

# De novo sequencing of the transcriptome reveals regulators of the floral transition in *Fargesia macclureana* (Poaceae)

Ying Li

International Center for Bamboo and Rattan <https://orcid.org/0000-0001-6005-1174>

Chunxia Zhang

Nanjing Forestry University

Kebin Yang

International Center for Bamboo and Rattan

Jingjing Shi

International Center for Bamboo and Rattan

Yulong Ding

Nanjing Forestry University

Zhimin Gao (✉ [gaozhimin@icbr.ac.cn](mailto:gaozhimin@icbr.ac.cn))

<https://orcid.org/0000-0003-4464-7159>

---

## Research article

**Keywords:** Transcriptome, Floral transition, Bamboo, Qinghai–Tibet Plateau

**Posted Date:** January 3rd, 2020

**DOI:** <https://doi.org/10.21203/rs.2.10521/v4>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published on December 30th, 2019. See the published version at <https://doi.org/10.1186/s12864-019-6418-2>.

# Abstract

**Background:** *Fargesia macclureana* (Poaceae) is a woody bamboo species found on the Qinghai–Tibet Plateau (QTP) approximately 2,000 ~ 3,800 m above sea level. It rarely blossoms in the QTP, but it flowered 20 days after growing in our lab, which is in a low-altitude area outside the QTP. To date, little is known regarding the molecular mechanism of bamboo flowering, and no studies of flowering have been conducted on wild bamboo plants growing in extreme environments. Here, we report the first de novo transcriptome sequence for *F. macclureana* to investigate the putative mechanisms underlying the flowering time control used by *F. macclureana* to adapt to its environment. **Results:** Illumina deep sequencing of the *F. macclureana* transcriptome generated 140.94 Gb of data, assembled into 99,056 unigenes. A comprehensive analysis of the broadly, specifically and differentially expressed unigenes (BEUs, SEUs and DEUs) indicated that they were mostly involved in metabolism and signal transduction, as well as DNA repair and plant-pathogen interactions, which may be of adaptive importance. In addition, comparison analysis between non-flowering and flowering tissues revealed that expressions of *FmFT* and *FmHd3a*, two putative *F. macclureana* orthologs, were differently regulated in NF- vs F- leaves, and carbohydrate metabolism and signal transduction were two major KEGG pathways that DEUs were enriched in. Finally, we detected 9,296 simple sequence repeats (SSRs) that may be useful for further molecular marker-assisted breeding. **Conclusions:** *F. macclureana* may have evolved specific reproductive strategies for flowering-related pathways in response to photoperiodic cues to ensure long vegetation growing period. Our findings will provide new insights to future investigations into the mechanisms of flowering time control and adaptive evolution in plants growing at high altitudes.

## Background

The flowering time is of crucial importance to ensure the reproductive success of flowering plants. Previous results have indicated that the floral transition is orchestrated by several parallel and interactive genetic pathways that are regulated by a variety of environmental and endogenous signals [1]. Many key genes and regulatory networks have been identified in herbaceous annual plants such as *Arabidopsis* [2, 3], rice [4], gourds [5], potato [6] and sorghum [7]. However, much less is known about such regulation in perennial plants. Despite the increasing attention on perennial dicotyledonous woody plants such as poplar [8, 9], eucalyptus [10] and citrus [11] species, to date, the molecular mechanism underlying floral regulation in monocotyledonous woody plants remains elusive. Furthermore, previous studies investigated flowering mainly by artificially altering the external signals (e.g. photoperiod and light intensity) and did not assess the impact of the original environment on the adaptive evolution of species-specific reproductive strategies.

Bamboo plants are an important group in the Bambusoideae subfamily of the monocotyledonous Poaceae. They exhibit a wide degree of variation in the timing (1-120 years) and nature (sporadic vs. gregarious) of flowering among species [12]. Sporadic flowering involves flowering in only a few isolated clumps, which set little or no seed and usually remain alive afterward [13]. In contrast, gregarious flowering involves all individuals of a species regardless of age and/or location within and among the populations at the same time, which is usually followed by death and seed setting [14]. And the simultaneous death of many individuals triggers serious ecological consequences, including changes in the population dynamics of neighboring plants, differences in soil properties, various effects on endangered animals that depend on bamboo [15], and the knock-on effects on human economies in many parts of the world [16]. Therefore, dissecting the regulators that

control the unique life history of bamboo may be of use for plant ecology and human society. However, to date, little is known regarding the molecular mechanisms of bamboo flowering, in part because of the sporadic occurrence of these flowering episodes and the long intervals between events.

Many genes have been identified as regulators of reproductive development in different bamboo species, including the MADS-box transcription factors [17-19], *CONSTANS (CO)* [20] and *FLOWERING LOCUS T (FT)* [21], among others. In addition, studies of sequenced transcriptomes have identified microRNAs related to floral development [22-24]. However, samples collected in these analyses were limited to mature spikelets or to different spikelets at different development stages. Thus, it is likely that dynamic changes in genes occurring at different development stages may be missing. In addition, the specific response of particular tissues to internal and external cues and how plants integrate these signals to regulate different phases of reproductive development (including the floral transition, florigen transport, and floral organ specification) has not yet been elucidated in bamboo. Furthermore, no studies of flowering have been conducted on wild bamboo plants growing in extreme environments.

Here, we took advantage of an unexpected flowering event in highland arrow bamboo, *Fargesia macclureana* [25], and performed the first *de novo* transcriptome analysis. This transcriptome includes data from six different tissues collected at different development stages, including inflorescences in the initial and peak flower stage (I- and P- spikelets), branchlets, and leaves from both flowering and non-flowering bamboo plants (F/NF-branchlets and F/NF-leaves). *F. macclureana* is a woody bamboo species found in areas 2,000 ~ 3,800 m above sea level on the Qinghai–Tibet Plateau (QTP) (Fig. 1), which is the highest and largest plateau in the world. The growth environment of the QTP is characterized by low temperature and low oxygen availability, reduced pathogen incidence, and intense radiation [26]. *F. macclureana* rarely blossoms in the QTP, but it flowered 20 days after growing in our lab, which is in a low-altitude area outside the QTP. Our goal is to use the transcriptomic data to gain a deeper understanding of the mechanisms underlying the control of flowering time and the adaptation of *F. macclureana* to the complex extreme conditions of the QTP. On one hand, we expect to detect regulatory hubs involved in the flowering mechanisms. On the other hand, we aim to discover signs of the adaptive evolutionary changes in *F. macclureana* in response to the harsh environmental conditions in the QTP, which may, in turn, provide a broader insight into the adaptive mechanisms for plants that grow at high altitudes.

## Results

### ***De novo* transcriptome assembly yielded 99,056 unigenes**

Illumina deep sequencing of the *F. macclureana* transcriptome generated 140.94 Gb of data, including 471,537,304 clean reads in 18 unique samples (Additional file 1: Table S1). The average Q20 (sequencing error rate less than 1%) and Q30 (sequencing error rate less than 0.1%) percentages were 95.64% and 89.95% respectively. The GC content of all samples ranged from 53.78% to 55.86%, with an average of 54.81%. Sample data were assembled into 289,122 transcript scaffolds, with an N50 and average length of 1,765 bp and 1,183 bp, respectively. The final *de novo* assembly included 99,056 unigenes, with an N50 and average length of 1,587 bp and 926 bp, respectively. Among these unigenes, 71.02% (70,354) were shorter than 1,000 bp and 12.06% (11,950) were longer than 2,000 bp (Table 1).

