Genome-wide mapping of the *HY5*-mediated gene networks in Arabidopsis that involve both transcriptional and post-transcriptional regulation

Huiyong Zhang[1–3], Hang He[1,2], Xuncheng Wang[1], Xiangfeng Wang[2], Xiaozeng Yang[3], Lei Li[3,*] and Xing W. Deng[1,2,*]

[1]*Peking–Yale Joint Center of Plant Molecular Genetics and Agrobiotechnology, National Laboratory of Protein Engineering and Plant Genetic Engineering, School of Life Sciences, Peking University, Beijing 100871, China,*
[2]*Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA, and*
[3]*Department of Biology, University of Virginia, Charlottesville, VA 22904, USA*

SUMMARY

LONG HYPOCOTYL 5 (HY5) is a basic leucine zipper transcription factor (TF) that functions downstream of multiple families of photoreceptors. Mutations in the *HY5* gene cause a myriad of aberrant phenotypes in Arabidopsis, including elongated hypocotyl, reduced accumulation of pigments, halted chloroplast development in greening hypocotyls, altered root morphology, and defective hormonal and stimulus responses. HY5 thus acts as an integrator that links various gene networks to coordinate plant development. Here we report the experimental mapping of *HY5*-mediated gene networks in Arabidopsis by integrating genomic loci occupied by HY5 and *HY5*-dependent gene expression profiles. Our results indicate that HY5 binds to over 9000 genes, detectably affecting the expression of over 1100 genes, either positively or negatively. Further, HY5 indirectly regulate many other genes through sub-networks mediated by other regulators. In particular, HY5 regulates eight miRNA genes that in turn control the transcript abundance of specific target genes. Over-expressing HY5-targeted miR408 resulted in phenotypes that are opposite to the *hy5* mutants. Together, our results reveal both transcriptional and post-transcriptional components of the *HY5*-mediated gene networks.

Keywords: HY5, tiling microarray, gene networks, microRNA, Arabidopsis.

INTRODUCTION

Differential gene expression associated with biological processes and developmental programs is regulated at many levels by various *trans*-acting factors. Transcription factors and miRNAs are two major classes of primary regulators for differential gene expression in eukaryotic organisms (Chen and Rajewsky, 2007). Transcription factors activate or repress transcription initiation, the first step of gene expression, by binding directly to promoters or enhancers in a sequence-specific manner. Elucidating the structure and function of transcription factors is one of the central schemes of molecular biology. Genetic studies have revealed that mutations in transcription factor genes often cause cellular or developmental defects. The availability of genome sequences has significantly improved the annotation of transcription factor genes and facilitated comparative analysis in multiple species (Riechmann *et al.*, 2000; Qu and Zhu, 2006).

In comparison, efforts to decipher miRNA pathways are more recent. Although mature miRNAs are only 20–24 nucleotides long, they are processed from much longer primary transcripts known as pri-miRNAs, via stem–loop-structured intermediates called pre-miRNAs (Bartel and Chen, 2004; Chen, 2005; Jones-Rhoades *et al.*, 2006; Voinnet, 2009). In higher plants, processing of pri-miRNAs and pre-miRNAs is performed in the nucleus, mainly by endonuclease DCL1. Two complementary short RNA molecules are generated as a consequence of these processing events. The duplex is then transported to the cytoplasm, where only the miRNA is integrated into the RNA-induced silencing complex (RISC) (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003). After integration into the RISC, miRNAs interact with their binding sites within target genes, and direct cleavage (German *et al.*, 2008) or translational repression (Brodersen *et al.*, 2008) of the target transcripts.

Following the initial discovery of miRNAs in the worm *Caenorhabditis elegans* (Lee *et al.*, 1993; Wightman *et al.*, 1993), much attention has focused on the identification of miRNAs and their target genes in various model organisms. Similar to transcription factors, miRNAs act *in trans* and bind their targets through short sequence motifs. Also similar to transcription factors, a single miRNA can regulate multiple target genes, and multiple miRNAs may act combinatorially to regulate the same gene (Chen and Rajewsky, 2007). This, together with the large number of miRNA genes, indicates that miRNAs affect a substantial proportion of the transcriptome. Indeed, dozens of miRNA–target gene pairs have been identified or proposed in Arabidopsis (Jones-Rhoades *et al.*, 2006; Fahlgren *et al.*, 2007; Alves *et al.*, 2009). It is now well-established that these gene circuits are crucial for many plant development processes and responses to environmental challenges (Jones-Rhoades *et al.*, 2006; García and Frampton, 2008; Voinnet, 2009).

Current knowledge indicates an intricate connection between miRNAs and transcription factors in the context of gene regulatory networks. On one hand, studies on miRNA targeting have revealed that genes encoding transcription factors are the most abundant miRNA target genes in both animals and plants (Jones-Rhoades *et al.*, 2006; Shalgi *et al.*, 2007). On the other hand, miRNAs are known to be transcribed by RNA polymerase II (Lee *et al.*, 2004), which indicates that miRNA transcription is subject to similar control mechanisms as protein-coding genes. Several bioinformatic studies have examined the global relationships between transcription factors and miRNAs (Shalgi *et al.*, 2007; Tsang *et al.*, 2007; Yu *et al.*, 2008; Re *et al.*, 2009). In *C. elegans*, a yeast-one-hybrid based transcription factor–miRNA network was experimentally mapped (Martinez *et al.*, 2008). It is clear from these studies there are indeed recurring interaction patterns between transcription factors and miRNAs, suggesting that combination of transcriptional and post-transcriptional mechanisms is an important design principle of gene regulatory networks (Yeger-Lotem *et al.*, 2004; Martinez *et al.*, 2008).

Genome-enabled high-throughput analysis of immunoprecipitated chromatin (ChIP-chip or ChIP-Seq) has been used for global discovery of transcription factor binding sites within protein coding genes (Cawley *et al.*, 2004; Sikder and Kodadek, 2005), but such efforts have generally not been extended to miRNA genes. We are interested in HY5, which acts downstream of multiple families of photoreceptors and promotes plant photomorphogenesis (Koornneef *et al.*, 1980; Oyama *et al.*, 1997; Ang *et al.*, 1998; Chattopadhyay *et al.*, 1998). Mutations in *HY5* cause aberrant light-mediated phenotypes in Arabidopsis, including an elongated hypocotyl, reduced chlorophyll/anthocyanin accumulation, and reduced chloroplast development in greening hypocotyls (Oyama *et al.*, 1997; Holm *et al.*, 2002). It is believed that HY5 plays a central role in the coordination of light signaling and appropriate gene expression (Quail, 2002; Sullivan and Deng, 2003). Furthermore, in the absence of HY5, Arabidopsis plants show altered root morphology and defects in hormonal and stimulus responses (Oyama *et al.*, 1997; Cluis *et al.*, 2004; Vandenbussche *et al.*, 2007), suggesting that HY5 could act as a signal transducer that links various signaling pathways to coordinate development. Therefore, elucidating the genome-wide binding pattern of HY5 in relation to various stimuli is critical to complete understanding of the regulatory gene network mediated by this important transcription factor.

In the present study, we performed genome-wide analysis of HY5 binding sites under both continuous white light (WL) and during light-to-dark transition (dark transition) using the ChIP-chip approach. We identified thousands of loci that are bound by HY5, including eight miRNA genes. Coupled with gene expression profiling using RNA-sequencing technology, our results indicate HY5 binding significantly affects the expression of over 1100 genes. In addition, HY5 indirectly regulates many other genes through sub-networks mediated by other regulators, such as miRNAs. Combining the regulatory branches mediated by HY5 and HY5-regulated miRNAs revealed several composite feed-forward loops (FFLs) for controlling specific target genes. Together, our analyses provide an example of the elucidation of gene networks that involve both transcriptional and post-transcriptional regulation in plants.

RESULTS

Profiling genome-wide HY5 binding sites in Arabidopsis

To identify *in vivo* HY5 binding sites at the genome scale in Arabidopsis, we performed a series of ChIP-chip assays (Figure 1a). The current study has several major improvements compared with a previous report (Lee *et al.*, 2007). We used an affinity-purified antibody against endogenous HY5 for ChIP instead of an antibody against the epitope tag fused to HY5, to reduce the risk of experimental artifacts associated with expressing an exogenous tagged protein. Using the HY5 antibody for ChIP, the proximal promoters of three well-studied *HY5* target genes, *CAB1*, *Rbcs1A* and *CHS*, were found to be enriched more than 400-fold over that of the control gene At4g26900, which is not bound by HY5 in wild-type plants (Figure S1a) but not in the loss-of-function *hy5-ks50* mutant (Figure S1b). This result indicates that ChIP using the HY5 antibody is highly specific, and should allow the identification of even weak HY5 binding sites.

