ACCESS
CellPress
Cell Reports
Resource

A Genome-Scale Resource for the Functional Characterization of *Arabidopsis* Transcription Factors

Jose L. Pruneda-Paz,[1,2,4,*] Ghislain Breton,[1,4,5] Dawn H. Nagel,[1,2,8] S. Earl Kang,[1,2] Katia Bonaldi,[1,2] Colleen J. Doherty,[1,2,6] Stephanie Ravelo,[1,7] Mary Galli,[3,9] Joseph R. Ecker,[3] and Steve A. Kay[1,2,8,*]
[1]Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[2]Center for Chronobiology, University of California, San Diego, La Jolla, CA 92093, USA
[3]Genomic Analysis Laboratory, Howard Hughes Medical Institute and The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[4]Co-first author
[5]Present address: Department of Integrative Biology and Pharmacology, University of Texas Health Science Center at Houston (UTHSC), Houston, TX 77030, USA
[6]Present address: Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695, USA
[7]Present address: Division of Rheumatology, Allergy, and Immunology, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA
[8]Present address: Molecular and Computational Biology Section, University of Southern California, Los Angeles, CA 90089, USA
[9]Present address: Waksman Institute of Microbiology, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA
*Correspondence: jprunedapaz@ucsd.edu (J.L.P.-P.), stevekay@usc.edu (S.A.K.)
http://dx.doi.org/10.1016/j.celrep.2014.06.033

SUMMARY

Extensive transcriptional networks play major roles in cellular and organismal functions. Transcript levels are in part determined by the combinatorial and overlapping functions of multiple transcription factors (TFs) bound to gene promoters. Thus, TF-promoter interactions provide the basic molecular wiring of transcriptional regulatory networks. In plants, discovery of the functional roles of TFs is limited by an increased complexity of network circuitry due to a significant expansion of TF families. Here, we present the construction of a comprehensive collection of *Arabidopsis* TFs clones created to provide a versatile resource for uncovering TF biological functions. We leveraged this collection by implementing a high-throughput DNA binding assay and identified direct regulators of a key clock gene (CCA1) that provide molecular links between different signaling modules and the circadian clock. The resources introduced in this work will significantly contribute to a better understanding of the transcriptional regulatory landscape of plant genomes.

INTRODUCTION

Transcription factors (TFs) are one of the largest functional classes of proteins encoded in eukaryotic genomes, often accounting for almost 8% of the total gene pool (Weirauch and Hughes, 2011). A genome-wide survey of binding sites for 119 human TFs using chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) revealed that even with a partial view of the TF ensemble (119 out of ~2,000), the fraction of DNA base-pairs that are involved in gene regulation is far greater than the protein-coding fraction (Bernstein et al., 2012). The results from this global survey further confirm that gene expression is regulated by the combinatorial effect of multiple TFs bound to gene regulatory regions. Despite their widespread importance, however, only a limited number of TFs have been characterized at the biochemical and molecular levels.

Pioneering work in sea urchin that combined genetics, transcript profiling, and localization of *cis*-regulatory modules established a gene regulatory network (GRN) model that mapped the regulatory logic of developmental control (Oliveri et al., 2008). Subsequent studies indicated that numerous transcription-based networks are sufficiently hardwired in the genome to be modeled (Davidson, 2006). The construction of circadian clock-regulated networks revealed a much more complex circuitry than was previously anticipated by forward genetic screens (Koike et al., 2012; Rey et al., 2011; Ueda et al., 2005). Several positive and negative feedback loops appear to have evolved in parallel, which might allow for multiple input signals to provide proper phasing and rhythmic synchrony of biological processes with the oscillating environment (Zhang and Kay, 2010). In addition to this redundant network architecture, most plant TF families have significantly expanded during evolution, potentially to diversify the mechanisms necessary to survive the myriad of challenges associated with direct exposure to the environment (Shiu et al., 2005). Thus, network and genetic redundancy likely accounts for a large proportion of the circadian and other transcriptional networks' resilience against forward genetic approaches in plants (Pruneda-Paz and Kay, 2010).

To overcome the difficulties associated with network and TF redundancy, novel approaches to physically map *cis*-regulatory networks have been developed (Bassel et al., 2012). Although TF-centered approaches such as ChIP-seq can reveal the extent

to which a particular TF is involved in genome-wide regulation, other techniques, such as high-throughput yeast one-hybrid (HT-Y1H), are promoter focused and can provide a survey of potential interactors for a single promoter. Reagents for the latter have been developed and applied successfully to study gene regulation in humans, flies (*Drosophila melanogaster*), and worms (*Caenorhabditis elegans*), and are starting to reveal the complexity of combinatorial gene regulation (Deplancke et al., 2004; Hens et al., 2011; Reece-Hoyes et al., 2011). Recently, we implemented this genomic strategy in plants (*Arabidopsis thaliana*, *Arabidopsis*) and identified transcriptional components of the circadian clock (Pruneda-Paz et al., 2009) as well as other plant physiological processes (Ito et al., 2012; Li et al., 2012; Niwa et al., 2013). In these studies, TF-DNA interactions identified in Y1H screens were confirmed in vivo by ChIP and the biological relevance of identified TFs was demonstrated by reverse-genetics approaches. Thus, combined with community-developed reagents such as homozygous insertion line collections (O'Malley and Ecker, 2010) and genome-editing strategies (Puchta and Fauser, 2014), HT-Y1H provides a powerful technique to explore the cohort of TFs that bind to any gene promoter. Specifically, large transcriptional regulatory networks such as the circadian clock can be comprehensively studied using this approach.

In our previous study, we took advantage of approaches developed for *C. elegans* (Deplancke et al., 2004, 2006), as well as previously generated *Arabidopsis* TF collections (Gong et al., 2004; Paz-Ares et al., 2002; Underwood et al., 2006; Yamada et al., 2003), to develop HT-Y1H screens for *Arabidopsis* (Pruneda-Paz et al., 2009). Since our initial arrayed clone collection included the open reading frames (ORFs) for only 186 circadian-regulated TFs (Pruneda-Paz et al., 2009), we decided to generate a genome-wide plasmid library encompassing the ORFs for all known *Arabidopsis* TFs (TF ORFeome). During the last decade, several efforts have been undertaken to annotate *Arabidopsis* TFs and create digital repositories (Guo et al., 2005; Kummerfeld and Teichmann, 2006; Riaño-Pachón et al., 2007; Riechmann et al., 2000), as well as to generate TF clone collections (Gong et al., 2004; Paz-Ares et al., 2002; Underwood et al., 2006; Yamada et al., 2003). Building from these resources, we constructed the most complete *Arabidopsis* TF ORFeome available to date. All clones in this collection were individually sequence validated, have the same reading frame and vector backbone, and are compatible with recombination-based cloning. Thus, quality and versatility are essential features of this ORFeome. This TF collection was used to develop automated HT-Y1H screens that uncovered players in the transcriptional networks underlying the *Arabidopsis* clock function. The resources presented here will provide reliable and universal reagents for exploring the regulatory landscape of the *cis*-regulome.

RESULTS

Construction of a TF Clone Collection Compatible with Recombination-Based Cloning

The structural redundancy inherent to complex positive/negative-feedback loops and the large expansion of plant TF families greatly hinders our ability to deconvolute the organization of transcription-based regulatory networks such as the *Arabidopsis* circadian clock. To bypass these limitations, we previously implemented a reverse-genetics approach that proved to be successful in other model organisms and identified a TF that regulates the *Arabidopsis* clock function (Deplancke et al., 2004; Pruneda-Paz et al., 2009). This alternative strategy takes advantage of TF-specific plasmid libraries and the Y1H system (Deplancke et al., 2004). Considering the simplicity and success of the approach and that our first TF collection only included circadian-regulated TFs, we hypothesized that additional clock components could be discovered with a larger gene collection. We therefore aimed to generate a complete sequence-validated *Arabidopsis* TF library to continue our exploration of the circadian system in a comprehensive and unbiased manner. To achieve this goal, we compiled a comprehensive list of *Arabidopsis* TFs and transcriptional regulators (hereafter globally referred to as TFs) by combining all TFs predicted in four independent databases (PlnTFDB, http://plntfdb.bio.uni-potsdam.de; DATF, http://datf.cbi.pku.edu.cn; DBD, http://www.transcriptionfactor.org; and REGIA Consortium, http://www.jicgenomelab.co.uk) (Guo et al., 2005; Kummerfeld and Teichmann, 2006; Paz-Ares et al., 2002; Riaño-Pachón et al., 2007). Because each database used alternative TF identification algorithms to survey the *Arabidopsis* genome, the TF gene predictions in each set did not fully overlap. To be as inclusive as possible, we included all TF encoding genes present in any of these databases (2,492 genes) in our initial cloning pipeline (Table S1). Grouping TFs by their occurrence in each data set revealed that approximately equal proportions of the genes are found in one or all databases (Tables S1 and S2). Additionally, based on specific literature searches, we included a small number of TFs (Tables S1 and S2; referred to as "other").

