The R-loop Atlas of Arabidopsis Development and Responses to Environmental Stimuli

Wei Xu#, Kuan Li#, Shuai Li#, Quancan Hou#, Yushun Zhang, Kunpeng Liu and Qianwen Sun

Tsinghua-Peking Joint Center for Life Sciences, and Center for Plant Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China.

# These authors contributed equally to this work.

*Correspondence: sunqianwen@mail.tsinghua.edu.cn.

Short title: The R-loop atlas of Arabidopsis.

One-sentence summary: This study provides the landscape of R-loops in Arabidopsis during development and stress responses, and lays the groundwork for future investigation of the diverse functions of R-loops.

Address correspondence to: sunqianwen@mail.tsinghua.edu.cn. 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 (www.plantcell.org) is: Qianwen Sun (sunqianwen@mail.tsinghua.edu.cn).

ABSTRACT

R-loops are a common chromatin feature with essential functions in multiple cellular processes and diseases. However, little is known about the dynamic patterns of R-loops in a given organism. Here, using our recently developed genome-wide R-loop profiling method, we generated a comprehensive atlas quantifying the R-loop patterns of Arabidopsis thaliana at 53 points during development and during responses to environmental stimuli. The R-loop patterns were fairly stable in plants at the vegetative stage and in response to different light spectra and other environmental stimuli. Notably, the R-loops showed turnover during the plant lifecycle, with patterns switching between generations. Importantly, R-loop dynamics were not strongly associated with RNA abundance, indicating that the mechanisms regulating R-loop formation
and RNA accumulation are independent. We also observed enrichment of
R-loops in transcription factor-binding regions, suggesting that R-loops could
function as potential cis transcriptional regulators. This study provides an
overview of R-loop dynamics in Arabidopsis during development and stress
responses, highlights the unique dynamics of R-loops in the flowering plant
Arabidopsis, and lays the groundwork for elucidating the functions of R-loops.

**Keywords**
R-loop dynamics, genome regulation, environmental stimuli

**INTRODUCTION**
R-loops are chromosomal structures composed of a DNA:RNA hybrid double
helix and a single-stranded DNA molecule. R-loops were previously thought to
be rare and their influence in the genome was trivial. However, an increasing
number of studies have revealed that R-loops are unique, prevalent,
multifunctional genomic structures that are present in organisms from
prokaryotes to eukaryotes (Ginno et al., 2012; Ginno et al., 2013; Sun et al.,
2013; Skourtì-Stathaki et al., 2014; Chen et al., 2015; Santos-Pereira and
Aguilera, 2015; Al-Hadid and Yang, 2016; Sanz et al., 2016; Wahba et al.,
2016; Chen et al., 2018). R-loops play essential roles in many biological
processes, such as DNA replication, DNA methylation, histone modification,
chromosome segregation, transcriptional regulation, DNA repair, and
immunoglobulin class switching recombination (Yu et al., 2003; Sun et al.,
2013; Skourtì-Stathaki et al., 2014; Chen et al., 2015; Ohle et al., 2016;
Grunseich et al., 2018; Kabeche et al., 2018; García-Muse and Aguilera, 2019).
Critical functions for R-loops in plants have also been recently revealed, such
as key roles in root development (Shafiq et al., 2017) and genome integrity and
stability (Yang et al., 2017; Yuan et al., 2019; Yang et al., 2020). Nevertheless,
most studies of plant R-loop function have focused on specific R-loop loci
(Santos-Pereira and Aguilera, 2015).
Several methods have been employed to characterize genome-wide R-loop distribution in mammals or budding yeasts, including DNA:RNA immunoprecipitation (DRIP) coupled to high-throughput sequencing (DRIP-seq) (Ginno et al., 2012), DNA:RNA in vitro enrichment coupled to high-throughput sequencing (DRIVE-seq) (Ginno et al., 2012), DNA:RNA immunoprecipitation followed by hybridization on tiling microarrays (DRIP-chip) (Chan et al., 2014), chromatin immunoprecipitation with antibody S9.6 followed by deep sequencing (S9.6 ChIP-seq) (El Hage et al., 2014), RNA:DNA immunoprecipitation followed by high-throughput sequencing (RDIP-seq) (Nadel et al., 2015), and S1 nuclease DNA:RNA immunoprecipitation with deep sequencing (S1-DRIP-seq) (Wahba et al., 2016). However, these methods cannot provide strand-specific information for R-loops, which is crucial for investigating the formation, regulation and function of R-loops. To solve this problem, several methods have been developed to profile R-loops in a strand-specific manner, including DNA:RNA immunoprecipitation followed by cDNA conversion coupled to high-throughput sequencing (DRIPc-seq) (Sanz et al., 2016; Sanz and Chedin, 2019), and by expressing a catalytically dead RNASEH1 followed by strand-specific amplification of immunoprecipitated DNA (R-ChIP) (Chen et al., 2018). We recently developed a novel R-loop mapping method, ssDRIP-seq (single-strand DNA ligation-based library preparation from DNA:RNA hybrid immunoprecipitation, followed by sequencing), a powerful technique for genome-wide R-loop profiling (Xu et al., 2017). ssDRIP-seq conveniently profiles R-loops in a strand-specific manner and produces reproducible, high-resolution data (Xu et al., 2017). This method has been successfully applied to many organisms, including both prokaryotes and eukaryotes (Yang et al., 2019; Yuan et al., 2019). Using ssDRIP-seq, we previously profiled R-loops in the Arabidopsis thaliana genome and discovered several unique features of R-loops in plants, particularly the high enrichment of antisense R-loops around transcription start sites (TSSs) (Xu et al., 2017).
However, a general understanding of R-loop formation and dynamics during development and in response to environmental stimuli in plants is still lacking. In this study, we profiled R-loops quantitatively in a large group of samples from the model plant Arabidopsis at different developmental stages, under different light and temperature conditions, and in the presence of various biotic and abiotic stresses (Figure 1A). By performing multidimensional analyses of the R-loop dynamics data set, we determined that R-loop distribution patterns were surprisingly conserved and stable and that R-loop dynamics were independent of RNA abundance. This study provides an initial overview of R-loop landscapes under various conditions in Arabidopsis and can set the grounds for further understanding of the diverse roles of R-loops in plant development and responses to environment stimuli as well as the structure and organization of eukaryotic genomes.

RESULTS

R-loop patterns are relatively conserved and stable

During various developmental stages, plant transcriptomes exhibit dramatic dynamic changes in specific tissues (Xiang et al., 2011; Liu et al., 2012). To explore the dynamic patterns and interactions of R-loops during development and environmental responses in Arabidopsis, we collected data from four sets of plants grown under various conditions, including different developmental stages, temperatures, light levels, and biotic and abiotic stimuli (Figure 1A and Supplemental Data Set 1). For the negative control, we treated genomic DNA (gDNA) with ribonuclease H (RNase H), a non-sequence-specific endonuclease that degrades RNAs within DNA:RNA hybrids. ssDRIP-seq is a high-efficiency method that can be used to profile and characterize R-loops in a strand-specific manner in almost any organism (Xu et al., 2017). All R-loop data obtained in this study were generated by ssDRIP-seq (Figure 1B and Supplemental Figure 1A). After data processing, more than 23 million reads were generated and mapped to the Arabidopsis
TAIR10 genome for each sample, and more than 20-fold coverage of the entire genome was obtained, except for the negative RNase H controls (Supplemental Data Set 1; also see the database The R-loop Atlas, http://bioinfor.kib.ac.cn/R-loopAtlas). Snapshots of R-loops showed that the signals from wR-loops (R-loops formed in the Watson strand of DNA) and cR-loops (R-loops formed in the Crick strand of DNA) were present in different regions of the DNA, with very little overlap, demonstrating a high degree of strand specificity (Figure 1C). To demonstrate how strand-specificity is achieved in ssDRIP-seq, we analyzed the sizes of gDNA fragments generated by enzyme cocktail. The results showed that restriction enzyme digestion fragmented the Arabidopsis genome into ~52 bp (Supplemental Figure 1B). As these enzymes have common properties that only cleave double-stranded DNA but not DNA:RNA hybrid (http://rebase.neb.com/cgi-bin/hybcombolist), intact R-loop structures with two short dsDNA wings were separated from gDNA after digestion. Short dsDNA wings are crucial for the detachment of ssDNA from R-loops, because the weak binding force of dsDNA wings make the ssDNA was washed away easily in washing steps (Supplemental Figure 1C). R-loop signals in the RNase H-treated samples were much weaker than those in the untreated samples, supporting the validity and specificity of ssDRIP-seq (Figure 1C).

Surprisingly, the R-loop patterns were relatively conserved under most conditions, except during generational switches such as flowering and germination and at 84 hours after recovery from long heat stress treatment (37°C for 30 h, LHSR84) (Figure 1C). Both the wR-loop and cR-loop peaks were distributed conservatively in different regions of the genome, and the signal intensity was more or less the same among most samples (Figure 1C). This is quite different from the results of RNA expression analysis and the changes in other chromosome structures, such as histone modifications, which show dramatic dynamic changes globally during plant development and in response to environmental stimuli (Brusslan et al., 2015). As also shown in
Figure 1D, based on Spearman rank correlation coefficients, the R-loop distribution patterns were relatively similar in most samples, except for a few groups, including flowers and germinating seedlings.

We then analyzed the average levels of R-loops on the total annotated genes. Consistent with previous observations (Xu et al., 2017), sense R-loops were highly enriched in gene bodies, while antisense R-loops were anchored around TSSs in every sample from all four sets of data (Figure 1E and Supplemental Figure 2A). Meanwhile, some samples from important developmental stages exhibited different levels of R-loops in gene bodies and/or promoter (TSS ± 150 bp) regions (Supplemental Figure 2B, 2C, and 2D). For example, compared to seedling leaves, the levels of both sense and antisense R-loops were reduced in the regions upstream of TSSs and downstream of TTSs (transcription termination sites) in flowers, whereas the levels of sense R-loops near TTSs and antisense R-loops around TSSs were reduced during germination (Supplemental Figure 2B). Different light conditions did not alter the levels of antisense R-loops, but there were dramatic differences in the levels of sense R-loops in white light (WL) and dark (DK, no light) (Supplemental Figure 2C). Finally, the levels of antisense R-loops around TSSs increased slightly during long heat stress treatment (37°C for 30 h), followed by a decline for 12 or 84 h after treatment (Supplemental Figure 2D), which suggested that this class of R-loops might play regulatory roles in plant responses to environmental temperature changes.