Based on a small set of light-regulated genes, previous studies concluded that HY5 binding activity was not affected during light-to-dark transition (Lee *et al.*, 2007). We performed a genome-wide comparison of HY5 binding between plants grown under continuous WL and in dark transition to identify regulated changes in light-responsive genes. To allow sufficient time for the manifestation of gene activity
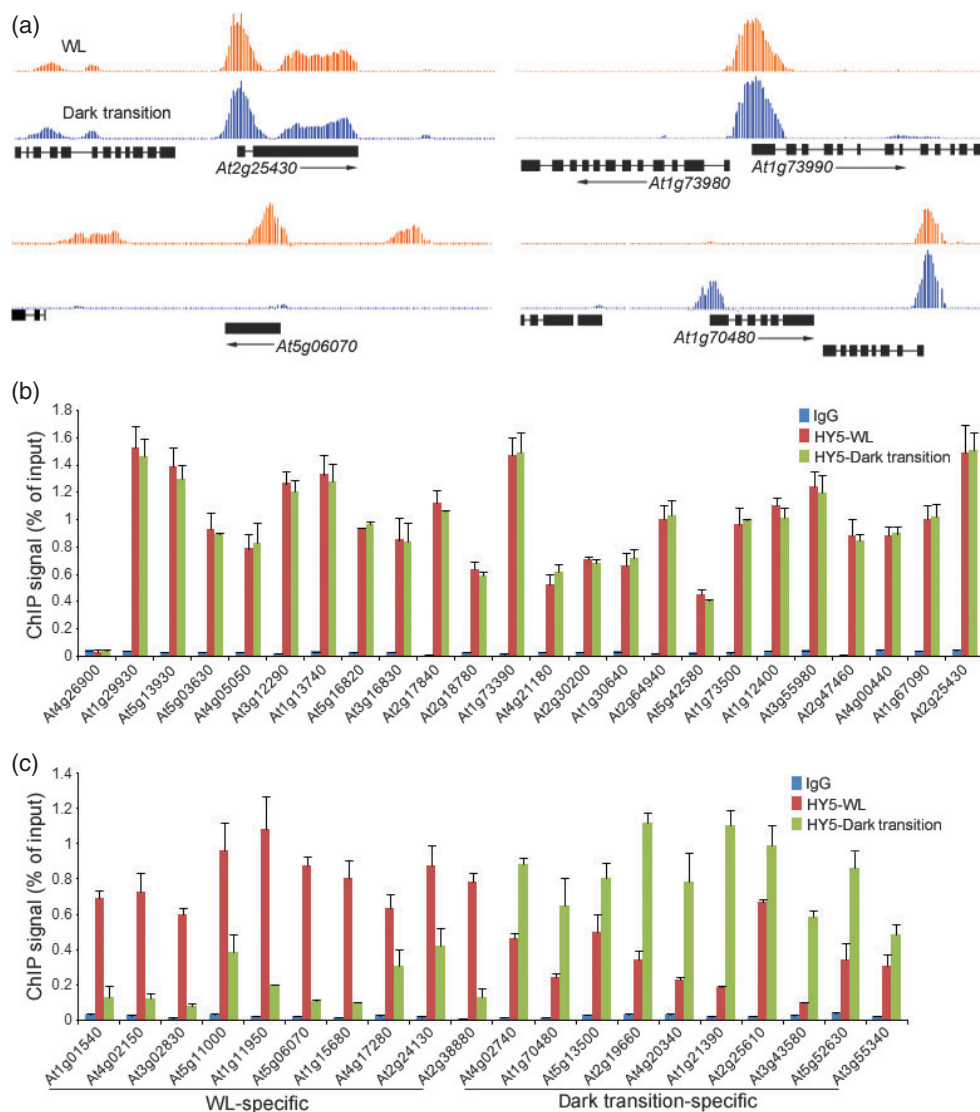
Figure 1. Analysis of HY5 binding loci using an Arabidopsis genome tiling array.

(a) Typical HY5 binding loci under WL and dark transition. The loci are visualized using the Affymetrix integrated genome browser (http://www.affymetrix.com/partners_programs/programs/developer/tools/download_igb.affx). Each vertical bar represents the signal intensity of an individual oligonucleotide probe anchored to genomic location. The structure, chromosomal location and transcriptional direction of the nearby genes are also shown.

(b) ChIP-qPCR confirmation of HY5 binding loci identified under both WL and dark transition. A negative control (At4g26900) and two positive controls [At1g29930 (CAB1) and At5g13930 (CHS)] were included.

(c) ChIP-qPCR confirmation of HY5 binding loci specifically detected under WL or dark transition. According to the results of the tiling array assay, the ten genes on the left showed specific enrichment of HY5 binding under WL, while the ten genes on the right showed specific enrichment of HY5 binding under dark transition.

changes, and to minimize the reduction in HY5 protein level, the experiment was performed after 8 hours of dark treatment (Figure S1c). The ChIP experiment was performed using wild-type plants under WL or dark transition as well as *hy5* mutant (*hy5-ks50*) plants grown under WL as a control.

For global identification of HY5 binding sites, we used the Affymetrix 1.0R Arabidopsis genome tiling array, which covers approximately 97% of the genome. This array has 35 bp resolution (Zhang *et al.*, 2006), which 14 times higher than that used in the previous study (Lee *et al.*, 2007). After

subtraction of the non-specific binding activity observed in the *hy5* background, 9852 high-confidence HY5-bound regions were identified under WL, with a total length of approximately 6.3 Mb, representing approximately 5.3% of the sequenced nuclear genome. Under dark transition, 9355 HY5-bound loci were identified, covering approximately 5.5 Mb and representing approximately 4.6% of the genome. To validate these array results, we performed a series of quantitative PCR analyses after the ChIP assay (ChIP-qPCR) on randomly selected HY5-occupied regions.

As shown in Figure 1(b), HY5 binding was confirmed by ChIP-qPCR for 21 of the 23 tested loci (91.3%) to which HY5 binds under both WL and dark transition. Of 13 selected regions that were found to be bound by HY5 specifically under WL, ten (76.9%) showed strong and reproducible signals in ChIP-qPCR analysis (Figure 1c). Of 15 regions specific to dark transition, ten (66.7%) were validated to be bound by HY5 specifically under dark transition (Figure 1c). These results indicate that HY5-occupied regions identified under both WL and dark transition have a lower false-positive discovery rate (8.7%).

Characterization of HY5 binding sites

HY5 binding sites were found in all genome components, including genic regions, transposable elements and inter-genic regions, although they were significantly depleted in the latter two groups (Figure 2a). To further assess the distribution of HY5 binding sites within the genic regions, we compared the position of HY5 binding relative to the transcription start site (TSS) of the nearest gene. A large proportion of the HY5 binding sites (39% in WL and 35% in dark transition) were located within the regions 1 kb upstream of the TSSs. When compared with random genomic loci, there is statistical enrichment for HY5 binding in these regions, as

well as the 5' UTRs, but not the coding regions or the 3' UTRs (Figure 2a). The pronounced clustering of HY5 binding around TSSs is clear when the frequency of binding sites against the distance from the TSS is plotted (Figure 2b). This analysis indicates that HY5 binding is centered near the TSS, with a notable bias toward the 5' UTR (Figure 2b). The preferential positioning of HY5 binding in these regions is consistent with a role for HY5 in classical transcriptional regulation of gene expression.

