JIPB Journal of Integrative Plant Biology

Research Article
https://doi.org/10.1111/jipb.12990

Immediate transcriptional responses of *Arabidopsis* leaves to heat shock

Min Liu[1†], Jiafu Zhu[1,2†] and Zhicheng Dong[1*]

1. Innovative Center of Molecular Genetics and Evolution, Guangzhou Higher Education Mega Center, School of Life Sciences, Guangzhou University, Guangzhou 510006, China
2. Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, South China Botanical Garden, the Chinese Academy of Sciences, Guangzhou 510650, China
†These authors contributed equally to this article.
*Correspondence: Zhicheng Dong (zc_dong@gzhu.edu.cn)

Min Liu

Zhicheng Dong

ABSTRACT

Plants have evolved efficient mechanisms for adapting to temperature fluctuations, known as heat stress response and heat stress memory. Although the transcriptional regulatory network of plant heat stress response has been established, little is known about the genome-wide transcriptional changes occurring within the first several minutes after heat shock. Here, we investigated the nascent RNA and mature messenger RNA (mRNA) from plant leaf tissues exposed to 5 min of heat shock treatment using global run-on sequencing and RNA sequencing methods. Only a small group of genes were up- or downregulated at both the nascent RNA and mRNA levels. Primed plants that were already exposed to mild heat stress exhibited a more drastic alteration at multiple transcriptional steps than naïve plants that had not experienced heat stress. Upon heat shock, we also observed the following: (i) engaged RNA polymerase II accumulated downstream of transcription start sites; (ii) 5′ pausing release was a rate-limiting step for the induction of some heat shock protein genes; (iii) numerous genes switched transcription modes; (iv) pervasive read-through was induced at terminators; and (v) heat stress memory occurs at multiple steps of the transcription cycle, such as at Pol II recruitment, 5′ pausing, elongation, and termination.

Keywords: Transcription, Nascent RNA biogenesis, Arabidopsis, Heat Shock, Stress Memory

Liu, M., Zhu, J., and Dong, Z. (2021). Immediate transcriptional responses of *Arabidopsis* leaves to heat shock. J. Integr. Plant Biol. **63:** 468–483.

INTRODUCTION

Unlike animals, plants are sessile organisms, that cannot change their location to avoid stresses, such as drought, salinity and extreme temperature, but they have evolved different adaptive ways to tolerate these stresses. Temperature is a fundamental factor affecting plant vegetative and reproductive development (Hedhly et al., 2009), such as hypocotyl elongation, flowering, pollination, and germination (Herrero and Johnson, 1980; Covell et al., 1986; Gray et al., 1998; Fitter and Fitter, 2002). Recent global climate change, especially global warming, has compromised crop yields (Lesk et al., 2016).

Biochemically, high temperature triggers the denaturation of protein, destabilization of the cell membrane and production of reactive oxygen species (ROS) (Hasanuzzaman et al., 2013). To protect cells from these adverse changes, plants alter their gene expression upon heat shock (HS). Genes involved in protein folding, osmotic balancing, and detoxification from oxidants are quickly induced. Heat shock proteins (HSPs), known as molecular chaperones, function in the stabilization of proteins and membranes to alleviate the damage caused by HS (Wang et al., 2004). Catalase and ascorbate peroxidase, ROS scavenging enzymes, are also induced by heat. Recently, a transcriptional regulatory network was established in which HSFA1s were demonstrated

to be the central transcription factors (TFs) in plants (Ohama et al., 2017). Upon HS, HSFA1s were proposed to directly transactivate downstream TF genes as well as *HSP*s, triggering a transcriptional cascade of heat stress response.

Moreover, abiotic stresses are usually repetitive, and an enhanced response to recurring stresses may promote plant fitness. Several studies indicated that a previous exposure to a stressor may alter a plant's subsequent stress response by producing faster and/or stronger reactions that may improve the plant's stress tolerance (see Bruce et al., 2007; Avramova, 2015 for reviews). Altered responses to consecutive stresses imply that plants exercise a form of "stress memory". Epigenetic mechanisms, such as chromatin structure remodeling, histone modification, microRNA, as well as certain TFs and Pol II are thought to be involved in the "stress memory" of plants (Ding et al., 2012; Sani et al., 2013; Stief et al., 2014; Mozgová et al., 2015; Brzezinka et al., 2016; Lämke et al., 2016; Liu et al., 2019).

However, previous studies used steady-state messenger RNA (mRNA) levels as a readout (Cortijo et al., 2017; Ohama et al., 2017; Liu et al., 2018; Li et al., 2019), which restrains our understanding of direct transcriptional heat stress response as well as heat stress memory, especially regarding the immediate transcriptional response to HS. In this study, we investigated the first wave of transcriptomic responses to HS at the nascent RNA level

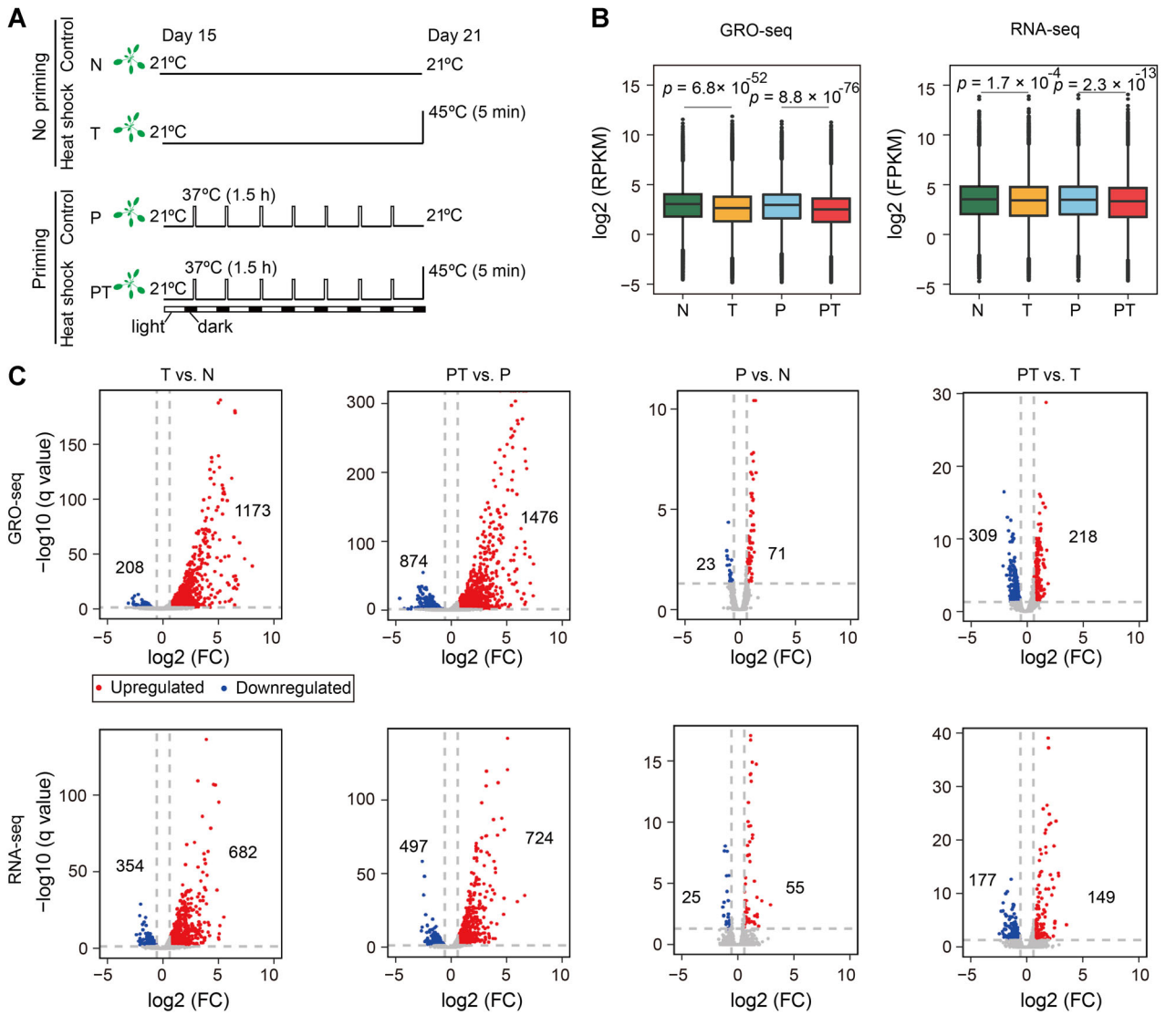

Figure 1. Heat shock triggers rapid and extensive changes in transcription

(A) Overview of four treatment setups. "N," "T," "P," and "PT" represent "Naïve," "Triggered," "Primed," and "Primed and Triggered," respectively. **(B)** Boxplots of the overall transcription levels among the four treatment groups at the nascent RNA level (left) and mature RNA level (right). The line in the box represents the median value, and the upper and lower whiskers represent 75% and 25% of the data, respectively. Two-sided *P* values for the Wilcoxon signed-rank test are presented in the plots. **(C)** Volcano plots for differential gene expression in different treatment groups at the nascent RNA level (upper) and messenger RNA (mRNA) level (below). "FC", represents "fold change". The red and blue dots represent up- and downregulated genes, respectively.

using global run-on sequencing (GRO-seq) (Zhu et al., 2018) on 5-min heat-shocked plants.

RESULTS

GRO-seq and RNA-seq revealed different changes in mRNA metabolism in very early responses to HS

To explore how *Arabidopsis* respond and adapt to HS at the transcriptional level, we treated 2-week old plants grown at 21°C under 16 h light/8 h dark cycles as four groups: Naïve (N), Triggered (T), Primed (P), and Primed and Triggered (PT) (Figure 1A). For N and T groups, plants were cultured at 21°C for one more week, and then leaf tissues were collected in dark at 21°C (N) or after a HS at 45°C for 5 min (T). While for P and PT groups, plants were exposed to a heat treatment for 1.5 h at 37°C before dawn each day for 1 week, and leaf tissues were collected in dark at 21°C (P) or after a HS at 45°C for 5 min (PT). After HS, the temperature of the leaf surface was about 28°C as measured by an infrared thermometer in comparison to 21°C from the control plant sets. Therefore, after the 5-min 45°C treatment, the T and PT plants experienced a quick ambient temperature increase but did not reach a lethal temperature. The tissues were then frozen with liquid nitrogen and stored in a −80°C freezer.

To elucidate the landscape of transcriptional change after the very early responses to HS, we adopted our recently published GRO-seq method (Zhu et al., 2018), which allowed for the analysis of multiple frozen samples at one time. GRO-seq revealed transcriptionally engaged RNA polymerase distribution from a sub-gene to a genome scale. We also performed poly(A) enrichment RNA-seq to measure the steady-state level of mature mRNA as comparison. Two biological replicates were reproducible (Tables S1 and S2; Figure S1).

