Diversification of heat shock transcription factors expanded thermal stress responses during early plant evolution

Ting-Ying Wu,1,* Kar Ling Hoh,1,2 Kulaporn Boonyaves,1,3 Shalini Krishnamoorthi1 and Daisuke Urano1,2,3

1 Temasek Life Sciences Laboratory, 1 Research Link, 117604, Singapore
2 Department of Biological Sciences, National University of Singapore, 117558, Singapore
3 Singapore-MIT Alliance for Research and Technology, Singapore

*Author for correspondence: tingying@tll.org.sg, lbnest21@gmail.com

T.-Y.W. and D.U. conceived the whole project. T.-Y.W. designed the experiments. D.U. supervised the project. T.-Y.W. and S.K. generated genetics mutants and performed basic characterization. T.-Y.W. and K.L.H. performed morphological characterization and H.S. phenotyping. T.-Y.W. collected RNA-seq samples. T.-Y.W. and K.B. performed metabolite extraction. T.-Y.W. and D.U. performed phylogenetic analysis. T.-Y.W. analyzed all data acquired with the input of KLH (image analysis) and K.B. (GC-MS). T.-Y.W. wrote the article with the input of all authors. D.U. and S.K. edited the article.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) is Ting-Ying Wu (tingying@tll.org.sg or lbnest21@gmail.com).

Abstract

The copy numbers of many plant transcription factor (TF) genes substantially increased during terrestrialization. This allowed TFs to acquire new specificities and thus create gene regulatory networks (GRNs) with new biological functions to help plants adapt to terrestrial environments. Through characterizing heat shock factor (HSF) genes MpHSFA1 and MpHSFB1 in the liverwort Marchantia polymorpha, we explored how heat-responsive GRNs widened their functions in M. polymorpha and Arabidopsis thaliana. An interspecies comparison of heat-induced transcriptomes and the evolutionary rates of HSFs demonstrated the emergence and subsequent rapid evolution of HSFB prior to terrestrialization. Transcriptome and metabolome analyses of M. polymorpha HSF-null mutants revealed that MpHSFA1 controls canonical heat responses such as thermotolerance and metabolic changes. MpHSFB1 also plays essential roles in heat responses, as well as regulating developmental processes including meristem branching and antheridiophore formation. Analysis of cis-regulatory elements revealed development- and stress-related TFs that function directly or indirectly downstream of HSFB. Male gametophytes of M. polymorpha showed higher levels of thermotolerance than female gametophytes, which could be explained by different expression levels of MpHSFA1U and MpHSFA1V on sex chromosome. We propose that the diversification of HSFs is linked to the expansion of HS responses, which enabled coordinated multicellular reactions in land plants.

Introduction

Simple gene regulatory mechanisms in ancestral organisms have expanded to higher complexity as species evolved. Such advancement in gene regulatory systems often occurs in a saltatory (discontinuous) manner, allowing for the rapid acquisition of new biological roles. Transcription factors (TFs) are crucial for the transcriptional regulation of numerous genes. Therefore, the numbers and combinations of TF
genes expressed in cells contribute to the diversity of gene regulatory networks (GRNs) (Chalancon et al., 2012; Marshall-Colón and Kliebenstein, 2019). cis-regulatory elements determine the architectures and functions of GRNs, which gradually evolved during early plant life (Erwin and Davidson, 2009; Stergachis et al., 2014; Wu et al., 2021).

Heat shock transcription factors (HSFs) are central regulators of heat responses such as protein biosynthesis and folding and metabolic changes in all eukaryotes. In addition to these conventional functions, plant HSFs have acquired broader roles in growth regulation and biotic stress responses (Scharf et al., 2012), possibly through gene duplications and subsequent mutations in angiosperms (Lehti-Shiu et al., 2017). Reflecting their broader roles, land plants contain multiple HSF genes (e.g. 21 HSF genes in Arabidopsis thaliana) that exhibit functional distinctions between knock-out phenotypes (Scharf et al., 2012). These distinct functions can be explained by tissue- and stress type-specific induction mechanisms as well as their diversified domain structures, which classify HSFs into three families: HSFA, B, and C (Andrási et al., 2021). Members of the HSFA subclass serve as master transcriptional activators of heat stress (HS) responses. Genetic ablation of Arabidopsis HSFA1 caused impaired heat-sensitive phenotypes and altered gene expression (Yoshida et al., 2011; Liu et al., 2011). In contrast, HSFB family member act as either co-activators repressors of HSF, depending on the species and/or experimental conditions (Bharti et al., 2004; Ikeda et al., 2011a; Fragkostefanakis et al., 2019a). In addition to HS responses, the ectopic expression of HSFB genes alters root development under normal growth conditions (Begum et al., 2013).

Morphological control accounts for the developmental adaptation strategies of land plants under HS or high ambient temperature. These unfavorable thermal conditions lead to elongated stems and hyponasty in some crops, in turn leading to a reduction in total biomass (Hasanuzzaman et al., 2013). HS also induces the accumulation of certain metabolites as protectants against high temperatures. More specifically, HS alters the core metabolic fluxes of glycolysis, the pentose phosphate pathway, the TCA cycle, and starch and amino acid biosynthesis pathways (Raza, 2020). Nevertheless, a thorough understanding of metabolic gene expression and metabolite accumulation under HS conditions and the involvement of HSFs is yet to be achieved, mainly due to the high gene redundancy in flowering plants (Lehti-Shiu et al., 2017).

Here, using interspecies meta-transcriptome analysis and HSF knockout mutants in Marchantia polymorpha, we analyzed the diversification and expansion of heat-responsive GRNs during aquatic to land plant evolution. Phylogenetic and evolutionary rate analyses of HSFs revealed the emergence of HSF subclasses at the base of Streptophyta and the strong accumulation of mutations in HSFB genes. Marchantia polymorpha HSF A1 and HSFB1 knockout mutants (Mpsfa1 and Mpsfb1) showed distinct transcriptomic defects under both HS and non-HS conditions, which can be explained by the different usage of cis-regulatory sequences from direct and/or indirect target genes. The Mpsfa1 mutants showed increased thermal sensitivity, such as thallus senescence, metabolic alterations, and growth retardation. The Mpsfb1 mutants showed altered expression profiles of genes related to developmental processes as well as typical heat responses. The unconventional role of Mpsfb1 was confirmed by examining a knockout mutant that displayed developmental defects in meristem and reductive organ formation. We also uncovered gender- and developmental stage-based differences in HS response in M. polymorpha, which could be explained by female- and male-specific Mpsfa1 genes on U and V chromosomes (Mpsfa1U and Mpsfa1V, respectively). Collectively, the diversification of HSFs in early land plants has expanded their gene regulatory and biological roles associated with developmental and metabolic responses to HS.

**Results**
The functions of HS-responsive genes expanded during early plant evolution

It is widely accepted that land plants acquired greatly advanced morphological and intracellular mechanisms to adapt to temperature fluctuations. To compare the HS-responsive mechanisms from green algae to angiosperms, we first analyzed public and in-house transcriptome datasets from a Streptophyta alga (Klebsormidium nitens), a bryophytic plant (M. polymorpha), three flowering plants (A. thaliana, tomato [Solanum lycopersicum] and rice [Oryza sativa]), and the yeast Saccharomyces cerevisiae (Supplemental Figure S1A and Supplemental Data Set 1). These species represent different taxa distributed along the plant evolutionary timescale. K. nitens is a close relative to land plants (Lewis and McCourt, 2004; de Vries and Archibald, 2018), M. polymorpha represents a bryophytic plant, and A. thaliana, S. lycopersicum, and O. sativa are dicot or monocot models that have been extensively studied (Kotak et al., 2007; Mittler et al., 2012).

The multi-taxa analysis yielded approximately 700 to approximately 2,600 HS-responsive genes in different species (Supplemental Figure S1A and Supplemental Data Set 1). Orthologous gene analysis of the HS-responsive genes (Supplemental Figure S1B and S1C) identified 10 and 70 HS-responsive orthogroups conserved in all six species and in three angiosperms, respectively (Figure 1, A and B, highlighted in red text). A higher overlap of HS-induced gene orthologs was observed within four land plants (the degree of overlap ranged from 0.14 to 0.3), which was greater than the median of the degree of overlap between all pairs of six species (median = 0.125), while HS-repressed genes tended to be conserved from K. nitens to angiosperms (Figure 1, C and D). HS-induced genes conserved in all six species and in three plants were associated with the gene ontology (GO) terms abiotic stress response and metabolic process (magenta text, Figure 1E) and regulation of gene expression, development, or reproduction (green text,
Figure 1E), respectively. On the other hand, HS-repressed genes tended to be associated with metabolic process and cell growth in all species examined (Supplemental Figure S1D).