## Most unigenes were functionally annotated and classified

A total of 47,306 unigenes were annotated (Additional file 2: Table S2). Of these, 45,516 (96.22%) unigenes were found to encode products that showed significant similarity to characterized proteins in the non-redundant protein sequence database (Nr) at an E-value threshold of  $10^{-5}$  (Table 2). We also found that 7,027 (15.45 %) unigenes showed similarity to genes found in rice, 11.33% were similar to those found in *Brachypodium distachyon*, and we also found a significant proportion of the unigenes that were similar to those found in *Setaria italica*, *Oryza brachyantha*, and *Zea mays* (Fig. 2a). We identified 24,847 (52.52%), 28,317 (59.86%) and 43,909 (92.82%) unigenes that showed significant matches to entries in the Swiss-Prot, Pfam, and eggnoG databases, respectively (Table 2). Many unigenes expressed in the *F. macclureana* transcriptome were functionally annotated as regulators of plant responses to evolutionarily important phenotypes, including membrane stabilization, heat stress response and pathogen defense (Additional file 2: Table S2).

## Functional annotation indicated that many unigenes were involved in metabolism and genetic information processing

We were able to annotate 13,128 unigenes (27.75% of the total) in 25 different categories of the COG (clusters of orthologous groups) classification database (Fig. 2b). Of these, the cluster for “General function prediction only” (3,277, representing 24.96% of the 13,128 unigenes annotated by this database) was the largest group, followed by “Replication, recombination and repair” (2,202, 16.77%), “Transcription” (1,571, 11.97%), and “Translation, ribosomal structure and biogenesis” (1,429, 10.88%). The “Signal transduction mechanisms”, “posttranslational modification, protein turnover, chaperones”, “carbohydrate and amino acid transport and metabolism” and “transport and metabolism” categories also contained a significant proportion of the annotated unigenes.

GO enrichment analysis indicated that these predicted unigenes were categorized into three main categories—i.e. biological process (BP), cellular component (CC), and molecular function (MF). As shown in Fig. 2c, for unigenes that were enriched in the BP category, they were mainly involved in biological processes related to reproduction, posttranslational modification and signal transduction; as for those in the CC category, they were mainly involved in cellular components related to membrane, ubiquitin ligase complex, mitochondrion, chloroplast and etc.; while for those in the MF category, they were mainly involved in molecular functions related to signaling transduction (e.g. “ATP binding”, “zinc ion binding”, “protein kinase activity”, and etc.) (Additional file 3: Table S3).

We also mapped 14,307 unigenes (representing 30.24% of the total) to six different KEGG subsystems, including metabolism, genetic information processing, environmental information processing, cellular processes, and organismal systems. As shown in Fig. 3, the majority of these unigenes (7,922, representing 66.17% of the 14,307 unigenes classified using KEGG annotations) were assigned to metabolic pathways, including carbohydrate metabolism, energy metabolism, and others. In addition, 4,024 unigenes (28.13%) were assigned to genetic information processing, including transcription, translation, and folding, and 474 unigenes (3.31%) were found to be related to membrane transport and signal transduction. We also found 707 genes (4.94%) that were related to transport and catabolism and 377 genes (2.64%) related to environmental adaptation.

## **Most BEUs were involved in genetic information processing, environmental adaptation and signal transduction**

As shown in the Venn diagram (Fig. 4a), we found nearly equal numbers of unigenes that were broadly and specifically expressed in I-spikelets, P-spikelets, F-branchlets, and F-leaves. COG analysis indicated that most BEUs were clustered in signal transduction mechanisms (T), replication, recombination and repair (L), and transcription (K), besides general function prediction only (R). GO enrichment analysis for these BEUs indicated that they were also mainly involved in reproduction, environmental adaptation and signal transduction, which was largely similar with that for all predicted unigenes (Additional file 4: Table S4-a).

KEGG enrichment analysis also indicated that these BEUs were mainly enriched in pathways related to environmental adaptation (including circadian rhythm, endocytosis, and plant-pathogen interactions), signal transduction (including plant hormone signal transduction, phosphatidylinositol signaling system, and inositol phosphate metabolism) and genetic information processing (including spliceosome, mRNA surveillance, and RNA transport and degradation; Additional file 4: Table S4-b).

## **The SEUs were mostly involved in carbohydrate metabolism, energy metabolism, and environmental adaptation**

As shown in Fig. 4a, we identified 10,653 unigenes that were specifically expressed in spikelets, including 5,528 and 5,025 unigenes in I- and P-spikelets, respectively. We also found 9,067 and 7,437 unigenes that were specifically expressed in F-branchlets and F-leaves, respectively. COG annotation indicated that the distribution patterns of SEUs among the 26 terms were similar, with the number of SEUs within each term varying among the three tissues (Fig. 4b).

The GO enrichment analysis indicated that these SEUs not only shared some common GO terms, but also had some particular ones. As shown in Fig. 4c and Additional file 4: Table S4-c, for those SEUs that were enriched in the BP category, they were broadly involved in several important biological processes, including “protein phosphorylation”, “regulation of flower development”, “protein ubiquitination”, “regulation of transcription, DNA-templated”, “reciprocal meiotic recombination” and “meiotic chromosome segregation”. In addition, SEUs in I- and P- spikelets were also involved in some processes related to reproduction; and those in F-branchlets were mainly involved in processes related to posttranslational modification; while those in F-leaves were mainly involved in processes related to plant-pathogen interaction. As for those in the CC category, they were broadly involved in several important cellular components, including “mitochondrion”, “plasma membrane” and “plastid”. In addition, SEUs in I- and P-spikelets were also involved in ribosome and mitochondria; and those in F-branchlets were mainly involved in endoplasmic reticulum and proteasome; and those in F-leaves were mostly involved in chloroplast. As for those in the MF category, they were broadly involved in several molecular functions, including “ATP binding”, “ubiquitin-protein transferase activity” and “protein tyrosine kinase activity”. In addition, SEUs in I- and P-spikelets were also involved in DNA and microtubule binding; those in F-branchlets were also enriched in oxidoreductases activities; and those in F-leaves were also enriched in enzymes involved in carbohydrate metabolism.

As shown in Additional file 5: Fig. S1, KEGG pathway analysis indicated that SEUs in I- and F-spikelets mainly mapped to the ribosome pathway, with those in F-branchlets mainly mapped to the ribosome, amino acid biosynthesis, and carbon metabolism pathways, and those in F-leaves mainly mapped to KEGG pathways related to energy metabolism (including oxidative phosphorylation, fatty acid metabolism, and

photosynthesis), environmental adaptation (e.g. proteasomes), genetic information processing, and various unrelated metabolic pathways (e.g. tryptophan metabolism, beta-alanine metabolism, and N-glycan biosynthesis).

### **DEUs were mostly involved in carbohydrate and energy metabolism, signal transition and environmental adaptation**

As shown in Table 3, many unigenes showed differential expressions across all 15 groups sampled. The number of DEUs in each sample pair ranged from 970 between I- vs P-spikelets to 13,577 in NF-leaves vs I-spikelets. For most pairwise comparisons, the number of up- and down-regulated DEUs was approximately the same, except for four groups, including I- vs P-spikelets, F-branchlets vs both I- and P- spikelets, and F-leaves vs P-spikelets.