We assumed that the density of a GRO-seq signal in gene body (GB) is a good quantitative measure for the gene's nascent RNA transcription activity. A short period of HS reduced the overall transcription of both the T and PT plants genome-wide, which was less obvious at the mRNA level (Figures 1B, S2). Actively transcribed protein coding genes (n = 18,432; genes with normalized reads density RPKM (reads per kilobase per million mapped reads) >1 in GRO-seq or FPKM (fragments per kilobase of exon model per million mapped fragments) >1 in RNA-seq at least in one treatment) were used to detect different expression genes (DEGs). Within a 5 min exposure to HS, 1,173 and 1,476 genes exhibited increased transcription in the T and PT plants, respectively, as determined by GRO-seq, while only 682 and 724 genes of mRNA expression levels were increased by 1.5-fold (Figure 1C; Table S3). Although there was a genome-wide reduction of transcription activity after HS, the number of significantly downregulated genes is smaller than that of upregulated genes (208 in T vs. N and 874 in PT vs. P at nascent RNA level; 354 in T vs. N and 497 in PT vs. P at mRNA level) (Figure 1C). Thus, it is suggested that the rate of down-regulation is slower than that of upregulation, and a longer HS would be expected to trigger more downregulated genes as detected in a previous study in *Arabidopsis* as well as in mammalian cells (Mahat et al., 2016; Cortijo et al., 2017). In addition, priming did not change the transcription activity a lot (P vs. N, Figure 1C). It is indicated that most genes triggered by a high temperature could recover after 24 h, especially at the nascent RNA level (Figure S3). Three genes (*AT1G07135*, *AT1G14200*, and *HSP70/AT3G12580*) were picked for further validation using quantitative reverse transcription polymerase chain reaction (PCR) at chromatin-associated RNA and mRNA levels (Figure S4).

Heat shock triggered transcriptional changes in different categories of genes

Alteration of nascent RNA synthesis precedes changes in mRNA levels. Therefore, GRO-seq is more sensitive than RNA-seq for detecting quick transcriptional changes in response to ambient temperature increases. Furthermore, only a small number of genes were up- or downregulated at both the nascent RNA and mRNA levels (Figure 2A).

We next asked which functional groups of genes were quickly changed. Gene Ontology (GO) enrichment analysis indicated that heat, protein folding, light, oxidative stress response genes as well as some other biotic/abiotic response genes were enriched in only the upregulated nascent RNA category or in both the nascent RNA and mRNA categories in T/N and PT/P (Figure 2B, upper panel). Stress-involved hormones, such as salicylic acid and jasmonic acid were also enriched. However, some other terms were enriched in only upregulated mRNA genes, such as response to auxin, carbohydrates, and amino acid metabolism. Therefore, both naïve and primed plants actively increased the transcription of stress-responsive genes upon HS, and upregulation of mRNA encoding auxin-responsive, carbohydrate, and amino acid metabolism genes might result from an increase in their stability.

Among the downregulated genes, only circadian rhythm and cytokinin-responsive genes were enriched for T/N, while genes involved in primary and secondary metabolism, hormone signaling, nutrients response, and so on were enriched for PT/P at nascent RNA level. Genes downregulated at mRNA level in both T/N and P/PT are associated with nucleotide, ribose phosphate, glucan, starch biosynthesis, and so on (Figure 2B, lower panel).

HSFA1s have been reported to be “master regulators” that are indispensable in the activation of transcriptional networks (Mishra et al., 2002; Liu et al., 2011; Yoshida et al., 2011). However, aside from *HSFA1E* being slightly upregulated at the nascent RNA level in T plants, we failed to detect transcription upregulation at either the nascent RNA level or the mRNA level for *HSFA1*s upon 5-min HS treatment (Figures 2C and S5). Upon HS, HSFA1s induce the expression of TF genes, such as *DREB2A*, *MBFC1*, *HSFA7C*, *HSFA2*, and *HSFB*s, to fine-tune the expression of HS-inducible genes (Yoshida et al., 2011). Most TFs activated by HFSA1s were significantly upregulated at both the nascent RNA level and the mRNA level (Figures 2C and S5). These results suggest that *HSF1A*s were activated at the

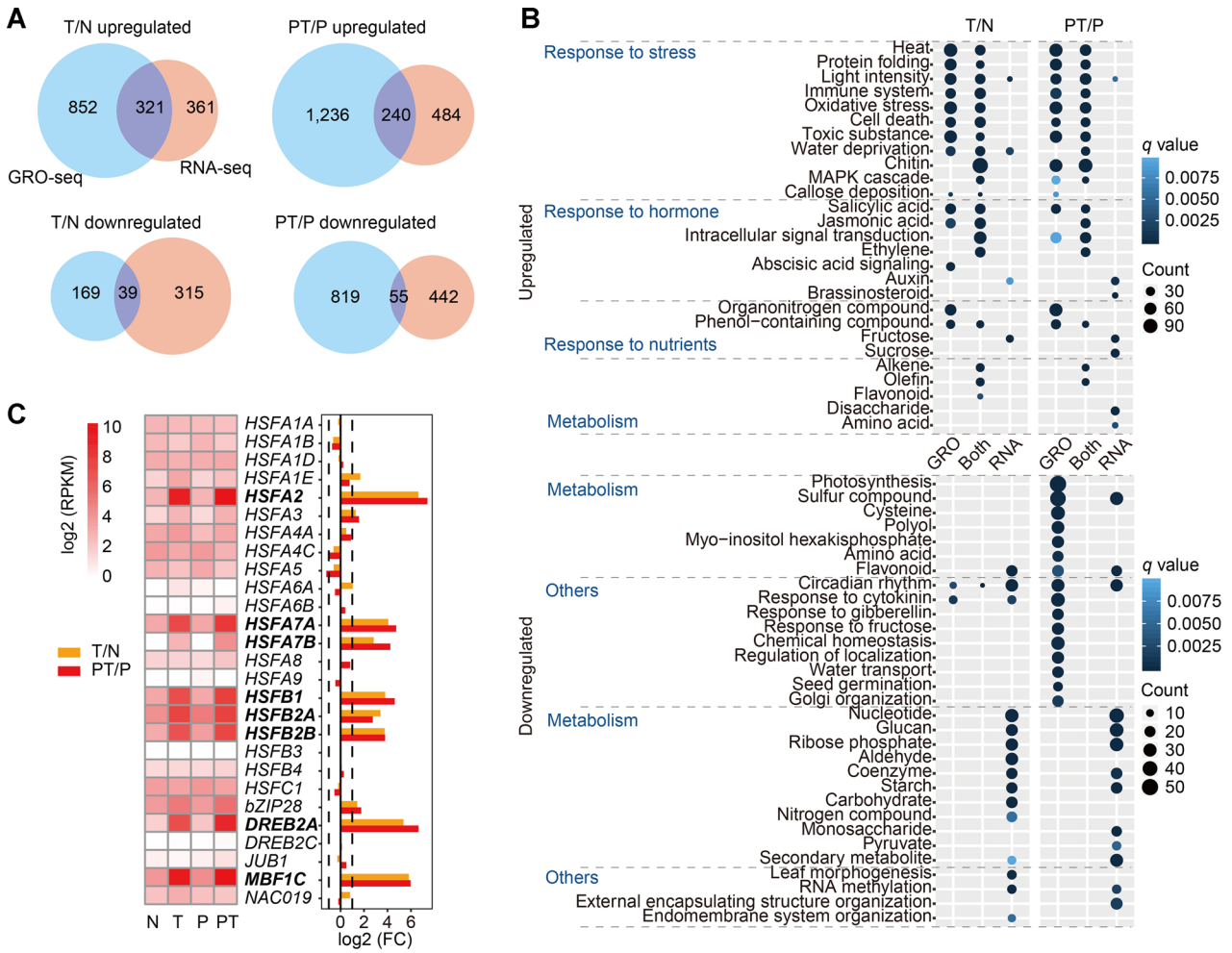

Figure 2. Different expression genes (DEGs) at nascent RNA and messenger RNA (mRNA) levels

(A) Venn diagrams of up- (upper) and downregulated (below) genes after heat shock (HS) as detected by global run-on sequencing (GRO-seq) and RNA-seq in T vs. N and PT vs. P. **(B)** Gene Ontology (GO) terms enriched in up- and downregulated genes among the different groups: "in only GRO-seq," "in both GRO-seq and RNA-seq," and "in only RNA-seq." The size of the dot represents the gene count. A hypergeometric test was used for statistical analysis, and the *P* values from the tests were converted to false discovery rate (FDR)-corrected *q* values. **(C)** The transcriptional activity of transcription factor genes involved in the transcriptional regulatory network in response to HS.

translation and/or post-translation levels but not at the transcription level upon 5-min HS treatment, which is consistent with previous findings (Yamada et al., 2007; Liu et al., 2008; Hahn et al., 2011; Ohama et al., 2016). To explore whether there are different TFs involved between primed and naïve plants upon HS, we adopted data of target genes of 387 TFs belonging to 25 families as detected by DNA affinity purification-seq (O'Malley et al., 2016). We found that the upregulated genes were targeted by each TF family to a similar extent, although the upregulated gene sets were somehow different in T vs. N and PT vs. P at the nascent RNA level (Figure S6).

Distinct types of heat stress memory genes in primed plants

Pre-exposure to biotic or abiotic stress can lead to enhanced defense responses to recurring stresses (Avramova, 2015). And stress memory can be stored at the transcriptional level (Ding et al., 2012; Ling et al., 2018; Liu et al., 2018). However, the previously reported stress memory genes are defined at the mRNA level. Transcriptional memory can be either sustained changes in expression or faster/stronger response when exposed to a second stimulus (Lämke and Bäurle, 2017). Thus, we considered genes that were up- or downregulated in PT vs. T and/or P vs. N at the nascent RNA and/or mRNA levels as heat stress memory genes. We surveyed the 845 genes that were up- or downregulated in PT vs. T and/or P vs. N, among which 585 and 373 genes were detected at the nascent RNA and mRNA levels, respectively (Figure 3A). However, only 113 genes were shared. Hierarchical clustering of these 113 memory genes reveals four major clusters (Figure 3B and C). Genes in cluster 1 (C1, $n = 25$) showed similar patterns between nascent RNA and mRNA. These genes are the well-known memory genes, including *HSFA7A* and several *HSPs* (Lämke et al., 2016; Grinevich et al., 2019), the mRNA of which were upregulated when plants were exposed to HS, and a higher upregulation