Given that HSFs evolved to form sub-families in land plants (Scharf et al., 2012; Wang et al., 2018) and that HSFs are the master regulators of HS responses in plants (Ohama et al., 2017), we hypothesized that the divergence of HS-induced genes was associated with the fixation and specialization of HSFs in plants. We observed that HSFA1 has maintained the original residues more strictly (turquoise, Figure 2A), while HSFB1 spans a large phylogenetic distance from its emergence to the separation of these proteins in Streptophyta species (Figure 2, A and B; P < 0.01). Analysis of the deduced sequences of ancestral HSFs (Supplemental Figure S2A) revealed that 3 and 10 residues in the HSFA and HSFB subclasses, respectively, mutated from the common HSF ancestor. The mutated residues were located in the α-helix 2 (H2) in the DNA binding domain (DBD), where the domain interacts with DNA (Supplemental Figure S2B). In general, members of the HSFA and HSFB subclasses contain conserved surface residues for DNA-binding in land plants.
Figure 2 HSFA1 and HSFB1 have different phylogenetic distances in plants. A, Phylogeny of plant HSF sequences. The maximum-likelihood tree was constructed using the DBD sequences of selected taxa. HSFA, HSFB, and HSFC groups are shown in different colors. Red text denotes MpHSFA1U, MpHSFA1V and MpHSFB1. Colored nodes indicate HSF sequences in A. thaliana (red), O. sativa (green), bryophyta (purple), and streptophycean algae (yellow). All sequences used to construct the phylogenetic trees are shown in Supplemental File 1 and 2. B, Duplication length from the HSFA and HSFB groups. **P < 0.01, determined by Mann–Whitney U Test. C, Heatmap (left panel) showing the enrichment (–log10(FDR)) of HS-induced TF binding motifs determined by Wilcoxon signed rank test using AME suites. Colors in the heatmap represent the significance determined by hypergeometric distribution followed by BH adjustment. Heat map showing the enrichment (–log10(FDR)) of predictive TF binding motifs determined by Wilcoxon signed rank test using AME suites. Gray cells denote not applicable (NA); red text represents significantly enriched TFs and motifs found in at least four species.
plants (Supplemental Figure S2C), although the H2 helix in HSFA underwent small but further evolutionary changes (Supplemental Figure S2D). The HSFC subclass was separated from HSFA in flowering plants, showing a large branch distance within the HSFA subtree (light green, Figure 2A). These results suggest that the plant HSFB and HSFC subclasses accumulated multiple mutations during the transition from aquatic to land plants and within angiosperms, which might be related to the increasing complexity of plant HS responses.

The number of TFs gradually increased from aquatic to terrestrial plants (Bowman et al., 2017; Wilhelmsson et al., 2017) (Supplemental Figure S2E). Gene number comparisons in aquatic and land plants revealed 40 TF families that are present in at least one green alga genome (Supplemental Figure S2E). Of these, 22 TF families including HSF multiplied in at least one species (Figure 2C, left), only the HSF family showed a conserved HS response in all five species except for S. lycopersicum (light green, Figure 2A). These results suggest that the plant HSFB and HSFC subclasses accumulated multiple mutations during the transition from aquatic to land plants and within angiosperms, which might be related to the increasing complexity of plant HS responses.

MphHSFs regulate distinct developmental processes and thermotolerance

Marchantia polymorpha are bryophytic plants that possess HSFA and HSFB (Wang et al., 2018) (Figure 2A, red text), and MphHSFA1 and MphHSFB1 are ubiquitously expressed in most tissues in Marchantia (Supplemental Figure S2F). To elucidate the functional differences between the two HSF subfamilies, we generated CRISPR/Cas9 knockout mutants for the single copy MphHSF genes (Mphhsfa1V and Mphhsfb1V) with different indel editing sites (Supplemental Figure S2G). We first analyzed the transcriptomic and metabolic profiles of the mutants under control and HS conditions (Figures 3 and 4); the morphological phenotypes and HS responses of the mutants are described in Figures 5–7.

Transcriptome analysis revealed over 3,000 differentially expressed genes (DEGs) in Mphhsfa1V and/or Mphhsfb1 under HS (Figure 3A, Supplemental Figure S3A, Supplemental Data Set 2). Under control conditions (Supplemental Figure S3B), 40 genes (6.9%) overlapped among Mphhsfa1V- and Mphhsfb1-specific genes, while 479 genes (82.6%) were specifically mis-regulated in Mphhsfb1. A large portion of HS-responsive genes detected in the wild type (Figure 1) still exhibited HS responses in Mphhsfa1V and Mphhsfb1 (blue columns, Figure 3A), although their induction or repression levels were smaller in the mutants (Figure 3B; P < 0.001). This suggests the importance of both MphHSF genes and the existence of an HSF-independent mechanism during thermal responses. Among the 1,124 HS-regulated DEGs in Mphhsfb1 and 1,076 in Mphhsfa1V, 473 (42%) and 300 (27.9%) HS-regulated DEGs overlapped with HS-responsive DEGs in wild-type Tak-1, respectively (Supplemental Figure S3B). These results suggest that although MphHSFA1 and MphHSFB1 have individual specialized functions, their functions in regulating HS responses are partially redundant.

GO enrichment analysis revealed that MphHSFA1V regulates canonical HS responses, such as the unfolded protein response, chaperone, and hormone signaling (Figure 3C, magenta text and D), whereas MphHSFB1’s functions are related to cell division, cell-wall modification, and transporters (Figure 3C, green text and E, false discovery rate (FDR) cutoff = 0.1). The requirements of HSF for both basal and HS-regulated gene expression were further confirmed by qRT-PCR of MphHSP90, MphHSP70, MphHSFA1V, and MphHSFB1 (Supplemental Figure S3, B and C). The roles of HSFA1 in transcriptional regulation appear to be conserved in M. polymorpha and A. thaliana, as shown by the significant overlap of the MphHSFA1V-specific genes with those of the Arabidopsis hsfa1 mutant (P < 1 × 10−7; Figure 3, F and G).

Promoter analysis (within 2-kb upstream of the transcription starting site) of these HS-responsive genes revealed the presence of binding motifs for AP2/EREBP, C2H2, C2C2-dof, and BBRBPC TFs (red text in Supplemental Figure S3D), with a higher probability than expected. Interestingly, heat shock elements (HSEs) were significantly enriched only in the promoter regions of Mphhsfa1V-specific genes (33% of Mphhsfa1V-specific genes) under non-HS conditions, while cis-regulatory motifs for Golden2 (G2)-like TFs and WRKY TFs were more highly enriched in Mphhsfb1-specific genes (Supplemental Figure S3D). We then analyzed cis-regulatory elements in the promoter regions of Mphhsfa1V-specific or Mphhsfb1-specific TF genes (Supplemental Figure S3E). HSE was predominantly enriched in Mphhsfa1V-specific TF genes, while NAM, ATAF and CUC (NAC), basic Helix-Loop-Helix (bHLH), and Teosinte branched1/Cincinnati/proliferating cell factor (TCP)-binding motifs were more significantly found in Mphhsfb1-dependent TF genes.

Transcriptional regulation occurs for both direct and indirect targets. To further understand the functions and motif sequence patterns of genes that are likely the direct target of MphHSFA1 and MphHSFB1, we narrowed the analysis to DEGs possessing HSE in their promoter regions. First, we analyzed DEGs downregulated by Mphhsfa1V and Mphhsfb1 mutations under control conditions (22 and 202 DEGs, respectively; Supplemental Figure S4A). Of these DEGs, 11 DEGs in Mphhsfa1V (11/22 = 50%) and 70 DEGs in Mphhsfb1 (70/202 = 34.7%) contained at least one HSE in their promoter regions (Supplemental Figure S4A). GO analysis revealed that the 11 Mphhsfa1V-specific DEGs were associated with the heat response, abiotic stress response, and protein folding, while the 70 Mphhsfb1-specific DEGs were
related to ion transport, homeostatic process, and morphogenesis (FDR < 0.05, Supplemental Figure S4B).

We then extracted HS-induced DEGs containing at least one HSE (Supplemental Figure S4C), as HSEs are generally required for the activation of HS-induced genes (Supplemental Figure S4C). The analysis identified HSE-containing DEGs: 205 genes (205/332 = 61.7%) misregulated specifically in MphasisV1 and 56 genes (56/130 = 43.1%) in MphasisV1, where the lower percentage in MphasisV1 suggested more indirect regulation by MphasisV1 (Supplemental Figure S4C). GO analysis revealed that DEGs not induced by HS in MphasisV1 were associated with canonical HS responses, such as protein folding and chaperone activities, while DEGs not induced in MphasisV1 were associated with the wounding response, metabolic process, and auxin transport (FDR < 0.05, Supplemental Figure S4D). We next analyzed HSE motif sequences in the MphasisV1- or MphasisV1-specific DEGs, which would represent HSFA1-binding and HSFB1-binding motifs, respectively (Supplemental Figure S4E, E and F). Twelve nonredundant HSEs were identified as HSFA1-binding motifs, nine of which contained a canonical HSE sequence (GAAAnTTCC) (Supplemental Figure S4E). On the other hand, six nonredundant HSEs were identified as HSFB1-
**Figure 4** Heat-induced changes in primary metabolism differ between Mphsfa1 and Mphsfb1. A, Thermal regulation of genes involved in primary metabolism in wild type and Mphsf plants. Dot plots show the median log 2 FC of heat-dependent induction (red plots) or suppression (blue plots) of metabolite-related genes in Mphsfa1V, Mphsfb1, and Tak-1 plants. Metabolite-related genes were categorized using the PlantCyc database. PlantCyc pathways with three or more DEGs in at least one genotype are shown. *P* < 0.05, **P** < 0.01, and ***P*** < 0.001, determined by Mann–
binding motifs, but only two sequences possessed the canonical HSE sequence (Supplemental Figure S4E). These results suggest that MpHSFA1 regulates canonical HS responses mainly through direct target genes, while MpHSFB1 controls specialized HS responses both from direct binding (wounding, metabolic process, and auxin transport) and indirect binding (cell wall biosynthesis and development) under HS conditions.