Several binding motifs, including G-box, C-box, hybrid CG- or CA-boxes, Z-box (Yadav *et al.*, 2002) and hybrid CT-box (Shin *et al.*, 2007), have been identified as light-responsive elements. It has been shown that HY5 specially recognizes these binding motifs (Song *et al.*, 2008). We found that these motifs appear at a higher frequency in the identified HY5 binding regions than in the whole genome (Figure 2c). We systematically tested the sequences surrounding the HY5-recognized motifs for enrichment of other known DNA motifs using the Arabidopsis *cis*-regulatory element database (Molina and Grotewold, 2005). This analysis revealed that a number of *cis*-elements for other transcription factors are enriched in HY5-bound regions (Figure 2d). Binding sites for 19 transcription factor families were significantly over-represented among *HY5* target
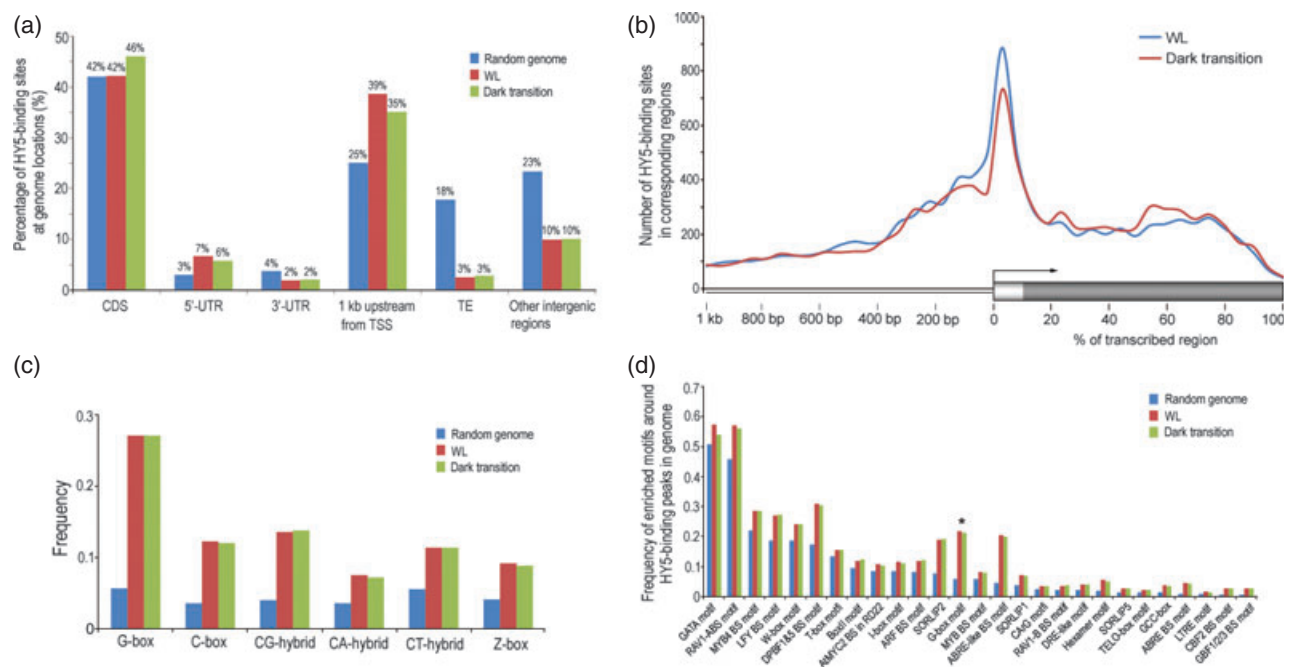
Figure 2. Distribution of HY5-binding loci within the Arabidopsis genome.
(a) Localization of HY5 binding sites (red or green) or randomly selected genomic regions (blue) in relation to annotated genome components. Percentages in the coding region (CDS), 5' and 3' UTRs, 1 kb region upstream of the TSS, transposable element (TE) and intergenic region are shown.
(b) Gene-level distribution of HY5 binding loci. Each gene (thick horizontal bar, shown as 2.5 kb, which is the mean length of Arabidopsis genes) was divided into 20 intervals, and the 1 kb region upstream of each gene (thin horizontal bar) was divided into 50 bp intervals. The number of genes that showed HY5 binding in each interval was plotted.
(c) Enrichment of known consensus HY5 recognition motifs within detected HY5 binding loci compared to the whole genome.
(d) Enriched recognition motifs of transcription factors in genomic regions flanking the HY5 binding loci. The G-box (indicated by an asterisk) is included as a control.

genes (Table S1). These include transcription factor families with functions in the control of hormone responses (AP2-EREPBP, BES1, GRAS and MYB-related), organ growth (LIM and TCP), environmental responses (HSF and WRKY), chromatin remodeling (HMG, PcG and PHD), and developmental processes (C2C2-Dof and C2H2). Furthermore, we grouped *HY5*-bound genes based on the occurrence of specific motifs within the HY5 binding sites (Figure 2d), and analyzed the genes using gene ontology (GO) analysis. Three significantly enriched GO categories were found for the co-binding factors: developmental, morphogenesis and hormone response pathways in which HY5 has been reported to play a role (Table S2). Together these results suggest that HY5 may regulate many targets through gene networks involving other transcriptional regulators.

Identification and characterization of *HY5*-regulated target genes

Mapping of HY5-bound regions relative to the annotated genes in the Arabidopsis genome indicated close proximity of a binding site to 9606 and 9027 genes under WL and dark transition, respectively (Figure 3a). We therefore considered

these as putative *HY5* target genes. Of these, approximately 60% (6836) showed no difference in HY5 binding profile between WL and dark transition. However, 2770 (23%) and 2191 (17%) genes showed specific enrichment of HY5 binding under either WL or dark transition, respectively (Figure 3a). When the two datasets were combined, it was found that HY5 could potentially bind to 11 797 genes in total. This list is nearly three times larger than previously reported (Lee *et al.*, 2007), attesting to the completeness of the current dataset.

To study the impact of HY5 binding on target gene expression, we used RNA sequencing to examine *HY5*-mediated gene expression changes between wild-type and *hy5* plants. From this analysis, we identified 2752 differentially expressed genes, of which 1390 were up-regulated and 1362 were down-regulated by HY5, respectively (Figure 3b). We then compared the differentially expressed genes from the RNA sequencing analysis with the *HY5* target genes from ChIP-chip analysis. The comparison identified 1173 genes as *HY5*-regulated target genes that were bound by HY5 and exhibited HY5-dependent expression (Figure 3b). Among these, 516 and 657 are negatively and positively regulated
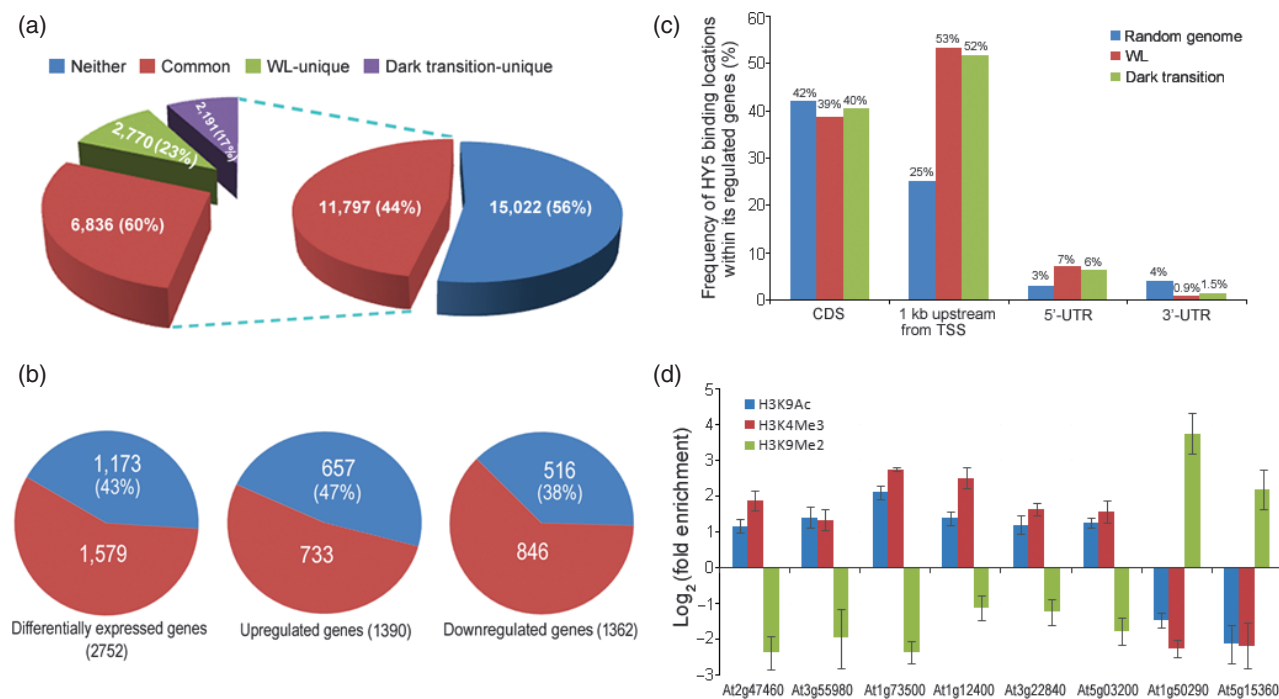
Figure 3. HY5 binding results in both activation and repression of target genes.
(a) Combining the WL and dark transition data results in a total of 11 797 *HY5*-bound genes in Arabidopsis. This dataset comprises 6836 genes bound by HY5 under both WL and dark transition, and 2770 and 2119 uniquely detected under WL and dark transition, respectively.
(b) Relationship between *HY5*-bound genes and *HY5*-dependent genes revealed by mRNA sequencing analysis. The blue areas represent the number of differentially expressed genes that also have HY5 binding sites and their percentage compared with total differentially expressed genes.
(c) Localization of HY5 binding sites (red or green) or randomly selected genomic regions (blue) in relation to annotated genome components. Percentages in the coding region (CDS), 1 kb region upstream of the TSS, and the 5′ and 3′ UTRs of *HY5*-regulated target genes are shown.
(d) Histone modifications at the HY5 binding loci of *HY5*-regulated target genes. ChIP was performed using antibodies against acetylated H3K9 (H3K9Ac), trimethyl H3K4 (H3K4Me3) and dimethyl H3K9 (H3K9Me2). Identical amounts of precipitated or input DNA were used for quantitative PCR. The first six genes are positively regulated by HY5 and the last two are negatively regulated by HY5.

by HY5, respectively (Figure 3b). An even larger proportion of these genes (53% under WL and 52% under dark transition) have the HY5 binding site located within the region 1 kb upstream of the TSS (compare Figure 3c with Figure 2a). Together, these results indicate that HY5 binding directly affects the expression of a large number of genes.