The Venn diagram of DEU sets shows that 5,494 unigenes were differentially expressed in F-branchlets/F-leaves vs I- and P-spikelets. For those DEUs that were up-regulated in spikelets, they are mainly mapped to KEGG pathways related to carbohydrate metabolism, plant-pathogen interactions and DNA repair (Fig. 5a). Notably, among the 970 DEUs identified between I- and P-spikelets, 916 up-regulated DEUs were mapped to KEGG pathways related to metabolic activity (Additional file 6: Table S5).

A total of 5,494 unigenes were differentially expressed in the DEU sets of spikelets/F-leaves vs F- branchlets. Upregulated DEUs in F-branchlets were mapped to KEGG pathways including phenylalanine metabolism, phenylpropanoid biosynthesis, ABC transporters, and flavone and flavonol biosynthesis (Fig. 5b). Those that were upregulated in F- and NF-leaves vs F- branchlets were mainly mapped to plant hormone signal transduction, homologous recombination, base excision repair, and mismatch repair (Additional file 6: Table S5). Notably, 3,275 (50.20% of the total) DEUs found between NF- and F-branchlets were upregulated; these were mainly mapped to KEGG pathways related to replication and recombination (Additional file 6: Table S5). Those that were downregulated were mainly mapped to carbon fixation and photosynthesis (Additional file 6: Table S5).

We also found that 6,966 (43.69% of the total) DEUs found in spikelets/F-branchlets vs F-leaves were up-regulated, and were mainly mapped to KEGG pathways related to carbohydrate metabolism (Fig. 5c). 2,492 (49.52%) DEUs in NF-vs F-leaves were up-regulated, and these were mainly mapped to starch and sucrose metabolism (Additional file 6: Table S5). In contrast, downregulated DEUs were mainly mapped to KEGG pathways related to photosynthesis (Additional file 6: Table S5).

### **Two putative *FT* orthologs were regulated differently in the circadian rhythm–plant pathway**

Among the 5,032 DEUs identified between NF- and F-leaves, 70 were mapped to the circadian rhythm–plant KEGG pathway (Additional file 7: Fig. S2) and 10 of them showed differential expressions (Additional file 8: Table S6). Notably, c109220.graph\_c0 and c110963.graph\_c4 were both annotated as *FT* orthologs: the former was a putative bamboo ortholog of *Heading date 3a* (Swissprot: PE=1 SV=1), and the latter was another ortholog of rice *FT*; we designated them as *FmHd3a* and *FmFT*, respectively.

As shown in Fig. 6a, protein sequence alignment indicated that both *FmFT* and *FmHd3a* had high amino acid sequence similarities (77.14%) with the known *FT*/*TFL1* proteins and had the critical amino acids of *FT*/*Hd3a*

proteins. For example, they both carry a conserved phosphatidylethanolamine-binding protein (PEBP) domain, D-P-D-x-P and G-x-H-R motifs, and the invariant histidine (asterisk), all of which are relevant to the ability of PEBP proteins to bind phosphoryl ligands, hence interfering with certain kinases and effectors [27-28]. Furthermore, all five proteins carry tyrosine-139 and tryptophan-143 (triangles) in the PEBP domain, two conserved sites in FT homologs acting as flowering inducers [29]. Whereas, our comparison also revealed that there were differences in amino acid sequences between FTs and Hd3as.

Additionally, phylogenetic analysis indicated that these ten proteins were subdivided into two distinct subgroups (Fig. 6b). FT and Hd3a proteins from *O. sativa* (OsFT, OsRFT1 and OsHd3a), *F. macclureana* (FmFT and FmHd3a) and *A. thaliana* (AtFT) were clustered in a branch. Furthermore, OsFT and FmFT were clustered together, just like OsHd3a and FmHd3a; they were both closer to each other than they were to AtFT.

Interestingly, two FT orthologs were regulated differently in NF- vs F- leaves. *FmHd3a* was significantly upregulated in F-leaves (FDR = 4.23,  $\log_2FC = 5.55$ ), while *FmFT* was significantly downregulated in F-leaves (FDR = 4.25E-07,  $\log_2FC = -4.81$ ). RT-qPCR analysis also showed that *FmHd3a* was significantly more highly expressed in I-/P- spikelets and F-leaves than in NF-leaves or NF- branchlets (Fig. 6c).

### **WGCNA results identified gene modules related to specific tissues**

WGCNA results showed that unigenes expressed in the six different tissues of flowering and nonflowering plants tested here clustered into 18 branches representing 18 different genetic modules (Additional file 9: Fig. S3a). Unigenes within each module were highly co-expressed, while those in different modules were co-expressed to a lower degree (Additional file 9: Fig. S3b). In six of the samples collected, we identified nine significant gene modules including 1,344 unigenes. Here, correlation coefficient of a module with a related trait > 0.7 was used as a threshold of significance (Additional file 9: Fig. S3c). Notably, these six tissues were more strongly divided into clades according to whether they were flowered or not rather than by the differences among tissues (Additional file 9: Fig. S3d).

In addition, the unigenes in gene modules relating to I- and P- spikelets were most strongly enriched in KEGG pathways related to carbohydrate metabolism, genetic information processing, and environmental information processing. In contrast, those related to F- and NF- branchlets were mostly enriched in KEGG pathways related to metabolism, plant hormone signal transduction, and genetic information processing. The gene modules related to F-leaves were enriched in pathways related to plant hormone signal transduction and protein processing, while the gene modules related to NF-leaves were enriched in KEGG pathways related to oxidative phosphorylation (Additional file 10: Table S7).

### **Identification of SSRs**

We detected a total of 9,296 SSRs in 7,668 unigenes longer than 1,000 bp (Additional file 11: Table S8). 1,628 (21.23%) unigenes contained more than one SSR. Mono-nucleotide repeats were the most common (46.28% of all SSRs) at a density of 71 SSRs per Mb, followed by tri- (26.32%) and di- (22.06%) nucleotide repeats, with densities of 40 and 32 SSRs per Mb, respectively (Fig. 7).

## **Discussion**

## Activated *Hd3a* expression probably accelerates flowering in *F. macclureana*

FT is a key floral regulator that controls the timing of flowering and seasonal growth cessation in response to light and the circadian clock in many plant species [8, 10, 30]. In this study, *FmHd3a*, a bamboo ortholog of rice *FT*, was significantly expressed only in tissue samples collected from flowering plants. In rice, *Hd3a* functions as a major photoperiodic flowering regulator and participates in the *OsGI-Hd1-Hd3a* module, which is similar to the *GI-CO-FT* module in *Arabidopsis* [31]. *Hd1* activates and suppresses *Hd3a* expression by promoting heading under the short day (SD) and long day (LD) conditions, respectively [32-33]. As *F. macclureana* rarely blossoms on the QTP, which experiences a long photoperiod with a low ratio of red to far-red light, it may have evolved specific reproductive strategies involving flowering-related pathways in response to photoperiodic cues to ensure long vegetation growing period. It is probably that the weak light intensity with a low proportion of blue light might activate *Hd3a* expression even in the LD conditions, thereby accelerating flowering. Notably, reproduction pathways play an important role in the mechanisms of plant adaptation to extreme environments. Previous studies showed that the phytochrome and flowering time regulatory protein 1 (PFT1) from *Crucihimalaya himalaica*, a close relative of *Arabidopsis* and *Capsella*, grows on the QTP, showed signs of positive selection for adaptive divergence [34]. We presume that the function of *Hd3a* in promoting flowering is likely to be conserved between bamboo and rice, because both of them belong to the Poaceae.