Uncoupling of sGB-R-loops and asTSS-R-loops

Considering the significant enrichment of sense R-loops in gene bodies (hereinafter referred to as sGB-R-loops) and antisense R-loops around TSSs (hereinafter referred to as asTSS-R-loops) (Figure 2A), we analyzed the enrichment and dynamics of these two classes of R-loops in all genes and transposable elements (TEs). The patterns and levels of sGB-R-loops in most
groups of data sets were similar, except for the following: flowers in the development group, DK samples in the light group, and LHS (long heat stress), LHSR12 (12 h recovery after long heat stress), and LHSR84 samples in the temperature group (Supplemental Figure 3). As shown in the scatter plots, the differences of R-loops between most treatment groups vs. control groups were not significant (Figure 2B to 2D). For example, different light conditions, a constant temperature of 17°C vs. 27°C, short-term heat or cold stress, long, cold stress, and SA (Salicylic acid), ABA (Abscisic acid) or UVC (Ultraviolet light in the C spectrum, 254 nm) treatment resulted in only a few genes with different R-loop levels, i.e., differential R-loop genes (DRGs, defined by a q value < 0.05 and a |log₂FC| > 1), suggesting that the levels of both sGB-R-loops and asTSS-R-loops were very stable under these conditions (Figure 2B and 2D; Supplemental Figure 4A to 4C). However, we found that when the plants were treated with long heat stress (whether they were allowed to recover or not; Figure 2B) or at different developmental stages (Figure 2C) or grown in the dark (Figure 2D), R-loops exhibited various differences, indicating that R-loops were dramatically altered under special conditions. These findings suggest that R-loops might perform important functions during plant growth and responses to environmental signals. The number of DRGs in each sample among both sGB-R-loops and asTSS-R-loops are similar (Figure 2B to 2D), revealing their synchronized dynamics. In addition, scatter plot analysis of the fold changes between the LHS group and the untreated control group (CTRL-7d) indicated that the dynamic ranges of the sGB-R-loops and asTSS-R-loops were also quite similar (Figure 2E). These similarities suggest the possible coupling of sGB-R-loop and asTSS-R-loop dynamics. However, linear correlation analysis of the dynamics between sGB-R-loops and asTSS-R-loops based on a comparison of LHS and CTRL-7d showed that the correlation coefficients were very low (R² = 0.021, Rho = 0.119; Figure 2E).
To visualize the entire pattern on a larger scale, we used all of the samples to analyze the coupling between sGB-R-loops and asTSS-R-loops by identifying overlapping DRGs among these groups. For this purpose, we used the \( sDI \) (sGB-R-loop Dynamics Index) and \( asDI \) (asTSS-R-loop Dynamics Index) to represent the variability of the sGB-R-loops and asTSS-R-loops, respectively, in a certain gene or TE (see Supplemental Figure 5A for details). According to our algorithm, when a gene or TE had an increased \( sDI \) or \( asDI \), this indicated that it had increased sGB-R-loop or asTSS-R-loop dynamics, respectively. The \( sDI \) and \( asDI \) values for the same gene were unrelated, as the Spearman rank correlation coefficient was only 0.1036 (Figure 2F, see Supplemental Figure 5B for details of the method), indicating that there was a very weak correlation between sGB-R-loop and asTSS-R-loop dynamics. Furthermore, for all genes and TEs, the total \( sDI \) value was higher than the total \( asDI \) value (Figure 2G), suggesting that levels of asTSS-R-loops were more stable than sGB-R-loops and revealing the unconnected and diverse regulatory mechanisms involved in the functioning of sGB-R-loops vs. asTSS-R-loops.

**R-loop turnover during generational switches and recovery from long heat stress**

The results in Figure 2B and 2C showed the dramatic R-loop changes during several developmental stages and after long heat stress recovery. Thus, we further analyzed these R-loop dynamics in detail. Cluster analysis showed that the R-loops in flower exhibited the most dissimilar pattern compared to other vegetative tissues, including root, seedling leaf, young leaf and old leaf, whereas germinating plants exhibited the second most dissimilar pattern (Supplemental Figure 3A). Notably, the dynamic pattern observed during germination was very close to that observed in flower, as their Euclidean distance was smaller than that between any other samples (Supplemental Figure 3A). Furthermore, strong overlap in the DRGs associated with the germination stage and flower compared to those in seedlings were also
observed (Supplemental Figure 4D). Altogether, these results reveal a turnover pattern in which R-loop levels showed global changes when the plant entered the reproductive stage from the vegetative stage. Most of these changes were restored to the patterns of the vegetative stage during seed germination.

To further explore the R-loop turnover pattern, fuzzy clusters derived from the sGB-R-loop and asTSS-R-loop DRGs during development were analyzed. We identified four clusters of sGB-R-loop and asTSS-R-loop DRGs (Figure 3A). Cluster 1 represents the flower-specific hyper-R-loop genes, which had increased R-loop levels in flowers but decreased R-loop levels in the germination stage, roots, seedlings, and young and old leaves. Cluster 2 represents leaf-specific hyper-R-loop genes, which exhibited increased R-loop levels in seedlings and young and old leaf than in other samples. Cluster 3 represents germination-specific hyper-R-loop genes. Cluster 4 represents vegetative period-specific hyper-R-loop genes. Of these four clusters, Clusters 2 and 4 represent genes with regeneration R-loop turnover patterns, while the other 2 clusters, 1 and 3, represent flower and germination-specific patterns, respectively.

Notably, no cluster showed high R-loop levels in germination and flower samples simultaneously. Moreover, sGB-R-loop DRGs and asTSS-R-loop DRGs from the same clusters showed surprisingly little overlap (Figure 3B), further suggesting that the regulatory mechanisms of sGB-R-loops and asTSS-R-loops are independent, as shown in Figure 2F. Furthermore, the ratios of genes, TEs, and TE genes (definition provided in the Methods) in sGB-R-loop and asTSS-R-loop DRGs from the same clusters were almost the same (Figure 3C).

To systematically explore the influence of different chromatin features on sGB-R-loop and asTSS-R-loop DRGs, we classified the DRGs into nine different chromatin states (Sequeira-Mendes et al., 2014) by sorting the genomic regions based on various epigenetic markers. The DRGs were further
grouped into the active state (1, 3, 6, 7), repressed state (2, 4, 5), and silent state (8, 9). We noticed that the DRGs of sGB-R-loop and asTSS-R-loop were correspondingly dispensed in different states with the same ratio (Figure 3D; Supplemental Figure 6A). The ratio of genes, TEs, and TE genes in a cluster was consistent with chromatin state distribution, where Cluster 1 contained many genes in the active state while Cluster 3 contained many TEs and TE genes in the silent state (Figure 3C and 3D; Supplemental Figure 6A).

We analyzed the dynamic patterns of R-loops in seedling samples grown under different temperature conditions. Heat map analysis revealed that the most DRGs were in the LHS, LHSR12, and LHSR84 samples (Supplemental Figure 3C), as mentioned above (Figure 1E; Supplemental Figure 2D). Therefore, we analyzed the R-loop dynamics of the DRGs in the CTRL-7d, LHS, LHSR12, and LHSR84 samples using Mfuzz. This analysis also revealed four clusters of sGB-R-loop and asTSS-R-loop DRGs (Supplemental Figure 6B). Cluster 1 showed a persistent increase in R-loop levels after long heat stress, while Cluster 2 showed a persistent decrease. The R-loop levels in Cluster 3 genes increased after long heat stress and then recovered. The SGB-R-loops and asTSS-R-loops exhibited different dynamic trends in Cluster 4 genes.

We then investigated the overlap between sGB-R-loop DRGs and asTSS-R-loop DRGs with similar dynamic trends. We detected a weak overlap between sGB-R-loop DRGs and asTSS-R-loop DRGs in the same clusters (Supplemental Figure 6C). The ratios of TE and TE genes and the distribution of chromatin states of sGB-R-loop and asTSS-R-loop DRGs from the corresponding clusters were also almost the same, except for Cluster 4 (Supplemental Figure 6D and 6E). These results indicate that R-loop turnover also occurred during the recovery from long heat stress and most sGB-R-loop and asTSS-R-loop DRGs shared similar dynamic patterns, but these did not occur in the same genes. These results, combined with the R-loop dynamics during development, suggest that sGB-R-loops and asTSS-R-loops share
similar dynamic patterns and that R-loops are more variable in heterochromatin regions. Consistent with the data described above, we also observed a pattern of uncoupling between sGB-R-loops and asTSS-R-loops. Next, we determined the total levels of R-loops in plants during different developmental stages and in the recovered LHS samples by immunostaining with S9.6 antibody. Consistent with our sequencing data, decreased R-loop levels were detected during the germination stage and in flowers (Supplemental Figure 7A). In addition, the levels of R-loops in plants subjected to long heat stress initially increased, followed by a decrease after 5 days of recovery (Supplemental Figure 7B). The presence of R-loop turnover during development and following recovery from long heat stress was also validated by DRIP-qPCR (Supplemental Figure 7C).

Unique R-loop dynamics in ONSEN and Pol III-transcribed genes
During our analysis of R-loop regulatory patterns, we observed some loci that showed specific R-loop dynamics. For instance, ONSEN is a subgroup of copia-type retrotransposons that are activated by long heat stress (Ito et al., 2011). Notably, the sGB-R-loop levels on ONSEN genes greatly increased (more than 100-fold) after long heat stress and decreased after recovery (Figure 4A). It was reported that extrachromosomal ONSEN DNA copies were synthesized after long heat stress (Ito et al., 2011), suggesting that the increased R-loop levels observed in the current study were due to increased DNA copy number. To address this, we performed DRIP-qPCR using total DNA as input. The results still showed a significant R-loop level increase for one ONSEN gene (AT1G11265, Figure 4A), supporting the finding that the changes in ssDRIP-seq signals in ONSEN genes were due to increased R-loops levels. These increases could be due to the intermediate state of complementary DNA synthesis during reverse transcription. RNA Pol III is an essential RNA polymerase involved in the transcription of several noncoding functional RNAs, such as tRNAs and snoRNAs.
(Abascal-Palacios et al., 2018). We previously demonstrated that Pol III-transcribed loci show only sense R-loop formation (Xu et al., 2017). To further explore the dynamics of R-loops in each tRNA and/or snoRNA gene, we analyzed all profiled samples in depth. R-loops in either tRNA or snoRNA genes were significantly altered in the DK (Figure 4B), germination, flower (Supplemental Figure 8A), LHSR12, and LHSR84 samples (Supplemental Figure 8B). Almost all DRGs were downregulated in these samples (Figure 4B; Supplemental Figure 8A, and 8B). The changes in the DK samples depended on the specific stages of R-loops in RNA Pol III-transcribed genes, as they changed dramatically, while other genes showed few or no differences in R-loops (Figure 2D).

We further analyzed the levels of sense and antisense R-loops in tRNAs or snoRNAs. As shown in the metaplots, only sense R-loops were enriched in tRNAs and snoRNAs; the sense R-loop levels in the DK samples decreased by more than 50% compared to the control (Figure 4C and 4D). These results indicate that the downregulation of sense R-loops occurred in almost every tRNA or snoRNA gene examined (Figure 4D). Furthermore, compared to seedling leaves, more than 300 tRNA and approximately 50 snoRNA genes exhibited decreased R-loop levels in flowers or during germination, while only ~20 tRNA or snoRNA genes exhibited changes in R-loop levels in young leaves (Figure 4E). In the DK samples, R-loop levels were reduced in 95 tRNA and 35 snoRNA genes (Figure 4E). Considering that there were much fewer DRGs in the DK samples than in the flower or germination samples, the percentage of Pol III-transcribed DRGs per total DRGs was much higher than that for the other samples (Figure 4F). Furthermore, more DRGs were present in Pol III-transcribed genes (Figure 4G), suggesting that Pol III-transcribed genes showed more dynamic changes in R-loops.

These analyses revealed similar R-loop dynamics in Pol III-transcribed genes in flower, germination, and DK samples. Hence, we determined whether the same Pol III-transcribed DRGs were regulated in these three groups of
samples. As shown in the Venn diagrams, there was a strong overlap among the Pol III-transcribed DRGs in the flower, germination, and DK samples (Supplemental Figure 8C). This result suggests that R-loop turnover in Pol III-transcribed genes occurs during development and that the presence of light after germination might be required to facilitate R-loop turnover in Pol III-transcribed genes.

Altogether, two groups of genes, ONSEN and Pol III-transcribed genes, exhibited clear differences in dynamics patterns compared to other genes, suggesting that complex and variable regulatory mechanisms influence R-loop dynamics in Arabidopsis.

Weak coupling between R-loop status and RNA abundance

Most instances of R-loop formation are co-transcriptional and dynamically generated and resolved (Sanz et al., 2016). To address whether R-loop status is coupled with fluctuations in RNA abundance, we examined samples subjected to long-term heat stress, as they exhibited dramatic changes in both RNA expression (Zhang et al., 2017) and R-loops (Figure 2B). We first compared the total RNA-seq and ssDRIP-seq data from LHS samples and controls in scatter plots. The dynamic ranges of sGB-R-loops and asTSS-R-loops were much smaller than that of RNA (Figure 5A and 5B). Meanwhile, we detected a weak linear correlation for changes in sGB-R-loop/asTSS-R-loop and RNA levels, as the $R^2$ value was 0.057 for sGB-R-loops and 0.005 for asTSS-R-loops (Figure 5A). Clustering analysis showed that the patterns of both sGB-R-loops and asTSS-R-loops were clearly different from the RNA expression patterns (Figure 5C). The differences in R-loops in the LHS vs. control samples were relatively small and gradually increased during recovery, while there were great differences in RNA abundance between the LHS and control samples, followed by a decrease in recovery (Figure 5C).
There were also several cases of the uncoupling of R-loop status and RNA levels (Figure 5D and 5E; Supplemental Figure 9). For example, the RNA level of \textit{HSP21} was greatly increased in the LHS samples (more than 100-fold vs. the control), but the levels of sense and antisense R-loops were unchanged (Figure 5D). For other protein-coding genes, only the RNA level was altered, while the R-loop level was stable (Figure 5D) or even reduced (Supplemental Figure 9A and 9B). Exceptions were observed at \textit{ONSEN} gene loci, in which R-loop and RNA levels increased simultaneously under LHS and then decreased after recovery (Figure 5E; Supplemental Figure 9C). However, this type of synchronous change in R-loop and RNA level was present only in the \textit{ONSEN} loci and not in other TEs or heat-stimulated gene loci (Figure 5E and 5F). Besides, the overlap between differential sense/antisense RNA genes and differential sGB-/asTSS-R-loop genes was analyzed. We also detected weak overlap between each pair of samples (Supplemental Figure 9D), confirming the weak coupling between R-loop status and RNA abundance.