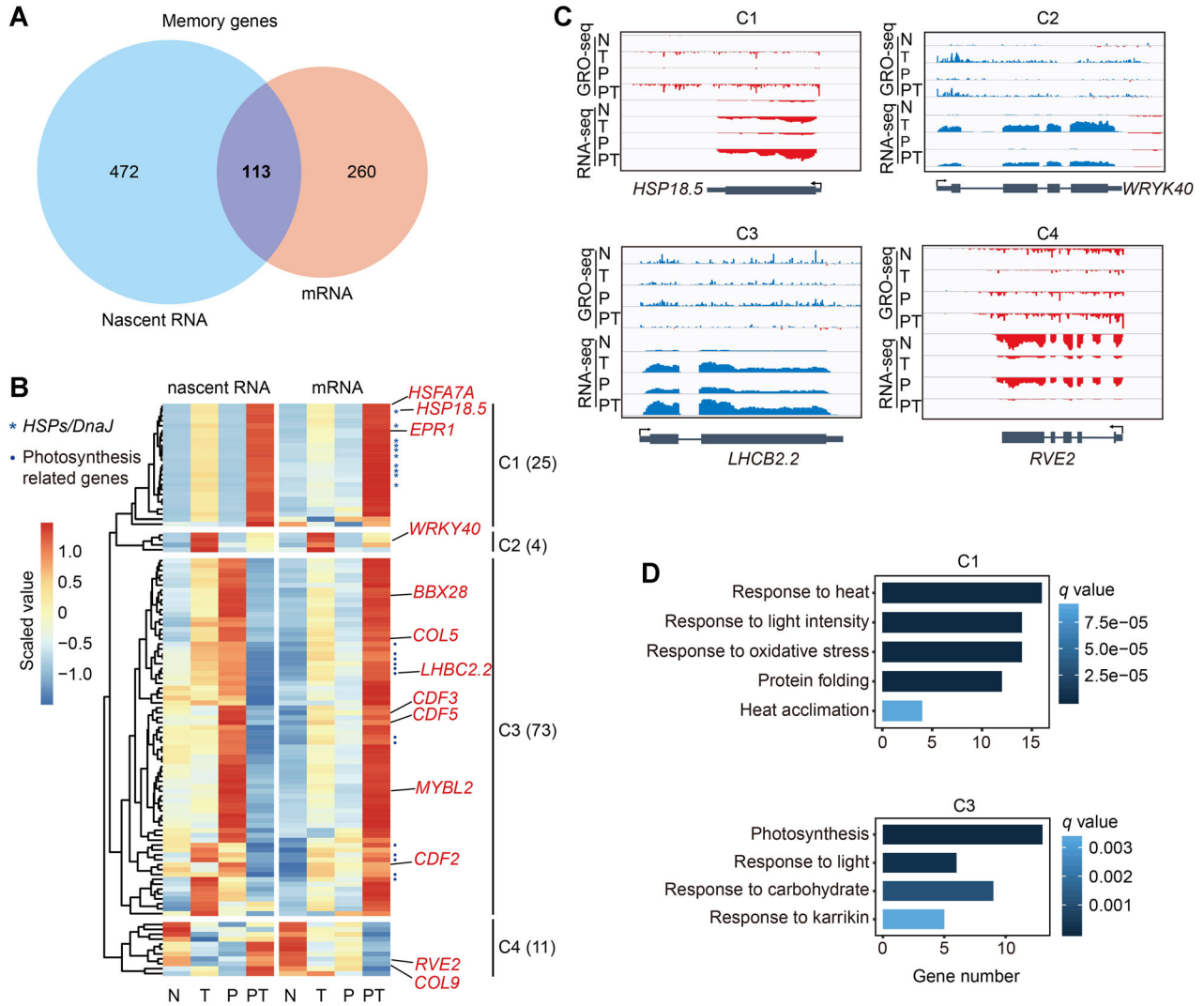

Figure 3. Heat stress memory genes

(A) Venn diagram of memory genes (up- or downregulated in P vs. N and/or PT vs. T) in global run-on sequencing (GRO-seq) and RNA-seq. **(B)** Heatmaps of the 113 common memory genes. Genes were clustered into four groups by hierarchical clustering method based on the scaled values of read density in the four treatment groups. **(C)** Screenshots of transcriptional activity of representative genes in four groups. **(D)** Gene Ontology (GO) terms enriched in clusters 1 and 3.

was observed when primed plants were exposed to HS. Gene Ontology enrichment analysis indicated that genes involved in response to heat, light intensity, oxidative stress, and protein folding were significantly enriched (Figure 3D). Genes in a relatively small cluster 2 (C2, $n = 4$) showed a lower upregulation in primed plants compared with naïve plants when exposed to HS. A TF involved in stress response, *WRKY40* was detected in this cluster. Genes in cluster 3 (C3, $n = 73$) showed different transcription pattern in nascent RNA and mRNA. Genes involved in photosynthesis, response to light and carbohydrate were significantly enriched (Figure 3D). Notably, some stress response TFs, such as *CDFs* and *MYBL2* (Wang et al., 2016; Corrales et al., 2017), were found in this cluster. Most of them were more actively transcribed in P and accumulated the highest mRNA level in PT, suggesting that they were transcriptionally induced at the transcription level within shorter than 5 min and subside at 5 min in primed plants. Therefore, these genes could be the very first wave genes in respond to HS. The cluster 4 (C4, $n = 11$) represents the genes downregulated at mRNA in PT with various transcription activity in four treatment groups. Six genes were picked from C3 and C4 for quantitative reverse transcription PCR validation, and most of them showed a similar expression pattern to GRO-seq or RNA-seq (Figure S7).

From the transcriptional activity of all 845 memory genes, we observed different patterns (Figure S8). Some genes showed concordant changes upon HS at nascent RNA and mRNA levels, while others are not. For the transcriptional memory genes ($n = 279$) that transcribed highest in PT at the nascent RNA level, we detected three mRNA patterns with the highest expression detected either in N, T, or PT,

respectively. However, GO term of response to heat is enriched in all three patterns, suggesting a complicate transcriptional and post-transcriptional regulation of heat stress memory genes.

Heat shock triggered the retention of Pol II at the transcription start site

GRO-seq not only detected the transcriptional activities of genes but also revealed changes in different transcriptional stages after HS. In addition to the transcriptional activity, a GRO-seq density peak infers a rate-limiting step of transcription, such as 5′ pausing downstream of the transcription start site. The 5′ pause index (5′ PI) is an indicator of the degree of Pol II stalling in the proximal promoter. Previously, we showed that approximately 38.4% of active genes pause at the proximal promoter region in young seedlings (Zhu et al., 2018). We then asked whether the 5′ pausing was altered by ambient temperature. To avoid interference from neighboring gene signals, we focused on genes that were >1 kb away from the neighboring gene on the same strand and had a length >1 kb. Meta-profile analyses showed that 5-min HS resulted in more Pol II accumulation in the proximal promoter region and less accumulation in the GB, especially for PT plants (Figure 4A). The overall 5′ PI was increased

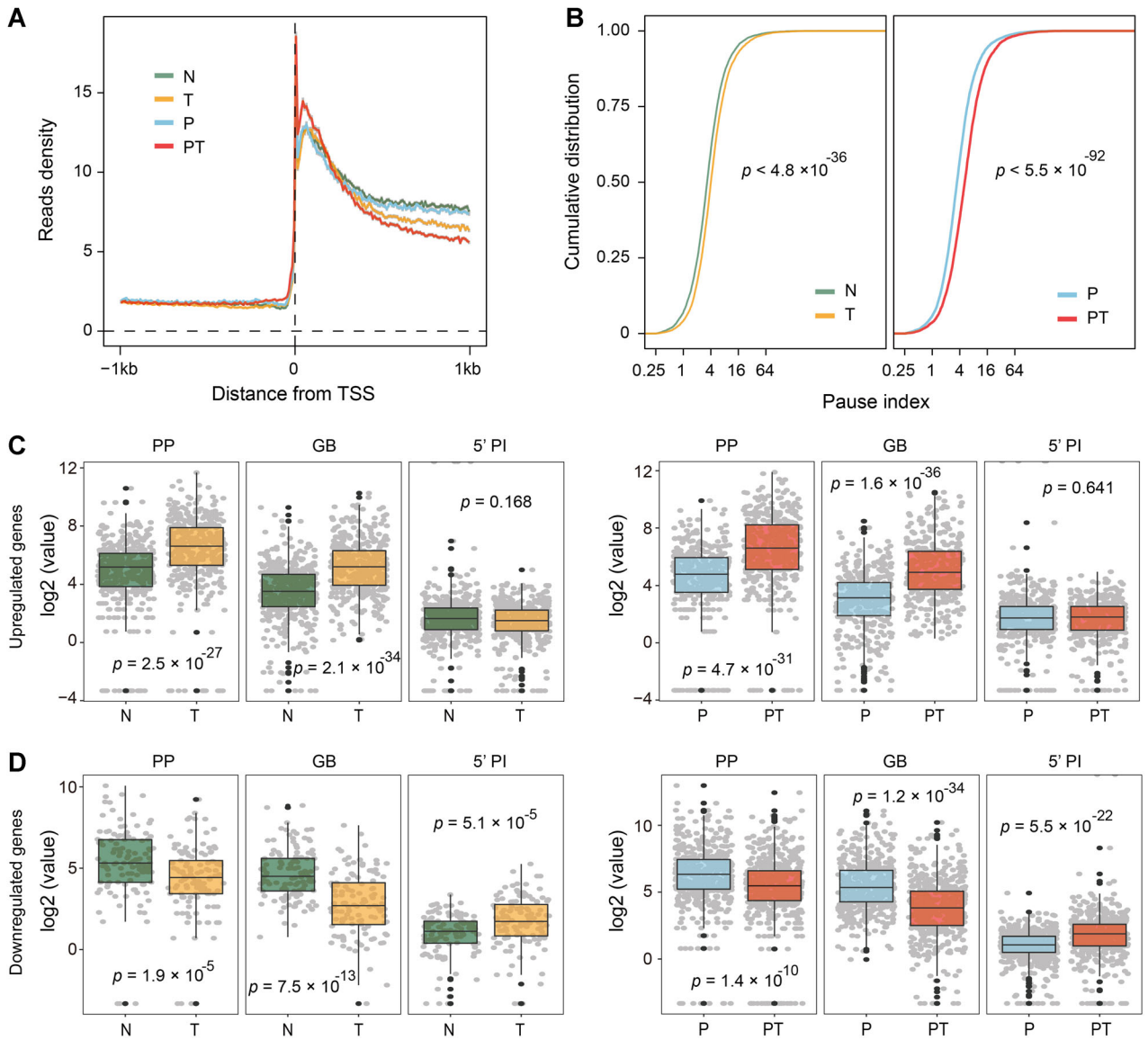

Figure 4. Promoter-proximal pausing release as a key step for transcriptional regulation upon heat shock (HS)

(A) Meta-profile of the read density around the transcription start site (TSS). Lines and shading represent the mean ± S.E.M. for each bin. **(B)** Comparison of the cumulative curves of the 5′ PI between N and T (left) and between P and PT (right). Two-sided *P* values of the Kolmogorov-Smirnov test are presented in the plots. **(C)** Comparison of the read density in the promoter-proximal region (PP), the read density in the gene body (GB), and the 5′ PI for upregulated genes under HS (N vs. T and P vs. PT). **(D)** Comparison of the read density in the PP region, read density in the GB, and PI for downregulated genes under HS (N vs. T, and P vs. PT).

significantly in T and PT plants (Figure 4B). For upregulated genes, the Pol II density in both the proximal promoter and the GB was increased in T and PT plants; however, this result was not observed for 5′ PI, indicating that the increase in Pol II recruitment and elongation contributes to transcriptional escalation (Figure 4C). For transcriptionally downregulated genes, the Pol II density in both the proximal promoter and the GB was decreased, while the 5′ PI was increased. Therefore, less Pol II was recruited to the proximal promoter and released from the pausing state for productive elongation, accounting for the transcriptional decrease (Figure 4D). Taken together, these results suggest that upon HS a portion of 5′ pausing Pol II fails to enter productive elongation.

