ORIGINAL RESEARCH

American Society of Plant Biologists · SOCIETY FOR EXPERIMENTAL BIOLOGY · WILEY

Conserved and variable heat stress responses of the Heat Shock Factor transcription factor family in maize and *Setaria viridis*

Zachary A. Myers[1] | Clair M. Wootan[1] | Zhikai Liang[1] | Peng Zhou[2] | Julia Engelhorn[3,4] | Thomas Hartwig[3,4] | Springer M. Nathan[1]

[1]Department of Plant and Microbial Biology, University of Minnesota, Minneapolis, MN, USA

[2]Chinese Academy of Agricultural Sciences, Institute of Crop Sciences, Beijing, China

[3]Heinrich-Heine University, Düsseldorf, Germany

[4]Max Planck Institute for Plant Breeding Research, Cologne, Germany

Correspondence
Springer, Nathan M, Department of Plant and Microbial Biology, University of Minnesota, Saint Paul MN 55108, USA.
Email: springer@umn.edu

Abstract

The Heat Shock Factor (HSF) transcription factor family is a central and required component of plant heat stress responses and acquired thermotolerance. The HSF family has dramatically expanded in plant lineages, often including a repertoire of 20 or more genes. Here we assess and compare the composition, heat responsiveness, and chromatin profiles of the HSF families in maize and *Setaria viridis* (*Setaria*), two model C4 panicoid grasses. Both species encode a similar number of HSFs, and examples of both conserved and variable expression responses to a heat stress event were observed between the two species. Chromatin accessibility and genome-wide DNA-binding profiles were generated to assess the chromatin of HSF family members with distinct responses to heat stress. We observed significant variability for both chromatin accessibility and promoter occupancy within similarly regulated sets of HSFs between *Setaria* and maize, as well as between syntenic pairs of maize HSFs retained following its most recent genome duplication event. Additionally, we observed the widespread presence of TF binding at HSF promoters in control conditions, even at HSFs that are only expressed in response to heat stress. TF-binding peaks were typically near putative HSF-binding sites in HSFs upregulated in response to heat stress, but not in stable or not expressed HSFs. These observations collectively support a complex scenario of expansion and subfunctionalization within this transcription factor family and suggest that within-family HSF transcriptional regulation is a conserved, defining feature of the family.

KEYWORDS
abiotic stress, gene regulation, heat stress, maize, *Setaria*, transcription factors

1 | INTRODUCTION

Environmental perception and responsiveness are key features of plant growth and development. Understanding the intricate ways in which such capabilities have evolved and are encoded could provide avenues to develop more resilient crop varieties, particularly considering ongoing warming trends that now regularly expose plants to excessive temperature for which they are often less well adapted. Robust molecular responses to heat stress are important for plant survival and recovery and are controlled in large part by the Heat Shock

Factor (HSF) transcription factor family. HSF families are typically small in animals and fungi, encoded by 1 and 4 genes in *Saccharomyces cerevisiae* and humans, respectively (Nakai et al., 1997; Sorger & Pelham, 1988), but are significantly expanded in plant species, ranging from as few as 16 in *Camellia sinensis* (Liu et al., 2016) to as many as 56 in wheat (Xue et al., 2014).

Many aspects of HSF biology are conserved across eukaryotes, including their DNA-binding unit consisting of a multimeric complex that recognizes inverted repeats of an *NGAAN* sequence (Perisic et al., 1989). Plant HSFs are typified by highly conserved DNA-binding and oligomerization domains, and are subclassified into three families: HSFA, HSFB, and HSFC. These families are assigned by the presence or absence of activator motifs and nuclear import and export signals, as well as the length of key feature linker sequences. Interestingly, the expansion of different HSF families in plant lineages has not been consistent, and several broad trends have been observed, including relative overexpansion of the *HSFC* family in monocots relative to eudicots and the lack of *HSFA9*, *HSFB3*, and *HSFB5* members in many monocots (Guo et al., 2016). These large and diverse families of plant HSFs are best known as mediators of heat shock responses through activation of molecular chaperones known as Heat Shock Proteins (HSPs) (Andrási et al., 2021; Nover et al., 2001), though they are also well-documented to regulate other molecular responses, including salt, drought, and osmotic stress in transgenic Arabidopsis (Bechtold et al., 2013; Ogawa et al., 2007; Yokotani et al., 2008).

Several studies have investigated the evolutionary forces driving plant HSF expansion and diversification, particularly in the terrestrial transition of early plants (Wang et al., 2018; Wu et al., 2022; Yang et al., 2014), but functional comparative analyses between species have remained less thoroughly explored. Work in Arabidopsis and tomato has identified that while some HSF TFs are expressed under normal growth conditions, many are instead induced in response to heat stress (Kotak et al., 2007; Ohama et al., 2017; von Koskull-Döring et al., 2007). The induction of some heat-responsive HSFs is lost in *hsf* mutants, and HSFA1b physically interacts with the *HSFA2* promoter in Arabidopsis, suggesting a complex within-family regulatory scheme (Yoshida et al., 2011). It is unclear to what extent stable and heat-responsive HSF accumulation has been conserved across plant species, and whether these regulatory properties vary by gene content, paralog/ortholog identity, or local chromatin features.

The HSF gene family provides an opportunity to assess the potential relationship between chromatin accessibility and gene expression, especially in the context of dynamic responses to the environment. There are a variety of recent approaches that have been developed to provide genome-wide profiles of chromatin accessibility and TF occupancy (Lu et al., 2018; Schmitz et al., 2022). The application of these approaches in several plant species have generated insights into putative *cis*-regulatory elements in plants and highlighted the role of accessible chromatin (Lu et al., 2017) or TF occupancy (Savadel et al., 2021) as approaches to map regulatory elements. There is less known about whether these patterns of chromatin

accessibility are highly dynamic at genes with strong responses to abiotic stress. Several studies have found examples of genes with changes in chromatin that are associated with expression changes in response to environmental stimuli (Raxwal et al., 2020; Reynoso et al., 2022; Sullivan et al., 2014; Zeng et al., 2019). However, the extent to which changes in chromatin accessibility are associated with changes in gene expression is unclear, especially for examples of rapid responses to a stress as is often observed for transcript responses to heat stress at the HSFs, which can occur within 10 min (Lohmann et al., 2004).

In this study, we utilize two C4 grasses, maize, and *S. viridis* (Setaria), to assess changes in composition and heat-responsive expression of HSF genes. Maize is an agronomically important crop with significant genetic and genomic resources (Hufford et al., 2021; Schnable et al., 2009), while *Setaria* is a more recently developed model system with the benefits of a small genome (Mamidi et al., 2020), transformability (Thielen et al., 2020; Weiss et al., 2020), and growth parameters well suited for greenhouse and growth chamber environments (Huang et al., 2016; Li & Brutnell, 2011). These two monocot species have a largely similar complement of HSF genes with some additional paralogs in maize as a result of a relatively recent whole-genome duplication event. Transcriptome analyses in both species reveal many examples of conserved heat-responsiveness for HSF expression as well as examples of divergent patterns of response to heat stress. The TF occupancy based on MNase-defined cistrome Occupancy Analyses-sequencing (MOA-seq (Savadel et al., 2021)) reveals quite limited changes in accessibility and occupancy in control and heat stress conditions, even at genes with highly upregulated expression in response to heat stress. These findings reveal that many of the highly responsive HSFs have TF occupancy even in control conditions and at very low expression levels, suggesting a poised chromatin state that can be rapidly activated.