Interestingly, *FmFT* and *FmHd3a* were regulated differently in NF- vs F- leaves, although they were both annotated as FT regulators in the circadian rhythm-plant pathway. Basically, FT sub-family protein acts as florigen; the diverse functions of the *FT* gene family in flowering regulation had been demonstrated in many different plant species [29]. For example, sugar beet BvFT1 repressed flowering and the divergence within three amino acids of an external loop in PEBP domain was demonstrated to be the major cause [35]. Similarly, a single amino acid exchange in PEBP domain was sufficient to convert FT to TFL1, an activator and a repressor [36-37]. In our study, five proteins carry invariant amino acids in positions where variation occurred; however, there were many differences in predicted amino acid sequences between them in PEBP domain. Especially for some positions, amino acids of AtFT and FmHd3a are the same, but different from that of FmFT (Fig. 6a). Possibly, FmFT acts as a repressor FT due to the conversion of certain amino acids in PEBP domain, while FmHd3a acts as an inducer one. Thus, we suspect that the floral transition of *F. macclureana* is regulated by a complex regulatory network in which at least two unique *FT* orthologs interact with the circadian clock pathways. However, how these circadian clock pathways mediate the activation of *FmHd3a* and *FmFT* in response to light signalling remains to be elucidated by future research.

Notably, we detected *FmHd3a* expressions in all four tissues collected from the flowering plants, but not in NF-leaves or NF-branchlets. We also noticed that all plants were sampled from the same provenance and grown in chambers under the same conditions after transplantation, but all non-flowering plants each had a longer section of rhizome than the flowering ones. Given that rhizome can provide nutrients for plants to maintain normal growth, thus, we hypothesize that the broken balance between vegetative and reproductive growth after transplantation result in the floral transition. Additionally, the DEUs upregulated in F-leaves relative to NF-leaves were mainly enriched in KEGG pathways related to starch, sucrose, and galactose metabolism, and corresponding down-regulated DEUs were mapped to the light and carbon fixation, plant circadian rhythm, and photosynthesis pathways. Therefore, we speculate that bamboo *FT* orthologs might be regulated by upstream regulators involved in carbohydrate metabolism.

## **Carbohydrate metabolism and signal transport may be major factors in floral transition, organogenesis, and death after flowering**

Bamboo exhibits excellent flexibility and fracture toughness, and so far, the presence of fibers within the bamboo culm was thought responsible for these remarkable mechanical properties [38]. And the development of plant fibers is accompanied by the of carbohydrate metabolism [39]. Perhaps this is the reason why many DEUs were involved in carbohydrate metabolism pathway. Interestingly, our results indicated that starch and sucrose metabolism was a major enriched KEGG pathway for the DEUs from several combination pairs, including NF- vs F-leaves, branchlets/leaves vs I- & P- spikelets. Transcripts and metabolic signatures of maize leaves have shown that the balance between transitory starch and sucrose is associated with the autonomous floral transition [40]. And in *Lilium*, carbohydrates have been found to be transported from the vascular bundles to floral organs during reproduction [41]. Yang et al., (2017) also reported that the deficiency in the resources in male flowers reduced pollen viability in *Tapiscia sinensis* due to biased carbohydrate transport toward the female flowers [42]. Therefore, we suspect there may be a correlation between DEUs related to starch and sucrose metabolism and arrow bamboo floral organ development.

In rice, excessive uridine 5'-diphosphoglucose-glucose (UDPG) can result in programmed cell death, accumulation of reactive oxygen species and an increase in the caspase-like activity [43] and inactivate starch synthase disrupted normal male reproduction by delaying programmed cell death in cotton [44], suggestive of a correlation between starch and sucrose metabolism and death. Bamboo flowering, especially in masting species, often causes plants to wilt and die after setting seed. It is possible that increased starch and sucrose metabolism might trigger the excessive accumulation of reactive oxygen species and result in the altered activity of key enzymes in important biological pathways.

In the present study, unigenes related to the signal transduction pathways were significantly upregulated in the tissues of flowering arrow bamboo plants. The transcriptomic profiles of *Posidonia oceanica* also showed a strong metabolic activation of hormones in the heat stress-induced flowering plants [45]. Signal transduction-related genes were also found to have undergone both significant positive selection and expansion events on the adaptive evolution of *Crucihimalaya himalaica* [34] and *cyanobacterium Trichormus* sp. NMC-1 [46] on the QTP. In the present study, unigenes related to the signal transduction pathways were significantly upregulated in the tissues of flowering arrow bamboo plants. We suspect that this may be due to the long distance transport of the FT proteins, which ensures floral promotion at the shoot apex [47], or the phytohormone signaling and calcium signaling, which play diverse roles in the specification of flower organs during arrow bamboo reproductive development [48-49].

## ***F. macclureana* has presumably evolved an integrated mechanism to adapt to the harsh environment of the QTP**

We detected the broad expressions of unigenes encoding putative Hsp proteins, such as *heat shock protein 70* (*Hsp70*) and *heat shock protein 90* (*Hsp90*). Both Hsp70 and Hsp90 are important for maintaining cellular protein homeostasis under stress conditions and they function by activating other targets [50-52]. It is likely that the sudden exposure to the higher temperature of the lab (i.e. outside the QTP) triggered the expressions of these unigenes. Warmer temperatures can greatly reduce flowering synchrony among individuals from 72 woody and herbaceous species [53]. It was also reported that *P. oceanica*, a highly clonal and long-lived

species, massively bloomed after a simulated heatwave [45]. Given the cold temperatures present at the high-altitude regions of the QTP [26], it is reasonable to presume that *F. macclureana* has developed into a heat-sensitive but not heat-tolerant bamboo species and flowering is probably a stress-induced response to the higher temperature in the lab.

DNA repair and disease-resistance pathways have been found to play a crucial role in the highland adaptation of Tibetan highland plants [46, 54]. In the present study, we detected significant expressions of many unigenes related to pathogen response that contained either a nucleotide binding (NB)-ARC domain or a leucine-rich repeat (LRR) domain, which were present in most resistance (R) proteins [55-57]. We also identified many differentially expressed unigenes that were significantly enriched in the DNA repair pathways. Since relatively fewer species of pathogenic microorganisms and intense UV radiation exist on the QTP [58], it is reasonable to presume that *F. macclureana* has evolved a relatively narrow range of pathogen specificity and specific DNA-repair mechanisms. Sudden exposure to the lab environment, which contains a heavier load of pathogens and weak light intensity, may have induced an innate defensive response of *F. macclureana*, and those that were enriched in the plant-pathogen interaction and DNA repair response pathways may be important for *F. macclureana* to cope with the new environment present in the lab. Although further studies are needed to investigate the molecular mechanisms responsible for the putative adaptive evolutionary changes, this study provides insights into how plants adapt to harsh and extreme environments.

## Conclusions

In the present study, we constructed a novel *de novo* transcriptome analysis for *F. macclureana*. Based on two major KEGG pathways of carbohydrate metabolism and signal transduction that DEUs were enriched in, as well as the different regulation of *FmFT* and *FmHd3a* in NF- vs F- leaves, we speculate that both environmental signals and physiological status have effect on floral transition in *F. macclureana*. Significant expressions of unigenes enriched in DNA repair and plant-pathogen interaction pathways may reflect the adaptation of *F. macclureana* to its high radiation and pathogen-specific environment on the QTP. We identified both similarities and differences in adaptive mechanisms (e.g., disease-resistance and DNA repair pathways) and stress-induced flowering mechanisms (e.g., carbohydrate metabolism and signal transduction pathways) among plants that grow at high altitudes. Although further experimental verification is needed, our results provide insight into the regulation of flowering time in highland bamboo as well as how this species adapts to harsh and extreme environments.