In our previous study, run-ons were performed at both 22°C and 30°C with the nuclei isolated from seedling grown at 21°C, followed by high-throughput sequencing of nascent RNA (Zhu et al., 2018). No significant difference of Pol II profiling was observed from GRO-seq between 22 °C and 30 °C run-on reactions. Further analysis showed that few genes were identified as differentially transcribed at higher temperature (30 °C) than the normal growth temperature (22 °C) (Figure S9A, B), exemplified as *HSP70* and *HSFA2* (Figure S9C, D). Therefore, that heat triggered gene transactivation and retention of pausing state of Pol II requires integrative cell function, which further suggests that Pol II and chromatin interaction upon HS needs more factors such as signaling transduction that might get lost during the nuclei isolation.

Pausing states transition among the four treatments

To monitor the changes in transcription modes from N/P to T/PT plants, we further classified genes into four categories (Figure 5A and Methods): type I (active, non-paused), active transcription in the GB without a significant enrichment peak in the proximal promoter region; type II (active, paused), active transcription in the GB with a significant pause in the proximal promoter region; type III (non-active, paused), significantly enriched in the proximal promoter region without active transcription in the GB; type IV (non-active, non-paused), inactive transcription without a significant enrichment in the proximal promoter region.

In N and P plants, the proportions of genes within each class were similar (Figure 5B). After exposure to heat, the number of type I genes was reduced from 5 246/5 183 (N/P) to 4 786/3 886 (T/PT), and the number of inactive genes (type III and IV) was increased from 5,666/5,784 (N/P) to 6,124/6,091 (T/PT), suggesting an overall suppressive effect of HS on gene transcriptional activity (Figure 5B), consistent with a global transcriptional downregulation upon HS (Figure 1B). Despite the decrease in the active gene numbers, the number of type II genes increased from 2,057 (P) to 3,047 (PT), more likely caused by an increase in 5′ PI.

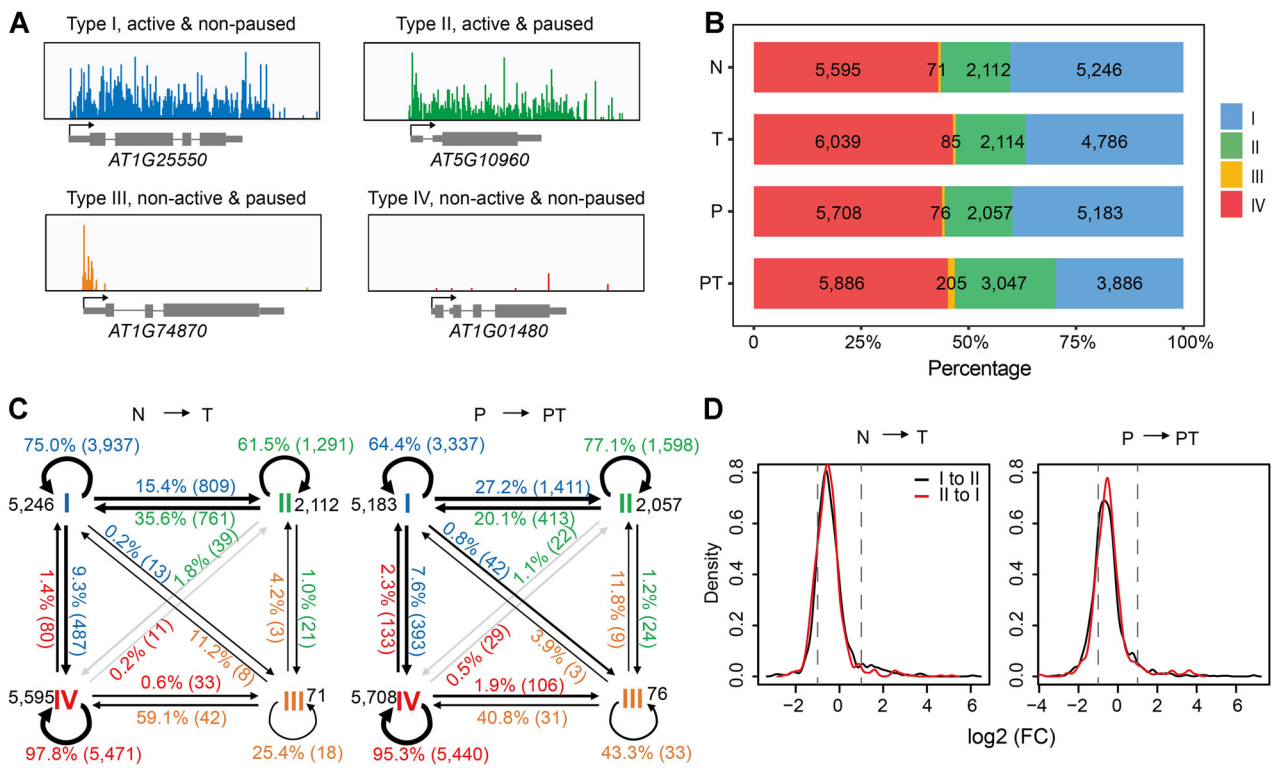

Figure 5. Pol II state transformation under heat shock

(A) Four types of genes were defined based on the relative read density of the promoter-proximal peak relative to the GB. **(B)** Number of each gene type in four treatment groups. **(C)** Proportions and numbers (in parentheses) of genes experienced transcription mode transition in the two treatment groups. **(D)** Density distribution of gene activity (in log2) fold change of type I to type II (black line) and type II to type I (red line) transition genes from N to T (left) and P to PT (right).

Next, we examined the transition among these four gene categories (Figure 5C). Except for type III, which had the fewest genes, more than 60% of the genes stayed in their original transcription type group. Most transitions (15.4%–35.6%) occurred between type I and II. Neither the I-to-II nor the II-to-I transition correlated with an increase or decrease at the transcriptional level (Figure 5D). Notably, transitions from inactive genes (type III and IV) to active genes (type I and II) were rare (Figure 5C), suggesting that inactive genes tend to stay inactive and that gene activation in heat stress response requires basal transcription.

Proximal promoter pausing and release is rate-limiting for the induction of some *HSP* genes

HSP genes were the quickest and most induced genes upon HS. Out of 82 active *HSP* genes, 38 and 43 were induced by at least 2-fold at the nascent RNA level in T vs. N and PT vs. P, respectively (Figure S10). In addition, 15 and 22 *HSP* genes were induced at the mRNA level in T vs. N and PT vs. P, respectively (Figure S10). Furthermore, small HSP, HSP70 and HSP90 family members were mostly induced, as exemplified by small *HSP/AT1G59860* (induced by 92.7 fold in T vs. N and 182.2 fold in PT vs. P), *HSP70/AT3G12580* (induced by 95.1 fold in T vs. N and 102.4 fold in PT vs. P), and *HSP90/AT5G52640* (induced by 84.6 fold in T vs. N and 97.6 fold in PT vs. P) (Figure S10). As 5′ pausing was originally discovered in *Drosophila HSP* genes (Gilmour and Lis, 1986; Rougvie and Lis, 1988), we next evaluated whether pausing release is a rate-limiting step for *HSP* gene induction in plants. We made different observations. First, no pausing was observed under either the non-HS (N and P) condition or the HS (T and PT) condition ($n = 45$, 54.9%), for example, *AT3G09440*, an *HSP70* family gene (Figure S11A). Second, pausing was detected in all conditions ($n = 16$, 19.5%), for example, *HSP70-1/AT5G02500* (Figure S11B). Third, states were transformed from non-active to active ($n = 15$, 18.3%), for example, *HSP23.6/AT4G25200* (Figure S11C). Fourth, pausing was also observed for HS and primed plants, but not for naïve plants ($n = 3$, 3.7%), for example, *HSP20-like/AT2G19310* (Figure S11D). Fifth, a pausing signal was observed in primed plants ($n = 3$, 3.7%), for example, *HSP70/AT3G12580*, and release from pausing might contribute to the stronger activation after HS (PT) (Figure S11E). Taken together, the pausing signals observed in some *HSP* genes indicate that pausing release is the rate-limiting step for their induction. Given that HSP gene families have undergone expansion in the plant kingdom, it is reasonable to speculate that plants adopt different transcriptional modes for different HSP members to address complicated environmental stresses.

Heat shock triggered pervasive transcriptional read-through at terminators

Recently, a thorough study on nuclear RNA profiling of mammalian cells under various conditions showed that almost all stresses induced transcriptional read-through at the terminators (Vilborg et al., 2017). Furthermore, Sun and colleagues reported that dehydration stress resulted in a longer 3′ untranslated region (UTR) in *Arabidopsis* (Sun et al., 2017), indicating the usage of downstream alternative polyadenylation sites. To determine whether HS affects transcription termination in plants, we profiled nascent RNA and mRNA around the annotated polyadenylation sites (PAS). As expected, a lower signal around the PAS (±250 bp) and a higher signal 500 bp downstream of the PAS were observed in both T and PT plants (Figure 6A, left panel, B), which suggests that a short HS time significantly decreases polyadenylation and termination efficiency. However, a prominent pattern change in the mRNA level was not observed (Figure 6A, right panel).

To quantitatively measure read-through events, we applied the read-through index, which is the ratio of read density between the PAS extension region (200–1 000 bp downstream of the PAS) and the GB region, using GRO-seq data and the ratio of read density in the PAS extension region to the gene's FPKM for RNA-seq (Figure 6C). As expected, the read-through index was extensively elevated due to more Pol II accumulation downstream of the PAS after 5 min HS compared with that in the non-HS plants at the nascent RNA level (Figure 6D, upper panel). However, little change was observed at the mRNA level (Figure 6D, bottom panel). It is possible that 5 min is not sufficient to observe read-through at the mRNA level.

