Transcriptomic Analysis of Changes in Gene Expression During Flowering Induction in Sugarcane Under Controlled Photoperiodic Conditions

João Ricardo Vieira Manechini[1], Paulo Henrique da Silva Santos[2], Elisson Romanel[3], Michael dos Santos Brito[4], Maximiliano Salles Scarpari[5], Stephen Jackson[6], Luciana Rossini Pinto[2,5] and Renato Vicentini[1]*

[1] Laboratório de Biologia de Sistemas, Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil, [2] Departamento de Genética e Melhoramento de Plantas, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual de São Paulo (UNESP), Jaboticabal, Brazil, [3] Laboratório de Genômica de Plantas e Bioenergia (PGEMBL), Departamento de Biotecnologia, Escola de Engenharia de Lorena (EEL), Universidade de São Paulo (USP), Lorena, Brazil, [4] Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo (UNIFESP), São José dos Campos, Brazil, [5] Centro de Cana, Instituto Agronômico de Campinas (IAC), Ribeirão Preto, Brazil, [6] School of Life Sciences, The University of Warwick, Coventry, United Kingdom

Flowering is of utmost relevance for the agricultural productivity of the sugarcane bioeconomy, but data and knowledge of the genetic mechanisms underlying its photoperiodic induction are still scarce. An understanding of the molecular mechanisms that regulate the transition from vegetative to reproductive growth in sugarcane could provide better control of flowering for breeding. This study aimed to investigate the transcriptome of +1 mature leaves of a sugarcane cultivar subjected to florally inductive and non-inductive photoperiodic treatments to identify gene expression patterns and molecular regulatory modules. We identified 7,083 differentially expressed (DE) genes, of which 5,623 showed significant identity to other plant genes. Functional group analysis showed differential regulation of important metabolic pathways involved in plant development, such as plant hormones (i.e., cytokinin, gibberellin, and abscisic acid), light reactions, and photorespiration. Gene ontology enrichment analysis revealed evidence of upregulated processes and functions related to the response to abiotic stress, photoprotection, photosynthesis, light harvesting, and pigment biosynthesis, whereas important categories related to growth and vegetative development of plants, such as plant organ morphogenesis, shoot system development, macromolecule metabolic process, and lignin biosynthesis, were downregulated. Also, out of 76 sugarcane transcripts considered putative orthologs to flowering genes from other plants (such as *Arabidopsis thaliana*, *Oryza sativa*, and *Sorghum bicolor*), 21 transcripts were DE. Nine DE genes related to flowering and response to photoperiod were analyzed either at mature or spindle leaves at two development stages corresponding to the early stage of induction and inflorescence primordia formation. Finally, we report a set of flowering-induced long non-coding RNAs and describe their level of conservation to other crops, many of which showed expression patterns correlated against those in the functionally grouped gene network.

Keywords: sugarcane, flowering, photoperiodism, transcriptome, artificial induction

INTRODUCTION

Flowering is an essential part of the life cycle for angiosperms. It influences the plant's adaptation to different environments, vegetative development, biomass accumulation, and grain production (Hill and Li, 2016). Flowering is coordinated by a diversity of genes organized in an intricate network of five gene pathways: photoperiod, vernalization, autonomous, gibberellin, and age (Wellmer and Riechmann, 2010; Srikanth and Schmid, 2011; Yamaguchi and Abe, 2012; Song et al., 2015; Hill and Li, 2016). These pathways, coupled with soil quality, water supply, and temperature, can restrain, promote, or disrupt flowering (Jackson, 2009; Hong and Jackson, 2015; Brambilla et al., 2017). Genes and metabolic pathways involved with flowering in grasses and in the model plant *Arabidopsis thaliana* share some similarity.

As an important cash crop, sugarcane (*Saccharum* spp. hybrid) supplies most of the global demand for sugar and ethanol and is impacted by flowering, which is highly undesired in the field as the transition from the vegetative to the reproductive stage ceases plant vertical growth and consumes its energetic reservoir, leading to short and underdeveloped stalks as well as a drastic reduction in extractable sugars available for the industry, hence reducing productivity (Berding and Hurney, 2005). On the other hand, flowering is desirable for plant breeding as crosses require both parent plants to have emitted a well-formed inflorescence at the same time. The difficulty of achieving flowering synchronism hampers breeding as sugarcane cultivars show high variability for flowering time (Glassop et al., 2014; Melloni et al., 2015).

The genus *Saccharum* accounts for a variety of species spread across several countries in five continents and possesses a wide range of photoperiodical behaviors (Moore and Berding, 2013) and multiple critical day lengths (CDL) required for flowering. Sugarcane cultivars, for instance, are considered as intermediate day length plants (IDP), meaning that they can flower around the CDL with many levels of floral induction, requiring a critical photoperiod of approximately 12 h 55 min with daily reductions of 45 s in artificial photoperiod regimes to promote flowering, characterizing a quantitative short day behavior (Moore and Berding, 2013; Glassop et al., 2014; Melloni et al., 2015). Photoperiod facilities have been used successfully in this crop to induce flowering (Melloni et al., 2015; Hale et al., 2017), allowing the control of important external factors that affect flowering, such as temperature, humidity, and day length, thus enabling the simulation of ideal conditions for flowering induction.

Photoperiodic flowering is controlled by a subset of genes involved with photo-perception, circadian rhythm, and molecular long distance signaling (McWatters et al., 2001; Putterill and Varkonyi-Gasic, 2016). At the terminal point of the photoperiodic flowering pathway lies the florigen protein known as FLOWERING LOCUS T (FT), a phosphatidylethanolamine-binding protein (PEBP) expressed in the phloem companion cells following induction by the CONSTANS (CO) protein (Suárez-López et al., 2001; Torti et al., 2012; Wang et al., 2016; Xu et al., 2016). When expressed, the FT protein travels to the shoot apical meristem (SAM), where it interacts with FLOWERING LOCUS-D (FD), a bZIP transcription factor, switching the plant from vegetative to reproductive development by activating the floral meristem identity genes (Andrés and Coupland, 2012). For that to happen, *CO* needs to be induced by the GIGANTEA (GI) protein, which is itself repressed by the gene *CYCLING DOF FACTOR 1* (*CDF1*). Hence, the latter must be repressed by the combined action of genes *FLAVIN-BINDING, KELCH REPEAT, F-BOX 1* (*FKF1*), *GI*, and *PSEUDO-RESPONSE REGULATOR* (*PRR*) *5, 7*, and *9* (Huq et al., 2000; Imaizumi et al., 2005; Mizuno and Nakamichi, 2005; Seaton et al., 2015; Yuan et al., 2019). These *PRR* genes can be activated by expression of a MYB-related transcription factor called LATE ELONGATED HYPOCOTYL (LHY), which, in turn, is activated by photoreceptor transcription factor PHYTOCHROME INTERACTING FACTOR 3 (PHY3) while under red light exposure (Martínez-García et al., 2000; Mizuno and Nakamichi, 2005; Lu et al., 2009). Aside from genes expressing their respective proteins, there is also post-translational regulation caused by interactions with other RNA species, such as microRNAs (miRNA) and long non-coding RNAs (lncRNA) (Yamaguchi and Abe, 2012; Henriques et al., 2017; Zhu et al., 2017). In Arabidopsis, Schmid et al. (2003) show the role of the microRNA miR172 precursor gene family in the photoperiodic induced flowering, mediating the accumulation of *AP2*-related gene mRNA. Over the last decade, lncRNA research has grown rapidly since the discovery of their ability to modulate gene expression levels. However, most of their functions remain unknown due limitations of *in silico* detection (Budak et al., 2020).

Despite the importance of sugarcane flowering for the sugar-energy sector, a deeper understanding of the regulations responsible for the photoperiodic induction process is still scarce (Coelho et al., 2013, 2014; Glassop and Rae, 2019), and the large-scale investigation of differentially expressed (DE) genes of a cultivar under controlled conditions of photoperiodic induction, in fact, still needs to be reported. Studies have been conducted with data mined from the Sugarcane Expressed Sequence Tag (EST) Project (SUCEST) database in order to identify putative flower-specific genes in sugarcane via homology to *Arabidopsis thaliana* flower development genes and proteins (Figueiredo et al., 2001), specifically the multigene family of MADS-box transcription factors and the gene family *APETALA2* (*AP2*), a key gene for flower development that acts as a promoter of early floral meristem identity (Dornelas and Rodriguez, 2001). Most of the studies of gene expression related to sugarcane flowering were carried out using plant material collected in the field after floral induction, a biological material that does not reflect the period before flowering or of flowering induction itself (Papini-Terzi et al., 2005; Medeiros et al., 2016).

The RNA sequencing technique (RNA-Seq) established prominence in high-throughput analysis of transcriptomes for literally every organism, allowing the measurement of transcript expression levels much more accurately when compared with other methods, e.g., Northern blot, ESTs, and microarrays (Wang et al., 2009; Nonis et al., 2014; Das et al., 2020). The RNA-Seq

technique can also detect small sequence variations, such as SNPs and regulatory elements, such as non-coding and lncRNAs (Vicentini et al., 2012; Ilott and Ponting, 2013; Cardoso-Silva et al., 2014; Veneziano et al., 2016). In the case of sugarcane cultivars, having a huge, aneupolyploid, and highly complex genome (Garsmeur et al., 2018; Mancini et al., 2018), the use of the RNA-Seq technique is also beneficial for bypassing certain challenges regarding data assembling and analysis.

In the present work, we describe the transcriptomic profiling of a mature leaf during photoperiodic induction of flowering of a sugarcane commercial cultivar. We identify 7,083 DE genes, of which 5,623 showed significant identity to other plant genes. Nine of these genes, related to flowering and response to photoperiod, identified in our RNA-Seq experiment, were analyzed via RT-qPCR either at mature or spindle leaves at two development stages corresponding to the early stage of induction and inflorescence primordia formation. We also investigate the presence and conservation of photoperiodic induced/repressed lncRNAs that may integrate the regulation module for flowering in sugarcane.

MATERIALS AND METHODS

Plant Material and Photoperiodic Treatments

The IAC sugarcane cultivar IACSP96-7569 has a regular flowering behavior under artificial photoperiodic regimes (Melloni et al., 2015). This cultivar was vegetatively propagated by single bud chips, planted into boxes filled with substrate (Plantmax®) and placed in a greenhouse for 28 days. Plantlets were transferred to 3.8-L tree pots filled with equal amounts of clay soil, sand, and substrate (Plantmax®) and placed randomly in two rail carts of an automated photoperiod facility at the Centro de Cana – IAC at Ribeirão Preto, São Paulo state, Brazil. After 7 months, plants were submitted to two different photoperiod treatments (one for each rail cart), either a constant long day (LD) photoperiod of 13 h and 30 min for non-inductive photoperiodic treatments (NIPT), or a short day (SD) photoperiodic treatment of 12 h and 50 min shortened by 45 s per day as an inductive photoperiodic treatment (IPT) until inflorescence emergence. Once a week, +1 mature leaf (first leaf with visible dewlap), the spindle leaves (furled immature leaves), and the SAM were collected from three plants (biological replicates) in each photoperiodic treatments (six plants total) at ZT10 (**Supplementary Figure 1**). The +1 mature leaf and the spindle leaves collected from the same plant were immediately frozen in liquid nitrogen and later stored at −80°C until RNA extraction, and the corresponding SAM was fixed in FAA 50% (formalin, acetic acid, and ethyl alcohol) for histological sectioning to confirm the apical meristem developmental stage. The IPT SAM samples, corresponding to the seventh week (45 days after the beginning of the photoperiodic treatment) showed start of transition from vegetative to floral meristem through histological assessment (**Supplementary Figure 2**). Hence, the +1 mature leaf samples of this time point, in both IPT and NIPT,

were selected for RNA-Seq analysis to assess DE transcripts between IPT and NIPT.

RNA Extraction, Quantification, Sequencing, and Data Set Quality Control

Total RNA extraction was performed by using the PureLink RNA Mini Kit (Thermo Fisher Scientific®) according to the manufacturer's recommendations. RNA Integrity Number (RIN ≥ 8.0) and quantity were evaluated using a 2,100 Bioanalyser (Agilent®) spectrophotometer. Libraries were assembled according to mRNA-Seq Sample Preparation v2 kit (Illumina®) recommendations. Briefly, mRNA was isolated and annealed to primers for single-stranded cDNA synthesis. The mRNA template was removed, and the second cDNA strand was synthesized, generating the double-stranded cDNA (dscDNA). The dscDNA fragments were isolated and the cDNA fragments ligated to the sequencing adapters. Fragments of 100 bp were amplified through PCR. Finally, the cDNA extracted from six biological replicates (three IPT and three NIPT) collected in the seventh week were taken to the Central Laboratory for High Performance Technologies (LaCTAD – UNICAMP, Campinas, São Paulo state, Brazil) for sequencing in a HiSeq 2,500 (Illumina®) sequencer. The FastQC v0.11.7 (Andrews, 2010) and NGS QC Toolkit v2.3.3 (Patel and Jain, 2012) software were used for quality check, adapter/barcode removal, filtering, and trimming procedures. The raw sequence data were deposited in the NCBI SRA database with the accession number SRP302030.