## Methods

### Tissues collection

The studied plant species is highland arrow bamboo (*Fargesia macclureana*), and it grows mainly as an underbrush of coniferous forest or coniferous and broad-leaved mixed forest, and sometimes forms a pure population in the QTP at an altitude of approximately 2,000 ~ 3800 m above sea level (Fig. 1). *F. macclureana* was formally identified by Stapleton in 1993 [25] and detailed explanations are provided in the volume 22 of Flora of China (<http://foc.iplant.cn/>) [59]. A voucher specimen of this material has been deposited in the Bamboo Research Institute of Nanjing Forestry University. A total of six seedlings of *F. macclureana* were

sampled from the wild with the permission of the local forestry department and collected from the Bayi District, Linzhi City, Tibet, China (29°46' 0.95" N, 94°44'46.36" E, altitude: ~2,200 m). They were dug out of the ground, each with a section of rhizome, which can provide nutrients for plants. Then they were all transferred to individual pots at the State Forestry Administration Key Open Laboratory at the International Center for Bamboo and Rattan in Beijing (N: 39°59' 17.52", E: 116°28'46.06", altitude ~34 m). During growth, the plants were maintained at  $28 \pm 1$  °C and 50-55% relative humidity under a 16/8 h (light/dark) photoperiod regimen with a light intensity of  $200 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ . All the seedlings were watered with a 1/3 B5 macronutrient nutrient solution three times a week. After twenty days, four seedlings which had shorter rhizomes flowered, while the other two didn't blossom until the time of sampling. One month later, we collected samples of six tissues for further *de novo* sequencing, including inflorescences in the initial flower stage (I-spikelets), inflorescences at the peak flower stage (P-spikelets), branchlets of the flowering plants (F-branchlets), leaves of the flowering plants (F-leaves), branchlets of the non-flowering plants (NF-branchlets) and leaves of the non-flowering plants (NF-leaves). We collected three independent replicates of each tissue type.

### **RNA extraction, quantification, and qualification**

Total RNA was extracted from each of the six unique tissues mentioned above using a RNeasy plant RNA extraction kit (Qiagen, Dusseldorf, Germany), and the extraction procedure was performed according to the manufacturer's instructions. RNA degradation and contamination were monitored using 1% agarose gels. RNA purity was checked using a NanoPhotometer<sup>®</sup> spectrophotometer (Implen GmbH, Munich, Germany). RNA concentration was measured using a Qubit<sup>®</sup> RNA Assay Kit and a Qubit<sup>®</sup>2.0 Fluorometer (Life Technologies, CA, USA). RNA integrity was assessed using an RNA Nano 6000 Assay Kit run on an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

### **Library preparation for transcriptome sequencing**

Library construction and RNA-Seq were performed by the Biomarker Biotechnology Corporation (Beijing, China). A total of 3  $\mu\text{g}$  RNA per sample was used for RNA preparation. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads, followed by fragmentation carried out using divalent cations at elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5 $\times$ ). First strand cDNA was synthesized using random hexamer primers and M-MuLV Reverse Transcriptase (RNase H<sup>-</sup>). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. The remaining overhangs were converted into blunt ends via the exonuclease and polymerase activities. Next, the 3' ends of the DNA fragments were adenylated and ligated to the NEBNext adaptors with hairpin loop structures to prepare samples for hybridization, this was to select cDNA fragments that are 150-200 bp in length. Library fragments were then purified using an Agencourt AMPure XP system (Beckman Coulter, Brea, CA, USA), and 3  $\mu\text{l}$  USER enzyme (New England Biolabs, Ipswich, MA, USA) was added to the size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by five min at 95°C before PCR. PCR was performed using Phusion High-Fidelity DNA polymerase (Thermo Fisher, Waltham, MA, USA), universal PCR primers, and the Index (X) Primer. Finally, the PCR products were purified using the AMPure XP system and library quality was assessed on the Agilent Bioanalyzer 2100.

### **Clustering and sequencing**

The clustering of index-coded samples was performed using a cBot Cluster Generation System and a TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA), and all experimental procedures were performed according to the manufacturer's instructions. After that, library preparations were sequenced on an Illumina HiSeq 2000 platform and paired-end reads were generated.

### ***De novo* transcriptome assembly**

Raw data in fastq format were first processed using in-house perl scripts. Clean data were obtained by removing low-quality reads and reads that contain the adapters or poly-N sequences. Meanwhile, we checked the quality of our unassembled read dataset by examining various measures including Q20, Q30, GC-content, and sequence duplication. All the downstream analyses were performed using high-quality clean data.

The transcriptome was assembled using clean reads from all libraries and samples. The assembly was produced using Trinity [60] with `min_kmer_cov` set to 2 and all other parameters set to their respective default values.

### **Functional annotation of the transcriptome**

Gene function was annotated using the following databases: Nr (NCBI non-redundant protein sequences), Pfam (Protein family), KOG/COG/eggNOG (Clusters of Orthologous Groups of proteins), Swiss-Prot (a database of manually annotated and reviewed protein sequences), KEGG (the Kyoto Encyclopedia of Genes and Genomes), and the GO (Gene Ontology) database.

### **Quantification of gene expression levels**

Gene expression levels were estimated using RSEM [61] for each sample: clean data were mapped back onto the assembled transcriptome, followed by a read count for each gene. The expression levels of unigenes were calculated and normalized using FPKM (fragments per kb per million fragments) [62].

### **Expression analysis of broadly and specifically expressed unigenes**

For all unigenes, those that were expressed in all six tissues were defined as broadly expressed unigenes (BEUs). Similarly, unigenes that were specifically expressed in only one tissue were defined as specifically expressed unigenes (SEUs). The identification of BEUs and SEUs was conducted by using tools on the BMKCloud platform (<http://www.biocloud.net>).

### **Weighted gene co-expression network analysis (WGCNA)**

WGCNA was performed on all unigenes identified using the *WGCNA* R package. We calculated the adjacency matrices and performed the topological overlap measures (TOMs), which show the degree of overlap in shared neighbors between pairs of genes in the network to define gene clusters in our transcriptome dataset.  $1 - \text{TOM}$  was used as a dissimilarity measure for hierarchical clustering and module detection. Modules of the clustered genes were then selected using the Dynamic Tree Cut algorithm as implemented by WGCNA. To identify modules that are significantly related to particular tissues, expression profiles of each module were summarized by a module eigengene (ME) and the correlations between the modules and corresponding tissues were calculated.

## Expression analysis of differentially expressed unigenes (DEUs)

Before analysis, we conducted a principal component analysis (PCA) and removed one replicate that showed an inconsistent expression pattern in the NF-branchlets and NF-leaves to ensure consistency in the expression patterns of unigenes between replicates (Additional file 12: Figure S4).

Expression analysis of the DEUs between pairs of tissues/groups was performed using the *DESeq* package (1.10.1) in R. *DESeq* provides statistical routines for identifying differential expression in the digital gene expression data using a model based on the negative binomial distribution. The resulting *P* values were adjusted using the Benjamini-Hochberg method for controlling the false discovery rate [63]. Here, uni-transcripts with an absolute value of  $\log_2$  ratio  $\geq 2$ , an FDR significance score  $< 0.01$ , and an adjusted *P*-value  $< 0.05$  were deemed to be differentially expressed.

## Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis

To understand the higher-level functions of the observed unigenes, we performed GO term annotation and KEGG pathway enrichment analysis using BMKCloud (<http://www.biocloud.net/>; [64]). We used KOBAS 2.0 [65] to test the statistical enrichment of differentially expressed genes in KEGG pathways. Pathways with *P* values  $< 0.05$  were considered significantly enriched.

## Protein-protein interactions (PPIs)

The DEU and SEU sequences were queried using BLASTX against the related species to predict PPIs that the DEUs and SEUs may be involved in. This search procedure was capable of identifying PPIs that may be similar to any others found in the STRING database (<http://string-db.org/>). These PPIs were then visualized using Cytoscape [66].

## Detection of SSRs

Picard-tools version 1.41 and samtools version 0.1.18 were used to sort data, remove duplicated reads, and merge the bam alignment results of each sample. SSRs were identified using MISA (<https://webblast.ipk-gatersleben.de/misa/>).

## Validation of *FmHd3a* transcript levels by RT-qPCR

To verify the expression pattern of the *FmHd3a*, we used RT-qPCR to assess the expressions of *FmHd3a* in six distinct tissues. First-strand cDNA was synthesized from total RNA extracted by using a reverse transcription system (Promega, Madison, WI, USA) following the manufacturer's instructions. Each RT-qPCR amplification was performed at least three times, and *NTB* and *TIP41* were used as internal controls [67]. Primers for these genes are listed in Additional file 13: Table S9. The relative expression levels of *FmHd3a* in different tissues were calculated using the  $2^{-\Delta\Delta CT}$  method [68]. The statistical significance of differences in the mean levels of expression was tested using a one-way ANOVA. Significant differences in transcript abundance between different tissues were then compared using Duncan's multiple range tests as implemented by SPSS version 17.0 (IBM SPSS, Chicago, USA). We considered mean differences at  $P < 0.05$  and  $P < 0.01$  to be statistically significant and highly statistically significant, respectively.

# Abbreviations

**BEUs:** Broadly expressed unigenes; **BP:** Biological process; **CC:** Cellular component; **COG:** Clusters of orthologous groups; **DEUs:** Differentially expressed unigenes; **F-branchlets:** Branchlets of the flowering plants; **F-leaves:** Leaves of the flowering plants; **FPKM:** Fragments per kb per million fragments; **GO:** Gene ontology; **I-spikelets:** Inflorescences in the initial flower stage; **KEGG:** Kyoto Encyclopedia of Genes and Genomes; **ME:** Module eigengene; **MF:** Molecular function; **NF-branchlets:** Branchlets of the non-flowering plants; **NF-leaves:** Leaves of the non-flowering plants; **Nr:** NCBI non-redundant protein sequences; **PCA:** Principal component analysis; **Pfam:** Protein family; **PPIs:** Protein-protein interactions. **P-spikelets:** Inflorescences at the peak flower stage; **QTP:** Qinghai–Tibet Plateau; **SEUs:** Specifically expressed unigenes; **SSRs:** Simple sequence repeats; **TOMs:** topological overlap measures. **WGCNA:** Weighted gene co-expression network analysis.

# Declarations

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Availability of data and materials

The RNA sequencing dataset generated during the current study have been submitted to NCBI Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>) with the accession number PRJNA544133.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was supported by the Special Funds for Fundamental Scientific Research on Professional Work Supported by International Center for Bamboo and Rattan (No. 1632019008) and the Sub-Project of National Science and Technology Support Plan of the Twelfth Five-Year in China (No. 2015BAD04B01). The funder is the corresponding author of this manuscript and he played an important role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Authors' contributions

Conceived and designed the experiments: ZM G. Performed the experiments: KB Y and JJ S. Analyzed the data: CX Z, YL D and Y L. Interpreted the results and wrote the paper: Y L. All authors have read and approved the final manuscript.

## Acknowledgements

Not applicable.

## References

1. Putterill J, Varkonyi-Gasic E. FT and florigen long-distance flowering control in plants. *Curr Opin Plant Biol.* 2016;33:77-82.
2. Wigge PA, Kim MC, Jaeger KE, et al. Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science.* 2005;309:1056-1059.
3. Liu H, Wang Q, Liu Y, et al. *Arabidopsis* CRY2 and ZTL mediate blue-light regulation of the transcription factor CIB1 by distinct mechanisms. *Proc Natl Acad Sci U S A.* 2013;110:17582-17587.
4. Du A, Tian W, Wei M, et al. The DTH8-Hd1 module mediates day-length-dependent regulation of rice flowering. *Mol Plant.* 2017;10:948-961.
5. Lin MK, Belanger H, Lee YJ, et al. FLOWERING LOCUS T protein may act as the long-distance florigenic signal in the cucurbits. *Plant Cell.* 2007;19:1488-1506.
6. Navarro C, Cruz-Oro E, Prat S. Conserved function of *FLOWERING LOCUS T (FT)* homologues as signals for storage organ differentiation. *Curr Opin Plant Biol.* 2015;23:45-53.
7. Wolabu TW, Zhang F, Niu L, et al. Three *FLOWERING LOCUS T*-like genes function as potential florigens and mediate photoperiod response in sorghum. *New Phytol.* 2016;210:946-959.
8. Bohlenius H, Huang T, Charbonnel-Campaa L, et al. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science.* 2006;312:1040-1043.
9. Hsu CY, Liu Y, Luthe DS, et al. Poplar *FT2* shortens the juvenile phase and promotes seasonal flowering. *Plant Cell.* 2006;18:1846-1861.
10. Klocko AL, Ma C, Robertson S, et al. *FT* overexpression induces precocious flowering and normal reproductive development in *Eucalyptus*. *Plant Biotechnol J.* 2016;14:808-819.
11. Munoz-Fambuena N, Nicolas-Almansa M, Martinez-Fuentes A, et al. Genetic inhibition of flowering differs between juvenile and adult *Citrus* trees. *Ann Bot.* 2019;123:483-490.
12. Zhou X, Ruan J, Wang G, et al. Characterization and identification of microRNA core promoters in four model species. *PLoS Comput Biol.* 2007;3:e37.
13. Bhattacharya S, Das M, Bar R, et al. Morphological and molecular characterization of *Bambusa tulda* with a note on flowering. *Ann Bot.* 2006;98:529-535.
14. Yuan JL, Yue JJ, Gu XP, et al. Flowering of woody bamboo in tissue culture systems. *Front Plant Sci.* 2017;8:1589.
15. Tian Z, Liu X, Fan Z, et al. The next widespread bamboo flowering poses a massive risk to the giant panda. *Biol Conserv.* 2019;234:180-187.
16. Giordano CV, Sánchez RA, Austin AT. Gregarious bamboo flowering opens a window of opportunity for regeneration in a temperate forest of Patagonia. *New Phytol.* 2009;181:880-889.
17. Shih MC, Chou ML, Yue JJ, et al. *BeMADS1* is a key to delivery *MADS*s into nucleus in reproductive tissues-*De novo* characterization of *Bambusa edulis* transcriptome and study of *MADS* genes in bamboo floral development. *BMC Plant Biol.* 2014;14:179.