Next, we established a pipeline to decide whether read-through transcription occurred in a gene induced by HS for GRO-seq (see Methods, Figure S12). To avoid the interference of neighboring genes, active genes (RPKM ≥ 3 in GB for all treatments) with distances from downstream genes >1 kb were selected for analysis, which resulted in 7 169 genes being included. Finally, 3,602 (50.2%) and 3,818 (53.2%) genes, with 2,610 genes in common, met the criteria for read-through in T vs. N and PT vs. P, respectively (Figure 7A, upper panel; Table S4). We hereby designated the 2 610 common set genes as faithful read-through genes. Meanwhile, the rest of the genes in both T vs. N and PT vs. P were defined as non-read-through genes (2,359). As indicated by the meta-profiles, Pol II accumulated more downstream of the PAS for genes with read-through under HS, although less Pol II accumulated within the PAS ± 250 bp region for both read-through and non-read-through genes (Figure 7A, middle panel). For another termination measurement, the PAS extension length (Figure S12 and Methods) of read-through genes after HS was also increased (Figure 7A, bottom panel). Notably, the P plants had a stronger read-through response than the N plants after HS at the nascent RNA level. After priming, more genes showed read-through after HS (Figure 7A, upper panel), and the read-through index of genes in PT plants was also significantly higher than that in T plants (Figure 7B). Therefore, plants became more prone to transcription read-through in response to HS after priming.

Although no significant differences in the PAS signals or nucleotide composition around the annotated PAS were

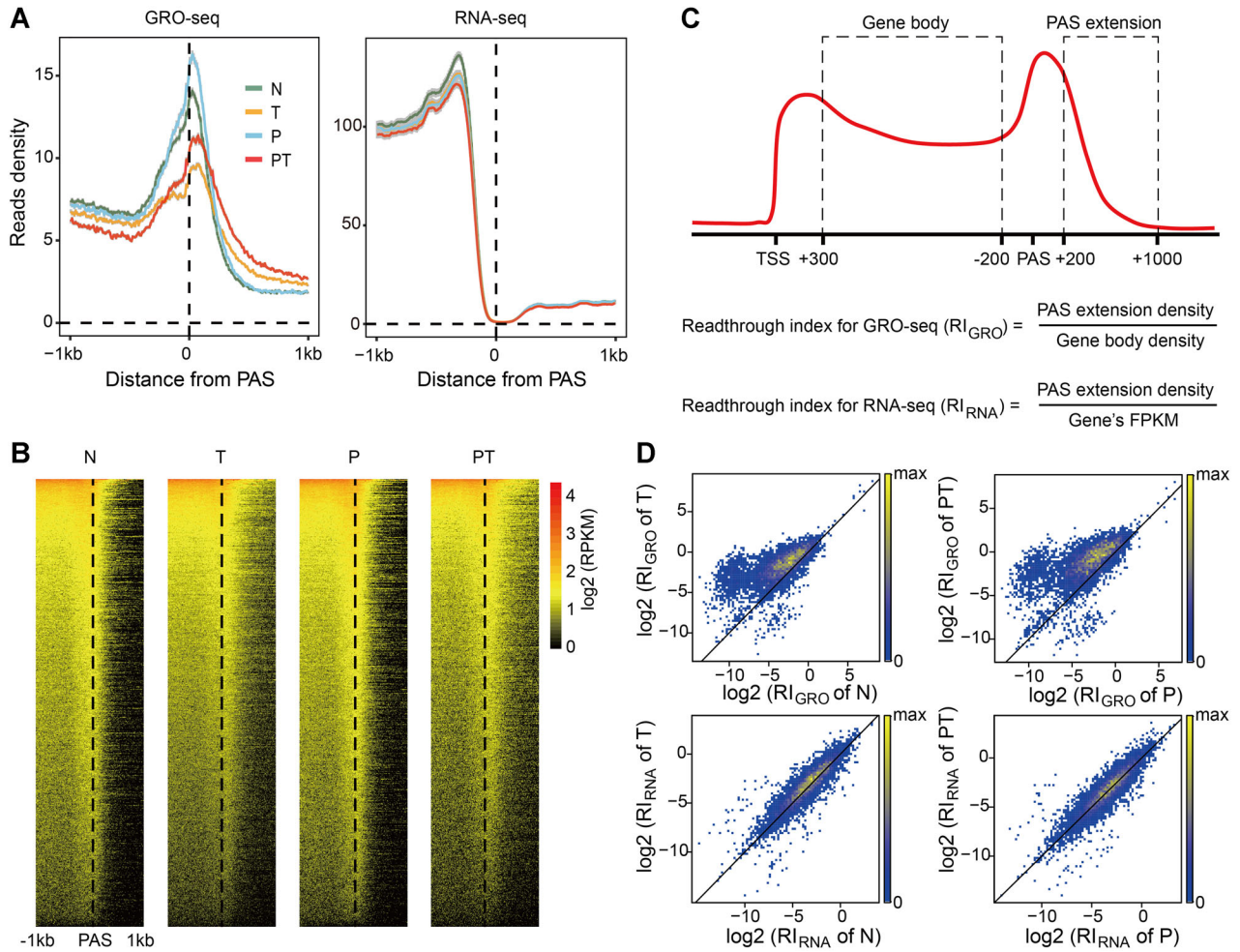

Figure 6. Heat shock triggers widespread transcriptional termination read-through

(A) Meta-profiles of read density around the polyadenylation site (PAS) for global run-on sequencing (GRO-seq) and RNA-seq. A total of 13 024 genes with lengths >1 kb and >1 kb from neighboring genes on the same strand were analyzed. Lines and shading represent the mean ± S.E.M. for each bin. **(B)** Heatmaps of nascent RNA read density distributions around the PAS for 6 455 active genes in at least one treatment that also had lengths >1 kb and were >1 kb away from neighboring genes on the same strand. Genes are ordered from high to low based on their gene activity. **(C)** Definition of the read-though index (RI) for GRO-seq and RNA-seq. **(D)** Scatter plots of the RI for GRO-seq (upper) and RNA-seq (bottom) between T vs. N (left) and PT vs. T (right). The color from yellow to blue represents the dot density from high to low.

observed between read-through and non-read-through genes (Figure S13), read-through genes were significantly longer than non-read-through genes, and the length of non-read-through genes was significantly shorter than that of overall genes (Figure 7C). Furthermore, read-through was more likely to be triggered in longer genes (Figure S14). In addition, the change in gene activity after HS was similar between the read-through and non-read-through genes (Figure S15A).

Read-through induced the 3′UTR extension of HS response genes

We used the same pipeline to calculate the PAS extension lengths of genes for RNA-seq data using the same gene sets as those used for GRO-seq. Genes with a PAS extension length in T (or PT) at least 200 bp longer than that in N (or P) were chosen as candidates for 3′ UTR extension. To further exclude false positives, we manually checked the read-through state of candidate genes in the genome browser. Finally, we found 79 genes with read-through (Table S5). Most of them showed read-through in both T and PT, and read-through was also detected at the nascent RNA level, such as *APX1*, while, for most genes with read-through at the nascent RNA level, no read-through was detected at the mRNA level, such as *AT4G08980* (Figure 7D). In addition, reverse transcription PCR was used to validate read-through under HS at the mRNA level. In treatment T and PT, a significant increasing of transcription was observed across and downstream PAS regions (Figures 7E and S16). Notably, a stronger signal was detected by primers downstream PAS (F3 and R3) than primers across PAS (F2 and R2), which suggests that there is only a small portion of mRNA with extended 3′ end, while the fate of downstream read-through RNA needs to be further investigated.

To further explore whether there were more genes with an extended 3′ UTR after a longer HS, we reanalyzed RNA-seq

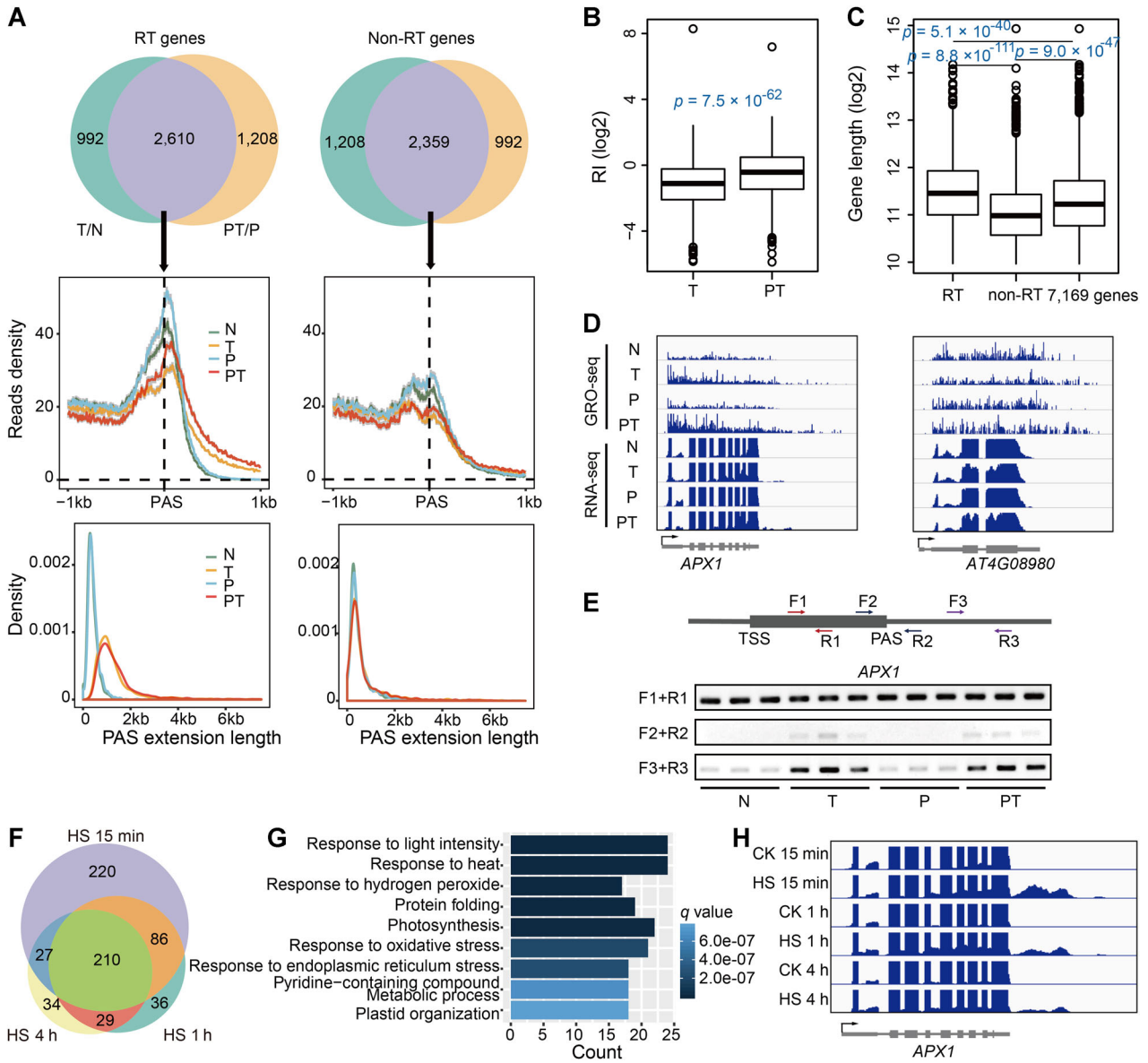

Figure 7. Characterization of genes with and without read-through (RT)