Analysis of the top 10 enriched motifs from MpHSFA1V overexpression chromatin immunoprecipitation (ChIP)-seq peaks validated HS-binding elements as the most significantly enriched motifs (Supplemental Figure S5, A and B, FDR = 10.1 × 10^{-11}). A few other motifs similar to ERF, bZIP, C2C2, and BBRPBC binding motifs were also identified (FDR = 5.08 × 10^{-11} \sim 1.40 \times 10^{-7}, Supplemental Figure S5A). Furthermore, mis-regulated genes in the Mphsfa1V mutants significantly overlapped with genes directly bound by HSFA1 (117, FDR = 1.15 \times 10^{-11}), suggesting they are direct regulatory targets of HSFA1 (Supplemental Figure S5C). GO analysis of the direct regulatory targets revealed functions associated with protein folding, response to stress, and response to HS (FDR < 0.05 as cut off, Supplemental Figure S5D). Collectively, these results possibly attribute both functional overlaps and distinctions between MpHSFA1V and MpHSFB1 to the binding preference of cis-regulatory motifs from direct or indirect target genes and TFs. These motifs are associated with the overlap of either the HSFA or HSFB subclass (Supplemental Figures S3E and S4) with the motifs of heat-induced TFs across five plant species (Figure 2C). This suggests that each HSF subclass acquired specialized functions by recruiting a different set of TFs in their GRNs to trigger HS-responsive cascades either commonly conserved in multiple species or only in specific plant lineages.

MpHSFs modulate primary metabolism in response to HS

HS modulates metabolic processes in various eukaryotes (Figures 1 and 2), which are likely mediated by HSFs in land plants (Figure 3). To further dissect the conserved HS responses, we extracted metabolic pathway genes from the PlantCyc database (Schläpfer et al., 2017; Figure 4A). Compared to Tak-1 plants, Mphpfa1V and Mphpfb1 plants exhibited a reduced number and extent of HS-induced or repressed metabolic genes involved in GO categories amino acid, hormone, secondary metabolism, fatty acid, nucleoside and nucleotide, and transport protein biosynthesis or degradation (Figure 4A, magenta text; P < 0.05). Extraction of PlantCyc metabolic genes from meta-transcriptome analysis revealed that similar metabolic adjustments under HS are conserved from bryophytic to flowering plants (Supplemental Figure S5E). Interestingly, the expression of metabolic genes involved in energy pathways such as photosynthesis and aerobic respiration were drastically altered in Mphpfa1 but not in Mphpfb1 plants (Figure 4A, green text), highlighting additional functional distinctions between the HSFA and HSFB subclasses.

These transcriptome-based inferences were confirmed by profiling primary metabolites by gas chromatography–mass spectrometry (GC–MS). The analysis yielded nearly 300 peaks regulated by HS, which contained 80 compounds annotated in NIST mass spectral library (Figure 4B; Supplemental Data Set 3). In line with the transcriptomic changes in anabolic pathway genes, HS resulted in higher accumulation of various sugars and amino acids such as alanine, proline, phenylalanine, and tyrosine (Figure 4B; P < 0.05, compared to non-HS control conditions). Compared to Tak-1, Mphpfa1V and Mphpfb1 plants exhibited higher steady-state levels of these amino acids under non-HS conditions, with lower levels of induction under HS (Figure 4B, ~70% and ~39% of measured amino acids that significantly accumulated in Tak-1 versus Mphpfa1V or Mphpfb1 plants, respectively). In contrast, in Mphpfa1V and Mphpfb1 plants, ~40% of the measured organic acids highly accumulated after 5 h of HS treatment. Of note, HS induced GABA accumulation only in Tak-1 and Mphpfb1 but not in Mphpfa1V, suggesting that Mphpfa1 plays unique roles in the metabolic HS response (Figure 4B; P < 0.05).

Metabolic adjustments are often linked to developmental strategies under stress conditions. Therefore, we quantified morphological changes and growth retardation in the Mphpf mutants (Figure 5, A–D). Under non-HS conditions, Mphpfa1V (Figure 5A) and Mphpfb1 (Figure 5B) plants exhibited a 1.5-fold (Figure 5C) and 4-fold (Figure 5D) smaller thallus area than the wild-type Tak-1 (P < 0.001), with early and slightly delayed senescence of the thallus, respectively (Supplemental Figure S6, A and C). In addition, the Mphpfb1 mutants, but not Mphpfa1V, showed a higher rhizoid-thallus ratio, a general indicator of stress in M. polymorpha (Supplemental Figure S6, B and D; P < 0.001). A thermostolerance assay in a 37°C growth chamber for 24 h revealed increased HS sensitivity of Mphpfa1V (Figure 5, E–G; P < 0.001) but not Mphpfb1 (Figure 5F). Nevertheless, Mphpfb1 mutants exhibited hypersensitivity to HS when they were subjected to an acute 37°C treatment for 2 h in a water bath (Figure 5, H and I). These results suggest that Mphpfa1 and Mphpfb1 are both required for thermostolerance but have distinct roles in morphological stress responses.
Two *MpHSFA1* genes encoded on the U and V chromosomes have comparable functions in thallus development and HS responses

*MpHSFA1* is encoded as a single-copy gene on the U and V chromosomes of female and male gametophytes (BC3 and Tak-1, respectively) (Bowman et al., 2017). We hypothesized that there are gender-based differences in thermal responses associated with the female- or male-specific *HSFA1*. A whole transcriptome comparison between BC3 and Tak-1 revealed more than 1,000 HS-responsive genes shared by the two gametophytes (Supplemental Figure S7, A and B and Supplemental Data Set 4). The HS-induced genes that overlapped between BC3 and Tak-1 were related to the GO terms heat response, light response, abiotic stress, and protein folding (All, Supplemental Figure S7B). In both gametophytes, temperature response genes showed acute expression changes...
within 1 h of HS, whereas the expression of metabolism process genes was altered later, at 5 h of HS treatment (Supplemental Figure S7B, FDR cut-off: < 0.1). These results indicate that female and male gametophytes have comparable transcriptomic and metabolic responses to HS.

The functions of MphHSFA1U were genetically confirmed by generating female Mphsa1U mutants (Figure 6 and Supplemental Figure S7, C–F). Compared to the parental line BC3, the Mphsa1U mutants exhibited growth retardation and a smaller green area (Figure 6, A and C, left; P < 0.01) and early senescence (Supplemental Figure S7, D and F) but no change in rhizoid coverage (Supplemental Figure S7, E and G). Resembling Mphsa1V mutants (Figure 3), the Mphsa1U mutants showed reduced HS tolerance. After 37°C treatment for 24 h with a recovery period of 14 days, 75%–100% of the thallus area remained green in BC3 female plants compared to 0%–50% in Mphsa1U mutants (Figure 6, B and C, right; P < 0.01). Transcriptome analysis revealed that nearly half of the HS-responsive genes were not induced or repressed in Mphsa1U (Supplemental
Figure S7, H and I). Some other HS-responsive genes were still induced in the mutant, but at a lower level (Supplemental Figure S7I; \( P < 0.001 \)). These transcriptomic changes in Mphsfa1U were highly similar to those in Mphsfa1V (Figure 6, D and E). Consistent with Mphsfa1V (Figure 3C), transcriptomic defects in Mphsfa1U were related to the GO terms heat response and protein folding (Figure 6F) and oxidative stress response and fungal response (Supplemental Figure S7K). Likewise, HS-responsive metabolic changes in Mphsfa1U were comparable to those of Mphsfa1V (Figure S8 and Figure 6G). In general, the Mphsfa1U and Mphsfa1V mutants exhibited greater changes in the accumulation of various amino acids and sugars after 5-h of HS treatment compared to BC3 and Tak-1 (Figure 6H, Supplemental Figure S7L). These results indicate that Mphsfa1 plays comparable roles in the transcriptomic and metabolic regulation of HS responses.

**Overexpression of MphHSFA1U or MphHSFA1V inhibits thallus branching and enhances thermotolerance**

MphHSFA1 knockouts resulted in growth retardation, a possible tradeoff between growth and stress responses (Figure 3). We next examined if the overexpression of MphHSFA1U or MphHSFA1V (Supplemental Figure S8A) alone without HS would cause any developmental anomalies. The thalli of MphHSFA1 overexpression (ox) lines were four- to five-fold smaller than BC3 and Tak-1 thalli (Supplemental Figure S8B and S8C; \( P < 0.01 \)). In addition, MphHSFA1 ox had longer plastochrons, causing the thalli to stretch, with fewer branches rather than appearing rounded (Supplemental Figure S8, C and D). Gender-based differences were observed in MphHSFA1-induced necrosis. Both MphHSFA1U and MphHSFA1V ox lines in the BC3 background exhibited necrosis of the growing points beginning on Day 14, while MphHSFA1 ox lines in the Tak-1 background showed no necrosis even on Day 28 (Supplemental Figure S8E). Despite the gender-based differences in necrotic phenotypes, all MphHSFA1 ox lines showed enhanced thermotolerance in both the BC3 and Tak-1 backgrounds (Supplemental Figure S8, F and G).

Transcriptomic profiles of MphHSFA1 ox lines under non-HS conditions would reflect the developmental and other biological processes controlled by HS. A comparison of such transcriptome data revealed roughly 100 genes commonly regulated by MphHSFA1U and MphHSFA1V overexpression and HS in both the BC3 and Tak-1 backgrounds (Supplemental Figure S9, A and B, Supplemental Data Set 2). However, their levels of induction and repression in response to MphHSFA1 ox were significantly smaller than those regulated by HS in the parental wild-type plants (Supplemental Figure S9, C and D). GO terms associated with temperature response and protein folding were consistently enriched in the overlapping DEGs (far right columns, Supplemental Figure S9E), although several HS-responsive genes were specifically induced in Tak-1 or BC3 plants (far left columns, Supplemental Figure S9E). Interestingly, HSEs were highly enriched in HS-responsive genes in both BC3 and Tak-1 (far left columns, Supplemental Figure S9F), but not in the genes controlled by MphHSFA1 ox under non-HS conditions. Taken together, MphHSFA1 overexpression alone only partially induced HS responses and transcriptomic changes and did not fully reproduce thermal responses, which might require the activation of additional thermal signaling pathways.