Overall, these results reveal the weak coupling of R-loop status and RNA level dynamics induced by LHS. We examined whether this phenomenon is prevalent in Arabidopsis. Analysis of our ssDRIP-seq data together with published RNA-seq data from roots, flowers, seedling leaves (GSE38612) (Liu et al., 2012) and germinating seedlings (GSE94712) (Kawakatsu et al., 2017) indeed revealed weak linear correlations between changes in R-loop and RNA levels in flower/root and seedling leaf tissue (Figure 5G and Supplemental Figure 10A), which is similar to our findings for heat stress. Fuzzy cluster analysis of RNA-seq data from the four samples mentioned above generated four clusters with trends similar to the R-loop clusters (Figure 3A) (Supplemental Figure 10B). However, although the dynamics patterns were similar, only weak overlaps were observed between the RNA clusters and sGB-/asTSS-R-loop clusters (Supplemental Figure 10C). In addition, the chromatin state enrichments of the four RNA clusters were very similar.
Altogether, our results indicate that R-loop dynamics are weakly coupled with RNA abundance dynamics genome-wide, an important finding for understanding R-loop dynamics and functions.

A small number of R-loops are correlated with R-loop regulators

A number of R-loop regulators have been investigated in yeast, animals and plants (Santos-Pereira and Aguilera, 2015). We wondered whether these regulators modulate R-loops genome-wide through their effects on RNA dynamics. To address this question, we analyzed the R-loop regulatory network induced by LHS, representing the dynamic relationship between the RNA levels of known and predicted R-loop regulators as well as the R-loop levels of all genes and TEs. Eleven R-loop regulators exhibited significant linear correlations with either sGB- or asTSS-R-loops (Figure 6A and Supplemental Figure 11A; Supplemental Table 1), suggesting that these R-loop regulators might be involved in regulating at least a portion of R-loop loci during long-term heat stress responses. These 11 R-loop regulators were sorted into two clusters that might target different R-loop loci groups (Figure 6A; Supplemental Figure 11A). This result suggests the presence of different R-loop regulatory pathways, resulting in differential R-loop dynamics. It was also noticed that more sGB-R-loop loci were correlated with R-loop regulators, while only a few asTSS-R-loop loci were correlated. This might be due to the higher stability of the asTSS-R-loops or the various R-loop regulation pathways. Overall, only a few R-loop loci showed significant linear correlations with known or predicted R-loop regulators, likely because the dynamic relationship between R-loop regulators and R-loops is nonlinear, a notion we could not analyzed with the current linear model (see Methods). Alternatively, perhaps some R-loop regulators with greater or more prevalent functions are still unknown.
After analyzing how R-loop regulators modulate R-loop levels, we examined how R-loops regulate RNA levels. First, genes and TEs that showed both RNA and R-loop level alternations by LHS were analyzed. We also found that most of the changed genes showed a positive correlation with sGB-R-loops and RNAs after exposure to LHS, but few asTSS-R-loops were changed along with RNAs (Figure 6B). After 12 or 84 h of recovery, a negative correlation of both sGB-R-loops and asTSS-R-loops with changed genes was found. Positively correlated loci for R-loops and RNAs were found for the TEs and TE genes, suggesting a differential R-loop regulation mechanism connecting TEs/TE genes and normal genes. Since asTSS-R-loops cannot be byproducts of sense RNA transcription, the potential direction of regulation should be from the asTSS-R-loop to RNA. Surprisingly, although some asTSS-R-loops showed a negative correlation with RNA levels, a number of asTSS-R-loops showed a positive correlation, suggesting that asTSS-R-loops might facilitate RNA transcription.

Although many simultaneously changed genes were found as mentioned above, most genes showed stable asTSS-R-loops levels even though the RNAs were greatly changed (Figure 5). Interestingly, further analysis revealed that the asTSS-R-loop levels in LHS-induced differentially expressed genes, which were sorted into four clusters based on their dynamic patterns (Supplemental Figure 11B), were much higher than those in genes with unchanged RNA levels (Figure 6C). These results indicate that transcriptionally dynamic genes had higher asTSS-R-loop levels, suggesting that asTSS-R-loops might play a role in the transcriptional regulation of nearby genes. We speculated that transcription factors (TFs) might recognize and associate with these asTSS-R-loops, as some TFs, such as zinc finger proteins, can bind to DNA:RNA hybrids in vitro (Shi and Berg, 1995).

In Arabidopsis, ABA-induced binding of a number of TFs was previously investigated by analyzing a comprehensive ChIP-seq data set (GSE80564)(Song et al., 2016). We therefore analyzed the relationship
between R-loops and TF binding sites during ABA treatment. We analyzed the
colocalization of R-loop peaks and TF binding regions. A permutation test
indicated that sense R-loop peaks showed insignificant or weak negative
colocalization relationships with TFs located in gene body regions, whereas
antisense R-loop peaks were significantly enriched in TF binding regions in
gene promoters (Figure 6D). Finally, metaplot analysis showed that
ABA-induced genes \textit{ABF1} and \textit{HB7} were positively enriched in antisense
R-loop peaks and negatively enriched in sense R-loop peaks (Figure 6E and
6F; Supplemental Figure 12). Together, these results demonstrate that
antisense R-loops and TFs are co-localized, raising the possibility of that gene
transcription is regulated through TF binding or recognition of DNA:RNA hybrid
structures near TSSs.
DISCUSSION

An R-loop atlas for eukaryotic systems

The essential functions of R-loops in different species have been investigated for many years. However, few genome-wide R-loop sequencing data sets are currently available, as wild-type lines of only a few species and a small number of mutants and transgenic lines have been analyzed in most studies, limiting our understanding of R-loops from a general perspective. Here, we profiled and characterized R-loop dynamic patterns in 53 Arabidopsis samples from different developmental stages and grown under different light and temperature conditions, as well as in the presence of many different environmental stimuli. Based on these samples, we built an R-loop atlas for higher plants. This R-loop atlas will greatly extend our understanding of R-loop biology.

R-loop patterns are relatively conserved and weakly coupled with changes in RNA abundance

At the genome level, the distribution and intensity of both wR-loop and cR-loop peaks were relatively similar among most samples. At gene level, the sense R-loops were highly enriched in gene bodies, and antisense R-loops were mainly detected around TSSs in all of the samples, which also showed a conserved pattern. These conserved R-loop distribution patterns suggest that R-loops, at least in conserved peak regions, represent a solid chromosome structure that contributes to chromosomal composition in addition to the basic structure of DNA and histones. The features of conserved R-loop patterns are quite different from those of other chromosomal structures, such as histone modifications, for which the distribution and intensity show dynamic changes globally or specifically at certain loci during plant development and in response to environmental stimuli (Brusslan et al., 2015). This raises the question of whether a R-loop structure influences chromatin remodeling via structural change such as histone modifications. It would be worth investigating whether
R-loop-containing regions are prone to chromatin expansion, considering the unique loose three-strand structure of R-loops. It is also possible that some conserved R-loops anchor R-loop readers to facilitate important cellular processes.

Although a steady-state RNA level does not represent an ongoing transcriptional process, fluctuations in RNA abundance reflect transcriptional activity changes. One important finding of this study is that R-loop dynamics are not strongly correlated with changes in RNA abundance. Our analyses of different samples revealed weak linear correlations between R-loops and RNA abundance, there were even negative correlations at some loci. In the case of long-term heat stress, the differences in R-loop levels between the LHS and control groups were quite small, whereas RNA rapidly but transiently accumulated to very high levels in response to long-term heat stress. The weak coupling of the relationship between R-loop and RNA dynamics could be partially attributed to the finding that the lifetime of RNA is short, averaging only two minutes (Baudrimont et al., 2017), as RNA is either translated to protein or degraded, whereas R-loop structures are relatively stable, with an average half-life of 10 minutes at promoter regions to 2 h at some terminator regions (Sanz et al., 2016). The independent dynamics and conserved patterns of R-loops suggest that organisms use unique mechanisms to establish, maintain, and resolve R-loop formation, supporting the idea that R-loops play important biological roles rather than merely being byproducts of transcription.

Specific loci with different R-loop dynamic patterns

Unlike RNA, the abundance of chromatin R-loops at specific loci is limited by DNA copy number. We speculate that the increase in R-loop levels at the ONSEN loci in the LHS samples resulted from high levels of reverse transcription of TE RNA for the following reasons. First, both R-loops and RNA were strongly activated simultaneously by LHS, which decreased after
recovery, revealing a synchronous change pattern. Second, the R-loop level in the ONSEN loci increased so much after heat stress that such an increase would require more gDNA copies than that contained in the input. It seems that the only way this could occur is using the RNA, as one gDNA template can be transcribed into many RNA copies, and TE can reverse transcribe these RNAs to form DNA:RNA hybrids. Such extrachromosomal DNA:RNA hybrids can also be detected by ssDRIP-seq or ssDRIP-qPCR. Third, the transcription of the other TE, ATHILA6A, was also activated by LHS, but it did not show an increase in R-loop levels. An analysis of the open reading frames in TEs showed that ATHILA6A lacks a reverse transcriptase domain, while the ONSEN loci contain all components necessary for reverse transcription, which supports our hypothesis.

In this study, we found that the sense R-loop levels in RNA Pol III-transcribed tRNA and snoRNA loci showed significant decreases in the flower, germinating seedlings, and DK samples. The Pol III-transcribed DRGs showed strong overlaps among these samples, suggesting that these unstable but conserved changes in R-loops might share the same regulators. Considering the conditions of these three samples, light might have been the factor that triggered these changes. Since the dynamic changes at these loci clearly occurred in chromatin, this indicates that R-loop levels showed dynamic changes under specific conditions.

**R-loop turnover during generational switches**

Another interesting discovery of the current study is R-loop turnover, with periodic and regular resolution and rebuilding of R-loops during the life cycle. In addition, R-loop levels on a cluster of genes increased after long-term heat stress and then decreased, and most sGB-R-loop and asTSS-R-loop DRGs shared similar dynamics patterns. These findings imply that these R-loops play conserved roles and that their dynamics are required for the normal plant life cycle and heat stress responses. It appears that light is required during seed
germination to rebuild R-loops on Pol III-transcribed genes rather than Pol II-transcribed genes. Considering the important role of tRNAs in germination, light-induced R-loop turnover in tRNAs might take place during the regulation of tRNA transcription.

R-loop regulation and R-loops as potential *in cis* regulators

Many R-loop regulators have been reported in different species, most of which are considered to be R-loop resolvers (Santos-Pereira and Aguilera, 2015). However, our analyses indicated that only a few R-loop loci showed significant linear correlations with the RNA levels of these R-loop regulators. A possible explanation is that the regulatory mechanism operates in a more complex manner and that the correlation between R-loops and regulators is not linear, as the regulatory mechanism operates in a complex nonlinear manner. Another possibility is that more important R-loop regulators that affect most R-loop loci in Arabidopsis have not yet been identified. As R-loops are conserved chromosome structures with stable patterns, although R-loops showed a genome-wide turnover pattern in both germinating and flowering tissues, some R-loops could be transgenerational and thus were not regulated by any regulator generation after generation. To test this hypothesis, more precise profiling in pollen, ovules, and single-celled zygotes should be performed.

