Comparative transcriptomic analysis of flowering induction in pitaya induced by supplementary lighting

CURRENT STATUS: POSTED

Rui Xiong
Hainan University

Liu Chengli
Hainan University

Min Xu
Hainan University

Shuang-Shuang Wei
Hainan University

Hua Tang
Hainan University

Corresponding Author
thtiger@163.com
ORCiD: https://orcid.org/0000-0001-9891-1694

DOI:
10.21203/rs.2.9246/v1

SUBJECT AREAS
Internal Medicine Specialties

KEYWORDS
Pitaya, supplementary lighting induced flowering, de novo RNA-seq analysis, differentially expressed genes, qRT-PCR
Abstract

Background

Pitayas are currently attracting considerable interest as a fruit with many health benefits. However, the lack of natural light after November in Hainan, China, severely restricts the production of pitaya in winter. To further explore the molecular mechanisms regulating flowering in pitaya, we used de novo RNA sequencing-based transcriptomic analysis for four stages of pitaya subjected to light induction.

Results

We assembled 68113 unigenes in total, comprising 29782 unigenes with functional annotations in the NR database, 20716 annotations in SwissProt, 18088 annotations in KOG, and 11059 annotations in KEGG. Comparison between different samples revealed different numbers of significantly differentially expressed genes (DEGs). A number of DEGs involved in energy metabolism-related processes and plant hormones were detected. Moreover, we discovered many CONSTANS-LIKE, FLOWERING LOCUS T and other DEGs involved in direct regulation of flowering, along with CDF and TCP, which function as typical transcription factor genes in the flowering process. At the transcriptomic level, we confirmed 13 DEGs with different functions in the time-course response to light-induced flowering by quantitative reverse-transcription PCR analysis.

Conclusions

These DEGs may include some key genes that control the floral-induction network, increasing our understanding of the molecular mechanism of floral regulation in pitaya. These findings will also aid the development of biotechnologies aimed at creating a variant of pitaya that is less sensitive to light conditions and blooms throughout the year.

1 Background

Pitaya (Hylocereus undulatus Britt) belongs to the Cactaceae family and Hylocereus genus, and is a tropical or subtropical fruit tree. The specific origin of the pitaya remains inconclusive. At present, the academic community generally believes that it originated in southern Mexico. Nowadays, it is geographically widespread, occurring in Israel, Japan’s Okinawa, China’s Guangxi, Hainan, and Taiwan, and Southeast Asia[1]. The earliest literature report on pitaya is from the General and Natural
History of the Indies, written by Fernández de Oviedo y Valdés and published in 1535, which clearly records that this fruit may be edible. The stem of the pitaya, with its waxy surface, can store a large amount of water, the leaves have metamorphosed into thorns on account of long-term environmental stresses, and the crassulacean acid metabolism (CAM) respiratory pathway allows the dragon fruit to endure extreme environments. The fruits are prized not only for their fine appearance and striking colors, but also for the potentially beneficial metabolic effect of Hylocereus polyrhizus fruit betacyanins (HPBN) on high-fat-associated human diseases. A number of studies have been undertaken to ascertain the nutritional value of pitaya, which is a rich source of organic acids, sugars, polyphenols, vitamin C, and antioxidant activity. The pitaya flower is spectacular: it elongates at dusk and blooms at night, and then withers after pollination in the daytime. Pitaya is an obvious long-day plant. In the northern hemisphere, most pitaya flowers bloom from May to the end of October as part of the summer season. However, flowering is severely restricted under the very reduced light levels in the winter in China, which then negatively affects the supply of pitaya fruits available to consumers. If we want to produce pitaya fruit in the winter season, when the light is inadequate, we have to supply sufficient light to promote flowering.

When the plant has passed from the juvenile to the adult phase and from the vegetative to the reproductive stage, the plant can be induced to flower. Flowering is regulated by an integrated network of chemical and genetic pathways, such as the photoperiod pathway, vernalization pathway, gibberellic acid (GA) pathway, and autonomous pathway. Various input signals activate the signal transduction pathways that control flowering time. The photoperiod pathway refers to the response to day length and quality of light. This has been shown to play a crucial role in controlling flowering in Arabidopsis. CONSTANS (CO), in the presence of light, has been proven to be a key regulator of flower promotion through the activation of FLOWERING LOCUS T (FT) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) under long-day (LD) conditions. FT then associates with FLOWERING LOCUS D (FD), and the FT-FD complex promotes expression of the downstream floral meristem identity genes, such as APETALA1 (API) and LEAFY (LFY), to induce flowering.

In Arabidopsis, plant hormone signaling (via GA, JA, auxins, etc.) is required for floral induction. GA
directly promotes $SOC1$ and $LFY$ expression and increases the level of $SOC1$ mRNA, which, in turn, activates the downstream genes $LFY$ and $AP1$ to induce flowering[12]. One study found that tomato plants with jai1-1 mutants exhibited delayed flower opening, showing that jasmonic acid (JA) acts as a positive regulator of flowering[13]. IAA (indole-3-acetic acid) is likely involved in the mechanism that controls the growth of the male gametophyte to the egg cell in the ovule, demonstrating auxin is a major controlling signal that synchronizes flower development in Arabidopsis[14].

Understanding the detailed flowering mechanism is important for an understanding of the plant growth cycle, especially in the transition from vegetative to reproductive development. For pitaya, few studies have been conducted on molecular mechanisms that promote flowering by light supplementation. Therefore, understanding gene regulation for flowering could facilitate the development of pitaya varieties with photoperiod insensitivity, which could be helpful for reducing energy costs and sustaining year-round production.

2 Results
2.1 Flowering induction of pitaya by supplementary light treatment in the winter season

These findings showed that sufficient supplementary light can induce pitaya flowering if the minimum air temperature is $\geq$ 15 °C. In this supplementary light experiment, NL represents the control group (no light, no flowering) (Fig. 1A), L0 represents no flowering in the light-treated group (Fig. 1B), L1 represents the flower bud stage in the light-treated group (Fig. 1C), and L2 represents 1 week after the bud stage in the light-treated group (Fig. 1D). About 20~25 days after the buds appeared, the pitaya flowers were in full bloom at night (Fig. 1E), signifying the success of this flower induction system with respect to pitaya fruit production in the winter season.

