



## Stress-induced endogenous siRNAs targeting regulatory intron sequences in *Brachypodium*

Hsiao-Lin V. Wang, Brandon L. Dinwiddie, Herman Lee, et al.

RNA published online December 5, 2014

---

**Supplemental Material** <http://rnajournal.cshlp.org/content/suppl/2014/12/01/rna.047662.114.DC1.html>

**P<P** Published online December 5, 2014 in advance of the print journal.

**Creative Commons License** This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

---

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To subscribe to *RNA* go to:  
<http://rnajournal.cshlp.org/subscriptions>

---

# Stress-induced endogenous siRNAs targeting regulatory intron sequences in *Brachypodium*

HSIAO-LIN V. WANG, BRANDON L. DINWIDDIE, HERMAN LEE, and JULIA A. CHEKANOVA

School of Biological Sciences, University of Missouri–Kansas City, Kansas City, Missouri 64110, USA

## ABSTRACT

Exposure to abiotic stresses triggers global changes in the expression of thousands of eukaryotic genes at the transcriptional and post-transcriptional levels. Small RNA (smRNA) pathways and splicing both function as crucial mechanisms regulating stress-responsive gene expression. However, examples of smRNAs regulating gene expression remain largely limited to effects on mRNA stability, translation, and epigenetic regulation. Also, our understanding of the networks controlling plant gene expression in response to environmental changes, and examples of these regulatory pathways intersecting, remains limited. Here, to investigate the role of smRNAs in stress responses we examined smRNA transcriptomes of *Brachypodium distachyon* plants subjected to various abiotic stresses. We found that exposure to different abiotic stresses specifically induced a group of novel, endogenous small interfering RNAs (stress-induced, UTR-derived siRNAs, or sutr-siRNAs) that originate from the 3' UTRs of a subset of coding genes. Our bioinformatics analyses predicted that sutr-siRNAs have potential regulatory functions and that over 90% of sutr-siRNAs target intronic regions of many mRNAs in *trans*. Importantly, a subgroup of these sutr-siRNAs target the important intron regulatory regions, such as branch point sequences, that could affect splicing. Our study indicates that in *Brachypodium*, sutr-siRNAs may affect splicing by masking or changing accessibility of specific *cis*-elements through base-pairing interactions to mediate gene expression in response to stresses. We hypothesize that this mode of regulation of gene expression may also serve as a general mechanism for regulation of gene expression in plants and potentially in other eukaryotes.

**Keywords:** smRNAs; abiotic stresses; splicing; splicing regulation; *Brachypodium*

## INTRODUCTION

Plants have evolved sophisticated mechanisms to adapt to environmental stresses, and abiotic stresses, such as heat, cold, and salinity, trigger changes in the expression of thousands of genes, with regulation at both the transcriptional and post-transcriptional levels (Borsani et al. 2005; Mahajan and Tuteja 2005; Hirayama and Shinozaki 2010; Mastrangelo et al. 2012; Staiger and Brown 2013). However, our understanding of the interactions among the regulatory networks that control plant gene expression in response to environmental changes remains limited.

One mechanism by which plants respond to environmental stress is by modifying gene expression through the activity of small RNAs (smRNAs) (Borsani et al. 2005; Sunkar et al. 2007; Ruiz-Ferrer and Voinnet 2009), including microRNAs (miRNAs) and small interfering RNAs (siRNAs), which differ in their biogenesis (Bartel 2004; Chapman and Carrington 2007; Chen 2009; Castel and Martienssen 2013). All smRNAs are incorporated into Argonaute (AGO) proteins; in the

AGO complex, miRNAs repress their targets at primarily post-transcriptional levels through cleavage or translational inhibition of their target RNAs (Bartel 2004; Chapman and Carrington 2007; Chen 2009) and siRNAs affect translation and cleavage of target RNAs similarly to miRNAs, but can also affect DNA methylation and histone modification of their targets (Matzke et al. 2009; Law and Jacobsen 2010; Wierzbicki et al. 2012).

Studies in *Arabidopsis* and other plants showed that miRNAs and siRNAs function in the responses to many different abiotic stresses (Katiyar-Agarwal et al. 2007; Sunkar et al. 2007; Ruiz-Ferrer and Voinnet 2009; Zhang et al. 2012; Wong et al. 2014). In *Brachypodium*, miRNAs also function in response to cold, dehydration, and pathogen infections, supporting the idea that smRNAs play a general role in protecting the plant from stresses (Zhang et al. 2009; Jeong et al. 2013). Recent work also described a large and diverse population of stress-responsive nat-siRNAs derived from

Corresponding author: [chekanovaj@gmail.com](mailto:chekanovaj@gmail.com)

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.047662.114>.

© 2014 Wang et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

*cis*-natural antisense transcripts (*cis*-NATs) of biotic and abiotic stress-challenged rice, suggesting that nat-siRNAs may contribute to plant responses to environmental stresses (Borsani et al. 2005; Katiyar-Agarwal et al. 2007; Lu et al. 2012; Zhang et al. 2012). In addition to regulating gene expression through RNA degradation and translational repression, the most prominent role of plant siRNAs is their function in siRNA-dependent DNA methylation (Matzke et al. 2009; Law and Jacobsen 2010; Wierzbicki et al. 2012). It was reported that abiotic stresses also induce changes in DNA methylation in various plant species (Steward et al. 2002; Boyko et al. 2007; Dowen et al. 2012), suggesting a link between responses to stresses and chromatin modifications in plants. More recently, 21 nt siRNAs induced in response to infection of *Arabidopsis* with the bacterial pathogen *Pseudomonas syringae* were implicated in triggering dynamic changes in DNA methylation (Dowen et al. 2012), suggesting that stress-triggered siRNAs could also possibly regulate gene expression by affecting chromatin compaction.

Eukaryotes also regulate gene expression by regulating splicing of pre-mRNAs. Regulation of splicing (and smRNA pathways) in response to environmental stresses is particularly important since stresses trigger rapid, global changes in transcriptomes leading to dramatic reprogramming of gene expression on many different levels and these post-transcriptional mechanisms can provide the rapid responses vital for survival of stresses (Palusa et al. 2007; Reddy 2007; Mastrangelo et al. 2012; Staiger and Brown 2013). Splicing, removal of introns and ligation of exons, is executed by the spliceosome, which consists of five small nuclear RNAs (snRNAs) and over 180 proteins, and pre-mRNA splicing is among the most highly regulated processes in eukaryotes (Hoskins et al. 2011; Braunschweig et al. 2013). Most eukaryotic genes also undergo alternative splicing (AS), in which the spliceosome selects different pairs of splice sites in a pre-mRNA transcript to produce different mRNAs (Black 2003). Indeed, recent high-throughput studies demonstrated that at least 95% of multiexon genes in human, over 60% of intron-containing genes in *Arabidopsis*, and over 40% of genes in rice produce alternatively spliced mRNAs (Wang and Brendel 2006; Pan et al. 2008; Filichkin et al. 2010; Marquez et al. 2012; Reddy et al. 2013). A study of the transcriptome of *Brachypodium* plants detected only a few alternatively spliced transcripts (Walters et al. 2013), but this number will most likely increase as more studies add data. Assembly of the spliceosome on pre-mRNAs occurs via step-wise recognition of the short sequences at the exon/intron boundaries by snRNAs through base-pairing interactions. While the three minimal core splicing motifs, the 5' splice site (5'ss), the 3' splice site (3'ss), and the branch point (BP) are present in every intron and are required for the splicing reaction, they are degenerate and lack sufficient information to determine the correct 5' and 3' pairs (Black 2003; Hoskins et al. 2011; Braunschweig et al. 2013). The additional information necessary for fidelity and efficiency of splicing process is also pro-

vided by numerous additional exonic/intronic *cis*-regulatory elements on the pre-mRNA molecule. These *cis*-elements, collectively termed the “splicing code,” are recognized by *trans*-acting regulatory proteins, which promote or repress splice-site selection depending on their binding location and the surrounding sequence context (Martinez-Contreras et al. 2007; Wang and Burge 2008; Long and Caceres 2009; Barash et al. 2010). RNA secondary structure also contributes to regulate splicing by affecting splice-site accessibility (Hiller et al. 2007; Shepard and Hertel 2008). In addition to authentic constitutive and alternative splice sites, eukaryotic genes also contain cryptic splice sites not normally used for splicing and a large excess of so-called “decoy” splice sites, the sequences that match the consensus splice-site motifs as well as authentic sites, yet are rarely or never used in splicing (Sun and Chasin 2000; Coté et al. 2001). Despite the large potential for errors due to (i) the degeneracy and variability of splice signals, (ii) the myriad possibilities for alternative splicing, (iii) the huge variety of positional and context-dependent regulatory *cis*-elements governing recognition of splice sites, and (iv) the presence of numerous cryptic and decoy splice sites, the splicing process appears to occur with very high fidelity. Therefore, the splicing machinery must be able to distinguish authentic splice sites and also respond to regulatory cues that affect splicing.

There is also an extensive crosstalk between splicing and other gene regulatory mechanisms (Luco et al. 2011; Braunschweig et al. 2013). Recent evidence suggests that the fidelity of splice-site selection is governed not only by the interaction of snRNPs and non-snRNP factors with pre-mRNA but also by the factors associated with chromatin and the transcriptional machinery (Braunschweig et al. 2013). For example, chromatin features, such as DNA and chromatin modifications that alter chromatin compaction, also regulate splicing by affecting RNA Polymerase II elongation rate, thus modulating the binding of effector proteins, which in turn affects splicing (Alló et al. 2009; Alló and Kornblihtt 2010; Dujardin et al. 2014). DNA modifications and chromatin compaction are regulated through smRNA pathways in many eukaryotes (Matzke et al. 2009; Wierzbicki et al. 2012; Castel and Martienssen 2013).

The crucial functions of splicing and smRNA pathways in regulation of gene expression in eukaryotes indicate that these pathways may connect; however, examples of these two pathways intersecting remain limited. A few cases of regulation of splicing by RNAi components have been reported in *Drosophila* and mammalian cells (Alló et al. 2009; Ameyar-Zazoua et al. 2012; Taliaferro et al. 2013). Ago proteins were also found to bind throughout the length of pre-mRNA transcripts in both smRNA-independent and smRNA-dependent manners, suggesting a connection between these two processes (Zisoulis et al. 2010; Leung et al. 2011; Taliaferro et al. 2013). In addition, the RNAi machinery and siRNAs were also shown to be involved in regulation of alternative splicing through epigenetic mechanisms (Alló et al. 2009;

Luco et al. 2011; Ameyar-Zazoua et al. 2012) and human RNAi componentsAGO1 andAGO2 were reported to affect alternative splicing by linking chromatin modifiers with the splicing machinery (Alló et al. 2009; Alló and Kornblihtt 2010; Ameyar-Zazoua et al. 2012).

Although work in animal systems has provided intriguing hints on the potential roles of smRNAs in splicing, our understanding of their role in plants remains unclear. Recent work reported that several *Arabidopsis* splicing mutants also exhibit defects in siRNA accumulation and DNA methylation (Zhang et al. 2013), while the *Arabidopsis* homolog of pre-mRNA splicing factor PRP3 affects DNA methylation without altering siRNAs level (Huang et al. 2013). The presence of miRNA binding sites within introns of *Arabidopsis* and rice genes also has been shown recently, suggesting that these miRNAs could participate in cleavage of pre-mRNAs (Meng et al. 2013). Reciprocally, expression of miRNAs was shown to be affected by alternative splicing of the transcripts that serve as precursors of these miRNAs in *Arabidopsis* (Yan et al. 2012; Jia and Rock 2013). In two cases, heat stress and abscisic acid triggered changes in miRNA expression through AS (Yan et al. 2012; Jia and Rock 2013). However, there is no evidence so far suggesting that small regulatory RNAs, such as miRNAs or siRNAs, can directly regulate splicing in plants.

Here, to investigate the role of smRNAs in stress responses, we examined the smRNA transcriptomes in *Brachypodium* plants challenged by various abiotic stresses. We found that a specific group of *Brachypodium* stress-responsive genes gives rise to a novel group of 24 nt smRNAs from their 3' UTRs and that these smRNAs have potential regulatory functions. Furthermore, our bioinformatics analyses indicate that over 90% of these smRNAs target regulatory regions within introns, including branch point sequences that may affect splice-site selection and fidelity of splicing.

## RESULTS

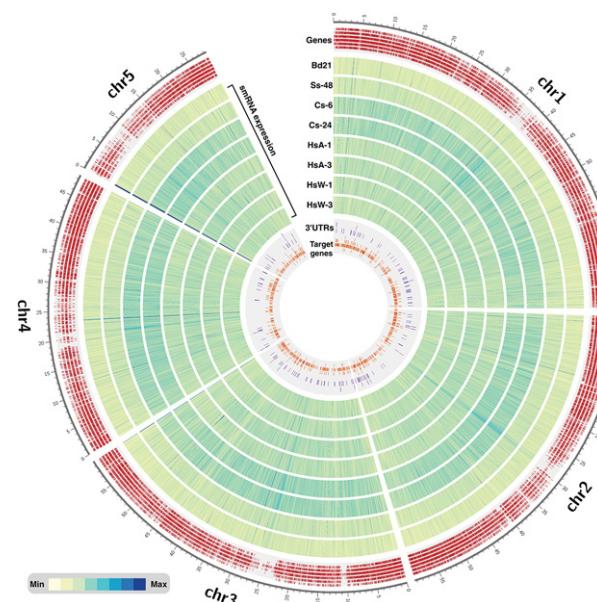
### Analysis of smRNA transcriptomes in *Brachypodium* plants under abiotic stress

To profile the populations of smRNAs in the model monocot *Brachypodium distachyon* and examine their regulation in response to abiotic stresses, we conducted high-throughput sequencing of smRNAs from plants exposed to four different abiotic stress conditions, cold, heat (air), heat (water immersion), and salt, in the wild-type *Brachypodium* cultivar Bd21. For our experiments we used information from the literature to select two time-points for stress durations, short and long, which differed for each stress: cold (6 and 24 h), heat-air (1 and 3 h), heat-water (1 and 3 h), and salt (48 h) (Vogel et al. 2005; Hong et al. 2008; Kumar et al. 2008). We generated small RNA libraries for Illumina sequencing from the leaves of *Brachypodium* plants subjected to stresses (Supplemental Table S1) and selected smRNAs between 15 and

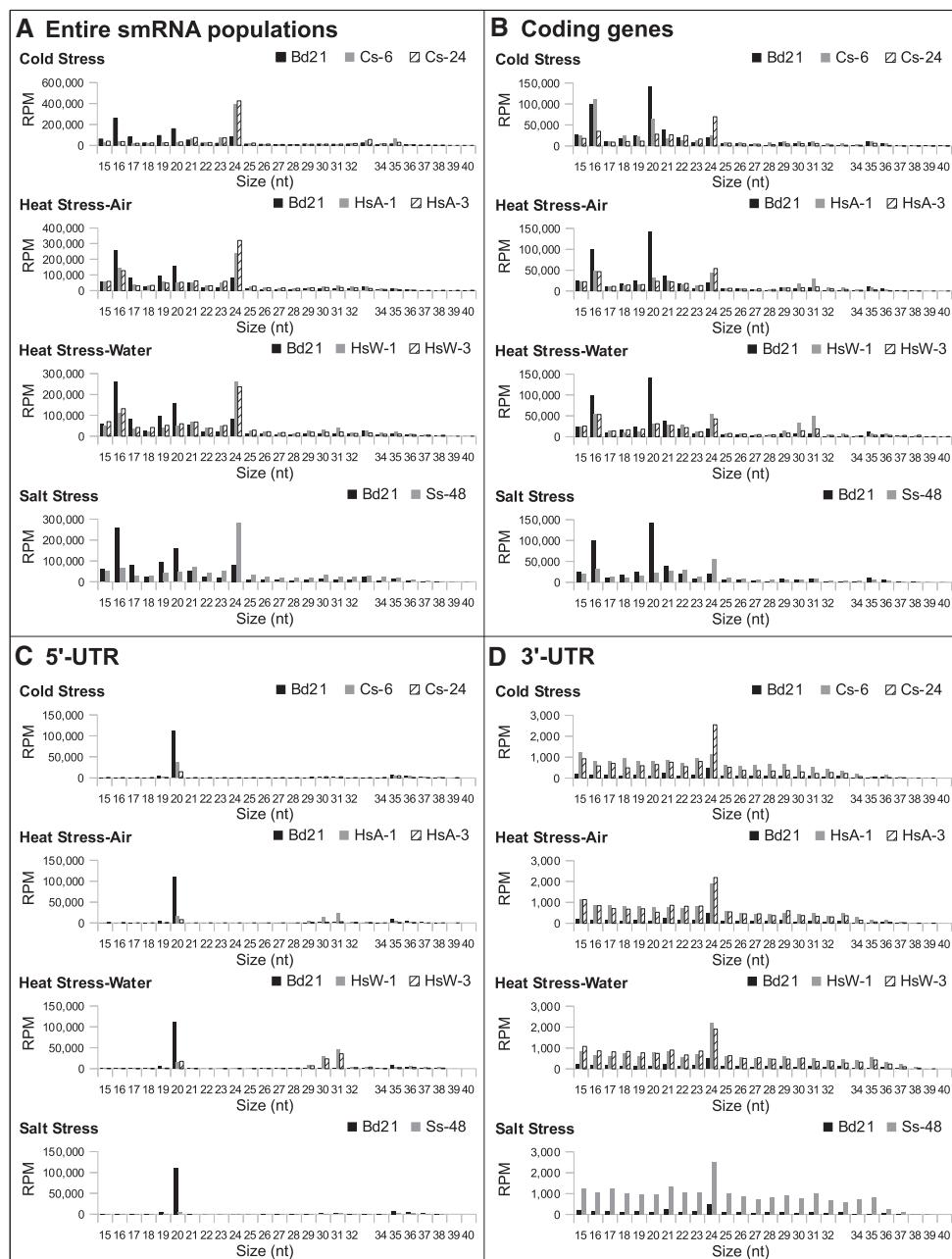
40 nt in length, which we mapped to the *Brachypodium* genome (Draper et al. 2001; International Brachypodium Initiative 2010; Brkljacic et al. 2011).

To characterize these smRNAs, we first analyzed and compared the genomic regions producing smRNAs in response to various stresses (Supplementary information; Supplemental Figs. S1, S2). We then plotted the smRNA expression data as a heatmap on all five chromosomes, compared the stressed smRNA expression data set with the smRNA expression from unstressed Bd21 plants (Fig. 1). This is the first graphical representation of stress-induced smRNA expression, on a whole-genome scale in *Brachypodium* plants. As expected, we observed that in both unstressed and stressed Bd21 plants smRNAs are highly expressed from the centromeric and pericentromeric regions, which are known to be silenced through smRNA-dependent DNA methylation in plants. Also, consistent with previously published data, we found that smRNAs are highly expressed specifically from the telomeric region on chromosome 5 (International Brachypodium Initiative 2010).

We then classified the smRNAs based on their size, level of expression, and genomic features. The majority of functional smRNAs in plants ranged from 20 to 25 nt. We found that the response to all stresses primarily affected the 20 and 24 nt smRNAs (Fig. 2A). We also observed a similar trend when we examined smRNAs that correspond to predicted



**FIGURE 1.** Heat map representation of smRNA expression in response to stresses. Bd21, unstressed *Brachypodium* plants; Ss-48, salt stress; Cs-6 or Cs-24, cold stress, 6 or 24 h; HsA-1 or HsA-3, heat stress by air incubation, 1 or 3 h; HsW-1 or HsW-3, heat stress by water immersion, 1 or 3 h. The outer track highlights the position of coding genes along the *Brachypodium* chromosomes. 3'-UTR track corresponds to the positions of genes (along the chromosomes) that express smRNAs from their 3' UTRs. The target genes track corresponds to the positions of genes (along the chromosomes) that are targeted by smRNAs.