2 | RESULTS

In order to facilitate targeted comparative genomic approaches of HSF genes between maize and *Setaria*, we identified and extracted predicted HSF protein sequences from the *S. viridis* A10 genome (Mamidi et al., 2020). In total, we identified 26 putative SvHSFs, including 14 HSFAs, seven HSFBs, and five HSFCs, which subdivide into a similar distribution of subfamilies as other related grasses (Figure 1, S1)(Nagaraju et al., 2015; Wang et al., 2009). Phylogenetic analysis of *Setaria* and previously identified maize HSF TFs (Lin et al., 2011; Zhang et al., 2020) identified conserved subfamilies between the two species but also uncovered a handful of compositional differences (Figure 1a). The HSF genes in maize have been given multiple gene identification numbers based on the multiple reference genome release versions of maize (Hufford et al., 2021; Jiao et al., 2017; Schnable et al., 2009) and have also been assigned different gene names in prior publications (Lin et al., 2011; Tello-Ruiz et al., 2016; Zhang et al., 2020) (Table S1). In this work,
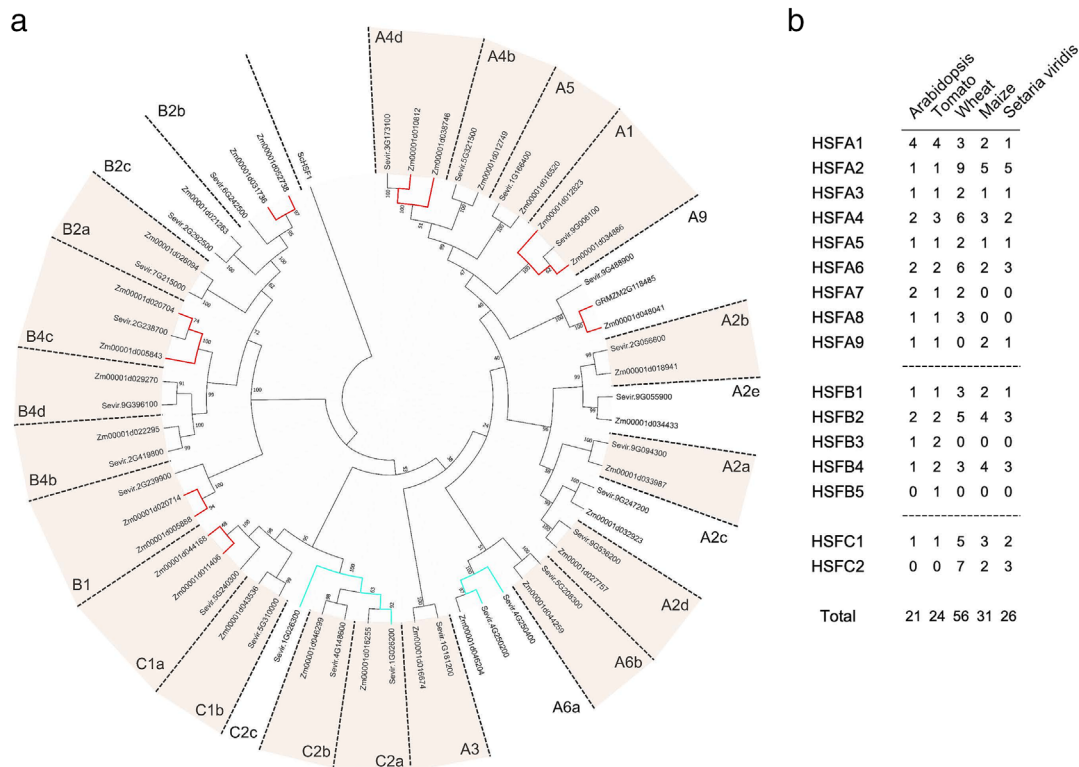
	Arabidopsis	Tomato	Wheat	Maize	Setaria viridis
HSFA1	4	4	3	2	1
HSFA2	1	1	9	5	5
HSFA3	1	1	2	1	1
HSFA4	2	3	6	3	2
HSFA5	1	1	2	1	1
HSFA6	2	2	6	2	3
HSFA7	2	1	2	0	0
HSFA8	1	1	3	0	0
HSFA9	1	1	0	2	1
HSFB1	1	1	3	2	1
HSFB2	2	2	5	4	3
HSFB3	1	2	0	0	0
HSFB4	1	2	3	4	3
HSFB5	0	1	0	0	0
HSFC1	1	1	5	3	2
HSFC2	0	0	7	2	3
Total	21	24	56	31	26

FIGURE 1 Phylogenetic relationship between the HSF families of maize and Setaria. (a) A neighbor-joining consensus tree was constructed to orthologous HSF subfamilies. Subfamilies are indicated on the outer ring and were previously assigned in maize (Zhang et al., 2020). Putative tandem duplicates are connected in cyan and retained maize1:maize2 paralogs are connected in red. Syntenic genes between Setaria and maize are shaded in tan. *Saccharomyces cerevisiae* HSF1 (ScHSF1) was included to root the tree. Numbers at branch points indicate % confidence in consensus assignment over 1000 bootstrap runs. (b). Composition of HSF subfamilies across Arabidopsis (Nover et al., 2001), tomato (Scharf et al., 2012), wheat (Xue et al., 2015), maize (Zhang et al., 2020), and Setaria.

we have utilized the gene names defined in Zhang et al., 2020 and have also provided the B73v4 (Jiao et al., 2017) gene identification numbers (other names and identifications are available in Table S1). All of the HSF subfamilies present in maize have putative orthologs in *Setaria*, with the majority of these identified as syntenic sets (Figure 1b, Table S1, S2). In a number of the HSF subfamilies, there are more maize genes than *Setaria* genes. This is likely due to a relatively recent whole-genome duplication event in maize (Schnable et al., 2011). Indeed, an analysis of genome collinear regions previously used to identify retained duplications present in the maize1 and maize2 subgenomes reveals seven HSF subfamilies that have retained maize1:maize2 paralogs (indicated by red branches in Figure 1a). The paralogs derived from this whole-genome duplication event account for the bulk of the difference in total HSF family size between the two species. We also noted that two *Setaria* subfamilies, SvHSFA6a (*Sevir.4G250200/Sevir.4G250400*) and SvHSFC2 (*Sevir.1G026200/Sevir.1G026300*), have undergone apparent tandem duplication events based on the genomic positions of these genes (indicated by cyan branches in Figure 1a, Figure S2).

2.1 | Gene expression in SvA10 in response to heat stress

Prior work has characterized genes that are upregulated in response to heat stress in maize seedlings using RNA-seq (Zhou et al., 2021). We conducted RNA-seq experiments on control and heat-stressed (1 h, 10°C increase over ambient [30°C]) *Setaria* seedlings, using the same growth stage, stress treatments, and sampling conditions used to generate the maize data (Zhou et al., 2021). We identified 1941 upregulated and 986 downregulated genes in response to a 1-h heat stress in *Setaria* (Figure 2a–b, Table S3). Gene Ontology (GO) enrichment analysis of all upregulated genes identified significant overrepresentation of several heat stress responsive categories (Figure S3, Table S4), suggesting an effective heat stress response. Based on the analysis of a set of 10,003 1:1 orthologs between the two species (Table S5), we found ~fivefold enrichment ($p < 3.46e$-103) for orthologous genes that are upregulated (requiring both a $\log_2$ fold change >1 and an adjusted p-value <.05) in both species in response to heat stress (Table S5). A comparison of the Heat Shock Protein (HSP) 1:1 orthologs in the two species reveals similar patterns
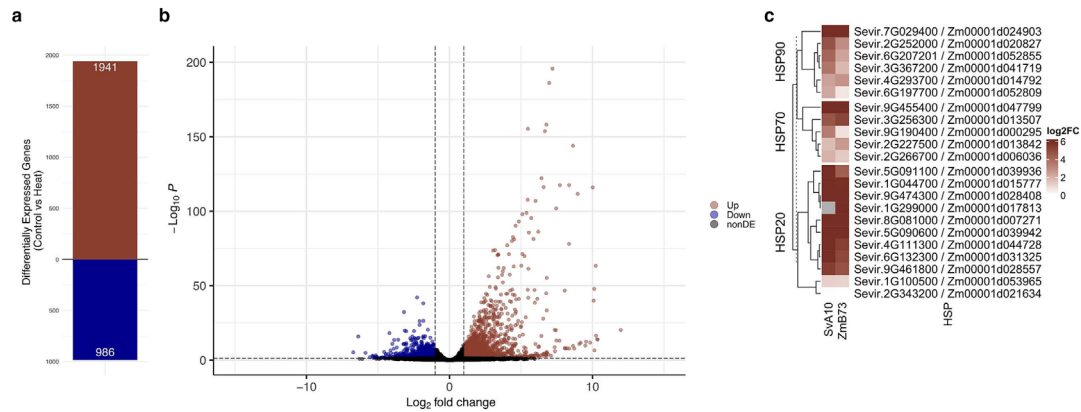
FIGURE 2 Identification of heat-responsive DEGs in *Setaria viridis*. (a) Total number of up- and downregulated DEGs in response to a 1 h, +10°C heat treatment. Cutoffs for DEG classification included an expression threshold of a least one count per million in one sample, a $\log_2$ fold change (heat/control) requirement of >1 (Up) or <-1 (Down), and an adjusted *p*-value <0.05. (b) Volcano plot showing $\log_2$ fold changes and -Log10 *P*-values for every expressed gene. Dashed lines indicate $\log_2$ fold change cutoffs (vertical lines) and *p*-value cutoffs (horizontal line). (c) Heatmap of all 1:1 Heat Shock Protein (HSP) orthologs that are significantly upregulated in either maize or Setaria, broken down by HSP class.

of upregulation in both species (Figure 2c). These results suggest overall similarities in the gene expression responses to heat stress in maize and *Setaria*.

2.2 | Comparison of heat stress responses for *Setaria* and maize HSFs

We examined expression levels of the HSF TFs in order to identify patterns of HSF responsiveness between the two species and differences within individual subfamilies in each species (Figure 3, S4). Members of seven HSF subfamilies (A2a, A2c, A6a, B1, B2b, B2c, and C2a) were significantly upregulated in both maize and *Setaria*, ranging from moderately upregulated in the C2a subfamily to very dramatically upregulated in the A6a subfamily. The most significant differentially expressed HSFs in either species were typically members of HSF subfamilies with conserved heat responsiveness between the two species, such as A2a (7.4–8.2 $\log_2$ fold change upregulated) and A6a (7.5–9.2 $\log_2$ fold change upregulated). Five subfamilies (A2b, A2d, A4b, A4d, and B2a) exhibited *Setaria*-specific responses to heat based on significant expression responses to a 1-h heat stress event, with subfamily B2a responding particularly strongly in *Setaria* (Figure 3). Only subfamily A2e exhibited a significant maize-specific response but was highly expressed in *Setaria* in both control and heat-stressed conditions (Figure 3, S4A). Subfamily A3 was the only subfamily that showed expression in maize but not *Setaria*; however, this gene is very lowly expressed in both species and barely met the 1CPM threshold to be assigned as expressed in maize (Figure S4). Subfamily A4d was the only subfamily that showed *Setaria*-specific expression and was moderately upregulated in *Setaria* in response to heat stress.