Figure 1 Supplementary light experiment. A: NL represents the control group (no flowering under no light); B: L0 represents the no-flowering stage in the light treatment group; C: L1 represents the flower bud stage after light treatment; D: L2 represents 1 week post bud stage under light treatment; E: flowering of pitaya induced by light treatment in the winter season.
2.2 Sequencing, de novo assembly, and annotation

To understand the underlying molecular events involved in the response to light-induced flowering, we sent pitaya samples from the NL, L0, L1, and L2 stages to the Gene Denovo Biotechnology Co. (Guangzhou, China) for de novo assembly of RNA-seq data. More than 39 Gb raw data reads were retrieved. The high-quality clean reads number of the three biological replicates for the NL sample were 45444076 (97.78%), 42546349 (97.75%), and 42956302 (97.82%); the clean reads for the L0 sample were 35518498 (97.83%), 38856464 (97.71%), and 49626346 (98.11%); the clean reads for the L1 sample were 42262290 (97.78%), 52312394 (97.94%), and 49023546 (98.01%); and the clean reads for the L2 sample were 49899592 (98.17%), 48213826 (98.08%), and 45191594 (98%) (Additional file 1). These findings indicated that the proportion of high-quality clean reads after filtration of each sample was relatively high, signifying that the sequencing quality was good. The de novo assembly resulted in a total of 60580077 bases, but only 68113 unigenes (not less than 200 nt) with an N50 of 1730 nt (Table 1). The average length of our assembly was 889 nt, of which the maximum length of unigene was 36165 nt, the minimum length 201 nt, and the GC (guanine-cytosine) percentage 41.96%. Figure 2 showed the scatter plot of size distribution of the unigenes. In order to search for homologous sequences, annotation analysis was performed on the unigenes using four public databases: KOG, KEGG, Swissprot, and NR. In total, 29959 unigenes (removing duplicates) were found in at least one of these databases (Table 2). There were 29782 unigenes with functional annotations in the NR database, 20716 annotations in SwissProt, 18088 annotations in KOG, and 11059 annotations in KEGG. We also observed that 8618 unigenes were annotated by all four major databases (Fig. 3). At the same time, we found only 6220, 75, 33, and 32 unique annotations in the NR, Swissprot, KOG, and KEGG databases, respectively.

Table 1 Results of assembly

| Sample Stage | Clean Reads (Gb) | Clean Reads % |
|--------------|-----------------|--------------|
| NL           | 45444076 (97.78%) | 42546349 (97.75%) | 42956302 (97.82%) |
| L0           | 35518498 (97.83%) | 38856464 (97.71%) | 49626346 (98.11%) |
| L1           | 42262290 (97.78%) | 52312394 (97.94%) | 49023546 (98.01%) |
| L2           | 49899592 (98.17%) | 48213826 (98.08%) | 45191594 (98%) |

Figure 2 Length distribution of all unigenes.
Table 2 Results of annotations in four databases

Figure 3 Venn diagram of the four database annotations.

2.3 Global analysis of differential expression profiling

The abundance of each gene was determined by counting reads per kilobase per million reads (RPKM) to infer the expression level. Therefore, this method could be used to directly compare the difference in gene expression among samples. The correlation between gene expression levels among the samples was a key criterion in determining whether the experiments were reliable and whether the samples chosen were suitable. If one sample had a high degree of similarity to another, the correlation value between them would be very close to 1. We calculated the correlation values between samples based on the RPKM results. According to the standard recommended by the ENCODE (Encyclopedia of DNA Elements) project[15], the square of the correlation value should be ≥0.92 (under ideal experimental conditions and with suitable samples). The heatmap of correlations for these samples is shown in Figure 4. From the figure, we can see that the coefficient between the three repetitions for the NL stage was found to be above 0.98, and the correlation between the three repetitions for the L0 stage was above 0.96. Similarly, the correlation between the three repetitions for the L1 and L2 stages was found to be above 0.99 and 0.97, respectively, signifying that the levels of expression between all the biological replicates were highly correlated. In other words, we found a low level of correlation between the different samples, indicating that the difference between these four samples was significant.

Figure 4 Relationship analysis of the samples. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage under light treatment.
In our study, comparison between different samples resulted in different numbers of DEGs (Fig 5). When NL and L0 were compared, we found only 51 up-regulated genes and 88 down-regulated genes, totaling 139 DEGs. A total of 12352 DEGs were obtained when we compared NL and L1, of which 8640 genes were up-regulated and 3712 genes were down-regulated. In other words, compared with NL, the number of DEGs in the L1 sample was found to be about nine times that in the L0 sample, where the number of up-regulated genes was 169 times that of L1 and the number of down-regulated genes was 41 times higher than that of L1. Similar results were found in other comparison groups. Gene ontology (GO) functional enrichment analysis was applied in our subsequent study in order to annotate the expression patterns of the DEGs to the chosen GO terms: molecular function, cellular component, and biological process. Since a gene often has multiple different functions, the same gene will appear under different terms and each histogram is statistically independent of any others.

In our research, 50 GO terms relating to molecular function, 30 GO terms relating to cellular component, and 158 GO terms relating to biological process were enriched with respect to the L0 stage, as compared to the NL stage, after light treatment (Additional file 2,3,4). Under the same light treatment, DEGs associated with the L1 stage were enriched by 399, 162, and 1042 GO terms relating to molecular function, cellular component, and biological process, respectively (Additional file 5,6,7). Moreover, findings regarding the GO analysis of other groups were similar to those described above. Most importantly, we found that the main enriched terms of all groups were the same regardless of ontologies. In particular, ‘cellular component’ mainly related to the cell part, membrane, and organelle; ‘molecular function’ mainly related to binding, catalytic activity, and transport activity; and ‘biological process’ mainly related to cellular, metabolic, single-organism transport, suggesting that DEGs primarily enriched in those terms may be associated with induction of pitaya blossoms. To further investigate the gene expression profiles, we made use of KEGG pathway analysis to determine some of the DEGs involved in important biochemical metabolic and signal transduction pathways during light-induced flowering of pitaya. For example, two auxin response factors (unigene0028447, unigene0002811) were up-regulated, two genes related to jasmonic acid (unigene0030164, unigene0033693) were up-regulated, and five genes related to starch and sucrose metabolism
(unigene0052631, unigene0036015, unigene0040385, unigene0041472, unigene0035017) were also up-regulated.

Figure 5 Gene statistics for DEGs in different samples. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage under light treatment.

2.4 Energy-related genes in the process of flowering induction

During our investigation of light-induced pitaya flowering, we focused on the differentially expressed genes in the NL, L0, and L1 stages. It is well known that plant flowering is a normal physiological process. Before a plant flowers, some essential organic nutrients accumulate in the plant so as to accommodate the significant uptake of nutrients during flowering. Soluble sugar and soluble protein are important nutrients that are closely associated with plant flowering. Large increases in ATP concentration initiated by changes in environmental conditions, especially by different kinds of light treatments, have already been observed[16]. Our analysis found that two up-regulated genes, unigene 0046195 (beta-glucosidase BoGH3B-like) and unigene 0014135 (ATP synthase), in the Krebs cycle of glucose metabolism, were involved in the L0 stage, but not the NL stage. Moreover, the expression levels of eight genes (unigene0006689, unigene0031088, unigene0044754, unigene0009228, unigene0045801, unigene0047216, unigene0050048, unigene0053090) involved in sucrose synthesis in L1 were found to have increased, and three of these genes were up-regulated in L1, but not in L0. This finding, which relates to energy metabolism, demonstrates that large amounts of sucrose are needed as an energy supply during the flowering process.