To investigate whether histone modification is part of the mechanism for *HY5*-mediated gene regulation, we examined permissive and inhibitory histone H3 modifications at HY5-bound regions for several target genes in both wild-type and *hy5* seedlings. Comparing the pattern between wild-type and *hy5* seedlings revealed that the level of permissive H3 histone modifications (acetylated H3K9 and tri-methylated H3K4) was increased, while that of the inhibitory histone H3 modification (di-methylated H3K9) was substantially decreased in HY5-bound regions for genes positively regulated by HY5 (Figure 3d). In contrast, the HY5 binding regions of the negatively regulated genes, At1g50290 and At5g15360, showed completely opposite histone modifications (Figure 3d). Thus, transcriptional activation and repression mediated by HY5 are associated with specific histone modifications at HY5 binding sites.

GO analysis showed that transcription factors are enriched among the *HY5*-regulated target genes (Figure S2). The 115 *HY5*-regulated transcription factor genes are listed in Table S3. These results suggest that *HY5*-regulated transcription factors could play important roles in the expression of HY5 targets and thus HY5-dependent developmental programs. For instance, genes related to the auxin, cytokinin, ethylene and jasmonic acid pathways were highly enriched in the *HY5*-regulated target genes (Figure S3). Consistently, various hormone-related transcription factors are also among the enriched genes in HY5 targets (Table S4). These results suggest that HY5 coordinates various hormonal responses during seedling development in response to light. Our data further support the notion that *HY5*-dependent gene regulation may be achieved through transcriptional cascades initiated by transcription factors that are among the *HY5*-regulated target genes.

HY5 directly regulates several miRNA genes

We found that approximately 10% of HY5-binding loci are within the intergenic regions between protein-coding genes (Figure 2a). This observation prompted us to investigate whether HY5 can regulate the expression of miRNA genes. We identified HY5 binding loci that are immediately upstream of the genomic region corresponding to an annotated miRNA hairpin structure. This analysis produced eight candidate miRNAs: miR156d, miR172b, miR402, miR408, miR775, miR858, miR869 and miR1888 (Figure 4a). In all cases except miR402, the HY5 binding peaks are approximately 300–800 bp upstream of the hairpin structure, and contain at least one copy of consensus HY5 binding motifs such as the G-box, GC-box and CT-box (Figures 2c

and 4b). To further validate the binding specificity, we performed ChIP-qPCR analysis of the putative HY5 binding loci in both the wild-type and *hy5* background. This experiment revealed significant ChIP products in wild-type either under WL or dark transition, but not in the *hy5* mutant (Figure 4c), confirming that the binding activity is indeed specific to HY5. Based on documented miRNA gene structures (Xie *et al.*, 2005; Megraw *et al.*, 2006; Zhou *et al.*, 2007), the physical distance between HY5 binding loci and the miRNA hairpin structure suggests that HY5 binds to the core promoter of these miRNA genes. Together, these results indicate that HY5 binds directly to the putative promoter region of multiple miRNA genes.

To test whether HY5 binding is required for proper expression of the miRNAs, we performed Northern blotting to compare their expression levels in wild-type and *hy5* seedlings. This analysis revealed that five of the miRNAs showed significantly decreased expression in the *hy5* mutant (Figure 4d), although we failed to detect expression of miR172b, miR869 or miR1888 in either wild-type or *hy5* seedlings (data not shown). Together, these results indicate that HY5 is primarily an activator for the miRNA genes that it regulates.

HY5 indirectly regulates miRNA target genes

The eight miRNAs regulated by HY5 were collectively predicted to target 108 genes using the program miRU (Zhang, 2005). Eighteen of these have been previously validated using a modified 5′ RACE assay that detects miRNA-guided target cleavage. Because miRNAs typically repress expression of their target genes, the true targets of the miRNAs regulated by HY5 should be increased in the *hy5* mutant. Based on this reasoning, we selected 24 predicted miRNA target genes (including seven validated genes), and performed quantitative RT-PCR to examine their expression level in both wild-type and *hy5* plants. As shown in Figure 5, expression of 21 validated or predicted targets significantly increased in *hy5*, which correlates with the decreased level of miRNAs in the mutant (Figure 4d). Thus, these potential miRNA target genes could be repressed by HY5 through activation of specific miRNAs.

As a first step to recapitulate the complex HY5 regulatory circuits involving miRNAs, we attempted to identify HY5-regulated composite feed-forward loops (FFLs) that are the most abundant motifs in transcription networks (Milo *et al.*, 2002; Shen-Orr *et al.*, 2002). We did this by integrating data sets involving HY5 binding loci, *HY5*-dependent gene expression and target genes of HY5-regulated miRNAs. This analysis identified a total of eight genes that are targeted by both HY5 and an miRNA that itself is regulated by HY5 (Figure 6a). Topologically, these genes could be regulated by an FFL that consists of HY5 and an miRNA. Depending on whether HY5 activates or represses the target, the loops can be divided into coherent and incoherent types with
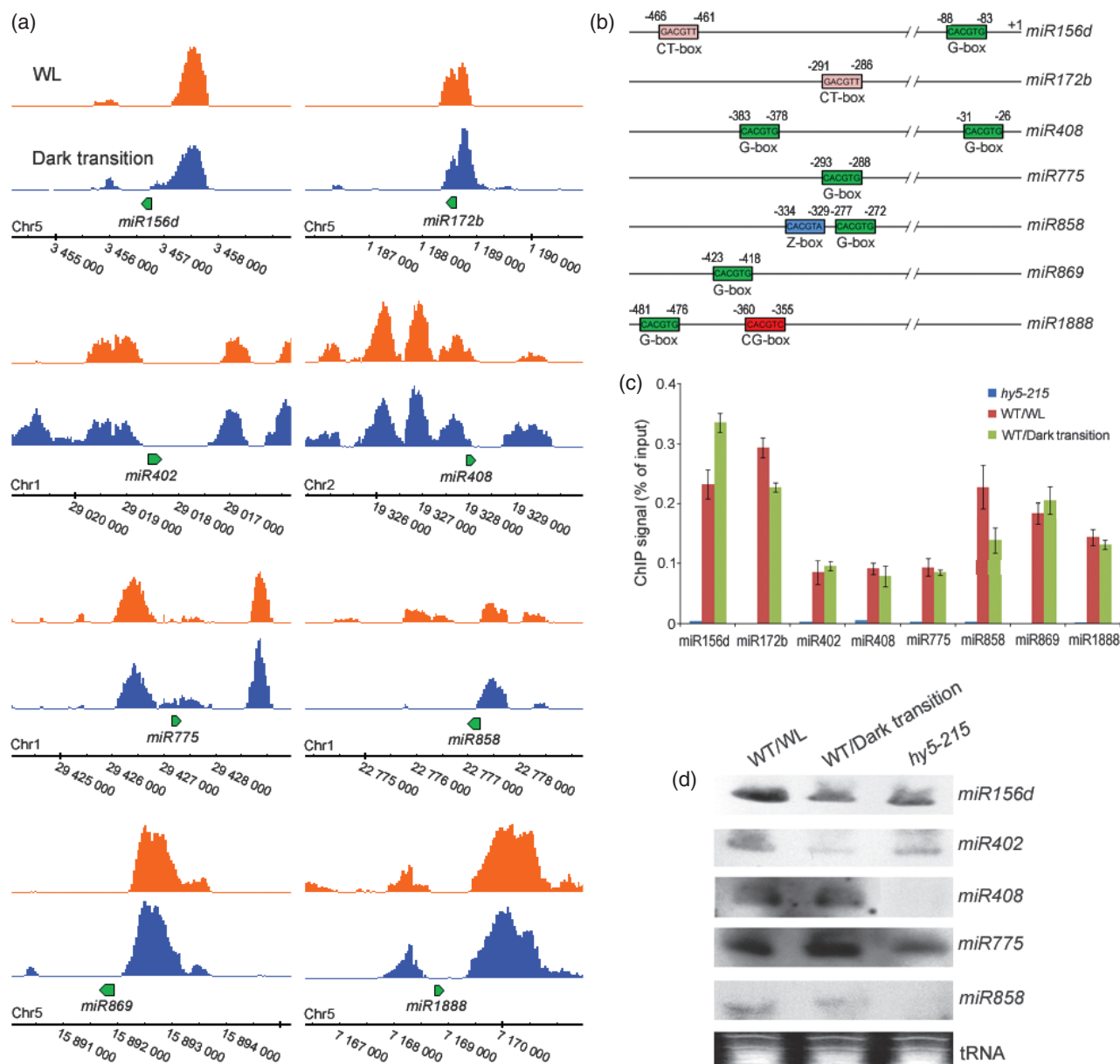
Figure 4. HY5 directly regulates the expression of miRNA genes.
(a) HY5 binding at the upstream region of miRNA genes was visualized using the Affymetrix integrated genome browser. Horizontal green arrows indicate the miRNA hairpin structures and transcriptional direction.
(b) Sequence analysis for conserved HY5 recognition motifs within the 500 bp regions upstream of the miRNA hairpin structures.
(c) Confirmation of HY5 binding to the miRNA promoter regions by ChIP coupled with quantitative PCR.
(d) Northern blot analysis of miRNA expression in wild-type seedlings grown under WL and dark transition and in *hy5-215* under WL.

distinctive dynamic properties (Figure 6b). Indeed, expression dynamics typical of genes regulated by FFLs can be observed under dark transition based on available expression profiles. For example, At4g29430 is co-regulated by HY5 and miR172b in an incoherent FFL. Its expression pattern under dark transition is illustrated in Figure 6(c). Interestingly, the pulse-like expression dynamics of this gene after dark treatment correlate with a significant decrease in HY5-binding activity (Figure 6d).