Reference-Based Mapping and DE Analysis

Transcriptome mapping and gene expression analysis were conducted using BowTie2 v2.3.4.2 (Langmead and Salzberg, 2012) and RSEM v1.3.0 (Li and Dewey, 2011), respectively, with default RSEM parameters for BowTie2, and using a reference sugarcane transcriptome for transcript mapping. Count matrices obtained by RSEM for each library (expected counts, non-normalized) were submitted to DE analysis with a DESeq2 (Love et al., 2014) (p-value < 0.05) default script, and the fold change values were internally transformed to $\log_2$ scale.

Transcriptome Annotation by Local Alignment and Phylogenetic Analysis

Transcript annotation was performed using BLAST v2.7.1+, retrieving information through similarity analysis with related organisms through a subtractive approach (**Supplementary Figure 3**). First, we chose four closely related grass protein databases: *Sorghum bicolor* v3.1.1 (McCormick et al., 2018), *Zea mays* v3 (Schnable et al., 2009), *Triticum aestivum* v2.2 (Mayer et al., 2014), and *Oryza sativa* v7.0 (Ouyang et al., 2007) and eudicot model plant *Arabidopsis thaliana* TAIR10 (Lamesch et al., 2012), available at JGI's Phytozome v12[1]. A BLASTx was performed from the taxonomically closest to the farthest organism from sugarcane with 20 hits per

---

[1] https://phytozome.jgi.doe.gov/pz/portal.html

transcript (e-value $\leq$ 1e-05). Hits and no hits were filtered at each BLASTx run. We considered as "hits" the transcripts that produced alignments that passed the filtering parameters. On the other hand, "no hits" are any transcript that either produced alignments that did not pass the filtering parameters or had no alignments produced during BLASTx. Hence, only "no hits" were carried further to the next BLASTx step with the next closest organism, following this order and filter parameters (alignment coverage and alignment identity, respectively): *S. bicolor* (60% and 50%), *Z. mays* (50% and 50%), *T. aestivum* (50% and 50%), *O. sativa* (50% and 50%), and *A. thaliana* (50% and 40%).

Phylogenetic analysis and putative orthology inference were performed by a pipeline (Bottcher et al., 2013) consisting of a series of local and global alignments of target protein sequences taken from KEGG circadian rhythm pathway (McWatters et al., 2001) against our transcriptome and other protein databases (same used previously for *S. bicolor*, *O sativa*, and *A. thaliana*) for phylogenetic relationship inference (**Supplementary Table 1**). First, target sequences in protein FASTA format were tBLASTn against our transcriptome (identity > 40% and coverage > 50%) capturing the first 15 best hits for each target. Then, these best hits were BLASTp against sorghum, rice, and Arabidopsis protein databases, collecting the best 40 sequences (e-value < 1e-05) for a global multiple alignment done with the software MAFFT (Yamada et al., 2016) using default parameters. Finally, these alignments were clustered under maximum likelihood phylogeny analysis conducted by the software phyML (Anisimova and Gascuel, 2006) with a WAG plus gamma substitution model and aLRT test. Results are phylogenetic trees, depicting the target sequences from sorghum, rice, Arabidopsis, and their respective sugarcane transcripts, which show statistically robust evidence for orthology inferring between sugarcane and other plant genes.

Quantitative Real-Time PCR (RT-qPCR) Assessment of Target Genes

Nine DE transcripts detected in the RNA-Seq experiment and involved in photoperiod response and flowering time were selected for relative gene expression analysis by quantitative real-time PCR (RT-qPCR). These transcripts not only were evaluated at the seventh week (corresponding to RNA-Seq experiment; 45 days of photoperiodic treatment) but also at the 13th week (plants with visible floral primordia formation, after 86 days of treatment; **Supplementary Figure 2**) in both mature and spindle leaves. Oligonucleotide primer pairs (**Table 1**) were designed from the respective transcript sequences by using PrimerQuest® Tool[2] adopting as conditions primer length from 17 to 22 bp; GC content from 35% up to 65%; melting temperature between 59°C and 65°C and amplicon size between 100 and 250 bp. Primer quality was assessed with NetPrimer[3] and primer pair efficiency by using the software LinRegPCR v7.5 (Ruijter et al., 2009). The cDNA synthesis was performed with the QuantiNova

[2]www.idtdna.com/SciTools

[3]www.premierbiosoft.com/netprimer

TABLE 1 | Sugarcane putative photoperiodic response and flowering time genes selected for gene expression analysis by RT-qPCR using mature and spindle leaves collected in the seventh and 13th weeks.

Gene (symbol)	Gene name	Primer sequence (5′- 3′)	Amplicon size (bp)
ScAGL7	*AGAMOUS LIKE-7*	F: GACGGTTCAGGCTCAGATT	152
		R: GCTTTAATGAACGCACACCTC	
ScAGL12	AGAMOUS LIKE-12	F: GAGATGGGCTATTCCTTCTGAC	149
		R: CTCCTGAAGGGCTATGGTTTAT	
ScCDF2	CYCLING DOF FACTOR 2	F: CTGTGATGGTGCCAGGTAAA	147
		R: GCACAAGTGGGTATGGAAATG	
ScCDF3	CYCLING DOF FACTOR 3	F: TCAGGTTTCGACTGGAATGG	151
		R: AAGGAGATGAGAAGGCAGAAAG	
ScEID1	EID1-like 1	F: TTCTGAGGACACAAAGGAAGAG	166
		R: CAAAGAGAAAGGCAGCTAGGA	
ScLHY	LATE ELONGATED HYPOCOTYL	F: GTGTCTCTCCACACAGAGTTAAA	161
		R: TTGTCCGCATCTACATCACTAC	
ScPRR1	PSEUDO-RESPONSE REGULATOR 1	F: CTCAAGCACATACACCACCA	153
		R: ATGCCGATGACCACACATT	
ScPRR5	PSEUDO-RESPONSE REGULATOR 5	F: ACAGAAGCAGAAACTGACTCG	102
		R: CCTTCAGTCTTACCAGTCCAAT	
ScPRR7	PSEUDO-RESPONSE REGULATOR 7	F: CAGTGGCAGTGGAAGTGAAA	149
		R: CATTGAGTCCGACACTGAAGTC	
ScUBQ1*	UBIQUITIN 1	F: AGCCTCAGACCAGATTCCAA	110
		R: AATCGCTGTCGAACTACTTGC	
ScTUB*	TUBULIN	F: CTCCACATTCATCGGCAACTC	237
		R: TCCTCCTCTTCTTCCTCCTCG	

Reference genes.

Reverse Transcription kit (QIAGEN Strasse 1, 40724 Hilden, Germany). The RT-qPCR assays were performed using a Bio-Rad IQ5 machine. The reaction was conducted in a final volume of 10 µl containing 5 µL of SYBR Green 2x from GoTaq® qPCR Master Mix Kit, (Promega, United States), 3 µL of cDNA (1:20 dilution), and primer pairs at their respective adjusted concentration. Amplification conditions were 95°C for 3 min, followed by 40 cycles of 10 s at 95°C and 30 s at 60°C, followed by a melting curve from 55°C to 95°C. For the RT-qPCR of the target genes, the three biological replicates sampled in each photoperiodic treatment (IPT and NIPT) were performed in triplicate (i.e., three technical replicates) and technical duplicates adopted for the primer efficiency and optimization analyses. Sugarcane *Ubiquitin1* (*ScUBQ1*) and *Tubulin* (*ScTUB*) were used as normalizers (**Table 1**). Relative expression data and statistical analysis were performed using the software REST 2009 (Pfaffl et al., 2002) with 2,000 iterations and differences considered significant when $P < 0.001$, $P < 0.01$, and $P < 0.05$. The IPT was considered "treated" and NIPT considered "untreated" for the purpose of expression calculations in the software.

Gene Ontology Terms Enrichment and Functional Annotation Analysis

Annotation and mapping of Gene Ontology (GO) terms were made with OmicsBox v1.1.164 (BioBam Bioinformatics, 2019) using RefSeq non-redundant release 49 database (Pruitt et al., 2012) to annotate the sugarcane transcriptome. For mapping of GO terms, the GOA database v2019.08 was used. Gene set enrichment was conducted with a Fisher's test with detailed results for all GO terms for biological processes. Gene sets had a maximum of 4,000 results with a false discovery ratio filter of 0.25. Enrichment networks were rendered with the software Cytoscape (Shannon et al., 2003) with ClueGO (Bindea et al., 2009) using annotated Arabidopsis gene names taken from BLASTx essays. Tree maps were created using REVIGO (Supek et al., 2011). Functional enrichment analysis were conducted with MapMan (Schwacke et al., 2019) using *A. thaliana* mapping "X4 Araport11 R1.0" and pertinent pathways related to flowering, photoperiod, and photosynthesis (downloaded from: https://mapman.gabipd. org/mapmanstore). Sugarcane transcripts annotated against the Arabidopsis TAIR10 database (Lamesch et al., 2012) were used to create an experimental file containing only DE transcripts from sugarcane converted to *A. thaliana* names.

Analysis of lncRNA

DE lncRNAs (lncRNAs) were identified *in silico* in two steps: First, we classified all "no hit" DE transcripts ($n = 1,460$) identified in our pipeline using CPAT (Wang et al., 2013) to verify the coding and ncRNA probabilities for "no hit" sequences. The program was trained with 400 randomized lncRNAs and 400 coding genes from *Zea mays*, and the threshold to consider ncRNA was > 0.3. For the non-coding set, we aligned transcripts against the *S. bicolor* genome using Sim4 (Florea et al., 1998). Second, we classified the lncRNA data set for conservation between other species by using data available in GreeNC (Gallart et al., 2016), which annotates lncRNAs present in genomes available in the Phytozome database, and BLASTn (e-value 1e-20) against the following data: *S. bicolor* ($n = 5305$), *Z. mays* ($n = 18,110$), *Setaria italica* ($n = 3492$), *O. sativa* ($n = 5237$), *Brachypodium distachyon* ($n = 5584$), *T. aestivum* ($n = 38,820$), and *A. thaliana* ($n = 3008$).
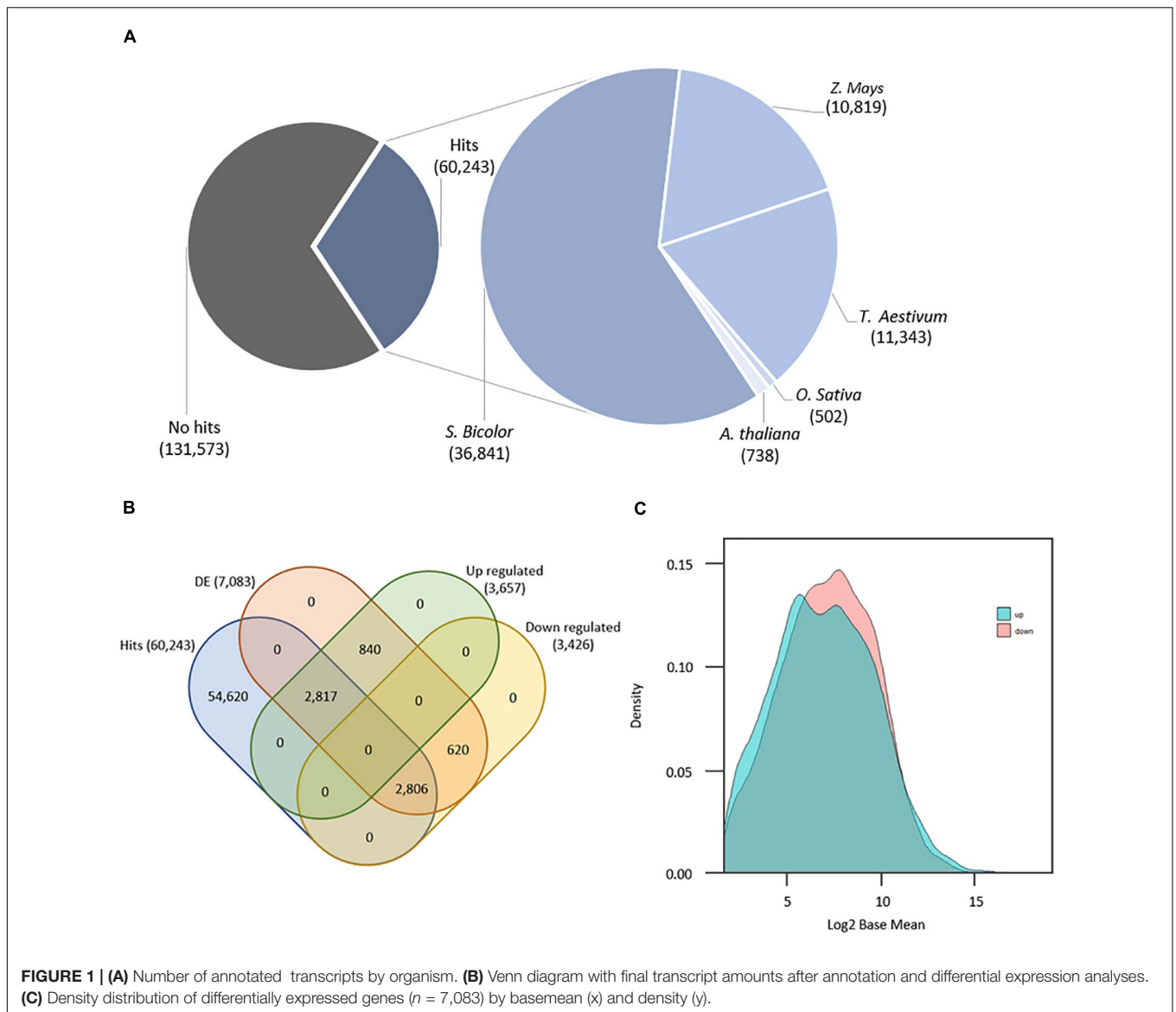
RESULTS