2.5 Plant hormone and signal transduction related genes

Three types of phytohormone—auxin, jasmonic acid (JA), and the brassinosteroids (BRs)—have been individually connected to floral timing, though they have not been extensively studied. These hormones can be transported as signal molecules in plants, which in turn produce some signal transduction cascades that direct a series of metabolic activities in plants.
Genetic evidence has revealed that in *Arabidopsis*, the polar transport of auxin controls flower formation and differentiation[17]. Genes regulating floral organ development and gynoecium vascularization have been discovered, indicating the probable involvement of auxins in flower development. In our study (Fig 6 and Additional file 8), we found an auxin-related gene (unigene 0029163) that was significantly up-regulated in L0, but not in NL. Twenty-three genes were significantly expressed in L1, of which two genes (unigene 0027380, unigene 0030569) were down-regulated and 21 genes were up-regulated. Studies have reported that in *Arabidopsis thaliana*, the BR biosynthesis-deficient *det2* mutant is delayed flowering by at least 10 days compared to the wild type. The level of endogenous BR is lower 10% than that of the wild type[18]. Compared with NL, only one up-regulated gene was expressed in L0 (unigene0026249), but six up-regulated DEGs (unigene0004470, unigene0025032, unigene0030389, unigene0038596, unigene0035058, unigene0042473) were expressed in L1. At the same time, compared with L0, eight up-regulated genes involved in the brassinosteroid metabolism pathway were expressed in L1. Generally speaking, JA is known to activate transcription factors (TFs) that trigger a large-scale process of various abiotic and biotic stresses, but also plays an important role in regulating flower opening. Jasmonate-ZIM domain (JAZ) proteins have been shown to regulate the level of JA, to counteract flower abscission in *Nicotiana attenuata* plants[19]. In our own study, we found one gene that was down-regulated (unigene 0021973) in L0, and four genes that were up-regulated (unigene 0030164, unigene 0032865, unigene 0034754, unigene 0049406) in L1. In addition, only one up-regulated gene (unigene0033693) was expressed in L1, but not in L0.

Figure 6 Heatmap of the DEGs associated with various plant hormones. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: stage of flowering 1 week after light treatment.

### 2.6 CO and FT or other directly regulating DEGs

In plants, seasonal changes in day length are perceived in leaves, which initiate long-distance
signaling that induces flowering at the shoot apex. In Arabidopsis, FLOWERING LOCUS T (FT) is now known as florigen, which is crucial for the proper timing of flowering, and CONSTANS (CO) is a key player in the activation of FT expression. In other words, FT acts partially downstream of CONSTANS (CO), which promotes flowering in response to long days. In this study (Fig. 7 and Additional file 9), relative to NL, in the light treatment group L0, one gene—the CONSTANS-related gene (unigene 0025406)—was found to be significantly down-regulated, and the FT gene was not significantly expressed. In comparison with NL, 13 CONSTANS-LIKE genes were significantly differentially expressed in L1, among which three genes were up-regulated (unigene002346462, unigene0027129, unigene0035118) and 10 genes were down-regulated (unigene0025406, unigene0027332, unigene0028118, unigene0029386, unigene0030366, unigene0034031, unigene0034048, unigene0035131, unigene0043102, unigene0050191). At the same time, two upregulated FT genes (unigene0001751, unigene0005649) were significantly differentially expressed. One FT gene (unigene0025148) was down-regulated, and six FT genes (unigene0002836, unigene0006372, unigene0025249, unigene0038103, unigene0042376, unigene0045680) were up-regulated in L1, but not in L0.

In addition to these well-known CO and FT genes, certain genes are closely related to the regulation of plant flowering processes. For example, the protein encoded by Hd3a, a rice ortholog of FT, moves from the leaf to the shoot apical meristem and induces flowering in rice, suggesting that the Hd3a protein may be the rice florigen[20]. In our analysis, an HD3a-like gene (unigene0019970), whose expression decreased under light treatment, was detected in L0. Maybe the HD3a-like gene was up-regulated during the rice flowering period but down-regulated during the pitaya flowering period.

The LEAFY gene is an important element in the transition from the vegetative to the reproductive phase, as LEAFY is both necessary and sufficient to initiate the growth of individual flowers. On long days, Arabidopsis plants flower soon after germination, in parallel with rapid up-regulation of LEAFY[21]. In this research, we found one LEAFY homolog gene (unigene0024655) to be significantly up-regulated in L1. We also found two genes encoding flowering time control protein to be involved in L1, and FPA (unigene0045587) to be down-regulated, while the reverse was true for the expression of
FCA (unigene0049096). One flowering-promoting factor 1-like gene (unigene0023129) was up-regulated in L1.

Light is one of the clock entrainment signals and can induce a plant’s circadian rhythm, acting as a bridge in the signaling network between the environmental stimulation and its internal processes. According to KEGG pathway analysis, some DEGs are associated with circadian rhythm in pitaya. One gene encoding Arabidopsis pseudo response regulator (APRR5)-like protein (unigene0045107) was found to be down-regulated in L0, but not in NL. Two down-regulated circadian clock associated 1-like genes (unigene0036436, unigene0049229), two down-regulated genes encoding EARLY FLOWERING 4-like protein (unigene0022870, unigene0039817), and an up-regulated gene encoding EARLY FLOWERING 1-like protein unigene0043786 were expressed in L1, but not in NL.

Figure 7 Heatmap of the DEGs associated with CO and FT. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage after light treatment.

2.7 Transcription factors control the course of flowering

Cycling DoF factor (CDF) proteins are part of the larger family of plant-specific DOF (DNA binding with one finger) transcription factors, which presumably includes a single C2-C2 zinc finger, a highly conserved DNA-binding domain[22]. Several Arabidopsis DoF transcription factors—CDF1, 2, 3, and 5—are implicated in circadian rhythms and photoperiodism, and their expression at elevated levels is sufficient to repress flowering[23]. In this study (Fig. 8 and Additional file 10), we found two CDF genes (unigene0033003, unigene0031426) that were down-regulated in L0, but not in NL. Similarly, three CDF genes (unigene0031426, unigene0026114, unigene0033003) were down-regulated, and two genes (unigene0029432, unigene0006016) up-regulated, in L1.

In the development and evolution of flowering plants, the MADS domain transcription factor plays an important role because it determines the establishment of flower morphology and flower development. The MADS-box gene contains a conserved MADS-box motif and is generally classified...
into a type I (type I) and a type II (type II) with a conserved K-, I-, and C-domain subfamily. In general, the type II subfamily is divided into ABCDE models during the development of the flower[24]. Compared with NL, there were no significantly differentially expressed MADS-box genes in L0. In contrast, one gene encoding AGAMOUS-like MADS-box protein AGL19 (unigene0027094) was down-regulated, and the other six genes encoding AGAMOUS-like MADS-box protein (unigene0018425, unigene0021927, unigene0025394, unigene0037726, unigene 0047376, unigene 0202108) were up-regulated, in L1.

Other transcription factors that can bind to the CO promoter include the TEOSINTE BRANCHED 1/ CYCLOIDEA/ PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR (TCP) transcription family[25]. Members of this family control multiple traits in diverse plant species, including flower and petal asymmetry[26]. We discovered no DEG-related TCP TFs in L0, while all six genes encoding TCP (unigene0005880, unigene0009351, unigene0025734, unigene0028812, unigene0030793, unigene0033768) were up-regulated in L1, but not in NL.

Figure 8 Heatmap of the three main TFs involved in pitaya flowering. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage after light treatment.