**Gender-specific developmental and thermotolerance phenotypes associated with MphHSFA1U and MphHSFA1V**

There was no clear gender-based differences in transcriptomic changes caused by HS (Supplemental Figure S7, A and B) or MphHSFA1 overexpression (Supplemental Figure S8) when we compared Tak-1 with BC3, which is a female introgression line of Tak-1. Nevertheless, BC3 plants were more sensitive to HS than Tak-1 plants (Figure 7, A and B; \( P < 0.001 \)). To further explore the differences in thermal sensitivity, we compared the transcript levels of sex chromosome genes using fragments per kilobase of transcript per million mapped reads (FPKM) values as a measure of expression levels (Figure 7C and Supplemental Figure S10A). Three and 12 genes out of 75 U chromosome and 99 V chromosome genes showed altered transcript levels after HS treatment (FDR \(< 0.1 \)), but this was independent of gender and Mphsfa1 knockout (Supplemental Figure S10A). Interestingly, however, the FPKM value of MphHSFA1V was 2- to 2.5-fold higher than that of MphHSFA1U under non-HS conditions (Figure 7C). In addition, genes downregulated specifically in Tak-1 plants, but not in BC3 or Mphsfa1 plants, were associated with DNA conformation changes (Supplemental Figure S10B), which may explain the gender-based differences in HS sensitivity (Ray et al., 2019; Montgomery et al., 2020).

Finally, we analyzed reproductive organ development as a haplotype-specific biological process. Both male and female Mphsfa1 mutants showed delayed antheridiophore or archegoniophore development compared to Tak-1 (Figure 7D and E) or BC3 (7F and 7G). Mphsfa1V and Mphsfa1U plants also exhibited more yellowish stems or reproductive organs after HS (Figure 7, D [HS] and 7, F [HS], white arrows), pointing to their roles in the development and thermotolerance of reproductive tissues. Surprisingly, Mphsfb1 plants produced no antheridiophores under the same condition (Figure 7D and E). Based on this observation, together with the results of transcriptomic analysis (Figure 3), we propose that MphHSFA1 controls reproductive organ formation as a part of a HS response, whereas MphHSFB1 directly regulates antheridiophore development. The data collectively suggest that MphHSFA1 is a master regulator of HS responses, while MphHSFB1 has more specialized functions in addition to temperature responses.

Land plants often face gradual temperature changes over a long period of time, which could also be regulated by MphHSFA1 and MphHSFB1. We therefore treated plants at
22°C, 29°C, and 32°C and monitored their growth over 21 days (Supplemental Figure S11, A–E). *M. polymorpha* plants were hypersensitive to 29°C, as shown by the significantly reduced plant size ($P < 0.011$) at this temperature compared to 22°C (Supplemental Figure S11F). In contrast, the thallus growth of *Tak-1* and *M. polymorpha* was similar between 22°C and 29°C. At 32°C, thallus growth was arrested after 5 days of treatment regardless of genotype. These results suggest that *M*PH*SFA1* is responsible for controlling wider temperature responses, spanning ambient temperature changes to HS responses, while *M*PH*SB1* plays important roles preferentially at higher temperatures.

Gemma cup formation was also promoted at 29°C in *Tak-1*, but not in *M. polymorpha*, which had a similar number of gemma cups at both 22°C and 29°C (Supplemental Figure S11G; $P = 0.018$). Female BC3 plants showed a severe growth defect ($P < 2 \times 10^{-16}$) at 29°C, where their thalli tended to curl and burrow into the agar medium (Supplemental Figure S11H).
Figure S12, A–C). The growth defect was more severe in Mphsfa1U, which were smaller at 29°C (Supplemental Figure S11; P = 0.021). The observed gender-associated difference aligns with the lower survival rate observed in BC3 plants compared to Tak-1 plants (Figure 7B).

Discussion

Terrestrialization of land plants is generally viewed as the most important event in the natural history of life. Successful terrestrialization required drastic changes in the morphological features and intracellular mechanisms of plants to help them adapt to harsh terrestrial environmental conditions, such as high temperature and water deficiency, a stark contrast to the likely stable aquatic environment. The fixation of HSFs and the subsequent rapid evolution of the HSFB subclass occurred during the early evolution of Streptophyta plants (Figures 1 and 2). This evolutionary mode contrasts to that of typical TFs including HSFA subclass members within the vascular plant clade (Scharf et al., 2012), in which the diversification of TFs is gradual, as supported by phylogenetic trends with small phylogenetic branch lengths. The saltation of protein sequences appears as an unbalanced long branch length in a phylogenetic tree. Such long branches, which were described previously in plant TF phylogenies (D), could have also augmented the evolution of ancestral green species to widen the physiological and developmental functions that are inherited by extant land plant species. The model for plant HSF diversification and functional expansion is summarized in Figure 7H.

The rapidly evolved HSFB subclass control morphological processes under non-HS conditions (Figures 5, 7, D and E) and the expression of various genes relates to development or cell wall biosynthesis (Figures 3 and 4). The developmental defects in Mphsfb1, such as defective antheridophore formation (Figure 7, D and E), could be mediated by auxin signaling pathways. The vegetative-to-reproductive phase transition in angiosperms requires auxin accumulation (Cheng et al., 2007; Krizek, 2011). Analysis of organ-specific transcriptomes in M. polymorpha suggested the co-expression of Mphsfb1 with auxin biosynthesis and perception genes such as TAA, YUCCA, and AUX/IAA in reproductive organs (Flores-Sandoval et al., 2018). These lines of evidence prompted the suggestion that auxin mediates some HSF-dependent developmental processes in M. polymorpha and in A. thaliana. Interestingly, genes for abiotic stress responses, defense responses, and acquired system resistance were anomalously expressed in Mphsfb1 even under non-HS conditions (Figure 3). Similar defense-related functions of HSFB were observed in S. lycopersicum, where an Slhsfb1-RNAi mutant showed altered proteome profiles associated with abiotic stress responses, defense responses, and cell-wall modification (Fragkostefanakis et al., 2019). In addition, Arabidopsis AtHSFB1 is known to regulate the growth-to-defense transition upon pathogen invasion (Pajerowska-Mukhtar et al., 2012). Hence, the defense-related functions of HSFB, as well as their developmental functions, were evolutionarily fixed prior to the separation of vascular and nonvascular plants.

Our transcriptome and metabolome analyses revealed the accumulation of amino acids such as proline, alanine, phenylalanine, and tyrosine under HS (Figure 4). Proline serves as a general stress protectant (Hare and Cress, 1997), while the aromatic amino acids phenylalanine and tyrosine are two key compounds used to produce phenylpropanoid-derived secondary metabolites (Vogt, 2010). The genes for phenylpropanoid pathways were enriched under HS in multiple species including Klebsormidium (labeled by the term cell structure in Supplemental Figure S5E). In addition, similar changes in metabolites and metabolic pathways were previously observed in heat-stressed Arabidopsis, tomato, and green algae (Fragkostefanakis et al., 2016; de Vries et al., 2020; Paupière et al., 2020). These observations indicate that the accumulation of phenylpropanoid compounds is a well-conserved HS response inherited from aquatic green species. Genes involved in energy production were also enriched under HS in multiple species, indicating the important role of primary metabolic changes that were conserved from aquatic to terrestrial plants (Supplemental Figure S5E). The accumulation of components in both the phenylpropanoid and energy pathways (amino acids, sugars, and organic acids) is regulated by both Mphsfa1 and Mphsfb1 (Figure 4). While the metabolomic profiles of HSF-null mutants in angiosperms are yet to be explored, our findings suggest that the metabolic functions of HSF existed in ancestral green algae and were retained in the HSFA and HSFB subclasses after the split.

This study also focused on gender-based differences in HS responses, which were not explained by the sequence diversity of Mphsfa1U and Mphsfa1V (Figure 7). Instead, transcriptional regulation of Mphsfa1V and Mphsfa1U may partially account for the differences (Figure 7). For example, Mphsfa1V (male), with high steady-state expression levels, contains 80 TF-binding motifs, including 4 HSEs, whereas Mphsfa1U (female), which is expressed at low levels, only contains 41 motifs with no HSE (within the 2-kb promoter region). Mphsfa1U and Mphsfa1V share CDS- and protein-sequence identities of 63.5% and 74.5%, respectively, with different exon-intron structures. We also found that Hsfal gene is located on the U chromosome in Marchantia inflexa, which separated from M. polymorpha more than 100 million years ago (Villarreal et al., 2016). However, since BC3 is an ~90% introgression line, there might be few SNPs that contribute to the different HS responses of male and female M. polymorpha (Bowman et al., 2017). While this concept must be confirmed experimentally, the long-term retention of the two HSFA genes on Marchantia sex chromosomes further supports the regulatory or functional distinctions between the gender-specific HSFAs.

Collectively, this study focused on HSFA and HSFB in M. polymorpha and revealed their common and subclass-
specific functions (Figures 3–7). While this and other previ-
sous studies mainly focused on the functions of HSF genes in
HS responses, our transcriptome and cis-regulatory element
analyses strongly suggested the interplay of these HSFs with
different types of TFs under HS. For example, besides the
cis-regulatory motifs (e.g. HSE) that could be found in com-
monly regulated genes in Mphpsfa1 and Mphpsfb1, cis-
regulatory motifs for HSFs and stress-related TFs were highly
enriched in Mphpsfa1-specific genes, while WRKY-binding
motifs were highly ranked among Mphpsfb1-specific genes
(Supplemental Figure S3 and S4). The different coupling
partners would have given rise to functional distinctions be-
tween GRNs downstream of HSFA and HSFB either directly
or indirectly. Duplicated genes are often maintained in the
genome for a long period of time once one of the duplicates
has acquired new important functions. This consequently
leads to the expansion of GRNs (Voordeckers et al., 2015).
The subsequent functional distinctions between the HSFA
and HSFB subclasses would exemplify typical evolutionary
processes that occurred in early plant life.