Quality Control of Data, Transcriptome Reference-Based Mapping, Gene Expression Analysis, and Annotation

A total of 12 libraries (6 biological samples × 2 technical replicates), encompassing 615,794,012 paired end reads with 100 base pairs (PHRED Score ≥ 30) were produced by RNA-Seq. Quality control of data (i.e., adapter/barcode removal and trimming) resulted in 562,091,180 (91.3% of the total data) high-quality paired-end reads ranging from 70 to 95 base pairs and PHRED score ≥ 38. Transcriptome reference-based mapping resulted in 112,584 mapped transcripts with mean length of 914 base pairs (length $< 200 = 26,826$ transcripts; $> 1.000 = 61,129$

transcripts), 63,028 transcripts containing open read frames, and CG content of 48.9% (**Supplementary Table 2**). Subtractive transcriptome annotation found a total of 60,243 (31.4% of the whole transcriptome) good-quality annotated transcripts (hits): 36,841 annotated from S. *bicolor*, 10,819 from *Z. mays*, 11,343 from *T. aestivum*, 502 from *O. sativa*, and 738 from *A. thaliana* (**Figure 1A**). Considering a previous study with sugarcane transcriptome annotation (Cardoso-Silva et al., 2014), the present methodology showed an increased number of annotated sugarcane transcripts from *S. bicolor* proteins and reduction for *O. sativa* proteins as the former is closely related to sugarcane. We found 7,083 transcripts with statistically significant DE comparing induced to non-induced treatments. From these, 3,657 are considered upregulated (51.64%), and 3,426 transcripts are considered downregulated (48.36%). There are 2,990 transcripts showing fold change ($\log_2$) greater or equal to 0.5, and 667 DE transcripts with fold change of less than 0.5 and greater than 0. Downregulated transcripts with fold change less than or equal to -0.5 summed 2,559 with 867 transcripts showing fold change between 0.5 and 0, meaning that there are more DE transcripts either for up- or downregulation at higher values of fold change. We have 5,623 annotated DE genes, being 2,817 upregulated and 2,806 downregulated (**Figure 1B**). From these, 2,180 upregulated transcripts had fold change greater or equal to 0.5 with 637 having fold change between 0.5 and 0. Also, 1,987 downregulated transcripts showed fold change less than or equal to -0.5, and 819 showed fold change between -0.5 and 0. Finally, transcript density by base mean (i.e., the average value of the normalized counts divided by size factors) of all 7,083 DE transcripts revealed that the downregulated group of transcripts have higher differential expression values than the upregulated group (**Figure 1C**).

The phylogenetic putative ortholog inference returned 76 transcripts as putative orthologs from *A. thaliana*, sorghum, and rice circadian rhythm and flowering genes as follows (number of putative ortholog transcripts in parenthesis): *ScAP2* (1), *ScATC* (2), *ScCDF2* (1), *ScCDF3* (1), *ScCHE* (3), *ScCKA3* (2), *ScCKA4* (2), *ScCKB2* (3), *ScCKB3* (3), *ScCKB4* (1) *ScCOP1* (4), *ScCRY1* (4), *ScCRY2* (4), *ScFKF1* (1), *ScFT* (5), *ScGI* (5), *ScLHY* (4), *ScPFT1* (5), *ScPHYa* (4), *ScPRR5* (2), *ScPRR7* (6), *ScSOC1* (5), *ScTOC1* (1), and *ScZTL* (7) (**Supplementary Table 3**). Because we are dealing with sugarcane genes, we renamed our transcripts according to Gray et al. (2009). We detected differential expression for 21 of these transcripts (**Figure 2A**), of which 14 (*ScAP2*, *ScCKA3*, three *ScCOP1*, *ScCRY1*, two *ScLHY1*, three *ScPRR7*, *ScSOC1*, and two *ScZTL*) were considered upregulated (or induced) and seven (*ScCKB3*, *ScCRY2*, *ScFKF1*, *ScPFT1*, *ScPRR5*, *ScPRR7*, and *ScTOC1*) considered downregulated (or repressed). Interestingly, most of these DE putative ortholog transcripts are related to signaling of light perception (*ScCRY1*, *ScCRY2*, and *ScPFT1*) and to the clock central oscillator (*ScZTL*, *ScPRR3*, *ScPRR5*, *ScPRR7*, and *ScLHY1*) of the circadian rhythm pathway indicating that, in this stage of photoperiodic treatment, the plant is still responding to the photoperiodic stimuli. The remaining 55 detected putative ortholog transcripts showed no DE (**Figure 2B**). Genes from the

FIGURE 1 | **(A)** Number of annotated transcripts by organism. **(B)** Venn diagram with final transcript amounts after annotation and differential expression analyses. **(C)** Density distribution of differentially expressed genes ($n = 7,083$) by basemean (x) and density (y).
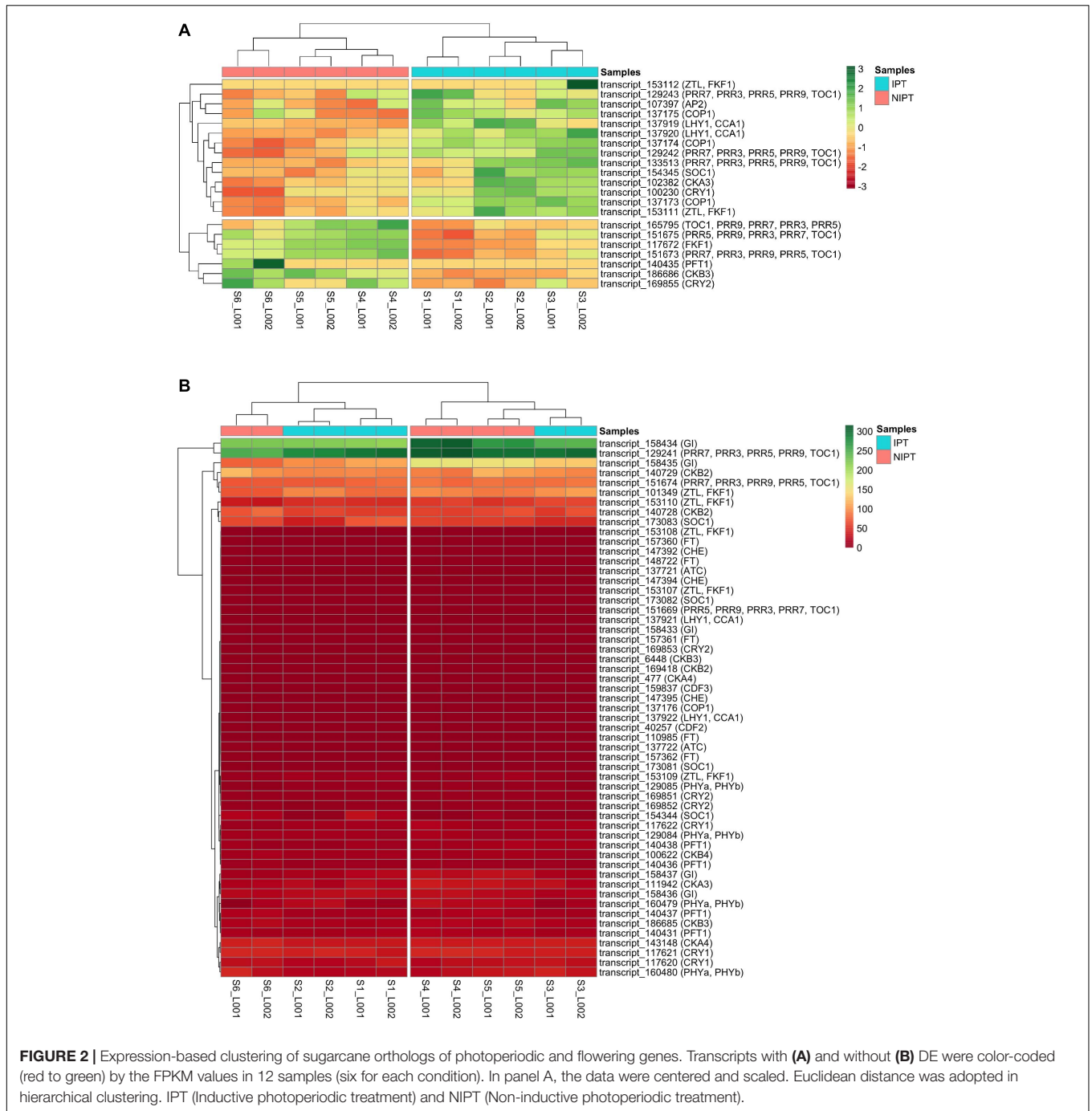
canonical flowering pathway, such as *ScFT* and *ScGI*, showed no DE, which was expected as these genes are more active at the start of the signaling stage in the mature leaf before signal translocation and consequent flowering (Kobayashi et al., 1999). These findings indicates that sugarcane, despite already being induced, still regulates the circadian clock pathway in response to photoperiod changes and that the time frame comprehending expression of FT proteins in sufficient quantity, its arrival, and hence accumulation at the SAM to the point of inducing cell differentiation from vegetative to reproductive, in sugarcane, can be short.

Assessment of Relative Gene Expression by RT-qPCR

Gene expression in induced leaves relative to non-induced leaves conducted via RT-qPCR showed that, for spindle leaves

collected in the seventh week, the DE genes with statistical significance were *ScCDF3* (0.57), *ScAGL7* (repressed, 0.45), and *ScAGL12* (1.84) (**Figure 3**). For the 13th week, the DE genes with a statistically significant difference were *ScPRR1* (4.66), *ScLHY* (2.57), *ScAGL12* (1.98), *ScEID1* (2.28), *ScCDF2* (2.89), *ScCDF3* (2.80), *ScPRR5* (1.83), *ScAGL7* (1.94), and *ScPRR7* (4.93), all of them induced. For mature leaves in the seventh week, the DE genes with a statistically significant difference were *ScPRR1* (repressed, 0.51), *ScCDF2* (2.44), and *ScCDF3* (1.75). Finally, for mature leaves of the 13th week, the genes that showed significant differential expression were *ScAGL12* (induced, 1.93), *ScPRR5* (1.56), and *ScPRR7* (induced, 1.90) (**Supplementary Table 4**).

The dissociation curves in spindle leaves for the seventh week all showed specific PCR amplification of the target gene. Efficiency (E) ranged from 1.37 to 2.03 and $R^2 = 0.99$. At the 13th week, the *ScPRR1* gene showed a bad dissociation curve, which

FIGURE 2 | Expression-based clustering of sugarcane orthologs of photoperiodic and flowering genes. Transcripts with **(A)** and without **(B)** DE were color-coded (red to green) by the FPKM values in 12 samples (six for each condition). In panel A, the data were centered and scaled. Euclidean distance was adopted in hierarchical clustering. IPT (Inductive photoperiodic treatment) and NIPT (Non-inductive photoperiodic treatment).

made its use impracticable. The rest of the genes showed a single peak, indicating specificity. Efficiency ranged from 1.68 to 1.97 and R^2 from 0.77 to 0.99. In the mature leaves, at the seventh week, all primers showed a single peak indicating amplification of a single product. Efficiency varied from 1.66 to 2.04, and R^2 was always maintained at 0.99. The dissociation curves in the mature leaf at the 13th week were also all reasonable, showing a single peak each. Efficiency was 1.68 to 2.2 and $R^2 = 0.99$ (**Supplementary Figures 4–7**).