While some *Arabidopsis* TF ORFeomes were previously generated (Gong et al., 2004; Paz-Ares et al., 2002; Underwood et al., 2006; Yamada et al., 2003), our goal was to build a comprehensive, sequence-validated, and homogeneous clone collection regarding the plasmid backbone, resistance marker, and coding sequence reading frame. Each TF coding sequence was amplified with its respective STOP codon and cloned in the Gateway-compatible vector pENTR/D (Life Technologies). This universal format simplifies bulk downstream usage of the collection to develop TF-centered experimental approaches. Primer pairs for the longest isoform of the 2,492 TFs were designed according to TAIR9 gene annotations (http://www.arabidopsis.org; Lamesch et al., 2012) and used throughout six independent rounds of cloning and sequencing (Figure 1A). To build upon previous TF collections, we used 1,728 clones from the Salk/Stanford/PGEC Consortium, TIGR, REGIA, and Yale collections (Gong et al., 2004; Paz-Ares et al., 2002; Underwood et al., 2006; Yamada et al., 2003) as PCR templates for the first round. This step yielded 746 pENTR/D-TF clones for the new TF ORFeome (Figure 1A). The remaining TF coding sequences (1,210) were amplified "de novo" using a cDNA pool isolated from 1-week-old seedlings collected over different times of the day. After the sixth cloning round, we completed a collection of 1,956 sequence-confirmed pENTR/D-TF clones representing 78.5%

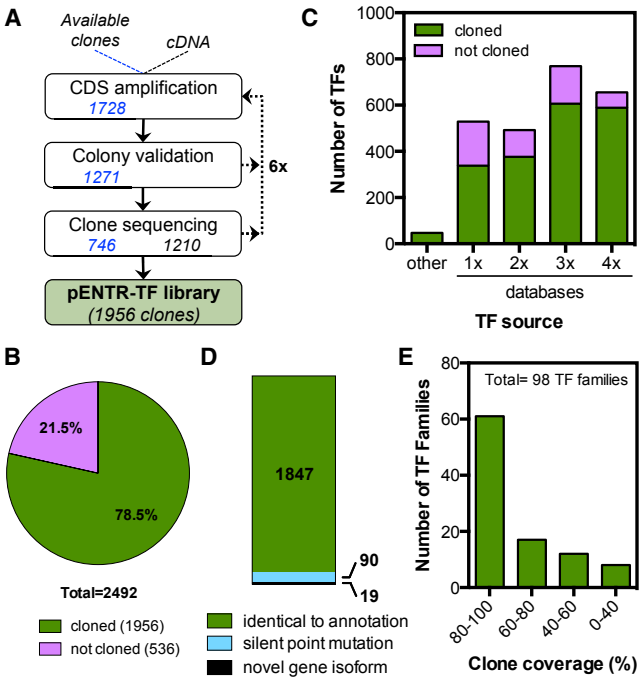

Figure 1. Construction of a Genome-Wide *Arabidopsis* TF Collection

(A) Cloning workflow. Each TF coding sequence was amplified from available clones or "de novo" from cDNA (the number of coding sequences obtained at each step is indicated in blue and black, respectively).

(B) Coverage of sequence-validated clones included in the collection.

(C) Frequency distribution of *Arabidopsis* TFs in four independent databases (cloned genes in each category are indicated in green; see also Table S1).

(D and E) Sequence quality (D) and TF family coverage (E) for the clones included in the collection.

of all *Arabidopsis* TFs (Figures 1A–1C; Table S3). Almost all TF genes included in the collection (99%) encode polypeptides predicted by the current *Arabidopsis* genomic annotation (TAIR10) (Figure 1D; Table S4). A fraction of these clones (4.6%) contain silent point mutations likely generated during gene amplification (Figure 1D; Tables S3 and S4). The remaining 1% of the genes included in the collection correspond to coding sequences with conserved mismatches (with respect to the current gene annotation) repeatedly observed across most cloning attempts, and therefore likely encode TF isoforms not previously identified (Figures 1D and S1A; Tables S3 and S4). The collection includes TFs for 93 out of 97 (96%) predicted *Arabidopsis* TF families and has a clone coverage higher than 50% for 86 (89%) of these TF families (Figures 1E and S1B; Table S5). Only four TF family singletons (hATP, LFY, NOT, and TBP) are not present in the collection, whereas other families, such as ARF, PHD, and SNF2, are significantly underrepresented (Figure S1B; Table S5). Cloning these TFs' coding sequences seems particularly difficult and may require alternative cloning strategies and/or expression systems.

Throughout the different cloning iterations, we systematically monitored each clone status at different steps, which allowed us to troubleshoot the cloning procedure for specific groups of genes. For example, cloning coding sequences longer than 2,000 bp required an adjustment of most protocols (for details, see the Supplemental Experimental Procedures). Out of the 536 missing clones, 135 were not found in any previous collection and never amplified from our cDNA pool, 210 were prone to multiple mutations and never found intact, and the rest could not be cloned for various reasons (Tables S6 and S7). It is worth noting that gene size does not seem to be a limiting factor for our cloning procedure (Table S6). Although a small subset of predicted *Arabidopsis* TFs remain to be cloned, the resource described here represents the largest eukaryotic sequence-validated TF ORF collection available to date. Several genes in our TF ORFeome are either missing (15.1%) or not specifically detected (3.3%) by the commonly used ATH1 DNA microarray platform (Affymetrix) (Figures S1C and S1D). Inclusion of such genes in functional genomic applications will certainly enhance our ability to gain more information and subsequently uncover their functions (individual clones and the entire collection are available through the *Arabidopsis* Biological Resource Stock Center, http://www.abrc.osu.edu).

Construction of a TF Collection Suitable for Y1H Screens

Tremendous progress has been made in defining the DNA binding protein landscape using tools such as TF-centered HT-Y1H (Deplancke et al., 2004; Hens et al., 2011; Reece-Hoyes et al., 2011). By taking advantage of the pENTR/D-TF collection's compatibility with recombination-based cloning (Gateway; Life Technologies), we transferred each TF into a Y1H-compatible destination plasmid (pDEST22) that carries the GAL4-Activation Domain (GAL4_AD) located 5′ to the TF insertion site. TFs cloned in pDEST22 are thus expressed in yeast as C-terminal fusions to the GAL4_AD. After three rounds of ORF transfer and destination plasmid validation by restriction analysis and gene-specific colony PCR, 100% of the 1,956 clones were successfully cloned into pDEST22 (Figure 2A). The pDEST22-TF collection was arrayed in 21 96-well plates. To efficiently manage the larger number of constructs, we developed a newly optimized version of our previous TF library screen protocol (Figure S2A). Briefly, we generated a modified yeast strain (YU) that is able to mate with the strain YM4271, which carries the *promoter::lacZ* reporter constructs (Figure S2B). Similarly to a recently described procedure (Gaudinier et al., 2011), we transformed each pDEST22-TF 96-well plate into the YU strain using a previously described HT yeast transformation protocol (Walhout and Vidal, 2001). Each plate included one empty well as a negative growth control and one pDEST22 empty plasmid control to set the basal reporter activity in the absence of DNA binding. After transformation, the 96-well plates were indexed into six 384-well plates. In this condensed format, the YU library strains were mated with promoter-specific YM4271 reporter strains and the β-galactosidase reporter activity was quantified (Figure S2A). HT-Y1H screens were initially tested and optimized using a fully automated liquid-handling robotic platform (details in Supplemental Experimental Procedures) and a temporary "bridge" TF library generated while construction of the final gold standard collection was underway (Li et al., 2012). Although automation significantly increases the screening throughput, the overall procedure can be performed manually and thus could be implemented in any laboratory.