**FIGURE 2.** Effect of abiotic stresses on *Brachypodium* smRNA populations. The distribution of smRNAs based on their length and expression level (RPM, reads per million). (A) The expression of *Brachypodium* smRNAs in response to various stresses; (B) the distribution of smRNAs mapped to *Brachypodium* coding genes; (C) the distribution of smRNAs mapped to 5' UTRs of *Brachypodium* coding genes; (D) the distribution of smRNAs mapped to 3' UTRs of *Brachypodium* coding genes. Bd21, unstressed *Brachypodium* plants; Cs, 6 or 24 h of cold stress; HsA, 1 or 3 h of heat stress by air incubation; HsW, 1 or 3 h of heat stress by water immersion; Ss-48, salt stress.

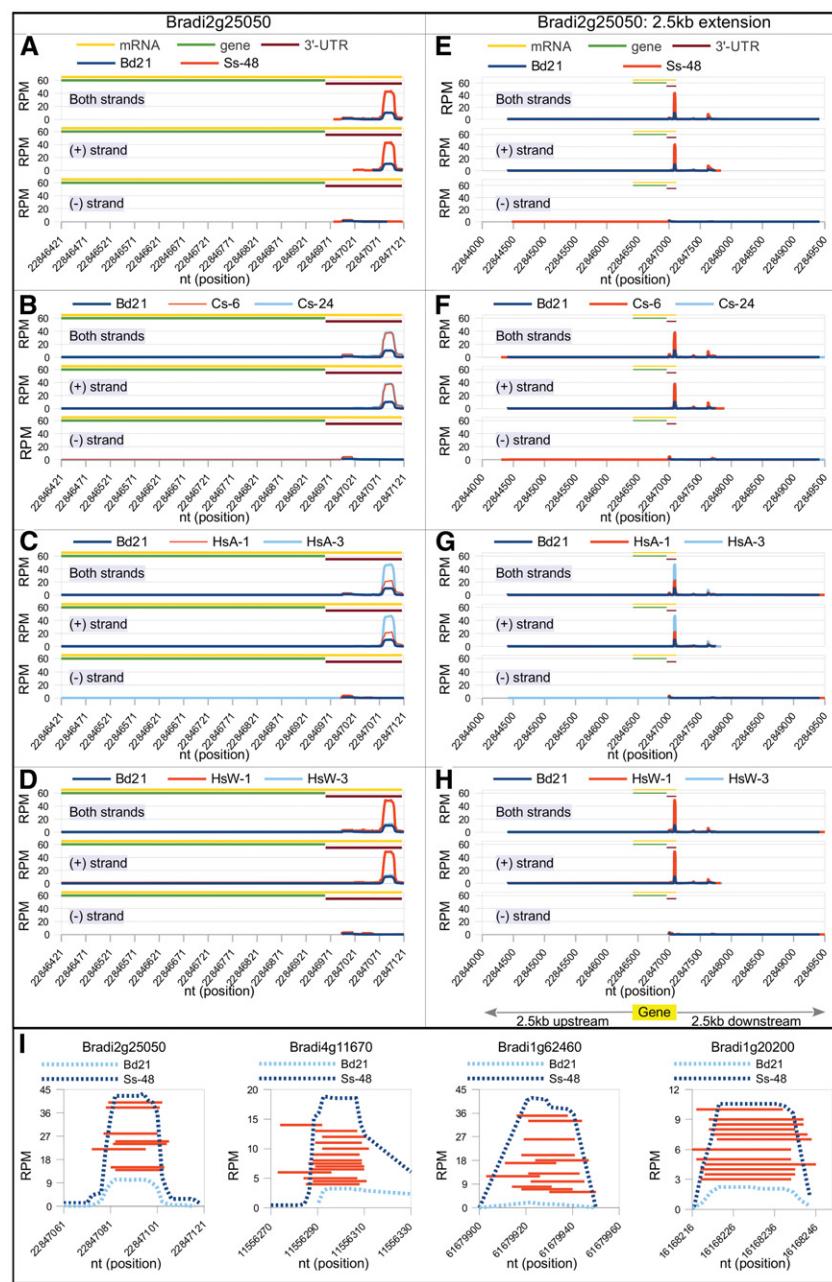
*Brachypodium* coding genes (Fig. 2B). To further investigate the origin of smRNAs corresponding to coding genes, we examined the smRNAs that map to 5' UTRs and 3' UTRs of *Brachypodium* coding genes. We found that, for all the stress conditions used in our study, the exposure to stress led to a significant decrease in 20 nt smRNAs that map to 5' UTRs (Fig. 2C), and a significant increase in smRNAs of all sizes that map to 3' UTRs of coding genes, with 24 nt smRNAs be-

ing the most abundant affected group (Fig. 2D). We next examined the populations of smRNAs that map to 5' UTRs and analyzed the 5' -UTR loci that exhibit significant decreases in production of 20 nt smRNAs. However, we found that only a few 5' UTRs showed this trend and the majority of affected 20 nt smRNAs correspond to *Bradi1g05790*, an unknown gene that shows sequence similarity only to putative chloroplast proteins in rice or maize and thus is likely specific to grasses.

### smRNAs produced from 3' UTRs of stress-responsive coding genes

We then decided to focus on smRNAs originating from the 3' UTRs of coding genes and the nature of the coding genes that give rise to them. To further examine the genes that show induction of smRNAs from their 3' UTRs in response to stresses, we used gene ontology (GO) classification, and chose for further, detailed analysis several specific loci that produce smRNAs in response to all the stresses used in this study (Supplemental Table S4). One of the examined loci was *Bradi2g25050*, a previously uncharacterized locus that encodes a member of the AP2-EREBP (ethylene-responsive element binding proteins) family of plant-specific transcription factors, which regulate development and abiotic stress responses in *Arabidopsis* (Song et al. 2005). To characterize the smRNAs produced from the *Bradi2g25050* 3' UTR, we mapped these smRNAs at 2-nt resolution to the coding region of *Bradi2g25050* (Fig. 3A–D), including an additional 2.5 kb of up- and downstream sequences (Fig. 3E–H). We found that all the smRNAs come from a short stretch of the *Bradi2g25050* 3' UTR. These smRNAs were produced only from the plus strand and were collinear with the *Bradi2g25050* transcript. Notably, all four different stress conditions used in our study triggered an identical pattern of smRNA production from the *Bradi2g25050* 3' UTR (Fig. 3A–H).

To examine the pattern of smRNA production from the 3' UTRs of other genes, we analyzed in detail the 3' UTRs of several additional genes that give rise to smRNAs in response to all applied stresses, *Bradi4g11670*, *Bradi1g62460*, and *Bradi1g20200* (Fig. 3I). All unique smRNAs originating from the 3' UTRs of these genes were mapped at 2-nt resolution to the 3' UTRs of their origin. We found that, similar to *Bradi2g25050*, all smRNAs originated only from a short stretch of each examined 3' UTR, 25 nt for the *Bradi2g25050* 3' UTR. Similarly, all smRNAs were produced only from the plus strand and were the same polarity as their predicted precursor



**FIGURE 3.** The pattern of smRNA production from the 3' UTR of *Bradi2g25050* in response to various stresses. (A–H) The region of *Bradi2g25050* 3' UTR that gives rise to smRNAs triggered by various stresses. (A–D) smRNAs produced from the *Bradi2g25050* locus were mapped to its coding region. The peak in 3' UTR corresponds to the position of smRNAs that are triggered by various stresses. (E–H) smRNAs produced from the *Bradi2g25050* region were mapped at a lower resolution to the genomic region comprising the coding region and additional 2.5 kb up- and downstream sequences. (A,E) Salt stress; (B,F) cold stress, 6 and 24 h; (C,G) heat stress air, 1 and 3 h; (D,H) heat stress by immersion in water, 1 and 3 h. These plots show the smRNAs are produced from *Bradi2g25050* in response to all applied stresses, and the pattern of smRNA production is identical for all stresses used in this study. (I) Characteristics of the smRNAs originating from the 3' UTRs of various genes in response to salt stress: *Bradi2g25050*, *Bradi4g11670*, *Bradi1g62460*, and *Bradi1g20200*. All unique smRNAs are shown mapped to 3' UTRs at 2-nt resolution. All unique smRNAs triggered by salt stress are displayed according to their location along their respective 3' UTRs. The position of smRNAs relative to each respective 3' UTR is indicated on the x-axis. The level of each smRNA expression (in RPM) in response to salt stress is indicated on the y-axis. The light blue dotted line corresponds to the position and expression of smRNAs found in unstressed Bd21 plants. The dark blue dotted line corresponds to the position and the level of expression of smRNAs triggered by high salinity.

transcripts (Fig. 3I). When we analyzed smRNAs originating from these four 3' UTRs, we found that vast majority of them are 24 nt in length.

### 3'-UTR-derived smRNAs with predicted targets

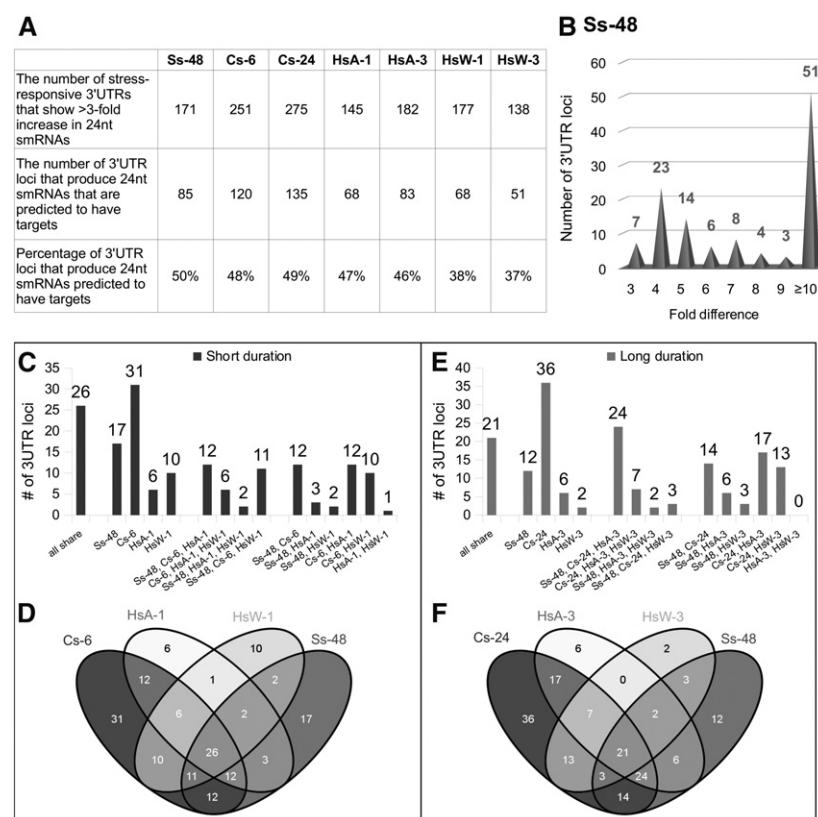
Using *Bradi2g25050* and the other genes described previously as examples, we set out to isolate all 3' UTRs genome-wide that serve as precursors to smRNAs and also exhibit the same pattern of smRNA production in response to stresses as these four genes. To that end, we first isolated all 3' UTRs that could give rise to smRNAs in a collinear fashion with their precursor transcripts. Since the majority of smRNAs found to be triggered by stresses from the 3' UTRs of the four example genes are 24 nt in length, we then calculated the abundance of 24 nt smRNAs from each individual 3' UTR. For the subsequent analysis, we retained only the 3' UTRs that produce 24 nt smRNAs both in response to stresses and in the unstressed Bd21 plants. Next, we compared the smRNA expression triggered by individual stresses from each 3' UTR with smRNA expression (from identical 3' UTRs) in unstressed Bd21 plants and retained only the 3' UTRs that exhibit at least a threefold increase in production of 24 nt smRNAs in response to each stress. We then reasoned that, if these stress-induced smRNAs are functional, they should have targets within genes elsewhere in the genome. Therefore, we searched for their potential targets on the basis of sequence complementarity. Using this logic, we isolated the group of 3' UTRs that produce smRNAs with perfect complementarity to transcripts of other genes.

Applying the algorithm written based on the set of rules outlined above to the whole genome, we identified a group of genes that, in response to stresses, can produce 24 nt smRNAs collinear with their precursor transcripts, and these 24 nt smRNAs have the potential to target products of other genes in *trans* (Fig. 4A; Supplemental Tables S5–S11). Our analysis indicates that some stress-induced, 3'-UTR-derived 24 nt smRNAs do not have complementary targets. As an example, salt stress triggers production of 24-nt smRNAs from 171 individual 3' UTRs of coding genes, but only 85 of these give rise to smRNAs with perfect complementarity to predicted targets elsewhere in the genome. We observed a very similar percentage

for all stresses used in this study. Notably, in all stress treatments, the majority of 3' UTRs that satisfied the stringent parameters used to isolate them exhibit sharp increases, >10-fold, in induction of 24 nt smRNAs in response to stresses in comparison to unstressed plants (Fig. 4B; Supplemental Fig. S3).

### Functional analysis of the genes producing smRNAs from their 3' UTRs

To obtain more information on the pathways that govern the plant responses to different stresses, we used the groups of 3' UTRs that satisfied all of the above criteria (from mRNAs that increase in response to stress and produce smRNAs from their 3' UTRs) in response to each individual stress and compared them between different stresses (Fig. 4C–F). This comparison was done separately for short (C,D) or long (E,F) stress time points. As a result we isolated a set of genes that have an increase in smRNA levels in response to all stresses,



**FIGURE 4.** Analysis of 3' UTRs of coding genes that give rise to smRNAs with predicted *trans*-targets. (A) Summary of the group of stress-responsive 3' UTRs that give rise to smRNAs with predicted *trans*-targets (see text for description of parameters). For example 85 salt-stress-responsive UTRs give rise to 24 nt smRNAs that are predicted to target the products of other coding genes in *trans*. (B) Fold difference in expression of smRNAs from 3' UTRs in response to salt stress. (C–F) Cross-comparison of loci that show significant increase in 24 nt smRNAs with predicted targets in response to various stresses. Venn diagram demonstrates the number of smRNA generating 3' UTRs that are either stress-specific or common among different stress treatments. The comparison was done separately for short (C,D) or long (E,F) stress treatment.

the sets of genes that overlap between different individual stresses, and the sets of genes that are specific to one individual type of stress (Fig. 4C–F; Supplemental Table S12). The overlap between affected 3' UTRs (Fig. 4D,F) suggests that similar smRNA pathways could be activated in response to different abiotic stresses (Fig. 4D,F). The overall trend was very similar for short or long stress treatments.

To obtain more information on the networks that govern the plant responses to stresses we also carried out the functional analysis of the genes producing smRNAs from their 3' UTRs in response to different stresses. To this end, we subjected the genes producing smRNAs from their 3' UTRs to GO analysis. This analysis was performed independently for each individual stress and stress duration. The resultant GO terms were then examined to find which GO terms were enriched, and the results were visualized using REVIGO and Cytoscape applications (Fig. 5A; Supplemental Fig. S4; Supplemental Table S4; Smoot et al. 2011; Supek et al. 2011). In the case of salt stress (Ss-48), a total of 171 coding genes were subjected to GO term analysis and 116 GO terms were used for the GO term enrichment (Fig. 5A). The GO analysis of the genes producing smRNAs from their 3' UTRs indicated that similar biological processes are enriched among all stress conditions and in short and long stress durations (Fig. 5A; Supplemental Fig. S4; Supplemental Table S4). The enriched biological processes include various binding activities and, in particular, protein, DNA, RNA, chromatin, and ATP binding. In addition to binding activities, the comparison of GO term enrichment between different kinds of stresses indicates that RNA polymerase II activity and various membrane channel activity were enriched among all stress conditions (Fig. 5A; Supplemental Fig. S4; Supplemental Table S4). Interestingly, we found that H3K9-specific histone methyltransferase activity was specifically enriched only following heat stress by water immersion (Supplemental Fig. S4E,F). These results suggested that genes producing smRNAs from their 3' UTRs are associated with functionally important protein coding genes and these genes may be part of an important mechanism regulating plant stress responses.

### 3'-UTR-derived smRNAs with predicted *trans*-targets are uniformly 24 nt long

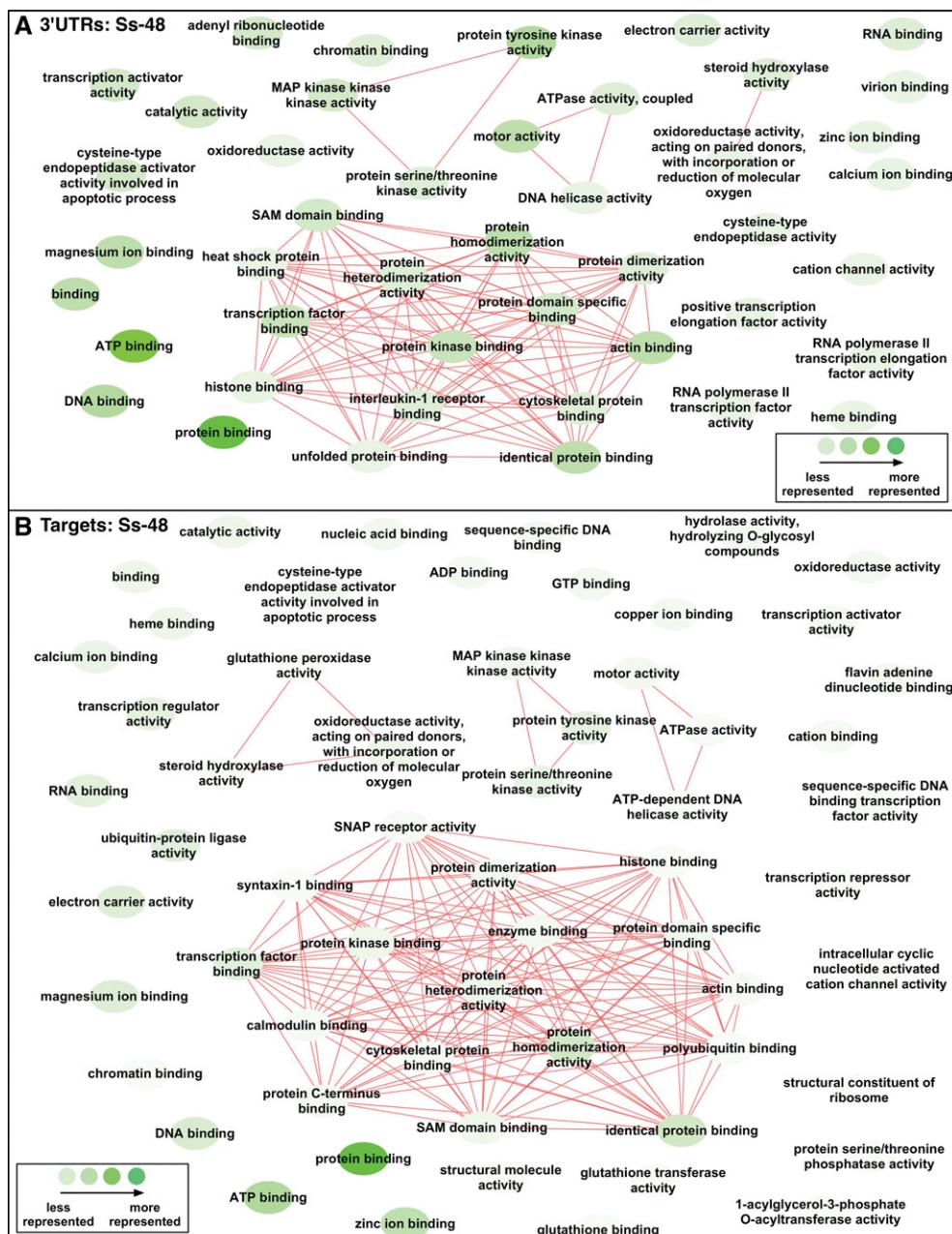
Our initial analysis of all stress-induced smRNAs that map to 3' UTRs indicated that the set of 24 nt smRNAs is the predominant group of smRNAs triggered by all stresses; however, smRNAs of other sizes also increased in abundance (Fig. 2D). In addition, while analyzing the smRNAs produced from 3' UTRs of genes, including *Bradi2g25050*, *Bradi4g11670*, *Bradi1g62460*, and *Bradi1g20200* (Fig. 3I), we also observed a few smRNAs of other sizes.