Gene Ontology Terms Enrichment and Functional Annotation Analysis

Due the reasonably large amount of annotated DE genes detected in our data, we initially resort to large-scale analysis to interpret the overall situation of expressed genes despite the eventual loss of detail. This analysis was conducted individually for the groups of induced (upregulated) and non-induced (downregulated) genes. The present Gene Ontology enrichment returned an extensive set of terms for
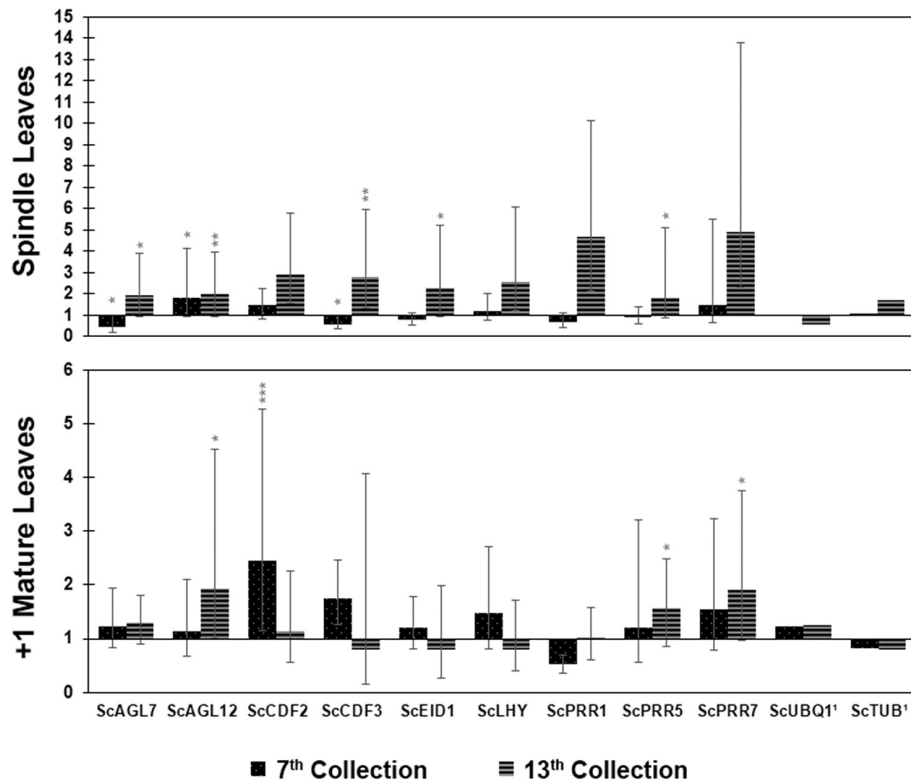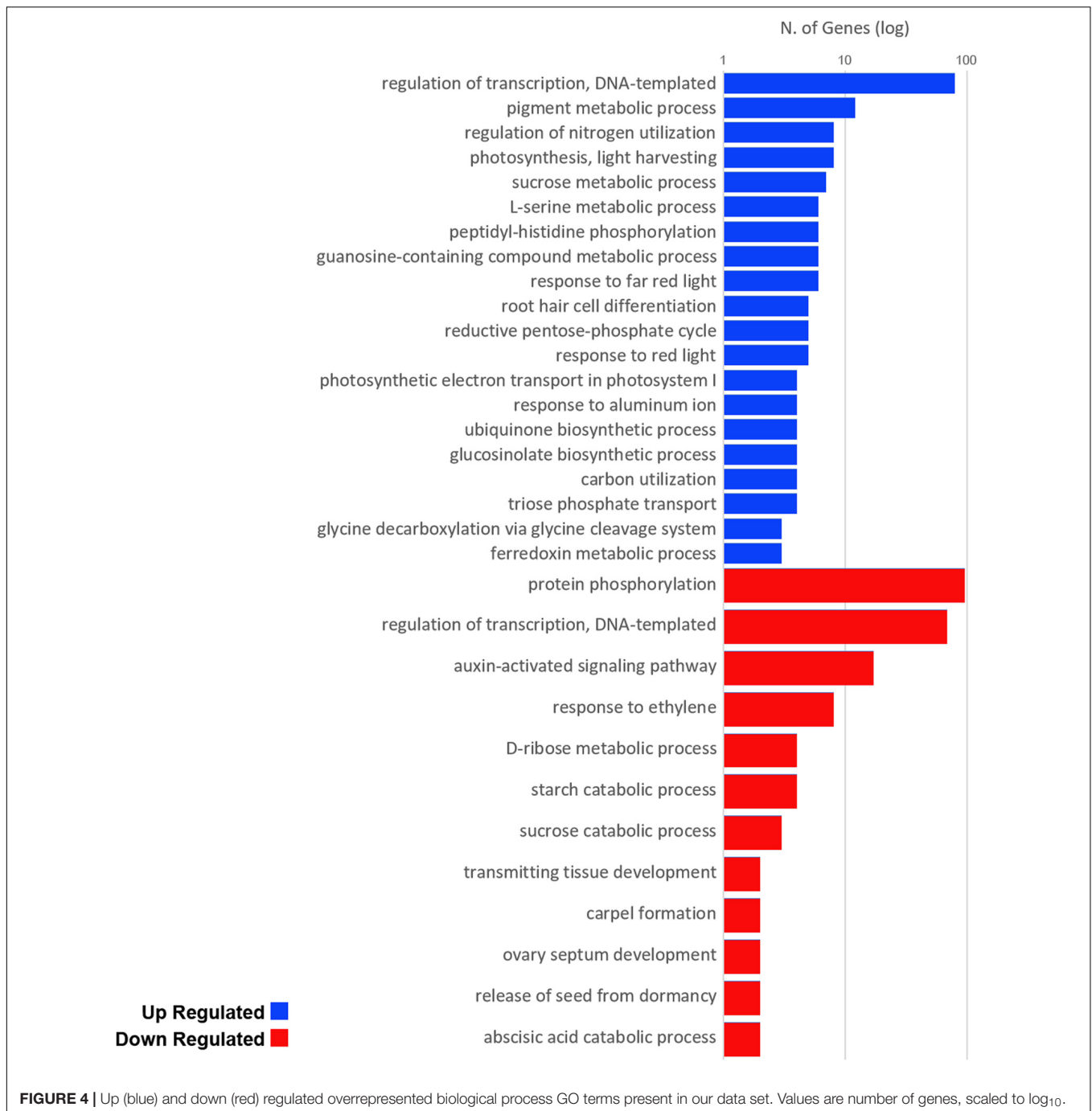
FIGURE 3 | Relative expression (induced/non-induced) of circadian clock genes at two time points following the start of the photoperiodic treatment; seventh week (45 days) and 13th week (85 days), from both spindle leaves and mature leaves. [1]Reference Genes. $P < 0.001$ (***); $P < 0.01$ (**); and $P < 0.05$ (*).
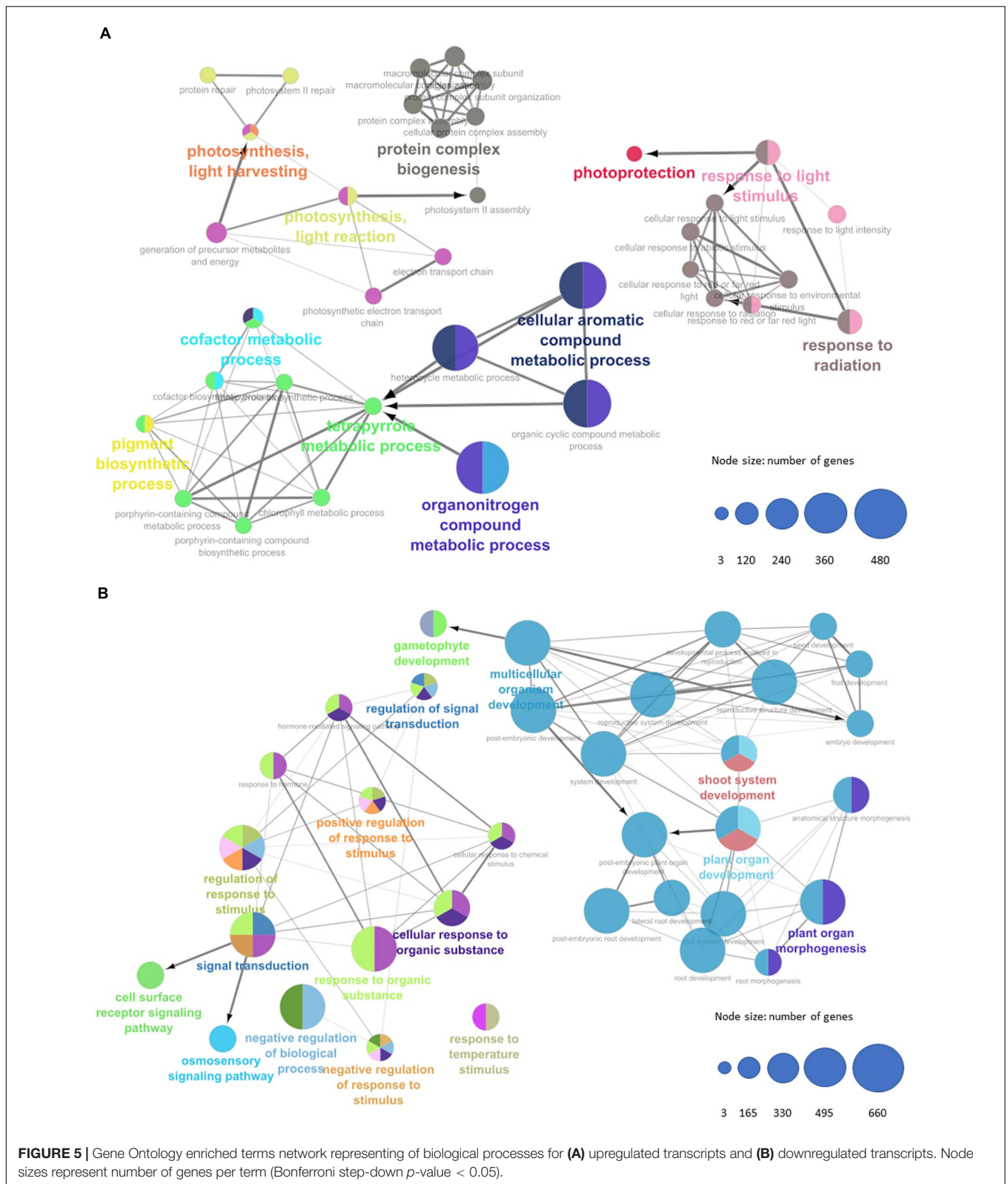
biological processes with some relation to the literature of flowering in monocots (Coneva et al., 2012; Minow et al., 2018). In the induced gene group, we found upregulation of overrepresented GO terms (**Figure 4**, with details provided at **Supplementary Table 5**) with important roles for plant structure development (pigment metabolic process, regulation of nitrogen, and carbon utilization), photorespiration, and energetic metabolism (sucrose, L-serine, and guanosine-containing compound metabolism, triose phosphate transport, ubiquinone biosynthesis, reductive pentose-phosphate cycle, glycine decarboxylation, and ferredoxin metabolism), and photosynthetic activities (photosynthesis/light harvesting, response to red and far-red light, and photosynthetic electron transport in photosystem I) as well as for other responses to external stimuli (glucosinolate biosynthesis and response to aluminum). On the other hand, we found downregulation of overrepresented GO terms related to vegetative growth (auxin-activated signaling pathway, response to ethylene, and transmitting tissue development), energetic metabolism (D-ribose metabolism, starch, and sucrose catabolism), reproductive structure formation (carpel and ovary septum development), and dormancy and seed germination regulation (release of seed from dormancy and abscisic acid catabolism). For a more general representation on how these terms relate, tree maps were generated for easier visualization of these overrepresented GO terms (**Supplementary Figure 8**).