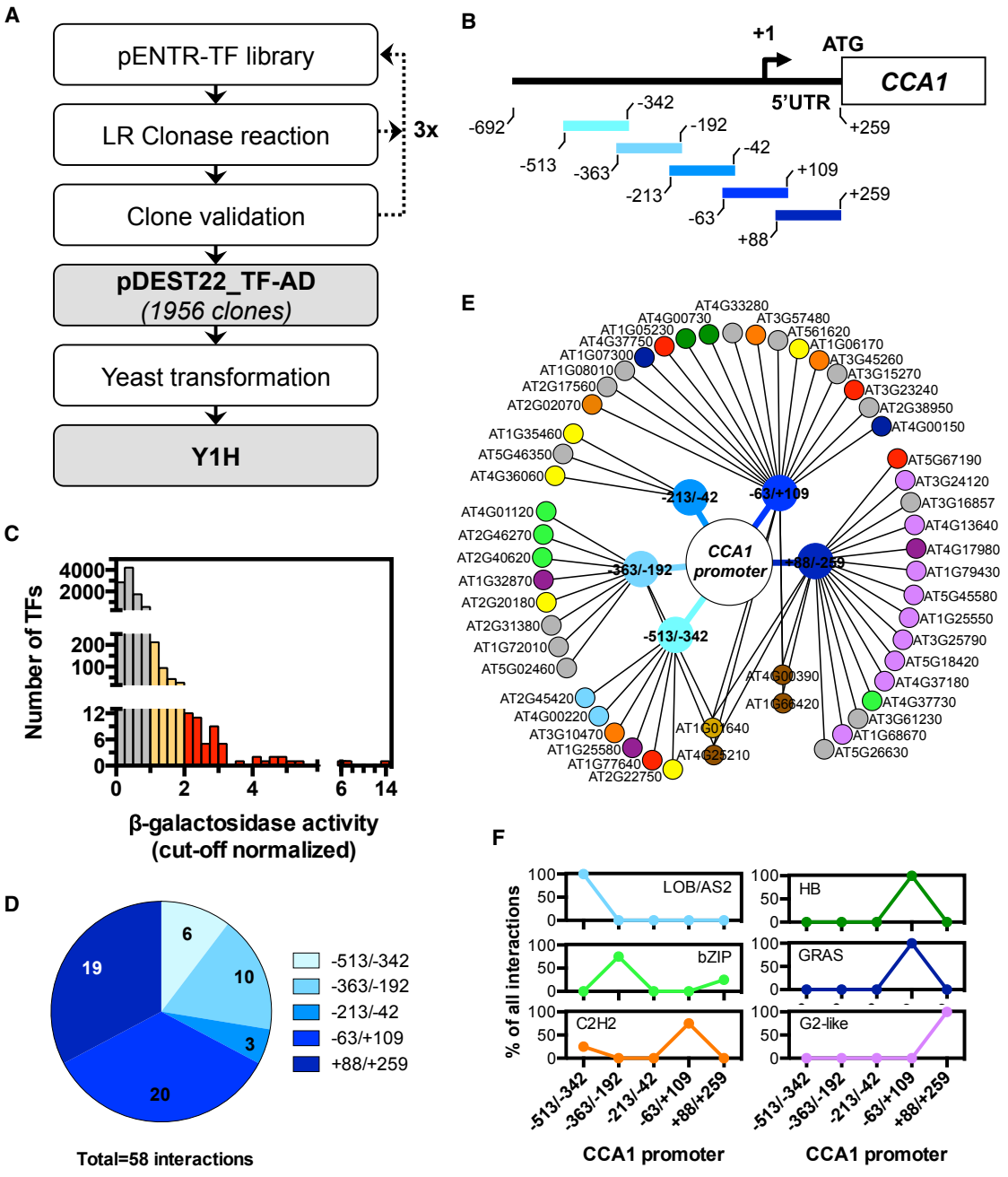

Figure 2. *CCA1* Promoter Y1H Screens with a Genome-Wide *Arabidopsis* TF Collection

(A) Workflow for the construction and validation of a "daughter" pDEST22-TF collection suitable for HT-Y1H screens.

(B) Schematic of the screened *CCA1* promoter fragment baits.

(C) Histogram of the HT-Y1H results obtained for all *CCA1* promoter fragments (Table S8). β-galactosidase activities obtained for each TF-promoter combination were normalized to the cutoff value established for each promoter fragment. Bars represent the number of TF-promoter combinations that resulted in cutoff normalized values below (gray) or above (orange and red) the cutoff value (bin width = 0.25). Red bars indicate high-confidence interactions.

(D) Distribution of high-confidence TF interaction hits for each *CCA1* promoter fragment (Table S8).

(E) *CCA1* promoter TF interaction network. Interacting TFs and the corresponding *Arabidopsis* gene identification (AGI) numbers are organized clockwise based on the β-galactosidase reporter activity obtained in the Y1H screen. Interactions by a single member of a TF family (gray) or multiple members of the bHLH (yellow), AP2-ERBP (red), NAC (purple), C2H2 (orange), LOB/AS2 (light blue), bZIP (light green), GRAS (blue), HB (green), G2-like (light purple), TRAF (light brown), or GeBP (brown) TF families are indicated (Table S8).

(F) Frequency distribution for the binding of multiple TF family members to each of the *CCA1* promoter fragments used in the HT-Y1H screen. TF families with high promoter fragment specificity are shown.

Uncovering Regulators of the *CCA1* Promoter Activity

In our original TF-centered Y1H screen, the promoter of a key *Arabidopsis* clock component, *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*), was scanned with a library encompassing 186 circadian-regulated *TFs* (Pruneda-Paz et al., 2009). Since this initial screen uncovered a core clock component, called CCA1 HIKING EXPEDITION (CHE), we decided to search for additional *CCA1* promoter regulators with the newly generated pDEST22-TF library. Since yeast promoters are rather compact (Dobi and Winston, 2007), instead of screening a long 1–2 kb *CCA1* promoter fragment, we previously opted to screen five short overlapping promoter segments of ~200 bp in length (Pruneda-Paz et al., 2009). To determine whether indeed this strategy increases the sensitivity of HT-Y1H screens, we analyzed the binding of CHE to either the five overlapping *CCA1* promoter fragments or the full-length *CCA1* promoter in a Y1H assay (Figure S2C). The binding of CHE could only be detected when a short *CCA1* promoter fragment was used, indicating that at least for some TFs, such as CHE, the screen sensitivity is significantly improved by reducing the distance between the *cis* element and the minimal promoter driving the *lacZ* reporter gene expression (Figure S2C). Based on this result, the five original *CCA1* promoter fragments were screened with the new TF collection (Figure 2B). The β-galactosidase reporter activity was determined for each promoter fragment-TF combination and negative control. Cutoff values were established at 4 SDs above the average reporter activity obtained for the pDEST22 empty plasmid controls (n = 21; details in the Supplemental Experimental Procedures). Using this criterion, we identified 15–160 TF interactions per promoter fragment, for a total of 431 interactions (318 unique TFs) that were statistically different from the pDEST22 control (Figure S2D; Table S8). Most of these interactions were specific for a single promoter fragment (73.3%), but some TFs interacted with two (19.8%), three (5.3%), four (1.3%), or all (0.3%) promoter baits (Figure S2E; Table S8). To further reduce the likelihood of false positives, we doubled the cutoff value (8 SDs above the average reporter activity obtained for the pDEST22 empty plasmid controls), which resulted in 58 high-confidence interactions by 52 unique TFs (Figures 2C–2E). Four of these TFs bound to two or three *CCA1* promoter fragments, accounting for ten high-confidence interactions (Figure 2E). The remaining 48 TFs showed increased β-galactosidase reporter activity with only one *CCA1* promoter bait fragment (Figure 2E; Table S8). A TF family analysis of these interactors revealed a wide array of different families and, in many cases, an enrichment of TFs from the same family on specific fragments (Figures 2E and 2F). An analysis of the TF family distribution for the 318 TFs that were found to interact with any of the five *CCA1* promoter fragments indicated a widespread TF family representation (Figure S2G). In addition, similar TF family binding fragment preferences were found in this larger set of *CCA1* promoter interactors (Figure S2H; Table S8).