Although we used threefold ratio in increase of 24 nt smRNA production in response to stresses as one of the criteria to isolate this subset of 3' UTRs, the calculations we

applied did not discriminate against 3' UTRs that generate smRNAs of sizes other than 24 nt. Thus, our analysis can also identify smRNAs of various sizes that are produced from this subset of 3' UTRs. To gain more insight into the nature of smRNAs triggered by stresses and how they may be produced from their precursor transcripts, we analyzed in detail the entire population of smRNAs of all sizes originating from the group of 3' UTRs isolated using the stringent criteria described above (Fig. 4A; Supplemental Tables S5–S11). Therefore, we used the entire population of smRNAs triggered from these 3' UTRs by each stress and profiled these smRNAs by size and expression level. To our surprise, we found that the smRNAs originating from this group of 3' UTRs are almost uniformly 24 nt long and dramatically increase in their abundance under all stress conditions (Fig. 6A–G). The only other group of smRNAs that had a very minor increase in abundance was a group of 21 nt smRNAs. We observed the same pattern in the response to all stresses used in our study. We also inspected all smRNAs triggered from each individual 3' UTR and found that smRNAs that differ in size from 24 nt were expressed only from a few loci. For example, in the case of salt stress, only three loci (*Bradi2g04380*, *Bradi3g45010*, and *Bradi3g53300*) from the total set of 85 isolated 3' UTRs (<4%), give rise to smRNAs of different sizes, while all other loci produce only 24 nt smRNAs in response to salt stress. This observation suggests that a small number of 3' UTRs that produce smRNAs of all sizes could lead to a visual distortion of the size distribution of the total population of smRNAs mapped to 3' UTRs, and thus could also explain the increase in abundance of 3'-UTR-derived smRNAs of different sizes observed in Figure 2D. The observation that smRNAs originating from this group of 3' UTRs are exclusively 24 nt in length also argues that these smRNAs are not the products of nonspecific mRNA degradation, and supports our hypothesis that these smRNAs have regulatory potential. Profiling these smRNAs by their first nucleotide revealed that the majority of them start with a 5'A, with 5'U being the second largest group (Fig. 7), suggesting that if these smRNAs participate in RNAi pathways they could preferentially be loaded into *Brachypodium* homologs of the Ago2 and Ago4 complexes to silence their targets (Mi et al. 2008).

To shed more light on the nature of smRNAs produced from the group of identified 3' UTRs and whether they could be miRNAs, we used mfold tools to examine the possible RNA secondary structures that could be formed by RNA corresponding to the 3' UTRs of these genes (Zuker 2003). We used the 3' UTRs of genes responsive to salt stress in this analysis and found that the RNA secondary structures predicted using bioinformatics RNA folding tools suggested that these loci are unlikely to be miRNA precursors (see Materials and Methods).

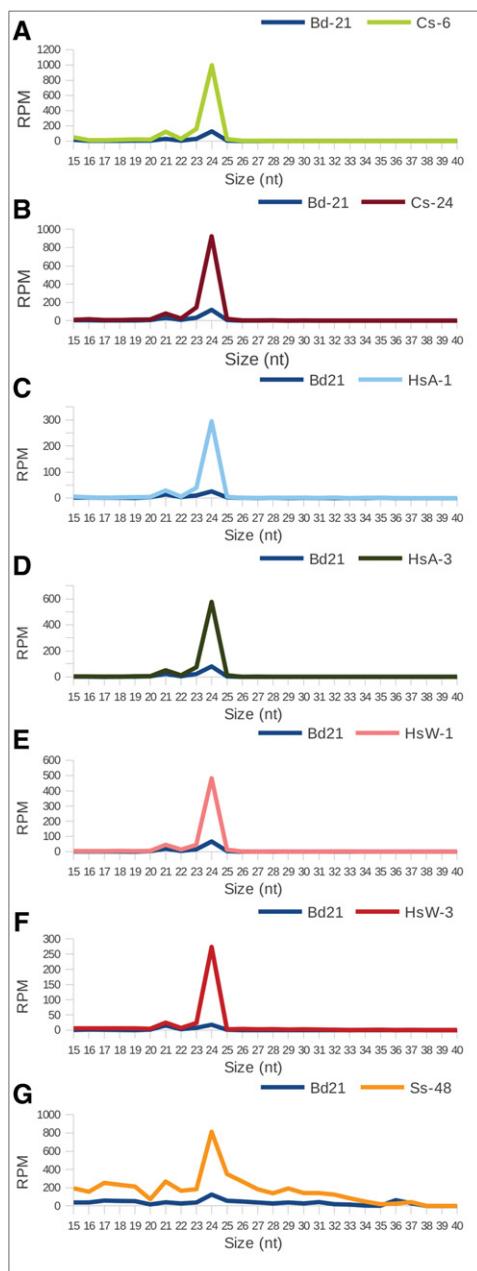
Thus, we isolated a set of stress-responsive genes that produce a novel group of smRNAs from short stretches of their 3' UTRs. These smRNAs are generated in collinear manner to



**FIGURE 5.** Enriched biological processes among genes producing smRNAs from their 3' UTRs and genes targeted in *trans* by sutr-siRNAs in response to salt stress (Ss-48). (A) The genes harboring smRNAs producing 3' UTRs were subjected to GO-term enrichment and the analysis was performed separately for each stress condition and duration of stress. The figure represents the enriched GO terms for genes that produce smRNAs in response to salt stress (Ss-48). A total of 171 3'-UTR loci and 116 GO-terms were used for the analysis. Similar functional categories were clustered together and connected within the same network using REVIGO and Cytoscape. Nodes (green circles) reflect the GO terms, and edges (red lines) connect similar functional categories. The node color indicates the frequency of the GO terms and corresponds to the color scale indicated in the figure. (B) The genes targeted by sutr-siRNA in *trans* were subjected to GO-term enrichment and the analysis was performed separately/independently for each stress condition and duration. The figure represents the GO-terms enriched in targeted coding genes in salt stress 48 h (Ss-48) condition and a total 363 coding genes and 480 GO-terms were used for the analysis. Similar functional categories clustered together and connected within the same network using REVIGO and Cytoscape. Nodes (green circles) reflect the GO-terms, and edges (red lines) connect the similar functional categories. The node color indicates the frequency of the GO terms and corresponds to the color scale indicated in the figure.

their precursor transcripts, are almost exclusively 24 nt in length and have perfect complementarity to transcripts of other *Brachypodium* coding genes. Since these 24 nt smRNAs triggered by stresses from these 3'-UTR loci are predicted to

target transcripts of other *Brachypodium* coding genes and thus have a potential to be functional, we termed them sutr-siRNAs (for stress-induced, UTR-derived siRNAs) and focused our analysis on their predicted targets.



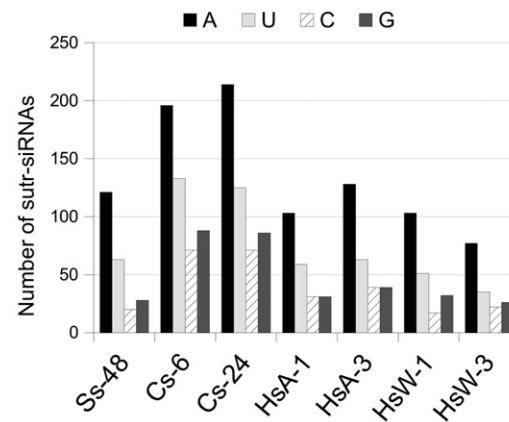
**FIGURE 6.** Characterization of smRNAs originating from 3' UTRs in response to stresses. (A–G) Size distribution and expression of smRNAs originating from the group of 3' UTRs isolated using set of parameters described in text. This group of 3' UTRs produces smRNAs that are predicted to target other coding genes in *trans*. The numbers of 3' UTRs used to profile smRNAs originating from them are listed in Figure 4A. For example, 85 salt-stress-responsive 3' UTRs, 120 cold-stress-responsive UTRs, etc., were used to profile smRNAs originating from them.

### Predicted targets of sutr-siRNAs

To isolate the genes that could be targeted by sutr-siRNAs, we used the sequence of each unique sutr-siRNA to search genome-wide for all *Brachypodium* coding genes that harbor sequences with reverse complementarity, allowing no mis-

matches. This identified a group of coding genes predicted to be targeted by various sutr-siRNAs in response to stresses (Table 1; Supplemental Tables S5–S11). As an example, we found that salt stress triggers the induction of 232 unique sutr-siRNAs that have perfect complementarity to the transcripts of 363 *Brachypodium* coding genes (Table 1; Supplemental Table S5).

To obtain more functional information on the networks that govern plant responses to different abiotic stresses, we carried out functional analysis of the genes targeted by sutr-siRNAs using GO classification. As for the 3' UTRs, GO classification was conducted separately for each individual stress and duration of stress and the resultant GO-terms for the sutr-siRNA target genes were then subjected to enrichment analysis and visualized using REVIGO and Cytoscape (Fig. 5B; Supplemental Fig. S5; Supplemental Table S13; Smoot et al. 2011; Supek et al. 2011). For example, in salt stress, (Ss-48), 363 *trans*-targeted coding genes were subjected to GO-term analysis and 480 GO-terms were used for the GO-term enrichment analysis (Ss-48) (Fig. 5B). Similar to enrichment of GO terms of the genes that produce sutr-siRNAs, GO classification of the genes targeted by sutr-siRNAs suggested that sutr-siRNAs target a broad spectrum of genes involved in different biological processes, although the majority of the GO-terms of genes targeted by sutr-siRNAs indicate that these target genes also involved in binding activities (Fig. 5B; Supplemental Fig. S5; Supplemental Table S13). Early and late time points of stress duration were very similar within each stress. Importantly, we found that many enriched GO-terms were involved in transcriptional activities and were enriched in all stresses. These results suggest that sutr-siRNAs potentially target functionally important coding genes, which further indicates the importance of sutr-siRNAs in the plant response to abiotic stresses.



**FIGURE 7.** The relative frequency of each 5' terminal nucleotide among sutr-siRNAs. Frequency distribution of the 5' terminal nucleotide for sutr-siRNAs originating in different stress conditions, Ss-48, salt stress; Cs, 6 or 24 h of cold stress; HsA, 1 or 3 h of heat stress by air incubation; HsW, 1 or 3 h of heat stress by water immersion. The numbers on the y-axis correspond to the numbers of smRNAs.

**TABLE 1.** Summary of sutr-siRNAs and their predicted target genes

		Ss-48	Cs-6	Cs-24	HsA-1	HsA-3	HsW-1	HsW-3
A. sutr-siRNAs <sup>a</sup>	The number of unique 24nt smRNA sequences with perfect complementarity to other genes	232	488	496	224	269	203	160
	The number of unique 24 nt smRNAs targeting sequences within introns	204	439	445	196	241	179	138
	Percentage of unique 24 nt smRNAs targeting sequences within introns	88%	90%	90%	88%	90%	88%	86%
B. Target genes <sup>b</sup>	The number of genes predicted to be targeted by sutr-siRNAs in <i>trans</i>	363	685	653	390	401	336	268
	The number of genes predicted to be targeted by sutr-siRNAs within introns	333	633	605	359	358	311	244
	Percentage of predicted target genes that are targeted within introns	92%	92%	93%	92%	89%	93%	91%

<sup>a</sup>Summary of sutr-siRNAs triggered by various stresses. About 90% of all sutr-siRNAs have complementarity to sequences within introns of coding genes.

<sup>b</sup>Summary of genes predicted to be targeted by sutr-siRNAs. Over 90% of target sites are located within intronic regions of these genes.

## Potential targeting mechanisms of sutr-siRNAs

Mechanistically, known smRNAs regulate gene expression by sequence-specific transcript degradation, translational inhibition or transcriptional repression (Bartel 2004; Chen 2009; Castel and Martienssen 2013). As a first step toward elucidating the mechanism of sutr-siRNA action, we examined their predicted target sites. We found that sutr-siRNA target sites fall into two main categories. Some sutr-siRNAs (fewer than 10% of total) have their predicted target sites located almost exclusively within the 3' UTRs of their target genes (Supplemental Table S14). This observation suggests that these sutr-siRNAs could participate in RNAi pathways and regulate the expression of their targets through translational repression or mRNA degradation, consistent with a function in *trans*-acting gene regulation. In contrast, and much to our surprise, we found that over 90% of all genes predicted to be targeted by sutr-siRNAs have the target site located within one of their introns. This trend was almost identical in all types of stresses used in our study (Table 1B). Reciprocally, ~90% of all sutr-siRNA sequences have complementarity to intronic sequences (Table 1A).

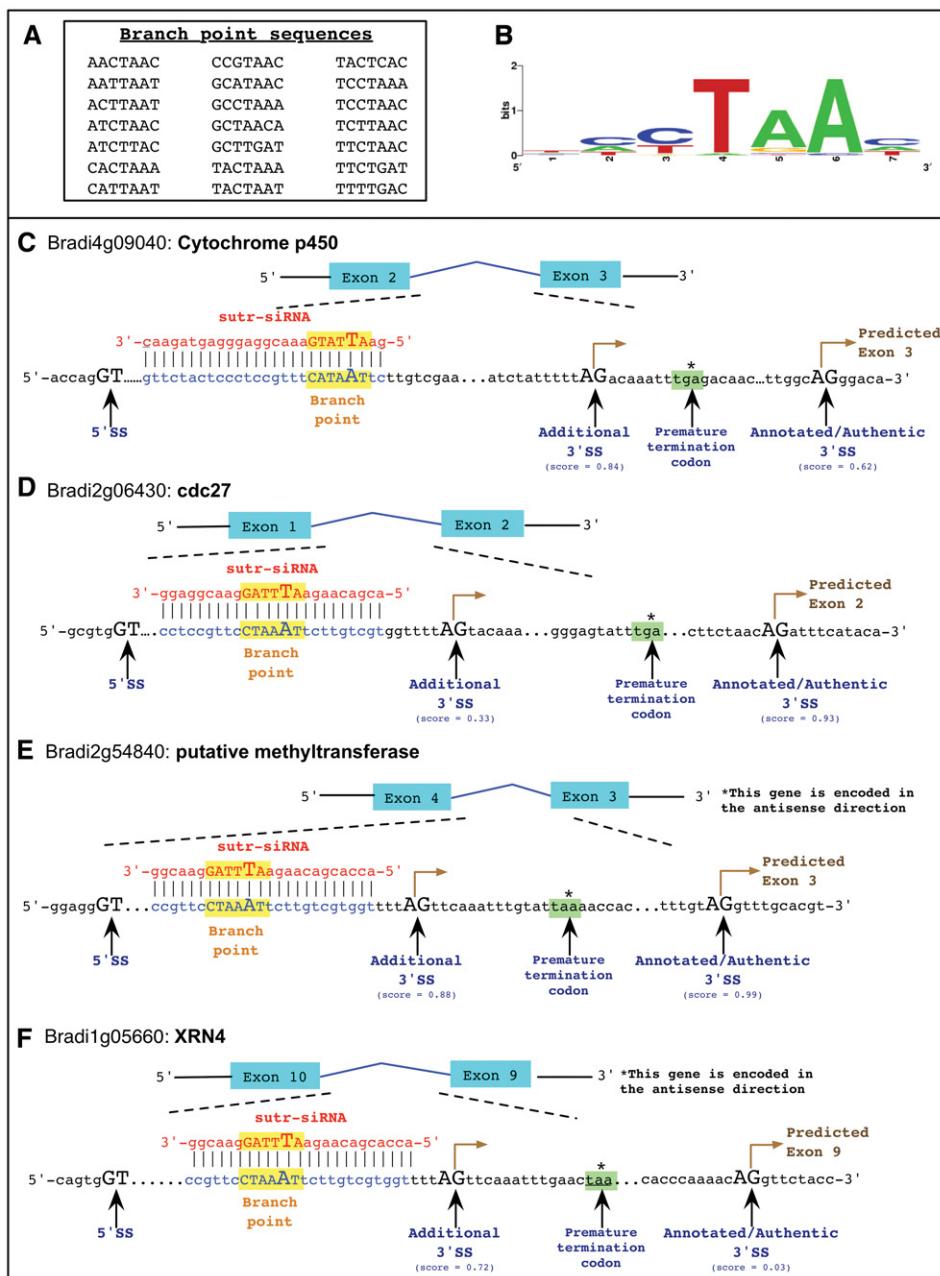
Since these sutr-siRNAs are predicted to target introns, and thus pre-mRNAs, we reasoned that they function in the nucleus and thus could be involved in the degradation of pre-mRNAs, in epigenetic regulation or in pre-mRNA processing. Therefore, we set out to investigate the features of the sutr-siRNAs that are predicted to target introns and chose the group of sutr-siRNAs triggered by salt stress for detailed analysis. The number of intronic target sites is higher than the number of sutr-siRNAs predicted to target them; therefore, we first examined whether different sutr-siRNAs carry common motifs or have other common features. We did not identify a specific common consensus motif shared by all of the sutr-siRNAs, but we found that sutr-siRNAs were uniformly purine rich, indicating that they target py-

rimidine-rich regions. Pyrimidine and polypyrimidine-rich tracts are located between important regulatory regions within introns, such as the branch point and 3' splice site (Reddy 2007). The branch point sequence in plants is located ~17–40 nt upstream of the 3' splice site, and the branch point and downstream 3' splice site are essential signals for efficient recognition of the 3' splice site (Reddy 2007; Reddy et al. 2012).

To investigate whether sutr-siRNAs can target any conserved intronic *cis*-elements that are required for splicing, we used the population of sutr-siRNAs triggered by salt stress for detailed analysis. The branch point sequence, with its conserved A residue, which is required for the catalysis in the first *trans*-esterification reaction in the first step of splicing, is among the most identifiable regulatory *cis*-elements within introns. In vertebrates, the branch point region is variable but has a more-identifiable consensus sequence, CURAY (Wang and Burge 2008). However, in plants the same branch point consensus sequence is not obvious and the sequences defining splice sites are degenerate (Reddy 2007; Reddy et al. 2012). Therefore, we used a database of *Arabidopsis* and rice branch point sequence motifs, which were predicted on the basis of training sets compiled using experimental data (Szcześniak et al. 2013), to interrogate the population of sutr-siRNAs that have target sites within introns.

Interestingly, we found that 30% of intron-targeting sutr-siRNAs carry sequences that are complementary to the predicted plant branch point sequences (Fig. 8A). We also identified the branch point motif that is favored by this group of sutr-siRNAs (Fig. 8B). The position of the branch point site relative to the polypyrimidine tract and the 3' splice site downstream from it are essential for it to function as a branch point in splicing reaction.

In order to find out whether the predicted intron branch point sequences targeted by these sutr-siRNAs could indeed be considered as functional branch points, we examined gene structures of a subset of target genes using the gene models



**FIGURE 8.** sutr-siRNAs that are predicted to target branch point sequences and gene structure of their targets. (A) Predicted intron branch point sequences that are complementary to sutr-siRNAs. (B) The features of BP sequences predicted to be targeted by the sutr-siRNAs. A position weight matrix for the branch point consensus in sutr-siRNAs. The height of each letter represents its frequency of occurrence. (C,D) Gene structure of sutr-siRNA target genes. (C) The structure of *Bradi4g09040*, encoding Cytochrome P450. The target site of the sutr-siRNA is located in intron 2 of the cytochrome P450 transcript. (D) The structure of *Bradi2g06430*, encoding DNA polymerase subunit Cdc27. The target site of sutr-siRNA is located in intron 1 of the Cdc27 transcript. (E) The structure of *Bradi2g54840*, encoding a putative methyltransferase. The target site of sutr-siRNA is located in intron 3 of the *Bradi2g54840* transcript. (F) The structure of *Bradi1g05660*, encoding 5'-3' exoribonuclease *XRN4*. The target site of the sutr-siRNA is located in intron 9 of the *XRN4* transcript. The annotated/authentic 3' splice site (3'SS) is the splice site used to produce full-length protein. Additional 3'SS is the splice site that could be used if the alternative branch point complementary to sutr-siRNA is chosen in splicing. In these examples the choice of alternative splice site would introduce a premature stop codon downstream from the splice site, resulting in either a short isoform or producing RNA substrate for nonsense-mediated degradation. (\*) Exon/intron numbering of *XRN4* (*Bradi1g05660*) and putative methyltransferase (*Bradi2g54840*) genes follows the genomic sequence annotation of the *Brachypodium* Sequence Consortium (International Brachypodium Initiative 2010), where the exon numbering is based not on the gene/transcript orientation, but on the DNA direction from 5' to 3'. Therefore, for genes encoded on the sense strand, the first exon (closest to the 5' end of the transcript) is "exon 1." For genes encoded on the antisense strand, however, the last exon (closest to the 3' end of the transcript) is "exon 1."

available from *Brachypodium* databases. In plants, the branch point (BP) sequences are located upstream of the polypyrimidine tracts and usually ~17–40 nt upstream of 3' splice site (Reddy 2007; Reddy et al. 2012). While some of the predicted BP sequences targeted by sutr-siRNAs did not have a clearly identifiable 3' splice-site consensus downstream from the BP sequences targeted by sutr-siRNAs, other BP sequences were found positioned next to the polypyrimidine tract and the 3' splice site downstream.