The GRNs underlying HS responses are initiated when
land plants and algae face rapid and drastic temperature
increases. These GRNs are centered around HSFs. A single
canonical HSF (CrHSF1) in Chlamydomonas reinhardtii is a
key regulator of HS that directly activates HSPs to protect
cells from mis-folded proteins (Schmollinger et al., 2013).
CrHSF1 possesses features typical of Class-A HSFs, suggesting
their conserved roles across algae and land plants (Schroda
et al., 2015). In addition, the expression of HS marker genes
such as HSP70, HSP90, and sHSPs was induced at 15 min of
HS treatment and lasted for 2 h in C. reinhardtii and K. nit-
tens (Monte et al., 2020), which is similar to our observations
in A. thaliana seedlings or M. polymorpha gemmalings at
36°C to 42°C (Supplemental Figure S4; Larkindale and
Vierling, 2008; Myouga et al., 2006). This indicates that rapid
transcriptional responses upon HS are similar in green algae
and young land plant seedlings. Considering that the aquatic
habitat is subject to prolonged HS due to the high heat ca-
pacity of water, transcriptomic comparisons of prolonged
HS responses between aquatic and land plants would be
of interest in the future. Mphpsfa1 and Mphpsfb1 likely
control plant responses at different temperature ranges
(Supplemental Figures S11 and S12). The activation of HSFs
and downstream TFs or genes at different temperature
ranges may thus be a part of their functional distinctions.

The Streptophyta alga K. nitens and the bryophyte M. pol-
ymorpha contain two HSFs belonging to Class A and Class
B. Our sequence analysis revealed that some charophyte al-
gae lack HSFA or HSFB (Figure 2), implying that the two
classes of HSFs were not evolutionarily fixed in some algal
lineages. Our study showed that Mphpsfa1 is a key regulator
of HS response, similar to CrHSF1 and other Class A HSFs in
vascular plants. On the other hand, Mphpsfb1 possesses
functions distinct from Class A HSFs as well as Class B HSFs
in land plants. Given the divergent functions demonstrated
in Marchantia, Arabidopsis, and tomato (Ikeda et al., 2011;
Fragkostefanakis et al., 2019), HSFB1 likely acquired special-
ized functions during the evolution from algae to land
plants to extend HS GRNs.

Materials and methods

Growth conditions and morphological
colorization

Marchantia polymorpha wild type (accession BC3 [female]
and Tak-1 [male]), Mphpsfa1 overexpression lines, and
Mphpsfa1 and Mphpsfb1 knockout mutants were grown on 1/
2 × Gamborg B5 medium (pH 5.5) with 1% agar under
continuous light (50–60 μmol m⁻²s⁻¹, 6500K Cool Daylight
LED, Philips) at 22°C. For morphological characterization,
gemmae were grown under a 16-h light/8-h dark photope-
eriod for 10 days at 22°C, and the gemmalings were then
transferred to continuous light at 22°C. Plant images were
taken daily from Days 10 to 21 for the analysis of plant area,
number of thalli, and aspect ratio of the ellipse covering the
thallus (ratio of the length of the shorter axis to the longer
axis). These morphological characters were extracted with
OpenCV2 (Pulli et al., 2012).

HS treatment of M. polymorpha

Ten to 20 wild-type and mutant gemmae from the respective
hsfa1 lines were grown on 1/2 × B5 (pH 5.5) medium
with 1% agar at 22°C under a 16-h/8-h light/dark period for
10 days. The 10-day-old plants were treated at 37°C for 24-
h with continuous light in a growth chamber and then re-
covered for 14 days at 22°C. Plants were imaged on Days 1
(right after treatment) and 14. Survival ratio was calculated
as the green pixel area/total thallus area. The green and thal-
lus pixels are defined as the H, S, and V values between [30,
20, 255] and [80, 255, 255], respectively. The proportion of
gemmalings that remained green on Day 1 and Day 14 were
recorded for each genotype. To study the tolerance to
harsh temperature changes, hsfb1, HSFA1U and HSFA1V ox
mutants grown in agar plates were immersed in a water
bath preheated to 37°C. Seven-day-old plants grown on 1/2
× B5 (pH 5.5) plates with 1% agar were immersed in a wa-
ter bath at 37°C for 2 h (without light) and immediately
transferred to a 22°C growth chamber with continuous light
for recovery. Photographs were taken and survival rates
were analyzed on Day 5 after HS.

For long-term HS, gemma from Mphpsfa1V, Mphpsfb1, Tak-
1, Mphpsfa1U, and BC3 plants were grown under a 16-h
light/8-h dark photoperiod for 7 days at 22°C. The gemmal-
ings were then transferred to growth chambers under a 16-
h light/8-h dark photoperiod at 22°C, 29°C, or 32°C for 21
days. Plant images were taken daily for the analysis of plant
area and number of gemma cups. Plant size was extracted
with OpenCV2 (Pulli et al., 2012).

HSF overexpression and knockout mutant lines in
M. polymorpha

All cloning and transformation were performed as previously
described (Wu et al., 2021). For overexpression mutant lines,
protein-coding sequences of MphsFA1U and MphsFA1V were amplified from BC3 and Tak-1 plants, respectively. The DNA fragments were inserted into the binary vector pMpGWB336 that expresses an N-terminal tagRFP-tagged protein under the control of the MpEF1α promoter. The plasmids were transformed into M. polymorpha thalli via Agrobacterium-mediated transformation (Ishizaki et al., 2016), and successful transformants were screened on 1/2 B5 medium plates with 100-μg mL^-1 cefotaxime and 0.5-μM chlorosulfuron. Transgene expression was confirmed in T2 generation plants using qPCR, and two independent ox lines were selected based on the transcript level (Supplemental Figure S7A). Gene locus identifiers and primers used for sequencing and cloning are listed in Supplemental Data Set 5.

Mphsfa1U, Mphsfa1V, and Mphsfb1 knockout mutants were generated using the clustered regularly interspaced short palindromic repeats/CRISPR Associated protein 9 (CRISPR/Cas9) method. Two gRNAs flanking the short DNA fragment of the target gene were designed and cloned into the pMPGE011 vector. Agrobacterium-mediated transformation of thalli was performed as previously described (Kubota et al., 2013), followed by the screening of transformants with 100-μg mL^-1 cefotaxime and 0.5-μM chlorosulfuron. For double transformation, thalli were co-transformed using two constructs with the same selectable marker. All experiments were conducted using the T2 generation. Detailed information about the Mphsfa1IV, Mphsfb1, Mphsfa1Vb1, and Mphsfa1U knockout mutants and gRNA sequences is provided in Supplemental Figures S2G, S6E, and S7C.

RNA sequencing and data analysis
For M. polymorpha, gemmalings were grown in 1/2 × B5 liquid medium for 7 days under continuous light and subjected to HS at 37°C for 0, 1, and 5 h. Total RNA was isolated from the gemmalings using an RNAeasy Plant Mini Kit (Qiagen, Hilden, Germany). For K. nitens, droplets of algal suspension from a fully grown plate were plated on C medium plates and incubated for 1 month. The algal plates were then immersed into a water bath at 37°C for 2 h. Library preparation, including polyA enrichment and sequencing of 150-bp paired-end reads on the NovaSeq platform, was performed at Novogene, Singapore.

Raw sequences from M. polymorpha and K. nitens were mapped to the reference genomes (M. polymorpha genome v.5.1 and K. nitens GCA_000708835.1_ASM70883v1 from NCBI) using STAR (v.2.5.3a) (Dobin et al., 2013) with default settings. Differentially expressed genes (DEGs) were identified with the R package DESeq2 (Love et al., 2014), followed by comparisons to the expression level at 0 h with the threshold of log2-fold change ≥ 1 and False Discovery Rate (FDR) < 0.1 using the “FdrShrink ashr” method. Arabidopsis orthologs of M. polymorpha and K. nitens genes were identified as previously described (Fischer et al., 2011) for GO analysis. GO term enrichment analysis was performed with the AgriGO v.2 database (bioinfo.cau.edu.cn/agriGO) using standard settings. Significantly enriched GO terms (FDR < 0.05) were visualized as dot plots using the R package ggplot2. The temperature used for RNA-sequencing was determined based on the induction levels of HSFA1V, HSFB1, and a heat shock protein gene (sHSP) at 22°C to 40°C (Supplemental Figure S3A). The three genes showed the highest expression at 37°C and decreased expression levels in the hsfa1V and hsfb1 mutants. For the analysis of sex chromosome genes, trimmed sequences from M. polymorpha were mapped to the M. polymorpha (v.3.1, marchantia.info) reference genome because the full gene annotations of both sex chromosomes were available in this version. The IDs were converted to v.6.1.