Although HT-Y1H screens provide an effective strategy for rapidly identifying potential regulators of any given gene promoter, the biological relevance of these regulators must be addressed in vivo. CCA1 is a key component of the *Arabidopsis* clock and its misexpression almost unequivocally results in altered clock function (Nagel and Kay, 2012). Thus, we anticipated that overexpression of *CCA1* promoter interactors would perturb the clock function. To perform a rapid validation assessment, we cotransformed *Arabidopsis* protoplasts with a plasmid carrying a clock reporter construct (*CCA1::LUC+*) and a plasmid carrying an overexpression construct for the potential regulators (*CsVMVS::TF*). After transformation, the protoplasts were incubated under constant light conditions and the luciferase activity was monitored every 2.5 hr over a period of 5 days. Luciferase traces were used to analyze the clock function (details in the Experimental Procedures) upon overexpression of the top 34 Y1H TF candidates (accounting for 38 of the 58 high-confidence *CCA1* promoter interactions; Table S8). Compared with an empty effector plasmid control, the amplitude of recorded luminescence rhythms was significantly reduced when 32 of these TFs were overexpressed, suggesting that all of them could regulate *CCA1* expression (Figure 3A; Table S8). A rhythmicity profile analysis showed that overexpression of 21 out of the 32 TFs that exhibited amplitude effects also resulted in arrhythmic *CCA1* promoter activity (Figures 3B, 3C, and S3A; Table S8). Our results indicate that the top TF candidates tend to result in stronger (i.e., arrhythmic) clock phenotypes, which suggests that HT-Y1H candidate prioritization based on the strength of the β-galactosidase reporter activity provides an effective way to predict biologically relevant interactions (Figure 3D). Analysis of microarray data sets did not reveal a strong expression correlation or anticorrelation between *CCA1* and most protoplast-validated *TF*s (Figures 3E and S3B). This result indicates that *CCA1* regulators are not necessarily coexpressed with *CCA1*, thus emphasizing the advantage of using HT-Y1H for TF discovery. The circadian clock is postulated to be a highly interconnected regulatory hub regulated by multiple environmental and endogenous signals (Pruneda-Paz and Kay, 2010). To gain insights into the possible biological role of these TFs, we wanted to examine their potential functional associations by performing a Gene Ontology (GO) analysis. However, since little is known about the function of these TFs, no GO annotation other than their classification as transcriptional regulators existed for the majority of them. Therefore, for the 28 protoplast-validated TFs that are specifically detected on the ATH1 arrays (Affymetrix), we searched for first-order interaction partners in a publicly available *Arabidopsis* protein interaction data set (http://interactome.dfci.harvard.edu/A_thaliana/; Arabidopsis Interactome Mapping Consortium, 2011; Figure S3C). This analysis revealed interaction networks for ten out of the 28 protoplast-validated TFs, and these networks were significantly enriched for a wide range of GO biological process terms (Figure 3F), consistent with the expansive role of the circadian clock. Altogether, these results suggest that a multitude of TFs likely mediate the regulation of *CCA1* by different endogenous and environmental signals.

Characterization of the Regulation of *Arabidopsis* Clock Function by FBH1

A recently characterized TF named FLOWERING BHLH 1 (FBH1) (AT1G35460) (Ito et al., 2012) was among the TFs that displayed clock phenotypes in transient *Arabidopsis* protoplast assays (Figures 3A–3C and S3A). Characterization of this TF likely has been limited by the fact that *FBH1*-associated probe sets on the ATH1 array exhibit nonspecific cross-hybridization

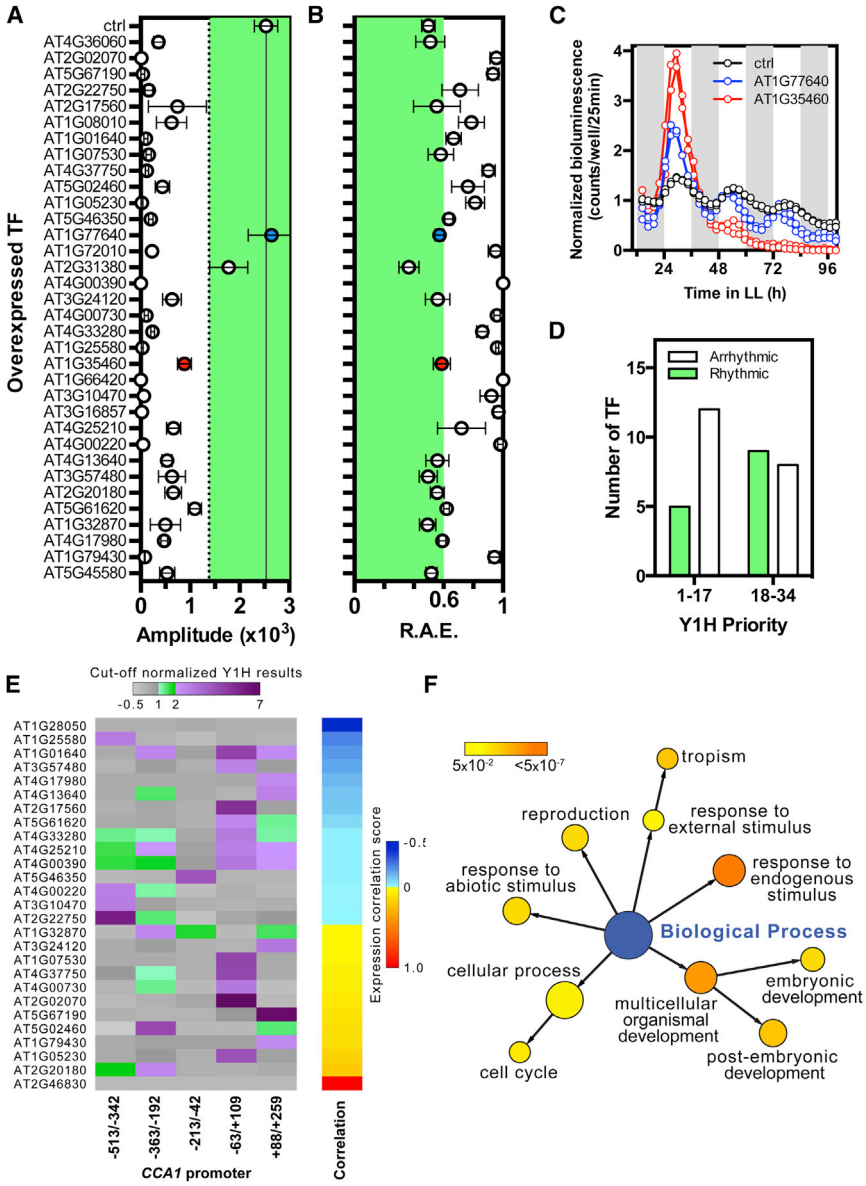

Figure 3. In Vivo Validation of Y1H Candidates

(A and B) Circadian clock phenotypes for TF overexpression in *Arabidopsis* protoplasts. TFs are indicated by the corresponding *Arabidopsis* gene identification numbers shown in (A) and were organized based on the β-galactosidase reporter activity (descending order). Each symbol represents the average value ± SEM for amplitude (A) or relative amplitude error (R.A.E.) (B) estimates of *CCA1* promoter-driven luciferase expression (*CCA1::LUC*+) in three independent experiments (n = 3–6). The cutoff limit was set at 3 SDs below the average amplitude value for the control (the green-shaded area indicates the interval of confidence for the control) (A). A relative amplitude error < 0.6 (green-shaded area) is indicative of rhythmic luciferase activity (B). Colored symbols (red and blue) indicate the examples shown in (C).

(C) Bioluminescence analysis of *CCA1::LUC*+ expression in *Arabidopsis* protoplasts transformed with *TF* overexpression plasmids (red and blue). Control traces (ctrl) correspond to protoplasts transformed with a control overexpression plasmid (black). Results are representative of three independent experiments.

(D) Frequency of rhythmic and arrhythmic clock phenotypes encountered in two groups of candidates based on the order of priority defined by the β-galactosidase reporter activity.

(E) Expression correlation analysis for *CCA1* and each TF validated in *Arabidopsis* protoplasts. Heatmaps representing the Y1H results defined by the β-galactosidase reporter activity obtained for each screened *CCA1* promoter fragment (left) and expression correlation scores (right) are shown. TFs are sorted according to their correlation score.

(F) Network of GO biological processes associated with TFs validated in *Arabidopsis* protoplasts. Node labels correspond to GO slim terms. Node colors denote the degree of correspondence to a GO slim term (enrichment p value < 0.05; n = 10 TFs).

