binding (Cistrome)
Protein-protein interaction (STRING)
Carbon metabolism
DUS2
E
Z score
Transcript
Trait-PCC
SbHIRD11 0.56
SbCAO 0.62
SbBRCA1 0.55
SbFAB1B 0.43
Sobic.009G207900.1 0.55
SbEMB2728 0.41
SbRPS7.2 0.42
SbAMK2 0.63
SbCRRL (DnaJ) 0.66
SbPPDK 0.67
PCC
DUS2
F
Sorghum time-series drought data
SbCRRL
SbPPDK
Control
Drought
BTX642: tolerant
RTX430: susceptible
Transcript abundance (Z-score)
Growing period (week)
G
Sorghum tissue atlas data
Head
Upper leaf
Mid leaf
Lower leaf
Shoot
Grain
Sheath
Panicle
Whorl
Floral initiation
Growing
Juvenile
Whorl
Whorl
Grain matured
Floral initiation
Growing
Juvenile
Control
Urea
Nitrate
Ammonium
Fertilized
SbPPDK
SbCRRL
SbDUS2
Z score
H
SbCRRL
SbPPDK
R = 0.85, 2e-04
R = 0.93, 8.6e-09
N = 11
N = 18
DHU density (Z-score)
R = -0.7, 0.008
R = -0.58, 0.0094
DHU frequency (Z-score)
Transcript abundance (Z-score)


A
T-DNA insertion
5'
AtDUS2
3'
AT3G49640.1P
DUS-FMN-like domain
348 AA
1
100
200
300
SbDUS2.1
AtDUS2.1
Coverage: 94% Idenitity: 79.81%
B
Col−0
dus2
p = 0.0099
5'6-DHU concentration (nM)
N = 6
N = 6
Col−0
dus2
C
p < 0.05
N = 65
N = 67
Germination rate (%)
Col−0
dus2
Germination time (Hours)
D
Control (22°C)
Col−0
dus2
Heat (37°C)
N = 18
N = 21
N = 21
N = 14
Rosette Area of individual plant (Pixel number: *10 4)
Survival rate (%)
Heat treatment (Days)
p < 0.05
Total time (Minutes)
Col−0
dus2
Control (22°C)
Heat (37°C)
E
Reproductive stage (36 days)
Col-0
dus2
N = 56
N = 52
Bud initiation rate (%)
Col−0
dus2
Days after true leaf emergence
F
Col−0
dus2
p = 0.12
p = 0.00016
Net CO2 assimilation (µmol CO2 m-2 s-1)
N = 10
N = 10
Control
Heat
p = 0.33
p = 0.0024
Stomatal conductance (mol H2O m-2 s-1)
N = 10
N = 10
Control
Heat
G
p = 0.072
p = 0.0042
DHU frequency
N = 10
N = 10
Control
Heat
H
Control
Heat
Col−0
dus2
Day 1
Day 2
Day 3
Day 4
PC2: 17% variance
PC1: 42% variance
I
Day 4 Up-regulated transcripts
Cation transport
Signal transduction
Calcium ion import
17 / 458
12 / 254
3 / 9
response to chemical
defense response
response to stress
Fatty acid
Jasmonic acid
Wounding
30 / 229
30 / 225
31 / 233
Response terms
−log10(FDR)
Day 4 Down-regulated transcripts
Response terms
Salt stress
Acid chemical
Water deprivation
Inorganic substance
Biotic stress
Lignin metabolic process
Cell division
Metabolic process
Nuclear division
Cell cycle
23 / 474
25 / 459
24 / 418
37 / 790
25 / 429
8 / 76
23 / 450
10 / 59
19 / 208
43 / 688
Col−0
dus2
−log10(FDR)