From this family-wide comparison, each of the HSF genes was classified into three categories based on expression levels in control and heat samples: (i) "upregulated" (significantly up in heat), (ii) "stable" (expressed in both conditions but not DE), or (iii) "not

expressed" (lower than 1CPM in both conditions) (Figure 3b). In both species, HSFA family members account for the bulk of stably expressed HSFs (80% in both species), while the majority of HSFC family members were not expressed in either control or heat-stressed conditions (60% in both species, Figure 3b). We observed twice as many stably expressed HSFs in maize compared to *Setaria* (10 and five stable HSFs, respectively), and four more upregulated HSFs in *Setaria* compared to maize (13 and nine HSFs, respectively).

2.3 | Chromatin and epigenomic feature comparison of Setaria and maize HSFs

In addition to comparing the compositional and regulatory features of the HSFs in maize and *Setaria*, we also sought to characterize nearby chromatin accessibility and TF binding in order to better understand how chromatin accessibility might be associated with expression responsiveness to heat stress. We generated global scale TF footprinting datasets from *S. viridis* ME034v plants subjected to control (1 h, 30C) and heat-stressed (1 h, +10C) conditions using MOA-seq (Savadel et al., 2021) and leveraged previously published chromatin accessibility (ATAC-seq, Assay for Transposase-Accessible Chromatin-sequencing) and H3K36 trimethylation Chromatin Immunoprecipitation-sequencing levels (H3K36me3 ChIP-seq) in *S. viridis* A10 plants (Lu et al., 2019). While we generated MOA-seq data for both control and heat-stressed tissue samples, the ATAC-seq and H3K36me3 ChIP-seq datasets were previously generated only for plants grown in control conditions.

Peak calling of the MOA-seq analysis identified 97,361 and 102,549 statistically significant MOA-seq peaks across the *Setaria* genome in control and heat-stressed conditions, respectively. MOA-seq experiments and matched RNA-seq experiments were conducted on the highly transformable ME034v genotype of *S. viridis* (in contrast to the A10 reference genotype used in the earlier RNA-seq data). We
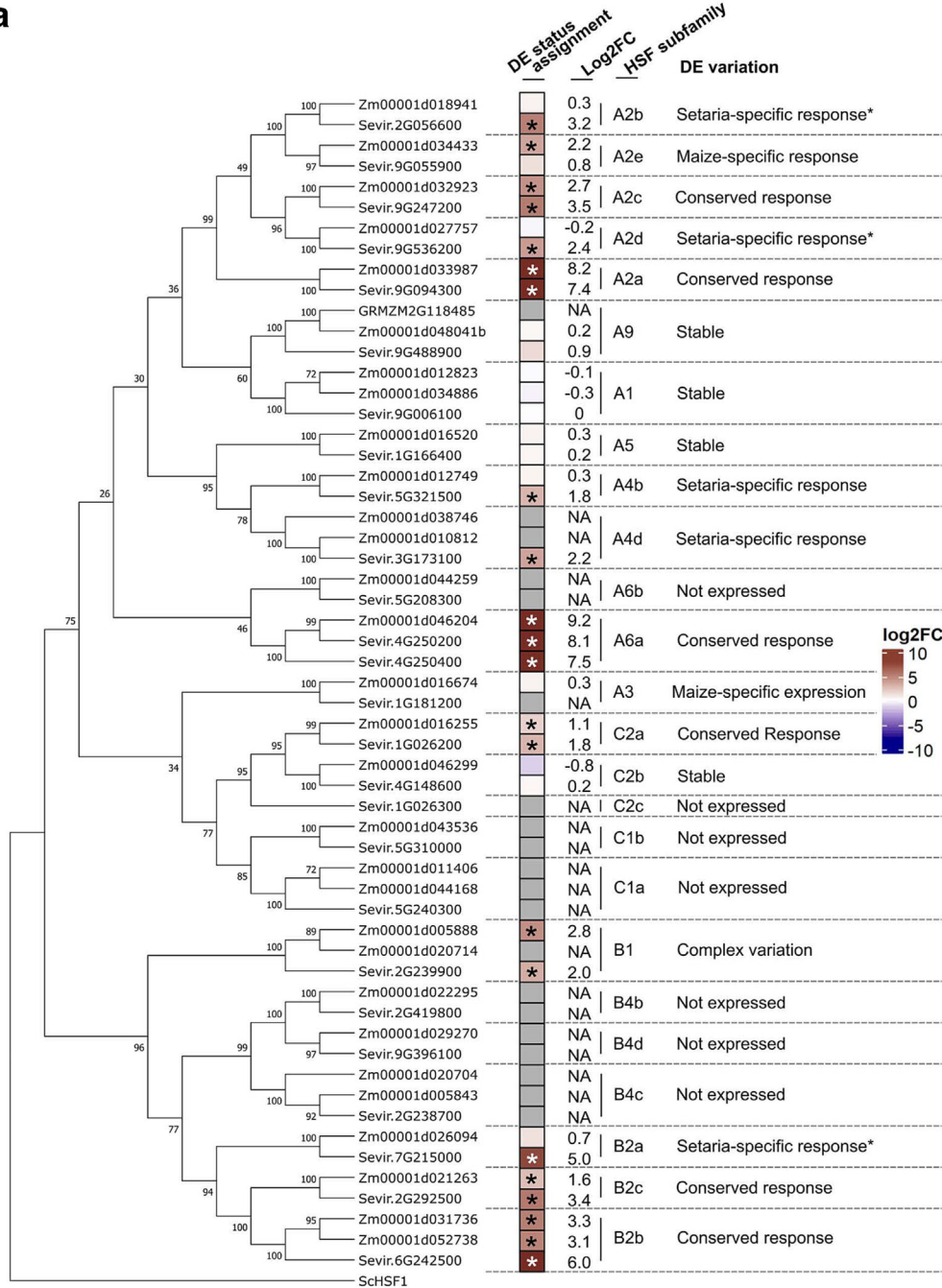
FIGURE 3 Legend on next page.

FIGURE 3 Subfamily-based comparison of HSF heat responsiveness between maize and Setaria. (a) Phylogenetic tree is reformatted but otherwise identical to that presented in Figure 1. Numbers at branch points indicate % confidence in consensus assignment over 1000 bootstrap runs. DE status assignment is presented as a heatmap of log$_2$ fold changes, with not express HSF (<1 CPM) colored gray with actual log$_2$ fold change values are included in the adjacent column. HSF subfamilies are indicated, as well as a brief explanation of the observe variation. Asterisks within the DE variation column indicated three subfamilies identified as Setaria-specific responsive where prior time course evidence suggests a missed heat-responsive pattern (e.g., these maize HSFs may have responded and retuned to prestress levels by the time tissue was collected). (b) Summary counts of HSF families in status assignment column), stable (expressed but not DE in heat, indicated as non-gray, no-asterisk cells), or not expressed (gray cells).
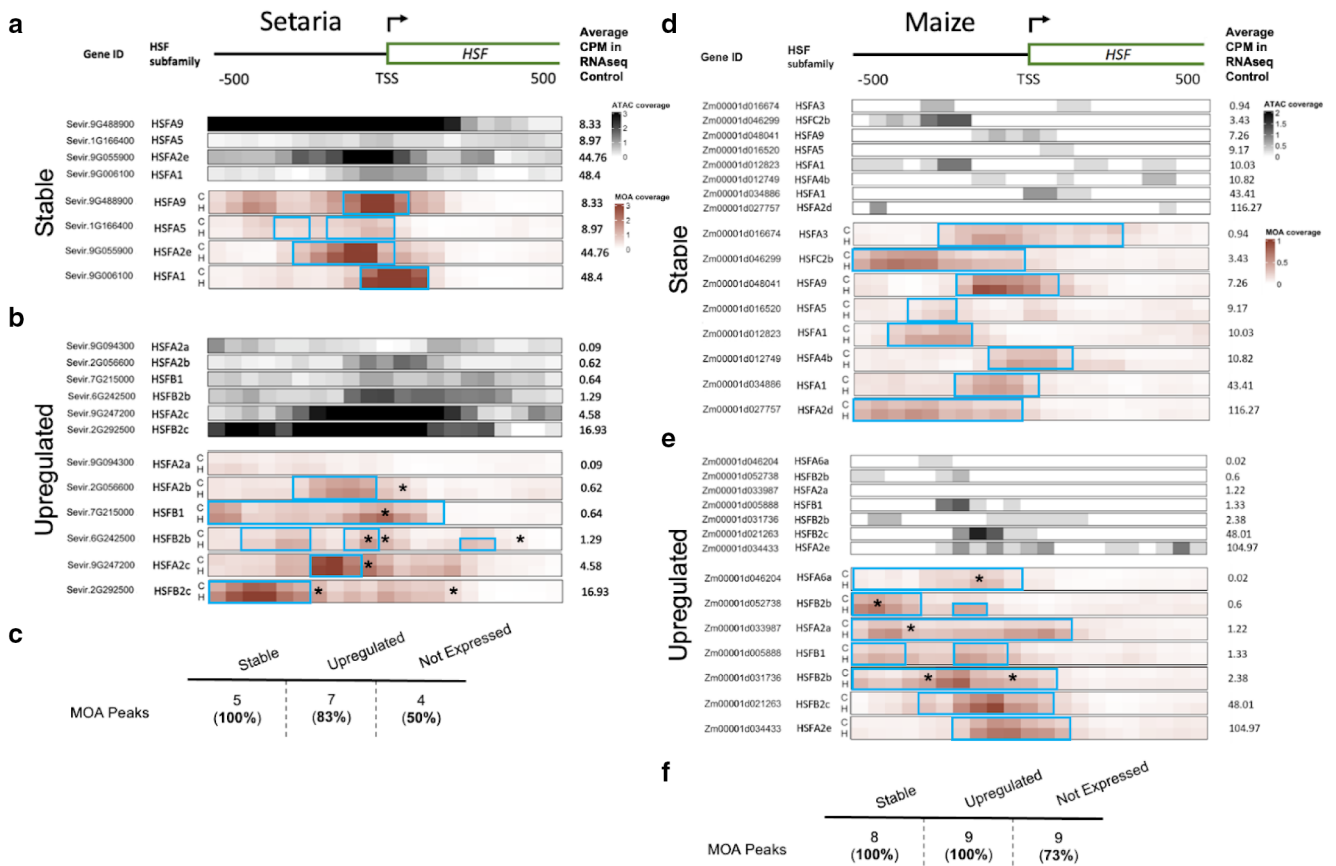