(Table S1). For this reason, we decided to further explore the function of FBH1 in the regulation of *CCA1*. Y1H assays performed in a 96-well format confirmed that FBH1 binds to the −213/−42 region of the *CCA1* promoter (Figure 4A). Two potential FBH1 binding sites, a canonical E box motif (CANNTG) and a noncanonical one (CANNNG), map to this region of the *CCA1* promoter (Figure 4B). To determine the functionality of these motifs, mutations were introduced and Y1H reporter strains carrying these mutations in the context of the −213/-42 region of the *CCA1* promoter were generated (Figure 4B). Reporter activation indicated that FBH1 binds preferentially to the noncanonical E box-like motif CACTAG (Figures 4B and 4C). To analyze the binding of FBH1 to the *CCA1* promoter in planta, we performed ChIP experiments using GFP-tagged *FBH1* overexpression lines. The same clock phenotypes were observed when tagged or untagged versions of FBH1 were overexpressed, suggesting that GFP-FBH1 retains its function (Figures S4D–S4F). The ChIP results confirmed that FBH1 specifically binds to the *CCA1* promoter in vivo (Figure 4D). To determine FBH1-associated clock phenotypes in planta, we obtained homozygous *FBH1*-overexpression/*CCA1::LUC*+ *Arabidopsis* lines (Figure S4A). A bioluminescence time-course expression analysis using these *Arabidopsis* lines indicated that the overall *CCA1* promoter activity was significantly reduced, suggesting that FBH1 is a repressor of *CCA1* (Figure 4E). In addition, while the period of clock-controlled oscillations remained unaltered in these lines, the phase of *CCA1* peak exhibited an ~1 hr advance (Figures 4E–4G, S4B, and S4C). Altogether, these results indicate that FBH1 is a clock component that directly and negatively regulates *CCA1*. Similar experiments will be required to characterize the function of the remaining *CCA1* regulators identified in this work.

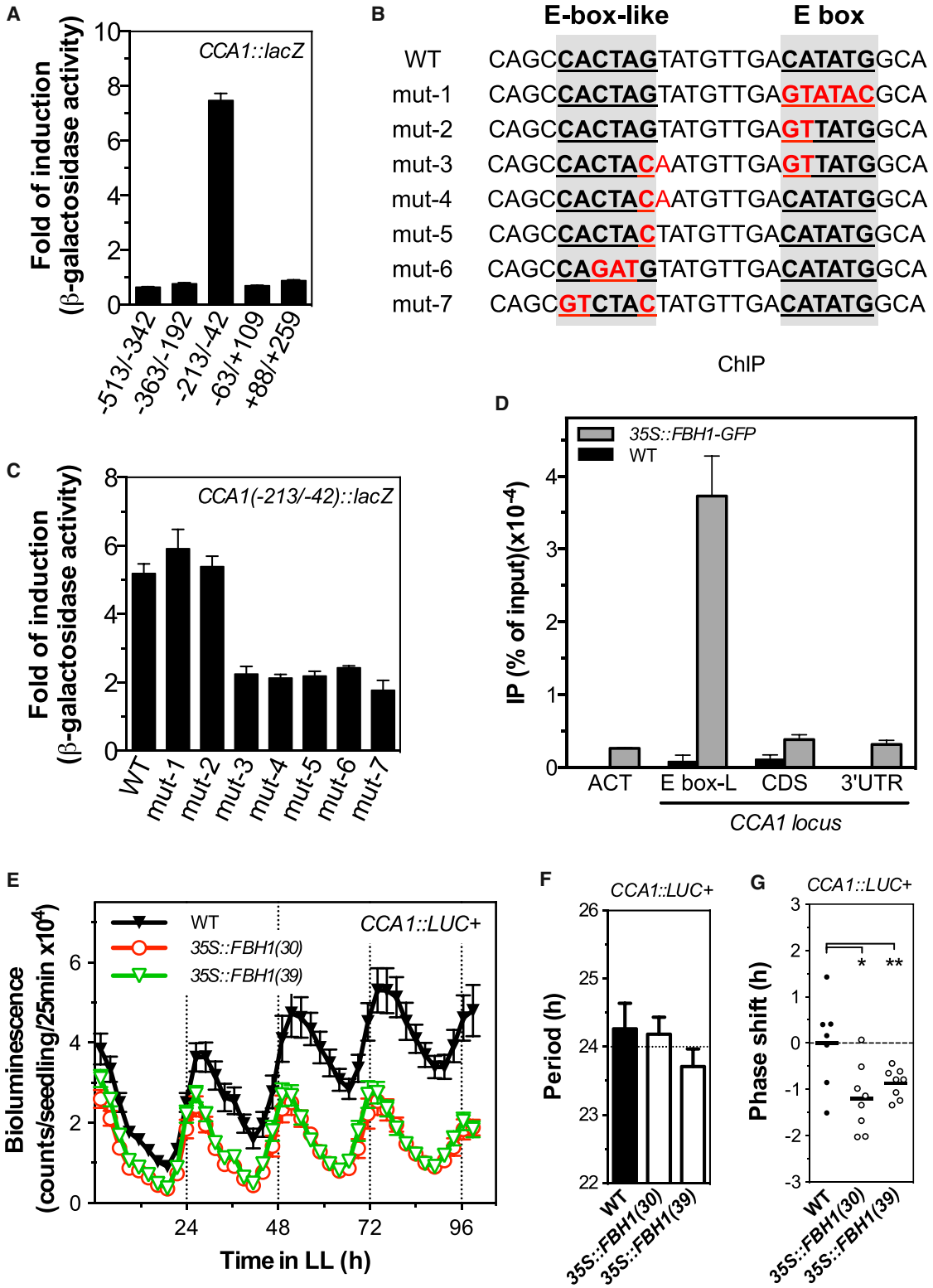

Figure 4. Characterization of the FBH1-*CCA1* Promoter Interaction

(A) Binding of FBH1 to different regions of the *CCA1* promoter in yeast. Bars represent the fold of induction in β-galactosidase activity for each of the DNA fragments indicated (n = 12).

(legend continued on next page)

DISCUSSION

By combining community-based gene annotation, previous ORFeome efforts, and our optimized pipeline for gene cloning and transfer, we created the most comprehensive eukaryotic TF ORF clone collection available to date (1,956 TFs). This collection is compatible with recombination-based cloning, which allows bulk transfer of TF coding sequences into any destination plasmid, such as those already created by the *Arabidopsis* community (Karimi et al., 2002; Rhee et al., 2003). The collection was conceived as a high-throughput resource to generate large genomic tools for the study of *Arabidopsis* TF function (Castrillo et al., 2011; Chang et al., 2013; Ou et al., 2011; Siggers et al., 2012; Wehner et al., 2011). One of our main goals in this project was to provide the highest clone quality. Indeed, the nucleotide sequence for 1,847 of the 1,956 clones included in the collection is identical to the current gene annotation (TAIR10). Despite our efforts, however, the collection is still missing 536 clones. A survey of our cloning pipeline history indicates that 25% of these genes were never detected by PCR (Table S6). These might correspond to pseudo-genes or genes that are not present in the cDNA pool used as PCR template. They could also be misannotated genes that are not amplifiable with the PCR primer set used. In addition, most of the remaining missing genes were successfully PCR amplified, but were never found as an intact or correct clone. It is possible that toxicity in *E. coli,* assuming that the cloned gene is translated, could account for this outcome. This seems to be the case, as almost 65% of these clones had deleterious point mutations, were consistently truncated at different locations, or could never be found at all (Table S6). Future attempts to complete the collection will need to consider alternative bacterial hosts or cloning and maintaining these constructs directly in yeast cells.