(A) Venn diagrams of RT and non-RT genes by comparing T/N and PT/P (upper panel) and meta-profiles of the read densities of RT and non-RT genes both in T/N and PT/P around the polyadenylation site (PAS) (middle panel). The bottom panel shows the frequency distributions of the PAS extension lengths for RT genes (left) and non-RT genes (right). For meta-profiles, lines and shading represent the mean ± S.E.M. for each bin. **(B)** Comparison of the read-through index of RT genes in T and PT. Two-sided *P* values for the Wilcoxon signed-rank test are presented in the boxplots. **(C)** Comparison of gene lengths among RT genes, non-RT genes and all genes. **(D)** Screenshots showing two examples of RT genes. The RT of AT5G64400 was detected at both the nascent RNA and messenger RNA (mRNA) levels. The RT of AT4G08980 was detected at only the nascent RNA level. **(E)** Reverse transcription polymerase chain reaction validation of APX1's read-through at mRNA level. Schematic shows position of regions analyzed (F1 + R1: the region between transcription start site (TSS) and PAS of annotation; F2 + R2: the region across PAS; F3 + R3: the region downstream of PAS). **(F)** Venn diagram showing the number of RT genes detected using RNA-seq data with different HS times (Cortijo et al., 2017). **(G)** Gene Ontology (GO) enrichment of the 210 common RT genes. The size of the dot represents the gene count. A hypergeometric test was used for statistical analysis, and the *P* values from the tests were converted to false discovery rate (FDR)-corrected *q* values. **(H)** An RT gene example from RNA-seq data (Cortijo et al., 2017).

data from a recent publication (Cortijo et al., 2017), where *Arabidopsis* plants were shifted from 17 °C to 27 °C and from 17 °C to 37 °C for 0.25, 1 and 4 h. We found that read-through was prominently induced when plants were shifted from 17 °C to 37 °C but not from 17 °C to 27 °C (Figure S17). The same pipeline was used to detect read-through genes. We identified 543, 361, and 300 candidate read-through genes after HS for 15 min, 1 h and 4 h, respectively (Figure 7F), with 210 genes in common (Table S6). Interestingly, these genes are involved in protein folding and response to light, heat, and oxidative stresses (Figure 7G), as exemplified by the HS-induced gene *APX1* (Figure 7H). Therefore, it is tempting to speculate that the 3′ UTR extension caused by read-through is a regulatory step in response to HS. Further, 94 out of the

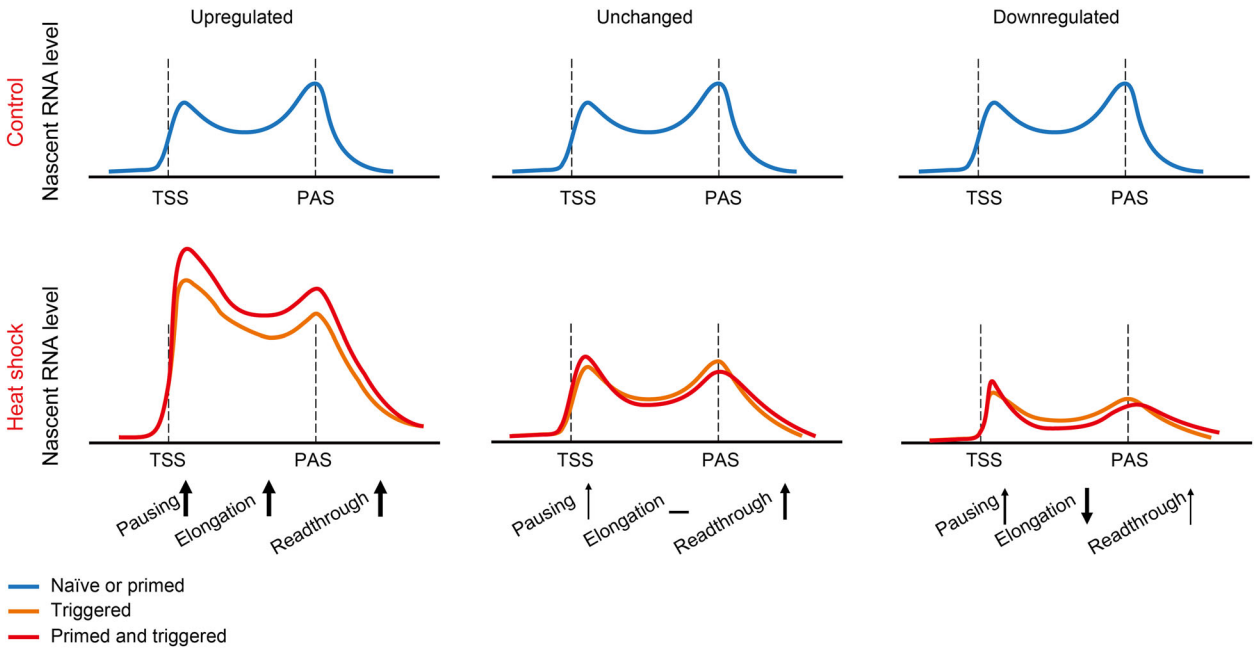

Figure 8. Transcription modes changed after heat shock (HS)

Genes were categorized into upregulated, unchanged, and downregulated groups. Transcription modes under non-HS and HS conditions are summarized from the meta-profiles of global run-on sequencing (GRO-seq) (Figures 4A and 6A). Proximal promoter pausing, elongation, and transcriptional read-through are indicated. The direction and thickness of arrows show the degree of increase or decrease. See details in text.

210 genes were overlapped with read-through genes detected by our GRO-seq data. Several HSP70 family genes are included (Table S7).

Previous studies have suggested that read-through might activate or repress the transcription of downstream neighboring genes (Crisp et al., 2018; Yu et al., 2019). Probably due to the short treatment time, an overall change of downstream gene transcription triggered by upstream read-through was not observed (Figure S15B).

DISCUSSION

Ambient temperature is one of the most fluctuating signals that sessile plants must address. Hence, plants have evolved prompt transcriptional responses to environmental temperature changes. Here, we show an unreported rapid and robust transcriptional response to HS by comparing the nascent RNA profiling of Naïve (N), Triggered (T), Primed (P), and Primed and Triggered (PT) plants using GRO-seq (Figure 8), which not only allowed the elucidation of very quickly altered genes but also the prediction of TFs involved. Performing the same analysis with mutants of ambient temperature signaling and/or transcriptional regulation would be informative.

Teves and Henikoff proposed that the reduced Pol II elongation upon HS is a major reason for transcription repression in *Drosophila* cells (Teves and Henikoff, 2011). We found that Pol II stalling in the proximal promoter region likely causes the genome-wide transcriptional downregulation, which agrees with two recent reports on mammalian cells transcriptional heat stress response using precision nuclear run-on sequencing (PRO-seq) (Mahat et al., 2016; Vihervaara et al., 2017). However, similar patterns of nascent RNA biogenesis obtained from GRO-seq at 22 °C and 30 °C rule out the possibility that Pol II aborted from the chromatin directly due to a temperature increase. It would be of great interests to know how plant cells stall Pol II at the proximal promoter upon HS.

Transcriptional read-through has been discovered as a common feature for mammalian cells experiencing various stresses and for plants experiencing dehydration treatment, at the mRNA level (Mahat et al., 2016; Sun et al., 2017; Vihervaara et al., 2017; Vilborg et al., 2017). In this study, read-through was revealed in more than 50% of genes by GRO-seq in 5-min HS plant leaf tissue (Figures 7A and 8), with 3′ UTR extensions from a small portion of those HS-induced read-through genes at both the 5-min and later HS mRNA levels. We also verified read-through at mRNA level by reverse transcription PCR and revealed different fates of transcriptional read-though RNA (Figure 7E). Taken together, these results suggest that HS induces a complicated and distinct change in the polyadenylation and termination upon HS. In addition, an alternative nascent RNA sequencing method, pNET-seq with anti-Ser2P antibody, is required to answer whether the HS trigger read-through detected by GRO-seq reflect the insufficient cleavage and polyadenylation or slower termination. Kindgren et al. (2018) recently reported that read-through of a long noncoding RNA interferes with the downstream gene transcription via RNA Pol II collision under cold conditions, which fine-tunes the

freezing tolerance. However, the biological significance and whether read-through occurs via signaling mechanisms or a passive physical change in the transcriptional machinery and/or chromatin remain to be elucidated.

Previous studies have shown that primed plants have a much more drastic response to drought, heat, and abiotic stresses (Bruce et al., 2007; Avramova, 2015). We demonstrate here that a genome-wide stress memory response occurred during nascent RNA biogenesis and discovered three prominent transcriptional changes induced by a very short HS (Figure 8): (i) RNA Pol II is blocked in the proximal promoter region, prohibiting productive elongation; (ii) overall downregulation in transcriptional activity; and (iii) pervasive transcriptional read-through at gene terminators. Hence, stress memory may occur at the transcription initiation, elongation, and termination steps (Figure 8). Given that read-through may affect the transcription of the downstream gene, transcriptional stress memory may start at one cycle and last until the next cycle.

Stress memory at the transcriptional level can be explained in two ways: first, P plants respond differently from N plants, for example, more Pol II stalls in the promoters of stress memory genes (Ding et al., 2012). Second, while P and N plants respond similarly, P plants transcribe more quickly than N plants, and N plants may catch up at a later time point. From the nascent RNA profiling results, we speculate that both scenarios might be true. Primed plants did a better job of prioritizing the transcriptomic changes upon HS than naïve plants. Mechanistically, Cortijo and colleagues proposed that HSFA1a promotes the fast exchange of H2A.Z for H2A at the transcriptional start sites of target genes in response to a mild temperature increase at 15 min, which allows for the fast transactivation of those genes (Cortijo et al., 2017). It is likely that H2A.Z is exchanged faster in PT plants than in T plants.

MATERIALS AND METHODS

Plant growth and HS treatment

Col-0 plants *Arabidopsis thaliana* were grown in soil at 21 °C under 16 h light/8 h dark cycles for 2 weeks followed by 1 week of heat stress priming or continuous growth as a control, and then were subjected to a heat shock. The plants were kept in the dark when treated with heat stress (37 °C, 1.5 h each day for 1 week) and HS (45 °C, 5 min once), and the control plants were also kept in the dark (21 °C). Leaf tissues from control and HS plants were collected with liquid nitrogen and stored at −80 °C. Two biological replicates were performed.

Nucleus run-on and GRO-seq library construction

Nuclear isolation, *in vitro* run-on, and nascent RNA purification followed by complementary DNA (cDNA) library construction were performed as previously reported (Zhu et al., 2018). The cDNA libraries were sequenced on the HiSeq2000 platform.