FIGURE 4 Chromatin accessibility analysis of the Setaria and maize HSFs. Chromatin accessibility (ATACseq) and TF binding (MOA-seq) datasets were used to generate a heatmap of coverage over a 1-kb region centered on the TSS of each (a–c) Setaria and (e–g) maize HSF. Expression data were used to facet the family into (a,e) stable, (b,f) upregulated, or (c,g) not expressed (<0.05) and are indicated as smaller boxes in the handful of cases where a condition-specific MOA peak was identified. (d,h) The number of MOA peak calls was summarized by expression category, where the number in parentheses indicates the proportion of genes that contained a significant MOA peak within the analyzed region.

focused on the regions surrounding the transcriptional start sites (TSS) of the *Setaria* HSFs, examining patterns of TF binding, chromatin accessibility, and H3K36 trimethylation levels (Figure 4a–c, S5–6), and excluded a subset of HSFs whose transcriptional response profiles varied between the two RNA-seq experiments (Figure S7). MOA peaks were identified in all stably expressed *Setaria* HSFs (Figure 4a,c), and in most cases were identified in both control and heat-stressed samples (Figure 4a). The lowly expressed HSFC2b/ *Sevir.4G148600* was the sole exception, where a significant MOA peak was only identified in control samples. Significant MOA peaks also overlapped regions of relatively higher ATAC coverage, though there was significant variation in local ATAC coverage across the HSF

family TSS regions. Upregulated *Setaria* HSFs were more variable in MOA peaks presence/absence and ATAC coverage, with ~83% of upregulated HSFs containing a MOA peak in both control and heat-stressed samples (Figure 4b,c). ATAC coverage was also generally higher in *Setaria* upregulated HSFs with higher average control CPM values (Figure 4b). HSFs without evidence for expression in our libraries tended to have dramatically lower coverage levels for both ATAC and MOA, with only 50% of the not expressed *Setaria* HSFs containing a significant MOA-seq peak (Figure S6). Interestingly, these MOA-seq peaks did not overlap with the few instances of deeper ATAC coverage in this expression category, while MOA-seq peaks in other expression categories did tend to overlap with higher ATAC coverage.

MYERS ET AL.

American Society
of Plant Biologists
Cultivating a better future through plant biology research

SEB
SOCIETY FOR EXPERIMENTAL BIOLOGY

WILEY

7 of 14

The presence of significant MOA peaks in control conditions was unexpected, particularly when considering they were observed in HSFs across a wide range of control expression levels, including several HSFs not expressed in control conditions (Figure 4b). Previous work identified cross-regulation within the HSF family, which prompted us to identify putative HSF binding sites within these same 1-kb regions. No putative binding sites were encoded within these regions in stably expressed HSFs, and a single putative binding site was identified over a significant MOA peak in a not expressed HSF (Figure 4a, S6). In contrast, 13 putative HSF binding sites were identified across 10 of the 13 upregulated *Setaria* HSFs, seven of which were within 50 bp of a significant MOA peak (Figure 4b).

A similar set of analyses were performed for the maize HSFs with one notable difference in sampling: The MOA-seq and matched RNA-seq data were generated for leaf tissue sampled after a 4 h $+10°C$ stress treatment rather than a 1-h treatment as in *Setaria*. Peak calling on these libraries identified 188,625 and 195,284 statistically significant MOA-seq peaks across the maize genome in control and heat-stressed conditions, respectively, nearly twice as many as identified in *Setaria*, which is not entirely surprising given the larger genome and higher number of genes in maize, as well as the potential effects of a longer heat stress duration in maize. Focusing on the region around each maize HSF TSS and faceting by expression category, and again excluding the small subset of HSFs whose transcriptional response profile varied between RNA-seq experiments (Figure S8), we saw many of the same trends in chromatin accessibility, TF binding, and putative HSF binding site distribution as observed in *Setaria* (Figure 4d–f, S6). All stably expressed maize HSFs contained a MOA-seq peak (Figure 4d), and unlike the variance seen in *Setaria*, all upregulated maize HSFs also contained a MOA-seq peak (Figure 4e). These MOA-seq peaks also tended to overlap regions with ATAC coverage, though ATAC coverage near HSF TSSs was generally lower and more restricted in maize than in Setaria. Not expressed HSFs were more likely to have a significant MOA peak in maize (73% in maize, 50% in *Setaria*), but coverage was again generally lower in this expression category than in stable or upregulated HSFs (Figure S6). Putative HSF binding sites were primarily identified around upregulated maize HSFs, and in each case was situated within a significant MOA peak (Figure 4e). Two putative HSF binding sites overlapping significant MOA peaks were identified in stably expressed HSFs, whereas the two putative HSF binding sites identified in not expressed HSFs were not near accessible or bound chromatin (Figure 4d, S6). Finally, there were several instances of obviously increased MOA coverage in maize when comparing control and heat-stressed samples (e.g., *Zm00001d048041*/HSFA9 and *Zm00001d052738*/HSFB2b), though in most cases, there were significant MOA-seq peaks in both conditions. One possible explanation for the evidence for altered MOA-seq peaks in maize but not *Setaria* could be the later sampling time in maize compared to *Setaria*.

We also examined H3K36me3 levels near HSF TSSs in both species, finding that stably expressed HSFs in both species were typically enriched for H3K36me3 coverage downstream of their TSS (Figure S5A,D). Upregulated HSFs in either species were more likely to have H3K36me3 coverage when more highly expressed in control conditions (Figure S5B,E), while HSFs with no evidence for expression in our libraries were typically void of H3K36me3 coverage (Figure S5C,F).

2.4 | Comparisons of chromatin properties at paralogs and orthologs/syntelogs

In addition to exploring expression category-wide trends, we also made targeted comparisons of paralogs and orthologs/syntelogs. There are two pairs of *Setaria* tandem duplicates that are present in *Setaria* but not maize. These two pairs of tandemly duplicated genes provide an opportunity to consider the patterns of chromatin and TF binding at duplicated genes. When examining the HSFC2a/HSFC2c pair, which split into different expression categories ("not expressed" and "upregulated"), we observed a clear reduction in both MOA and ATAC coverage in HSFC2c compared to HSFC2a and the presence of a significant MOA-seq peak in HSFC2a but not HSFC2c (Figure 4b–c, S9). Taken alongside the complete lack of expression of HSFC2c in our datasets, it seems that HSFC2c may be in the process of pseudogenization.

The tandem duplicate pair of HSFA6a members were both significantly upregulated in heat stress conditions and had very low or undetectable expression in control conditions (Figure 3a, S4); however, this was only observed in the A10 genotype RNA-seq data, and no expression was detected in either control or heat-stressed conditions in the ME034v genotype. Both genes lack evidence for H3K36me3, as expected due to the lack of expression in control conditions (Figure S5). While the promoters of these two *Setaria* HSFA6a members have nearly completely diverged, we did identify a small stretch of sequence similarity between the two that overlapped both a significant MOA peak and a putative HSF binding sites in each (Figure 4b, S9), despite a lack of evidence for expression. *Sevir.4G250200* had strong ATAC coverage that overlapped a MOA-seq peak, but there was little ATAC-seq coverage at *Sevir.4G250400* (Figure 5b, S9).

Targeted analysis of syntenic sets of maize and Setaria HSFs identified both conserved and divergent patterns of transcriptional heat responsiveness, chromatin accessibility, and TF binding. A comparison of the consistently upregulated HSFB2b subfamily, which is encoded by a paralog pair in maize and a single locus in *Setaria*, identified similar patterns of chromatin accessibility and TF binding (Figure 4b,e). While ATAC signal was relatively low around the two maize HSFB2b TSSs, it was overlapping a significant MOA peak and putative HSF binding site in both paralogs (Figure 4f). In *Setaria* HSFB2b, a similar trend was observed where a region of accessible chromatin partially overlapped with a significant MOA peak and putative HSF binding site (Figure 4b).

The HSFB1 subfamily is also encoded by a retained m1:m2 paralog in maize and a single locus in *Setaria* (Figure 1). This subfamily was assigned to different expression categories, with one upregulated and one not expressed HSFB1 in maize, and an upregulated HSFB1 in

Setaria (Figure 4, S4). Interestingly, MOA signal was similar in all three HSFB1 members, with a significant MOA peak near each TSS; however, the two upregulated HSFB1s have open chromatin regions overlapping their MOA peaks, while the not expressed HSFB1 has lower ATAC coverage over its significant MOA peak (Figure 4b,e). The lack of open and quickly accessible chromatin, such as that seen in the two upregulated HSFB1 members, might contribute to the lack of heat-responsive expression observed in the not expressed maize HSFB1.