Finally, we investigated whether HY5-regulated miRNA expression is relevant to its function. To this end, we generated transgenic plants expressing the stem–loop precursor of miR408 driven by the enhanced CaMV 35S promoter, which accumulated mature miR408 at much higher levels than wild-type plants (Figure 7a). Although the accumulated levels of pigments including both chlorophyll and anthocyanin are lower in the *hy5* plants than in the wild-type, the transgenic plants showed a significant
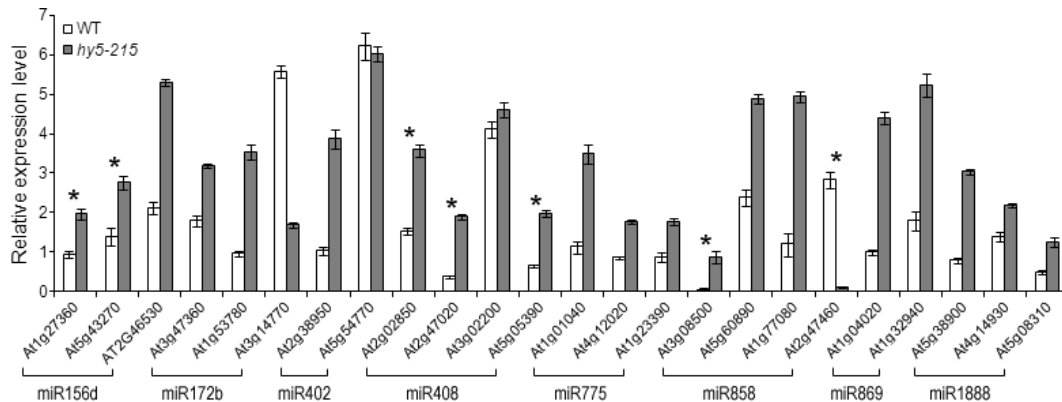
Figure 5. Analysis of transcript abundance of the predicted miRNA target genes.
Expression levels of 24 predicted miRNA target genes were determined by quantitative RT-PCR and normalized to those of the control *Actin7* in both wild-type and *hy5-215* seedlings. The asterisk indicates miRNA target genes validated in the literature.
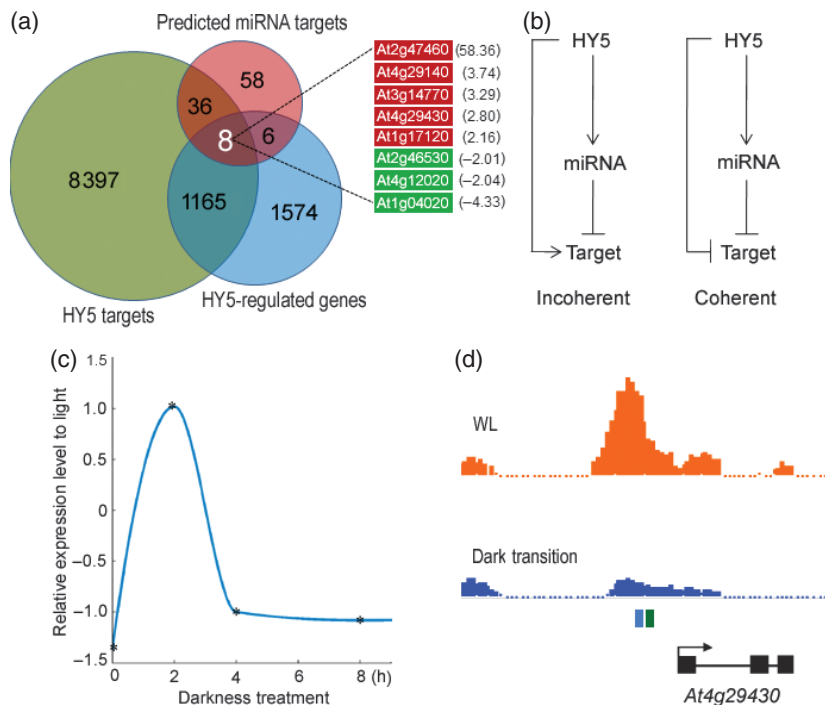
Figure 6. HY5 and its target miRNAs form composite FFLs.
(a) Venn diagram showing the overlap among the *HY5* target genes, *HY5*-regulated genes and predicted targets of the miRNAs that are regulated by HY5. *HY5*-regulated target genes that are also targeted by HY5-regulated miRNAs are listed, together with the expression fold-change between wild-type and *hy5*. Red indicates genes that are positively regulated by HY5. Green indicates genes that are negatively regulated by HY5.
(b) Structure of the two types of three-gene FFLs involving both HY5 and HY5-regulated miRNAs. The loops can be divided into coherent and incoherent types, depending on whether HY5 activates or represses the target.
(c) The pulse-like expression dynamics of At4g29430 in response to darkness. At4g29430 is co-regulated by HY5 and miR172b in an incoherent FFL. Expression data of At4g29430 (labeled with star) was obtained from the Arabidopsis Information Resource (deposit number ExpressionSet_206). The fitted curve was interpolated using MATLAB (http://www.mathworks.com/products/matlab).
(d) HY5-binding profiles at the promoter region of At4g29430 under WL and dark transition. Two putative motifs were found at the upstream HY5-bound region. Green bar represents G-box motif, and light blue bar represents Z-box motif.

increase in chlorophyll and anthocyanin content compared to wild-type (Figure 7b–d). These results indicate that over-expression of HY5-regulated miR408 leads to phenotypic changes opposite to those caused by the *hy5* mutation.

DISCUSSION

Proper gene expression associated with biological processes and developmental programs is regulated in the context of gene networks. Transcriptional gene networks
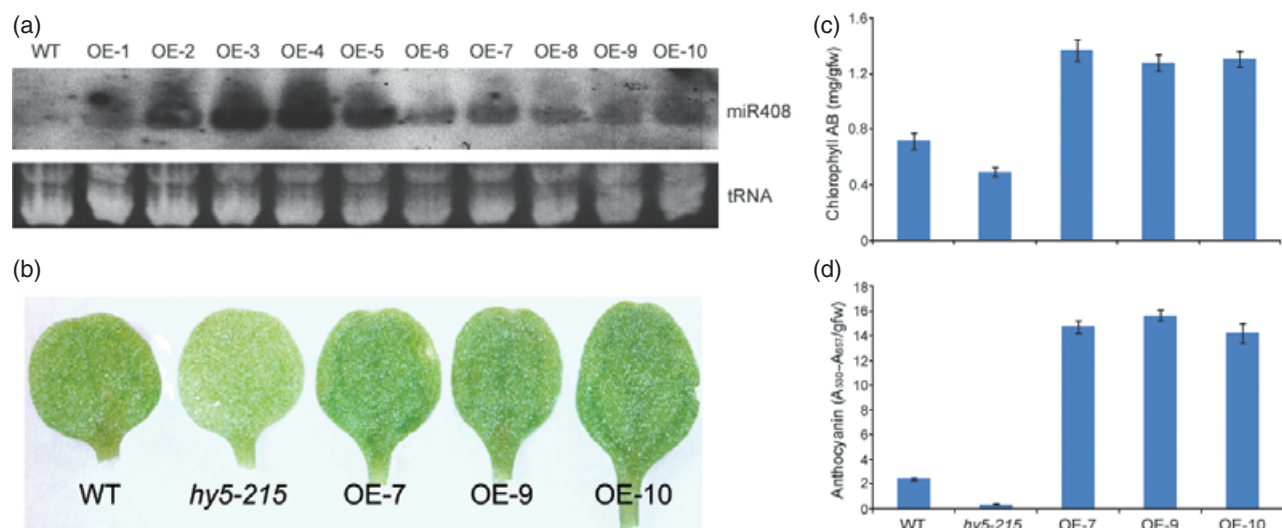
Figure 7. Over-expression of miR408.
(a) RNA gel-blot detection of miR408 in wild-type (WT) and ten independent transgenic lines.
(b) Cotyledons of 7-day-old seedlings of wild-type, *hy5-215* and three independent 35S::miR408 transgenic plants (lines OE-7, 9 and 10).
(c,d) Chlorophyll *a* and *b* and anthocyanin content in 7-day-old seedlings of wild-type, *hy5-215* and 35S::miR408 transgenic plants (lines OE-7, 9 and 10) grown under WL. Error bars represent standard deviation (*n* = 5).

mediated by transcription factors have been extensively mapped in eukaryotes from yeast (Harbison *et al.*, 2004), *Caenorhabditis elegans* (Deplancke *et al.*, 2006; Vermeirssen *et al.*, 2007) and *Drosophila melanogaster* (Sandmann *et al.*, 2007) to mammals (Boyer *et al.*, 2005). Although studies on these mapped networks are still incomplete, they have already generated insights into overall network architecture and design principles. In particular, these studies demonstrated the importance of basic structural units referred to as network motifs, which occur more often than expected on the basis of random connections (Milo *et al.*, 2002; Shen-Orr *et al.*, 2002). Conserved network motifs in diverse species thus represent evolutionarily successful mechanisms for precise and timely gene regulation.