We submitted all GO-enriched terms for functional clustering and annotation with ClueGO to uncover more significant term clusters with a higher degree of term specificity. GO networks revealed thousands of interconnected elements; hence, a p-value cutoff filter (<0.05) had to be implemented to capture more significant term clusters (**Figure 5**). Photosynthetic activity terms are still present at the induced group of genes after cutoff, showing photosynthesis, light harvesting, response to light stimulus, and response to radiation. This cluster induces the protein complex biogenesis through modulation of the photosystem II assembly. In turn, response to light stimulus, light intensity, and radiation seem to be inducing genes related to photoprotection. A third cluster of significantly enriched GO terms for the group of induced genes shows activity of a substantial number of genes (more than 480 genes) related to chemical compound and pigment metabolisms, such as tetrapyrrole and chlorophyl, respectively. As for the non-induced group of genes, two major clusters are portrayed: the first one (left side of **Figure 5B**) shows GO terms related to signal transduction, response to stimuli, response to chemical stimuli, and response to hormones, which could demonstrate some relation between repression of these response mechanisms during the change from the vegetative to the reproductive stage. Hence, as shown, this cluster may be responsible to modulate cell surface receptor signaling and osmosensory signaling pathways.

FIGURE 4 | Up (blue) and down (red) regulated overrepresented biological process GO terms present in our data set. Values are number of genes, scaled to $\log_{10}$.
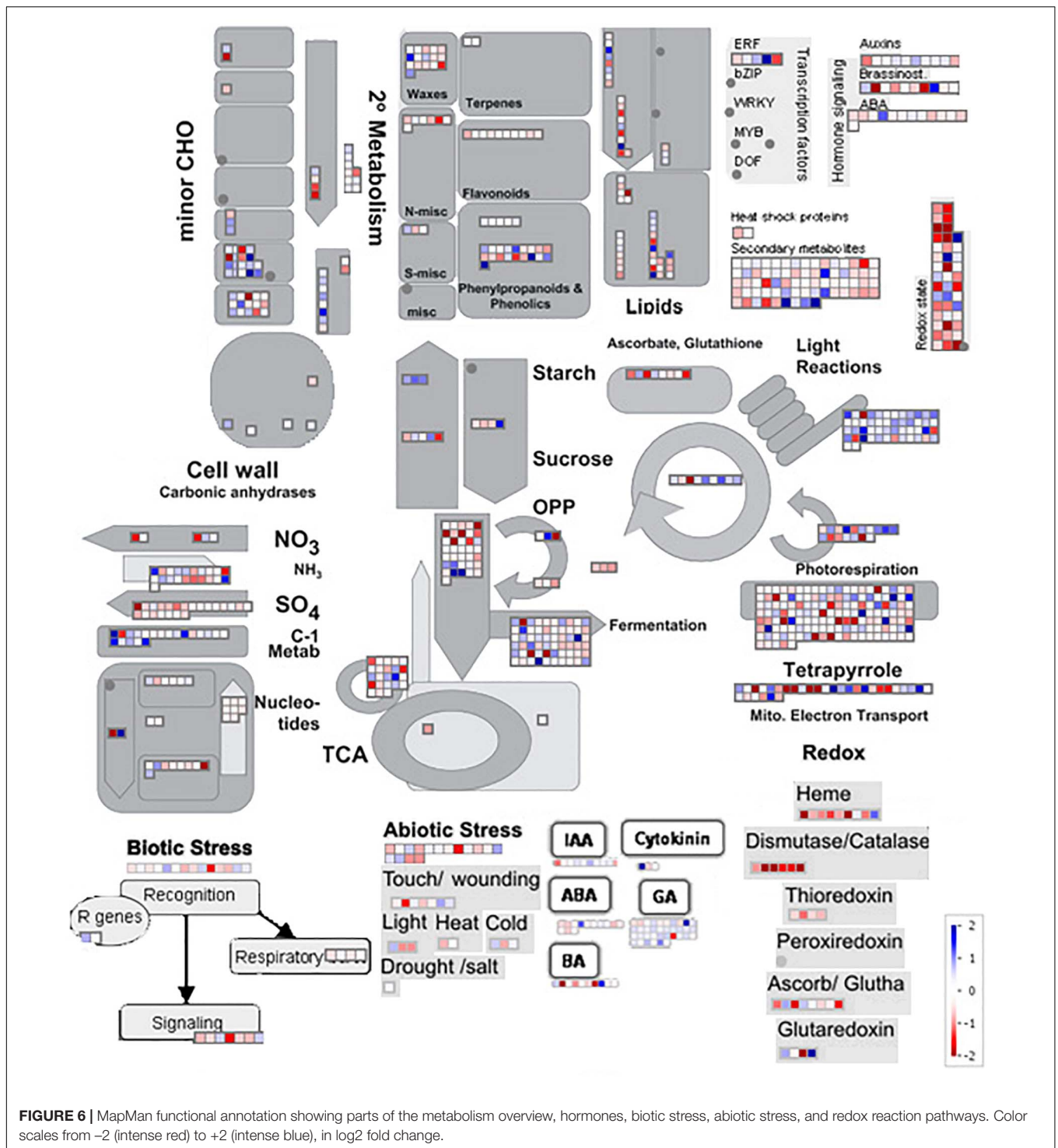
The second major cluster (right side of **Figure 5B**) depicts terms related to shoot system development, plant organ morphogenesis, and plant organ development. This cluster might also be involved in the modulation of the post-embryonic plant organ development and as well in gametophyte development, which is expected to be found in +1 mature leaves after floral induction. Furthermore, this can also be an indicator that the time frame between the induction of the FT protein and its critical accumulation in the SAM is short in sugarcane.

In general, the functional annotation made with MapMan shows expression changes in known gene pathways in plants (**Figure 6**). The effects of the flowering inductive photoperiod on sugarcane extends throughout several metabolic pathways and regulatory modules of the plant. There are significant gene expression alterations for some pathways, such as minor carbohydrate, starch/sucrose, lipids, and cell wall metabolisms, and hormones. Light reactions and photorespiration pathway alterations are predominant, and 2° metabolism, redox reactions, fermentation, tetrapyrrole

FIGURE 5 | Gene Ontology enriched terms network representing of biological processes for **(A)** upregulated transcripts and **(B)** downregulated transcripts. Node sizes represent number of genes per term (Bonferroni step-down *p*-value < 0.05).

metabolism and hormone signaling showed a considerable amount of DE genes. Interestingly, there are some small expression changes for biotic and abiotic stresses: signaling for
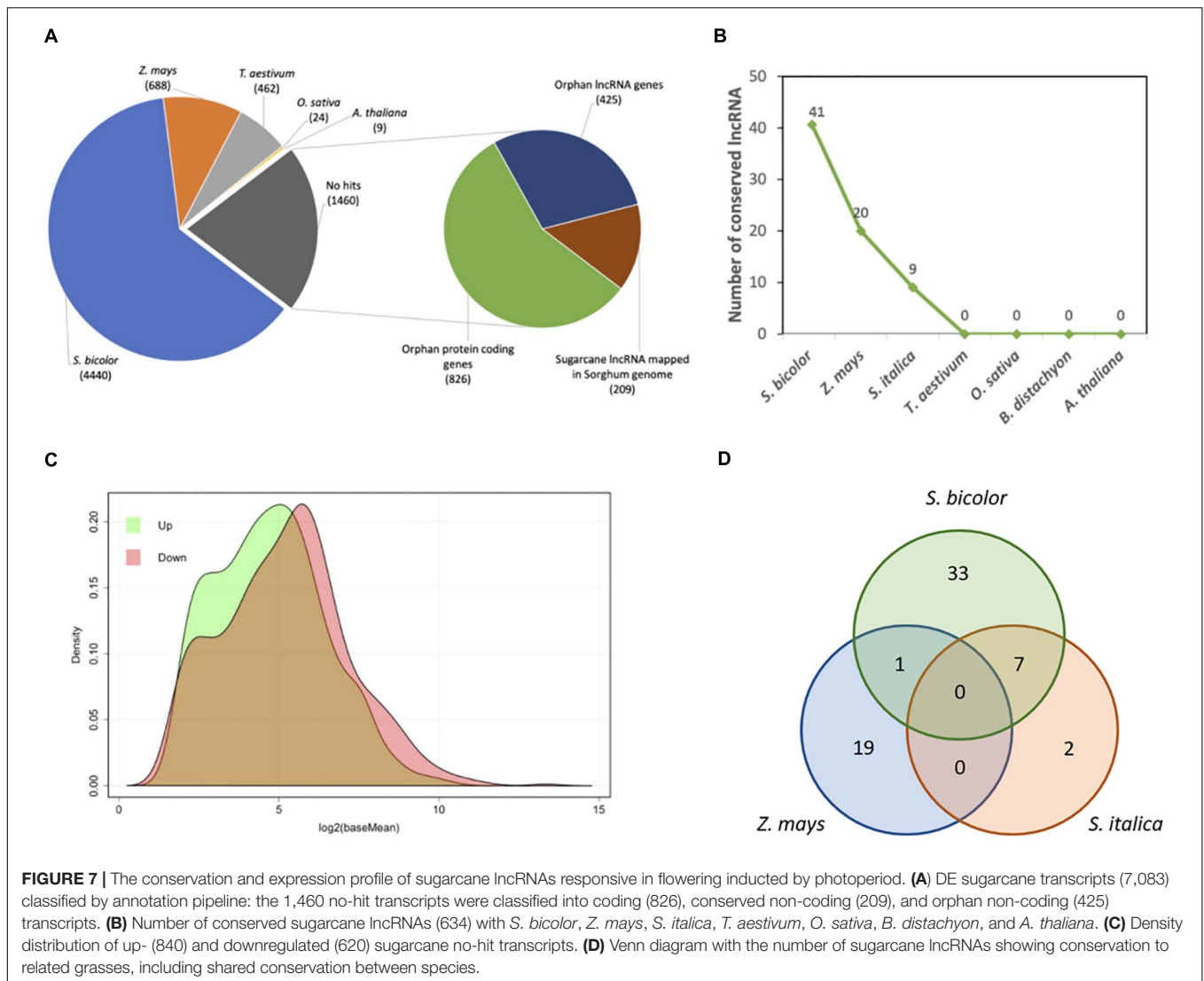
biotic stress is mostly downregulated, and light and cold biotic stress responses show mixed results though small (log2 fold change < 1 and >−1).

FIGURE 6 | MapMan functional annotation showing parts of the metabolism overview, hormones, biotic stress, abiotic stress, and redox reaction pathways. Color scales from –2 (intense red) to +2 (intense blue), in log2 fold change.

Sugarcane lncRNAs Responsive in Flowering Induction

We have also made an *in silico* analysis to identifying sugarcane flowering induction responsive lncRNAs. Such RNAs are characteristic for lacking or having small open reading frames and being longer than 200 nucleotides. A long non-coding

annotation of sugarcane floral induction transcriptome was performed over all 1,460 DE "no hits" transcripts. The BLASTn, Sim4, and CPAT were used in this pipeline (**Figure 7A**). Sim4 was used to find putative non-annotated genes/lncRNA (ncRNA), CPAT to verify the coding and ncRNA probabilities, and BLASTn to identify conserved sequences. After applying similarity filters, we found 634 sugarcane lncRNAs responsive to flowering

FIGURE 7 | The conservation and expression profile of sugarcane lncRNAs responsive in flowering inducted by photoperiod. **(A)** DE sugarcane transcripts (7,083) classified by annotation pipeline: the 1,460 no-hit transcripts were classified into coding (826), conserved non-coding (209), and orphan non-coding (425) transcripts. **(B)** Number of conserved sugarcane lncRNAs (634) with *S. bicolor*, *Z. mays*, *S. italica*, *T. aestivum*, *O. sativa*, *B. distachyon*, and *A. thaliana*. **(C)** Density distribution of up- (840) and downregulated (620) sugarcane no-hit transcripts. **(D)** Venn diagram with the number of sugarcane lncRNAs showing conservation to related grasses, including shared conservation between species.

induction (**Figure 7A**). From those, 62 (**Supplementary Table 6**) are shared with other grass species (**Figure 7B**). As some lncRNAs may be involved with flowering induction, these are relevant data for further studies (**Figure 7C**). The lncRNA annotation procedures reported a low level of sequence conservation against known sorghum or other grass transcripts (**Figure 7B**) but a high level of conservation in the sorghum genome (**Figure 7A**) was detected, showing that 209 transcripts could be unknown sorghum lncRNAs. Interestingly, there are seven conserved lncRNA sugarcane transcripts in common with *S. bicolor* and *S. italica* and only one with sorghum and *Z. mays* (**Figure 7D**). There are no conserved sugarcane transcripts in common with *Z. mays* and *S. italica*.