2.8 Validation of the selected DEGs by quantitative RT-PCR analysis

We consider the control of floral development genes or transcription factors that regulate downstream genes to be more important than other aspects of plant flowering processes. One cyclic DoF factor (CDF) transcription factor gene in each of the L0 and L1 samples was detected by quantitative RT-PCR. One gene encoding response regulator-like APRR5 related to circadian rhythm involved in course in L0 and two genes encoding circadian clock associated protein involved in L1 were selected to aid the detection of mRNA levels by quantitative RT-PCR. One and two CONSTANS-LIKE (CO) genes in L0 and L1, respectively, were chosen for real-time quantitative PCR. Furthermore, one flowering locus T (FT) gene and its ortholog encoding HEADING DATE 3A-like (HD3a-like), involved in L1 and L0
respectively, were detected by quantitative RT-PCR. In addition, two TCP TF genes and a gene encoding phytochrome B in L1 were also selected for quantitative RT-PCR (Table 1). The UBQ gene was used as a reference gene[27]. The quantitative RT-PCR (qRT-PCR) results from the L0 and L1 stages are shown in Figures 9 and 10. The expression profiles of all 13 detected genes revealed a similar trend and consistent results between qRT-PCR and RNA-seq. Three DEGs of L0 displayed significant down-regulation in plants under light conditions compared with NL, including unigene0031426 (cycling DoF factor 1), unigene0025406 (zinc finger protein CONSTANS-LIKE 2), and unigene0019970 (HEADING DATE 3A-like); however, unigene0045107 (response regulator-like APRR5) that concerning circadian clock revealed no remarkable decrease in its mRNA level. In the L1 stage, all nine DEGs showed significant differential expression through qRT-PCR. Two genes exhibiting functional homology to circadian clock associated 1 in the L1 stage showed a significant reduction in transcript levels under light treatment. A similar situation was found for two genes encoding TCP transcription factor, for which the mRNA abundance was substantially enhanced in the L1 sample following light treatment. Interestingly, two CO genes revealed an opposite expression in the L1 stage: one was up-regulated and in the other the mRNA level was reduced. This finding indicates that different CO gene families may be subjected to different types of regulation, resulting in opposite types of expression. In addition, CDF3-like and PHYB were found to be down-regulated, while FT displayed a higher mRNA level during L1 light treatment than NL, as determined by qRT-PCR. Figure 9 Relative expression of the floral development genes or transcription factors involved in the flowering process for the NL and L0 samples was determined by quantitative RT-PCR analyses. Expression was normalized to that of UBQ. The transcript levels from the NL sample were set at 1. The blue columns represent the expression levels determined by qRT-PCR (left axis), while the lines represent the gene expression levels determined by RNA-seq (right axis). The transcript abundances of the genes encoding cyclic DoF factor 1 protein (A), zinc finger protein CONSTANS-LIKE 2 (B), HEADING DATE 3A-like (C), and response regulator-like APRR5 (D) were determined and compared across the time course of light-induced treatment. Data represent means ± SD for the three replicates (n=3). The lowercase letters above the columns indicate a significant difference at p≤0.05.
between the columns by T test using SPSS statistical software. Data in columns with the same letters showed no significant difference ($p>0.05$).

Figure 10. Relative expression of the floral development genes or transcription factors involved in the flowering process for the NL and L1 samples was determined by quantitative RT-PCR analyses. Expression was normalized to that of $UBQ$. The transcript levels from the NL sample were set at 1. The blue or red columns represent the expression levels determined by qRT-PCR (left axis), while the lines represent the gene expression levels determined by RNA-seq (right axis). The transcript abundances of the genes encoding cyclic DoF factor 3-like protein (A), CONSTANS-LIKE 15 (B), CONSTANS-LIKE 4 (C), circadian clock associated1 (D, E), flowering locus T (F), phytochrome B (G), transcription factor TCP4 (H), and transcription factor TCP15 (I) were determined and compared across the time course of light-induced treatment. Data represent means ± SD for the three replicates ($n=3$). The lowercase letters above the columns indicate a significant difference at $p \leq 0.05$ between the columns by T test using SPSS statistical software. Data in columns with the same letters showed no significant difference ($p>0.05$).

2.9 Tissue specific expression of key genes

In this study, in order to further explore whether the genes associated with the flowering process of Pitaya are tissue-specific, different growth stages of stem, flower and fruit in Figure 11 were selected for specific analysis. We selected $CDF-3$ like(Unigene0026114) and $CO-15$ like(Unigene0035118) for tissue-specific analysis in Figure 12. The results showed that $CDF-3$ like had a certain expression level in various tissues and various periods, while its expression was increased in pistils and styles and highest in small green peel compared with in the old stem. The expression level of $CO-15$ like in flowers and fruits is higher than that of old stems, and the specificity is not obvious.

Figure 11 Different growth periods and tissues of pitaya.1: old stem,2: young stem,3: small flower bud,4: large flower bud, 4A: pistil,4B: stamen,4C: style,4D: flower scales,5: small green fruit, 6: big green fruit, 7: red pitaya, 8E: small green peel, 8F: small green flesh, 9: big green fruit, 9G: big green
peel, 9H: big green flesh, 10: red fruit, 10I: red peel, 10J: red flesh

Figure 12 The result of CDF-3 like (Unigene0026114) and CO-15 like (Unigene0035118) tissue-specific expression. Expression was normalized to that of UBQ. The transcript levels from the old stem were set at 1.

3 Discussion
Some genes involved in light-induced flowering have been extensively studied in other plants, especially the model plant Arabidopsis. However, no previous research in this area has been conducted for the pitaya, so it is necessary to study the relationship between flowering and the key genes regulating this process.

Hormone regulation, signaling, and homeostasis are very important factors in the flowering process. It is well known that hormonal signaling, involving hormones such as ethylene (ET), abscisic acid (ABA), and auxin, has an important role in regulating flowering time through the regulation of FLC, CO, and FT. A strict control of auxin homeostasis and the maintenance of an appropriate level of IAA are essential for normal growth and development. In our study, some genes (unigene0036729, unigene0037786) encoding auxin transporter-like protein were substantially up-regulated after light treatment. Initiation of floral primordia not only requires a local maximum level of auxin, but also the activity of AUXIN RESPONSE FACTOR[28]. In our analysis, many genes (unigene0004150, unigene0021847, unigene0032942, and so on) encoding auxin-responsive protein also displayed up-regulation. Defects along the jasmonic acid (JA) signaling pathway affect not only floret opening and anther dehiscence, but also spikelet development, in rice[29]. In our study, we observed that some genes encoding JAZ-like protein (e.g., MYC2 TFs) were involved in the JA signaling pathway.