We previously demonstrated that antisense R-loops are enriched around TSSs in Arabidopsis. These R-loops, which are known as asTSS-R-loops, have not yet been observed in any other species. The dynamics of asTSS-R-loops are independent of the dynamics of sense R-loops, suggesting that asTSS-R-loops have unique regulatory mechanisms and functions. The *in cis* dynamic relationship between sGB-R-loops and RNA is not always unidirectional. On one hand, R-loops in genes might affect RNA transcription. On the other hand, although not common, it is still possible that the RNA transcriptional changes in some genes directly affect R-loop levels. Because
asTSS-R-loops cannot be byproducts of RNA transcription, the potential direction of regulation is likely to be from asTSS-R-loops to RNA. In addition to the enrichment pattern of asTSS-R-loops, although some asTSS-R-loops showed a negative correlation with RNA levels, a number of asTSS-R-loops showed a positive correlation. These observations prompted us to propose that asTSS-R-loops play a regulatory role in RNA transcription, that is asTSS-R-loops might facilitate RNA transcription of nearby genes. Surprisingly, TFs were enriched around asTSS-R-loop regions, suggesting a new working mechanism involved in TF activity and providing novel insights into R-loops and gene regulation.
METHODS

Plant materials and growth condition

All Arabidopsis thaliana plants are Columbia ecotype (Col-0). Plants were cultured on 1/2 MS medium (Murashige and Skoog salt base; Sigma, USA; 0.25% phytagel, 1% sucrose) or soil under long-day conditions at 22°C with a 16 h of light/8 h of dark cycle, if not otherwise stated. The details of plant materials used in this study are listed in Supplemental Data Set 1.

ssDRIP-seq library construction

Nuclei were isolated from 2~5 g of Arabidopsis seedlings or tissues, followed by SDS/proteinase K digestion at 37°C for 6~12 hours. Genomic DNA was extracted by the phenol-chloroform method and precipitated with 1 volume isopropanol at room temperature. DNA was fragmented at 37°C for 6~12 hours using endonucleases (Ddel, Msel, Mbol and NlaIII; New England Biolabs; 0.1 U/μl final concentration for each enzyme). The negative control was treated with RNase H (New England Biolabs, 0.5 U/μl finial concentration) at 37°C overnight. DRIP was performed as described previously (Xu et al., 2017), with 3.5 μg gDNA input for each sample. The S9.6 antibody was purified from HB-8730 (ATCC).

The DRIPed DNA was sonicated to ~250 bp with an M220 Focused-ultrasonicator (Covaris). The sonicated DNA was used to construct ssDRIP-seq library by using the Accel-NGS 1S Plus DNA Library Kit (Swift Biosciences) following instructions from the manufacturer. The libraries were checked on an Agilent BioAnalyzer, followed by sequencing on an Illumina HiSeq X Ten system. For each biological replicate, more than 23 million mapped reads were obtained from each ssDRIP-seq library except the RNase H negative control (Supplemental Data Set 1).

The ssDRIP-seq data were deposited in the NCBI Sequence Read Archive (SRA) database under accession number GSE116232.
RNA-seq

CTRL-7d (7-day-old Col-0 seedlings), LHS (7-day-old Col-0 seedlings, 37°C 30 hours), LHSR12 (7-day-old Col-0 seedlings, recovery 12 hours after LHS), and LHSR84 (7-day-old Col-0 seedlings, recovery 84 hours after LHS) were used for total RNA-seq. NEB Next Ultra Directional RNA Library Prep Kit for Illumina (NEB, Ipswich, USA) was used to prepared RNA-seq library in a strand-specific manner. The RNA-seq libraries were sequenced on an Illumina HiSeq X Ten system. The RNA-seq data were deposited in the NCBI Sequence Read Archive (SRA) database under accession number GSE118338.

Sequencing data processing

For ssDRIP-seq, reads were aligned to the TAIR10 genome with Bowtie 2 using default settings (Langmead and Salzberg, 2012), with all duplicates removed by Picard tools (http://broadinstitute.github.io/picard). The total mapped reads (non-strand R-loops) were divided into forward (wR-loops, representing an R-loop formation containing single-strand DNA on the Watson strand and an DNA:RNA hybrid on the Crick strand) and reverse reads (cR-loops) (Xu et al., 2017) by using samtools (Li et al., 2009).

DESeq2 (Love et al., 2014) was used to analyze R-loop differences between samples. The FDR (false discovery rate)(Benjamini and Hochberg, 1995) method was used to calculate adjusted p value (q value), and the q values and log2FC values were used to determine significant difference. MACS2 (Feng et al., 2012) was used to identify peaks with default parameters.

For visualization, the aligned reads files (BAM) were converted to normalized coverage files (bigWig) with 1 bp bins using bamCoverage from deepTools (Ramirez et al., 2016). Normalization was performed using bamCoverage from deepTools, with read coverage normalized to 1× sequencing depth (also known as Reads Per Genomic Content) with a re-normalization by shuffled peaks (total R-loop peaks were shuffled randomly, and the 95% mean of...
R-loop signal from each sample on shuffled peaks was used as denominator) to eliminate the disturbances from abnormally high value regions. Snapshots of the data were constructed using the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Spearman correlation coefficients were calculated with plotCorrelation from deepTools using 500 bp bins. Heatmaps were generated with computeMatrix from deepTools. Metaplots were generated with deepTools and R scripts, representing the median, mean (arithmetic mean), or 95% mean (with both the top and bottom 2.5% data points discarded to reduce the disturbance from extreme values) of read coverage or other types of data over the chosen features of interest. Scatter plots were generated by MATLAB or R scripts.

For RNA-seq, reads were mapped to the TAIR10 genome using STAR (Dobin et al., 2013) with the following options --outMultimapperOrder Random --outFilterIntronMotifs RemoveNoncanonicalUnannotated. Reads counting in genes and TEs performed by using HTseq (Anders et al., 2015) with the parameter -s reverse -f bam -r pos -m union. Part of genes or TEs share the same positions, so these loci were extracted and counted reads using following parameter -s reverse -f bam -r pos -m intersection-nonempty --nonunique all. Differential expression gene analysis was performed by using DESeq2 (Love et al., 2014) and NOISeq (Tarazona et al., 2012). DESeq2 was used to analyze RNA-seq data of LHS set, while the q values and log2FC values were used to determine significant difference. NOISeq was used to analyze the RNA-seq data of germination (GSE94712), root, leaf, and flowers (GSE38612) which lacking replications, and the gene which prob value > 0.9 was defined as a differentially expressed gene.

Fuzzy cluster analysis was performed by using Mfuzz (Kumar and M, 2007). For R-loop cluster analysis, all DRGs between any two samples were analyzed. For RNA cluster analysis, all DEGs between any two samples were analyzed. Network analysis was performed by following steps. First, the union of DEGs and DRGs between any two samples from LHS group were chosen to perform
further analysis. The Pearson correlation coefficients between the RNA-seq and ssDRIP-seq signal on the union genes mentioned above were calculated and a threshold of 0.75 was used. Cytoscape (Shannon et al., 2003) was used to visualize the networks.

TEs (transposable elements) were annotated by TAIR10_transposable_elements (https://www.arabidopsis.org/download_files/Genes/TAIR10_genome_release/TAIR10_transposable_elements/TAIR10_Transposable_Elements.txt). TE genes (transposable element genes) were annotated by TAIR10_GFF3_genes (https://www.arabidopsis.org/download_files/Genes/TAIR10_genome_release/TAIR10_gff3/TAIR10_GFF3_genes.gff).

**DRIP-qPCR**

DRIP-qPCR were performed as described (Xu et al., 2017). The primers used in this study were listed in the Supplemental Table 2.

**Immunofluorescence**

All plant tissues were cross-linked with 4% formaldehyde at 4°C for 60 min and chopped with a razor in NEB buffer (10 mM Tris-HCl pH 9.5, 10 mM KCl, 500 mM sucrose, 0.10% β-mercaptoethanol, 0.10% Triton X-100) for 30 min. Homogenate was filter purified with 100 μm nylon filter and centrifuged at 2000 rpm at 4°C. The nuclear pellet was gently suspended in NEB 2 buffer (10 mM Tris-HCl pH 9.5, 10 mM KCl, 125 mM sucrose, 0.10% β-mercaptoethanol, 0.10% Triton X-100), and layered on the top of NEB 3 buffer (10 mM Tris-HCl pH 9.5, 10 mM KCl, 850 mM sucrose, 0.10% β-mercaptoethanol, 0.10% Triton X-100). After centrifugation at top speed at 4°C, nuclear pellet was resuspended in 50 μl NEB1 buffer and drop 30 μl nuclear pellet onto the slide. The slide was dried at room temperature, and re-fix the nuclei with 4% formaldehyde in PBST for 1 min. After washing 3 times with 1x PBST, nuclei on slide were blocked with 3% BSA in PBST and incubated with S9.6 anti-body at 4°C overnight. After
washing 3 times with 1x PBST, the slide was incubated with secondary Goat
Anti-mouse antibody (Alexa Fluor 488) for 1 hours at 37°C. The slide was
washed with 1x PBST and add a drop of ProLong Gold antifade reagent with
DAPI, and then covered the slide with cover glass and analyzed under
confocal microscope.

Accession Numbers

The ssDRIP-seq, RNA-seq data and processed files are available in NCBI's
Gene Expression Omnibus under accession number GSE116232 and
GSE118338, detailed information of tools and software used in this study are
listed in Supplemental Table 3. The database of The R-loop Atlas can be found
in http://bioinfor.kib.ac.cn/R-loopAtlas.

Supplemental Data

Supplemental Figure 1. Introduction of ssDRIP-seq.
Supplemental Figure 2. Overview of R-loop atlas in Arabidopsis.
Supplemental Figure 3. Dynamic patterns of sGB-R-loops and
asTSS-R-loops.
Supplemental Figure 4. R-loop dynamics in light, SA, UVC, and ABA groups.
Supplemental Figure 5. Algorithm of sDI (sGB-R-loop Dynamics Index) and
asDI (asTSS-R-loop Dynamics Index).
Supplemental Figure 6. Cluster analysis of R-loops changed in LHS group.
Supplemental Figure 7. Validation of R-loop turnover by immunostaining and
DRIP-qPCR.
Supplemental Figure 8. Unique R-loop dynamics on tRNA and snoRNA
genes.
Supplemental Figure 9. Uncoupling between R-loop and RNA expression
changes induced by LHS.
Supplemental Figure 10. Uncoupling between R-loop and RNA expression
changes during development.
Supplemental Figure 11. Network of R-loop regulators on asTSS-R-loop, and cluster analysis of RNA expression in LHS group.

Supplemental Figure 12. Co-localization between antisense R-loops and TF binding sites.

Supplemental Table 1. List of R-loop regulators analyzed in this study.

Supplemental Table 2. DRIP-qPCR primers used in this study.

Supplemental Table 3. Information of software used in this study.

Supplemental Data Set 1. Information of ssDRIP-seq data, condition descriptions, and R-loop peaks used in this study.
ACKNOWLEDGEMENTS

We greatly appreciate all The Sun Lab members for the useful discussions. This work was funded by grants from the National Natural Science Foundation of China (grant nos. 91740105, 31822028 and 91940306 to Q. Sun, 31900302 to W. Xu, and 31971333 to K. Liu), and the Ministry of Science and Technology of China (2016YFA0500800 to Q. Sun). W. Xu and K. Liu are supported by the postdoctoral fellowships from Tsinghua-Peking Joint Center for Life Sciences. The Sun Lab is supported by Tsinghua-Peking Joint Center for Life Sciences and the 1000 Young Talent Program of China.

AUTHORS CONTRIBUTIONS

Q. Sun conceived, designed, and supervised the experiments. W. Xu, S. Li, and Q. Hou were responsible for the R-loop profiling experiments. W. Xu and K. Li were responsible for the data analysis. K. Liu was responsible for the data analysis of R-loops on tRNA or snoRNA genes. Y. Zhang was responsible for the immunofluorescence experiments. Q. Sun and W. Xu wrote the manuscript with contributions from S. Li and Q. Hou. All authors read and approved the final manuscript.
**Figure Legends**

**Figure 1. The landscape of R-loops in Arabidopsis.**

(A) Overview of the Arabidopsis samples used in this study. Four sets of samples, including plants exposed to different light, temperature, and stress conditions and at different developmental stages, are shown in the left panel. Specific samples from each set are shown in the right three panels.