We then chose several sutr-siRNA target genes, *Bradi4g09040*, *Bradi2g06430*, *Bradi2g54840*, and *Bradi1g05660* for detailed analysis (Fig. 8C–F). In these genes, branch point sequences targeted by sutr-siRNAs are located within 16–31 nt upstream of the potential 3' splice sites (Fig. 8C–F). Based on the protein domain sequence alignment and GO analysis *Bradi4g09040*, *Bradi2g06430*, *Bradi2g54840*, and *Bradi1g05660* encode Cytochrome P450 (the ortholog of *Arabidopsis* AT2G26170), DNA polymerase subunit Cdc27 (the ortholog of *Arabidopsis* AT1G78650), a putative methyltransferase (the ortholog of *Arabidopsis* AT2G39750), and 3'-5' exoribonuclease XRN4 (the ortholog of *Arabidopsis* AT1G54490), respectively.

We then examined the structures of these genes, using the gene models available from *Brachypodium* databases. When we examined the intron regions in the vicinity of the sutr-siRNAs target sites in these target genes in detail and compared them with the annotated gene structures, we found that additional splice sites could be predicted in these regions. These additional predicted splice sites have high similarity to the known consensus splice sites in human and *Drosophila* and in all cases sutr-siRNAs targeted the BP of the splice site located upstream of the annotated splice site (Fig. 8C–F).

Bioinformatics analysis suggests that the selection of the splice sites predicted to be targeted by sutr-siRNAs in the examined genes would result in the introduction of a premature stop codon downstream from this 3' splice site; this stop codon can activate nonsense-mediated mRNA decay (NMD) and thus down-regulate the expression of the affected gene (Brogna and Wen 2009).

The *Brachypodium* consortium has annotated the structure of these genes based on EST and RNA-seq data (Draper et al. 2001; International Brachypodium Initiative 2010; Brkljacic et al. 2011). To determine whether alternatively spliced transcripts of these four genes were identified experimentally, we examined the annotated gene structures from the *Brachypodium* consortium and the expression data and found that no alternatively spliced transcripts of these genes were identified experimentally in *Brachypodium* plants grown under normal conditions, nor in plants challenged by various abiotic stresses (International Brachypodium Initiative 2010; Priest et al. 2014). A recent genome-wide study of the alternative-splicing landscape experimentally identified many novel splice junctions *Arabidopsis* (Marquez et al. 2012). Thus, we used the information from this study and TAIR 9 to cross-examine the splicing pattern of the closest orthologs

of these *Brachypodium* genes in *Arabidopsis*. Similarly, we did not find isoforms that were spliced at the 3' splice site targeted by the sutr-siRNAs. The position of the sutr-siRNAs targeted 3' splice sites upstream of the annotated major splice site in these genes also suggests that these splice sites could fall into the category of “decoy” splice sites, sequences that have similar degrees of similarity to consensus sequences as authentic splice sites, but that are rarely or never spliced (Sun and Chasin 2000; Coté et al. 2001). Thus, the selection of the 3' splice sites targeted by sutr-siRNAs in the examined genes will most likely result in an aberrant PTC-containing transcript that could be turned over by NMD. The base-pairing interactions between sutr-siRNAs and BP sequences targeted by sutr-siRNAs may potentially prevent the 3' splice sites downstream from them from being recognized as the authentic splice sites.

Perhaps the most intriguing gene predicted to be targeted by sutr-siRNAs is *Bradi1g05660*, the *Brachypodium* ortholog of *Arabidopsis* XRN4, the functional homolog of the yeast XRN1 (Kastenmayer and Green 2000; Olmedo et al. 2006). XRN4 is a 5'-3' exoribonuclease that degrades uncapped mRNAs and in *Arabidopsis* is known to act as silencing suppressor and a regulator of developmental and biotic stress response pathways (Belostotsky 2004; Gazzani et al. 2004; Olmedo et al. 2006; Gregory et al. 2008). XRN4 also plays an important role in the heat-sensing pathway in *Arabidopsis* (Merret et al. 2013). The heat shock response triggers the quick and global reprogramming of gene expression on many different levels and 25% of the *Arabidopsis* transcriptome undergoes rapid XRN4-dependent degradation in response to heat shock. Thus, the function of XRN4 is required for the thermotolerance of plants to long exposure to high temperatures, with *xrn4* mutant plants losing their ability to adapt to and survive heat stress.

We hypothesize that under stress conditions, the broader role of sutr-siRNAs could be to increase the fidelity of the splicing reaction, by masking or changing accessibility to specific cryptic, decoy, or alternative splice sites through base-pairing interactions, or by affecting the overall secondary structure of pre-mRNAs. Therefore, sutr-siRNAs could ensure that plants under stress conditions use the authentic splice site and produce functional protein to allow the plant to defend against abiotic stress conditions.

## DISCUSSION

Plants use a sophisticated array of mechanisms to respond to the changes in their environments upon exposure to different abiotic stresses. These mechanisms involve coordinated induction of various signals that trigger alterations in gene expression networks, which, together, allow plants to survive in a variety of environmental conditions (Mahajan and Tuteja 2005; Staiger and Brown 2013; Mastrangeli et al. 2012). Various smRNAs regulate many of these processes (Chen 2009). Here, we used *Brachypodium* plants as a model system

to investigate how the smRNA transcriptome responds to various abiotic stresses.

We isolated the specific group of *Brachypodium* stress-responsive genes that respond to stresses by giving rise, from their 3' UTRs, to a novel group of smRNAs with regulatory potential (Fig. 4A). Some of these genes generate smRNAs in response to all stresses or in response to more than one stress (Fig. 4C–F). The identical response of some genes to different stresses suggests that similar smRNA pathways could be activated in response to these abiotic stresses. Other genes responded only to a specific stress, suggestive of stress-specific smRNA pathways (Fig. 4C–F). Importantly, over half of these isolated 3' UTRs exhibit >10-fold increase in levels of smRNAs in response to stress (Fig. 4B; Supplemental Fig. S3A–F).

Our analyses also demonstrated that production of these 24 nt smRNAs from 3' UTRs of stress-responsive genes follows a specific, striking pattern. These smRNAs originate only from a very short stretch of each 3' UTR and are collinear with their precursor transcripts (Fig. 3A–H). Also, all of these smRNAs have predicted targets within transcripts of other coding genes (Table 1; Fig. 4A). Although smRNAs of the same polarity as their precursor transcripts might arise as degradation products of their precursor RNAs, our data suggest that sutr-siRNAs do not fall into this category. First, for nonspecific degradation products, we would expect smRNAs originating from the entire mRNA, but the sutr-siRNAs are produced only from a specific, very short stretch of 25–30 nt of each 3' UTR (Fig. 3A–I). Second, for nonspecific degradation products, we would expect to observe smRNAs of various sizes. However, with the exception of a few 21 nt smRNAs, we observed only 24 nt sutr-siRNAs from these selected 3' UTRs (Fig. 6A–G). Therefore, this specific group of stress-responsive 3' UTRs identified here serves as precursors to only one type of smRNAs, which we termed sutr-siRNAs.

### Possible mechanisms of sutr-siRNA biogenesis

Plants have large and diverse endogenous smRNA populations in two main categories, miRNAs and siRNAs, based on the structures of their precursors and their biogenesis (Bartel 2004; Xie et al. 2004; Chapman and Carrington 2007). miRNAs originate from miRNA precursors that form imperfect RNA stem-loops. The very diverse populations of plant siRNAs originate from perfect double-stranded RNA (dsRNA) precursors that can be formed by various mechanisms (Xie et al. 2004). To shed more light on a possible mechanism of sutr-siRNAs biogenesis, we used the set of all salt-stress-responsive genes to analyze possible RNA secondary structures that could be formed by their transcripts. However, none of the secondary structures predicted using bioinformatics RNA folding tools met the parameters required for being considered as miRNA precursors (Meyers et al. 2008) (see Materials and Methods). Although we are aware of the limitations of mfold, our data suggest that these

loci are unlikely to be miRNA precursors. The sutr-siRNAs are 24 nt long and in *Arabidopsis* DCL3 generates the majority of 24 nt siRNAs (Xie et al. 2004), suggesting that if this extends to *Brachypodium*, the biogenesis of sutr-siRNAs might require the *Brachypodium* homolog of DCL3, encoded by *Bradi3g29290*. Inverted repeats could be one of the sources of dsRNA serving as precursor for DCL proteins to generate smRNAs. However, when we analyzed the regions in the vicinity of sutr-siRNA producing sites, we found no DNA sequences nearby that could be considered reverse complementary to the sutr-siRNAs, indicating that sutr-siRNAs are unlikely to be produced from the dsRNA precursors formed by inverted repeats. In addition, based on the *Brachypodium* genome annotation, no *cis*-NAT pairs could be predicted in locations giving rise to sutr-siRNAs. Therefore, dsRNA precursors for sutr-siRNAs must originate through different mechanisms. Most sutr-siRNAs start with a 5'A (Fig. 7), suggesting that if these smRNAs participate in RNAi pathways they could preferentially be loaded into *Brachypodium* homologs of Ago2 and Ago4 complexes, encoded by *Bradi5g21800* and *Bradi2g10370*, to silence their targets (Mi et al. 2008). Further experimental work and the use of different experimental approaches will be needed to determine the mechanism of sutr-siRNA biogenesis.

### sutr-siRNA targets

We also identified the group of genes predicted to be targeted by sutr-siRNAs. GO classification of the predicted target genes suggested that sutr-siRNAs target a broad spectrum of genes involved in different processes (Fig. 5B; Supplemental Fig. S5; Supplemental Table S13). To our surprise, over 90% of the target genes are predicted to have the target site located in one of their introns (Table 1). The bioinformatics analysis of intron-targeting sutr-siRNAs indicated that they are complementary to intronic polypyrimidine tracts and a third of them carry sequences that are complementary to plant branch point sequences, which are among the most important intronic *cis*-elements determining the choice of splice site. When we analyzed the structures of the predicted target genes, we found that these sutr-siRNAs were targeting not the major annotated splice sites, but rather the additional potential splice sites located upstream of the annotated major splice sites in the target introns (Fig. 8C–F). Our bioinformatics analysis also indicates that the choice of the splice sites targeted by the sutr-siRNAs could potentially lead to short, alternatively spliced transcripts and result in the introduction of a stop codon downstream from these sutr-siRNAs targeted 3' splice sites, which could make these transcripts substrates for nonsense-mediated decay (Brogna and Wen 2009). The fact that most of the sutr-siRNAs are complementary to intronic regulatory regions suggests that they may play a role in the regulation of splicing under stress conditions.

Both pre-mRNA splicing and smRNA pathways in plants are powerful mechanisms involved in regulation of gene

expression in all conditions, including abiotic stresses (Sun-kar et al. 2007; Ali and Reddy 2008; Staiger and Brown 2013). However, a direct link between pre-mRNA splicing and small RNA pathways has so far remained elusive. The possible connection between splicing and RNAi pathways has been explored in other organisms and several examples of a regulation of alternative splicing through smRNAs in other eukaryotes have been reported recently (Alló et al. 2009; Alló and Kornblihtt 2010; Ameyar-Zazoua et al. 2012). Among the most direct connections between RNAi machinery and splicing was the finding that human RNAi components AGO1 and AGO2 can link chromatin modifiers with the splicing machinery (Alló et al. 2009; Alló and Kornblihtt 2010; Ameyar-Zazoua et al. 2012). In this case, Ago proteins were shown to facilitate spliceosome recruitment by modulating the elongation rate of RNA polymerase II, thereby affecting alternative splicing. Recent work also demonstrated that *Drosophila* Ago-2 protein binds G-rich regions within introns, suggesting that it also may be involved in regulation of splicing of specific transcripts, providing another possible link between components of RNAi machinery and splicing (Taliaferro et al. 2013). These examples suggest that siRNAs and the RNAi pathway are involved in splicing and alternative splicing in these organisms, thus connecting splicing with epigenetic modifications.

### Potential mechanism of action of sutr-siRNAs that target introns

Similar to other smRNAs, sutr-siRNAs could function in sequence-specific transcript degradation or participate in epigenetic regulation of gene expression by modulating DNA methylation in response to stresses (Chen 2009; Dowen et al. 2012). The approaches used in our study differ from the approaches used to link *Drosophila* and mammalian RNAi pathways in alternative-splicing decisions by connecting splicing with epigenetic modifications. Thus, we cannot rule out the possibility that sutr-siRNAs are involved in connecting splicing with epigenetic modifications or in degradation of pre-mRNAs. The use of different experimental approaches will be needed to distinguish among all possibilities. However, the fact that many sutr-siRNAs appear to target splice sites that are potentially cryptic or decoy splice sites argues that these sutr-siRNAs may function in either masking these splice sites or preventing other splicing *cis* determinants from being recognized by the splicing machinery.

The main splicing signals for pre-mRNA processing are degenerate and numerous additional intronic and exonic *cis*-regulatory elements function as recruiters for *trans*-acting splicing factors that recognize these *cis*-signals on pre-mRNAs (Goren et al. 2006; Wang and Burge 2008; Barash et al. 2010; Reddy et al. 2012). The recognition of the correct splice sites is critical and the commitment to splice at a particular splice site is believed to occur during the step-wise cotranscriptional assembly of the spliceosomal machinery

on the pre-mRNA. The formation of the commitment and presplicing complexes appears to be reversible and accumulating evidence also indicates that these steps could be subject to regulation more often than other steps (Chen and Manley 2009; Hoskins et al. 2011). It is intriguing to speculate that stress-triggered sutr-siRNAs with complementarity to regulatory *cis*-elements within introns may function in a manner that differs from the canonical roles of smRNAs in transcript degradation or epigenetic regulation of gene expression. The splice sites targeted by sutr-siRNAs could potentially produce either alternatively spliced RNAs or aberrant transcripts that become possible substrates for nonsense-mediated decay (Fig. 8C–F) and thus down-regulate the expression of the genes. Therefore, the base-pairing of sutr-siRNAs to alternative branch point sequences or other *cis*-elements could prevent them from being recognized by the splicing machinery through inhibiting the first steps of splicing at the target splice site and ensuring that the machinery chooses the correct splice site. Indeed, one of the sutr-siRNA targets encodes 5'-3' exoribonuclease XRN4, one of the key regulators in plant stress responses. Therefore, in this case, sutr-siRNAs could ensure that plants under stress conditions use the correct splice site and produce functional XRN4 protein to allow the plant to survive abiotic stress conditions.

Redirecting of splicing events through the use of single-stranded oligomers has been explored widely in therapeutic and experimental settings (Dominski and Kole 1993; Hua et al. 2010; Kole et al. 2012; Liu et al. 2012). Synthesized splice-switching oligonucleotides can direct pre-mRNA splicing by binding to intronic elements and blocking access to the transcript by the spliceosome and other splicing factors (Dominski and Kole 1993; Liu et al. 2012). In another example, the endogenous snoRNA HBII-52 and the smRNAs processed from it were reported to function in regulation of alternative splice-site selection through base-pairing-dependent RNA-RNA interactions in human and mouse tissue cultures (Kishore and Stamm 2006; Kishore et al. 2010), providing evidence that endogenous smRNAs also can modulate splicing. Thus, one could also envision the existence of other endogenous smRNAs that bind to either cryptic splice sequences or other *cis*-regulatory elements to ensure the preferential use of one splice site relative to other splice sites, thereby ensuring the fidelity of splicing reaction, which may become particularly important under stress conditions. In this case, smRNAs would be expected to have an advantage over proteins due to their complementarity to specific RNA targets.

Based on our observations, we favor the hypothesis that under stress conditions, sutr-siRNAs may affect the selection of splice sites by either masking *cis* determinants of specific cryptic or alternative splice sites, or by affecting pre-mRNA overall secondary structure, which can also affect the accessibility of *cis*-elements and splice sites (Hiller et al. 2007). To our knowledge, our finding represents the first genome-wide identification of smRNAs that have a potential to

regulate splicing of pre-mRNAs in response to abiotic stresses. Production of sutr-siRNAs in our study was triggered specifically by various abiotic stresses. Why would sutr-siRNAs be needed during stress conditions? Stresses trigger global reprogramming of gene expression on different levels and many protein factors are also sequestered by chaperone proteins to protect them from misfolding or degradation during stresses (Mahajan and Tuteja 2005). Therefore, it is possible that during stress it becomes more important to rely on base-pairing interactions through smRNAs. It is worth noting that one of the parameters we applied for isolating the specific group of 3' UTRs producing sutr-siRNAs was the condition that they express 24 nt smRNAs both in response to stresses and in unstressed plants. Indeed, 24 nt smRNAs were expressed in unstressed plants (Fig. 6A–G), suggesting that sutr-siRNAs may function in plants grown under normal conditions as well as in stress responses.

It remains to be determined if a similar group of smRNAs exists in different plant species or other organisms, and further experimental work will be needed to determine the significance of this observation. Also, to fully understand the mechanism of sutr-siRNA action, the splice isoforms of the sutr-siRNA target genes produced in wild-type plants will need to be compared with the splice isoforms produced in the organisms that do not express sutr-siRNAs during stress treatments. To address these questions requires an experimental system permitting elimination of sutr-siRNA during stress treatment. This can be accomplished through either use of mutants that eliminate or suppress the expression of a specific sutr-siRNA, or by using a different experimental system, such as mammalian or *Drosophila* cells, where experiments can be conducted *in vitro* in splicing extracts (by eliminating sutr-siRNAs with the antisense oligonucleotides that titrate the sutr-siRNAs). Our study, conducted in *Brachypodium*, provides a framework for researchers studying mechanisms of splicing regulation in other eukaryotes.

## MATERIALS AND METHODS

### Plant material and abiotic stress conditions

A community standard diploid inbred line of *Brachypodium distachyon*, Bd21, was used for all experimental treatments. The palea and lemma were carefully peeled off from the mature seeds with fine forceps. The stripped seeds were sterilized by soaking in a solution of 0.615% sodium hypochlorite supplemented with 0.1% Tween 20 for 10 min with occasional rocking. The sterilized seeds were thoroughly rinsed three times with sterile double-distilled water and placed on Murashige and Skoog (1/2 MS)-agar plates (1/2 MS-agar plates hereafter) (4.3 g/L MS salts with vitamins, 3% sucrose, pH 5.8, and 0.4% phytagel). The seeds/plates were cold-treated at 4°C for 2 d to synchronize germination, followed by germination in a light incubator at 22°C with a daily photoperiodic cycle of 16 h light and 8 h dark. Germinated seedlings were transplanted into soil and grown in a light incubator at 22°C with a daily

photoperiodic cycle of 16 h light and 8 h dark. For heat stress treatments, plants were either incubated at 42°C for 1 or 3 h in the temperature and light controlled growth chamber, or immersed in 42°C water for 1 or 3 h. For cold-stress treatment, plants were incubated at 4°C for 6 or 24 h. For the salt-stress treatment, the soil of plants was soaked with 300 mM NaCl until saturation and plants were grown for 48 h in a light incubator. The leaves of unstressed and stress treated plants were collected and frozen in liquid nitrogen before RNA extraction.

### Library construction and smRNA sequencing

Total RNA was isolated from the leaves of unstressed and stress treated plants using TRIzol (Life Technologies) according to the manufacturer's protocol. The RNA samples were used for sequencing library construction using the Small RNA sample Prep v1.5 kit and TruSeq Small RNA Sample Prep kit (Illumina) according to the manufacturer's instructions, as described previously (Shin et al. 2013). The smRNA libraries were sequenced using the Illumina Genome Analyzer II (by DNA Core Facility, University of Missouri) according to the manufacturer's instructions.