3 | DISCUSSION

3.1 | Broad strokes of similarities and differences in HSFs between Setaria and maize

A comparison of the full suite of HSF genes in maize and *Setaria* revealed many similarities and several key differences. We found many of the previously reported trends in maize HSF family composition (Lin et al., 2011; Zhang et al., 2020) to hold true in *Setaria*, including expanded HSFA2 and HSFC subfamilies (relative to dicots) and the absence of the HSFB3 and HSFB5 subfamilies. In addition to expansion of the HSFA2 subfamily, regulation of most HSFA2 members in *Setaria* and maize seemed markedly different than that reported in Arabidopsis and tomato, where eight of the 10 HSFA2 members in *Setaria* and maize were expressed in control conditions (average control CPM >1, Figure 3, S4). One possible explanation for the relative heat tolerance of maize and *Setaria* compared to that of Arabidopsis and tomato could be attributed to expression of some HSFA2 TFs in non-heat stress conditions, providing a basal level of heat tolerance.

Compositional differences between the HSF families of *Setaria* and maize were fully explained by either retained maize1:maize2 paralogs or tandem duplications in *Setaria*. The expression responses of HSFs to heat stress were more varied between the two species. While many HSFs have similar or marginally different responses, there are examples such as HSFA4d and HSFB2a that exhibit heat-responsive expression in *Setaria* alone. Relatively little is known about the function of these HSF subfamilies. The HSFA4d subfamily has been explored somewhat in rice, where overexpression of OsHSFA4d exhibited increased heat tolerance in transgenic rice (Yamanouchi et al., 2002), and overexpression of the related wheat HSFA4a led to improved cadmium tolerance in transgenic rice (Shim et al., 2009). The HSFB2 subfamily has been observed as heat stress induced in pigeon pea (*Cajanus cajan*) and implicated in defense responses against botrytis in cucumber (*Cucumis sativus*) through induction of pathogenesis-related defense proteins in response to heat stress (Kharisma et al., 2022; Ramakrishna et al., 2022). Maize HSFB2a was not differentially expressed in the 1-h heat stress data used here, but a previous study suggested that it was transiently induced at 30-min post heat stress before returning to its prestress expression level (Zhou et al., 2021), suggesting a potentially different temporal response between the two species.

3.2 | HFSA6 subfamily diversification in the grasses

The HSFA6 family has been cursorily examined in wheat, where it was found to be heat induced and able to drive enhanced heat responsiveness when ectopically induced alongside a heat stress (Xue et al., 2015). With this in mind, it is interesting to note that *Setaria* has retained a tandemly duplicated copy of HSFA6a, and there appears to be a genotype-specific expression pattern when comparing expression data between the A10 and ME034v genotypes (Figure S7). While one report has identified low-level expression of the HSFA6a tandem duplicates in ME034v (Thielen et al., 2020), multiple following reports have not identified any expression of either HSFA6a in either control (Read et al., 2022) or heat-stressed samples of ME034v (Figure S7, (Anderson et al., 2021)). Across plant lineages, the HSFA6 family appears to typically be encoded by two or three members (Guo et al., 2016), though there are no identified members in apple or carrot (Giorno et al., 2012; Huang et al., 2015), and as many as six members in wheat (Xue et al., 2014). Despite the significant level of induction in response to heat observed in maize and *Setaria*, we can find no reports of Arabidopsis HSFA6a being heat responsive, either in the literature or through the Arabidopsis eFP browser (Winter et al., 2007). Further, we discovered that the tandem duplication event in *Setaria* appeared to also encompass a putative HSF TFBS approximately 200-bp upstream of each HSFA6a coding sequence, though the sequences surrounding these HSF TFBSs had nearly completely diverged (Figure S6). We identified overlap of significant MOA peaks and putative HSF binding sites upstream of the HSFA6a TSSs in both species (Figure 4e, S9), and the narrow conservation of this site in *Setaria* suggests that these sites might be functionally important; however, it is difficult to reconcile the lack of expression in ME034v with the presence of significant MOA-seq peaks. These sites could represent an instance of within-family regulation of the HSFs that is somehow rendered nonfunctional in the ME034v genotype. The concept of HSFs promoting expression of other HSFs in response to heat stress has been suggested before (Liu et al., 2013; Nishizawa-Yokoi et al., 2011), though the extent to which this relationship occurs within the larger family remains unknown.

3.3 | Use of chromatin properties to predict expression responses in a gene family

The expansion of TF families has provided opportunities for specialization and subfunctionalization of individual family members. Divergent patterns of gene expression for members of gene families can be one mechanism of subfunctionalization. HSFs often exhibit responsiveness to heat stress events, but this is not necessarily observed for all members of the gene family. In many species, researchers have used gene-specific approaches (such as quantitative real-time PCR) or transcriptome profiling following a heat stress to monitor which HSF genes exhibit response to heat stress. We were interested in assessing whether chromatin accessibility or chromatin modifications could be

used to predict the heat-responsive expression for different members of a gene family.

Prior studies (Lu et al., 2019; Ricci et al., 2019) have documented genome-wide patterns for a number of histone modifications and chromatin accessibility in maize and *Setaria* seedling leaf tissue that was not exposed to a heat stress. We examined these features to see whether they might facilitate predicting HSF heat responsiveness. Some chromatin features, such as H3K36 trimethylation, were highly associated with genes that are expressed in ambient conditions, including the stably expressed HSFs as well as upregulated HSFs that are expressed at moderate levels in control conditions. In addition, the majority of the genes that exhibit expression in ambient or heat stress conditions contain accessible regions near the TSS. However, we did not find evidence for specific chromatin features that were found in the HSFs that are heat responsive, suggesting limited potential to use these chromatin properties to predict responsiveness. Chromatin accessibility was often found at many of the upregulated HSFs, even though several of these genes are not expressed in ambient conditions. This may reflect the fact that many of these HSFs are rapidly (<30 min) induced in response to a heat stress event and chromatin accessibility is necessary to allow this rapid activation.

The MOA-seq datasets were generated for both heat and control conditions in maize and *Setaria* to assess potential changes in chromatin accessibility and TF binding at genes with expression that is increased or activated in response to heat stress. In order to align with previous datasets generated in our lab, the MOA-seq data were generated after a 1-h time stress for *Setaria* but a 4-h heat stress in maize. We found quite limited changes in MOA-seq profiles in *Setaria* control and heat-stressed plants despite highly distinct transcript levels for many HSF genes. The genes that are highly upregulated tended to already have MOA-seq peaks in control samples suggesting potential TF binding even prior to activation. We did not note novel peaks near these upregulated genes despite the highly altered transcript abundance. One possible explanation is that these HSFs are being actively repressed, allowing for a quick gene expression change in response to derepression at these positions, similar to the regulation of auxin responsive genes following the derepression of Aux/IAA repressors (reviewed in (Chapman & Estelle, 2009; Teale et al., 2006)). In maize plants subjected to a 4-h heat stress, we did find several examples of increased MOA-seq peaks at some heat-responsive genes. In many cases, the maize genes have a MOA-seq peak in control conditions, but the strength of the MOA-seq signal is increased in the 4-h heat stress sample. This could reflect increased TF binding and occupancy following heat stress and may suggest that these changes require several hours to become fully apparent, suggesting that the chromatin accessibility changes happen after changes in transcription. Prior studies that have compared chromatin accessibility in control and stress conditions have reported similar findings of limited numbers of novel accessible regions but more examples of quantitative shifts in chromatin accessibility (Lee & Bailey-Serres, 2019; Maher et al., 2018; Raxwal et al., 2020; Reynoso et al., 2021).

While our analyses of chromatin properties were focused on the HSF gene family, there are further opportunities to compare heat-responsive expression and chromatin in the full set of heat-responsive genes in maize and *Setaria*. Understanding the potential role of chromatin accessibility and TF footprinting to document *cis*-regulatory elements and predict expression responses will provide a roadmap for understanding the evolution of gene expression responses to abiotic stress events.

4 | MATERIALS AND METHODS

4.1 | Plant materials and growth conditions, sampling conditions

Setaria A10 and ME034v plants were grown in growth chambers at 30/20°C under 12-h light/12-h dark cycles for either 12 (RNA-seq samples) or 20 (MOA-seq and MOA-paired RNA-seq samples) days before harvesting leaves. Seeds were sown directly onto wet soil (PGX Gro-Mix) and bottom watered every other day until just prior to treatment or collection. All heat stress treatments were conducted approximately 3 h after lights were turned on, and plants were moved into either a 40°C incubator or a 30°C incubator for 1 h. After treatment, leaves were collected in paper envelopes and snap-frozen in liquid nitrogen. Leaf 3 of 12-day-old A10 samples used in RNA-seq analyses, while both L3 and L4 were harvested for 20-day-old ME034v MOA-seq samples to increase total biomass collected.

For the maize MOA-seq experiments reported here, maize B73 seedlings were grown for 3 weeks in growth chambers at 30/20°C under 12-h light/12-h dark cycles. Approximately, 3 h after lights were turned on, plants were exposed to either a 40°C heat stress or left at 30°C for 4 h. Leaves 3 and 4 were collected in paper envelopes and snap-frozen in liquid nitrogen, collecting a total of ~10 g of fresh tissue.