Motif enrichment analysis
AME from MEME suite (Bailey et al., 2009) was used to identify enriched TF binding motifs in the 2-kb region upstream from the transcription start site of genes assigned to each DEG group containing genotype-specific genes (Supplemental Figure S3D) or HS-induced TF genes (Figure 2C and Supplemental Figure S3E). The genes were compared with nonheat-responsive genes (Supplemental Figure S3D) or nonheat-responsive TFs (Figure 2C and Supplemental Figure S3E) to calculate the enrichment (adjust P-value) by Wilcoxon Rank sum test, adjusting by Benjamini–Hochberg (BH) post-hoc test. The parameters were set as follows: --verbose 1 --scoring avg --method fisher --hit-lo-fraction 0.25 --eval-score-report-threshold 10.0 --control --shuffle --kmer 2. The Arabidopsis DAP-Seq database (O’Malley et al., 2016) was used as a reference.

RNA extraction, cDNA synthesis, and qRT-PCR
Seven-day-old gemmalings grown in liquid 1/2 × B5 medium were treated at 37°C for 0, 1, and 5 h with continuous light. After treatment, plant samples were harvested, immediately frozen in liquid N₂, and stored at –80°C until further use. Total RNA was extracted from M. polymorpha gemmalings using an RNAeasy Plant Mini Kit (Qiagen, Hilden, Germany). First-strand cDNA was synthesized using the GoScript Reverse Transcription System (Promega, USA). qRT-PCR was performed as previously described (Wu et al., 2020). Data were analyzed according to the manufacturer’s instructions using CFX Manager v3.1 (Bio-Rad, USA). Gene expression levels were normalized using the expression of M. polymorpha ACTIN 7 (MpACT7). The primers used are listed in Supplemental Data Set 5.

ChIP and ChIP-sequencing
Approximately 1.5 g of 7-day-old gemmaling tissue from Tak-1 and the MphsFA1V overexpression lines were cross-linked with 1% formaldehyde in a vacuum chamber twice for 5 min. Subsequently, the chromatin DNA was isolated and sonicated as described previously (Wu et al., 2018). The DNA was immunoprecipitated with 25-μL anti-tagRFP magnetic beads (rtd-20, Chromotek, Germany) into 1 mL of solution containing 1.1% Trion X-100, 1.2-mM EDTA, 16.7-mM Tris–HCl, pH 8.0, 167-mM NaCl, and protease inhibitor (cComplete Protease Inhibitor Cocktail, Roche) (Wu et al., 2018).
The precipitated DNA was purified with a Qiaquick PCR purification kit (Qiagen, Germany) and processed further for ChIP-sequencing. Library preparation and sequencing of 150-bp paired-end reads were performed on the NovaSeq platform in Novogene, Singapore. Two biological replicates of Tak-1 and MpHSFA1V overexpression lines were included. Bowtie2 (v2.3.5) software (Langmead and Salzberg, 2012) was used for mapping using default settings, except that we discarded reads matching multiple loci, as they might introduce error signals by repeat counting. Peak calling was performed with the model-based analysis software MACS2 (Zhang et al., 2008) using Tak-1 data as a control. Two biological replicates were peak-called separately, and only the peaks that overlapped in both replicates with an irreproducible discovery rate < 0.05 were included in further analysis. The ChiPseeker (Yu et al., 2015) package in R was used for downstream analyses including peak annotation and genomic feature visualization. GO analysis was carried out using the AgriGO v.2 database. Motif discovery analysis was performed in AME from MEME suite (Bailey et al., 2009). The enriched motifs were compared with the DAP-seq database (O’Malley et al., 2016).

Interspecies comparison of HS-induced transcriptomic changes

A publicly available microarray dataset was retrieved from GEO (Supplemental Figure S1A; Liu et al., 2011; Liu and Chang, 2013; Zhang et al., 2013; Mishra et al., 2016; Cortijo et al., 2017; Keller et al., 2017; Alibhal et al., 2018; Nuño-Cabanes et al., 2020). The microarray data were processed with quantile normalization and background correction using the robust multi-array average (RMA) with the R packages “affy” and “gcma” (Wu et al., 2004; Gautier et al., 2004). The “limma” package (Smyth, 2005) was used to compute log2-transformed values of transcripts in a linear model. DEGs from the corresponding control group were identified using log2-fold change (FC) value > 1 and FDR < 0.05 as the threshold. Affymetrix probe IDs were then converted to AGI locus identification numbers on the BAR website (Waese et al., 2017). Probe IDs with no match or ambiguously matching multiple loci were discarded.

RNA-seq datasets for S. cerevisiae, A. thaliana, O. sativa, and S. lycopersicum were obtained from the GEO database (Supplemental Figure S1A). The raw fastq files were mapped to the respective genomes using STAR mapper to generate raw read counts. The datasets were first filtered by array quality weight. Datasets with moderated-t-statistics that were out of range of other datasets were eliminated. All data were processed using the package LIMMA (Smyth, 2005) to minimize the variations caused by different algorithms used for data pre-processing. Affymetrix and Agilent microarray data were background corrected using the Robust MultiChip Average (RMA) method, while RNA-Seq data were treated with voom, where the mean–variance relationship of the log-read counts was estimated and given a precision weight (Law et al., 2014). The data were then normalized using quantile normalization and converted into log-scale before entering the limma empirical Bayes analysis pipeline. To access the data quality after processing, the distance between array, array intensity distribution, and mean–variance trend were evaluated. The genome versions used in this study were as follows: S. cerevisiae S288C, A. thaliana-TAIR10, K. niten-GCA_000708835.1_ASM70883v1, S. lycopersicum-SL 3.0, M. polymorpha-v3.1 and v5.1, O. sativa-IRGSP 1.0.

Orthologous gene analysis

Whole proteome sequences of S. cerevisiae, K. niten, A. thaliana, O. sativa, S. lycopersicum and M. polymorpha were downloaded in FASTA format from the following databases: S. cerevisiae: S288C from Saccharomyces genome database (https://www.yeastgenome.org/), K. niten: GCA_000708835.1_ASM70883v1 from NCBI (https://www.ncbi.nlm.nih.gov/), A. thaliana: TAIR10 from TAIR database (https://www.arabidopsis.org/), O. sativa and S. lycopersicum: IRGSP 1.0 and SL3.0 from Phytozome v13 (https://phytozome-next.jgi.doe.gov/), and M. polymorpha: v5.1 from MarpolBase (https://marchantia.info/).

The proteome sequences from primary transcripts were used to construct OrthoMCL groups by sorting the sequences to pre-determined orthogroups (v6.1) with BLASTp (https://orthomcl.org/orthomcl/) (Fischer et al., 2011). The full list of orthogroups in each species is provided in Supplemental Data Set 1. In total, we identified 4,767 (S. cerevisiae), 11,246 (K. niten), 7,804 (M. polymorpha), 24,794 (O. sativa), 14,689 (A. thaliana) and 10,892 (S. lycopersicum) orthogroups in six species from a total of 495,339 pre-defined orthogroups in OrthoMCL v6.1. The genes assigned to the OrthoMCL groups were used for further analysis. The HS responses of these orthologous genes were analyzed and presented as the degree of overlap (calculated as |A ∩ B|/ min(|A|, |B|), where |A| and |B| represent the number of stress-related genes in stress A and B, respectively, Ma et al., 2014) of orthologous genes between species.

Phylogenetic analysis

HSF DBD domain sequences from A. thaliana and Oryza sativa spp. Japonica, Picea abies, Ginkgo biloba, Azolla filiculoides, Salvinia cucullata, Selaginella moellendorffii, Ceratodon purpureus (transcriptome), Physcomitrella patens, M. polymorpha, Spirogyra pratensis, Coleochaete orbicularis, Chara brauni, Mesotaenium endlicherianum, and Klebsormidium flaccidum were collected from the Tapscan database and JGI PhycoCosm (https://phyocosm.jgi.doe.gov/phyocosm/home) (Wilhelmsson et al., 2017). HSF sequences from Chlorophyta and Rhodophyta species; Chlamydomonas reinhardtii, Volvox carteri, Chlorella sp. NC64A, Chlorella vulgaris, Porphyra umbilicalis, Cyanidioschyzon merolae, Chondrus crispus, and Galdieria sulphuraria were included as outliers to construct a phylogenetic tree. The HSF sequences were preprocessed by removing redundant sequences that showed 90% or higher identity to other sequences. The remaining sequences were aligned using MAFFT version 7,
following the removal of gapped positions containing \( \geq 10\% \) gaps. MpHSFA1U, MpHSFA1V, and MpHSFB1 sequences were manually added to the multiple sequence alignment, as they were removed due to high similarity to other entries. The nonredundant set of HSF sequences (110 sequences with 89 positions, Supplemental Files 1 and 2) was used for phylogenetic analysis by IQ-TREE multicore version v2.1.2. The phylogenetic tree was constructed with the LG + R5 substitution model and 1,000 ultrafast bootstrap iterations. The model was selected based on Bayesian Information Criterion. For duplication length analysis (Vosseberg et al., 2021), phylogenetic distance to the duplication node uniting HSFA and HSFB branches was calculated for each HSFA or HSFB sequence. This distance was divided by the stem length of HSFA or HSFB branches (0.083 or 0.1924, respectively, Figure 2A). The duplication length was negatively log-transformed and presented in a bar histogram with frequency density. The common ancestor sequence was reconstructed using phylm (Guindon et al., 2005).

**Structure homology modeling**

The structures of MpHSFA1 and MpHSFB1 DBD were modeled using Modeller software version 9.23 (Eswar et al., 2006). A template atomic structure of animal HSF1 was obtained from the Protein Data Bank (PDB, ID = 5dsu; Neudegger et al., 2016). The protein sequences of MpHSFA1 and MpHSFB1 DBD were aligned with the HsHSF1 sequence in Modeller. Five structures each of HSFA1 and HSFB1 were modeled automatically with default settings. Modeled structures were assessed by the discrete optimized protein energy (DOPE) score, and the structures with the lowest DOPE values were used for further analyses. The PyMOL program was used to color atoms and generate the final images.