GRO-seq data processing

GRO-seq data were processed as described previously (Zhu et al., 2018). Briefly, raw reads were first trimmed by cutadapt (Martin, 2011), followed by filtering of ribosomal RNA and aligning to the TAIR10 version of the *Arabidopsis* genome using STAR (Dobin et al., 2013). Only uniquely mapped reads were retained. BEDTools was used to count the number of reads in each region (Quinlan and Hall, 2010). To avoid the interference of promoter-proximal pausing, the transcriptional activities of genes are presented as the read density in the GB, and differential transcription activity was evaluated using the R package DESeq2 (Love et al., 2014). DEGs were defined by the following criteria: they had to show more than 1.5-fold up- or down-regulation, and the false discovery rate (FDR) adjusted q value calculated by DESeq2 had to be less than 0.05. The read density for each region was calculated by normalizing the read count to the library size and mapable length (RPKM).

The 5′ pause index of each gene was calculated using the read density of the promoter-proximal peak (50 bp bin with the highest read density within the region −150 ~ +150 bp from the transcription start site (TSS)) divided by the read density of the GB (+301 bp from the TSS ~ PAS), as described in a previous study (Zhu et al., 2018). Significantly paused genes were identified using Fisher's exact test to compare the density of reads in the promoter-proximal peak to the density of reads in the GB compared to a uniform distribution of all reads based on the number of mapable bases. The P values from the tests were converted to FDR-corrected q values, and a cutoff of $q = 0.01$ and 5′ PI ≥ 3 was used for significance.

RNA-seq analysis

Total RNA from the same tissue as GRO-seq was isolated using TRIzol reagent (Thermo Fisher) and then treated with DNase I. Poly(A) RNA was purified using Dynabeads® Oligo $(dT)_{25}$ (Thermo Fisher). RNA-seq libraries were prepared using the NEBNext Ultra RNA Library Prep Kit (New England BioLabs) according to the manual and sequenced on the HiSeq2000 platform by Shanghai Hanyu Biotech. After removing adaptor sequences and low-quality data, the sequences were aligned to the *Arabidopsis* genome using STAR, and the mapped reads in each gene were quantified using the program htseq-count from the python package HTSeq (Anders et al., 2015). Differential gene expression was evaluated using the R package DESeq2. Since paired-end sequencing was applied to RNA-seq, gene expression was normalized as FPKM.

Gene Ontology

To survey the biological functions of groups of genes, Gene Ontology enrichment was performed using the Bioconductor package clusterProfiler (Yu et al., 2012).

Read-through index and PAS extension length

To search for genes with Pol II extension after PAS, we first defined a read-through index, which was calculated using

read density of the PAS extension region (+201 ~ +1 000 bp from PAS) divided by the read density of the GB (+301 bp from TSS ~ −200 bp from PAS). In addition, we adopted a computational pipeline to calculate the PAS extension length of the gene. The downstream region of the PAS of each gene was divided into 100 bp windows shifted by 50 bp, and the read density in each window was calculated. Once there were three continuous windows (~150 bp) with a read density <1 RPKM, Pol II transcription was considered to be finished. The extension length was calculated as 50×*n*, in which "*n*" was the *n*th window when Pol II was considered to be finished (Figure S12).

Genes with read densities of GB ≥3 RPKM in all treatments and distances from downstream genes >1 kb were kept for read-through analysis. After HS, genes with a read-through index fold change ≥1.5 (vs. control) and a PAS extension length 200 bp longer than that under control conditions were considered to have transcriptional termination read-through.

Poly(A) signals and nucleotide composition around the PAS

The AATAAA-like motif, that is, the canonical AATAAA signal and 18 single-nucleotide mutations, were considered poly(A) signals. The sequences 50 nt upstream of the TAIR10-annotated PAS were extracted for poly(A) signal analysis. The sequences 100 nt up- and downstream of the annotated PAS were extracted for nucleotide composition analysis.

Quantitative reverse transcription PCR and RT-PCR validation

Chromatin-associated RNA was isolated as described previously (Zhu et al., 2020), and then cDNA was synthesized from 0.2 to 0.5 μg of chromatin RNA using SuperScript™ III Reverse Transcriptase (Invitrogen) by following the protocol. Total RNA was isolated using TRIzol isolation reagent, and first-strand cDNA was synthesized from 1 μg of total RNA using the PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa, Cat. #RR047Q). Quantitative reverse transcription PCR was run in three biological replicates and two technical replicates. For the normalization of gene expression, the *PROTEIN PHOSPHATASE 2A SUBUNIT A3* (*PP2AA3*) (AT1G13320) was used as an internal standard (Czechowski et al., 2005).

To validate transcriptional read-through at mRNA level under HS, reverse transcription PCR was conducted for selected genes in three regions: the region between TSS and PAS, the region across PAS and the region downstream of PAS. The same cDNA of quantitative reverse transcription PCR was used. PCR reaction was subjected to 35 cycles of denaturing (15 s, 94 °C), annealing (20 s, 58 °C), and extension (20 s, 72 °C). The oligonucleotide sequence of primers used in this study are listed in Table S8.

Reanalysis of public RNA-seq data

Public RNA-seq data from the Gene Expression Omnibus (accession number GEO: GSE79355) were downloaded. In this study, plants were shifted from 17 °C to 27 °C for 15 min, 1 h, and 4 h with two replicates, and from 17 °C to 37 °C for 15 min, 1 h, and 4 h with one replicate. These RNA-seq data were processed using the same pipeline as that described above.

ACCESSION NUMBERS

The sequencing data and processed files are available at the Gene Expression Omnibus under accession number GSE128698.

ACKNOWLEDGEMENTS

We thank Dr. Zhe Wu for sharing the chromatin-associated RNA isolation protocol. This work was supported by the National Natural Science Foundation of China (31900463 to M.L. and 31871289, 31471165 to Z.D.).

CONFLICT OF INTEREST

The authors declare they have no conflicts of interest to this work.

AUTHOR CONTRIBUTIONS

Z. D., M. L., and J. Z. designed the research; J. Z. and M. L. performed the research; M. L. analyzed the data; Z. D., M. L., and J. Z. wrote the paper. All authors read and approved the manuscript.

Edited by: Yijun Qi, Tsinghua University, China

Received Apr. 23, 2020; **Accepted** Jul. 8, 2020; **Published** Jul. 9, 2020

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics **31:** 166–169.

Avramova, Z. (2015). Transcriptional 'memory' of a stress: transient chromatin and memory (epigenetic) marks at stress-response genes. Plant J. **83:** 149–159.

Bruce, T.J.A., Matthes, M.C., Napier, J.A., and Pickett, J.A. (2007). Stressful "memories" of plants: Evidence and possible mechanisms. Plant Sci. **173:** 603–608.

Brzezinka, K., Altmann, S., Czesnick, H., Nicolas, P., Gorka, M., Benke, E., Kabelitz, T., Jähne, F., Graf, A., Kappel, C., and Bäurle, I. (2016). *Arabidopsis* FORGETTER1 mediates stress-induced chromatin memory through nucleosome remodeling. eLife **5:** e17061.

Corrales, A.-R., Carrillo, L., Lasierra, P., Nebauer, S.G., Dominguez-Figueroa, J., Renau-Morata, B., Pollmann, S., Granell, A., Molina, R.-V., Vicente-Carbajosa, J., and Medina, J. (2017). Multifaceted role of cycling DOF factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in *Arabidopsis*. Plant, Cell Environ. **40:** 748–764.

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

Covell, S., Ellis, R.H., Roberts, E.H., and Summerfield, R.J. (1986). The influence of temperature on seed germination rate in grain legumes. I. Acomparison of chickpea, lentil, soybean and cowpea at constant temperatures. J. Exp. Bot. **37:** 705–715.

Crisp, P.A., Smith, A.B., Ganguly, D.R., Murray, K.D., Eichten, S.R., Millar, A.A., and Pogson, B.J. (2018). RNA polymerase II read-through promotes expression of neighboring genes in SAL1-PAP-XRN retrograde signaling. Plant Physiol. **178:** 1614–1630.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol. **139:** 5.

Ding, Y., Fromm, M., and Avramova, Z. (2012). Multiple exposures to drought 'train' transcriptional responses in *Arabidopsis*. Nat. Commun. **3:** 740.

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

Fitter, A.H., and Fitter, R.S.R. (2002). Rapid changes in flowering time in british plants. Science **296:** 1689.

Gilmour, D.S., and Lis, J.T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. Mol. Cell. Biol. **6:** 3984–3989.

Gray, W.M., Östin, A., Sandberg, G., Romano, C.P., and Estelle, M. (1998). High temperature promotes auxin-mediated hypocotyl elongation in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **95:** 7197–7202.

Grinevich, D.O., Desai, J.S., Stroup, K.P., Duan, J., Slabaugh, E., and Doherty, C.J. (2019). Novel transcriptional responses to heat revealed by turning up the heat at night. Plant Mol. Biol. **101:** 1–19.

Hahn, A., Bublak, D., Schleiff, E., and Scharf, K.-D. (2011). Crosstalk between Hsp90 and Hsp70 chaperones and heat stress transcription factors in tomato. Plant Cell **23:** 741–755.

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

Hedhly, A., Hormaza, J.I., and Herrero, M. (2009). Global warming and sexual plant reproduction. Trends Plant Sci. **14:** 30–36.

Herrero, M.P., and Johnson, R.R. (1980). High temperature stress and pollen viability of maize. Crop Sci. **20:** 796–800.

Kindgren, P., Ard, R., Ivanov, M., and Marquardt, S. (2018). Transcriptional read-through of the long non-coding RNA SVALKA governs plant cold acclimation. Nat. Commun. **9:** 4561.

Lämke, J., and Bäurle, I. (2017). Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. Genome Biol. **18:** 124.

Lämke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. The EMBO J. **35:** 162–175.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. Nature **529:** 84.

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

Ling, Y., Serrano, N., Gao, G., Atia, M., Mokhtar, M., Woo, Y.H., Bazin, J., Veluchamy, A., Benhamed, M., Crespi, M., Gehring, C., Reddy, A.S.N., and Mahfouz, M.M. (2018). Thermopriming triggers splicing memory in *Arabidopsis*. J. Exp. Bot. **69:** 2659–2675.

Liu, H.-c, Lämke, J., Lin, S.-y, Hung, M.-J., Liu, K.-M., Charng, Y.-y, and Bäurle, I. (2018). Distinct heat shock factors and chromatin modifications mediate the organ-autonomous transcriptional memory of heat stress. Plant J. **95:** 401–413.

Liu, H.-C., Liao, H.-T., and Charng, Y.-Y. (2011). The role of class A1 heat shock factors (HSFA1s) in response to heat and other stresses in *Arabidopsis*. Plant, Cell Environ. **34:** 738–751.

Liu, H.-T., Gao, F., Li, G.-L., Han, J.-L., Liu, D.-L., Sun, D.-Y., and Zhou, R.-G. (2008). The calmodulin-binding protein kinase 3 is part of heat-shock signal transduction in *Arabidopsis thaliana*. Plant J. **55:** 760–773.