Because of the importance of miRNAs in gene regulation, mapping and analysis of transcription factor–miRNA interactions should provide much-needed insights into integrated gene networks involving both transcriptional and post-transcriptional regulation. We sought to identify the network mediated by HY5 using genome-wide ChIP-chip analysis (Figure 1). Consistent with the important roles of HY5 in many plant processes, we identified HY5 binding regions for over 9000 genes under both WL and dark transition (Figure 2). We also performed RNA sequencing analysis, which showed that 2752 genes were either up- or down-regulated in null *hy5* mutant plants compared to wild-type (Figure 3), suggesting that HY5 acts as both a transcriptional activator and repressor. Combining the ChIP-chip and RNA sequencing data revealed that 1173 genes were both bound and regulated by HY5 (Figure 3), representing the proportion of gene networks consisting of *HY5*-regulated target genes.

The *HY5* gene was shown previously to be a positive regulator of photomorphogenesis of young seedlings (Koornneef *et al.*, 1980; Ang and Deng, 1994; Chory *et al.*, 1994). Detailed characterization of the *hy5* mutants also revealed a myriad of phenotypic defects related to the development of hypocotyls, roots and chloroplast (Oyama *et al.*, 1997). Many of these defects were attributed to the role of HY5 in the elongation and proliferation of specific cells (Oyama *et al.*, 1997). The *HY5*-regulated target genes identified in our work include a few that encode proteins related to cell elongation, proliferation and chloroplast development (Table S5). For example, we found that HY5 directly represses the expression of two transcription factor genes (*ARF2* and *ATHB-2*) that are involved in regulation of cell proliferation. In addition, several genes related to cell proliferation were also found to be repressed by HY5 either directly or indirectly. Not only are these results consistent with the notion that HY5 probably controls the initiation of lateral roots by repressing cell proliferation in wild-type plants (Oyama *et al.*, 1997), but provide excellent candidate genes for future analyses aimed at fully elucidating HY5-regulated processes.

Combining the ChIP-chip and RNA sequencing data also revealed that HY5 binding at most genes is not sufficient to cause changes in their expression level, and that over half of the genes for which proper expression is dependent on HY5 are not actually bound by HY5 (Figure 3). These results indicate that HY5 probably co-regulates many target genes with other transcriptional regulators through integrated subprograms. We showed that HY5 targets eight miRNA genes, and positively regulates their expression by binding directly to the upstream regulatory region (Figure 4). The validated and predicted target genes of these miRNAs are generally

negatively regulated by HY5 (Figure 5). These results suggest that some of the sub-programs in the *HY5* gene network are mediated by miRNAs. In support of this view, we found that transgenic plants over-expressing HY5-regulated miR408 accumulated higher levels of pigments compared with wild-type plants (Figure 7). This phenotype is the exact opposite to that caused by the *hy5* mutation (Figure 7). Collectively, our results indicate that the miRNA branches in the *HY5* gene networks may control specific development processes.

The miRNA genes have been shown to be transcribed by RNA polymerase II (Lee *et al.*, 2004), which indicates that miRNA transcription is subject to similar control mechanisms as protein-coding genes. In Arabidopsis, the promoter regions of a large number of miRNAs have either been mapped by 5′ RACE analysis of the pri-miRNA (Xie *et al.*, 2005) or computationally identified (Zhou *et al.*, 2007). These promoters have been searched successfully for DNA elements that bind transcription factors using position weight matrices of 99 known binding motifs (Megraw *et al.*, 2006). By mapping and validating HY5 occupation sites at miRNA genes (Figure 4), our results reinforce the view that miRNA genes are regulated similarly to protein-coding genes at the transcription level. Furthermore, our results revealed the integrative nature of gene networks, consisting of both transcriptional and post-transcriptional regulation, which certainly warrants further analyses.

In transcription networks, FFLs are abundant network motifs, and provide a mechanism to ensure proper transcriptional responses to environmental or developmental signals and protect the cells against environmental fluctuations (Milo *et al.*, 2002; Shen-Orr *et al.*, 2002; Cohen and Rosner, 2009). To recapitulate the complex regulatory circuits mediated by both HY5 and HY5-regulated miRNAs, we identified eight genes in HY5-regulated composite FFLs which involved miRNAs that were themselves regulated by HY5 (Figure 6). These results again suggest the importance of mapping regulatory gene networks in understanding biological pathways. Taken together, the current study provides a comprehensive profile of HY5 binding in response to a changing light environment. Our results also provide an example of elucidation of gene circuits in plants that involve both transcriptional and post-transcriptional regulations. Similar studies for other transcription factors should provide more mechanistic insight into the coordinated and adaptable control of gene batteries that underpin plant development and responses to environmental challenges.

EXPERIMENTAL PROCEDURES

Plant materials and growth conditions

The wild-type plants used in this work were *Arabidopsis thaliana* ecotype Col-0. Mutants defective in HY5 were either *hy5-ks50* or *hy5-215* as described previously (Oyama *et al.*, 1997; Lee *et al.*, 2007). Unless stated otherwise, the seeds were placed on MS medium (Sigma-Aldrich, http://www.sigmaaldrich.com/) and incubated at 4°C in the dark for 4 days, after which they were exposed to continuous white light (WL, 170 μmol m^{-2} sec^{-1}) at 22°C for 4 days. The plants were then incubated either under continued light or transferred to the dark, and harvested at the indicated times thereafter.

ChIP-chip assay

Chromatin isolation was performed using 4-day-old whole seedlings grown under WL or after 8 h in darkness as described by Bowler *et al.* (2004). The resuspended chromatin pellet was sonicated at 4°C using a Diagenode Bioruptor (http://www.diagenode.com/) set at high intensity for 10 min (intervals of 30 s on, 30 s off). The DNA was sheared to a mean size of approximately 500 bp. Chromatin was immunoprecipitated, washed, reverse cross-linked, amplified and hybridized as described in the Affymetrix Chromatin Immunoprecipitation Assay Protocol Revision 3 (http://www.affymetrix.com/). A polyclonal HY5 antibody purified from anti-HY5 rabbit IgG using Sepharose 4B beads (GE Healthcare, http://www.gelifesciences.com) coupled with purified recombinant HY5 protein was used. 'Flow-through' IgG lacking the HY5 antibody was used as a negative control. An aliquot of untreated sonicated chromatin was reverse cross-linked, and used as a total input DNA control for ChIP-PCR experiments. Four biological replicates for each condition were hybridized to tiling arrays, resulting in a total of 16 chips.

Raw microarray data were processed using three software packages [TileMap (Ji and Wong, 2005), Cisgenome (http://www.biostat.jhsph.edu/~hji/cisgenome) and Affymetrix tiling array software (Ghosh *et al.*, 2006)] to identify peak location and calculate peak intensity. Only those peaks that were identified by all three methods were considered as true binding sites. We used the following algorithm to assign the binding sites to genes. First, genes overlapping with HY5 binding sites were selected, and the one with the shortest distance between the TSS and the HY5 binding site was assigned to it. For genes that did not overlap with a binding site, a gene was assigned when the binding site was located within 1000 bp upstream or 500 bp downstream of the gene. When more than one gene was assigned to the same binding site, the gene with the shortest distance between its transcription start or end site and the binding site was selected.

ChIP assay for histone modifications

ChIP was performed as described previously (Gendrel *et al.*, 2005). Briefly, 4-day-old seedlings grown under white light or 8 h darkness were harvested and fixed in 1% formaldehyde for 15 min in a vacuum. The reaction was terminated by adding glycine to a final concentration of 0.125 M, and incubated in a vacuum for another 5 min. After chromatin isolation, histone-bound DNA was fragmented into a mean size of approximately 500 bp by sonication. After removing the cellular debris, a small aliquot of each sample was stored for reverse cross-linking as a direct total input DNA control. Other chromatin solutions were diluted 10-fold, and precleared by incubating with protein A–Sepharose Fast Flow beads (Sigma) for 1 h at 4°C. Immunoprecipitation was performed overnight at 4°C under rotation. Antibodies specific for various histone modifications (anti-H3K4me3, anti-H3K9ac and anti-H3K9me2) were obtained from Upstate Biotechnology (http://www.upstatebiotech.com). An equal amount of sample without antibody was used as a mock control. The samples were exposed to protein A beads for 1.5 h at 4°C to pull down protein–DNA complexes. After washing for 10 min each with low-salt wash buffer (150 mM NaCl,

0.1% SDS, 1% Triton X-100, 2 mM EDTA and 20 mM Tris/HCl, pH 8.0), high-salt wash buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA and 20 mM Tris/HCl, pH 8.0), LiCl wash buffer (0.25 M LiCl, 1% Nonidet P-40 (Sigma), 1% sodium deoxycholate, 1 mM EDTA and 10 mM Tris/HCl, pH 8.0) and TE buffer (10 mM Tris/HCl, pH 8.0, and 1 mM EDTA), the immunocomplexes were eluted from the protein A beads. The samples were reverse cross-linked by incubation at 65°C overnight, and then treated with proteinase K at 45°C for 1 h. DNA was extracted using phenol–chloroform, and precipitated using ethanol containing a one-tenth volume of 3 M NaOAc and 1 μg of glycogen.