### smRNA data analysis

Data processing was done using available tools and custom in-house UNIX shell programming as described previously (Mi et al. 2008; Shin et al. 2013). The raw sequences in Illumina GAIx sequencing reads were trimmed removing adapter using "fastx\_clipper" in the FASTX-Toolkit (version 0.0.13) (Blankenberg et al. 2010) and smRNAs with lengths between 15 and 40 nt were selected and mapped to the *Brachypodium* genomic sequences (Bd v1.0 version) (International Brachypodium Initiative 2010) using BOWTIE (version 0.12.7) (Langmead et al. 2009). Reads that failed to perfectly map to the nuclear genome with no mismatches were discarded. Each library was normalized to the total number of mapped reads multiplied by 10<sup>6</sup> (rpm, reads per million).

Classification of small RNAs was performed by BEDTools (v2.10.0) (Quinlan and Hall 2010) and in-house UNIX shell programming using the following databases: Bd v1.0 annotations for protein coding features downloading from www.Brachypodium.org, miRBase (release 18) (Kozomara and Griffiths-Jones 2010) for mature miRNA annotations. Some smRNAs match more than one annotation category; therefore the sum of the numbers is bigger than the total input number. The small RNA reads of 15–40 nt length were calculated and plotted versus the sum of their normalized reads per million (rpm).

The data of smRNA transcriptomes under various abiotic stresses challenges were plotted as heatmap on all five chromosomes (total of eight smRNA transcriptome libraries). The circular heat map visualization of smRNA transcriptome was drawn using Circos (Krzywinski et al. 2009). smRNA expression is represented in 10-kb blocks and the maximum value of the heat map is calibrated to the Bd21 (un-stress treated library). The outer annotation track highlights the position of coding genes (Genes). In addition to the smRNA transcriptomes, the genomic locations of 3' UTRs that serve as precursors to stress-induced sutr-siRNAs and the targeted genes of sutr-siRNAs were displayed on all five chromosomes in the same circular map.

## The analysis of smRNA producing genomic clusters

The analysis of smRNA producing genomic clusters was done using custom in-house UNIX shell programming. For clustering analysis, *Brachypodium* chromosomes were binned into nonoverlapping 500 nt clusters and an expression of 20–25 nt smRNAs was quantified in each cluster (Lee et al. 2012). *Brachypodium* consists of 1,084,598 nonoverlapping fixed size 500 bp clusters covering the complete nuclear genome (chromosome 1–5). The expression of 20–25 nt smRNAs mapped to each 500 nt genomic cluster was normalized to the total number of mapped reads multiplied by  $10^6$  (rpm, reads per million) and expression of smRNAs (in RPMs) produced from each individual 500 bp cluster was then cross-compared between unstressed plants and plants subjected to stresses.

## Calculation for genome-wide isolation of 3' UTRs that produce sutr-siRNAs

Calculation for genome-wide isolation of 3' UTRs that produce sutr-siRNAs was done using custom in-house UNIX shell programming applying the following rules: First, isolate all 3' UTRs that could give rise to smRNAs in a collinear fashion with their precursor transcripts. Second, calculate the abundance of 24 nt smRNAs in each individual 3' UTR and retain only the group of 3' UTRs that exhibit 24 nt smRNA expression both in response to stresses and in unstressed Bd21 plants. Third, retain only 3' UTRs that exhibit at least a threefold increase in expression of 24 nt smRNAs in response to stress. Fourth, isolate the group of 3' UTRs that produce smRNAs with perfect complementarity to transcripts of other genes.

## The analysis of RNA secondary structure

The analysis of RNA secondary structure of sutr-siRNAs producing 3' UTRs was done using mfold (Zuker 2003). The loci that were found to harbor any tandem or inverted repeat sequences were excluded from further analysis according to the criteria for annotating plant miRNAs (Meyers et al. 2008). For the 3'-UTR loci that did not contain any repetitive sequences, the flanking regions (300 bp upstream and 20 bp downstream, 150 bp upstream and 20 bp downstream, 150 bp upstream and 150 bp downstream, 20 bp upstream and 150 bp downstream, and 20 bp upstream and 300 bp downstream) of the sutr-siRNA producing sites were subjected to RNA secondary structure folding analysis (Lu et al. 2008). The flanking sequences were extracted from Bd21 v1.0 genomic sequences (International Brachypodium Initiative 2010) and RNA folded using mfold using the default parameters (Zuker 2003). The two lowest folding energies for each secondary structure were then inspected manually. The folds were discarded as putative pri-miRNA if the secondary structure revealed the miRNA/miRNA\* duplex had more than four mismatches, showed asymmetric bulges with more than two bases or present more than one time, a free energy of  $\geq 23$  kcal/mol, or the miRNA/miRNA\* duplex being located  $> 10$  bases from the terminal loop (Lai et al. 2003; Meyers et al. 2008).

## Isolation and analysis of sutr-siRNA targets

The *Brachypodium* genome annotations (v1.0 and v1.2) were used to classify the genomic features of sutr-siRNA targeting sites (International

Brachypodium Initiative 2010) and the analysis was performed by BEDTools (v2.10.0) (Quinlan and Hal 2010) and in-house UNIX shell programming. The predicted branch point sequences for both *Arabidopsis thaliana* and *Oryza sativa* (brachydownload.gz) were downloaded from Database of Plant Splice Sites and Splicing Signals (<http://lemur.amu.edu.pl/share/ERISdb/home.html>) (Szczęśniak et al. 2013). NCBI BLASTN was used to search for sequences complementary to the 24-nt sutr-siRNAs smRNA sequences. The nucleotide frequency of branch point sequences was calculated and graphically displayed using Web Logo (Crooks et al. 2004).

## Gene ontology analyses

The brachy\_v1.0 GO terms library was downloaded from the Munich information center for protein sequence (MIPS) (<http://mips.helmholtz-muenchen.de/>) (International Brachypodium Initiative 2010). GO terms were retrieved for both stress-induced 3'-UTR loci and the *trans*-target coding genes, and the complete list of GO terms is available in Supplemental Tables S4 and S13. The GO terms were further analyzed using simple clustering algorithm provided by REVIGO with standard parameters (Supek et al. 2011), which the REVIGO software is used to summarize the GO-terms in all stress conditions separately.

For each stress condition data set, the enriched biological processes are analyzed using REVIGO software and the resultant network generated by REVIGO was uploaded into Cytoscape (Smoot et al. 2011; Supek et al. 2011) for further editing and visualization modifications before the images are exported as final enrichment maps. The colors of the circles represent the various levels of GO-term enrichment, where the darker shade represents more represented GO-terms in specific data sets than the lighter one.

## DATA DEPOSITION

The smRNA sequencing data used in this study are available from the Gene Expression Omnibus (GEO) user accession number GSE55217.

## SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

## ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation (NSF) (Grant #0724168), the Missouri Life Sciences Research Board (MLSRB) (Grant #13230), the US Department of Agriculture (USDA) (Grant #35301-18207), and the National Institutes of Health (NIH) (#GM073872) to J.A.C.

Received August 14, 2014; accepted November 4, 2014.

## REFERENCES

Ali GS, Reddy ASN. 2008. Regulation of alternative splicing of pre-mRNAs by stresses. *Curr Top Microbiol Immunol* **326**: 257–275.

Alló M, Kornblihtt AR. 2010. Gene silencing: Small RNAs control RNA polymerase II elongation. *Curr Biol* **20**: R704–R707.

Alló M, Buggiano V, Fededa JP, Petrillo E, Schor I, la Mata de M, Aguirre E, Plass M, Eyras E, Elela SA, et al. 2009. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* **16**: 717–724.

Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, Morozova N, Fenouil R, Descotes N, Andrau JC, et al. 2012. Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol* **19**: 998–1004.

Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ. 2010. Deciphering the splicing code. *Nature* **465**: 53–59.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.

Belostotsky D. 2004. mRNA turnover meets RNA interference. *Mol Cell* **16**: 498–500.

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**: 291–336.

Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A; Galaxy Team. 2010. Manipulation of FASTQ data with Galaxy. *Bioinformatics* **26**: 1783–1785.

Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* **123**: 1279–1291.

Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, Kovalchuk I. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants (virus-induced plant genome instability). *Nucleic Acids Res* **35**: 1714–1725.

Braunschweig U, Guerousov S, Plocik AM, Graveley BR, Blencowe BJ. 2013. Dynamic integration of splicing within gene regulatory pathways. *Cell* **152**: 1252–1269.

Brkljacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, Brutnell T, Sibout R, Bevan M, Budak H, et al. 2011. *Brachypodium* as a model for the grasses: today and the future. *Plant Physiol* **157**: 3–13.

Brogna S, Wen J. 2009. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16**: 107–113.

Castel SE, Martienssen RA. 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* **14**: 100–112.

Chapman EJ, Carrington JC. 2007. Specialization and evolution of endogenous small RNA pathways. *Nat Rev Genet* **8**: 884–896.

Chen X. 2009. Small RNAs and their roles in plant development. *Annu Rev Cell Dev Biol* **25**: 21–44.

Chen M, Manley JL. 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomic approaches. *Nat Rev Mol Cell Biol* **10**: 741–754.

Coté J, Dupuis S, Jiang Z, Wu JY. 2001. Caspase-2 pre-mRNA alternative splicing: identification of an intronic element containing a decoy 3' acceptor site. *Proc Natl Acad Sci* **98**: 938–943.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14**: 1188–1190.

Dominski Z, Kole R. 1993. Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc Natl Acad Sci* **90**: 8673–8677.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci* **109**: E2183–E2191.

Draper J, Mur LA, Jenkins G, Ghosh-Biswas GC, Bablak P, Hasterok R, Routledge AP. 2001. *Brachypodium distachyon*. A new model system for functional genomics in grasses. *Plant Physiol* **127**: 1539–1555.

Dujardin G, Lafaille C, de la Mata M, Marasco LE, Muñoz MJ, Le Jossic-Corcos C, Corcos L, Kornblihtt AR. 2014. How slow RNA polymerase II elongation favors alternative exon skipping. *Mol Cell* **54**: 683–690.

Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong WK, Mockler TC. 2010. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res* **20**: 45–58.

Gazzani S, Lawrenson T, Woodward C, Headon D, Sablowski R. 2004. A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science* **306**: 1046–1048.

Goren A, Ram O, Amit M, Keren H, Lev-Maor G, Vig I, Pupko T, Ast G. 2006. Comparative analysis identifies exonic splicing regulatory sequences—the complex definition of enhancers and silencers. *Mol Cell* **22**: 769–781.

Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR. 2008. A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev Cell* **14**: 854–866.

Hiller M, Zhang Z, Backofen R, Stamm S. 2007. Pre-mRNA secondary structures influence exon recognition. *PLoS Genet* **3**: e204.

Hirayama T, Shinozaki K. 2010. Research on plant abiotic stress responses in the post-genome era: past, present and future. *Plant J* **61**: 1041–1052.

Hong S-Y, Seo P, Yang MS, Xiang F, Park CM. 2008. Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. *BMC Plant Biol* **8**: 112.

Hoskins AA, Friedman LJ, Gallagher SS, Crawford DJ, Anderson EG, Wombacher R, Ramirez N, Cornish VW, Gelles J, Moore MJ. 2011. Ordered and dynamic assembly of single spliceosomes. *Science* **331**: 1289–1295.

Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR. 2010. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev* **24**: 1634–1644.

Huang CF, Miki D, Tang K, Zhou HR, Zheng Z, Chen W, Ma ZY, Yang L, Zhang H, Liu R, et al. 2013. A pre-mRNA-splicing factor is required for RNA-directed DNA methylation in *Arabidopsis*. *PLoS Genet* **9**: e1003779.

International Brachypodium Initiative. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**: 763–768.

Jeong DH, Schmidt SA, Rymarquis LA, Park S, Ganssmann M, German MA, Accerbi M, Zhai J, Fahlgren N, Fox SE, et al. 2013. Parallel analysis of RNA ends enhances global investigation of microRNAs and target RNAs of *Brachypodium distachyon*. *Genome Biol* **14**: R145.

Jia F, Rock CD. 2013. *MIR846* and *MIR842* comprise a cistronic *MIRNA* pair that is regulated by abscisic acid by alternative splicing in roots of *Arabidopsis*. *Plant Mol Biol* **81**: 447–460.

Kastenmayer JP, Green PJ. 2000. Novel features of the XRN-family in *Arabidopsis*: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. *Proc Natl Acad Sci* **97**: 13985–13990.

Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. 2007. A novel class of bacteria-induced small RNAs in *Arabidopsis*. *Genes Dev* **21**: 3123–3134.

Kishore S, Stamm S. 2006. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* **311**: 230–232.

Kishore S, Khanna A, Zhang Z, Hui J, Balwierz PJ, Stefan M, Beach C, Nicholls RD, Zavolan M, Stamm S. 2010. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum Mol Genet* **19**: 1153–1164.

Kole R, Krainer AR, Altman S. 2012. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* **11**: 125–140.

Kozomara A, Griffiths-Jones S. 2010. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**: D152–D157.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**: 1639–1645.

Kumar K, Rao KP, Sharma P, Sinha AK. 2008. Differential regulation of rice mitogen activated protein kinase kinase (MKK) by abiotic stress. *Plant Physiol Biochem* **46**: 891–897.

Lai EC, Tomancak P, Williams RW, Rubin GM. 2003. Computational identification of *Drosophila* microRNA genes. *Genome Biol* **4**: R42.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**: 204–220.

Lee T-F, Gurazada SGR, Zhai J, Li S, Simon SA, Matzke MA, Chen X, Meyers BC. 2012. RNA polymerase V-dependent small RNAs in *Arabidopsis* originate from small, intergenic loci including most SINE repeats. *Epigenetics* **7**: 781–795.

Leung AK, Young AG, Bhutkar A, Zheng GX, Bosson AD, Nielsen CB, Sharp PA. 2011. Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat Struct Mol Biol* **18**: 237–244.

Liu J, Hu J, Corey DR. 2012. Expanding the action of duplex RNAs into the nucleus: redirecting alternative splicing. *Nucleic Acids Res* **40**: 1240–1250.

Long JC, Caceres JF. 2009. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* **417**: 15–27.

Lu C, Jeong DH, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D, et al. 2008. Genome-wide analysis for discovery of rice microRNAs reveals natural anti-sense microRNAs (nat-miRNAs). *Proc Natl Acad Sci* **105**: 4951–4956.

Lu T, Zhu C, Lu G, Guo Y, Zhou Y, Zhang Z, Zhao Y, Li W, Lu Y, Tang W, et al. 2012. Strand-specific RNA-seq reveals widespread occurrence of novel *cis*-natural antisense transcripts in rice. *BMC Genomics* **13**: 721.

Luco RF, Alló M, Schor IE, Kornblith AR, Misteli T. 2011. Epigenetics in alternative pre-mRNA splicing. *Cell* **144**: 16–26.

Mahajan S, Tuteja N. 2005. Cold, salinity and drought stresses: an overview. *Arch Biochem Biophys* **444**: 139–158.

Marquez Y, Brown JW, Simpson C, Barta A, Kalyna M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res* **22**: 1184–1195.

Martinez-Contreras R, Cloutier P, Shkreta L, Fisette JF, Revil T, Chabot B. 2007. hnRNP proteins and splicing control. *Adv Exp Med Biol* **623**: 123–147.

Mastrangelo AM, Marone D, Laidò G, De Leonardi AM, De Vita P. 2012. Alternative splicing: enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci* **185–186**: 40–49.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ. 2009. RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* **21**: 367–376.

Meng Y, Shao C, Ma X, Wang H. 2013. Introns targeted by plant microRNAs: a possible novel mechanism of gene regulation. *Rice (N Y)* **6**: 8.

Merret R, Descombin J, Juan YT, Favory JJ, Carpentier MC, Chaparro C, Chang YY, Deragon JM, Bousquet-Antonelli C. 2013. XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell Rep* **5**: 1279–1293.

Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, et al. 2008. Criteria for annotation of plant microRNAs. *Plant Cell* **20**: 3186–3190.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. 2008. Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**: 116–127.

Olmedo G, Guo H, Gregory BD, Nourizadeh SD, Aguilar-Henonin L, Li H, An F, Guzman P, Ecker JR. 2006. ETHYLENE-INSENSITIVE5 encodes a 5'→3' exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proc Natl Acad Sci* **103**: 13286–13293.

Palusa SG, Ali GS, Reddy AS. 2007. Alternative splicing of pre-mRNAs of *Arabidopsis* serine/arginine-rich proteins: regulation by hormones and stresses. *Plant J* **49**: 1091–1107.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40**: 1413–1415.

Priest HD, Fox SE, Rowley ER, Murray JR, Michael TP, Mockler TC. 2014. Analysis of global gene expression in *Brachypodium distachyon* reveals extensive network plasticity in response to abiotic stress. *PLoS One* **9**: e87499.

Quinlan AR, Hal IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Reddy AS. 2007. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu Rev Plant Biol* **58**: 267–294.

Reddy AS, Rogers MF, Richardson DN, Hamilton M, Ben-Hur A. 2012. Deciphering the plant splicing code: experimental and computational approaches for predicting alternative splicing and splicing regulatory elements. *Front Plant Sci* **3**: 18.

Reddy AS, Marquez Y, Kalyna M, Barta A. 2013. Complexity of the alternative splicing landscape in plants. *Plant Cell* **25**: 3657–3683.

Ruiz-Ferrer V, Voinnet O. 2009. Roles of plant small RNAs in biotic stress responses. *Annu Rev Plant Biol* **60**: 485–510.

Shepard PJ, Hertel KJ. 2008. Conserved RNA secondary structures promote alternative splicing. *RNA* **14**: 1463–1469.

Shin J-H, Wang HL, Lee J, Dinwiddie BL, Belostotsky DA, Chekanova JA. 2013. The role of the *Arabidopsis* exosome in siRNA-independent silencing of heterochromatic loci. *PLoS Genet* **9**: e1003411.

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**: 431–432.

Song CP, Agarwal M, Ohta M, Guo Y, Halfter U, Wang P, Zhu JK. 2005. Role of an *Arabidopsis* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. *Plant Cell* **17**: 2384–2396.

Staiger D, Brown JW. 2013. Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* **25**: 3640–3656.

Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. 2002. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem* **277**: 37741–37746.

Sun H, Chasin LA. 2000. Multiple splicing defects in an intronic false exon. *Mol Cell Biol* **20**: 6414–6425.

Sunkar R, Chinnusamy V, Zhu J, Zhu JK. 2007. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci* **12**: 301–309.

Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**: e21800.

Szcześniak MW, Kabza M, Pokrzywa R, Gudyś A, Makałowska I. 2013. ERIISdb: a database of plant splice sites and splicing signals. *Plant Cell Physiol* **54**: e10.

Taliaferro JM, Aspden JL, Bradley T, Marwha D, Blanchette M, Rio DC. 2013. Two new and distinct roles for *Drosophila* Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression. *Genes Dev* **27**: 378–389.

Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF. 2005. Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *Plant J* **41**: 195–211.

Walters B, Lum G, Sablok G, Min XJ. 2013. Genome-wide landscape of alternative splicing events in *Brachypodium distachyon*. *DNA Res* **20**: 163–171.

Wang BB, Brendel V. 2006. Genomewide comparative analysis of alternative splicing in plants. *Proc Natl Acad Sci* **103**: 7175–7180.

Wang Z, Burge CB. 2008. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**: 802–813.

Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, Ecker JR, Tang H, Pikaard CS. 2012. Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the *Arabidopsis* epigenome. *Genes Dev* **26**: 1825–1836.

Wong J, Gao L, Yang Y, Zhai J, Arikit S, Yu Y, Duan S, Chan V, Xiong Q, Yan J, et al. 2014. Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. *Plant J* **79**: 928–940.

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2**: E104.

Yan K, Liu P, Wu CA, Yang GD, Xu R, Guo QH, Huang JG, Zheng CC. 2012. Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in *Arabidopsis thaliana*. *Mol Cell* **48**: 521–531.