DISCUSSION

The results presented here start to shed some light on the changes in gene expression when sugarcane plants switch from vegetative to reproductive development following photoperiodic induction. This has not previously been done for sugarcane. We report more than 112,000 mapped sugarcane transcripts, of which 7,083 show significant DE in response to an IPT. Transcripts were annotated following a stringent protocol and robust statistical methods, aiming to find as much high-quality information as possible, which yielded more than 60,000 fully annotated transcripts. Also, we captured 76 transcripts as being putative real orthologs of flowering genes from different plant species in sugarcane with 21 being responsive to photoperiodic induction as well as 565 lncRNAs also responsive to the florally inductive photoperiod, which have many applications in further molecular studies of the subject.

Plants such as *Arabidopsis thaliana* (thale cress), *Glycine max* (soybean), *Oryza sativa* (rice), *Zea mays* (maize), *Sorghum bicolor* (sorghum), *Triticum aestivum* (wheat), and *Hordeum vulgare* (barley) have been intensively studied in the last decade in regard to their control of flowering (Matsubara et al., 2014; Song et al., 2015; Wolabu et al., 2016; Brambilla et al., 2017) and

are still used as models for investigating flowering pathways in related plant species. However, novel molecular functions and interactions can still be discovered in less well-studied organisms as in the case of sugarcane. Despite advances in flowering time regulation in grasses, such as *S. bicolor* (Lin et al., 2016), *Z. mays* (Coneva et al., 2012; Minow et al., 2018), *S. italica* (Doust et al., 2017), and *Brachypodium distachyon* (Higgins et al., 2010), the sugarcane photoperiod pathway still waits for gene identification and functional characterization.

Most florally induced genes are expressed in a tissue-specific manner. Once external stimuli are perceived by the right sensors, the pathway is activated to promote the transition from vegetative to reproductive development in plants (Kiełbowicz-Matuk and Czarnecka, 2014). Our real-time qPCR assays show that, in mature leaves collected in the seventh week (floral induction stage), the *ScPRR1* gene was repressed. In *Arabidopsis*, the *AtPRR* gene family consists of five members (*AtPRR9*, *AtPRR7*, *AtPRR5*, *AtPRR3*, and *AtPRR1/TOC1*) (Yamamoto et al., 2003) that are involved in the function of the circadian clock (Mizuno and Nakamichi, 2005). *AtPRR* transcripts start to accumulate after dawn in a specific order (*AtPRR9* – *AtPRR7* – *AtPRR5* – *AtPRR3* – *AtPRR1*) with an interval of 2 to 3 h each, forming a loop expression profile (Makino et al., 2001). Glassop and Rae (2019) found that, in the 24-h cycle, the *ScPRR1* gene in LD was highly expressed in spindle leaves and +1 mature leaf, reaching its maximum expression peak with about 10 h into the light regime, followed by a decrease after 2 h in the dark period. In the present study, with a similar time of collection (16 h 00, ZT10), *ScPRR1* appeared to be repressed in both spindle and mature leaves in the situation of shortening day length. In mature leaves at the 13th week, the *ScPRR* family genes that were significantly expressed were *ScPRR5* and *ScPRR7*, both induced and more expressed in the SD than the LD light regime. It should be noted that in the seventh week, the *ScPRR* feedback loop is at its end because it has a specific order of expression, and *ScPRR1* is the last to show a peak of expression. On the other hand, in the 13th week (floral primordium), the "loop" is at the halfway point because, in this sampling point, *ScPRR7* and *ScPRR5* are being expressed. In spindle leaves, at the 13th week, the *ScPRR7* and *ScPRR5* genes are significantly induced and follow the same pattern of expression as in mature leaves of the 13th week.

The *ScLHY* gene was significantly induced in spindle leaves of the 13th week as in the work by Glassop and Rae (2019). However, according to the authors, the peak of *ScLHY* expression in a 24-h cycle occurs mainly in the dark period, differing from our time of collection, which explains why its expression is at the beginning of the ascent in the present work. In *Arabidopsis*, AGAMOUS-LIKE 12 (*AtXAL1*) is a member of the MADS-box family, which is one of the key components in the floral induction and flowering development network. Tapia-López et al. (2008) suggest that *AtXAL1* is an upstream regulator for *AtSOC*, *AtFT*, and *AtLFY*. In sugarcane, *ScAGL12* was significantly induced in mature leaves collected in the 13th week and in spindle leaves of the seventh and 13th weeks. Tapia-López et al. (2008) also note via *in situ* hybridization of the *AtXAL1* mRNA that the gene is specifically expressed in *Arabidopsis* vascular tissue. There, in

the phloem's companion cells, the mRNA of the *AtFT* gene is expressed and translated, so the *AtFT* protein is translocated to the apical meristem, triggering flowering (Corbesier et al., 2007). Nonetheless, if *AtXAL1* acts upstream of the *AtFT* gene, this observation could only be valid for mature leaves in the 13th week because, in the spindle leaves, the vascular system is not yet fully developed. There is a redundancy between the *AtDof* (DNA-binding with one finger) transcription factors (*AtCDF1*, *AtCDF2*, *AtCDF3*, and *AtCDF5*) as negative regulators of *AtCO* at flowering time in *Arabidopsis* (Imaizumi et al., 2005; Fornara et al., 2009). The *ScCDF3* gene was significantly repressed in spindle leaves of the seventh week and mature leaves in the 13th week, inferring that the former tissue, when induced by photoperiod, is potentially inducing CONSTANS (hence FT) to promote flowering in the nearby SAM. Conversely, in seventh week mature leaves and in spindle leaves of the 13th week, flowering induction is expected to be halted. As for *ScCDF2*, there is induction in both tissues and times; however, there is a 2.5-fold increase in expression in seventh week mature leaves, inferring that *ScCDF2* is repressing flowering through repression of CONSTANS, as well as a threefold increase in expression of the gene in spindle leaves of the 13th week, in which flowering is not expected. This behavior of the *Dof* transcription factors seem to be somehow conserved between sugarcane and Arabidopsis despite the great evolutionary distance. However, the rice *Dof* family, comprising more than 30 elements (Li et al., 2009) act as flowering inductors. The *AtEID1* gene is related to a negative regulation through protein degradation (ubiquitin-dependent proteolysis) of *AtPhyA* dependent pathways (Büche et al., 2000) and appears significantly induced in spindle leaves at the 13th week (when light exposure is the lowest) although it is repressed in the seventh week. In Arabidopsis, the absence of *AtEID1* allows *AtPhyA* stability, responding as a far-red light receptor under continuous irradiation (Dieterle et al., 2001). For sugarcane, this relation between *ScEID1* and *ScPhyA* seems to be maintained.

In a recent study from our group (Santos et al., 2021), we tested reference genes against sugarcane DE orthologs of *AtPIL5*, a gene involved with seed germination, leaf senescence, gibberellin pathway, and light perception, and *AtLHP1* (aka *TFL2*, *TERMINAL FLOWER 2*, Feng and Lu, 2017), which is an epigenetic regulator of *FT* and *FLC* in Arabidopsis. It was found that, for mature leaves at the photoperiodic flowering induction stage (equivalent to this study's seventh week), *ScPIL5* is induced, and *ScLHP1* is repressed. Interestingly, the induction of *ScPIL5* may infer a different function of the gene between Arabidopsis and sugarcane because, in the former, *AtPIL5* (*AtPIF1*) represses chlorophyl synthesis in dark periods while suppressing hypocotyl elongation during light time, which is not observed in our Gene Ontology analysis (**Figure 5A**). Also, the repression of *ScLHP1* is related to diverse gene functions between Arabidopsis and sugarcane (Coelho et al., 2014) as, in the former, this gene is responsible for the maintenance of floral meristem identity during flowering, and for the latter, it may also have yet unknown functions for photoperiodic response of flowering in grasses.

To get a general overview of the effects of theinductive photoperiod in sugarcane, we performed large-scale Gene

Ontology enrichment terms analysis for DE annotated transcripts and assessed the overrepresented terms for biological processes (**Figure 4**). As expected, we found several processes and functions linked, directly or indirectly, to plant structure development, photorespiration, energetic metabolism, and photosynthetic activities. Although it is not possible to infer the exact moment of vegetative/reproductive transition, positive changes in the expression of genes involved in biological processes related to photoperiod and light responses indicates that these pathways are being induced. Upregulation of genes involved in pigment metabolism and photosystem I electron transport suggests that the plant may be working under some form of light-deprivation stress. Downregulation of genes involved in hormone pathways, such as the auxin, ethylene, and abscisic acid, may be a result of the change from vegetative to reproductive development and production of the associated reproductive structures. Interestingly, carpel formation, ovary septum development, and seed release from dormancy were also found in downregulated transcripts.

In the attempt to look closer at the large-scale GO term analysis, we conducted a network clustering of the enriched terms and selected the most significant clusters related to circadian rhythm and flowering. By doing so, it is possible to see hierarchies of GO terms at a higher level of detail (**Figure 5**). Response to light stimuli (light intensity and quality) upregulated together with photoprotection genes, suggesting once more that the plant may be under a kind of light-deprivation stress. Upregulation of light harvesting and light reaction processes can also be involved in the response of the plant to light deprivation, promoting repair of the photosystem II in the attempt to harvest more light in a scenario of decreasing daytime. These processes are also directly involved in the network for protein complex biogenesis and generation of energy, which suggests that the plant might be investing its energy reservoir in the increase of pigment biosynthesis through the tetrapyrrole metabolic pathway, which accounts for chlorophyll biosynthetic processes (**Figure 5A**). Henceforth, this data raises a question regarding the existence of crosstalk between the photoperiodic induction of flowering and responses to light-deprivation stress. The need for consistency of the flowering signal throughout the flowering stage in sugarcane is of utmost importance given that the interruption of light stimuli causes sugarcane to abort flowering, leading to floral reversion and malformed inflorescence (Glassop et al., 2014). On the other hand, a wide range of significant GO term clusters of downregulated transcripts is also observed (**Figure 5B**). Curiously, response to temperature stimulus is downregulated under IPT, whereas the temperature pathway is often combined with the photoperiod for flowering induction under the *FLC/SOC1/FUL* (*FLOWERING LOCUS C*, *SUPPRESSOR OF CONSTANS1*, and *FRUITFULL*) gene regulation module for eudicots (Bouché et al., 2016), suggesting that maybe there is a thinner interdependency between photoperiodic and temperature pathways for flowering in sugarcane. Downregulation of responses to chemical stimuli and hormones also constitutes a wide GO cluster, suggesting that, although the plant is repressing hormone pathways (i.e., auxin, ethylene, and ABA), there is a desensitization of cellular responses to such compounds, allied to downregulation of signal transduction in order to ensure the change from the vegetative development stage to the reproductive stage. Last, plant organ morphogenesis and gametophyte development are interestingly repressed under the inductive photoperiodic treatment. This may be related to halting vegetative shoot development following induction, which would precede floral organ initiation.

Functional annotation and metabolic pathway enrichment data supports the hypothesis that light-deprivation stress occurs during the photoperiodic induction of flowering sugarcane, together with desensitization of hormonal responses (**Figure 6**) of auxin, ABA, and 6-BenzylAdenine (BA). Moreover, Cytokinin and GA metabolism appear to be slightly upregulated. Abiotic stresses such as wounding, light, and cold responses, show small changes, but most of them seem to be repressed. As detected in the GO analysis, tetrapyrrole metabolism shows large changes, either induction or repression of some genes, which may infer accentuated modulation in pigment and aromatic compound metabolism. Photorespiration and light reactions also show big changes in gene expression, mostly induced, inferring that the plant is undergoing changes in how light (intensity and quality) is perceived, together with an increase in energy production.

Although our group have previously identified 1,446 sugarcane lncRNA expressed in leaves (Cardoso-Silva et al., 2014), this study reveals a set of 634 sugarcane lncRNAs DE under the photoperiodic condition. Furthermore, we show that 62 (approx. 10%) of these sugarcane transcripts share some level of conservation with gene *loci* from other grasses. Unfortunately, to date, most of them were not characterized, and a thorough study of this subset might prove invaluable in unveiling novel regulatory elements in grasses. Recently, a collection and rich resource of 6,510 lncRNAs were identified in Arabidopsis (Zhao et al., 2018). So far, few were functionally involved in flowering time regulation, such as *COLDAIR* and *COOLAIR* related to the vernalization-mediated epigenetic repression of a MADS-box *AtFLC* (Chekanova, 2015; Budak et al., 2020). Besides this, the overexpression of seven intronic lncRNA from Arabidopsis MADS-box genes activates the expression of their host genes (Liu et al., 2020). On the other hand, however, only FLC homologs antisense transcripts present in various grass species, called in *Brachypodium* spp. as *BdCOOLAIR1* and *BdCOOLAIR2*, were functionally characterized regulating *in cis* the rate of a MADS-box *BdODDSOC2* expression (Jiao et al., 2019). These data suggest further lncRNA physical localization and their co-expression pattern together with its host genes in control flowering time.