Light signals perceived by phytochrome A (phyA) and cryptochrome (cry) photoreceptors stabilize CO protein on long days[30]. In our research, some CO-like genes were up-regulated, while others were down-regulated, under the same conditions, indicating that some CO genes belonging to specific
families may be repressed by certain proteins or miRNAs in pitaya. Photoreceptors regulate CO stability and act antagonistically to generate daily rhythms in CO abundance[31]. Moreover, a phyB mutation has been shown to cause an increase in CO protein abundance in the early phase of the photoperiod[31], indicating that phyB mutations also accelerate flowering. So, we can conclude that the amount of CO is positively regulated by the photoreceptors cry1, cry2, and phyA, but the mRNA expression of CO is inhibited by phyB. A similar conclusion was drawn from our own study. The gene encoding phyB (unigene0002396) was down-regulated by qPCR, and the expression pattern was found to be the same as that for transcriptome sequencing, suggesting that PHYB may negatively regulate CO throughout the pitaya flowering process.

The circadian clock gene, which is a key transcriptional activator of the photoperiodic pathway, CO, is essential for proper photoperiodic flowering. In the current model of the Arabidopsis circadian clock, most components function as repressors[32]. Fujiwara et al. demonstrated that the circadian clock proteins LATE ELONGATED HYPOCOTYL (LHY) and CIRCADIAN CLOCK-ASSOCIATED1 (CCA1) repressed floral transition, and in particular repressed FT expression under long-day conditions[33]. In our analysis, the transcriptome level of most DEGs associated with circadian rhythm were declined, consistent with previous reports.

The FT gene was first identified in Arabidopsis thaliana, in which it plays a major role in the photoperiod pathway and in initiating the flowering signal in the apical meristem[34]. Unlike many other floral regulators, the deduced sequence of the FT protein suggests that it does not directly control transcription or transcript processing[35]. FT expression is affected by a large number of factors, some of which act as repressors and others as positive regulators. The former class includes several MADS box TFs[36], AP2-like proteins and B-box proteins, while the latter class includes CO[37], NUCLEAR FACTOR Y, and GIGANTEA (GI)[38]. This finding was also revealed in our own study; except for one down-regulated FT gene (unigene0025148), all the other genes were up-regulated. CYCLING DoF FACTOR 1 (CDF1) and its homologs play an important role in floral transition by repressing the expression of floral activator genes such as CO and FT in Arabidopsis[39]. One study indicated that the CDF repression operates through the formation of a CDF-TPL transcriptional
complex, which binds the CO and FT promoters to reduce their expression levels throughout the morning, for seasonal flowering[40]. The results of our quantitative RT-PCR analysis revealed down-expression of two CDF-like genes (unigene0031426, unigene0026114) but up-expression of a CO-like gene 15 (unigene0027129) and FT gene (unigene0005649). This finding demonstrated that CDF may repress CO and FT transcriptome expression during the pitaya flowering process. Simultaneously, we observed that the expression levels of a few CDF-like genes were higher than the expression level of NL, which may explain why some CO and FT genes in our study were down-regulated.

The TCP gene family encodes transcription factors that control diverse aspects of the developmental and growth traits of plants. According to the structure of the conserved region, they are classified into classes I and II in Arabidopsis[41], orchestrating the switch from vegetative to reproductive to senescence stages in the life cycle. They may be upstream and activate of SOC1 to modulate flowering[41]. TCP4 is directly associated with CO promoters through TCP binding sites, and promotes CO expression around dusk, together with other CO activators (FLOWERING BHLH)—FBHs[42]. Furthermore, these TCPs not only interact with the flower-repressing CDFs, but physically interact with FBHs to facilitate CO transcription[43]. Our research demonstrated that two transcription factor TCP4 genes (unigene0009351, unigene0033768), and some other TCP genes, such as TCP 15, TCP 18, were up-regulated. Moreover, the expression level of CO-like genes increased, indicating that TCP TFs may activate CO transcript expression in pitaya induced to flower through supplementary lighting.

Figure 13 shows the partial regulatory pathways of flowering based on the genes selected in this study and reference[44]. CO acts on the FT promoter region and is a major positive regulator of FT[31]. CDF binds to the CO promoter region to inhibit CO expression, and CDF is inhibited by F-box protein (Flavinbindingkelch-repeat 1, FKF1), because FKF1 can abolish CDF inhibition of CO and induce flowering[45]. This study also identified and verified the down-regulation of FKF1 transcription by qPCR, which is related to the biological state of flowering pitaya. This pathway map only speculates partial network on the flowering of pitaya, and the complete regulatory network and verification needs further exploration.
Figure 13 Pathways regulating the floral transition in *Hylocereus undulatus* Britt. Photoperiod promote the floral transition. Arrows indicate activation and lines ending with a vertical line indicate repression. Genes in red type indicate up-regulation and genes in green type indicate down-regulation. Genes in blue type indicate no significant expression. Genes in black type were not found by transcriptome data.

4 Conclusion
In summary, this study revealed comparative transcriptomic profiles under *supplementary lighting* and control treatments as investigated by de novo RNA-seq and quantitative RT-PCR. At the transcriptomic level, we detected several DEGs in different developmental stages that performed different functions in the time-course response to *supplementary-lighting-induced flowering*. Such DEGs may interact with each other, and some key genes may regulate the network of floral-induction pathways. These findings will enhance our understanding of the molecular mechanisms involved in flowering induction in pitaya and aid the development of biotechnologies aimed at creating a variant of pitaya that is less sensitive to light conditions and blooms naturally in winter, through genetic modification or targeted gene mutation. Biotechnological developments such as this are needed to reduce energy costs and sustain the pitaya industry in the winter in Hainan, China.

5 Methods

**5.1 Plant material and experimental design**

The plant material, 'Jindu No.1' red flesh pitaya, was cultivated at the Enhong Agricultural Technology Company in Dongfang city, Hainan province. One-year-old pitaya plants were selected for the experiment. One group of plants (the light-treated group) received 5 hours of supplementary light every night from 6.30 p.m. to 11.30 p.m, between January 1 and March 31, 2018. The control group of plants received no supplementary light. Besides the different supplementary light conditions, all other variables (such as watering frequency and fertilization use) were the same for both groups.
This research performed the lighting treatment experiment and obtained samples at four different stages: NL, represents control group (no light, no flowering); L0 (no flowering under light treatment); L1, (flower bud stage under light treatment); L2, (1 week post bud stage under light treatment). We used 15W LED mixed light bulbs as the light source. The light spectrum ratio of red light to blue light was 7:1, and the major wavelength was $\lambda = 580\, \text{nm}$.

5.2 Sample preparation and sequencing
We performed three biological replicates for each stage, each involving four plant samples for mixing well. Every sample was selected to be 10 cm in length from the top of the stem and then included 2 cm of tissue around the base of the thorn. These samples were then frozen immediately using liquid nitrogen and stored at $-80\, ^\circ\text{C}$ before being sent to Gene Denovo Biotechnology Co. (Guangzhou, China) for de novo and RNA-seq using the Illumina HiSeqTM 4000 platform. Raw data were filtered by removing reads that included adapter, poly-N$>10\%$, and low-quality reads in order to obtain clean reads as a foundation for further analysis. De novo assembly of the transcriptome was carried out using the short-read assembly program Trinity[46]. The rRNA-free, high-quality clean reads were mapped to the reference transcriptome using the short-read alignment tool Bowtie2[47], employing its default parameters. The RPKM (reads per kb per million reads) measure was used to determine the expression level of each gene, which was calculated as the fold change in expression of a gene expressed in different samples.