**Entropy measurement at each position in the MSA**

The Shannon entropy and JSD scores were used to estimate the conservation of residues for each position in plant HSF sequences (108 sequences with 89 positions, see above for multiple sequence alignment). Shannon entropy and JSD scores were calculated with the BLOSUM62 substitution matrix for estimating substitution probabilities and the Protein Residue Conservation Prediction program for background frequency of residue variability (Capra and Singh, 2007). The JSD score is presented in pseudo-color scale on the HSF structures. The Shannon entropy at each position is presented in Seq-Logo format.

**Metabolite extraction and metabolic profiling by GC–MS**

Seven-day-old *M. polymorpha* gemmalings (~50 mg) were collected and immediately frozen in liquid nitrogen prior to storage at \(-80^\circ\)C right after HS treatment at 37°C for 0, 1, and 5 h. The samples were then homogenized using a mortar and pestle precooled with liquid nitrogen and extracted in 700-\(\mu\)L methanol, and 60-\(\mu\)L of internal standard (0.2-mg nabitol mL\(^{-1}\) H\(_2\)O) was subsequently added as a quantification standard. The extraction, derivatization, standard addition, and sample injection were conducted as described previously (Lisec et al., 2006). Untargeted GC–MS analysis was performed using a previously described system (Avin-Wittenberg et al., 2015). In brief, samples were separated using a DB-35ms Ultra Inert column (30 m × 0.25 mm × 0.25 \(\mu\)m, Agilent, USA) with Agilent 7200 Q-TOF GC–MS coupled to Agilent 7890B Gas Chromatograph by Chemical, Molecular, and Materials Analysis Centre (CMMAC), National University of Singapore. Raw MS data were extracted using the MassHunter Profiler suite (B.08.00, Agilent), and resulting peaks were aligned and mapped to the NIST database (2017) using Mass Profiler Professional software (v 8.0, Agilent) with default setting. Student’s t test (\(P < 0.05\)) was applied to generate differentially accumulated compounds in HS versus control conditions in at least one pair-wise comparison.

**Statistical analysis**

Statistical analysis and graph construction were performed in R (v 4.0.4). The results of statistical analysis are provided in Supplemental Data Set 6.

**Accession numbers**

Sequence information from this article can be found in the Marchantia database (https://marchantia.info/) or under the following accession numbers: MpHSFA1U: MpUg00290, MpHSFA1V: MpVg00470, and MpHSFB1: Mp4g12230.

Raw sequencing data are available at the NCBI website under accession number GSE178776.

**Supplemental data**

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Orthology analysis of HS-responsive genes.

**Supplemental Figure S2.** Structural analysis of HSFA1 and HSFB1 and identification of *hsfa1* and *hsfb1* mutants in Marchantia.

**Supplemental Figure S3.** Transcriptomic changes and putative cis-regulatory elements (pCREs) in response to HS in Marchantia.

**Supplemental Figure S4.** Analysis of HSE-containing genes in *Mphsfa1V* and *Mphsfb1* mutants.

**Supplemental Figure S5.** Binding motifs of MpHSFA1.

**Supplemental Figure S6.** Characterization of *Mphsfa1V* and *Mphsfb1* mutants.

**Supplemental Figure S7.** Identification and characterization of *Mphsfa1U* mutants.

**Supplemental Figure S8.** Characterization of *HSFA1U* and *HSFA1V* ox lines in Marchantia.

**Supplemental Figure S9.** RNA-seq analysis of *HSFA1ox* lines in the Tak-1 and BC3 backgrounds.

**Supplemental Figure S10.** Transcriptomic changes in genes in BC3 and Tak-1 plants and in U and V chromosomes.

**Supplemental Figure S11.** Sensitivity of *Mphsfa1V* and *Mphsfb1* mutants at 29°C.
Supplemental Figure S12. Sensitivity of Mphsa1U mutants and BC3 plants at 29°C.

Supplemental Data Set 1. Orthogroups and orthologous HS-responsive DEGs from the selected taxa.

Supplemental Data Set 2. Differentially expressed genes in wild type and hsf mutant plants.

Supplemental Data Set 3. Metabolite contents under HS and control conditions.

Supplemental Data Set 4. Genes differentially expressed between three wild-type accessions of M. polymorpha (Tak-1, Tak-2, and BC3).

Supplemental Data Set 5. Primers used in this study.

Supplemental Data Set 6. Results of statistical analysis.

Supplemental Data File 1. Sequence alignment of HSF proteins used for phylogenetic analysis.

Supplemental Data File 2. Phylogenetic tree file for Figure 2A.

Funding
This study was supported by the Agency for Science, Technology and Research (A*STAR) Singapore under the industry alignment fund pre-positioning program; High Performance Precision Agriculture (HPPA) system (A19E4a0101) and by the Singapore-MIT Alliance for Research and Technology. Disruptive & Sustainable Technologies for Agricultural Research (A*STAR) Singapore under the industry alignment fund pre-positioning program.

Conflict of interest statement.
The authors have no conflicts of interest.

References
Albihal WS, Irabonosi O, Blein T, Persad R, Chernukhin I, Crespi M, Bechtold U, Mullineaux PM (2018) Arabidopsis HEAT SHOCK TRANSCRIPTION FACTOR A1b regulates multiple developmental genes under benign and stress conditions. J Exp Bot 69: 2847–2862
Andrási N, Pettkó-Szandtner A, Szabados L (2021) Diversity of plant heat shock factors: regulation, interactions, and functions. J Exp Bot 72: 1558–1575
Avin-Wittenberg T, Bajdzienko K, Wittenberg G, Alseekh S, Tohge T, Bock R, Giavalisco P, Fernie AR (2015) Global Analysis of the Role of Autophagy in Cellular Metabolism and Energy Homeostasis in Arabidopsis Seedlings under Carbon Starvation. Plant Cell 27: 306–322
Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li VW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37: W202–W208
Begum T, Reuter R, Schöffl F, Bharti K, Von Koskull-Döring P, Bharti S, Kumar P, Tintschl-1, Tak-2, and BC3). Erwin DH, Davidson EH (2009) The evolution of hierarchical gene regulatory networks. Nat Rev Genet 10: 141–148

Bharti K, Von Koskull-Döring P, Buning J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21

Bowman JL et al. (2013) Overexpression of AtHsfB4 induces specific effects on root development of Arabidopsis. Mech Dev 130: 54–60

Bharti K, Von Koskull-Döring P, Bharti S, Kumar P, Tintsch-Körbitzer A, Treuter E, Nover L (2004) Tomato heat stress transcription factor HsfB1 represents a novel type of general transcription coactivator with a histone-like motif interacting with the plant CREB binding protein ortholog HAC1. Plant Cell 16: 1521–1535

Bowman JT et al. (2017) Insights into land plant evolution garnered from the Marchantia polymorpha genome. Cell 171: 287–304.e15

Capra JA, Singh M (2007) Predicting functionally important residues from sequence conservation. Bioinformatics 23: 1875–1882

Chalandon G, Ravarani CNJ, Balaji S, Martínez-Arias A, Aravind L, Jothi R, Babu MM (2012) Interplay between gene expression noise and regulatory network architecture. Trends Genet 28: 221–232

Cheng Y, Dai X, Zhao Y (2007) Auxin synthesized by the YUCCA flavin monoxygenases is essential for embryogenesis and leaf formation in Arabidopsis. Plant Cell 19: 2430–2439

Cortijo S, Charoeasawan V, Brestovitsky A, Buning R, Ravarani C, Rhodes D, van Noort J, Jaeger KE, Wigge PA (2017) Transcriptional regulation of the ambient temperature response by H2A.Z nucleosomes and HSF1 transcription factors in Arabidopsis. Mol Plant 10: 1258–1273

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21

Flores-Sandoval E, Romani F, Bowman JL (2018) Co-expression and transcriptome analysis of marchantia polymorpha transcription factors supports class c afars as independent actors of an ancient auxin regulatory module. Front Plant Sci 9: 1345

Fragkostefanakis S, Simm S, El-Shershaby A, Hu Y, Bublak D, Mesihovic A, Darm K, Mishra SK, Tschiersch B, Theres K et al. (2019) The repressor and co-activator HsfB1 regulates the major heat stress transcription factors in tomato. Plant Cell Environ 42: 874–890

Fragkostefanakis S, Mesihovic A, Simm S, Paupière MJ, Hu Y, Paul P, Mishra SK, Tschiersch B, Theres K, Bovy A et al. (2016) HsfA2 controls the activity of developmentally and stress-regulated heat stress protection mechanisms in tomato male reproductive tissues. Plant Physiol 170: 2461–2477

Gautier L, Cole P, Bolstad BM, Iriarry RA (2004) Affy - Analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307–315

Gautier L, Chee M, Zhang Z, Venter E, Church GM (2001) waveplots: a sequence logo generator. Genome Res 11: 1446–1455

Guindon S, Lethiec F, Duroux P, Gascuel O (2006) PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res 33: W557–W559

Hare PD, Cress WA (1997) Metabolic implications of stress-induced proline accumulation in plants. Plant Growth Regul 21: 79–102

Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M (2013) Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. Int J Mol Sci 14: 9643–9684

Ikeda M, Mitsuda N, Ohme-Takagi M, Tsuchimoto M (2011) Arabidopsis HsfB1 and HsfB2b act as repressors of the expression of heat-inducible HsfSs but positively regulate the acquired thermotolerance. Plant Physiol 157: 1243–1254

Ishizaki K, Nishihama R, Yamato KT, Kohchi T (2016) Molecular genetic tools and techniques for Marchantia polymorpha research. Plant Cell Physiol 57: 262–270