(B) Brief workflow of ssDRIP-seq. First, the genomic DNA is extracted from Arabidopsis samples. After DNA fragmentation with restriction enzymes, only DNA:RNA hybrids are pulled down by S9.6 antibodies which specifically recognize and bind to DNA:RNA hybrids of various lengths. Finally, the library is prepared using an ssDNA-based strand-specific library kit and sequenced on an Illumina HiSeq X Ten System.

(C) IGV snapshots of the DRIP-seq signals from all samples (one of two replicates is shown). The wR-loops (Watson strand R-loops) are shown in red, while the cR-loops (Crick strand R-loops) are shown in blue. Samples are represented by different colors, and the detailed information is shown at the bottom.

(D) Heatmap of Spearman’s rank correlation coefficients between two samples. Samples are distinguished by color, as shown in Figure 1C.

(E) Metaplots of sense (left) and antisense (right) R-loop signals from samples at different developmental stages (upper) and under different light (middle) and temperature (lower) conditions, centered on total genes with 1 kb up- and downstream extensions. Line, 95% mean; shadow, 95% confidence interval.

**Figure 2. Dynamics of sGB-R-loops and asTSS-R-loops.**

(A) Diagram of an sGB-R-loop (sense gene body R-loop) and an asTSS-R-loop (antisense transcription start site R-loop). The sGB-R-loop is defined as a sense R-loop on the entire gene body, from the TSS (transcription start site) to the TTS (transcription termination site). The asTSS-R-loop is defined as an antisense R-loop around TSS, ± 150 bp.
(B) Scatter plots of the levels of sGB-R-loops (left panel) and asTSS-R-loops (right panel) per total genes and TEs. Samples exposed to different temperature treatments were compared with the control (CTRL-7d). Normalized reads counts are shown as \( \log_{10}(n+1) \). Red dot: q value < 0.05, \( \log_2 \) fold change > 1. Blue dot: q value < 0.05, \( \log_2 \) fold change < -1. Gray dot: other.

(C) Scatter plots of the levels of sGB-R-loops (middle panel) and asTSS-R-loops (right panel) on total genes and TEs (transposable elements, annotated by TAIR10_transposable_elements). Samples from different developmental stages were compared with seedling leaf tissue. Plots of changes in sGB-R-loop levels between flowers and seedling leaves are magnified (left panel) to show the changes in sGB-R-loop levels on TEs, TE genes (transposable element genes, annotated by TAIR10_GFF3_genes), and other genes. Normalized reads counts are shown as \( \log_{10}(n+1) \). Red dot: q value < 0.05, \( \log_2 \) fold change > 1. Blue dot: q value < 0.05, \( \log_2 \) fold change < -1. Gray dot: other.

(D) Scatter plots of sGB-R-loop (left) and asTSS-R-loop (right) levels on total genes and TEs. DK (dark) was compared with CTRL-7d. Normalized reads counts are shown as \( \log_{10}(n+1) \). Red dot: q value < 0.05, \( \log_2 \) fold change > 1. Blue dot: q value < 0.05, \( \log_2 \) fold change < -1. Gray dot: other.

(E) Scatter plot of asTSS-R-loop/sGB-R-loop \( \log_2 \) fold change between LHS and CTRL-7d on total genes and TEs. \( R^2 \) and Rho values (Spearman's rank correlation coefficients) are shown on the plot.

(F) *Dynamics Index* matrix of sGB-R-loops and asTSS-R-loops. \( sDI \) and \( asDI \) were calculated following the algorithm shown in Supplemental Figure 5. Genes assigned to the same \( sDI \) and \( asDI \) are located in the same dot, and the number of genes in the same dot is represented by color and size. The Rho value (Spearman's rank correlation coefficients) of all genes and TEs is shown.
(G) Left, box plot of $sDI$ and $asDI$ for all genes and TEs. Means are represented by crosses. Right, accumulation curves of $sDI$ and $asDI$ from all genes and TEs.

**Figure 3. Cluster analysis of sGB-R-loops and asTSS-R-loops during development.**

(A) Fuzzy cluster analysis of ssDRIP-seq signals in DRGs. The union of DRGs from any two samples of developmental stage set was analyzed. Line plots show standardized ssDRIP-seq signals, with individual gray lines representing individual loci and orange lines representing the values for the cluster center.

(B) Venn diagrams of the DRG clusters from Figure 3A. Light steel blue, sG-R-loops. Light coral, asTSS-R-loops.

(C) Proportional bar charts of TE genes (orange), TEs (blue), and other genes (gray) in the four DRG clusters from Figure 3A. Left, sG-R-loops. Right, asTSS-R-loops.

(D) Proportional bar charts of the chromatin states of the four DRG clusters from Figure 3A. The chromatin states were classified into three types: active state (1, 3, 6, 7), repressed state (2, 4, 5), and silent state (8, 9). State 1, marked with H3K4me2, H3K4me3, H3 acetylation, H3K36me3, often associated with transcribed regions. State 2, relative to state 1, but additionally enriched with H3K27me3. State 3, enriched in H3K4me1, H2Bub, H3K36me3 and H3K4me2/3, representing a transcription elongation signature. State 4, similar with state 2, but accompanied with lower levels of active transcription marks. State 5, corresponding to Polycomb-regulated chromatin. State 6, located in intragenic regions and slightly enriched in H2A.Z and H3K4me1. State 7, annotated as an intragenic state, enriched in H3K4me1, H2Bub, and H3K36me3. State 8, AT-rich heterochromatic state. State 9, GC-rich heterochromatic state.

**Figure 4. Genomic loci with abnormal R-loop dynamic patterns.**
(A) R-loops on ONSEN genes activated by long heat stress. Upper, The R-loop levels of CTRL-7d (CTL, green), LHS (red), LHSR12 (light blue), and LHSR84 (dark blue) on chromosome 1. Middle, IGV snapshots of R-loop signals on two ONSEN genes, AT1G11265 (left) and AT1G58140 (right). Lower, DRIP-qPCR of two loci in AT1G11265. Graphs shown as mean ± SEM of three technical replicates (circle) and are representative of two independent experiments.

(B) R-loop dynamics on tRNA (left) or snoRNA (right) genes. All samples from the light set were normalized to CTRL-7d. Gray dot, q value ≥ 0.05 in any sample. Colored dot, DRG (q value < 0.05) in the corresponding samples.

(C) Metaplots of sense (left) and antisense (right) R-loop levels of CTRL-7d (light blue) and DK (dark blue) on tRNA (upper) or snoRNA (lower) genes. Data shown as 95% mean.

(D) Heat maps of sense and antisense R-loop signals of CTRL-7d (left) and DK (right) on total tRNA (upper) or snoRNA (lower) genes.

(E) Statistics of differential R-loop tRNA (coral) or snoRNA (dark blue) genes from flower, germinating seedling, young leaf, and DK tissue. Flower, germination, and young leaf tissue were compared with seedling leaf tissue while DK was compared with CTRL-7d.

(F) Percentage of differential R-loop tRNA (left) or snoRNA (right) genes in total differential R-loop genes. Background, percentage of tRNA or snoRNA genes in total genes. Flower (light coral), germination (light blue), and young leaf (green) tissue were compared with seedling leaf tissue while DK (dark blue) was compared with CTRL-7d.

(G) Percentage of DRGs in total genes, sorted by tRNA, snoRNA, and other genes. Coral, differential R-loop tRNA genes/total tRNA genes. Dark blue, differential R-loop snoRNA genes/total-snoRNA genes. Gray, other differential R-loop genes/total other genes. Flower, germination and young leaf samples were compared with seedling leaf tissue, while DK was compared with CTRL-7d.
Figure 5. Uncoupling between R-loop dynamics and RNA expression changes.

(A) Scatter plots of RNA/sGB-R-loop (left) and RNA/asTSS-R-loop (right) log$_2$ fold change (LHS/CTRL-7d) on total genes and TEs. $R^2$ and Rho values (Spearman’s rank correlation coefficients) are shown on the plots.

(B) Fold changes (log$_2$) of RNA (green), sGB-R-loop (red), and asTSS-R-loop (blue) intensity (LHS/CTRL-7d).

(C) Heatmaps of RNA (upper), sGB-R-loop (middle), and asTSS-R-loop (lower) differences on total genes and TEs between each pair from the LHS set based on Euclidean distance.

(D) IGV snapshots of ssDRIP-seq and RNA-seq signals from the LHS set for protein-encoding genes HSP21 and AT3G01260, with the Watson and Crick strand values split.

(E) IGV snapshots of ssDRIP-seq and RNA-seq signals of the LHS set for AT1G11265 (ONSEN family gene) and AT2G29165 (other Ty1-Copia family gene), with the Watson and Crick strand values split.

(F) Scatter plot of log$_2$ fold changes (LHS/CTRL-7d) in RNA (x-axis) and sGB-R-loop (y-axis) for ATCOPIA78 (red),ATHILA6A (blue), and other TEs (gray). Linear trend lines and corresponding $R^2$ values are shown.

(G) Scatter plots of RNA/sGB-R-loop/asTSS-R-loop log$_2$ fold change values (upper, flower/seedling leaf. lower, root/seedling leaf) on total genes and TEs. $R^2$ values are shown on the plots.

Figure 6. Regulatory network and potential roles of R-loops in regulating gene expression.

(A) Network of the sGB-R-loop regulatory pathway by some known or predicted R-loop regulators. Green or red hexagons, RNA nodes (regulators). Blue hexagons, R-loop nodes (regulated). RECQL3, RECO-like helicase 3; RH17, DEAD-box RNA helicase 17; RH20, DEAD-box RNA helicase 20;
SEN1L, Senataxin-like helicase 1; RNH1A, RNase H1A; RNH1C, RNase H1B; TOP1ALPHA, Topoisomerase 1α; TOP2, Topoisomerase II; THO5B, a THO/TREX complex component for RNA transport; SPT16, a component of facilitates chromatin transcription (FACT) complex; BRCA1, Breast cancer susceptibility protein 1.

**B** Scatter plots of genes or TE5s in the LHS group with changes in both RNA and R-loop levels. X-axis, RNA log₂FC. Y-axis, sGB-R-loop (left panel) or asTSS-R-loop (right panel) log₂FC. All genes or TE5s in which R-loop and RNA levels changed simultaneously (q value < 0.05 and |log₂FC| > 1) are shown. Red dot, gene. Blue, TE gene. White, TE fragment.

**C** Violin plots of asTSS-R-loop levels (log₂(n+1)) of genes with different dynamics patterns (see Supplemental Figure 11B), 95% confidence interval. The median (white dot) and interquartile range (gray shadow) are shown.

**D** Permutation test of co-localization between sense or antisense R-loop peaks of the CTRL-EtOH-4h samples (shown in Supplemental Data Set 1) and ChIP-seq peaks of transcription factors of the EtOH-treated samples (GSE80564).

**E** Metaplots of ABA-induced ABF1 (left) and HB7 (right) ChIP-seq signal fold changes on sense (light blue) or antisense (light red) R-loop peaks, mean.

**F** Snapshots of ssDRIP-seq signals of the EtOH-treated sample and ABF1 and HB7 ChIP-seq signals of the EtOH- and ABA-treated samples.
Supplemental Figure 1. Introduction of ssDRIP-seq. (Support Figure 1)

(A) Workflow of ssDRIP-seq. 1, the nuclei were isolated from Arabidopsis samples. 2, after extraction from nuclei, the genomic DNA was fragmented by using restriction enzyme cocktails. 3, only the DNA:RNA hybrid of R-loop structure was pulled down by S9.6 antibodies which specifically recognizes and binds DNA:RNA hybrids of various lengths, while another ssDNA from R-loop formation was eluted out. 4, the library was prepared by using a ssDNA-based strand-specific library kit. First, the DRIPed DNA:RNA hybrid was sonicated to ~250 bp fragments and denatured to a ssDNA and a ssRNA. Then 3’end of the ssDNA was ligated to a single stranded adapter by Adaptase. Using a primer paired with the adapter, the ssDNA-adapter (Adp1) was extended to form dsDNA. Another adapter (Adp2) was then ligated to the 5’end of the generated dsDNA. After indexing PCR, standard library was cleaned up by beads. 5, NGS sequencing by Illumina HiSeq X ten system. 6, Data processing, including alignment, split bam as strand, normalization and deep analysis.