5.3 Identification of DEGs
In our study, genes with a false discovery rate (FDR) $\leq 0.05$ and an absolute value of Log2Ratio $\geq 1$ in a comparison were determined to be significantly differentially expressed genes (DEGs) that could be subjected to enrichment analysis of GO (gene ontology) functions and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. All DEGs were mapped to GO terms in the Gene Ontology database (http://www.geneontology.org/), and gene numbers were calculated for every term. Significantly enriched GO terms were defined by a corrected $p$ value $\leq 0.05$. Pathway enrichment analysis, which can be calculated using the same formula as that in GO analysis, identified significantly enriched metabolic pathways or signal transduction pathways among DEGs. Our raw data and processed data
have been uploaded in the Gene Expression Omnibus (GSE125083) at the NCBI

5.4 Quantitative RT-PCR

Total RNA was separately extracted from four groups, using the CTAB (cetyltrimethylammonium bromide) method. The resulting RNA (1 µg) was treated with gDNA Eraser to remove contaminating genomic DNA and synthesized to the first-strand complementary DNA using a PrimeScript RT reagent kit with gDNA Eraser (TaKaRa) according to the manufacturer’s instructions. Primer pairs for quantitative RT-PCR were designed using Primer Premier 5.0 (PREMIER Biosoft International, Palo Alto, California, USA). The cDNA reverse-transcription products were used as templates for qRT-PCR. An Applied Biosystems 7500 Real-Time PCR System (Life Technologies, USA) was used for quantitative RT-PCR, and the reaction consisted of 10 µL 2×SYBR Premix Ex Taq II, 0.4 µL 50×ROX Reference Dye or Dye II, 2 µL cDNA, 6 µL sterilized purified water, 0.8 µL forward primer (10 µM), and 0.8 µL reverse primer (10 µM) in a total volume of 20 µL. No-template controls were also prepared for each primer pair. The relative quantification method (Delta-Delta cycle threshold) was used to evaluate quantitative variation between the replicates examined. We selected 13 DEGs for quantitative RT-PCR (Table 3).

Table 3 Selected genes used for quantitative RT-PCR

List Of Abbreviations
API: APETALA1; BR: brassinosteroids; CDF: cycling DoF factor; CO: CONSTANS; cry: cryptochrome;
DEG: differentially expressed gene; FD: FLOWERING LOCUS D; FDR: false discovery rate; FT:
FLOWERING LOCUS T; GA: gibberellin; GC: guanine-cytosine; GO: gene ontology; JA: jasmonic acid;
JAZ: jasmonate-ZIM domain (proteins); LD: long day; LFY: LEAFY; phyA: phytochrome A; RPKM: reads per kilobase per million reads; SOC1: SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1; TCP:
TEOSINTE BRANCHED 1/ CYCLOIDEA/ PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR; FKF:flavin-binding kelch repeat F-box protein 1; PRR: pseudo response regulator; FBH :FLOWERING BHLH; TF:
transcription factor.

Declarations

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Competing interests**

The authors declare that they have no competing interests.

**Author contributions**

TH* conceived the study and participated in the design of all experiments. XR# performed the qRT-PCR analysis and statistical analysis and wrote the manuscript. LCL and XM participated in sample preparation and discussions about the experiments. All authors read and approved the final manuscript.

**Funding**

This work was supported by the Key Research and Development Project of Hainan Province (ZDYF2018080), the national pitaya breeding joint project of China (NYLHGG201801), and Research of Pitaya Breeding Technology and Pitaya Planting Demonstration Base Construction (2018-1596-4).

**Acknowledgements**

We would like to thank the Charlesworth Author Services for language editing services.

**References**

1. Lim TK: **Hylocereus undatus**. In: *Edible Medicinal and Non-Medicinal Plants: Volume 1, Fruits*. Dordrecht: Springer Netherlands; 2012: 650-655.

2. Tian X, Hongli LI, Yun HE, Hong Q, Wenbin HU, Qiong LI: Research Progress of Pitaya. *Northern Horticulture* 2015, **18**(18):188-193.

3. Li D, Arroyave Martinez MF, Shaked R, Tel-Zur N: Homozygote Depression in Gamete-Derived Dragon-Fruit (Hylocereus) Lines. *Frontiers in plant science*
4. Song H, Chu Q, Yan F, Yang Y, Han W, Zheng X: Red pitaya betacyanins protects from diet-induced obesity, liver steatosis and insulin resistance in association with modulation of gut microbiota in mice. Journal of gastroenterology and hepatology 2016, 31(8):1462-1469.

5. Suh DH, Lee S, Heo do Y, Kim YS, Cho SK, Lee S, Lee CH: Metabolite profiling of red and white pitayas (Hylocereus polyrhizus and Hylocereus undatus) for comparing betalain biosynthesis and antioxidant activity. Journal of agricultural and food chemistry 2014, 62(34):8764-8771.

6. Song H, Chu Q, Xu D, Xu Y, Zheng X: Purified Betacyanins from Hylocereus undatus Peel Ameliorate Obesity and Insulin Resistance in High-Fat-Diet-Fed Mice. Journal of agricultural and food chemistry 2016, 64(1):236-244.

7. Nerd A, Mizrahi Y: Reproductive biology of cactus fruit crops. In., vol. 18: John Wiley & Sons, Inc.; 1997: 321-346.

8. Jiang YL, Liao YY, Lin TS, Lee CL, Yang WJ: The Photoperiod-regulated Bud Formation of Red Pitaya (Hylocereus sp.). Hortscience 2012, 47(8):1063-1067.

9. Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T: Photoperiodic flowering: time measurement mechanisms in leaves. Annual review of plant biology 2015, 66:441-464.

10. Hepworth SR, Federico V, Dean R, Aidyn M, George C: Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. Embo Journal 2014, 21(16):4327-4337.

11. Pan R, Xu L, Wei Q, Wu C, Tang W, Oelmuller R, Zhang W: Piriformospora indica promotes early flowering in Arabidopsis through regulation of the photoperiod and gibberellin pathways. 2017, 12(12):e0189791.
12. Cheng H, Qin L, Lee S, Fu X, Richards DE, Cao D, Luo D, Harberd NP, Peng J: **Gibberellin regulates Arabidopsis floral development via suppression of DELLA protein function.** *Development* 2004, **131**(5):1055-1064.

13. Niwa T, Suzuki T, Takebayashi Y, Ishiguro R, Higashiyama T: **Jasmonic acid facilitates flower opening and floral organ development through the upregulated expression of SIMYB21 transcription factor in tomato.** 2018, **82**(2):292-303.

14. Aloni R, Aloni E, Langhans M, Ullrich CI: **Role of auxin in regulating Arabidopsis flower development.** *Planta* 2006, **223**(2):315-328.

15. De SN: **The ENCODE project.** *Nature methods* 2012, **9**(11):1046.

16. Bonzon M, Hug M, Wagner E, Greppin H: **Adenine nucleotides and energy charge evolution during the induction of flowering in spinach leaves.** *Planta* 1981, **152**(3):189-194.

17. Reinhardt D, Pesce ER, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C: **Regulation of phyllotaxis by polar auxin transport.** *Nature* 2003, **426**(6964):255-260.