Julca I, Ferrari C, Flores-Tornero M, Proost S, Lindner AC, Hackenberg D, Steinbachová L, Michaelidis C, Pereira SG, Misra CS et al. (2021) Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. Nat Plants 7: 1143–1159

Keller M, Hu Y, Mesihovic A, Fragkostefanakis S, Schleiff E, Simm S (2017) Alternative splicing in tomato pollen in response to heat stress. DNA Res 24: 205–217

Kotak S, Larkindale J, Lee U, von Koskull-Döring P, Vierling E, Scharf KD (2007) Complexity of the heat stress response in plants. Curr Opin. Plant Biol 10: 310–316

Krizek BA (2011) Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family. J Exp Bot 62: 3311–3319
Kubota A, Ishizaki K, Hosaka M, Kohchi T (2013) Efficient Agrobacterium-mediated transformation of the liverwort Marchantia polymorpha using regenerating thalli. Biosci Biotechnol Biochem 77: 167–172

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359

Larkindale J, Vierling E (2008) Core genome responses involved in acclimation to high temperature. Plant Physiol 146: 323–324

Law CW, Chen Y, Shi W, Smyth GK (2008) Diversity, expansion, and evolutionary novelty of plant DNA-binding transcription factor families. Biochim Biophys Acta – Gene Regul Mech 1860: 3–20

Lewis LA, McCourt RM (2004) Green algae and the origin of land plants. Am J Bot 91: 1535–1556

Lisse J, Schauer N, Kopka J, Willmitzer L, Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc 1: 387–396

Liu HC, Chang Y (2013) Common and distinct functions of Arabidopsis class A1 and A2 heat shock factors in diverse abiotic stress responses and development. Plant Physiol 163: 276–290

Liu HC, Liao HT, Chang Y (2011) The role of class A1 heat shock factors (HSFAs) in response to heat and other stresses in Arabidopsis. Plant Cell Environ 34: 738–751.

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

Ma C, Xin M, Feldmann KA, Wang X (2014) Machine learning–based differential network analysis: a study of stress-responsive transcriptomes in Arabidopsis. Plant Cell 26: 520–537

Marshall-Colón A, Kliebenstein DJ (2019) Plant networks as traits and hypotheses: moving beyond description. Trends Plant Sci 24: 840–852

Mishra U, Rai A, Kumar R, Singh M, Pandey HP (2016) Gene expression analysis of Solanum lycopersicum and Solanum habrochaites under drought conditions. Genomics Data 9: 40–41

Mittler R, Finka A, Goloubinoff P (2012) How do plants feel the heat? Trends Biochem Sci 37: 118–125

Monte I, Kneesaw S, Franco-Zorrilla JM, Chini A, Zamarreno AM, Garcia-Mina JM, Solano R (2020) An ancient COI1-independent function for reactive electrophilic oxylipins in thermotolerance. Curr Biol 30: 962–971.e3

Montgomery SA, Tanizawa Y, Yamato KT, Liu C, Ed F, Correspondence EB (2020) Chromatin organization in early land plants reveals an ancestral association between H3K27me3, transposons, and constitutive heterochromatin. Curr Biol 30: 573–588.e7

Myouga F, Motohashi R, Kuromori T, Nagata N, Shinozaki K (2006) An Arabidopsis chloroplast-targeted Hsp101 homologue, APG6, has an essential role in chloroplast development as well as heat-stress response. Plant J 48: 249–260

Neudegger T, Vergheze J, Hayer-Hartl M, Hartl FU, Bracher A (2016) Structure of human heat-shock transcription factor 1 in complex with DNA. Nat Struct Mol Biol 23: 140–146

Nuno-Cabanes C, Ugidos M, Tarazona S, Martin-Exposito M, Ferrer A, Rodriguez-Navarro S, Conesa A (2020) A multi-omics dataset of heat-shock response in the yeast RNA binding protein Mip6. Sci Data 7: 69

O’Malley RC, Huang SSC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. Cell 165: 1280–1292

Ohama N, Sato H, Shinozaki K, Yamaguchi-Shinozaki K (2017) Transcriptional regulatory network of plant heat stress response. Trends Plant Sci 22: 53–65

Pajerowska-Mukhtar KM, Wang W, Tada Y, Oka N, Tucker CL, Fonseca JP, Dong X (2012) The Hsf-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. Curr Biol 22: 103–112

Paupière MJ, Tikunov Y, Schleiff E, Bovy A, Fragnostefanakis S (2020) Reprogramming of tomato leaf metabolome by the activity of heat stress transcription factor HsfB1. Front Plant Sci 11: 1–10

Pulli K, Baksheev A, Kornyakov K, Erhunov V (2012) Mobile computer-vision technology will soon become as ubiquitous as touch interfaces. Commun ACM 55: 61–69

Ray J, Munn PR, Vihervaara A, Lewis JJ, Ozer A, Danko CG, Lis JT (2019) Chromatin conformation remains stable upon extensive transcriptional changes driven by heat shock. Proc Natl Acad Sci USA 116: 19431–19439

Raza A (2020) Metabolomics: a systems biology approach for enhancing heat stress tolerance in plants. Plant Cell Rep 41: 741–763

Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochim Biophys Acta – Gene Regul Mech 1819: 104–119

Scheppe P, Zhang P, Wang C, Kim T, Banf M, Chae L, Dreher K, Chvala AK, Nilo-Poyanco R, Bernard T et al. (2017) Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. Plant Physiol 173: 2041–2059

Schmollinger S, Schulz-Raffelt M, Strenkert D, Veyel D, Vallon O, Schroda M (2013) Dissecting the heat stress response in Chlamydomonas by pharmaceutical and RNAi approaches reveals conserved and novel aspects. Mol Plant 6: 1795–1813

Schroda M, Hemme D, Mühltäus T (2015) The Chlamydomonas heat stress response. Plant J 82: 466–480

Smyth GK (2005) limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer, pp 397–420

Stergachis AB, Nep S, Sandstrom R, Haugen E, Reynolds AP, Zhang M, Byron D, Canfield T, Stelhing-Sun S, Lee K et al. (2014) Conservation of trans-acting circuitry during mammalian regulatory evolution. Nature 515: 365–370

Villarreal A JC, Crandall-Stotler BJ, Hart ML, Long DG, Forrest LL (2016) Divergence times and the evolution of morphological complexity in an early land plant lineage (Marchantiaipsida) with a slow molecular rate. New Phy tol 209: 1734–1746

Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3: 2–20

Voordeckers K, Pougach K, Verstrepen KJ (2015) How do regulatory networks evolve and expand throughout evolution? Curr Opin Biotechnol 34: 180–188

Vosseberg J, van Hoof JJE, Marcet-Houben M, van Vlimmeren A van, Wijk LM, Gabaldón T, Snel B (2021) Timeing the origin of eukaryotic cellular complexity with ancient duplications. Nat Ecol Evol 5: 92–100

de Vries J, Archibald JM (2018) Plant evolution: landmarks on the path to terrestrial life. New Phy tol 217: 1428–1434

de Vries J, de Vries S, Curtis BA, Zhou H, Penny S, Feussner K, Voordeckers K, Pougach K, Verstrepen KJ et al. (2020) Heat stress response in the closest algal relatives of land plants reveals conserved stress signaling circuits. Plant J 103: 1025–1048

Waese J, Fan J, Pasha A, Yu H, Fucile G, Shi R, Cumming M, Kelley LA, Sternberg MJ, Krishnakumar V et al. (2017) ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant biology. Plant Cell 29: 1806–1821

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. Front Plant Sci 9: 1–11

Wilhelmsen PKI, Mühlich C, Ullrich KK, Rensing SA (2017) Comprehensive genome-wide classification reveals that many plant-specific transcription factors evolved in streptophyte algae. Genome Biol Evol 9: 3384–3397
Wu HW, Deng S, Xu H, Mao HZ, Liu J, Niu QW, Wang H, Chua NH (2018) A noncoding RNA transcribed from the AGAMOUS (AG) second intron binds to CURLY LEAF and represses AG expression in leaves. New Phytol 219: 1480–1491

Wu T-Y, Goh H, Azodi CB, Krishnamoorthi S, Liu M-J, Urano D (2021) Evolutionarily conserved hierarchical gene regulatory networks for plant salt stress response. Nat Plants 7: 797–799

Wu TY, Krishnamoorthi S, Goh H, Leong R, Sanson AC, Urano D (2020) Crosstalk between heterotrimeric G protein-coupled signaling pathways and WRKY transcription factors modulating plant responses to suboptimal micronutrient conditions. J Exp Bot 71: 3227–3239

Wu Z, Irizarry RA, Zhang L, Wu C, Carta R, Baggerly K, Coombes KR (2004) Preprocessing of oligonucleotide array data [5] (multiple letters). Nat Biotechnol 22: 656–658

Yoshida T, Ohama N, Nakajima K, Kidokoro S, Mizoi J, Nakashima K, Maruyama K, Kim JM, Seki M, Todaka D et al. (2011) Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. Mol Genet Genomics 286: 321–332

Yu G, Wang LG, He QY (2015) ChiP seeker: an R/Bioconductor package for ChiP peak annotation, comparison and visualization. Bioinformatics 31: 2382–2383

Zhang X, Rerksiri W, Liu A, Zhou X, Xiong H, Xiang J, Chen X, Xiong X (2013) Transcriptome profile reveals heat response mechanism at molecular and metabolic levels in rice flag leaf. Gene 530: 185–192

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W et al. (2008) Model-based analysis of ChiP-Seq (MACS). Genome Biol 9: R137