18. Wang X, Zhang X, Zhao L, et al. Morphology and quantitative monitoring of gene expression patterns during floral induction and early flower development in *Dendrocalamus latiflorus*. *Int J Mol Sci*. 2014;15:12074-12093.
19. Zhang Y, Tang D, Lin X, et al. Genome-wide identification of *MADS-box* family genes in moso bamboo (*Phyllostachys edulis*) and a functional analysis of *PeMADS5* in flowering. *BMC Plant Biol*. 2018;18:176.
20. Chiou T. The role of microRNAs in sensing nutrient stress. *Plant Cell Environ*. 2007;30:323-332.
21. Hisamoto Y, Kashiwagi H, Kobayashi M. Use of flowering gene *FLOWERING LOCUS T (FT)* homologs in the phylogenetic analysis of bambusoid and early diverging grasses. *J Plant Res*. 2008;121:451-461.
22. Gao J, Ge W, Zhang Y, et al. Identification and characterization of microRNAs at different flowering developmental stages in moso bamboo (*Phyllostachys edulis*) by high-throughput sequencing. *Mol Genet Genom*. 2015;290:2335-2353.
23. Wysocki WP, Ruiz-Sanchez E, Yin Y, et al. The floral transcriptomes of four bamboo species (Bambusoideae; Poaceae): support for common ancestry among woody bamboos. *BMC Genomics*, 2016;17:384.
24. Ge W, Zhang Y, Cheng Z, et al. Main regulatory pathways, key genes and microRNAs involved in flower formation and development of moso bamboo (*Phyllostachys edulis*). *Plant Biotechnol J*. 2017;15:82-96.
25. Stapleton CMA. *Fargesia macclureana*, a Tibetan bamboo in Europe. *Bamboo Soc. (GB) Newsl*. 1993;17:17.
26. Cao M, Jin Y, Liu N, et al. Effects of the Qinghai-Tibetan Plateau uplift and environmental changes on phylogeographic structure of the Daurian Partridge (*Perdix dauuricae*) in China. *Mol Phylogenet Evol*. 2012;65:823-830.
27. Karlgren A, Gyllenstrand N, Källman T, et al. Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol*; 2011;156(4):1967-1977.
28. Banfield MJ and Brady RL. The structure of *Antirrhinum* centroradialis protein (CEN) suggests a role as a kinase regulator. *J Mol Biol*. 2000;297(5):1159-1170.
29. Wickland DP and Hanzawa Y. The *FLOWERING LOCUS T/TERMINAL FLOWER 1* gene family: functional evolution and molecular mechanisms. *Mol Plant*. 2015;8(7):983-997.
30. Abe M, Kobayashi Y, Yamamoto S, et al. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science*. 2005;309:1052-1056.
31. Song YH, Shim JS, Kinmonth-Schultz HA, et al. Photoperiodic flowering: time measurement mechanisms in leaves. *Annu Rev Plant Biol*. 2015;66:441-464.
32. Yano M, Katayose Y, Ashikari M, et al. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. *Plant Cell*. 2000;12:2473-2484.
33. Hayama R, Yokoi S, Tamaki S, et al. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature*. 2003;422:719-722.
34. Zhang T, Qiao Q, Novikova PY, et al. Genome of *Crucihimalaya himalaica*, a close relative of *Arabidopsis*, shows ecological adaptation to high altitude. *Proc Natl Acad Sci U S A*. 2019;116: 7137-7146.
35. Pin PA, Benlloch R, Bonnet D, et al. An antagonistic pair of *FT* homologs mediates the control of flowering time in sugar beet. *Science*, 2010;330(6009):1397-1400.

36. Hanzawa Y, Money T, Bradley D. A single amino acid converts a repressor to an activator of flowering. *Proc Natl Acad Sci U S A*, 2005;102(21): 7748-7753.
37. Ahn JH, Miller D, Winter VJ, et al. A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *EMBO J*, 2006;25(3):605-614.
38. Habibi MK, Lu Y. Crack propagation in bamboo's hierarchical cellular structure. *Sci Rep*. 2014;4:5598.
39. Sun W, Gao Z, Wang J, et al. Cotton fiber elongation requires the transcription factor GhMYB212 to regulate sucrose transportation into expanding fibers. *New Phytol*. 2019;222:864-881.
40. Coneva V, Guevara D, Rothstein SJ, et al. 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*. 2012;63:5079-5092.
41. Castro AJ, Clement C. Sucrose and starch catabolism in the anther of *Lilium* during its development: a comparative study among the anther wall, locular fluid and microspore/pollen fractions. *Planta*, 2007;225:1573-1582.
42. Yang K, Zhou X, Wang Y, et al. Carbohydrate metabolism and gene regulation during anther development in an androdioecious tree, *Tapiscia sinensis*. *Ann Bot*. 2017;120:967-977.
43. Xiao G, Zhou J, Lu X, et al. Excessive UDPG resulting from the mutation of UAP1 causes programmed cell death by triggering reactive oxygen species accumulation and caspase-like activity in rice. *New Phytol*. 2018;217:332-343.
44. Min L, Zhu L, Tu L, et al. Cotton *GhCKI* disrupts normal male reproduction by delaying tapetum programmed cell death via inactivating starch synthase. *Plant J*. 2013;75:823-835.
45. Marín-Guirao L, Entrambasaguas L, Ruiz JM, et al. Heat-stress induced flowering can be a potential adaptive response to ocean warming for the iconic seagrass *Posidonia oceanica*. *Mol Ecol*, 2019;28: 2486-2501.
46. Qiao Q, Huang Y, Qi J, et al. The genome and transcriptome of *Trichormus* sp. NMC-1: insights into adaptation to extreme environments on the Qinghai-Tibet Plateau. *Sci Rep*. 2016;6: 29404.
47. Turck F, Fornara F, Coupland G. Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Ann Rev Plant Biol*. 2008;59:573-594.
48. Galli M, Liu Q, Moss BL, et al. Auxin signaling modules regulate maize inflorescence architecture. *Proc Nat Acad Sci USA*. 2015;112:13372-13377.
49. Yuan Z, Zhang D. Roles of jasmonate signalling in plant inflorescence and flower development. *Curr Opin Plant Biol*. 2015;27:44-51.
50. Mayer MP. Hsp70 chaperone dynamics and molecular mechanism. *Trends Biochem Sci*. 2013;38:507-514.
51. Schopf FH, Biebl MM, Buchner J. The HSP90 chaperone machinery. *Nat Rev Mol Cell Biol*. 2017;18:345-360.
52. Biebl MM, Buchner J. Structure, function, and regulation of the Hsp90 machinery. *CSH Perspect Biol*. 2019;DOI: 10.1101/cshperspect.a034017.
53. Zohner CM, Mo L, Renner SS. Global warming reduces leaf-out and flowering synchrony among individuals. *Elife*. 2018; DOI: 10.7554/eLife.40214.