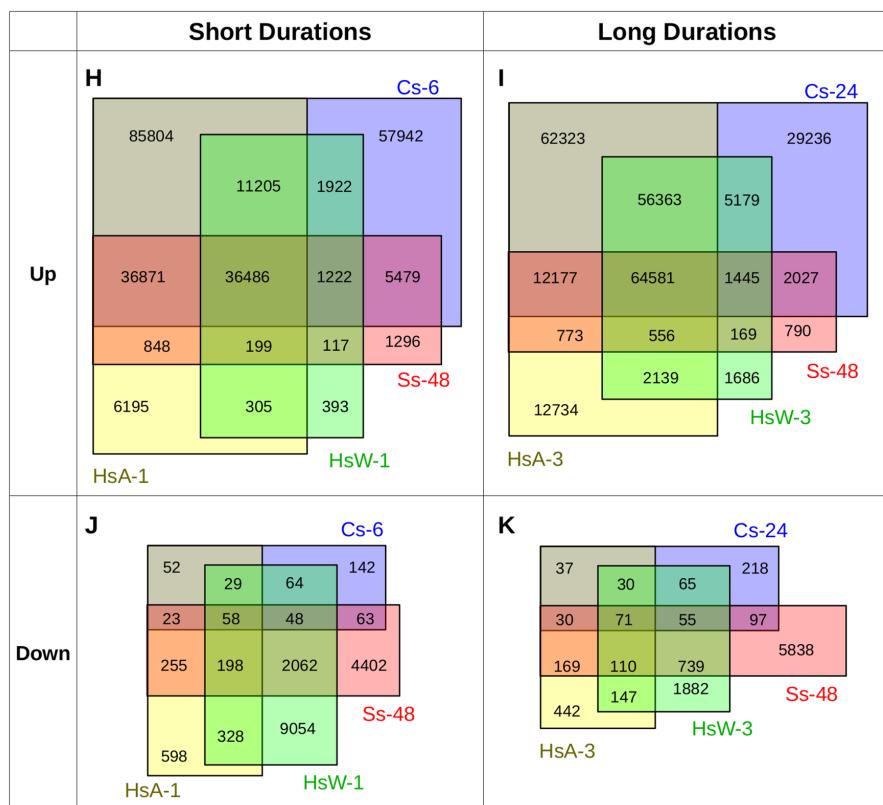
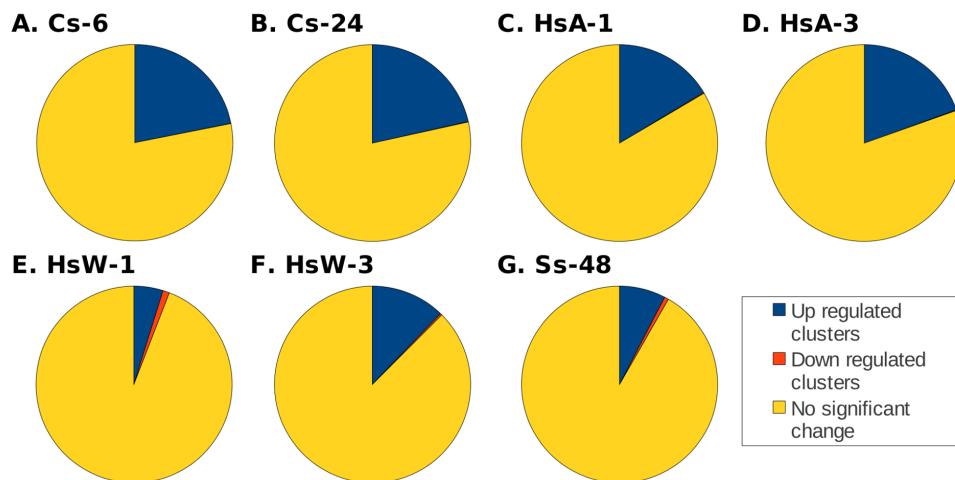
Zhang J, Xu Y, Huan Q, Chong K. 2009. Deep sequencing of *Brachypodium* small RNAs at the global genome level identifies microRNAs involved in cold stress response. *BMC Genomics* **10**: 449.

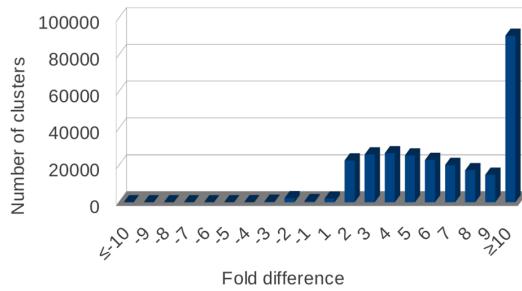
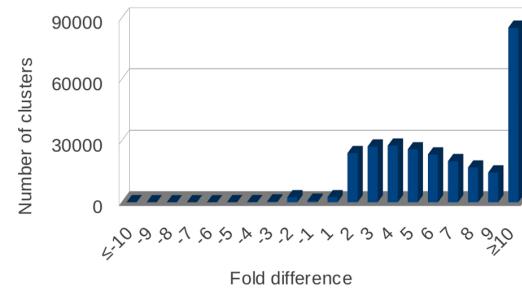
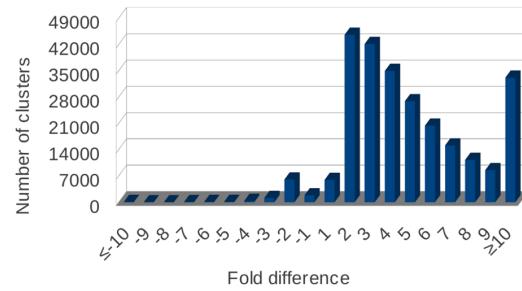
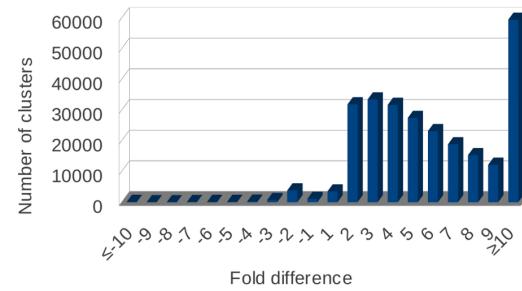
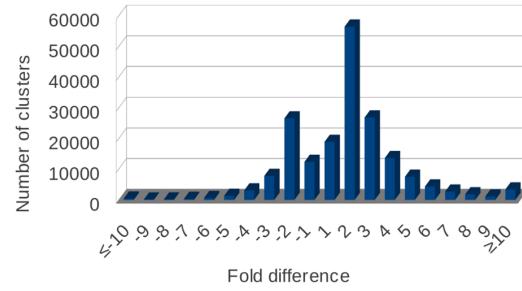
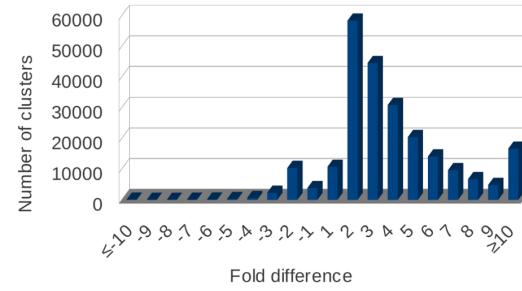
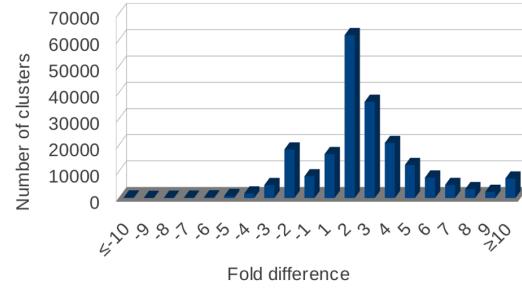
Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L, Niu D, Chen Z, Leung C, et al. 2012. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol* **13**: R20.

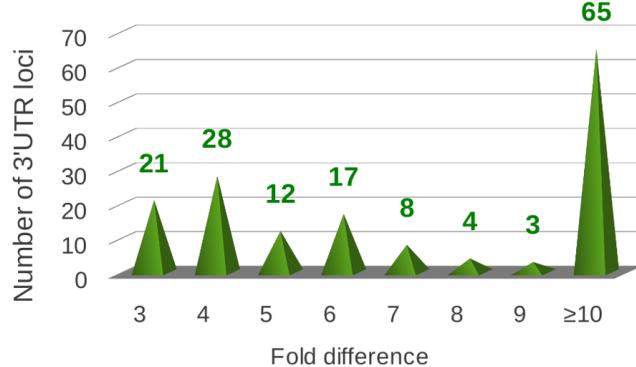
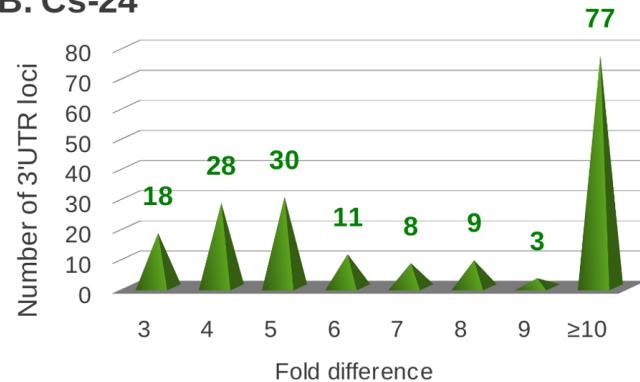
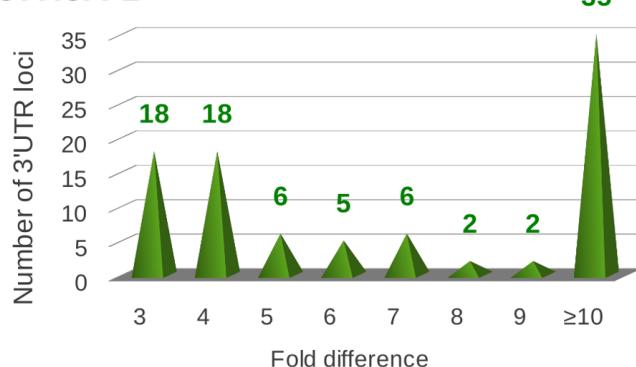
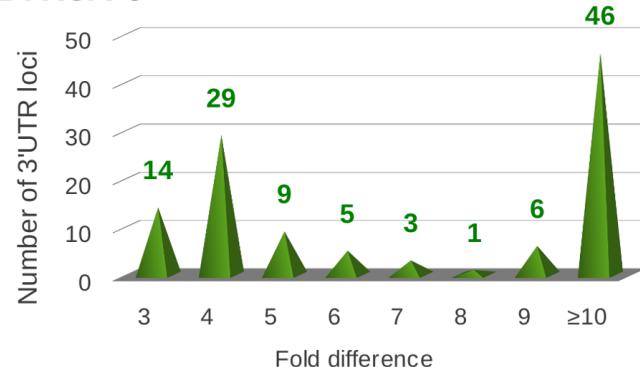
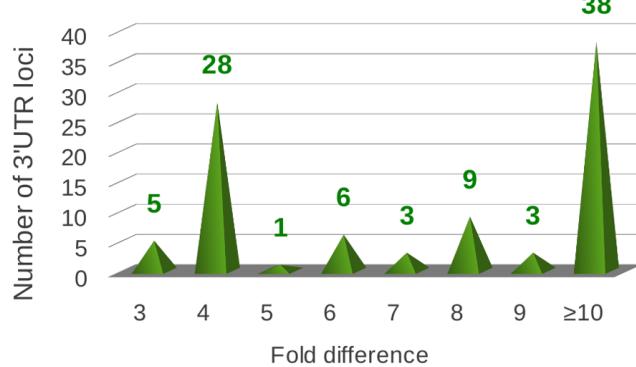
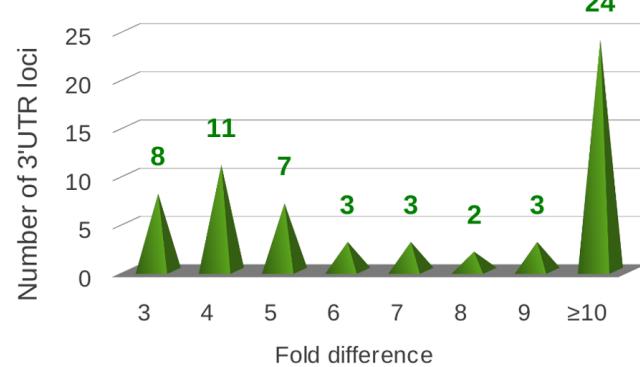
Zhang CJ, Zhou JX, Liu J, Ma ZY, Zhang SW, Dou K, Huang HW, Cai T, Liu R, Zhu JK, et al. 2013. The splicing machinery promotes RNA-directed DNA methylation and transcriptional silencing in *Arabidopsis*. *EMBO J* **32**: 1128–1140.

Zisouliis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW. 2010. Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat Struct Mol Biol* **17**: 173–179.

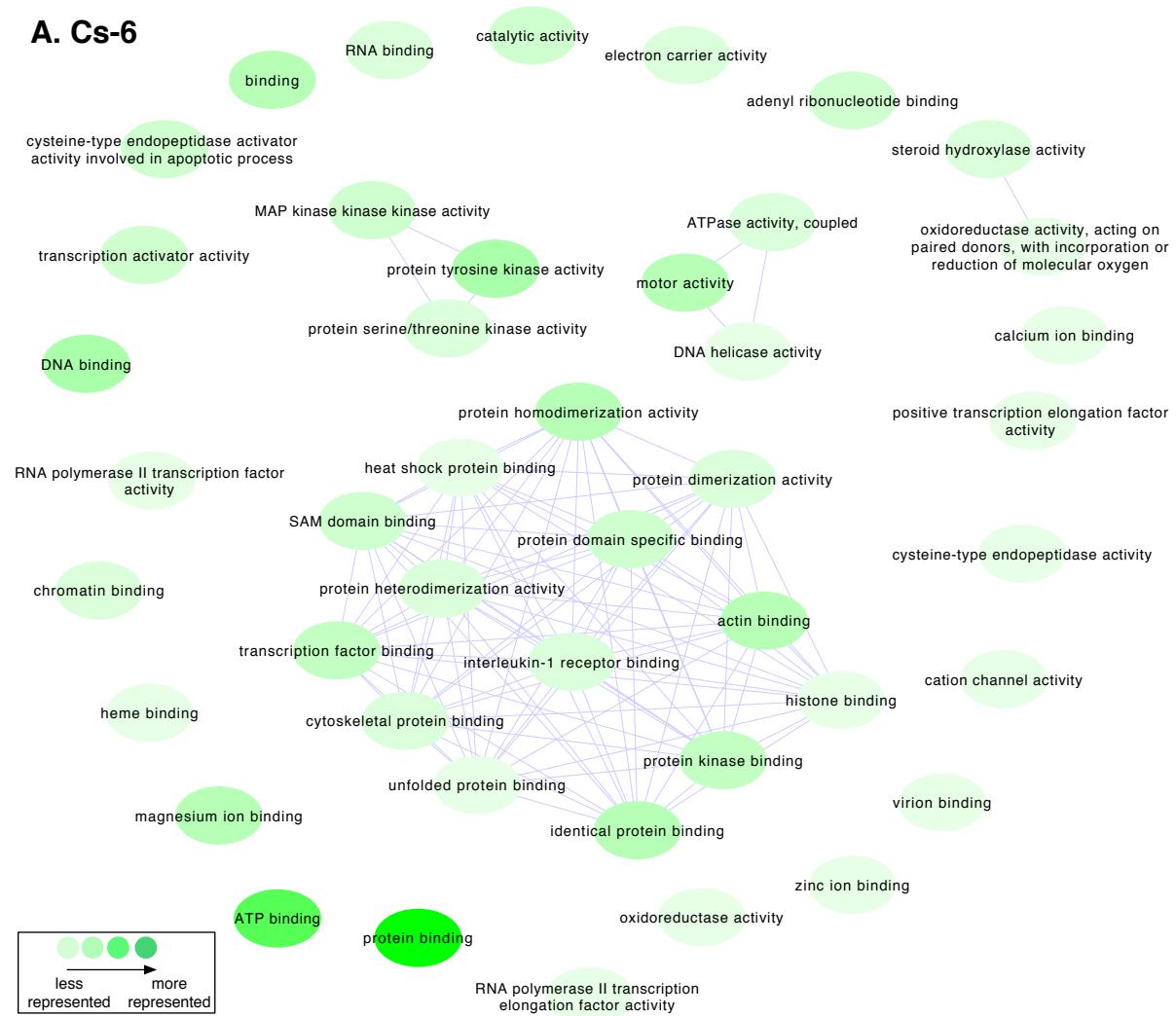
Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.



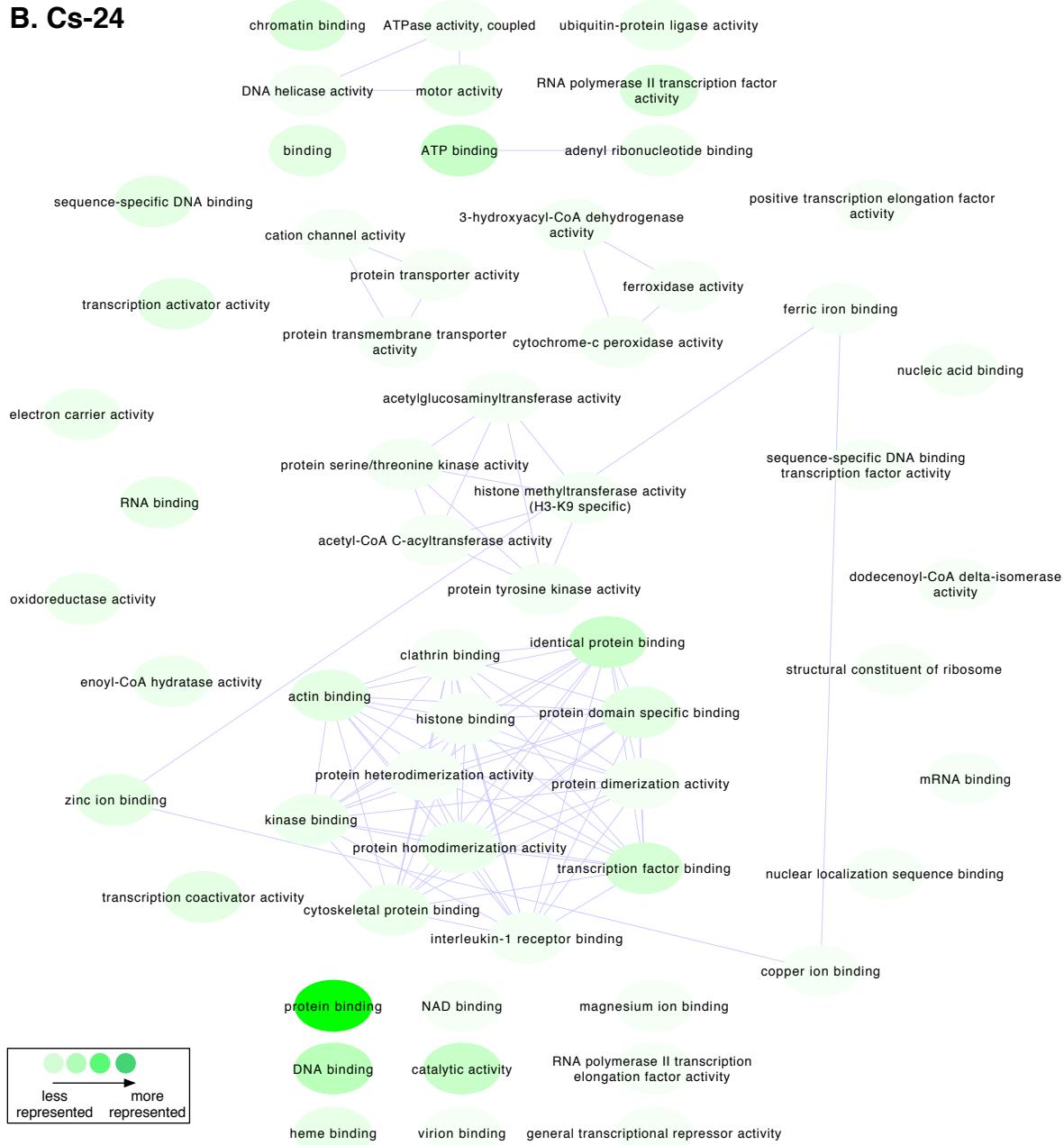
**A. Cs-6****B. Cs-24****C. HsA-1****D. HsA-3****E. HsW-1****F. HsW-3****G. Ss-48**