Search for transcription factor binding sites and GO analysis

The significant HY5 binding peaks (false discovery rate <0.001) were extended by 250 bp on both sides to identify the flanking genomic regions according to the Arabidopsis reference genome (version 7; http://www.arabidopsis.org/). These regions were then searched for perfect matches to the consensus transcription factor binding motif sequences using a Perl script. The consensus motif sequences were obtained from the AGRIS and Transfac databases (Davuluri *et al.*, 2003; Wingender *et al.*, 1996). The sequences surrounding the centers of HY5 binding sites (500 bp around each binding peak) were analyzed using AtcisDB (a database of predicted whole-genome *cis*-binding motifs in Arabidopsis) compared to the 1000 bp regions flanking the centers of HY5 binding sites on either side. GO analysis was performed using BiNGO (Maere *et al.*, 2005).

Quantitative PCR

The relative abundance of regions of interest in immunoprecipitated DNA was measured by quantitative PCR using an ABI 7500 system and Power SYBR Green PCR master mix (Applied Biosystems, http://www.appliedbiosystems.com/). ChIP values were normalized as a percentage of the input DNA. The enrichment of specific regions in ChIP was expressed relative to the respective values of the control region (exon 2 of the *ACT7* gene). Primer sequences are listed in Table S6.

Total RNA extraction and transcript detection

Using the same seedling samples that were used for the ChIP assay, total RNA was extracted using TRIzol reagent (Invitrogen, http://www.invitrogen.com) and treated with RNase-free DNase I (Invitrogen). RNA was reverse transcribed using SuperScript II reverse transcriptase (Invitrogen). The resulting cDNA was analyzed by quantitative PCR. Primer sequences are listed in Table S6, and *Actin7* was used as an endogenous control to normalize the relative expression level. The data were analyzed using ABI 7500 system SDS software (Applied Biosystems). The results of at least three independently performed experiments were analyzed.

For miRNA blotting, 20 μg total RNA was separated on 12% denaturing polyacrylamide gels containing 7 M urea, and transferred to Hybond N+ nylon membrane (GE Healthcare, http://www.gehealthcare.com). Blots were hybridized overnight with respective miRNA-complementary oligonucleotides that had been end-labeled using Digoxigenin-11-2′,3′-dideoxy-uridine-5′-triphosphate (Roche, http://www.roche.com) at 42°C in moderate CG hybridization buffer (1% w/v BSA, 0.5 M Na$_2$HPO$_4$, 15% v/v formamide, 1 mM EDTA, 7% w/v SDS) after pre-hybridization for at least 2 h. Blots were washed three times at 42°C in 1× SSC and 0.1% SDS for 10 min each time. This was followed by incubation with digoxigenin washing buffer at room temperature for 10 min, and then incubation for 30 min in blocking solution (Roche). The blots were then incubated with anti-digoxigenin-AP antibody (Roche) diluted in 1× blocking solution for 30 min, and washed twice in 1×

washing buffer for 15 min each time. Blots were equilibrated in detection buffer (Roche) for 5 min, and incubated for 15 min with CDP-*Star* solution (Roche) before exposure to X-ray films.

RNA sequencing library construction and data processing

DNase I-treated total RNA was prepared from wild-type and *hy5-215* seedlings grown under WL. The mRNA fraction was extracted from total RNA using Dynabeads oligo(dT)$_{25}$ (Invitrogen) according to the manufacturer's instructions. First- and second-strand cDNA were generated using Superscript II and random hexamer primers. Double-stranded cDNA was fragmented by nebulization, and used for mRNA library construction according to the manufacturer's instructions (Illumina, http://www.illumina.com). The software Bowtie (Langmead *et al.*, 2009) was used to map the sequence reads to the Arabidopsis genome, and only uniquely mapped reads were used in subsequent analyses. Reads mapped to exonic regions of annotated gene models were normalized against the length of the transcript. To compare gene expression between wild-type and the *hy5-215* mutant, length-normalized read density was quantile-normalized. A twofold difference was used as the threshold for identification of differentially expressed genes.

Chlorophyll and anthocyanin measurements

Measurement of chlorophyll and anthocyanin contents was performed essentially as described previously (Chory *et al.*, 1989, 1991). Briefly, for chlorophyll, 7-day-old seedlings grown under WL were harvested, weighed, frozen in liquid nitrogen, and ground to a fine powder. Total chlorophyll was extracted into 80% acetone. Chlorophyll *a/b* contents were calculated using MacKinney's specific absorption coefficients, for which chlorophyll $a = 12.7(A_{663}) - 2.69(A_{645})$ and chlorophyll $b = 22.9(A_{645}) - 4.48(A_{663})$ (A indicates absorption at each wavelength). The total chlorophyll content is expressed as micrograms of chlorophyll per gram seedlings. For anthocyanin, 7-day-old seedlings were harvested, weighed, homogenized, and incubated overnight in 0.3 ml of 1% HCl in methanol at 4°C. After addition of 0.2 ml of water, anthocyanin was extracted using an equal volume of chloroform. The quantity of anthocyanins was determined by spectrophotometric measurement of the aqueous phase (A_{530}–A_{657}), and normalized to the total fresh weight of tissue used in each sample.

Generation of transgenic plants over-expressing miR408

To produce the 35S::miR408 construct, the genomic sequence containing pre-miR408 was PCR-amplified from Arabidopsis Col-0 genomic DNA using Pfusion DNA polymerase (New England Biolabs, http://www.neb.com). The primers used were 5′-GGACTCGAGGCCAATTTCAAAGGTTAGATTG-3′ and 5′- CACGGATCCATTGAGGATATGATGCAGGGAG-3′. PCR product was cloned into the 35S-pKANNIBAL vector (Li *et al.*, 2010) between the *Xho*I and *Bam*HI restriction sites, and sequenced. Plants were transformed using the standard floral dipping method (Clough and Bent, 1998). Transformants were selected on medium containing 20 mg L^{-1} BASTA (bioWORLD, http://www.bio-world.com).

Accession number

The original data have been deposited in the National Institutes of Health Gene Expression Omnibus database under accession number GSE24974.

tiling array hybridization. This work was supported by grants from the National Institutes of Health (GM047850), the Ministry of Science and Technology of China (2009DFB30030) and the Ministry of Agriculture of China (2009ZX08012-021B).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Association of HY5 with the proximal promoters of *CAB1*, *RbcS1A* and *CHS*, confirmation of Arabidopsis mutant *hy5-ks50* as a good negative control, and *HY5* expression levels after various durations of dark treatment.

Figure S2. GO analysis of *HY5*-regulated direct target genes.

Figure S3. *HY5*-regulated target genes related to specific hormonal pathways.

Table S1. Significantly over-represented transcription factor families in HY5 targets.

Table S2. Enriched TRANSFAC motifs around HY5 binding regions involved in a biological function.

Table S3. *HY5*-regulated target genes encoding transcription factors.

Table S4. HY5-regulated hormone-related transcription factor genes.

Table S5. HY5-regulated cell growth and chloroplast development genes.

Table S6. Oligonucleotide sequences for the primer pairs used in ChIP-qPCR or quantitative RT-PCR.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

Alves, L. Jr, Niemeier, S., Hauenschild, A., Rehmsmeier, M. and Merkle, T. (2009) Comprehensive prediction of novel microRNA targets in *Arabidopsis thaliana*. *Nucleic Acids Res.* **37**, 4010–4021.

Ang, L.H. and Deng, X.W. (1994) Regulatory hierarchy of photomorphogenic loci: allele-specific and light-dependent interaction between the HY5 and COP1 loci. *Plant Cell*, **6**, 613–628.

Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A. and Deng, X.W. (1998) Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. *Mol. Cell*, **1**, 213–222.

Bartel, D.P. and Chen, C.Z. (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* **5**, 396–400.

Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A.V., Tariq, M. and Paszkowski, J. (2004) Chromatin techniques for plant cells. *Plant J.* **39**, 776–789.

Boyer, L.A., Lee, T.I., Cole, M.F. et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, **122**, 947–956.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L. and Voinnet, O. (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, **320**, 1185–1190.