18. Li J, Li Y, Chen S, An L: **Involvement of brassinosteroid signals in the floral-induction network of Arabidopsis.** *Journal of experimental botany* 2010, **61**(15):4221-4230.

19. Oh Y, Baldwin IT, Galis I: **A jasmonate ZIM-domain protein NaJAZd regulates floral jasmonic acid levels and counteracts flower abscission in Nicotiana attenuata plants.** *PloS one* 2013, **8**(2):e57868.

20. Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K: **Hd3a protein is a mobile flowering signal in rice.** *Science* 2007, **316**(5827):1033-1036.

21. Blazquez MA, Soowal LN, Lee I, Weigel D: **LEAFY expression and flower initiation**
22. Yanagisawa S: **Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants.** *Plant & cell physiology* 2004, **45**(4):386-391.

23. Yang J, Yang MF, Zhang WP, Chen F, Shen SH: **A putative flowering-time-related Dof transcription factor gene, JcDof3, is controlled by the circadian clock in Jatropha curcas.** *Plant science : an international journal of experimental plant biology* 2011, **181**(6):667-674.

24. Song GQ, Walworth A, Zhao D, Hildebrandt B, Leasia M: **Constitutive expression of the K-domain of a Vaccinium corymbosum SOC1-like (VcSOC1-K) MADS-box gene is sufficient to promote flowering in tobacco.** *Plant cell reports* 2013, **32**(11):1819-1826.

25. Mimida N, Kidou S, Iwanami H, Moriya S, Abe K, Voogd C, Varkonyi-Gasic E, Kotoda N: **Apple FLOWERING LOCUS T proteins interact with transcription factors implicated in cell growth and organ development.** *Tree physiology* 2011, **31**(5):555-566.

26. Balsemao-Pires E, Andrade LR, Sachetto-Martins G: **Functional study of TCP23 in Arabidopsis thaliana during plant development.** *Plant physiology and biochemistry : PPB* 2013, **67**:120-125.

27. Fengxia Y, Xiaopeng W, Guoli G, Jin T: **Cloning and sequence analysis of houskeeping genes Actini and UBQ from pitaya** *Guizhou Agricultural Science* 2013, **41**(09):1-4.(in Chinese).

28. Yamaguchi N, Wu MF, Winter C, Berns M, Nole-Wilson S, Yamaguchi A, Coupland G, Krizek B, Wagner D: **A Molecular Framework for Auxin-Mediated Initiation of Flower Primordia.** *Developmental cell* 2013, **24**(3):271-282.
Yuguo X, Yi C, Tatsiana C, Mulder PPJ, Jeroen H, Angela H, Adamantia A, Corinne M, Jean-Benoit M, Ludovico D: **OsJAR1 is required for JA-regulated floret opening and anther dehiscence in rice.** *Plant molecular biology* 2014, **86**(1-2):19-33.

Mockler T, Yang H, Yu X, Parikh D, Cheng YC, Dolan S, Lin C: **Regulation of photoperiodic flowering by Arabidopsis photoreceptors.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(4):2140-2145.

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G: **Photoreceptor regulation of CONSTANS protein in photoperiodic flowering.** *Science* 2004, **303**(5660):1003-1006.

Shim JS, Kubota A, Imaizumi T: **Circadian clock and photoperiodic flowering in Arabidopsis: CONSTANS is a hub for signal integration.** *Plant physiology* 2017, **173**(1):5-15.

Fujiwara S, Oda A, Yoshida R, Niinuma K, Miyata K, Tomozoe Y, Tajima T, Nakagawa M, Hayashi K, Coupland G: **Circadian clock proteins LHY and CCA1 regulate SVP protein accumulation to control flowering in Arabidopsis.** *The Plant cell* 2008, **20**(11):2960-2971.

Manoharan RK, Han JS, Vijayakumar H, Subramani B, Thamilarasan SK, Park JI, Nou IS: **Molecular and Functional Characterization of FLOWERING LOCUS T Homologs in Allium cepa.** *Molecules* 2016, **21**(2):217; 210.3390/molecules21020217.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C et al: **FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis.** *Science* 2007, **316**(5827):1030-1033.

Hu JY, Zhou Y, He F, Dong X, Liu LY, Coupland G, Turck F, de Meaux J: **miR824-**
Regulated AGAMOUS-LIKE16 Contributes to Flowering Time Repression in Arabidopsis. *The Plant cell* 2014, **26**(5):2024-2037.

37. Cao S, Kumimoto RW, Gnesutta N, Calogero AM, Mantovani R, Holt BF, 3rd: A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in Arabidopsis. *The Plant cell* 2014, **26**(3):1009-1017.

38. Yu Y, Liu Z, Wang L, Kim SG, Seo PJ, Qiao M, Wang N, Li S, Cao X, Park CM et al: WRKY71 accelerates flowering via the direct activation of FLOWERING LOCUS T and LEAFY in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* 2016, **85**(1):96-106.

39. Noguero M, Atif RM, Ochatt S, Thompson RD: The role of the DNA-binding One Zinc Finger (DOF) transcription factor family in plants. *Plant science : an international journal of experimental plant biology* 2013, **209**:32-45.

40. Goralogia GS, Liu TK, Zhao L, Panipinto PM, Groover ED, Bains YS, Imaizumi T: CYCLING DOF FACTOR 1 represses transcription through the TOPLESS co-repressor to control photoperiodic flowering in Arabidopsis. *The Plant journal : for cell and molecular biology* 2017, **92**(2):244-262.

41. Lucero LE, Manavella PA, Gras DE, Ariel FD, Gonzalez DH: Class I and Class II TCP Transcription Factors Modulate SOC1-Dependent Flowering at Multiple Levels. *Molecular plant* 2017, **10**(12):1571-1574.

42. Kubota A, Ito S, Shim JS, Johnson RS: TCP4-dependent induction of CONSTANS transcription requires GIGANTEA in photoperiodic flowering in Arabidopsis. 2017, **13**(6):e1006856.

43. Lenhard M, Liu J, Cheng X, Liu P, Li D, Chen T, Gu X, Sun J: MicroRNA319-regulated TCPs interact with FBHs and PFT1 to activate CO transcription and control
flowering time in Arabidopsis. *PLoS genetics* 2017, **13**(5):e1006833.

44. Zhang JZ, Ai XY, Sun LM, Zhang DL, Guo WW, Deng XX, Hu CG: Transcriptome profile analysis of flowering molecular processes of early flowering trifoliate orange mutant and the wild-type [Poncirus trifoliata (L.) Raf.] by massively parallel signature sequencing. *BMC genomics* 2011, **12**(63):https://doi.org/10.1186/1471-2164-1112-1163.

45. Young Hun S, Smith RW, To BJ, Millar AJ, Takato I: FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 2012, **336**(6084):1045-1049.

46. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Xian A, Fan L, Raychowdhury R, Zeng Q: Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature biotechnology* 2011, **29**(7):644-652.

47. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J: SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 2009, **25**(15):1966-1967.