Future activities to better understand flowering induction of sugarcane under photoperiodic treatments might focus on the large-scale investigation of different time points of induction aiming to verify key genes activities, compared with already published gene regulatory networks from other plants, and how treatments modulate these networks. Data from circadian rhythm experiments in grasses are not abundant and would greatly contribute to the subject. Furthermore, integration of post-translational regulatory elements on expression data for gene

regulatory networks, such as lncRNAs, might provide a deeper understanding of flowering in sugarcane.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI, BioProject ID: PRJNA691914 and the following link: http://www.ncbi.nlm.nih.gov/bioproject/691914.

AUTHOR CONTRIBUTIONS

LP and RV: conceptualization and project administration. JM, PS, and RV: data curation and formal analysis. LP, SJ, and RV: funding acquisition and supervision. JM, PS, ER, MB, MS, SJ, LP, and RV: investigation. JM, PS, MS, LP, SJ, and RV: methodology. LP, MS, and RV: resources. JM, LP, and RV: visualization and writing – original draft preparation. JM, PS, ER, LP, SJ, and RV: writing – review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.635784/full#supplementary-material

Supplementary Figure 1 | Schema of the sugarcane leaf structure, emphasizing the spindle leaves (top red circle), and the first leaf from the top to the bottom of the stalk with clearly visible dewlap (bottom red circle) also known as +1 mature leaf (adapted from Van Dillewijn, C. Botany of sugarcane. Waltham: Chronica Botanica, 1952. 371). **(A)** sagittal section of the upper culm: peeled, **(B)** spindle leaves dissected from the central part of the upper culm, **(C)** shoot apical meristem dissected from the bottom part of the upper culm, and **(D)** +1 mature leaf with red arrow showing the dewlap.

Supplementary Figure 2 | Morphology of the shoot apical meristem of sugarcane cultivar IACSP96-7569: **(A)** seventh week, 45 days in inductive conditions; **(B)** seventh week, 45 days in non-inductive conditions; **(C)** 13th week, 85 days in inductive conditions; and **(D)** 13th week, 85 days in non-inductive conditions (Santos et al., 2021).

Supplementary Figure 3 | Subtractive approach conducted to annotate sugarcane transcripts. All 191,816 transcripts were considered for annotation using successive BLASTx steps with sorghum, corn, wheat, rice, and Arabidopsis protein databases. Alignments produced in each step were filtered for coverage and identity between query and subject sequences, taking "hit" transcripts out of the initial bulk and carrying further "no hits" for the next BLASTx and repeating this procedure until the last organism, Arabidopsis.

Supplementary Figure 4 | Dissociation curves for primer pairs from essays on spindle leaves of the seventh week. E, primer efficiency; R^2, Correlation coefficient.

Supplementary Figure 5 | Dissociation curves for primer pairs from essays on spindle leaves of the 13th week. E, primer efficiency; R^2, correlation coefficient.

Supplementary Figure 6 | Dissociation curves for primer pairs from essays on mature leaves of the seventh week. E, primer efficiency; R^2, correlation coefficient.

Supplementary Figure 7 | Dissociation curves for primer pairs from essays on mature leaves of the 13th week. E, primer efficiency; R^2, correlation coefficient.

Supplementary Figure 8 | Tree maps representing enriched categories of biological processes for **(A)** upregulated transcripts and **(B)** downregulated transcripts.

Supplementary Table 1 | Target sequences. Gene names (and *Arabidopsis thaliana* TAIR9 entry) used for orthology inference.

Supplementary Table 2 | Main results and metrics of the reference-based mapping of RNA-Seq transcripts.

Supplementary Table 3 | Complete list with results from the phylogenetic analysis. Sugarcane transcripts were clustered closely to respective target *A. thaliana* sequences with correspondence to *S. bicolor* and *O. sativa* homologous genes.

Supplementary Table 4 | Gene relative expression conduced with RT-qPCR essays for two time points (seventh week, 45 days of treatment and 13th week, 85 days of treatment) and two tissues (spindle leaves and mature leaves).

Supplementary Table 5 | List of overrepresented enriched GO terms (FDR < 0.05) from sugarcane photoperiodic flowering induction transcriptome.

Supplementary Table 6 | List of conserved sugarcane lncRNAs responsive to flowering induction and its orthologs in *S. bicolor*, *Z. mays*, and *S. italica* (sorted by highest to lowest $\log_2$ fold change).

REFERENCES

Andrés, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nat. Rev. Genet.* 13, 627–639. doi: 10.1038/nrg3291

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data.* Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Anisimova, M., and Gascuel, O. (2006). Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. *Syst. Biol.* 55, 539–552. doi: 10.1080/10635150600755453

Berding, N., and Hurney, A. P. (2005). Flowering and lodging, physiological-based traits affecting cane and sugar yield: What do we know of their control mechanisms and how do we manage them? *F. Crop. Res.* 92, 261–275. doi: 10.1016/j.fcr.2005.01.015

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi: 10.1093/bioinformatics/btp101

BioBam Bioinformatics (2019). *OmicsBox - Bioinformatics Made Easy*. Available online at: https://www.biobam.com/omicsbox

Bottcher, A., Cesarino, I., Brombini, dos Santos, A., Vicentini, R., Mayer, J. L. S., et al. (2013). Lignification in sugarcane: Biochemical characterization, gene discovery, and expression analysis in two genotypes contrasting for lignin content. *Plant Physiol.* 163, 1539–1557. doi: 10.1104/pp.113.225250

Bouché, F., D'Aloia, M., Tocquin, P., Lobet, G., Detry, N., and Périlleux, C. (2016). Integrating roots into a whole plant network of flowering time genes in Arabidopsis thaliana. *Sci. Rep.* 6, 1–12. doi: 10.1038/srep29042

Brambilla, V., Gomez-Ariza, J., Cerise, M., and Fornara, F. (2017). The importance of being on time: Regulatory networks controlling photoperiodic flowering in cereals. *Front. Plant Sci.* 8:665. doi: 10.3389/fpls.2017.00665

Büche, C., Poppe, C., Schäfer, E., and Kretsch, T. (2000). eid1 : A New Arabidopsis Mutant Hypersensitive in Phytochrome A–Dependent High-Irradiance Responses. *Plant Cell* 12, 547–558. doi: 10.1105/tpc.12.4.547

Budak, H., Kaya, S. B., and Cagirici, H. B. (2020). Long Non-coding RNA in Plants in the Era of Reference Sequences. *Front. Plant Sci.* 11:276. doi: 10.3389/fpls.2020.00276

Cardoso-Silva, C. B., Costa, E. A., Mancini, M. C., Balsalobre, T. W. A., Canesin, L. E. C., Pinto, L. R., et al. (2014). De Novo Assembly and Transcriptome Analysis of Contrasting Sugarcane Varieties. *PLoS One* 9:e88462. doi: 10.1371/journal.pone.0088462

Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* 27, 207–216. doi: 10.1016/j.pbi.2015.08.003

Coelho, C. P., Costa-Netto, A. C., Colasanti, J., and Chalfun-Junior, A. (2013). A proposed model for the flowering signaling pathway of sugarcane under photoperiodic control. *Genet. Mol. Res.* 12, 1347–1359. doi: 10.4238/2013.April.25.6

Coelho, C. P., Minow, M. A. A., Chalfun-Júnior, A., and Colasanti, J. (2014). Putative sugarcane FT/TFL1 genes delay flowering time and alter reproductive architecture in Arabidopsis. *Front. Plant Sci.* 5:1–12. doi: 10.3389/fpls.2014.00221

Coneva, V., Guevara, D., Rothstein, S. J., and Colasanti, J. (2012). Transcript and metabolite signature of maize source leaves suggests a link between transitory starch to sucrose balance and the autonomous floral transition. *J. Exp. Bot.* 63, 5079–5092. doi: 10.1093/jxb/ers158

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., et al. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316, 1030–1033. doi: 10.1126/science.1141752

Das, S., Mcclain, C. J., and Rai, S. N. (2020). Fifteen years of gene set analysis for high-throughput genomic data: A review of statistical approaches and future challenges. *Entropy* 22, 1–23. doi: 10.3390/E22040427

Dieterle, M., Zhou, Y. C., Schäfer, E., Funk, M., and Kretsch, T. (2001). EID1, an f-box protein involved in phytochrome a-specific light signaling. *Genes Dev.* 15, 939–944. doi: 10.1101/gad.197201

Dornelas, M. C., and Rodriguez, A. P. M. (2001). A genomic approach to elucidating grass flower development. *Genet. Mol. Biol.* 24, 69–76. doi: 10.1590/s1415-47572001000100011

Doust, A. N., Mauro-Herrera, M., Hodge, J. G., and Stromski, J. (2017). The C4 model grass setaria is a short day plant with secondary long day genetic regulation. *Front. Plant Sci.* 8:62. doi: 10.3389/fpls.2017.01062

Feng, J., and Lu, J. (2017). LHP1 could act as an activator and a repressor of transcription in plants. *Front. Plant Sci.* 8:2041. doi: 10.3389/fpls.2017.02041

Figueiredo, R. C., Brito, M. S., Figueiredo, L. H. M., Quiapin, A. C., Vitorelli, P. M., Silva, L. R., et al. (2001). Dissecting the sugarcane expressed sequence tag (SUCEST) database: Unraveling flower-specific genes. *Genet. Mol. Biol.* 24, 77–84. doi: 10.1590/s1415-47572001000100012

Florea, L., Hartzell, G., Zhang, Z., Rubin, G. M., and Miller, W. (1998). A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* 8, 967–974. doi: 10.1101/gr.8.9.967

Fornara, F., Panigrahi, K. C. S., Gissot, L., Sauerbrunn, N., Rühl, M., Jarillo, J. A., et al. (2009). Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. *Dev. Cell* 17, 75–86. doi: 10.1016/j.devcel.2009.06.015

Gallart, A. P., Pulido, A. H., De Lagrán, I. A. M., Sanseverino, W., and Cigliano, R. A. (2016). GREENC: A Wiki-based database of plant lncRNAs. *Nucleic Acids Res.* 44, D1161–D1166. doi: 10.1093/nar/gkv1215

Garsmeur, O., Droc, G., Antonise, R., Grimwood, J., Potier, B., Aitken, K., et al. (2018). A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* 9, 51–55. doi: 10.1038/s41467-018-05051-5

Glassop, D., and Rae, A. L. (2019). Expression of sugarcane genes associated with perception of photoperiod and floral induction reveals cycling over a 24-hour period. *Funct. Plant Biol.* 46, 314–327. doi: 10.1071/FP18136

Glassop, D., Rae, A. L., and Bonnett, G. D. (2014). Sugarcane Flowering Genes and Pathways in Relation to Vegetative Regression. *Sugar. Tech.* 16, 235–240. doi: 10.1007/s12355-013-0284-z

Gray, J., Bevan, M., Brutnell, T., Buell, C. R., Cone, K., Hake, S., et al. (2009). A recommendation for naming transcription factor proteins in the grasses. *Plant Physiol.* 149, 4–6. doi: 10.1104/pp.108.128504

Hale, A. L., White, P. M., Webber, C. L., and Todd, J. R. (2017). Effect of growing media and fertilization on sugarcane flowering under artificial photoperiod. *PLoS One* 12:e81639. doi: 10.1371/journal.pone.0181639

Henriques, R., Wang, H., Liu, J., Boix, M., Huang, L. F., and Chua, N. H. (2017). The antiphasic regulatory module comprising CDF5 and its antisense RNA FLORE links the circadian clock to photoperiodic flowering. *New Phytol.* 216, 854–867. doi: 10.1111/nph.14703

Higgins, J. A., Bailey, P. C., and Laurie, D. A. (2010). Comparative genomics of flowering time pathways using brachypodium distachyon as a model for the temperate Grasses. *PLoS One* 5:e10065. doi: 10.1371/journal.pone.0010065

Hill, C. B., and Li, C. (2016). Genetic architecture of flowering phenology in cereals and opportunities for crop improvement. *Front. Plant Sci.* 7:906. doi: 10.3389/fpls.2016.01906

Hong, Y., and Jackson, S. (2015). Floral induction and flower formation-the role and potential applications of miRNAs. *Plant Biotechnol. J.* 13, 282–292. doi: 10.1111/pbi.12340

Huq, E., Tepperman, J. M., Quail, P. H., Qin, M.-M., and Quail, P. H. (2000). GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9789–9794. doi: 10.1073/PNAS.170283997

Ilott, N. E., and Ponting, C. P. (2013). Predicting long non-coding RNAs using RNA sequencing. *Methods* 63, 50–59. doi: 10.1016/j.ymeth.2013.03.019