4.2 | RNA isolation, library prep, and sequencing, DEG calling

RNA was isolated from snap-frozen and ground leaf tissue using the Qiagen RNeasy kit, following manufacturer's instructions without modification. Sequencing libraries were prepared using the standard TruSeq Stranded mRNA library protocol and sequenced on a NovaSeq S4 flow cell as 150-bp paired-end reads, producing at least 20 million reads for each sample. Both library construction and sequencing were carried out at the University of Minnesota Genomics Center. Sequencing reads were then processed through the nextflow-core RNA-seq pipeline (Ewels et al., 2020) for initial quality control and raw read counting. In short, reads were trimmed using Trim Galore! and aligned to the *S. viridis* reference genome (version 2.1, (Mamidi et al., 2020)) using the variant-aware aligner Hisat2 (Kim et al., 2015). Uniquely aligned reads were then counted per feature by feature-Counts (Liao et al., 2014). Differentially expressed genes were called through the DESeq2 package in R (Love et al., 2014), masking genes for expression by requiring genes to have >1 CPM in at least one

sequenced library. The maize MOA-paired RNA-seq data were handled in the same way but was aligned to the maize B73v5 reference genome (Hufford et al., 2021).

4.3 | Data visualization

Most figures were generated in R v4.1.1 (R Core Team [2021]), making heavy use of the ggplot2 and patchwork packages. Volcano plots were generated through the EnhancedVolcano package (https://github.com/kevinblighe/EnhancedVolcano). Heatmaps were generated through the ComplexHeatmap package (https://jokergoo.github.io/ComplexHeatmap-reference/book/). Figures were composed in Inkscape (https://inkscape.org/) and GIMP (https://www.gimp.org/).

4.4 | MOA-seq data generation and processing

Setaria and maize MOA-seq libraries were constructed upon a previous protocol (Savadel et al., 2021), using NEBNext® Ultra™ II DNA Library Prep with Sample Purification Beads and NEBNext Multiplex Oligos for Illumina Unique Dual Index Primer Pairs. Notably, 1.5 ug of purified MOA DNA was used as input, and five PCR cycles were used to reduce duplicates in the library preparation process. Illumina libraries were converted to BGI's DNBSEQ™ libraries and sequenced on one lane of the DNBSEQ-T7 platform at BGI, and China Indexes of reference genomes were built using STAR (v2.7.9a)(Dobin et al., 2013), and raw MOA-seq data reads were preprocessed using SeqPurge (v2019_09)(Sturm et al., 2016). Paired reads with overlapping regions were merged into single-end reads using bbmerge (version 38.18) (Bushnell et al., 2017). Processed MOA reads were aligned to the relative *Setaria* or maize reference genome using STAR. Alignment fragments with less than 81 bp and MAPQ as 255 were kept for analysis. Format conversion and the calculation of the average mapped fragment length (AMFL) was done using SAMtools (v1.9) (Li et al., 2009). For peak calling MOA bam files were used with MACS3 (v3.0.0a7) using the following parameters: -s and --min-length "AMFL," -max-gap "2x AMFL," -nomodel, -extsize "AMFL," -keep-dup all, -g "effective genome size". For normalization, the effective genome size was calculated per sample using unique-kmers.py (https://github.com/dib-lab/khmer/) with AFML and B73 AGPv4 fasta as inputs.

4.5 | Phylogenetic tree construction, MSAs

Phylogenetic trees were generated through MEGA (Kumar et al., 2018), using full-length amino acid sequences for each primary gene model for both maize and *Setaria* HSFs. *Setaria* HSFs were identified and confirmed through three approaches: (i) extraction of genes assigned to the HSF-type PFAM ID PF00447 in the *S. viridis* A10 v2.1 and *S. viridis* ME034v genome assemblies (Mamidi et al., 2020; Thielen et al., 2020); (ii) comprehensive searching through the HSF-centered HEATSTER database (Scharf et al., 2012); and (iii) querying the S *viridis* A10 v2.1 genome through BLAST with all previously identified maize HSFs. neighbor-joining trees were constructed using 1000 bootstrap iterations to produce the final consensus trees presented here (Saitou & Nei, 1987). Multiple sequence alignments were generated in Geneious Prime (v11.0.11) using the MUSCLE alignment algorithm with default parameters and up to 1000 bootstrap iterations (Edgar, 2004). As presented in the main text, full amino acid sequence MSAs were trimmed down to the DNA-binding domain region and/or oligomerization domain region in some instances.

4.6 | Identification of syntenic genes and paralogs

Syntenic genes and sets were pulled from an extension of a previous analysis (Schnable et al., 2011; Zhang et al., 2017). Identification of all 1:1 orthologs between maize and *Setaria* was performed through OrthoFinder v2.5.2 (Emms & Kelly, 2019), using primary amino acid sequences for the Maize B73v4 assembly (Jiao et al., 2017) and the *Setaria* A10 v2.1 assembly (Mamidi et al., 2020). OrthoFinder was run with default parameters, and 1:1 orthogroups were extracted to construct a final 1:1 ortholog table.

4.7 | Data availability and use of previously published datasets

Previously published ATAC and H3K36 trimethylation ChIP datasets (Lu et al., 2019) were retrieved from the NCBI Sequence Read Archive, accession number GSE128434, through use of the sra toolkit (https://github.com/ncbi/sra-tools). Raw reads were adapter trimmed with Trimmomatic (Bolger et al., 2014), quality controlled with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and aligned to either the Maize B73v4 or *Setaria* A10v2 genome with bowtie2 (Langmead & Salzberg, 2012). Resulting SAM files were converted to BAM format, sorted, and indexed through samtools (Danecek et al., 2021). 50 bp tile signal was calculated through the bamCoverage tool in deepTools v3.5.0 suite (Ramírez et al., 2016), using bedtools closest (Quinlan & Hall, 2010) to extract the tiles corresponding to the 1-kb region centered on each HSF TSS and TTS. MOA-seq data processing was performed as described above before applying the same bamCoverage and bedtools approach to extract tile-based MOA signal.

Setaria RNA-seq and MOA-seq datasets have been deposited at the NCBI BioProject and Short Read Archive, accession number PRJNA852218. Maize RNA-seq was retrieved from PRJNA747925, and maize MOA-seq is available from PRJNA849202. RNA-seq and MOA-seq datasets were analyzed as described above.

AUTHOR CONTRIBUTIONS

Zachary Myers and Nathan M. Springer conceived and designed the experiments. Zachary Myers and Peng Zhou collected, processed, and analyzed the RNA-sequencing data. Julia Engelhorn and Thomas

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Sequence data from this article can be found in the EMBL/GenBank data libraries under accession numbers provided in Supporting Information Table 1.

ORCID

Zachary A. Myers https://orcid.org/0000-0002-7225-0290
Clair M. Wootan https://orcid.org/0000-0002-8792-514X
Zhikai Liang https://orcid.org/0000-0002-9963-8631
Peng Zhou https://orcid.org/0000-0001-5684-2256
Julia Engelhorn https://orcid.org/0000-0002-6680-3012
Thomas Hartwig https://orcid.org/0000-0002-2707-2771
Springer M. Nathan https://orcid.org/0000-0002-7301-4759

PEER REVIEW

The peer review history of this article is available in the Supporting Information for this article.

REFERENCES

Anderson, C. M., Mattoon, E. M., Zhang, N., Becker, E., McHargue, W., Yang, J., Patel, D., Dautermann, O., McAdam, S. A. M., Tarin, T., Pathak, S., Avenson, T. J., Berry, J., Braud, M., Niyogi, K. K., Wilson, M., Nusinow, D. A., Vargas, R., Czymmek, K. J., ... Zhang, R. (2021). High light and temperature reduce photosynthetic efficiency through different mechanisms in the C4 model *Setaria viridis*. *Communications Biology*, 4, 1092. https://doi.org/10.1038/s42003-021-02576-2

Andrási, N., Pettkó-Szandtner, A., & Szabados, L. (2021). Diversity of plant heat shock factors: Regulation, interactions, and functions. *Journal of Experimental Botany*, 72, 1558–1575. https://doi.org/10.1093/jxb/eraa576

Bechtold, U., Albihlal, W. S., Lawson, T., Fryer, M. J., Sparrow, P. A. C., Richard, F., Persad, R., Bowden, L., Hickman, R., Martin, C., Beynon, J. L., Buchanan-Wollaston, V., Baker, N. R., Morison, J. I. L., Schöffl, F., Ott, S., & Mullineaux, P. M. (2013). Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b overexpression enhances water productivity, resistance to drought, and infection. *Journal of Experimental Botany*, 64, 3467–3481. https://doi.org/10.1093/jxb/ert185

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge—Accurate paired shotgun read merging via overlap. *PLoS ONE*, 12, e0185056. https://doi.org/10.1371/journal.pone.0185056

Chapman, E. J., & Estelle, M. (2009). Mechanism of auxin-regulated gene expression in plants. *Annual Review of Genetics*, 43, 265–285. https://doi.org/10.1146/annurev-genet-102108-134148

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, 10, 1–4. https://doi.org/10.1093/gigascience/giab008

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

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20, 238. https://doi.org/10.1186/s13059-019-1832-y

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, 38, 276–278. https://doi.org/10.1038/s41587-020-0439-x

Giorno, F., Guerriero, G., Baric, S., & Mariani, C. (2012). Heat shock transcriptional factors in Malus domestica: Identification, classification and expression analysis. *BMC Genomics*, 13, 639. https://doi.org/10.1186/1471-2164-13-639

Guo, M., Liu, J.-H., Ma, X., Luo, D.-X., Gong, Z.-H., & Lu, M.-H. (2016). The plant heat stress transcription factors (HSFs): Structure, regulation, and function in response to abiotic stresses. *Frontiers in Plant Science*, 7, 114. https://doi.org/10.3389/fpls.2016.00114