(B) Left, the distribution of fragment sizes after digestion by restriction enzymes used in this study. Right, the means and medians of fragment sizes in 5 chromosomes.

(C) Upper, comparison between DRIP-seq and ssDRIP-seq. ssDRIP-seq uses 4 high frequency restriction enzymes to cut dsDNA into short fragments (average size, 52 bp). This ensures the efficient DNA fragmentation and keeps genomic R-loop structures intact and with two short dsDNA wings after separating from gDNA. Short dsDNA wings are crucial for the detachment of ssDNA from R-loop, because the weak binding force of dsDNA wings make the ssDNA easily eluted away in washing steps. After ssDNA library construction, the ssDNA of DNA:RNA hybrid was sequenced. Lower, a snapshot of ssDRIP-seq signal and cut sites of restriction enzymes used in ssDRIP-seq.
Supplemental Figure 2. Overview of R-loop atlas in Arabidopsis. (Support Figure 1)

(A) Metaplots of sense (left) and antisense (right) R-loop signals from ABA groups (upper), or SA and UVC groups (lower), centered on total genes with 1 kb up- and down-stream extension. Line, 95% mean; shadow, 95% confidence interval.

(B) Log\(_2\) fold change of sense (left) and antisense (right) R-loop levels of samples from different developmental stages, normalized to seedling leaf. The metaplots was centered on total genes with 1 kb up- and down-stream extension.

(C) Log\(_2\) fold change of sense (left) and antisense (right) R-loop levels of samples from light set, normalized to CTRL-7d. The metaplots was centered on total genes with 1 kb up- and down-stream extension.

(D) Log\(_2\) fold change of sense (red) and antisense (blue) R-loop levels of samples from long heat stress and recovery groups, normalized to CTRL-7d. The metaplots was centered on total genes with 1 kb up- and down-stream extension.

Supplemental Figure 3. Dynamic patterns of sGB-R-loops and asTSS-R-loops. (Support Figure 2)

(A) Heatmaps of R-loop level differences on total genes and TEs between each pair from developmental stage set based on Euclidean distance. Left, sG-R-loops. Right, asTSS-R-loops.

(B) Heatmaps of R-loop level differences on total genes and TEs between each pair from light set based on Euclidean distance. Left, sG-R-loops. Right, asTSS-R-loops.

(C) Heatmaps of R-loop level differences on total genes and TEs between each pair from temperature set based on Euclidean distance. Left, sG-R-loops. Right, asTSS-R-loops.
Supplemental Figure 4. R-loop dynamics in light, SA, UVC, and ABA groups. (Support Figure 2)

(A) Scatter plots of sGB-R-loop (left panel) and asTSS-R-loop (right panel) level on total genes and TEs. Samples under different light conditions were compared with CTRL-7d. Normalized reads counts were shown as log_{10}(n+1). Red dot: q value < 0.05, log_2 fold change > 1. Blue dot: q value < 0.05, log_2 fold change < -1. Gray dot: others.

(B) Scatter plots of sGB-R-loop (left panel) and asTSS-R-loop (right panel) level on total genes and TEs. Sample treated by SA or UVC was compared with CTRL-14d. Normalized reads counts were shown as log_{10}(n+1). Red dot: q value < 0.05, log_2 fold change > 1. Blue dot: q value < 0.05, log_2 fold change < -1. Gray dot: others.

(C) Scatter plots of sGB-R-loop (left panel) and asTSS-R-loop (right panel) level on total genes and TEs. Sample treated by ABA was compared with EtOH. Normalized reads counts were shown as log_{10}(n+1). Red dot: q value < 0.05, log_2 fold change > 1. Blue dot: q value < 0.05, log_2 fold change < -1. Gray dot: others.

(D) Intersections of DRGs of germination (sky blue) and flower (bisque) on TE gene, TE and other genes.

Supplemental Figure 5. Algorithm of sDI (sGB-R-loop Dynamics Index) and asDI (asTSS-R-loop Dynamics Index). (Support Figure 2)

(A) Work flow of algorithm of sDI and asDI. In brief, for sDI, 1) every two samples in the same group were compared by DESeq and output sGB-R-loop DRGs. 2) if a certain gene was a DRG in this comparison, it would be assigned as “1”, if not, “0”. 3) calculate the sum from all the comparisons. 4) the sum of all the summation from 3) in all groups was defined as sDI. The same way was used for calculating asDI, except the DESeq analysis was performed on asTSS-R-loops.
(B) Demonstration of *Dynamics Index* matrix. Left panel, seven genes, gene A to G, were assigned *sDI* and *asDI* values. These genes were shown in a Cartesian coordinate system, x-axis, *sDI*, y-axis, *asDI*. If different genes assigned the same *sDI* and *asDI* values, they will share the same dot. Thus, color and size were chosen for representing the number of genes in the same dot. Right panel, the distribution of genes in Cartesian coordinate system was analyzed by Spearman's rank correlation coefficient. If the genes were enriched along the diagonal, a higher Rho value would be got, meaning positive correlation. If the genes distributed symmetrically, a lower Rho value (close to zero) would be got, meaning no correlation.

**Supplemental Figure 6. Cluster analysis of R-loops changed in LHS group. (Support Figure 3)**

(A) Bar chart of enrichments (log₂ fold change) of 4 DRG clusters from Figure 3A on each chromatin states, compared with total genes and TEs.

(B) Fuzzy cluster analysis of ssDRIP-seq signal in DRGs. The union of DRGs from any two samples of LHS set was analyzed. Line plots show standardized ssDRIP-seq signal, with individual grey lines representing individual locus, and the orange line representing the values for the cluster center. Upper, sG-R-loops. Lower, asTSS-R-loops.

(C) Venn diagrams between 4 DRG clusters from Figure S5B. Light steel blue, sG-R-loops. Light coral, asTSS-R-loops.

(D) Proportional bar charts of TE gene (orange), TE (blue) and other gene (grey) in 4 DRG clusters from Figure S5B. Upper, sG-R-loops. Lower, asTSS-R-loops.

(E) Proportional bar charts of 4 DRG clusters from Figure S5B on each chromatin states, compared with total genes and TEs.

**Supplemental Figure 7. Validation of R-loop turnover by immunostaining and DRIP-qPCR. (Support Figure 3)**
(A) S9.6 immunostaining in germination, root, old leaf, flower, and RNase H treatment (negative control). Bar = 2 μm.

(B) S9.6 immunostaining in CTRL-7d, LHS, LHSR12, LHSR84, and RNase H treatment (negative control). Bar = 2 μm.

(C) Snapshots of the ssDRIP-seq data in representative genomic regions are shown in the left panel. Regions marked by blue bars were validated by DRIP-qPCR, and the results are shown in the right panel.

Supplemental Figure 8. Unique R-loop dynamics on tRNA and snoRNA genes. (Support Figure 4)

(A) R-loop dynamics on tRNA (left) or snoRNA (right) genes. All samples from developmental stage set were normalized to seedling leaf. Grey dot, q value ≥ 0.05 in any sample. Colored dot, DRG (q value < 0.05) in corresponding samples.

(B) R-loop dynamics on tRNA (left) or snoRNA (right) genes. All samples from temperature set were normalized to CTRL-7d (CTL). Grey dot, q value ≥ 0.05 in any sample. Colored dot, DRG (q value < 0.05) in corresponding samples.

C, Intersections of germination/seedling leaf, flower/seedling leaf, and DK/CTRL-7d differential R-loop genes on tRNA (left) or snoRNA (right).

Supplemental Figure 9. Uncoupling between R-loop and RNA expression changes induced by LHS. (Support Figure 5)

(A) IGV snapshots of ssDRIP-seq (upper) and RNA-seq (lower) signals of LHS set on protein-encoding genes, AT1G01060 and AT4G16146, split by Watson and Crick strand. wR-loop, Watson R-loop. cR-loop, Crick R-loop. wRNA, Watson RNA. cRNA, Crick RNA.

(B) IGV snapshots of ssDRIP-seq (upper) and RNA-seq (lower) signals of LHS set on protein-encoding genes, AT2G22990 and AT1G01010, split by Watson and Crick strand. wR-loop, Watson R-loop. cR-loop, Crick R-loop. wRNA, Watson RNA. cRNA, Crick RNA.
(C) IGV snapshots of ssDRIP-seq (upper) and RNA-seq (lower) signals of LHS set on other ONSEN family genes, AT1G21945, AT1G48710, AT1G58140, AT3G32415, AT3G59720, AT3G61330 and AT5G13205, split by Watson and Crick strand. wR-loop, Watson R-loop. cR-loop, Crick R-loop. wRNA, Watson RNA. cRNA, Crick RNA.

(D) Venn diagrams between DEGs (light blue) and DRGs (light coral) of LHS, LHSR12, or LHSR84, compared with CTRL-7d.

Supplemental Figure 10. Uncoupling between R-loop and RNA expression changes during development. (Support Figure 5)

(A) IGV snapshots of ssDRIP-seq (upper) and RNA-seq (lower) signals during development.

(B) Fuzzy cluster analysis of RNA-seq signal in DEGs among germination (GSE94712), root, leaf, and flowers (GSE38612). The union of DEGs from any two samples was analyzed. Line plots show standardized RNA-seq signal, with individual grey lines representing individual locus, and the orange line representing the values for the cluster center.

(C) Venn diagrams between the DRG and DEG clusters from Figure 3A and S9B. Light orange, RNA. Light steel blue, sG-R-loops. Light coral, asTSS-R-loops.

(D) Bar chart of enrichments (log₂ fold change) of 4 DEG clusters from Figure S9B on each chromatin states, compared with total genes and TEs.

Supplemental Figure 11. Network of R-loop regulators on asTSS-R-loop, and cluster analysis of RNA expression in LHS group. (Support Figure 6)

(A) Network of asTSS-R-loop regulation pathway by part of known or predicted R-loop regulator. Green or red hexagons, RNA nodes (regulator). Blue hexagons, R-loop nodes (be regulated). RECQL3, RECQ-like helicase 3; RH17, DEAD-box RNA helicase 17; RH20, DEAD-box RNA helicase 20; SEN1L, Senataxin-like helicase 1; RNH1A, RNase H1A; RNH1C, RNase H1B;
TOP1ALPHA, Topoisomerase 1α; TOP2, Topoisomerase II; THO5B, a THO/TREX complex component for RNA transport; SPT16, a component of facilitates chromatin transcription (FACT) complex; BRCA1, Breast cancer susceptibility protein 1.

(B) Fuzzy cluster analysis of RNA-seq signal in DEGs. The union of DEGs from any two samples of developmental stage set was analyzed. Line plots show standardized RNA-seq signal, with individual grey lines representing individual locus, and the orange line representing the values for the cluster center.

Supplemental Figure 12. Co-localization between antisense R-loops and TF binding sites. (Support Figure 6)

Snapshots of ssDRIP-seq signals of the EtOH sample and ABF1 and HB7 ChIP-seq signals of EtOH and ABA samples.

Supplemental Table 1. List of R-loop regulators analyzed in this study.

Supplemental Table 2. DRIP-qPCR primers used in this study.

Supplemental Table 3. Information of software used in this study.

Supplemental Data Set 1. Information of ssDRIP-seq data, condition descriptions, and R-loop peaks used in this study.
Reference

Abascal-Palacios, G., Ramsay, E.P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. Nature 553, 301-306.

Al-Hadid, Q., and Yang, Y. (2016). R-loop: an emerging regulator of chromatin dynamics. Acta Biochim Biophys Sin (Shanghai).

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

Baudrimont, A., Voegeli, S., Viloria, E.C., Stritt, F., Lenon, M., Wada, T., Jaquet, V., and Becskei, A. (2017). Multiplexed gene control reveals rapid mRNA turnover. Science advances 3, e1700006.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society 57, 289-300.

Brusslan, J.A., Bonora, G., Rus-Canterbury, A.M., Tariq, F., Jaroszewicz, A., and Pellegrini, M. (2015). A Genome-Wide Chronological Study of Gene Expression and Two Histone Modifications, H3K4me3 and H3K9ac, during Developmental Leaf Senescence. Plant physiology 168, 1246-1261.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet 10, e1004288.