By following an iterative procedure like the one used to obtain the pENTR/D-TF clones and applying a stringent quality-control procedure after recombination-based cloning, we successfully transferred all TF coding sequences into a Y1H-compatible plasmid (pDEST22). This achievement allowed us to generate the most comprehensive TF collection for promoter DNA binding protein profiling available to date. Although most previous Y1H screens in eukaryotes were performed using large promoter fragments, we found that the use of shorter overlapping fragments provides increased sensitivity to the assay (Figure S2C). This experimental design revealed a total of 318 TFs that accounted for 431 interactions in the *CCA1* promoter region used in the screen. Although a similar number of these TFs (n = 68–160) were found to bind to most promoter baits, a significantly lower number (n = 15) was found for one particular promoter fragment (−213/−42). This is likely due to the elevated basal β-galactosidase activity detected with this reporter strain, which only allows the detection of TFs that strongly induce expression of the *lacZ* reporter. How many TFs can interact in vivo with a particular promoter? Additional large-scale deep screenings of specific promoter fragments using HT-Y1H are likely to provide this answer in the near future. Recent results from comprehensive monitoring of DNA binding for more than 100 TFs by ChIP-seq indicate that promoters are bound by several different TFs, each of which binds to thousands of sites in the genome, and that complex combinatorial gene regulation may be the norm in eukaryotes (Bernstein et al., 2012). This is in line with the concept of "billboard" enhancers, developed in studies of fly enhancers, which posits that a flexible group of diverse DNA binding proteins that aggregate depending on specific conditions suffice to provide time- and condition-dependent gene regulation (Arnosti and Kulkarni, 2005). In such a scenario, we could imagine that promoters containing hundreds of different TF binding sites might, depending on the tissue, developmental stage, or time of day, help to coalesce a specific combination of TFs to support a specific biological function. It is certainly premature to provide definitive answers, but tools such as the TF ORFeome presented here will certainly help in advancing toward this goal.

By clustering the 58 high-confidence *CCA1* promoter interactors by TF family, we observe that some of them display a significant binding preference for a specific *CCA1* promoter fragment (Figure 2E). This trend is maintained across all 318 interactors in our study (Table S8). For example, most interactions by G2-like TFs (63%) are detected with the +88/+259 *CCA1* promoter bait strain (Figure S2H). Similarly, 53% of all TCP interactions are detected with the −363/−192 *CCA1* promoter bait strain (Figure S2H). Most of these TCPs (eight out of ten) belong to the same clade, class I TCP TFs, suggesting that binding specificity is driven by the conserved DNA binding domain among these TFs and the presence of a specific *cis* element in the −363/−192 *CCA1* promoter region. In fact, we recently characterized the binding of one of these TFs, named CHE, to a canonical TCP class I binding site (GGNCCCAC) centered at nucleotides −227/−228 of the *CCA1* promoter (Figure S3B; Pruneda-Paz et al., 2009). Our results also indicate that some TF families display low promoter fragment specificity (Figure 2F).

(B) E box and E-box-like motifs present in the *CCA1* promoter region −243/−42. Wild-type (WT) and mutated versions (mut-1 to mut-7) used in Y1H assays are shown (base changes are indicated in red).
(C) Binding of FBH1 to the −243/−42 region of the *CCA1* promoter in yeast. WT and E box/E box-like mutant (B) promoter fragments fused to the *lacZ* gene were used as promoter baits. Bars represent the fold of induction in β-galactosidase activity for each of the DNA fragments indicated (n = 6).
(D) Binding of FBH1 to the *CCA1* promoter in vivo. ChIP assays were performed with *35S::FBH1-GFP* and wild-type *CCA1::LUC*+ (WT) seedlings. Immunoprecipitated DNA was quantified by real-time quantitative PCR with primers specific for the E-box-like locus in the *CCA1* promoter (Ebox-L) and for the CCA1 coding (CDS), *CCA1* 3′ UTR (3′ UTR), and ACTIN (ACT) coding regions as controls. Results were normalized to the input DNA (n = 3 independent experiments).
(E) Bioluminescence analysis of *CCA1::LUC*+ expression in *FBH1* overexpression lines (*35S::FBH1*) (n = 8). WT traces correspond to *CCA1::LUC*+ seedlings. Results are representative of at least three independent experiments.
(F) Period estimates of luciferase expression for the control and overexpression lines shown in (E). Each bar represents the average period value (hr) (WT, 24.27 ± 0.37; *35S::FBH1(30)*, 24.19 ± 0.25; *35S::FBH1(39)*, 23.71 ± 0.26 [±SD]).
(G) Phase shifts of luciferase expression for the control and overexpression lines shown in (E). Each symbol represents one seedling and the line is the average phase-shift value (hr) (*$p < 0.02$, **$p < 0.03$). Values represent means ± SEM (A, C, and D) or ± SD (D).

For example, six GeBP TFs account for 20 interactions (3.3 interactions per TF) that include all of the *CCA1* promoter fragments tested (Figure S2H). This interaction promiscuity is likely due to a low DNA binding affinity, broad binding preference, and/or lack of binding partners, as we did not identify common 5–10 nt motifs between two or more *CCA1* promoter fragments (data not shown). On the other hand, 20 AP2 TFs were also bound to all but one of the *CCA1* promoter fragments tested (Figure S2H). However, in this case, these TFs account for only 28 interactions (1.4 interactions per TF), suggesting that these AP2 TFs likely share a high DNA binding specificity, but across the family, different clades may recognize alternative *cis* elements. Some TF families are not present in our interactor data set, suggesting that the *CCA1* promoter regions tested in this work do not encode the corresponding TF binding site. This seems to be the case for ARF TFs, as the canonical AUX-RE (TGTCTC) is not present in the analyzed *CCA1* promoter fragments. However, it is also possible that the binding of some TFs cannot be detected, either because they are not properly expressed in yeast or because they require plant-specific interaction partners or posttranslational modifications. Altogether, these results indicate that in addition to identifying TFs, HT-Y1H provides insights into TF's DNA binding modes. Exploration of genome-wide TF binding preferences, as well as TF structural and phylogenetic analyses, will be required before we can further assess these observations (Badis et al., 2009; Weirauch and Hughes, 2011).

Secondary testing of top-priority HT-Y1H hits in a transient protoplast system revealed that most interactors could perturb *CCA1* expression and the overall clock function in vivo (Figures 3A and 3B). Interestingly, most of these TFs do not share a high expression correlation with *CCA1* and are associated with a wide range of biological functions (Figures 3E and 3F). These results suggest that a multitude of input signals are necessary for proper clock regulation and converge on the regulation of *CCA1*. This is in line with the critical role of CCA1 in resetting the clock, as well as a number of recent reports indicating that the circadian clock function is modulated by many signals in addition to light and temperature (Haydon et al., 2013; Hong et al., 2013; Pruneda-Paz and Kay, 2010). The increased regulatory complexity suggested by our results would improve network robustness while providing fine-tuning capability and ultimately would allow optimal synchronization of endogenous rhythms with the periodic changes in external environmental conditions. Further characterization in *Arabidopsis* seedlings of one of the TFs identified in our HT-Y1H screen, FBH1, confirmed that this is a regulator of *CCA1*. FBH1 binds preferentially to an E box-like motif (CACTAG) centered 405/406 nt upstream of the coding sequence start site. This motif contains only one mismatch with the recently published E box motif (CACTTG) bound by FBH1 on the promoter of the flowering regulator CONSTANS (Ito et al., 2012). Interestingly, our screen did not indicate any binding activity for FBH2 (AT4G09180), a homolog of FBH1, suggesting that the E box-like motif found in the *CCA1* promoter may be preferentially targeted by FBH1. Alternatively, the binding of FBH1/FBH2 may be enhanced by the presence of a cluster containing three tandem E box motifs within the *CO* promoter. Although FBH1 was described as a *CO* activator, our experiments suggest that it functions as a repressor of *CCA1*. It is possible that the different promoter contexts at the *CCA1* and *CO* loci are responsible for this discrepancy (Gordân et al., 2013). Screening the *CO* promoter with the expanded TF ORFeome presented here could reveal insightful differences between the potential regulators for both genes. Interestingly, the clock phase advance observed in *FBH1* overexpression lines is consistent with a similar phenotype observed in overexpression lines for a recently described *CCA1* repressor (i.e., CHE) (Pruneda-Paz et al., 2009). Although more work is needed to further confirm the potential *CCA1* promoter interactome suggested by the experiments in *Arabidopsis* protoplasts, the overall results suggest that multiple plant-signaling modules fine-tune the clock function through some of the TFs uncovered in this work.

Regulation of gene expression relies on the coordinated interaction of specific combinations of TFs bound to gene promoter regions. Globally, all TF-DNA interactions determine the basic architecture of gene regulatory networks that control fundamental aspects of cellular and organismal functions. This work introduces a valuable community resource for investigating the regulatory landscape of any *Arabidopsis* gene. Moreover, it provides the foundation for developing approaches that will ultimately reveal the molecular circuitry that regulates plant transcriptomes.

EXPERIMENTAL PROCEDURES

Details regarding the materials and methods used in this work are available in the Supplemental Experimental Procedures.