Huang, P., Shyu, C., Coelho, C. P., Cao, Y., & Brutnell, T. P. (2016). *Setaria viridis* as a model system to advance millet genetics and genomics. *Frontiers in Plant Science*, 7, 1781. https://doi.org/10.3389/fpls.2016.01781

Huang, Y., Li, M.-Y., Wang, F., Xu, Z.-S., Huang, W., Wang, G.-L., Ma, J., & Xiong, A.-S. (2015). Heat shock factors in carrot: Genome-wide identification, classification, and expression profiles response to abiotic stress. *Molecular Biology Reports*, 42, 893–905. https://doi.org/10.1007/s11033-014-3826-x

Hufford, M. B., Seetharam, A. S., Woodhouse, M. R., Chougule, K. M., Ou, S., Liu, J., Ricci, W. A., Guo, T., Olson, A., Qiu, Y., Della Coletta, R., Tittes, S., Hudson, A. I., Marand, A. P., Wei, S., Lu, Z., Wang, B., Tello-Ruiz, M. K., Piri, R. D., ... Dawe, R. K. (2021). De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science*, 373, 655–662. https://doi.org/10.1126/science.abg5289

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., Campbell, M. S., Stein, J. C., Wei, X., Chin, C. S., Guill, K., Regulski, M., Kumari, S., Olson, A., Gent, J., Schneider, K. L., Wolfgruber, T. K., May, M. R., Springer, N. M., ... Ware, D. (2017). Improved maize reference genome with single-molecule technologies. *Nature*, 546, 524–527. https://doi.org/10.1038/nature22971

Kharisma, A. D., Arofatullah, N. A., Yamane, K., Tanabata, S., & Sato, T. (2022). Regulation of defense responses via heat shock transcription factors in Cucumis sativus L. against *Botrytis cinerea*. *Journal of General Plant Pathology*, 88, 17–28.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12, 357–360. https://doi.org/10.1038/nmeth.3317

Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., & Scharf, K.-D. (2007). Complexity of the heat stress response in

plants. *Current Opinion in Plant Biology*, *10*, 310–316. https://doi.org/10.1016/j.pbi.2007.04.011

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549. https://doi.org/10.1093/molbev/msy096

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Lee, T. A., & Bailey-Serres, J. (2019). Integrative analysis from the epigenome to translatome uncovers patterns of dominant nuclear regulation during transient stress. *Plant Cell*, *31*, 2573–2595. https://doi.org/10.1105/tpc.19.00463

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, P., & Brutnell, T. P. (2011). *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. *Journal of Experimental Botany*, *62*, 3031–3037. https://doi.org/10.1093/jxb/err096

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Lin, Y.-X., Jiang, H.-Y., Chu, Z.-X., Tang, X.-L., Zhu, S.-W., & Cheng, B.-J. (2011). Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. *BMC Genomics*, *12*, 76. https://doi.org/10.1186/1471-2164-12-76

Liu, J., Sun, N., Liu, M., Liu, J., Du, B., Wang, X., & Qi, X. (2013). An autoregulatory loop controlling Arabidopsis HsfA2 expression: Role of heat shock-induced alternative splicing. *Plant Physiology*, *162*, 512–521. https://doi.org/10.1104/pp.112.205864

Liu, Z.-W., Wu, Z.-J., Li, X.-H., Huang, Y., Li, H., Wang, Y.-X., & Zhuang, J. (2016). Identification, classification, and expression profiles of heat shock transcription factors in tea plant (*Camellia sinensis*) under temperature stress. *Gene*, *576*, 52–59. https://doi.org/10.1016/j.gene.2015.09.076

Lohmann, C., Eggers-Schumacher, G., Wunderlich, M., & Schöffl, F. (2004). Two different heat shock transcription factors regulate immediate early expression of stress genes in Arabidopsis. *Molecular Genetics and Genomics*, *271*, 11–21. https://doi.org/10.1007/s00438-003-0954-8

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, Z., Hofmeister, B. T., Vollmers, C., DuBois, R. M., & Schmitz, R. J. (2017). Combining ATAC-seq with nuclei sorting for discovery of *cis*-regulatory regions in plant genomes. *Nucleic Acids Research*, *45*, e41. https://doi.org/10.1093/nar/gkw1179

Lu, Z., Marand, A. P., Ricci, W. A., Ethridge, C. L., Zhang, X., & Schmitz, R. J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nature Plants*, *5*, 1250–1259. https://doi.org/10.1038/s41477-019-0548-z

Lu, Z., Ricci, W. A., Schmitz, R. J., & Zhang, X. (2018). Identification of *cis*-regulatory elements by chromatin structure. *Current Opinion in Plant Biology*, *42*, 90–94. https://doi.org/10.1016/j.pbi.2018.04.004

Maher, K. A., Bajic, M., Kajala, K., Reynoso, M., Pauluzzi, G., West, D. A., Zumstein, K., Woodhouse, M., Bubb, K., Dorrity, M. W., Queitsch, C., Bailey-Serres, J., Sinha, N., Brady, S. M., & Deal, R. B. (2018). Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. *Plant Cell*, *30*, 15–36. https://doi.org/10.1105/tpc.17.00581

Mamidi, S., Healey, A., Huang, P., Grimwood, J., Jenkins, J., Barry, K., Sreedasyam, A., Shu, S., Lovell, J. T., Feldman, M., Wu, J., Yu, Y., Chen, C., Johnson, J., Sakakibara, H., Kiba, T., Sakurai, T., Tavares, R., Nusinow, D. A., … Kellogg, E. A. (2020). A genome resource for green millet *Setaria viridis* enables discovery of agronomically valuable loci. *Nature Biotechnology*, *38*, 1203–1210. https://doi.org/10.1038/s41587-020-0681-2

Nagaraju, M., Sudhakar Reddy, P., Anil Kumar, S., K. Srivastava, R., Kavi Kishor, P. B., & Rao, D. M. (2015). Genome-wide scanning and characterization of Sorghum bicolor L. heat shock transcription factors. *Current Genomics*, *16*, 279–291. https://doi.org/10.2174/1389202916666150313230812

Nakai, A., Tanabe, M., Kawazoe, Y., Inazawa, J., Morimoto, R. I., & Nagata, K. (1997). HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. *Molecular and Cellular Biology*, *17*, 469–481. https://doi.org/10.1128/mcb.17.1.469

Nishizawa-Yokoi, A., Nosaka, R., Hayashi, H., Tainaka, H., Maruta, T., Tamoi, M., Ikeda, M., Ohme-Takagi, M., Yoshimura, K., Yabuta, Y., & Shigeoka, S. (2011). HsfA1d and HsfA1e involved in the transcriptional regulation of HsfA2 function as key regulators for the Hsf signaling network in response to environmental stress. *Plant and Cell Physiology*, *52*, 933–945. https://doi.org/10.1093/pcp/pcr045

Nover, L., Bharti, K., Döring, P., Mishra, S. K., Ganguli, A., & Scharf, K. D. (2001). Arabidopsis and the heat stress transcription factor world: How many heat stress transcription factors do we need? *Cell Stress & Chaperones*, *6*, 177–189. https://doi.org/10.1379/1466-1268(2001)006<0177:AATHST>2.0.CO;2

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

Ohama, N., Sato, H., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2017). Transcriptional regulatory network of plant heat stress response. *Trends in Plant Science*, *22*, 53–65. https://doi.org/10.1016/j.tplants.2016.08.015

Perisic, O., Xiao, H., & Lis, J. T. (1989). Stable binding of Drosophila heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell*, *59*, 797–806. https://doi.org/10.1016/0092-8674(89)90603-X

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ramakrishna, G., Singh, A., Kaur, P., Yadav, S. S., Sharma, S., & Gaikwad, K. (2022). Genome wide identification and characterization of small heat shock protein gene family in pigeon pea and their expression profiling during abiotic stress conditions. *International Journal of Biological Macromolecules*, *197*, 88–102. https://doi.org/10.1016/j.ijbiomac.2021.12.016

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, *44*, W160–W165. https://doi.org/10.1093/nar/gkw257

Raxwal, V. K., Ghosh, S., Singh, S., Katiyar-Agarwal, S., Goel, S., Jagannath, A., Kumar, A., Scaria, V., & Agarwal, M. (2020). Abiotic stress-mediated modulation of the chromatin landscape in Arabidopsis thaliana. *Journal of Experimental Botany*, *71*, 5280–5293. https://doi.org/10.1093/jxb/eraa286

Read, A., Weiss, T., Crisp, P. A., Liang, Z., Noshay, J., Menard, C. C., Wang, C., Song, M., Hirsch, C. N., Springer, N. M., & Zhang, F. (2022). Genome-wide loss of CHH methylation with limited transcriptome changes in *Setaria viridis* DOMAINS REARRANGED METHYLTRANSFERASE (DRM) mutants. *The Plant Journal*, *111*, 103–116. https://doi.org/10.1111/tpj.15781

Reynoso, M., Borowsky, A., Pauluzzi, G., Elaine Yeung, Jianhai Zhang, Elide Formentin, Joel Velasco, Sean Cabanlit, Christine Duvenjian, Matthew J. Prior, Garo Z. Akmakjian, Roger B. Deal, Neelima R. Sinha, Siobhan M. Brady, Thomas Girke, Julia Bailey-Serres (2021) *SSRN Journal*. Cell Press. Available at: https://papers.ssrn.com/abstract=3908788 [Accessed May 5, 2022], https://doi.org/10.2139/ssrn.3908788

Reynoso, M. A., Borowsky, A. T., Pauluzzi, G. C., Yeung, E., Zhang, J., Formentin, E., Velasco, J., Cabanlit, S., Duvenjian, C., Prior, M. J., Akmakjian, G. Z., Deal, R. B., Sinha, N. R., Brady, S. M., Girke, T., & Bailey-Serres, J. (2022). Gene regulatory networks shape developmental plasticity of root cell types under water extremes in rice. *Developmental Cell*, *57*, 1177–1192.e6. https://doi.org/10.1016/j.devcel.2022.04.013

Ricci, W. A., Lu, Z., Ji, L., Marand, A. P., Ethridge, C. L., Murphy, N. G., Noshay, J. M., Galli, M., Mejía-Guerra, M. K., Colomé-Tatché, M., Johannes, F., Rowley, M. J., Corces, V. G., Zhai, J., Scanlon, M. J., Buckler, E. S., Gallavotti, A., Springer, N. M., Schmitz, R. J., & Zhang, X. (2019). Widespread long-range *cis*-regulatory elements in the maize genome. In *Nat plants* (Vol. 5) (pp. 1237–1249). Nature Publishing Group. https://doi.org/10.1038/s41477-019-0547-0

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*, 406–425.