Chen, L., Chen, J.-Y., Huang, Y.-J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., He, S., Zhou, Y., Abdel-Wahab, O., Zhang, D.-E., and Fu, X.-D. (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. Molecular Cell.

Chen, P.B., Chen, H.V., Acharya, D., Rando, O.J., and Fazzio, T.G. (2015). R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nat Struct Mol Biol 22, 999-1007.

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.

El Hage, A., Webb, S., Kerr, A., and Tollervey, D. (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet 10, e1004716.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nature Protocols 7, 1728-1740.

García-Muse, T., and Aguillera, A. (2019). R Loops: From Physiological to Pathological Roles. Cell.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chedin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell 45, 814-825.

Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I., and Chedin, F. (2013). GC skew at the 5’ and 3’ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res 23, 1590-1600.

Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A.,
Lanman, T., Chen, K., Schindler, A.B., Edwards, N., Ray-Chaudhury, A., Yao, J., Lehky, T., Piszczech, G., Crain, B., Fischbeck, K.H., and Cheung, V.G. (2018). Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. Molecular Cell.

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472, 115-119.

Kabeche, L., Nguyen, H.D., Buisson, R., and Zou, L. (2018). A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 359, 108-114.

Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. Genome Biol 18, 171.

Kumar, L., and M, E.F. (2007). Mfuzz: a software package for soft clustering of microarray data. Bioinformatics 2, 5-7.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huerta, C., and Chua, N.H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. The Plant cell 24, 4333-4345.

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

Nadel, J., Athanasiadou, R., Lemetre, C., Wijetunga, N.A., P, O.B., Sato, H., Zhang, Z., Jeddeloh, J., Montagna, C., Golden, A., Seoighe, C., and Greally, J.M. (2015). RNA-DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin 8, 46.

Ohle, C., Tesorero, R., Schermann, G., Dobrev, N., Sinning, I., and Fischer, T. (2016). Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. Cell.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research 44, W160-165.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature biotechnology 29, 24-26.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat Rev Genet 16, 583-597.

Sanz, L.A., and Chedin, F. (2019). High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. Nature protocols 14, 1734-1755.

Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X.Q., and Chedin, F. (2016). Prevalent, Dynamic, and Conserved R-Loop Structures
Associate with Specific Epigenomic Signatures in Mammals. Molecular cell 63, 167-178.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. The Plant cell 26, 2351-2366.

Shafiq, S., Chen, C., Yang, J., Cheng, L., Ma, F., Widemann, E., and Sun, Q. (2017). DNA Topoisomerase 1 Prevents R-loop Accumulation to Modulate Auxin-Regulated Root Development in Rice. Mol Plant 10, 821-833.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13, 2498-2504.

Shi, Y., and Berg, J.M. (1995). Specific DNA-RNA hybrid binding by zinc finger proteins. Science 268, 282-284.

Skourtli-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N.J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature 516, 436-439.

Song, L., Huang, S.C., Wise, A., Castanon, R., Nery, J.R., Chen, H., Watanabe, M., Thomas, J., Bar-Joseph, Z., and Ecker, J.R. (2016). A transcription factor hierarchy defines an environmental stress response network. Science 354.

Sun, Q., Csorba, T., Skourtli-Stathaki, K., Proudfoot, N.J., and Dean, C. (2013). R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. Science 340, 619-621.

Tarazona, S., García, F., Ferrer, A., Dopazo, J., and Conesa, A. (2012). NOIseq: a RNA-seq differential expression method robust for sequencing depth biases. EMBnet.journal 17, 18.

Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., and Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes & development 30, 1327-1338.

Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseeuw, E., Cao, Y., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G., and Datla, R. (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in Arabidopsis. Plant physiology 156, 346-356.

Xu, W., Xu, H., Li, K., Fan, Y., Liu, Y., Yang, X., and Sun, Q. (2017). The R-loop is a common chromatin feature of the Arabidopsis genome. Nat Plants.

Yang, X., Liu, Q.L., Xu, W., Zhang, Y.C., Yang, Y., Ju, L.F., Chen, J., Chen, Y.S., Li, K., Ren, J., Sun, Q., and Yang, Y.G. (2019). m(6)A promotes R-loop formation to facilitate transcription termination. Cell research 29, 1035-1038.

Yang, Z., Li, M., and Sun, Q. (2020). RHON1 Co-transcriptionally Resolves R-Loops for Arabidopsis Chloroplast Genome Maintenance. Cell Reports 30, 243-256.e245.

Yang, Z., Hou, Q., Cheng, L., Xu, W., Hong, Y., Li, S., and Sun, Q. (2017). RNase H1 Cooperates with DNA Gyrase to Restrict R-Loops and Maintain Genome Integrity in Arabidopsis Chloroplasts. Plant Cell 29, 2478-2497.

Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at
immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol 4, 442-451.

Yuan, W., Zhou, J., Tong, J., Zhuo, W., Wang, L., Li, Y., Sun, Q., and Qian, W. (2019). ALBA protein complex reads genic R-loops to maintain genome stability in Arabidopsis. Science advances 5, eaav9040.

Zhang, S.S., Yang, H., Ding, L., Song, Z.T., Ma, H., Chang, F., and Liu, J.X. (2017). Tissue-Specific Transcriptomics Reveals an Important Role of the Unfolded Protein Response in Maintaining Fertility upon Heat Stress in Arabidopsis. The Plant cell 29, 1007-1023.
Figure 1
Figure 1. The landscape of R-loops in Arabidopsis.

(A) Overview of the Arabidopsis samples used in this study. Four sets of samples, including plants exposed to different light, temperature, and stress conditions and at different developmental stages, are shown in the left panel. Specific samples from each set are shown in the right three panels.

(B) Brief workflow of ssDRIP-seq. First, the genomic DNA is extracted from Arabidopsis samples. After DNA fragmentation with restriction enzymes, only DNA:RNA hybrids are pulled down by S9.6 antibodies which specifically recognize and bind to DNA:RNA hybrids of various lengths. Finally, the library is prepared using an ssDNA-based strand-specific library kit and sequenced on an Illumina HiSeq X Ten System.

(C) IGV snapshots of the DRIP-seq signals from all samples (one of two replicates is shown). The wR-loops (Watson strand R-loops) are shown in red, while the cR-loops (Crick strand R-loops) are shown in blue. Samples are represented by different colors, and the detailed information is shown at the bottom.

(D) Heatmap of Spearman's rank correlation coefficients between two samples. Samples are distinguished by color, as shown in Figure 1C.

(E) Metaplots of sense (left) and antisense (right) R-loop signals from samples at different developmental stages (upper) and under different light (middle) and temperature (lower) conditions, centered on total genes with 1 kb up- and downstream extensions. Line, 95% mean; shadow, 95% confidence interval.
**Figure 2**

A. Diagram showing gene, TSS, sGB-R-loop, asTSS-R-loop, and developmental stages.

B. Scatter plots comparing sGB-R-loop and asTSS-R-loop under different temperature conditions (17°C, 27°C).

C. Table showing gene and TE counts across different developmental stages and temperature treatments.

D. Scatter plots comparing sGB-R-loop and asTSS-R-loop across different developmental stages and temperature treatments.

E. Scatter plot showing R² and Rho values for LHS/CTRL-7d and asTSS-R-loop.

F. Scatter plot showing the correlation between sDI and asDI with a Rho value of 0.1036.

G. Bar charts comparing the counts of sDI and asDI with a total gene and TE.
Figure 2. Dynamics of sGB-R-loops and asTSS-R-loops.

(A) Diagram of an sGB-R-loop (sense gene body R-loop) and an asTSS-R-loop (antisense transcription start site R-loop). The sGB-R-loop is defined as a sense R-loop on the entire gene body, from the TSS (transcription start site) to the TTS (transcription termination site). The asTSS-R-loop is defined as an antisense R-loop around TSS, ± 150 bp.

(B) Scatter plots of the levels of sGB-R-loops (left panel) and asTSS-R-loops (right panel) per total genes and TEs. Samples exposed to different temperature treatments were compared with the control (CTRL-7d). Normalized reads counts are shown as log10(n+1). Red dot: q value < 0.05, log2 fold change > 1. Blue dot: q value < 0.05, log2 fold change < -1. Gray dot: other.

(C) Scatter plots of the levels of sGB-R-loops (middle panel) and asTSS-R-loops (right panel) on total genes and TEs (transposable elements, annotated by TAIR10_transposable_elements). Samples from different developmental stages were compared with seedling leaf tissue. Plots of changes in sGB-R-loop levels between flowers and seedling leaves are magnified (left panel) to show the changes in sGB-R-loop levels on TEs, TE genes (transposable element genes, annotated by TAIR10_GFF3_genes), and other genes. Normalized reads counts are shown as log10(n+1). Red dot: q value < 0.05, log2 fold change > 1. Blue dot: q value < 0.05, log2 fold change < -1. Gray dot: other.

(D) Scatter plots of sGB-R-loop (left) and asTSS-R-loop (right) levels on total genes and TEs. DK (dark) was compared with CTRL-7d. Normalized reads counts are shown as log10(n+1). Red dot: q value < 0.05, log2 fold change > 1. Blue dot: q value < 0.05, log2 fold change < -1. Gray dot: other.

(E) Scatter plot of asTSS-R-loop/sGB-R-loop log2 fold change between LHS and CTRL-7d on total genes and TEs. R2 and Rho values (Spearman's rank correlation coefficients) are shown on the plot.

(F) Dynamics Index matrix of sGB-R-loops and asTSS-R-loops. sDI and asDI were calculated following the algorithm shown in Supplemental Figure 4. Genes assigned to the same sDI and asDI are located in the same dot, and the number of genes in the same dot is represented by color and size. The Rho value (Spearman's rank correlation coefficients) of all genes and TEs is shown.

(G) Left, box plot of sDI and asDI for all genes and TEs. Means are represented by crosses. Right, accumulation curves of sDI and asDI from all genes and TEs.
Figure 3

Graphical representation of R-loop changes and cluster analysis for sGB-R-loop and asTSS-R-loop.

Panel A: R-loop changes for different time points (germination, root, seedling, leaf, old leaf, flower) across clusters.

Panel B: Venn diagrams showing the overlapping clusters for sGB-R-loop and asTSS-R-loop.

Panel C: Bar charts illustrating the percentage distribution of TE, TE gene, and other gene categories across clusters for sGB-R-loop and asTSS-R-loop.

Panel D: Heatmaps representing the percentage distribution of states (state1 to state9) across clusters for sGB-R-loop and asTSS-R-loop.
**Figure 3. Cluster analysis of sGB-R-loops and asTSS-R-loops during development.**

**A** Fuzzy cluster analysis of ssDRIP-seq signals in DRGs. The union of DRGs from any two samples of developmental stage set was analyzed. Line plots show standardized ssDRIP-seq signals, with individual gray lines representing individual loci and orange lines representing the values for the cluster center.

**B** Venn diagrams of the DRG clusters from Figure 3A. Light steel blue, sG-R-loops. Light coral, asTSS-R-loops.

**C** Proportional bar charts of TE genes (orange), TEs (blue), and other genes (gray) in the four DRG clusters from Figure 3A. Left, sG-R-loops. Right, asTSS-R-loops.

**D** Proportional bar charts of the chromatin states of the four DRG clusters from Figure 3A. The chromatin states were classified into three types: active state (1, 3, 6, 7), repressed state (2, 4, 5), and silent state (8, 9). State 1, marked with H3K4me2, H3K4me3, H3 acetylation, H3K36me3, often associated with transcribed regions. State 2, relative to state 1, but additionally enriched with H3K27me3. State 3, enriched in H3K4me1, H2Bub, H3K36me3 and H3K4me2/3, representing a transcription elongation signature. State 4, similar with state 2, but accompanied with lower levels of active transcription marks. State 5, corresponding to Polycomb-regulated chromatin. State 6, located in intragenic regions and slightly enriched in H2A.Z and H3K4me1. State 7, annotated as an intragenic state, enriched in H3K4me1, H2Bub, and H3K36me3. State 8, AT-rich heterochromatic state. State 9, GC-rich heterochromatic state.
Figure 4

Figure 4
Figure 4. Genomic loci with abnormal R-loop dynamic patterns.