Tables

Table 1 Results of assembly

| Gene number | GC percentage | N50 | Max length | Min length | Average length | Total assembled bases |
|-------------|---------------|-----|------------|------------|-----------------|-----------------------|
| 68113       | 41.96         | 1730| 36165      | 201        | 889             | 60580077              |

Table 2 Results of annotations in four databases

| Total unigenes | Nr     | Swissprot | KOG     | KEGG | Genes with annotation | Genes without Annotation |
|----------------|--------|-----------|---------|------|------------------------|--------------------------|
| 68113          | 29782  | 20716     | 18088   | 11059| 29959                  | 38164                   |

Table 3 Selected genes used for quantitative RT-PCR
| Gene ID               | Function                               | Forward primer (5’→3’) | Reverse primer (5’→3’) |
|----------------------|----------------------------------------|------------------------|------------------------|
| unigene0031426       | Cyclic dof factor 1                     | GAGCAAGAGAAGATGGGGA    | GCAAAATGGCGGAACAGGAA   |
| unigene0025406       | Zinc finger protein CONSTANS-LIKE 2    | TTTGTTCGGGATGGATCTGGA  | TTGGGTGGAATTTGAATAAGG  |
| unigene0019970       | HEADING DATE 3A-like                   | GGCAACCAACTGAGCGGAAGA  | GCCATCCATTGTGACGAA     |
| unigene0045107       | Response regulator-like APRRS           | AAGACCGCACCACCTTTTC    | CATTCCCACATACGACCCAC   |

**NLvsL1**

| Gene ID               | Function                               | Forward primer (5’→3’) | Reverse primer (5’→3’) |
|----------------------|----------------------------------------|------------------------|------------------------|
| unigene0026114       | Cyclic dof factor 3-like                | GGTCGTGGAAGATGAACATG   | CTATAAAAGCGGGCTATCGAG  |
| unigene0027129       | CONSTANS-LIKE 15                       | TGGGAAGAAGCACCACCTCA   | AGCGGGACACCACGGAATAA   |
| unigene0028118       | CONSTANS-LIKE 4                        | GCTCTTCGGCGAGCTATGAA   | TGAAGTCCAGTCAGTGGTAAGG |
| unigene0036436       | Circadian clock associated1            | CCTCAAGATCCCTAATCGAC   | CCAAGAAGACCCGAAAAAC    |
| unigene0049229       | Circadian clock associated1            | CCAAGCATGCGAGCAGTTAAG  | GCTGGTTGAGTTTGGGTAAGAT |
| unigene0005649       | Flowering locus T                      | AGTGGGAGATCGGAATGGTAG  | GCTTTGGAATTCTCAGTACCA  |
| unigene0002396       | Phytochrome B                          | GGTCTATTTCCTTTCCCACCT  | AGAGAGCTGCTCCATCCGACT  |
| unigene0009351       | Transcription factor TCP4              | CACTGGAGGAGGTGATTATG    | GGATGGAGGTCGTTTGC      |
| unigene0005880       | Transcription factor TCP15             | GCCGCTGTGGATATACCTGCA  | GCCCGACCTGTATCCTGAG    |

**Figures**

![Image A](image-a.png)  ![Image B](image-b.png)  ![Image C](image-c.png)  ![Image D](image-d.png)
Supplementary light experiment. A: NL represents the control group (no flowering under no light); B: L0 represents the no-flowering stage in the light treatment group; C: L1 represents the flower bud stage after light treatment; D: L2 represents 1 week post bud stage under light treatment; E: flowering of pitaya induced by light treatment in the winter season.
Figure 2

Length distribution of all unigenes.
Figure 3

Venn diagram of the four database annotations
Figure 4

Relationship analysis of the samples. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage under light treatment.
Gene statistics for DEGs in different samples. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage under light treatment.

Figure 5
Heatmap of the DEGs associated with various plant hormones. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: stage of flowering 1 week after light treatment.
Figure 7

Heatmap of the DEGs associated with CO and FT. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage after light treatment.
Figure 8

Heatmap of the three main TFs involved in pitaya flowering. NL: stage of no light; L0: stage of no flowering after light treatment; L1: flower bud stage after light treatment; L2: 1 week post bud stage after light treatment.
Relative expression of the floral development genes or transcription factors involved in the
flowering process for the NL and LO samples was determined by quantitative RT-PCR
analyses. Expression was normalized to that of UBQ. The transcript levels from the NL
sample were set at 1. The blue columns represent the expression levels determined by qRT-
PCR (left axis), while the lines represent the gene expression levels determined by RNA-seq
(right axis). The transcript abundances of the genes encoding cyclic DoF factor 1 protein (A),
zinc finger protein CONSTANS-LIKE 2 (B), HEADING DATE 3A-like (C), and response
regulator-like APRR5 (D) were determined and compared across the time course of light-
induced treatment. Data represent means ± SD for the three replicates (n=3). The
lowercase letters above the columns indicate a significant difference at p≤0.05 between the
columns by T test using SPSS statistical software. Data in columns with the same letters
showed no significant difference (p>0.05).
Relative expression of the floral development genes or transcription factors involved in the flowering process for the NL and L1 samples was determined by quantitative RT-PCR analyses. Expression was normalized to that of UBQ. The transcript levels from the NL sample were set at 1. The blue or red columns represent the expression levels determined by qRT-PCR (left axis), while the lines represent the gene expression levels determined by RNA-seq (right axis). The transcript abundances of the genes encoding cyclic DoF factor 3-like protein (A), CONSTANS-LIKE 15 (B), CONSTANS-LIKE 4 (C), circadian clock associated1 (D, E), flowering locus T (F), phytochrome B (G), transcription factor TCP4 (H), and transcription factor TCP15 (I) were determined and compared across the time course of light-induced treatment. Data represent means ± SD for the three replicates (n=3). The
lowercase letters above the columns indicate a significant difference at $p \leq 0.05$ between the columns by T test using SPSS statistical software. Data in columns with the same letters showed no significant difference ($p > 0.05$).

Figure 11

Different growth periods and tissues of pitaya. 1: old stem, 2: young stem, 3: small flower bud, 4: large flower bud, 4A: pistil, 4B: stamen, 4C: style, 4D: flower scales, 5: small green fruit, 6: big green fruit, 7: red pitaya, 8E: small green peel, 8F: small green flesh, 9: big green fruit, 9G: big green peel, 9H: big green flesh, 10: red fruit, 10I: red peel, 10J: red flesh
Figure 12

The result of CDF-3 like(Unigene0026114) and CO-15 like(Unigene0035118) tissue-specific expression. Expression was normalized to that of UBQ. The transcript levels from the old stem were set at 1.
Pathways regulating the floral transition in Hylocereus undulatus Britt. Photoperiod promote the floral transition. Arrows indicate activation and lines ending with a vertical line indicate repression. Genes in red type indicate up-regulation and genes in green type indicate down-regulation. Genes in blue type indicate no significant expression. Genes in black type were not found by transcriptome data.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
Additional file 3.docx
Additional file 2.docx
Additional file 6.docx
Additional file 8.docx
Additional file 7.docx
Additional file 9.docx
Additional file 10.docx
Additional file 4.docx
Additional file 5.docx
Additional file 1.docx