Construction of a Gateway-Compatible *Arabidopsis* TF Collection

Arabidopsis TF coding sequences were PCR amplified and cloned in the pENTR/D Gateway donor vector (Life Technologies). All clones included in the collection (1,956) were sequence validated (Table S2).

Y1H Screens with a Genome-Wide TF Collection

To generate a TF collection suitable for Y1H screens, all TF coding sequences were transferred from pENTR/D to pDEST22 (Life Technologies) via Gateway recombination-based cloning (Life Technologies). In pDEST22-TF clones, the GAL4-activation domain is fused to the N terminus of the TF. The pDEST22-TF library was arrayed in a 96-well format and transformed into the YU yeast strain (generated in this work) as previously described (Walhout and Vidal, 2001). For Y1H screens, the resulting yeast collection was indexed into 384-well plates. In this format, YU yeast TF strains (MATα) were mated with YM4271 yeast strains (MATa) carrying chromosomally integrated *CCA1::lacZ* reporter constructs. After mating, diploid cells were enriched and the β-galactosidase activity for each well was determined as described elsewhere (Yeast Protocols Handbook, Clontech), but with modifications that allowed the assay to be performed in 384-well plates using a robotic platform (Biocel1200; Agilent Technologies). β-galactosidase activities were normalized to the average value obtained for control wells in which the pDEST22-TF plasmid was replaced by the pEXP-AD (empty pDEST22) control plasmid (control normalized). The binding cutoff values for each *CCA1* promoter fragment were set at 4 and 8 SDs above the mean control-normalized value obtained for all control wells. To compare results across all *CCA1* promoter fragment screens, control-normalized values were normalized to the cutoff value for each *CCA1* promoter fragment data set (cutoff normalized). A compiled list of all *CCA1* promoter-interacting candidates was generated using cutoff-normalized values (Table S8).

Isolation and Transformation *of Arabidopsis* Protoplasts

Candidate TF coding sequences were transferred from pENTR/D to the TF-overexpression vector pCsVMV-GW generated in this work using the pCsVMV-PP2C-AmiR vector backbone (Kim and Somers, 2010) via Gateway

recombination-based cloning (Life Technologies). *Arabidopsis* protoplasts were isolated from 4- to 5-week-old Columbia ecotype (Col-0) seedlings using a procedure adapted from previously published protocols (Kim and Somers, 2010; Wu et al., 2009) and aliquoted in white 96-well plates (Evergreen Scientific). The protoplasts were cotransformed according to a previously published procedure (Wehner et al., 2011), with a TF-overexpression (or control) vector (5 μg per 4 kb) and the reporter vector pOmegaCCA1-LUC_SK+ (Kim and Somers, 2010) (1 μg/4 kb) that contains a *CCA1::LUC*+ reporter construct. After plasmid cotransformation, the cells were resuspended in a solution containing 5% fetal bovine serum (Sigma) and 50 μM luciferin (Biosynth) as described previously (Kim and Somers, 2010), and 96-well plates were covered with a transparent plastic lid. Bioluminescence imaging was performed as described below.

Plant Materials and Growth Conditions

Arabidopsis thaliana (*Arabidopsis*) seedlings used in this work were from the Columbia ecotype (Col-0). The CCA1::LUC+ reporter line was previously described (Pruneda-Paz et al., 2009). To generate *FBH1* overexpression lines, the pENTR/D-FBH1 construct (plate U21 well D01; Table S2) was used to transfer the *FBH1* coding sequence into the pB7WG2 binary vector (Karimi et al., 2002) via Gateway recombination-based cloning (Life Technologies). Similarly, the FBH1-YFP tag overexpression vector pE104-FBH1 was created by transferring the FBH1 coding sequence from pENTR/D-FBH1 into the pEarley104 binary vector (Earley et al., 2006). These binary vectors were transferred into the *Arabidopsis* CCA1::LUC+ line by *Agrobacterium*-mediated transformation, as described previously (Zhang et al., 2006). Unless otherwise stated, plants were grown in Murashige-Skoog medium (1.5% agar) supplemented with 3% sucrose under 12 hr light (70 $\mu mol \times m^{-2} \times s^{-1}$)/12 hr dark cycles (LD) at 22°C. ChIP assays were performed as previously described (Pruneda-Paz et al., 2009).

Bioluminescence Detection and Data Analysis

Seeds were stratified for 2–3 days at 4°C and grown in Murashige-Skoog medium (1.5% agar) supplemented with 3% sucrose under 12 hr light (70 $\mu mol \times m^{-2} \times s^{-1}$)/12 hr dark cycles; LD) at 22°C. After 7 days, plates were transferred to constant light (70 $\mu mol \times m^{-2} \times s^{-1}$; LL) and sprayed with 1 mM luciferin (Biosynth), and the emitted luminescence was analyzed every 2.5 hr for 5 days using a digital CCD camera (Hamamatsu). Similarly, transformed *Arabidopsis* protoplasts were incubated in LD for 36 hr (one light and two dark periods), transferred to LL, and imaged as described above. Images were processed using Metamorph imaging software (Molecular Devices) and data were analyzed by fast Fourier transform-nonlinear least squares (FFT-NLLS) (Plautz et al., 1997) using the interface provided by the Biological Rhythms Analysis Software System (BRASS, available at http://millar.bio.ed.ac.uk) (Southern et al., 2006).

Resource Distribution

Both clone collections, pENTR/D-TF and pDEST22-TF, are available at the *Arabidopsis* Biological Resource Center (ABRC; http://abrc.osu.edu).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.06.033.

AUTHOR CONTRIBUTIONS

J.L.P.-P., G.B., S.E.K., and S.R. generated and validated pENTR-TF and pDEST22-TF clones. J.R.E. and M.G. provided clones for the initial library construction. J.L.P.-P. and G.B. optimized and performed Y1H screens. K.B. performed protoplast transient assays. C.J.D. performed bioinformatic analyses. J.L.P.-P. and D.H.N. performed FBH1 characterization. J.L.P.-P. and G.B. compiled and analyzed data for all experiments. J.L.P.-P., G.B., and S.A.K. analyzed results and wrote the manuscript. All authors edited the manuscript.

ACKNOWLEDGMENTS

We thank Elsebeth Kolmos, Marian Nohales, and Sabrina Sanchez for critical readings of the manuscript; Robert Schmitz for helpful discussions and technical tips; Dr. David Somers for providing the pOmegaCCA1-LUC_SK+, pCsVMV-PP2C-AmiR, and pCsVMV-AmiR plasmids; Kim Jeongsik for advice on the protoplast isolation protocol; and Leann Ortmann and Jasmine King for technical assistance. J.R.E. is a Howard Hughes Medical Institute and Gordon and Betty Moore Investigator (GMBF3034). The research reported in this publication was supported by the National Institute of General Medical Sciences, NIH, under grants R01GM056006, R01GM067837 and RC2GM092412 to S.A.K. and R01GM056006 to J.L.P.-P. as a coinvestigator.

Received: April 22, 2014
Revised: May 28, 2014
Accepted: June 19, 2014
Published: July 17, 2014

REFERENCES

Arabidopsis Interactome Mapping Consortium (2011). Evidence for network evolution in an Arabidopsis interactome map. Science *333*, 601–607.

Arnosti, D.N., and Kulkarni, M.M. (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? J. Cell. Biochem. *94*, 890–898.

Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X., et al. (2009). Diversity and complexity in DNA recognition by transcription factors. Science *324*, 1720–1723.

Bassel, G.W., Gaudinier, A., Brady, S.M., Hennig, L., Rhee, S.Y., and De Smet, I. (2012). Systems analysis of plant functional, transcriptional, physical interaction, and metabolic networks. Plant Cell *24*, 3859–3875.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Castrillo, G., Turck, F., Leveugle, M., Lecharny, A., Carbonero, P., Coupland, G., Paz-Ares, J., and Oñate-Sánchez, L. (2011). Speeding cis-trans regulation discovery by phylogenomic analyses coupled with screenings of an arrayed library of Arabidopsis transcription factors. PLoS ONE *6*, e21524.

Chang, K.N., Zhong, S., Weirauch, M.T., Hon, G., Pelizzola, M., Li, H., Huang, S.S., Schmitz, R.J., Urich, M.A., Kuo, D., et al. (2013). Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. Elife *2*, e00675.