(A) R-loops on ONSEN genes activated by long heat stress. Upper, The R-loop levels of CTRL-7d (CTL, green), LHS (red), LHSR12 (light blue), and LHSR84 (dark blue) on chromosome 1. Middle, IGV snapshots of R-loop signals on two ONSEN genes, AT1G11265 (left) and AT1G58140 (right). Lower, DRIP-qPCR of two loci in AT1G11265. Graphs shown as mean ± SEM of three technical replicates (circle) and are representative of two independent experiments.

(B) R-loop dynamics on tRNA (left) or snoRNA (right) genes. All samples from the light set were normalized to CTRL-7d. Gray dot, q value ≥ 0.05 in any sample. Colored dot, DRG (q value < 0.05) in the corresponding samples.

(C) Metaplots of sense (left) and antisense (right) R-loop levels of CTRL-7d (light blue) and DK (dark blue) on tRNA (upper) or snoRNA (lower) genes. Data shown as 95% mean.

(D) Heat maps of sense and antisense R-loop signals of CTRL-7d (left) and DK (right) on total tRNA (upper) or snoRNA (lower) genes.

(E) Statistics of differential R-loop tRNA (coral) or snoRNA (dark blue) genes from flower, germinating seedling, young leaf, and DK tissue. Flower, germination, and young leaf tissue were compared with seedling leaf tissue while DK was compared with CTRL-7d.

(F) Percentage of differential R-loop tRNA (left) or snoRNA (right) genes in total differential R-loop genes. Background, percentage of tRNA or snoRNA genes in total genes. Flower (light coral), germination (light blue), and young leaf (green) tissue were compared with seedling leaf tissue while DK (dark blue) was compared with CTRL-7d.

(G) Percentage of DRGs in total genes, sorted by tRNA, snoRNA, and other genes. Coral, differential R-loop tRNA genes/total tRNA genes. Dark blue, differential R-loop snoRNA genes/total-snoRNA genes. Gray, other differential R-loop genes/total other genes. Flower, germination and young leaf samples were compared with seedling leaf tissue, while DK was compared with CTRL-7d.
Figure 5
Figure 5. Uncoupling between R-loop dynamics and RNA expression changes.

(A) Scatter plots of RNA/sGB-R-loop (left) and RNA/asTSS-R-loop (right) log2 fold change (LHS/CTRL-7d) on total genes and TEs. R² and Rho values (Spearman's rank correlation coefficients) are shown on the plots.

(B) Fold changes (log2) of RNA (green), sGB-R-loop (red), and asTSS-R-loop (blue) intensity (LHS/CTRL-7d).

(C) Heatmaps of RNA (upper), sGB-R-loop (middle), and asTSS-R-loop (lower) differences on total genes and TEs between each pair from the LHS set based on Euclidean distance.

(D) IGV snapshots of ssDRIP-seq and RNA-seq signals from the LHS set for protein-encoding genes HSP21 and AT3G01260, with the Watson and Crick strand values split.

(E) IGV snapshots of ssDRIP-seq and RNA-seq signals of the LHS set for AT1G11265 (ONSEN family gene) and AT2G29165 (other Ty1-Copia family gene), with the Watson and Crick strand values split.

(F) Scatter plot of log2 fold changes (LHS/CTRL-7d) in RNA (x-axis) and sGB-R-loop (y-axis) for ATCOPIA78 (red), ATHILA6A (blue), and other TEs (gray). Linear trend lines and corresponding R² values are shown.

(G) Scatter plots of RNA/sGB-R-loop/asTSS-R-loop log2 fold change values (upper, flower/seedling leaf. lower, root/seedling leaf) on total genes and TEs. R² values are shown on the plots.
Figure 6
Figure 6. Regulatory network and potential roles of R-loops in regulating gene expression.

(A) Network of the sGB-R-loop regulatory pathway by some known or predicted R-loop regulators. Green or red hexagons, RNA nodes (regulators). Blue hexagons, R-loop nodes (regulated). RECQL3, RECQ-like helicase 3; RH17, DEAD-box RNA helicase 17; RH20, DEAD-box RNA helicase 20; SEN1L, Senataxin-like helicase 1; RNH1A, RNase H1A; RNH1C, RNase H1B; TOP1ALPHA, Topoisomerase 1α; TOP2, Topoisomerase II; THO5B, a THO/TREX complex component for RNA transport; SPT16, a component of facilitates chromatin transcription (FACT) complex; BRCA1, Breast cancer susceptibility protein 1.

(B) Scatter plots of genes or TE s in the LHS group with changes in both RNA and R-loop levels. X-axis, RNA log2FC. Y-axis, sGB-R-loop (left panel) or asTSS-R-loop (right panel) log2FC. All genes or TEs in which R-loop and RNA levels changed simultaneously (q value < 0.05 and |log2FC| > 1) are shown. Red dot, gene. Blue, TE gene. White, TE fragment.

(C) Violin plots of asTSS-R-loop levels (log2(n+1)) of genes with different dynamics patterns (see Supplemental Figure 10B), 95% confidence interval. The median (white dot) and interquartile range (gray shadow) are shown.

(D) Permutation test of co-localization between sense or antisense R-loop peaks of the CTRL-EtOH-4h samples (shown in Supplemental Data Set 1) and ChIP-seq peaks of transcription factors of the EtOH-treated samples (GSE80564).

(E) Metaplots of ABA-induced ABF1 (left) and HB7 (right) ChIP-seq signal fold changes on sense (light blue) or antisense (light red) R-loop peaks, mean.

(F) Snapshots of ssDRIP-seq signals of the EtOH-treated sample and ABF1 and HB7 ChIP-seq signals of the EtOH- and ABA-treated samples.
Abascal-Palacios, G., Ramsay, E.P., Beuron, F., Morris, E., and Vannini, A (2018). Structural basis of RNA polymerase III transcription initiation. Nature 553, 301-306.

Al-Hadid, Q., and Yang, Y. (2016). R-loop: an emerging regulator of chromatin dynamics. Acta Biochim Biophys Sin (Shanghai).

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

Baudriment, A., Voegeli, S., Viloria, E.C., Stritt, F., Lenon, M., Wada, T., Jaquet, V., and Becskei, A. (2017). Multiplexed gene control reveals rapid mRNA turnover. Science advances 3, e1700006.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society 57, 289-300.

Brusslan, J.A., Bonora, G., Rus-Canterbury, A.M., Tariq, F., Jaroszewicz, A., and Pellegrini, M. (2015). A Genome-Wide Chronological Study of Gene Expression and Two Histone Modifications, H3K4me3 and H3K9ac, during Developmental Leaf Senescence. Plant physiology 168, 1246-1261.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y., Luo, Z., Hanra, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet 10, e1004288.

Chen, P.B., Chen, H.V., Acharya, D., Rando, O.J., and Fazzio, T.G. (2015). R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nat Struct Mol Biol 22, 999-1007.

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.

El Hage, A, Webb, S., Kerr, A, and Tollervey, D. (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet 10, e1004716.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nature Protocols 7, 1728-1740.

García-Muse, T., and Aguilera, A. (2019). R Loops: From Physiological to Pathological Roles. Cell.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chedin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell 45, 814-825.

Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I., and Chedin, F. (2013). GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res 23, 1590-1600.
Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A., Lanman, T., Chen, K., Schindler, A.B., Edwards, N., Ray-Chaudhury, A., Yao, J., Lehly, T., Piszczek, G., Crain, B., Fischbeck, K.H., and Cheung, V.G. (2018). Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. Molecular Cell.

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472, 115-119.

Kabeche, L., Nguyen, H.D., Buisson, R., and Zou, L. (2018). A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 359, 108-114.

Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. Genom Biol 18, 171.

Kumar, L., and M, E.F. (2007). Mfuzz: a software package for soft clustering of microarray data. Bioinformation 2, 5-7.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huertero, C., and Chua, N.H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. The Plant cell 24, 4333-4345.

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

Nadel, J., Athanasiadou, R., Lometre, C., Wijetunga, N.A., P, O.B., Sato, H., Zhang, Z., Jeddeloh, J., Montagna, C., Golden, A., Seoighe, C., and Greally, J.M. (2015). RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin 8, 46.

Ohle, C., Tesorero, R., Schermann, G., Dobrev, N., Sinning, I., and Fischer, T. (2016). Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. Cell.

Ramirez, F., Ryan, D.P., Gruning, B., Bharadwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research 44, W160-165.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Gutman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature biotechnology 29, 24-26.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat Rev Genet 16, 583-597.

Sanz, L.A., and Chedin, F. (2019). High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. Nature protocols 14, 1734-1755.
Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X.Q., and Chedin, F. (2016). Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Molecular cell 63, 167-178.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. The Plant cell 26, 2351-2366.

Shafiq, S., Chen, C., Yang, J., Cheng, L., Ma, F., Widemann, E., and Sun, Q. (2017). DNA Topoisomerase 1 Prevents R-loop Accumulation to Modulate Auxin-Regulated Root Development in Rice. Mol Plant 10, 821-833.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Min, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13, 2498-2504.

Shi, Y., and Berg, J.M. (1995). Specific DNA-RNA hybrid binding by zinc finger proteins. Science 268, 282-284.

Skourti-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N.J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature 516, 436-439.

Song, L., Huang, S.C., Wise, A., Castanon, R., Nery, J.R., Chen, H., Watanabe, M., Thomas, J., Bar-Joseph, Z., and Ecker, J.R. (2016). A transcription factor hierarchy defines an environmental stress response network. Science 354.

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., and Dean, C. (2013). R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. Science 340, 619-621.

Tarazona, S., Garcia, F., Ferrer, A., Dopazo, J., and Conesa, A. (2012). NOIseq: a RNA-seq differential expression method robust for sequencing depth biases. EMBnet.journal 17, 18.

Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., and Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes & development 30, 1327-1338.

Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseeuw, E., Cao, Y., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G., and Datla, R. (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in Arabidopsis. Plant physiology 156, 346-356.

Xu, W., Xu, H., Li, K., Fan, Y., Liu, Y., Yang, X., and Sun, Q. (2017). The R-loop is a common chromatin feature of the Arabidopsis genome. Nat Plants.

Yang, Z., Hou, Q., Cheng, L., Xu, W., Hong, Y., Li, S., and Sun, Q. (2017). RNase H1 Cooperates with DNA Gyrase to Restrict R-Loops and Maintain Genome Integrity in Arabidopsis Chloroplasts. Plant Cell 29, 2478-2497.
Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol 4, 442-451.

Yuan, W., Zhou, J., Tong, J., Zhuo, W., Wang, L., Li, Y., Sun, Q., and Qian, W. (2019). ALBA protein complex reads genic R-loops to maintain genome stability in Arabidopsis. Science advances 5, eaav9040.

Zhang, S.S., Yang, H., Ding, L., Song, Z.T., Ma, H., Chang, F., and Liu, J.X. (2017). Tissue-Specific Transcriptomics Reveals an Important Role of the Unfolded Protein Response in Maintaining Fertility upon Heat Stress in Arabidopsis. The Plant cell 29, 1007-1023.
### The R-loop Atlas of Arabidopsis Development and Responses to Environmental Stimuli
Wei Xu, Kuan Li, Shuai Li, Quancan Hou, Yushun Zhang, Kunpeng Liu and Qianwen Sun
*Plant Cell*; originally published online February 19, 2020; DOI 10.1105/tpc.19.00802

This information is current as of May 6, 2020

| Supplemental Data | /content/suppl/2020/02/19/tpc.19.00802.DC1.html |
|-------------------|-----------------------------------------------|
| Permissions       | https://www.copyright.com/ccc/openurl.do?sid=pd_huw1532298X&issn=1532298X&WT.mc_id=pd_huw1532298X |
| eTOCs             | Sign up for eTOCs at: http://www.plantcell.org/cgi/alerts/ctmain |
| CiteTrack Alerts  | Sign up for CiteTrack Alerts at: http://www.plantcell.org/cgi/alerts/ctmain |
| Subscription Information | Subscription Information for *The Plant Cell* and *Plant Physiology* is available at: http://www.aspb.org/publications/subscriptions.cfm |