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., and Kay, S. A. (2005). Plant science: FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science* 309, 293–297. doi: 10.1126/science.1110586

Jackson, S. D. (2009). Plant responses to photoperiod. *New Phytol.* 181, 517–531. doi: 10.1111/j.1469-8137.2008.02681.x

Jiao, F., Pahwa, K., Manning, M., Dochy, N., and Geuten, K. (2019). Cold induced antisense transcription of FLOWERING LOCUS C in distant grasses. *Front. Plant Sci.* 10, 1–9. doi: 10.3389/fpls.2019.00072

Kiełbowicz-Matuk, A., and Czarnecka, J. (2014). *Interplays of Plant Circadian Clock and Abiotic Stress Response Networks*. Nertherland: Elsevier Inc, doi: 10.1016/B978-0-12-800876-8.00020-5

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. *Science* 286, 1960–1962. doi: 10.1126/science.286.5446.1960

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210. doi: 10.1093/nar/gkr1090

Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth1923

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, D., Yang, C., Li, X., Gan, Q., Zhao, X., and Zhu, L. (2009). Functional characterization of rice OsDof12. *Planta* 229, 1159–1169. doi: 10.1007/s00425-009-0893-7

Lin, Y.-R., Upadhyaya, H. D., Zhou, C., Hash, C. T., Clemente, T., Cuevas, H. E., et al. (2016). The Evolution of Photoperiod-Insensitive Flowering in Sorghum,

A Genomic Model for Panicoid Grasses. *Mol. Biol. Evol.* 33, 2417–2428. doi: 10.1093/molbev/msw120

Liu, Z. W., Zhao, N., Su, Y. N., Chen, S. S., and He, X. J. (2020). Exogenously overexpressed intronic long noncoding RNAs activate host gene expression by affecting histone modification in Arabidopsis. *Sci. Rep.* 10, 1–12. doi: 10.1038/s41598-020-59697-7

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lu, S. X., Knowles, S. M., Andronis, C., Ong, M. S., and Tobin, E. M. (2009). Circadian clock associated1 and late elongated hypocotyl function synergistically in the circadian clock of arabidopsis. *Plant Physiol.* 150, 834–843. doi: 10.1104/pp.108.133272

Makino, S., Matsushika, A., Kojima, M., Oda, Y., and Mizuno, T. (2001). Light response of the circadian waves of the APRR1/TOC1 quintet: When does the quintet start singing rhythmically in Arabidopsis? *Plant Cell Physiol.* 42, 334–339. doi: 10.1093/pcp/pce036

Mancini, M. C., Cardoso-Silva, C. B., Sforça, D. A., and Pereira de Souza, A. (2018). Targeted Sequencing by Gene Synteny," a New Strategy for Polyploid Species: Sequencing and Physical Structure of a Complex Sugarcane Region. *Front. Plant Sci.* 9:397. doi: 10.3389/fpls.2018.00397

Martínez-García, J. F., Huq, E., and Quail, P. H. (2000). Direct targeting of light signals to a promoter element-bound transcription factor. *Science* 288, 859–863. doi: 10.1126/science.288.5467.859

Matsubara, K., Hori, K., Ogiso-Tanaka, E., and Yano, M. (2014). Cloning of quantitative trait genes from rice reveals conservation and divergence of photoperiod flowering pathways in Arabidopsis and rice. *Front. Plant Sci.* 5:193. doi: 10.3389/fpls.2014.00193

Mayer, K. F. X., Rogers, J., Dole, J., Pozniak, C., Eversole, K., Feuillet, C., et al. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* 345, 1251788–1251788. doi: 10.1126/science.1251788

McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2018). The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tpj.13781

McWatters, H. G., Roden, L. C., and Staiger, D. (2001). Picking out parallels: Plant circadian clocks in context. *Philos. Trans. R. Soc. B Biol. Sci.* 356, 1735–1743. doi: 10.1098/rstb.2001.0936

Medeiros, A. L., Furtado, C. M., Leite, F. S., Souto, V. S., de Setta, N., Van Sluys, M.-A. A., et al. (2016). Molecular Genetic Dissection of Sugarcane Flowering under Equatorial Field Conditions. *Trop. Plant Biol.* 9, 252–266. doi: 10.1007/s12042-016-9175-2

Melloni, M. L. G., Melloni, M. N. G., Scarpari, M. S., Garcia, J. C., Landell, M. G. A., and Pinto, L. R. (2015). Flowering of Sugarcane Genotypes under Different Artificial Photoperiod Conditions. *Am. J. Plant Sci.* 06, 456–463. doi: 10.4236/ajps.2015.63051

Minow, M. A. A., Ávila, L. M., Turner, K., Ponzoni, E., Mascheretti, I., Dussault, F. M., et al. (2018). Distinct gene networks modulate floral induction of autonomous maize and photoperiod-dependent teosinte. *J. Exp. Bot.* 69, 2937–2952. doi: 10.1093/jxb/ery110

Mizuno, T., and Nakamichi, N. (2005). Pseudo-response regulators (PRRs) or true oscillator components (TOCs). *Plant Cell Physiol.* 46, 677–685. doi: 10.1093/pcp/pci087

Moore, P. H., and Berding, N. (2013). ""Flowering,"," in *Sugarcane: Physiology, Biochemistry, and Functional Biology*, eds P. H. Moore and F. C. Botha (Chichester, UK: John Wiley & Sons Ltd), 379–410. doi: 10.1002/9781118771280

Nonis, A., De Nardi, B., and Nonis, A. (2014). Choosing between RT-qPCR and RNA-seq: A back-of-the-envelope estimate towards the definition of the break-even-point. *Anal. Bioanal. Chem.* 406, 3533–3536. doi: 10.1007/s00216-014-7687-x

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., et al. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res.* 35, D883–D887. doi: 10.1093/nar/gkl976

Papini-Terzi, F. S., Rocha, F. R., Vêncio, R. Z. N., Oliveira, K. C., Felix, J., de, M., et al. (2005). Transcription profiling of signal transduction-related genes in sugarcane tissues. *DNA Res.* 12, 27–38. doi: 10.1093/dnares/12.1.27

Patel, R. K., and Jain, M. (2012). NGS QC toolkit: A toolkit for quality control of next generation sequencing data. *PLoS One* 7:e30619. doi: 10.1371/journal.pone.0030619

Pfaffl, M. W., Horgan, G. W., and Dempfle, L. (2002). Relative expression software tool (REST(C)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 30:36e. doi: 10.1093/nar/30.9.e36

Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012). NCBI Reference Sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Res.* 40, D130–5. doi: 10.1093/nar/gkr1079

Putterill, J., and Varkonyi-Gasic, E. (2016). FT and florigen long-distance flowering control in plants. *Curr. Opin. Plant Biol.* 33, 77–82. doi: 10.1016/j.pbi.2016.06.008

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den Hoff, M. J. B., et al. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37:e45. doi: 10.1093/nar/gkp045

Santos, P. H., da, S., Manechini, J. R. V., Brito, M. S., Romanel, E., Vicentini, R., et al. (2021). Selection and validation of reference genes by RT-qPCR under photoperiodic induction of flowering in sugarcane (Saccharum spp.). *Sci. Rep.* 11:4589. doi: 10.1038/s41598-021-83918-2

Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D., et al. (2003). Dissection of floral induction pathways using global expression analysis. *Development* 130, 6001–6012. doi: 10.1242/dev.00842

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534

Schwacke, R., Ponce-Soto, G. Y., Krause, K., Bolger, A. M., Arsova, B., Hallab, A., et al. (2019). MapMan4: A Refined Protein Classification and Annotation Framework Applicable to Multi-Omics Data Analysis. *Mol. Plant* 12, 879–892. doi: 10.1016/j.molp.2019.01.003

Seaton, D. D., Smith, R. W., Song, Y. H., MacGregor, D. R., Stewart, K., Steel, G., et al. (2015). Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Mol. Syst. Biol.* 11:776. doi: 10.15252/msb.20145766

Shannon, P., Markiel, A., Owen Ozier, Z., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 303, 2498–2504. doi: 10.1101/gr.1239303.metabolite

Song, Y. H., Shim, J. S., Kinmonth-Schultz, H. A., and Imaizumi, T. (2015). Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. *Annu. Rev. Plant Biol.* 66, 441–464. doi: 10.1146/annurev-arplant-043014-115555

Srikanth, A., and Schmid, M. (2011). Regulation of flowering time: All roads lead to Rome. *Cell. Mol. Life Sci.* 68, 2013–2037. doi: 10.1007/s00018-011-0673-y

Suárez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Lett. Nat.* 5:410.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800

Tapia-López, R., García-Ponce, B., Dubrovsky, J. G., Garay-Arroyo, A., Pérez-Ruíz, R. V., Kim, S. H., et al. (2008). An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in Arabidopsis. *Plant Physiol.* 146, 1182–1192. doi: 10.1104/pp.107.108647

Torti, S., Fornara, F., Vincent, C., Andrés, F., Nordström, K., Göbel, U., et al. (2012). Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. *Plant Cell* 24, 444–462. doi: 10.1105/tpc.111.092791

Veneziano, D., Di Bella, S., Nigita, G., Laganà, A., Ferro, A., and Croce, C. M. (2016). Noncoding RNA: Current Deep Sequencing Data Analysis Approaches and Challenges. *Hum. Mutat.* 37, 1283–1298. doi: 10.1002/humu.23066

Vicentini, R., Del Bem, L. E. V., Van Sluys, M. A., Nogueira, F. T. S. S., and Vincentz, M. (2012). Gene Content Analysis of Sugarcane Public ESTs Reveals Thousands of Missing Coding-Genes and an Unexpected Pool of Grasses Conserved ncRNAs. *Trop. Plant Biol.* 5, 199–205. doi: 10.1007/s12042-012-9103-z

Wang, H., Pan, J., Li, Y., Lou, D., Hu, Y., and Yu, D. (2016). The DELLA-CONSTANS Transcription Factor Cascade Integrates Gibberellic Acid and

Photoperiod Signaling to Regulate Flowering. *Plant Physiol.* 172, 479–488. doi: 10.1104/pp.16.00891

Wang, L., Park, H. J., Dasari, S., Wang, S., Kocher, J. P., and Li, W. (2013). CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Res.* 41, 1–7. doi: 10.1093/nar/gkt006

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484

Wellmer, F., and Riechmann, J. L. (2010). Gene networks controlling the initiation of flower development. *Trends Genet.* 26, 519–527. doi: 10.1016/j.tig.2010.09.001

Wolabu, T. W., Zhang, F., Niu, L., Kalve, S., Bhatnagar-Mathur, P., Muszynski, M. G., et al. (2016). Three FLOWERING LOCUS T-like genes function as potential florigens and mediate photoperiod response in sorghum. *New Phytol.* 210, 946–959. doi: 10.1111/nph.13834

Xu, F., Li, T., Xu, P.-B., Li, L., Du, S.-S., Lian, H.-L., et al. (2016). DELLA proteins physically interact with CONSTANS to regulate flowering under long days in Arabidopsis. *FEBS Lett.* 590, 541–549. doi: 10.1002/1873-3468.12076

Yamada, K. D., Tomii, K., and Katoh, K. (2016). Application of the MAFFT sequence alignment program to large data—reexamination of the usefulness of chained guide trees. *Bioinformatics* 32, 3246–3251. doi: 10.1093/bioinformatics/btw412

Yamaguchi, A., and Abe, M. (2012). Regulation of reproductive development by non-coding RNA in Arabidopsis: To flower or not to flower. *J. Plant Res.* 125, 693–704. doi: 10.1007/s10265-012-0513-7

Yamamoto, Y., Sato, E., Shimizu, T., Nakamich, N., Sato, S., Kato, T., et al. (2003). Comparative Genetic Studies on the APRR5 and APRR7 Genes Belonging to the APRR1/TOC1 Quintet Implicated in Circadian Rhythm, Control of Flowering Time, and Early Photomorphogenesis. *Plant Cell Physiol.* 44, 1119–1130. doi: 10.1093/pcp/pcg148

Yuan, N., Balasubramanian, V. K., Chopra, R., and Mendu, V. (2019). The photoperiodic flowering time regulator fkf1 negatively regulates cellulose biosynthesis. *Plant Physiol.* 180, 2240–2253. doi: 10.1104/pp.19.00013

Zhao, X., Li, J., Lian, B., Gu, H., Li, Y., and Qi, Y. (2018). Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-018-07500-7

Zhu, M., Zhang, M., Xing, L., Li, W., Jiang, H., Wang, L., et al. (2017). Transcriptomic Analysis of Long Non-Coding RNAs and Coding Genes Uncovers a Complex Regulatory Network That Is Involved in Maize Seed Development. *Genes* 8:274. doi: 10.3390/genes8100274