Savadel, S. D., Hartwig, T., Turpin, Z. M., Vera, D. L., Lung, P. Y., Sui, X., Blank, M., Frommer, W. B., Dennis, J. H., Zhang, J., & Bass, H. W. (2021). The native cistrome and sequence motif families of the maize ear. *PLoS Genetics*, *17*, e1009689. https://doi.org/10.1371/journal.pgen.1009689

Scharf, K.-D., Berberich, T., Ebersberger, I., & Nover, L. (2012). The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochimica et Biophysica Acta*, *1819*, 104–119. https://doi.org/10.1016/j.bbagrm.2011.10.002

Schmitz, R. J., Grotewold, E., & Stam, M. (2022). *Cis*-regulatory sequences in plants: Their importance, discovery, and future challenges. *Plant Cell*, *34*, 718–741. https://doi.org/10.1093/plcell/koab281

Schnable, J. C., Springer, N. M., & Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 4069–4074. https://doi.org/10.1073/pnas.1101368108

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., … Wilson, R. K. (2009). The B73 maize genome: Complexity, diversity, and dynamics. *Science*, *326*, 1112–1115. https://doi.org/10.1126/science.1178534

Shim, D., Hwang, J.-U., Lee, J., Lee, S., Choi, Y., An, G., Martinoia, E., & Lee, Y. (2009). Orthologs of the class A4 heat shock transcription factor HsfA4a confer cadmium tolerance in wheat and rice. *Plant Cell*, *21*, 4031–4043. https://doi.org/10.1105/tpc.109.066902

Sorger, P. K., & Pelham, H. R. (1988). Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. *Cell*, *54*, 855–864. https://doi.org/10.1016/S0092-8674(88)91219-6

Sturm, M., Schroeder, C., & Bauer, P. (2016). SeqPurge: Highly-sensitive adapter trimming for paired-end NGS data. *BMC Bioinformatics*, *17*, 208. https://doi.org/10.1186/s12859-016-1069-7

Sullivan, A. M., Arsovski, A. A., Lempe, J., Bubb, K. L., Weirauch, M. T., Sabo, P. J., Sandstrom, R., Thurman, R. E., Neph, S., Reynolds, A. P., Stergachis, A. B., Vernot, B., Johnson, A. K., Haugen, E., Sullivan, S. T., Thompson, A., Neri, F. V. III, Weaver, M., Diegel, M., … Stamatoyannopoulos, J. A. (2014). Mapping and dynamics of regulatory DNA and transcription factor networks in A. thaliana. *Cell Reports*, *8*, 2015–2030. https://doi.org/10.1016/j.celrep.2014.08.019

Teale, W. D., Paponov, I. A., & Palme, K. (2006). Auxin in action: Signalling, transport and the control of plant growth and development. *Nature Reviews. Molecular Cell Biology*, *7*, 847–859. https://doi.org/10.1038/nrm2020

Tello-Ruiz, M. K., Stein, J., Wei, S., Preece, J., Olson, A., Naithani, S., Amarasinghe, V., Dharmawardhana, P., Jiao, Y., Mulvaney, J., Kumari, S., Chougule, K., Elser, J., Wang, B., Thomason, J., Bolser, D. M., Kerhornou, A., Walts, B., Fonseca, N. A., … Ware, D. (2016). Gramene 2016: Comparative plant genomics and pathway resources. *Nucleic Acids Research*, *44*, D1133–D1140. https://doi.org/10.1093/nar/gkv1179

Thielen, P. M., Pendleton, A. L., Player, R. A., Bowden, K. V., Lawton, T. J., & Wisecaver, J. H. (2020). Reference genome for the highly transformable *Setaria viridis* ME034V. *G three*, *10*, 3467–3478. https://doi.org/10.1534/g3.120.401345

von Koskull-Döring, P., von Koskull-Döring, P., Scharf, K.-D., & Nover, L. (2007). The diversity of plant heat stress transcription factors. *Trends in Plant Science*, *12*, 452–457. https://doi.org/10.1016/j.tplants.2007.08.014

Wang, C., Zhang, Q., & Shou, H.-X. (2009). Identification and expression analysis of OsHsfs in rice. *Journal of Zhejiang University. Science. B*, *10*, 291–300. https://doi.org/10.1631/jzus.B0820190

Wang, X., Shi, X., Chen, S., Ma, C., & Xu, S. (2018). Evolutionary origin, gradual accumulation and functional divergence of heat shock factor gene family with plant evolution. *Frontiers in Plant Science*, *9*, 71. https://doi.org/10.3389/fpls.2018.00071

Weiss, T., Wang, C., Kang, X., Zhao, H., Gamo, M. E., Starker, C. G., Crisp, P. A., Zhou, P., Springer, N. M., Voytas, D. F., & Zhang, F. (2020). Optimization of multiplexed CRISPR/Cas9 system for highly efficient genome editing in *Setaria viridis*. *Plant Journal*, *104*, 828–838. https://doi.org/10.1111/tpj.14949

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, *2*, e718. https://doi.org/10.1371/journal.pone.0000718

Wu, T.-Y., Hoh, K. L., Boonyaves, K., Krishnamoorthi, S., & Urano, D. (2022). Diversification of heat shock transcription factors expanded thermal stress responses during early plant evolution. *Plant Cell*, *34*, 3557–3576. https://doi.org/10.1093/plcell/koac204

Xue, G.-P., Drenth, J., & McIntyre, C. L. (2015). TaHsfA6f is a transcriptional activator that regulates a suite of heat stress protection genes in wheat (Triticum aestivum L.) including previously unknown Hsf targets. *Journal of Experimental Botany*, *66*, 1025–1039. https://doi.org/10.1093/jxb/eru462

Xue, G.-P., Sadat, S., Drenth, J., & Lynne McIntyre, C. (2014). The heat shock factor family from *Triticum aestivum* in response to heat and other major abiotic stresses and their role in regulation of heat shock protein genes. *Journal of Experimental Botany*, *65*, 539–557. https://doi.org/10.1093/jxb/ert399

Yamanouchi, U., Yano, M., Lin, H., Ashikari, M., & Yamada, K. (2002). A rice spotted leaf gene, *Spl7*, encodes a heat stress transcription factor protein. *Proceedings of the National Academy of Sciences*, *99*, 7530–7535. https://doi.org/10.1073/pnas.112209199

Yang, Z., Wang, Y., Gao, Y., Zhou, Y., Zhang, E., Hu, Y., Yuan, Y., Liang, G., & Xu, C. (2014). Adaptive evolution and divergent expression of heat stress transcription factors in grasses. *BMC Evolutionary Biology*, *14*, 147. https://doi.org/10.1186/1471-2148-14-147

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

Yoshida, T., Ohama, N., Nakajima, J., Kidokoro, S., Mizoi, J., Nakashima, K., Maruyama, K., Kim, J. M., Seki, M., Todaka, D., Osakabe, Y., Sakuma, Y., Schöffl, F., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2011). Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Molecular Genetics and Genomics*, *286*, 321–332. https://doi.org/10.1007/s00438-011-0647-7

Zeng, Z., Zhang, W., Marand, A. P., Zhu, B., Buell, C. R., & Jiang, J. (2019). Cold stress induces enhanced chromatin accessibility and bivalent histone modifications H3K4me3 and H3K27me3 of active genes in potato. *Genome Biology*, *20*, 123. https://doi.org/10.1186/s13059-019-1731-2

Zhang, H., Li, G., Fu, C., Duan, S., Hu, D., & Guo, X. (2020). Genome-wide identification, transcriptome analysis and alternative splicing events of Hsf family genes in maize. *Scientific Reports*, *10*, 8073. https://doi.org/10.1038/s41598-020-65068-z

Zhang, Y., Ngu, D. W., Carvalho, D., Liang, Z., Qiu, Y., Roston, R. L., & Schnable, J. C. (2017). Differentially regulated orthologs in sorghum and the subgenomes of maize. *Plant Cell*, *29*, 1938–1951. https://doi.org/10.1105/tpc.17.00354

Zhou, P., Enders, T. A., Myers, Z. A., Magnusson, E., Crisp, P. A., Noshay, J. M., Gomez-Cano, F., Liang, Z., Grotewold, E., Greenham, K., & Springer, N. M. (2021). Prediction of conserved and variable heat and cold stress response in maize using *cis*-regulatory information. *The Plant Cell*, *34*, 514–534. https://doi.org/10.1093/plcell/koab267/6420714

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.