Davidson, E.H. (2006). The Regulatory Genome: Gene Regulatory Networks in Development and Evolution (Oxford: Academic Press).

Deplancke, B., Dupuy, D., Vidal, M., and Walhout, A.J. (2004). A gateway-compatible yeast one-hybrid system. Genome Res. *14* (10B), 2093–2101.

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A.M., Grove, C.A., Martinez, N.J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J.S., Hope, I.A., et al. (2006). A gene-centered C. elegans protein-DNA interaction network. Cell *125*, 1193–1205.

Dobi, K.C., and Winston, F. (2007). Analysis of transcriptional activation at a distance in Saccharomyces cerevisiae. Mol. Cell. Biol. *27*, 5575–5586.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. *45*, 616–629.

Gaudinier, A., Zhang, L., Reece-Hoyes, J.S., Taylor-Teeples, M., Pu, L., Liu, Z., Breton, G., Pruneda-Paz, J.L., Kim, D., Kay, S.A., et al. (2011). Enhanced Y1H assays for Arabidopsis. Nat. Methods *8*, 1053–1055.

Gong, W., Shen, Y.P., Ma, L.G., Pan, Y., Du, Y.L., Wang, D.H., Yang, J.Y., Hu, L.D., Liu, X.F., Dong, C.X., et al. (2004). Genome-wide ORFeome cloning and analysis of Arabidopsis transcription factor genes. Plant Physiol. *135*, 773–782.

Gordân, R., Shen, N., Dror, I., Zhou, T., Horton, J., Rohs, R., and Bulyk, M.L. (2013). Genomic regions flanking E-box binding sites influence DNA binding

specificity of bHLH transcription factors through DNA shape. Cell Rep *3*, 1093–1104.

Guo, A., He, K., Liu, D., Bai, S., Gu, X., Wei, L., and Luo, J. (2005). DATF: a database of Arabidopsis transcription factors. Bioinformatics *21*, 2568–2569.

Haydon, M.J., Mielczarek, O., Robertson, F.C., Hubbard, K.E., and Webb, A.A. (2013). Photosynthetic entrainment of the Arabidopsis thaliana circadian clock. Nature *502*, 689–692.

Hens, K., Feuz, J.D., Isakova, A., Iagovitina, A., Massouras, A., Bryois, J., Callaerts, P., Celniker, S.E., and Deplancke, B. (2011). Automated protein-DNA interaction screening of Drosophila regulatory elements. Nat. Methods *8*, 1065–1070.

Hong, S., Kim, S.A., Guerinot, M.L., and McClung, C.R. (2013). Reciprocal interaction of the circadian clock with the iron homeostasis network in Arabidopsis. Plant Physiol. *161*, 893–903.

Ito, S., Song, Y.H., Josephson-Day, A.R., Miller, R.J., Breton, G., Olmstead, R.G., and Imaizumi, T. (2012). FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. Proc. Natl. Acad. Sci. USA *109*, 3582–3587.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. *7*, 193–195.

Kim, J., and Somers, D.E. (2010). Rapid assessment of gene function in the circadian clock using artificial microRNA in Arabidopsis mesophyll protoplasts. Plant Physiol. *154*, 611–621.

Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science *338*, 349–354.

Kummerfeld, S.K., and Teichmann, S.A. (2006). DBD: a transcription factor prediction database. Nucleic Acids Res. *34* (Database issue), D74–D81.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. *40* (Database issue), D1202–D1210.

Li, L., Ljung, K., Breton, G., Schmitz, R.J., Pruneda-Paz, J., Cowing-Zitron, C., Cole, B.J., Ivans, L.J., Pedmale, U.V., Jung, H.S., et al. (2012). Linking photoreceptor excitation to changes in plant architecture. Genes Dev. *26*, 785–790.

Nagel, D.H., and Kay, S.A. (2012). Complexity in the wiring and regulation of plant circadian networks. Curr. Biol. *22*, R648–R657.

Niwa, M., Daimon, Y., Kurotani, K., Higo, A., Pruneda-Paz, J.L., Breton, G., Mitsuda, N., Kay, S.A., Ohme-Takagi, M., Endo, M., and Araki, T. (2013). BRANCHED1 interacts with FLOWERING LOCUS T to repress the floral transition of the axillary meristems in Arabidopsis. Plant Cell *25*, 1228–1242.

O'Malley, R.C., and Ecker, J.R. (2010). Linking genotype to phenotype using the Arabidopsis unimutant collection. Plant J. *61*, 928–940.

Oliveri, P., Tu, Q., and Davidson, E.H. (2008). Global regulatory logic for specification of an embryonic cell lineage. Proc. Natl. Acad. Sci. USA *105*, 5955–5962.

Ou, B., Yin, K.Q., Liu, S.N., Yang, Y., Gu, T., Wing Hui, J.M., Zhang, L., Miao, J., Kondou, Y., Matsui, M., et al. (2011). A high-throughput screening system for Arabidopsis transcription factors and its application to Med25-dependent transcriptional regulation. Mol. Plant *4*, 546–555.

Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B., Gilmartin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R., et al. (2002). REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana. Comp. Funct. Genomics *3*, 102–108.

Plautz, J.D., Straume, M., Stanewsky, R., Jamison, C.F., Brandes, C., Dowse, H.B., Hall, J.C., and Kay, S.A. (1997). Quantitative analysis of Drosophila period gene transcription in living animals. J. Biol. Rhythms *12*, 204–217.

Pruneda-Paz, J.L., and Kay, S.A. (2010). An expanding universe of circadian networks in higher plants. Trends Plant Sci. *15*, 259–265.

Pruneda-Paz, J.L., Breton, G., Para, A., and Kay, S.A. (2009). A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. Science *323*, 1481–1485.

Puchta, H., and Fauser, F. (2014). Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J. *78*, 727–741.

Reece-Hoyes, J.S., Barutcu, A.R., McCord, R.P., Jeong, J.S., Jiang, L., MacWilliams, A., Yang, X., Salehi-Ashtiani, K., Hill, D.E., Blackshaw, S., et al. (2011). Yeast one-hybrid assays for gene-centered human gene regulatory network mapping. Nat. Methods *8*, 1050–1052.

Rey, G., Cesbron, F., Rougemont, J., Reinke, H., Brunner, M., and Naef, F. (2011). Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. PLoS Biol. *9*, e1000595.

Rhee, S.Y., Beavis, W., Berardini, T.Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., et al. (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res. *31*, 224–228.

Riaño-Pachón, D.M., Ruzicic, S., Dreyer, I., and Mueller-Roeber, B. (2007). PlnTFDB: an integrative plant transcription factor database. BMC Bioinformatics *8*, 42.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., et al. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science *290*, 2105–2110.

Shiu, S.H., Shih, M.C., and Li, W.H. (2005). Transcription factor families have much higher expansion rates in plants than in animals. Plant Physiol. *139*, 18–26.

Siggers, T., Chang, A.B., Teixeira, A., Wong, D., Williams, K.J., Ahmed, B., Ragoussis, J., Udalova, I.A., Smale, S.T., and Bulyk, M.L. (2012). Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. *13*, 95–102.

Southern, M.M., Brown, P.E., and Hall, A. (2006). Luciferases as reporter genes. Methods Mol. Biol. *323*, 293–305.

Ueda, H.R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., Iino, M., and Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. Nat. Genet. *37*, 187–192.

Underwood, B.A., Vanderhaeghen, R., Whitford, R., Town, C.D., and Hilson, P. (2006). Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations. Plant Biotechnol. J. *4*, 317–324.

Walhout, A.J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods *24*, 297–306.

Wehner, N., Weiste, C., and Dröge-Laser, W. (2011). Molecular screening tools to study Arabidopsis transcription factors. Front Plant Sci *2*, 68.

Weirauch, M.T., and Hughes, T.R. (2011). A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. Subcell. Biochem. *52*, 25–73.

Wu, F.H., Shen, S.C., Lee, L.Y., Lee, S.H., Chan, M.T., and Lin, C.S. (2009). Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. Plant Methods *5*, 16.

Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., et al. (2003). Empirical analysis of transcriptional activity in the Arabidopsis genome. Science *302*, 842–846.

Zhang, E.E., and Kay, S.A. (2010). Clocks not winding down: unravelling circadian networks. Nat. Rev. Mol. Cell Biol. *11*, 764–776.

Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W., and Chua, N.H. (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat. Protoc. *1*, 641–646.