Cawley, S., Bekiranov, S., Ng, H.H. et al. (2004) Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell*, **116**, 499–509.

Chattopadhyay, S., Ang, L.H., Puente, P., Deng, X.W. and Wei, N. (1998) Arabidopsis bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. *Plant Cell*, **10**, 673–683.

Chen, X. (2005) MicroRNA biogenesis and function in plants. *FEBS Lett.* **579**, 5923–5931.

Chen, K. and Rajewsky, N. (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* **8**, 93–103.

Chory, J., Peto, C., Feinbaum, R., Pratt, L. and Ausubel, F. (1989) *Arabidopsis thaliana* mutant that develops as a light-grown plant in the absence of light. *Cell*, **58**, 991–999.

Chory, J., Nagpal, P. and Peto, C.A. (1991) Phenotypic and genetic analysis of *det2*, a new mutant that affects light-regulated seedling development in Arabidopsis. *Plant Cell*, **3**, 445–459.

Chory, J., Elich, T., Li, H.M. et al. (1994) Genes controlling Arabidopsis photomorphogenesis. *Biochem. Soc. Symp.* **60**, 257–263.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

Cluis, C.P., Mouchel, C.F. and Hardtke, C.S. (2004) The Arabidopsis transcription factor HY5 integrates light and hormone signaling pathways. *Plant J.* **38**, 332–347.

Cohen, E.E. and Rosner, M.R. (2009) MicroRNA-regulated feed forward loop network. *Cell Cycle*, **8**, 2477–2478.

Davuluri, R.V., Sun, H., Palaniswamy, S.K., Matthews, N., Molina, C., Kurtz, M. and Grotewold, E. (2003) AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinformatics*, **4**, 25.

Deplancke, B., Mukhopadhyay, A., Ao, W. et al. (2006) A gene-centered *C. elegans* protein–DNA interaction network. *Cell*, **125**, 1193–1205.

Fahlgren, N., Howell, M.D., Kasschau, K.D. et al. (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of *MIRNA* genes. *PLoS ONE*, **2**, e219.

García, P. and Frampton, J. (2008) Hematopoietic lineage commitment: miRNAs add specificity to a widely expressed transcription factor. *Dev. Cell*, **14**, 843–853.

Gendrel, A.V., Lippman, Z., Martienssen, R. and Colot, V. (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat. Methods*, **2**, 213–218.

German, M.A., Pillay, M., Jeong, D.H. et al. (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* **26**, 941–946.

Ghosh, S., Hirsch, H.A., Sekinger, E., Struhl, K. and Gingeras, T.R. (2006) Rank-statistics based enrichment-site prediction algorithm developed for chromatin immunoprecipitation on chip experiments. *BMC Bioinformatics*, **7**, 434.

Harbison, C.T., Gordon, D.B., Lee, T.I. et al. (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature*, **431**, 99–104.

Holm, M., Ma, L.G., Qu, L.J. and Deng, X.W. (2002) Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. *Genes Dev.* **16**, 1247–1259.

Ji, H. and Wong, W.H. (2005) TileMap: create chromosomal map of tiling array hybridization. *Bioinformatics*, **21**, 3629–3636.

Jones-Rhoades, M.W., Bartel, D.P. and Bartel, B. (2006) MicroRNAS and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **57**, 19–53.

Khvorova, A., Reynolds, A. and Jayasena, S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell*, **115**, 209–216.

Koornneef, M., Rolff, E. and Spruitt, C.J.P. (1980) Genetic control of light-inhibited hypocotyl elongation in *Arabidopsis thaliana* L. Heynh. *Z. Pflanzenphysiol.* **100**, 147–160.

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

Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75**, 843–854.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H. and Kim, V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060.

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I. and Deng, X.W. (2007) Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *Plant Cell*, **19**, 731–749.

Li, Y., Zhang, Q., Zhang, J., Wu, L., Qi, Y. and Zhou, J.M. (2010) Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. *Plant Physiol.* **152**, 2222–2231.

Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448–3449.

Martinez, N.J., Ow, M.C., Barrasa, M.I., Hammell, M., Sequerra, R., Doucette-Stamm, L., Roth, F.P., Ambros, V.R. and Walhout, A.J. (2008) A *C. elegans* genome-scale microRNA network contains composite feedback motifs with high flux capacity. *Genes Dev.* **22**, 2535–2549.

Megraw, M., Baev, V., Rusinov, V., Jensen, S.T., Kalantidis, K. and Hatzigeorgiou, A.G. (2006) MicroRNA promoter element discovery in Arabidopsis. *RNA*, **12**, 1612–1619.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. and Alon, U. (2002) Network motifs: simple building blocks of complex networks. *Science*, **298**, 824–827.

Molina, C. and Grotewold, E. (2005) Genome wide analysis of Arabidopsis core promoters. *BMC Genomics*, **6**, 25.

Oyama, T., Shimura, Y. and Okada, K. (1997) The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes Dev.* **11**, 2983–2995.

Qu, L.J. and Zhu, Y.X. (2006) Transcription factor families in Arabidopsis: major progress and outstanding issues for future research. *Curr. Opin. Plant Biol.* **9**, 544–549.

Quail, P.H. (2002) Photosensory perception and signalling in plant cells: new paradigms? *Curr. Opin. Cell Biol.* **14**, 180–188.

Re, A., Corá, D., Taverna, D. and Caselle, M. (2009) Genome-wide survey of microRNA-transcription factor feed-forward regulatory circuits in human. *Mol. Biosyst.* **5**, 854–867.

Riechmann, J.L., Heard, J., Martin, G. *et al.* (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*, **290**, 2105–2110.

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V. and Furlong, E.E. (2007) A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* **21**, 436–449.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.

Shalgi, R., Lieber, D., Oren, M. and Pilpel, Y. (2007) Global and local architecture of the mammalian microRNA–transcription factor regulatory network. *PLoS Comput. Biol.* **3**, e131.

Shen-Orr, S.S., Milo, R., Mangan, S. and Alon, U. (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* **31**, 64–68.

Shin, J., Park, E. and Choi, G. (2007) PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in Arabidopsis. *Plant J.* **49**, 981–994.

Sikder, D. and Kodadek, T. (2005) Genomic studies of transcription factor-DNA interactions. *Curr. Opin. Chem. Biol.* **9**, 38–45.

Song, Y.H., Yoo, C.M., Hong, A.P. *et al.* (2008) DNA-binding study identifies C-box and hybrid C/G-box or C/A-box motifs as high-affinity binding sites for STF1 and LONG HYPOCOTYL5 proteins. *Plant Physiol.* **146**, 1862–1877.

Sullivan, J.A. and Deng, X.W. (2003) From seed to seed: the role of photoreceptors in Arabidopsis development. *Dev. Biol.* **260**, 289–297.

Tsang, J., Zhu, J. and van Oudenaarden, A. (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol. Cell*, **26**, 753–767.

Vandenbussche, F., Habricot, Y., Condiff, A.S., Maldiney, R., Van der Straeten, D. and Ahmad, M. (2007) HY5 is a point of convergence between cryptochrome and cytokinin signalling pathways in *Arabidopsis thaliana*. *Plant J.* **49**, 428–441.

Vermeirssen, V., Barrasa, M.I., Hidalgo, C.A., Babon, J.A., Sequerra, R., Doucette-Stamm, L., Barabasi, A.L. and Walhout, A.J. (2007) Transcription factor modularity in a gene-centered *C. elegans* core neuronal protein–DNA interaction network. *Genome Res.* **17**, 1061–1071.

Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669–687.

Wightman, B., Ha, I. and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell*, **75**, 855–862.

Wingender, E., Dietze, P., Karas, H. and Knüppel, R. (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* **24**, 238–241.

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A. and Carrington, J.C. (2005) Expression of Arabidopsis *MIRNA* genes. *Plant Physiol.* **138**, 2145–2154.

Yadav, V., Kundu, S., Chattopadhyay, D., Negi, P., Wei, N., Deng, X.W. and Chattopadhyay, S. (2002) Light regulated modulation of Z-box containing promoters by photoreceptors and downstream regulatory components, COP1 and HY5, in Arabidopsis. *Plant J.* **31**, 741–753.

Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S., Milo, R., Pinter, R.Y., Alon, U. and Margalit, H. (2004) Network motifs in integrated cellular networks of transcription-regulation and protein–protein interaction. *Proc. Natl. Acad. Sci. USA*, **101**, 5934–5939.

Yu, X., Lin, J., Zack, D.J., Mendell, J.T. and Qian, J. (2008) Analysis of regulatory network topology reveals functionally distinct classes of microRNAs. *Nucleic Acids Res.* **36**, 6494–6503.

Zhang, Y. (2005) miRU: an automated plant miRNA target prediction server. *Nucleic Acids Res.* **33**, W701–W704.

Zhang, X., Yazaki, J., Sundaresan, A. *et al.* (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell*, **126**, 1189–1201.

Zhou, X., Ruan, J., Wang, G. and Zhang, W. (2007) Characterization and identification of microRNA core promoters in four model species. *PLoS Comput. Biol.* **3**, e37.