**A. Cs-6****B. Cs-24****C. HsA-1****D. HsA-3****E. HsW-1****F. HsW-3**

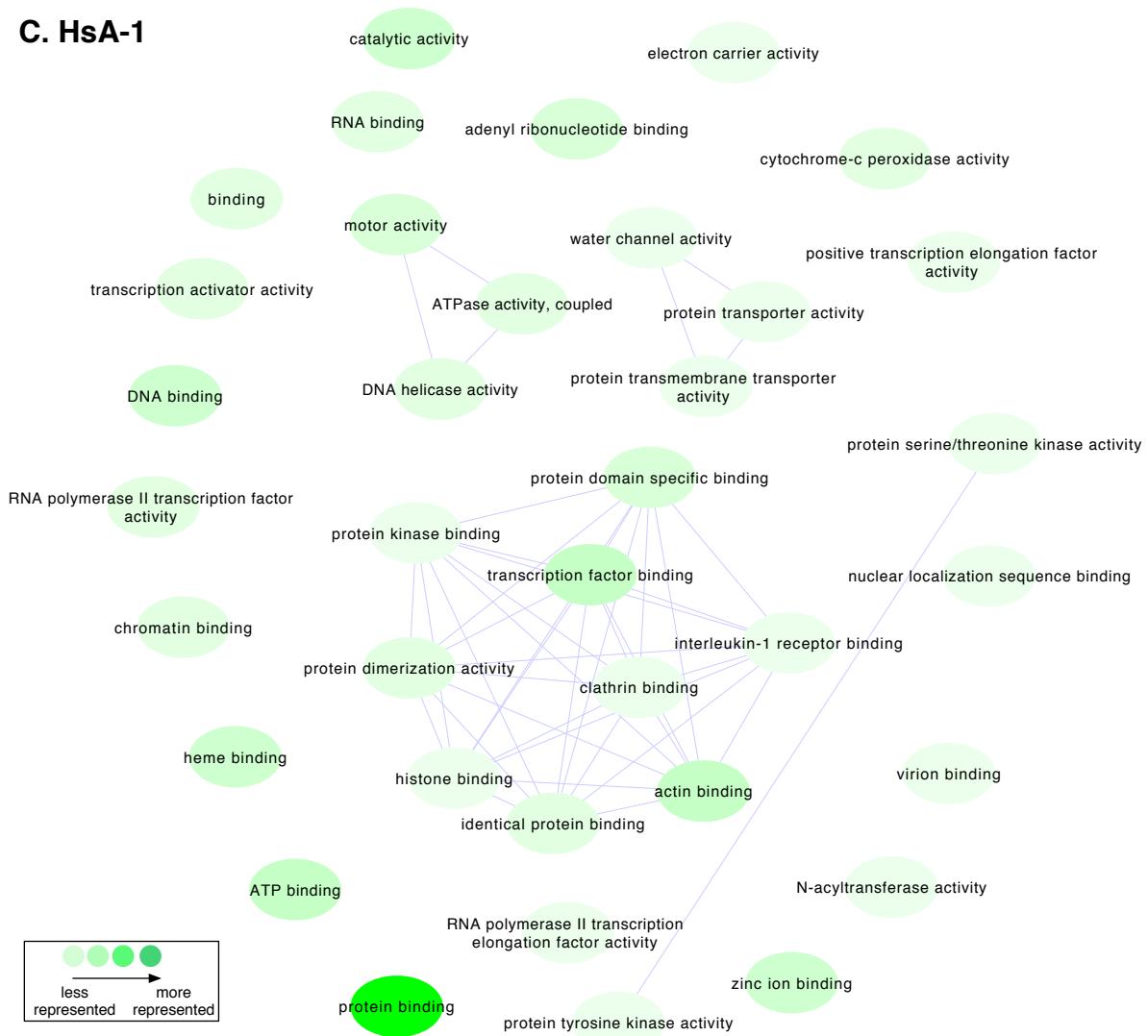
# A. Cs-6



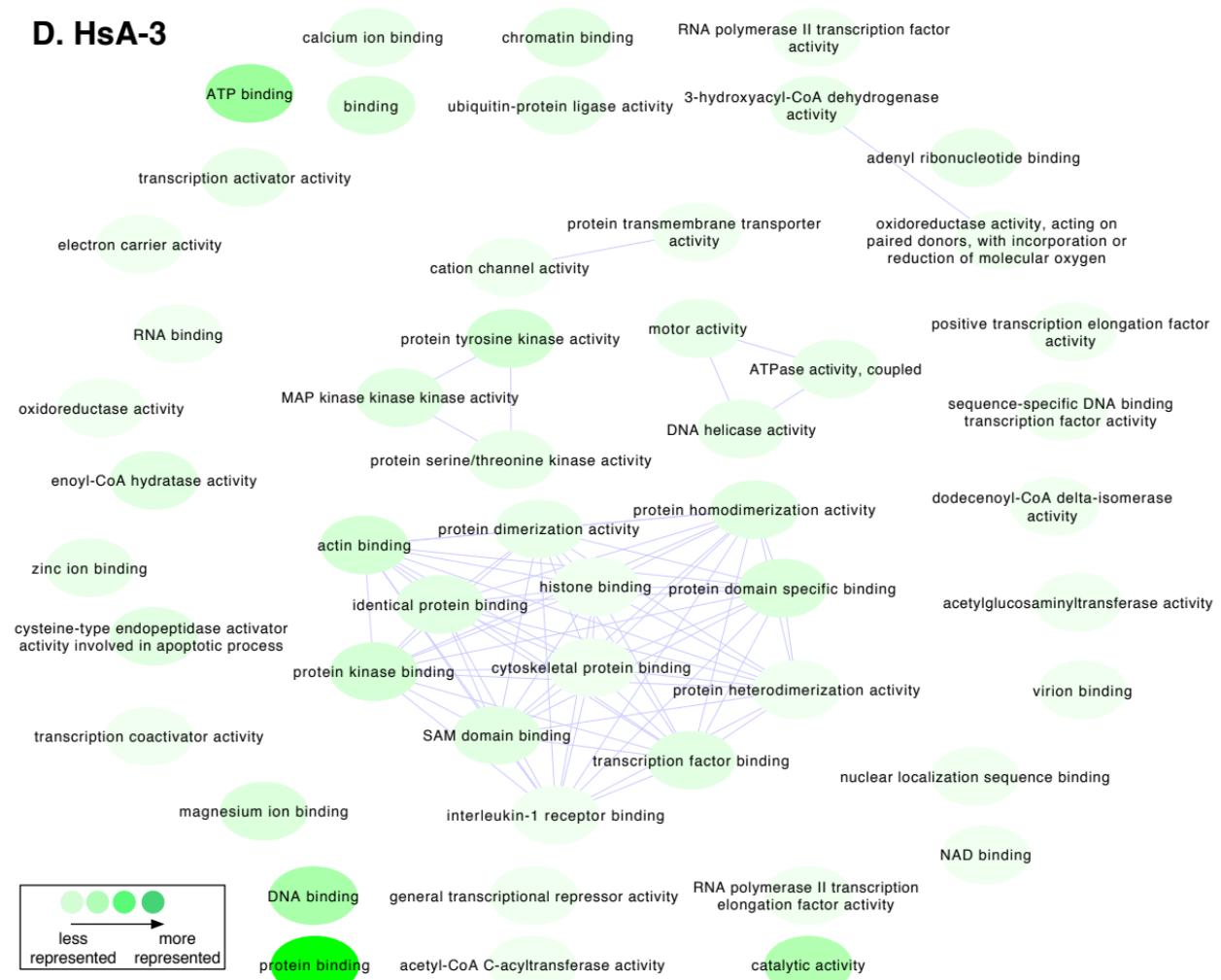
## B. Cs-24



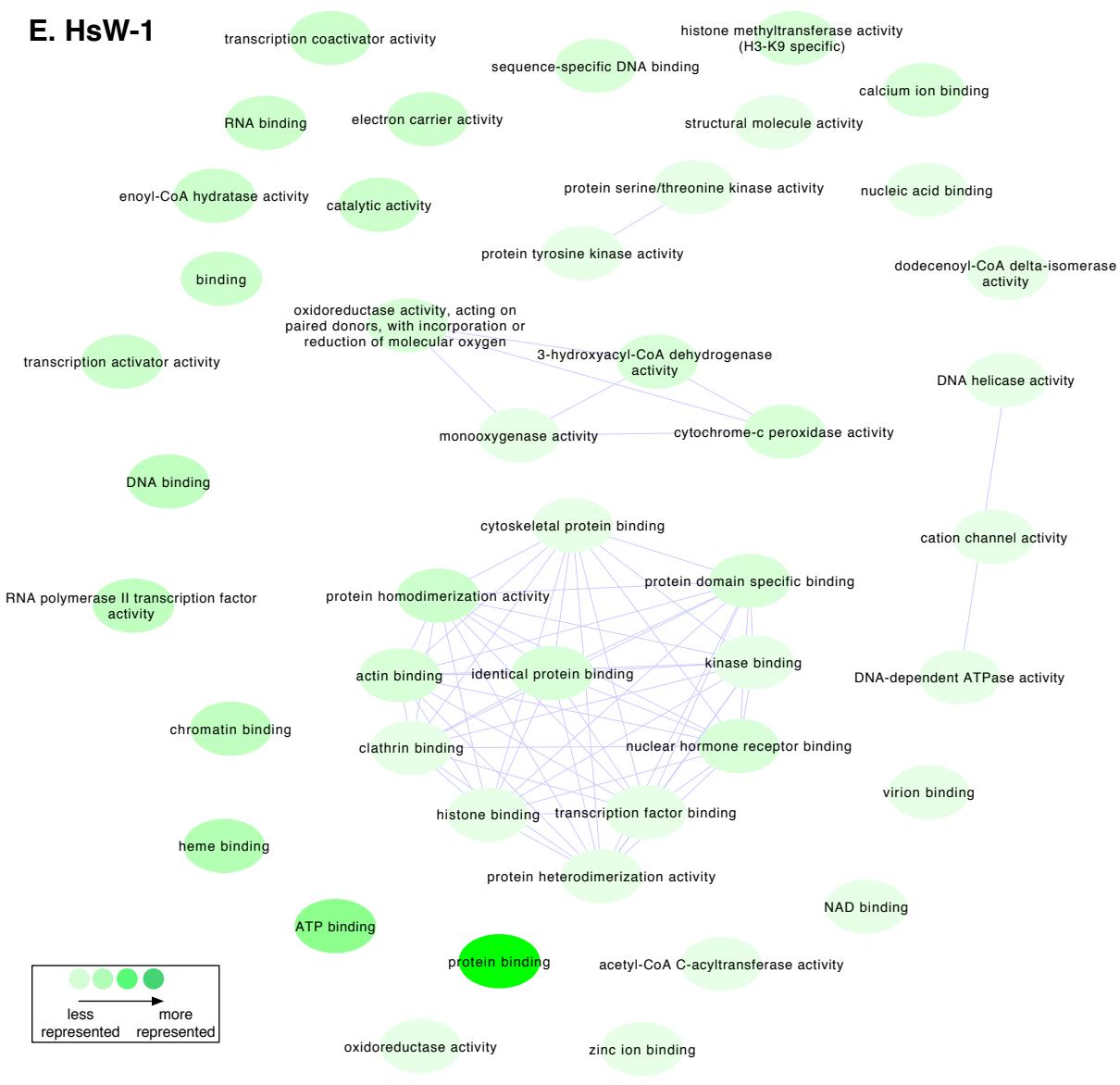
### C. HsA-1



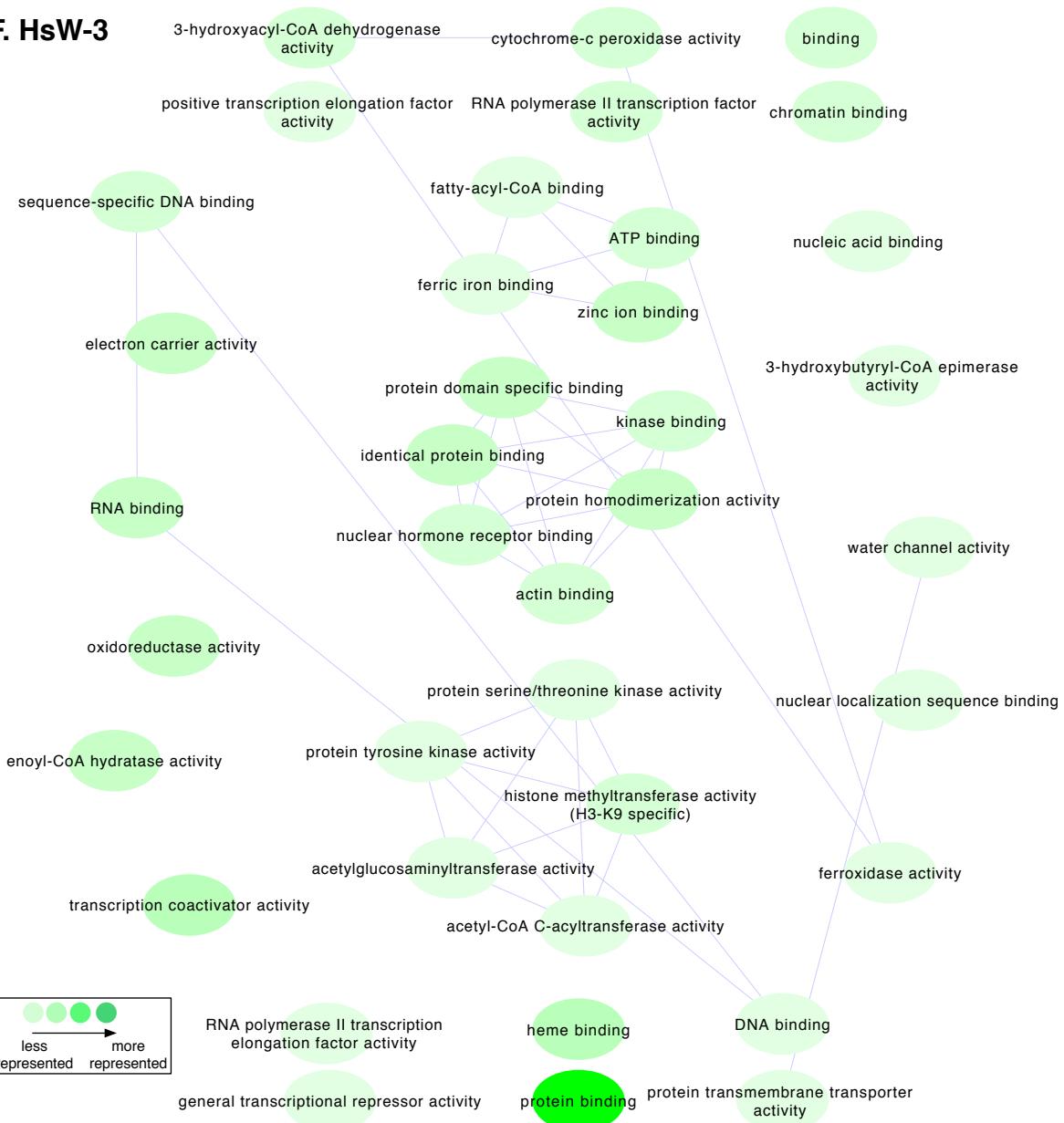
# D. HsA-3



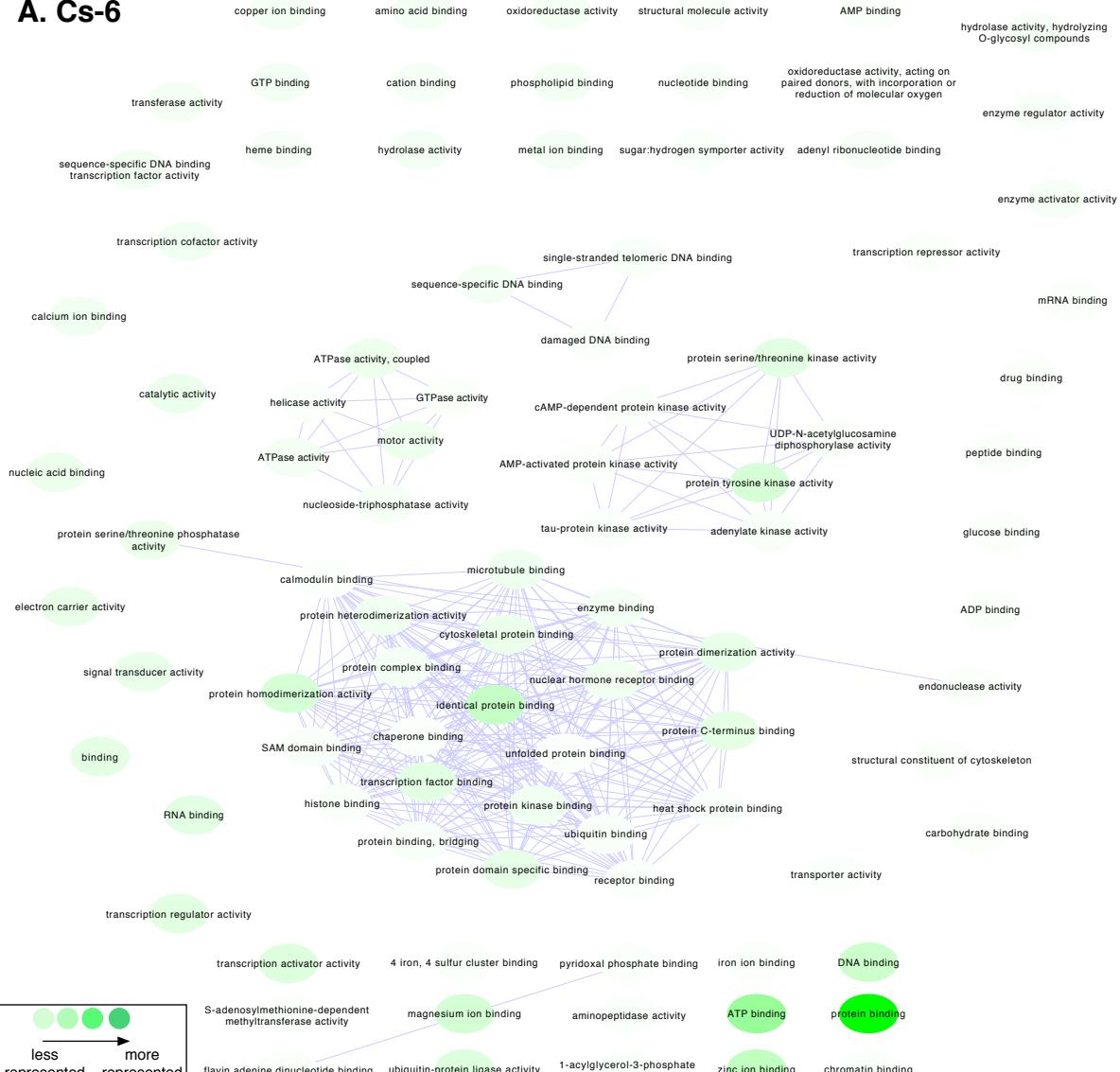
# E. HsW-1



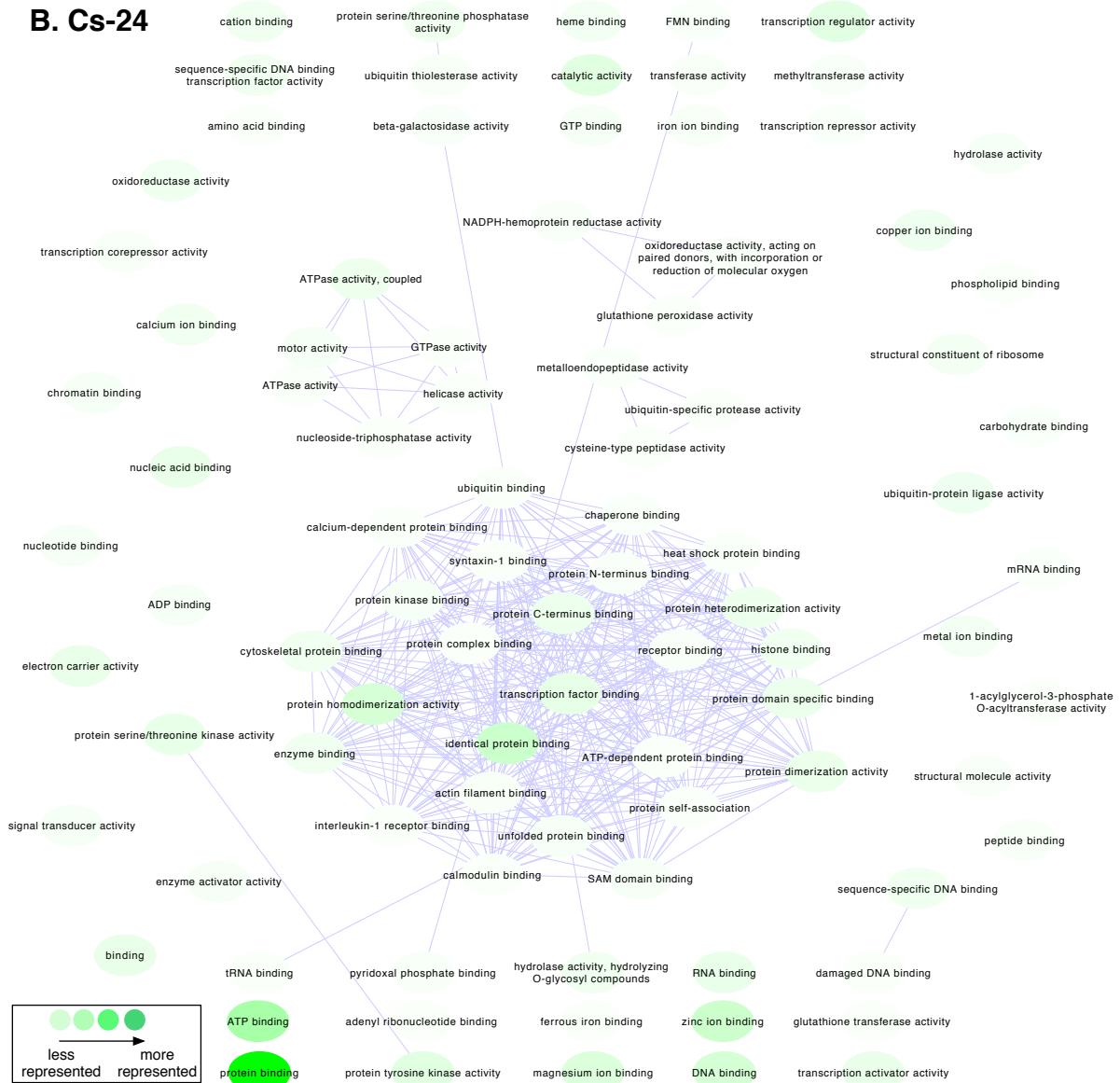
# F. HsW-3



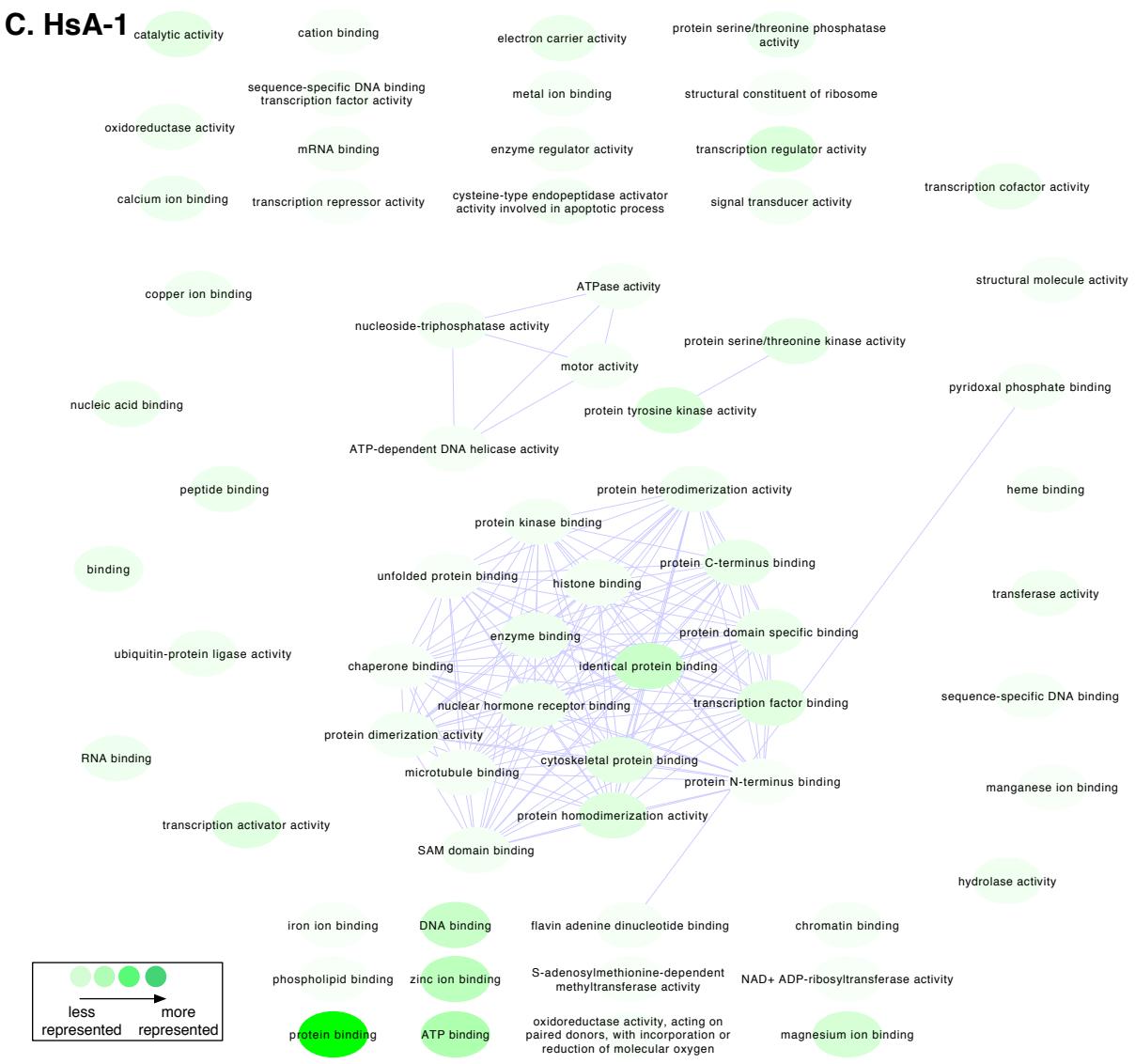
# A. Cs-6



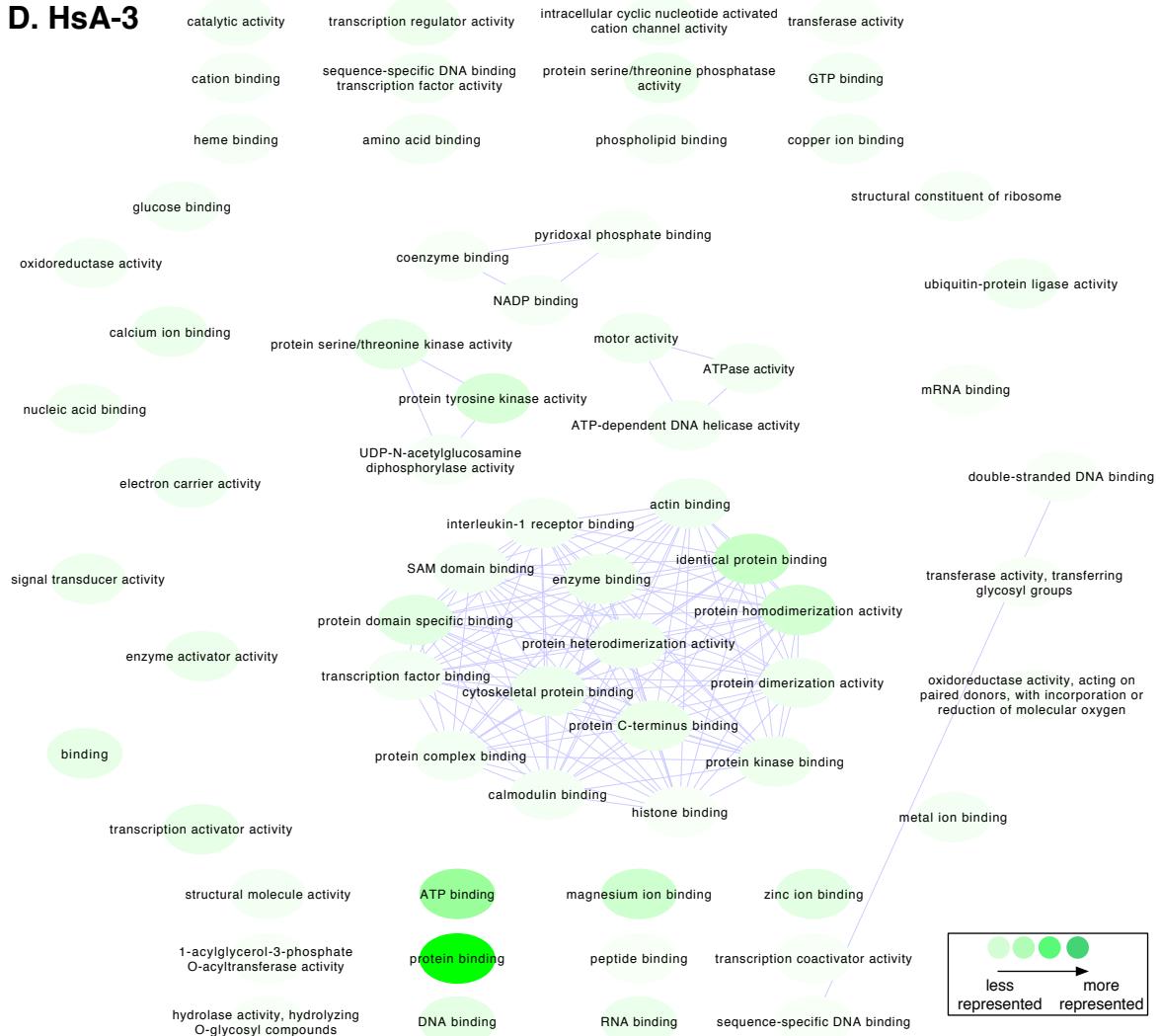
## B. Cs-24



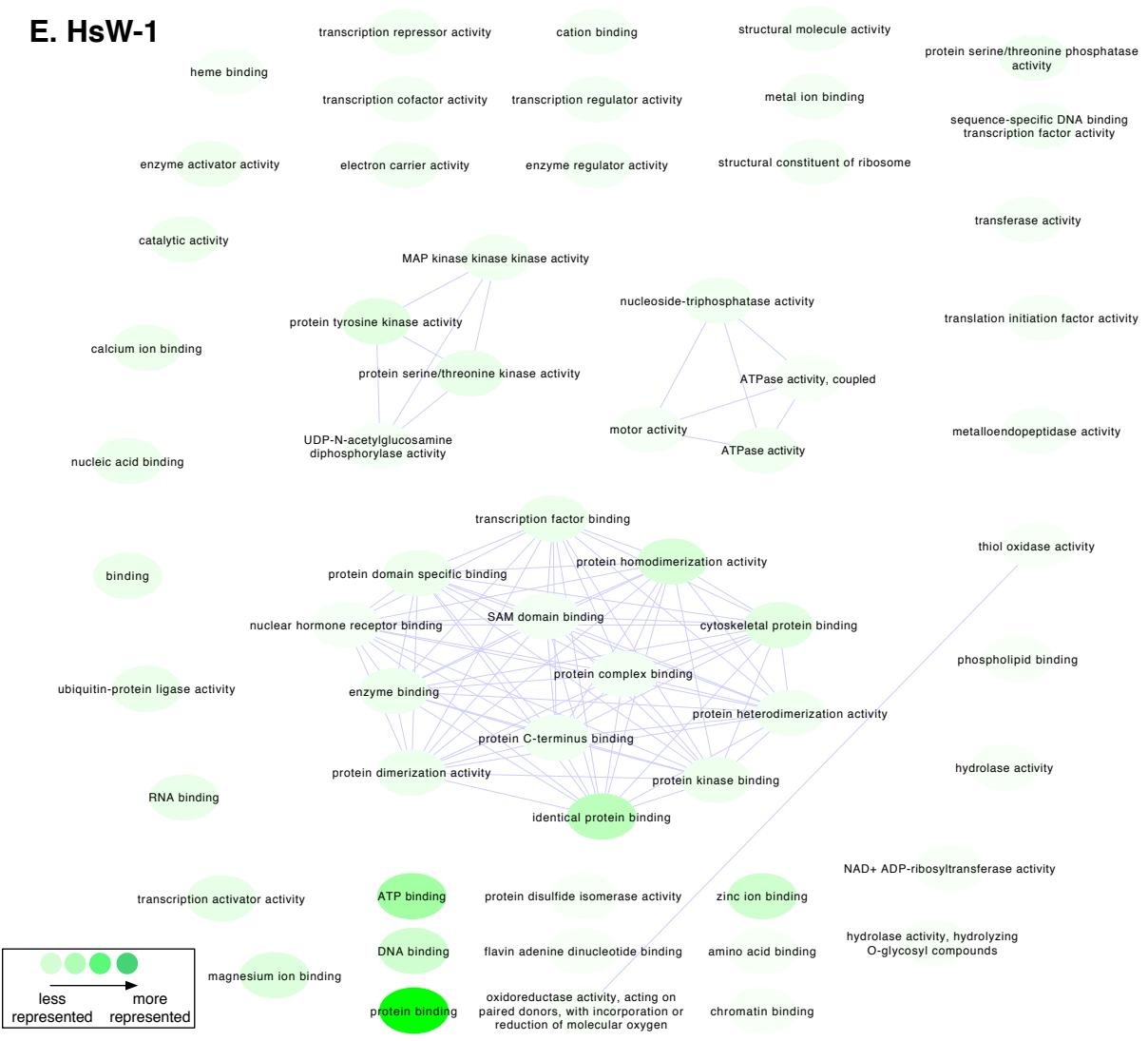
# C. HsA-1



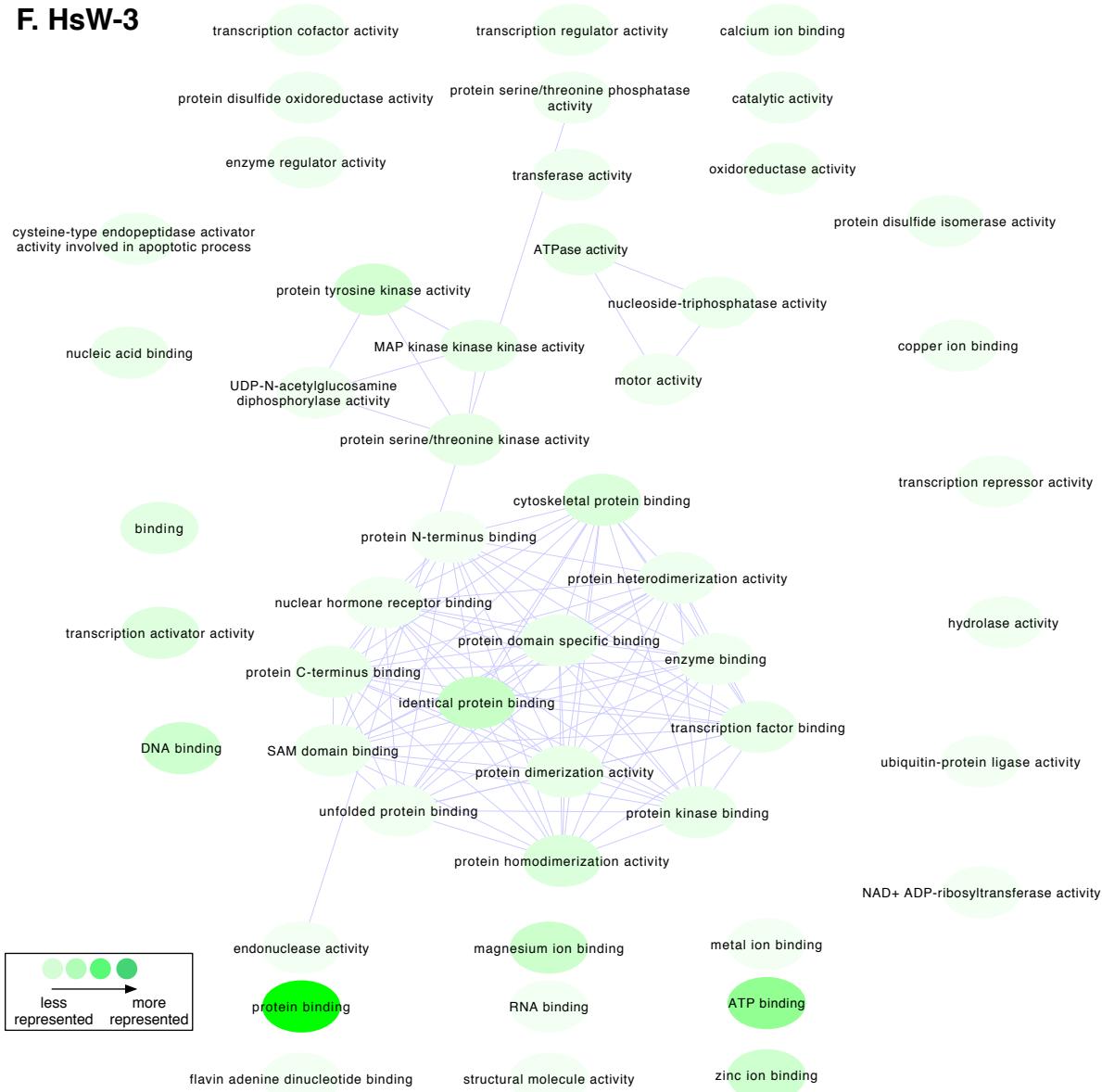
# D. HsA-3



# E. HsW-1



# F. HsW-3



## SUPPLEMENTARY INFORMATION

### Comparative analysis of genomic clusters producing smRNAs in response to various stresses

We also analyzed and compared the genomic regions producing smRNAs in response to various stresses. *Brachypodium* chromosomes were binned into non-overlapping 500 bp clusters and smRNAs were mapped to these genomic clusters. Since the majority of functional smRNAs in plants range from 20 to 25 nt in length, only populations of 20-25 nt smRNAs were used in this analysis. The expression of smRNAs produced from each cluster was then calculated and the expression of smRNAs from each cluster was compared between unstressed plants and plants subjected to various stresses. Most of the genomic clusters did not exhibit significant changes in smRNA production in response to stresses (based on 3-fold difference in expression), and the majority of clusters that exhibited changes in smRNA expression in response to stresses showed increases in smRNA production, while a relatively small number of genomic clusters produced fewer smRNAs in response to all stresses (Additional file 3: Figure S1 A-G, Additional file 1: Table S3, Additional file 4: Figure S2). It is worth noting that a high number of genomic clusters showed a greater than a 10-fold increase in smRNA production in response to cold stress (Additional file 4: Figure S2).

We then cross-compared the genomic clusters that exhibited either increases or decreases in smRNA expression between different stresses. The area-proportional Venn diagrams show the number of smRNA generating clusters that are affected in either a stress-specific way or affected similarly in response to various stress treatments (Additional file 3: Figure S1 H-K). The comparison was done separately for short (Additional file 3: Figure S1 H and J) or long (Additional file 3: Figure S1 I and K) stress treatments. Many genomic clusters showed increased smRNA production in all stresses used in our study. However, cold stress and heat stress (by air) affected the production of smRNAs from many identical genomic clusters (depicted as overlap between affected clusters in the area-proportional Venn diagrams), suggesting that similar smRNA pathways could be activated in response to these two abiotic stresses (Additional file 3: Figure S1 H-K).