Liu, J., Feng, L., Gu, X., Deng, X., Qiu, Q., Li, Q., Zhang, Y., Wang, M., Deng, Y., Wang, E., He, Y., Bäurle, I., Li, J., Cao, X., and He, Z. (2019). An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in *Arabidopsis*. Cell Res. **29:** 379–390.

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

Mahat, D.igB., Salamanca, H.H., Duarte, Fabiana, M., Danko, Charles, G., and Lis, J.ohnT. (2016). Mammalian heat shock response and mechanisms underlying its genome-wide transcriptional regulation. Mol. Cell **62:** 63–78.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. **17:** 10–12.

Mishra, S.K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., and Scharf, K.-D. (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes Dev. **16:** 1555–1567.

Mozgová, I., Wildhaber, T., Liu, Q., Abou-Mansour, E., L'Haridon, F., Métraux, J.-P., Gruissem, W., Hofius, D., and Hennig, L. (2015). Chromatin assembly factor CAF-1 represses priming of plant defence response genes. Nat. Plants **1:** 15127.

O'Malley, R.onanC., Huang, S., shan, C., Song, L., Lewsey, M.athewG., Bartlett, A., Nery, J.osephR., Galli, M., Gallavotti, A., and Ecker, J. osephR. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. Cell **165:** 1280–1292.

Ohama, N., Kusakabe, K., Mizoi, J., Zhao, H., Kidokoro, S., Koizumi, S., Takahashi, F., Ishida, T., Yanagisawa, S., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2016). The transcriptional cascade in the heat stress response of *Arabidopsis* is strictly regulated at the level of transcription factor expression. Plant Cell **28:** 181–201.

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

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26:** 841–842.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell **54:** 795–804.

Sani, E., Herzyk, P., Perrella, G., Colot, V., and Amtmann, A. (2013). Hyperosmotic priming of *Arabidopsis* seedlings establishes a long-term somatic memory accompanied by specific changes of the epigenome. Genome Biol. **14:** R59.

Stief, A., Altmann, S., Hoffmann, K., Pant, B.D., Scheible, W.-R., and Bäurle, I. (2014). *Arabidopsis* miR156 regulates tolerance to recurring environmental stress through *SPL* transcription factors. Plant Cell **26:** 1792–1807.

Sun, H.X., Li, Y., Niu, Q.W., and Chua, N.H. (2017). Dehydration stress extends mRNA 3' untranslated regions with noncoding RNA functions in *Arabidopsis*. Genome Res. **27:** 1427–1436.

Teves, S.S., and Henikoff, S. (2011). Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. Genes Dev. 25.

Vihervaara, A., Mahat, D.B., Guertin, M.J., Chu, T.Y., Danko, C.G., Lis, J.T., and Sistonen, L. (2017). Transcriptional response to stress is pre-wired by promoter and enhancer architecture. Nat. Commun. 8.

Vilborg, A., Sabath, N., Wiesel, Y., Nathans, J., Levy-Adam, F., Yario, T. A., Steitz, J.A., and Shalgi, R. (2017). Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. Proc. Natl. Acad. Sci. USA **114:** E8362–E8371.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci. **9:** 244–252.

Wang, Y., Wang, Y., Song, Z., and Zhang, H. (2016). Repression of MYBL2 by both microRNA858a and HY5 Leads to the activation of anthocyanin biosynthetic pathway in *Arabidopsis*. Mol. Plant **9:** 1395–1405.

Yamada, K., Fukao, Y., Hayashi, M., Fukazawa, M., Suzuki, I., and Nishimura, M. (2007). Cytosolic HSP90 regulates the heat shock response that is responsible for heat acclimation in *Arabidopsis thaliana*. J. Biol. Chem. **282:** 37794–37804.

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., and Yamaguchi-Shinozaki, K. (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, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS: J. Integr. Biol. **16:** 284–287.

Yu, X., Martin, P.G.P., and Michaels, S.D. (2019). BORDER proteins protect expression of neighboring genes by promoting 3′ Pol II pausing in plants. Nat. Commun. **10:** 4359.

Zhu, D., Mao, F., Tian, Y., Lin, X., Gu, L., Gu, H., Qu, L.J., Wu, Y., and Wu, Z. (2020). The features and regulation of co-transcriptional splicing in *Arabidopsis*. Mol. Plant **13:** 278–294.

Zhu, J., Liu, M., Liu, X., and Dong, Z. (2018). RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*. Nat. Plants **4:** 1112–1123.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: http://onlinelibrary.wiley.com/doi/10.1111/jipb.12990/suppinfo

Figure S1. Principle component analysis (PCA) of different samples for global run-on sequencing (GRO-seq) (left) and RNA-seq (right)

Figure S2. Boxplots of the transcriptional level for non-upregulated genes ($n = 16\,665$) among the four treatment groups at the nascent RNA level

(A) Reads count of genes was normalized to total library. **(B)** To avoid the bias caused by the significant increase in upregulated genes, reads count of genes was normalized to total library excluding reads on upregulated genes. The line in the box represents the median value, and the upper and lower whiskers represent 75% and 25% of the data, respectively. Two-sided *p* values for the Wilcoxon signed-rank test are presented in the plots.

Figure S3. Venn diagram of upregulated genes (**A** and **C**) and downregulated genes (**B** and **D**) in T vs. N and P vs. N for global run-on sequencing (GRO-seq) (**A** and **B**) and RNA-seq (**C** and **D**)

Figure S4. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) validation of nascent RNA (**A**) and messenger RNA (mRNA) (**B**)

(A) and mRNA (B). Two biological replicates were performed for GRO-seq and RNA-seq. Three biological replicates were performed for qRT-PCR. All data represent the mean value of biological replicates. Error bars represent the standard deviations.

Figure S5. Messenger RNA (mRNA) expression levels of transcription factor (TF) genes involved in transcriptional regulatory network in response to heat shock (HS)

Heatmap (left) shows the expression level of genes in four treatments and barplot (right) represents the fold change of genes between different treatment groups.

Figure S6. Upregulated genes and the percentages targeted by different transcription factor (TF) families **(A)**

Venn diagram of upregulated genes in T vs. N and PT vs. P. **(B)** The radar map indicates the percentage of upregulated genes upon HS in T vs. N (orange) and PT. vs. P (red) targeted by each transcription factor family. Target genes of each TF family are derived from DAP-seq.

Figure S7. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) validation of nascent RNA and messenger RNA (mRNA) for stress memory genes in cluster 3 (**A**) and cluster 4 (**B**)

Two biological replicates were performed for GRO-seq and RNA-seq. Three biological replicates were performed for qRT-PCR. All data represent the mean value of biological replicates. Error bars represent the standard deviations.

Figure S8. Clusters of 845 memory genes

Hierarchical clustering of 845 genes based on nascent RNA transcriptional activity reveals five major clusters. And then genes in each cluster were grouped by hierarchical clustering based on mRNA level. Significantly enriched GO terms were showed for each group.

Figure S9. No difference of transcriptional activity was detected when global run-on sequencing (GRO-seq) was done under different run-on temperatures (**22°C** and **30°C**)

(A) Comparison of reads density near TSS and PAS with different run-on temperature. **(B)** MA-plot of gene activity under 22 °C vs. 30 °C. No DEGs was detected. **(C)** and **(D)** IGV screenshots of *HSP70* and *HSFA2*.

Figure S10. Gene activities of *HSP* genes at the nascent RNA level (left) and the messenger RNA (mRNA) level (right)

Heatmap shows the expression level of genes in four treatment groups and barplot represents the fold change of genes between different treatment groups.

Figure S11. Screenshots of nascent RNA distribution along some *HSP* genes

Figure S12. Pipeline to calculate the polyadenylation site (PAS) extension length

The downstream PAS was divided into 100 bp bins shifted by 50 bp. The read density of each bin was calculated. Once there were three continuous bins with an RPKM <1, Pol II transcription was considered to be stopped. The PAS extension length was calculated by multiplying n by 50.

Figure S13. No significant difference in poly(A) signals or nucleotide composition was observed between 2,610 RT and 2,359 non-RT genes

(A) Percentages of the occurrence and distribution of AATAAA-like motifs across 50-nt upstream sequences of annotated PAS sites. Statistical significance was tested by Fisher's exact test. **(B)** Nucleotide compositions of 100 nt up- and downstream flanking sequences of annotated PAS regions of RT genes (left) and non-RT genes (right).

Figure S14. Boxplots indicating the ranges of relative polyadenylation site (PAS) extension lengths of genes that have different lengths in T vs. N (upper) and PT vs. P (bottom)

The bottom and top edges of the box represent the first and third quartile values, respectively, and the bold line inside the box represents the median value.

Figure S15. Comparison of the fold changes of gene activity between RT and non-RT genes (**A**) and the downstream genes (**B**) in T vs. N (left) and PT vs. P (right)

The bottom and top edges of the box represent the first and third quartile values, respectively, and the bold line inside the box represents the median value. Two-sided *p* values for the Wilcoxon signed-rank test are presented in the plots.

Figure S16. Reverse transcription polymerase chain reaction validation of read-through at messenger RNA (mRNA) level

(A) Screenshots showing two RT genes at nascent and mRNA levels. **(B)** RT-PCR validation of readthrough. Schematic shows position of regions analyzed (F1+R1: the region between the annotated TSS and PAS; F2+R2: the region across PAS; F3+R3: the region downstream of PAS).

Figure S16. Transcription read-through was detected at the messenger RNA (mRNA) level by transferring plants from 17°C to 37°C but not from 17°C to 27°C

(A) Comparison of genes' RI between control (CK) and heat stress (HS) under different heat stress treatments. The upper panel shows the treatment of plants from 17 °C to 37 °C and the bottom panel shows the

treatment of plants from 17 °C to 27 °C. Genes with RI > 0.01 in either treatment were used in analysis. Two-sided *p* values for the Wilcoxon signed-rank test are presented in the plots. **(B)** Screenshot of the read densities along RNA-seq genes from IGV. Regions with transcription readthrough were highlighted by red boxes.

Table S1. The number of raw reads and final reads of global run-on sequencing (GRO-seq) and RNA-seq

Table S2. Pearson correlation of global run-on sequencing (GRO-seq) and RNA-seq biological replicates. Reads were mapped to the genome, and unique reads were binned in 500 bp windows

Table S3. Gene expression levels detected by global run-on sequencing (GRO-seq) and RNA-seq

Table S4. RT genes detected by global run-on sequencing (GRO-seq) under 5-min heat shock (HS)

Table S5. RT genes detected by RNA-seq under 5-min heat shock (HS).

Table S6. RT genes detected by RNA-seq from public data.

Table S7. RT genes detected by both global run-on sequencing (GRO-seq) from our study and RNA-seq from public data

Table S8. The oligonucleotide sequence of primers used in this study