54. Fang O, Zhang QB. Tree resilience to drought increases in the Tibetan Plateau. *Global Change Biol.* 2019;25:245-253.
55. Li J, Huang H, Zhu M, et al. A plant immune receptor adopts a two-step recognition mechanism to enhance viral effector perception. *Mol Plant*, 2019;12:248-262.
56. Shao ZQ, Xue JY, Wang Q, et al. Revisiting the Origin of Plant NBS-LRR Genes. *Trends Plant Sci.* 2019;24:9-12.
57. Takken FL, Albrecht M, Tameling WI. Resistance proteins: molecular switches of plant defence. *Curr Opin Plant Biol.* 2006;9:383-390.
58. Zhang XJ, Yao TD, Ma XJ, et al. Microorganisms in a high altitude glacier ice in Tibet. *Folia Microbiol.* 2002;47:241-245.
59. Li D, Stapleton C. Bambuseae. In: *Flora of China*. Edited by Wu ZY, Raven PH, Hong DY, Vol. 22. Beijing: Science Press and St. Louis USA: Missouri Botanic Garden Press. 2006:1-180.
60. Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29:644-652.
61. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 2011;12:323.
62. Mortazavi A, Williams BA, McCue K, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5:621-628.
63. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. *J R Stat Soc B.* 1995;57:289-300.
64. Kanehisa M. KEGG Bioinformatics resource for plant genomics and metabolomics. *Methods Mol Biol.* 2016;1374:55-70.
65. Xie C, Mao X, Huang J, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucl Acids Res.* 2011;39:W316-322.
66. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13:2498-2504.
67. Fan C, Ma J, Guo Q, et al. Selection of reference genes for quantitative real-time PCR in bamboo (*Phyllostachys edulis*). *PLoS One.* 2013;8:e56573.
68. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta Ct}$  Method. *Methods.* 2001;25:402-408.

## Tables

Table 1 Length range of transcripts and unigenes identified in the transcriptome of *F. macclureana*.

Length Range	Transcripts	Unigenes
200-300	36,390 (12.59%)	25,291 (25.53%)
300-500	47,515 (16.43%)	21,257 (21.46%)
500-1,000	78,453 (27.13%)	23,806 (24.03%)
1,000-2,000	77,456 (26.79%)	16,752 (16.91%)
2,000+	49,308 (17.05%)	11,950 (12.06%)
Total number	289,122	99,056
Total length	341,956,623	91,685,618
N50 length	1,765	1,587
Mean length	1,182.74	925.59
Total number	289,122	99,056
Total length	341,956,623	91,685,618
N50 length	1,765	1,587
Mean length	1,182.74	925.59

Table 2 Statistics of annotation analysis of unigenes

Anno_Database	Annotated_Number	percentage	300<=length<1,000	length>=1,000
COG_Annotation	13,128	27.75	3,261	7,515
GO_Annotation	34,055	71.99	10,659	17,855
KEGG_Annotation	14,307	30.24	4,550	7,397
KOG_Annotation	23,492	49.66	6,863	12,779
Pfam_Annotation	28,317	59.86	7,823	16,896
Swissprot_Annotation	24,847	52.52	7,450	14,500
eggNOG_Annotation	43,909	92.82	14,040	21,568
Nr_Annotation	45,516	96.22	15,031	22,271
All_Annotated	47,306	100.00	15,602	22,437

Table 3. Differentially expressed unigenes (DEUs; Fold change > 2; FDR < 0.01) among tissues of *F. macclureana*. DEUs\_total: the total number of DEUs; DEUs\_up (%): the number (and percentage) of up-regulated DEUs; DEUs\_down (%): the number (and percentage) of down-regulated DEUs).

Number	Group	DEUs_total	DEUs_up (%)	DEUs_down (%)
1	I-spikelets vs P-spikelets	970	916 (94.43)	54 (5.57)
2	F-branchlets vs I-spikelets	4,970	3,046 (61.29)	1,924 (38.71)
3	F-branchlets vs P-spikelets	5,124	3,338 (65.14)	1,786 (34.86)
4	F-branchlets vs F-leaves	8,467	3,967 (46.85)	4,500 (53.15)
5	F-leaves vs I-spikelets	12,829	6,791 (52.93)	6,038 (47.07)
6	F-leaves vs P-spikelets	10,791	6,625 (61.39)	4,166 (38.61)
7	NF-branchlets vs I-spikelets	11,628	6,135 (52.76)	5,493 (47.24)
8	NF-branchlets vs P-spikelets	10,809	5,893 (54.52)	4,916 (45.48)
9	NF-branchlets vs F-branchlets	6,524	3,275 (50.20)	3,249 (49.80)
10	NF-branchlets vs F-leaves	11,670	5,902 (50.57)	5,768 (49.43)
11	NF-branchlets vs NF-leaves	3,853	1,946 (50.51)	1,907 (49.49)
12	NF-leaves vs I-spikelets	13,577	6,921 (50.98)	6,656 (49.02)
13	NF-leaves vs P-spikelets	11,718	6,130 (52.31)	5,588 (47.69)
14	NF-leaves vs F-branchlets	11,659	5,606 (48.08)	6,053 (51.92)
15	NF-leaves vs F-leaves	5,032	2,492 (49.52)	2,540 (50.48)

## Supplementary File Legends

**Additional file 1: Table S1.** Statistic of sequencing and assembly data.

**Additional file 2: Table S2.** 47,306 unigenes were annotated and their predicted functions.

**Additional file 3: Table S3.** The top10 enriched GO terms in three main categories for all unigenes identified.

**Additional file 4: Table S4.** The most enriched GO terms and KEGG pathways for the specially expressed unigenes (SEUs) and broadly expressed unigenes (BEUs) in all tissues collected from flowering plants.

**Additional file 5: Figure S1.** KEGG annotation of unigenes that were specifically expressed in P-spikelets (a), F-branchlets (b) and F-leaves (c) of arrow bamboo flowering plants. The size of dots is proportional to the number of unigenes.

**Additional file 6: Table S5.** KEGG enrichment of differentially expressed unigenes (DEUs) between different tissues.

**Additional file 7: Figure S2.** Hub unigenes in regulatory networks of flowering identified based on analysis of DEUs among tissues. Unigenes c109220.graph\_c0 and c110963.graph\_c4, showing differential expressions between NF-leaves and F-leaves, are both bamboo orthologs of *FLOWERING LOCUS T (FT)*, which was marked with a red square; while unigenes down-regulated were marked with green squares.

**Additional file 8: Table S6.** Expressions of 10 unigenes that were differentially expressed between F-leaves vs NF-leaves and mapping into the circadian phythm-plant KEGG pathway.

**Additional file 9: Figure S3.** Weighted gene co-expression network analysis (WGCNA) of all unigenes identified in the transcriptome of *F. macclureana*. **(a)** The phylogenetic tree diagram and the heat map related to the traits. This diagram is divided into three parts: the cluster tree of gene system, the module color of corresponding genes, and the correlation between genes related to each trait in tested samples and its module. The redder the color, the more positive the correlation; conversely, blue is negatively correlated. **(b)** Gene co-expression network heatmaps drawn by randomly selected 1,500 genes, in which the left and the upper sides are the symmetrical system clustering tree of gene network/module, and the lower right area indicates the dissimilarity between genes, and the smaller the value is, the darker the color is. **(c)** Module and trait correlation heat map showing the relationship between a module and a given trait. The closer the correlation between a shape and a module is to the absolute value of 1, it is likely that this trait is related to the module gene work. **(d)** Systematic clustering tree of samples based on unigenes expressions.

**Additional file 10: Table S7.** KEGG enrichment of unigenes in nine significant gene module relating to P-spikelets.

**Additional file 11: Table S8.** 9,296 SSRs identified from 7,668 unigenes.

**Additional file 12: Figure S4.** Principal component analysis (PCA) of unigenes expressions for 18 samples collected from inflorescences in the initial and peak flower stage (I- and P- spikelets), branchlets and leaves of flowering and non-flowering bamboo plants (F/NF-branchlets and F/NF-leaves).

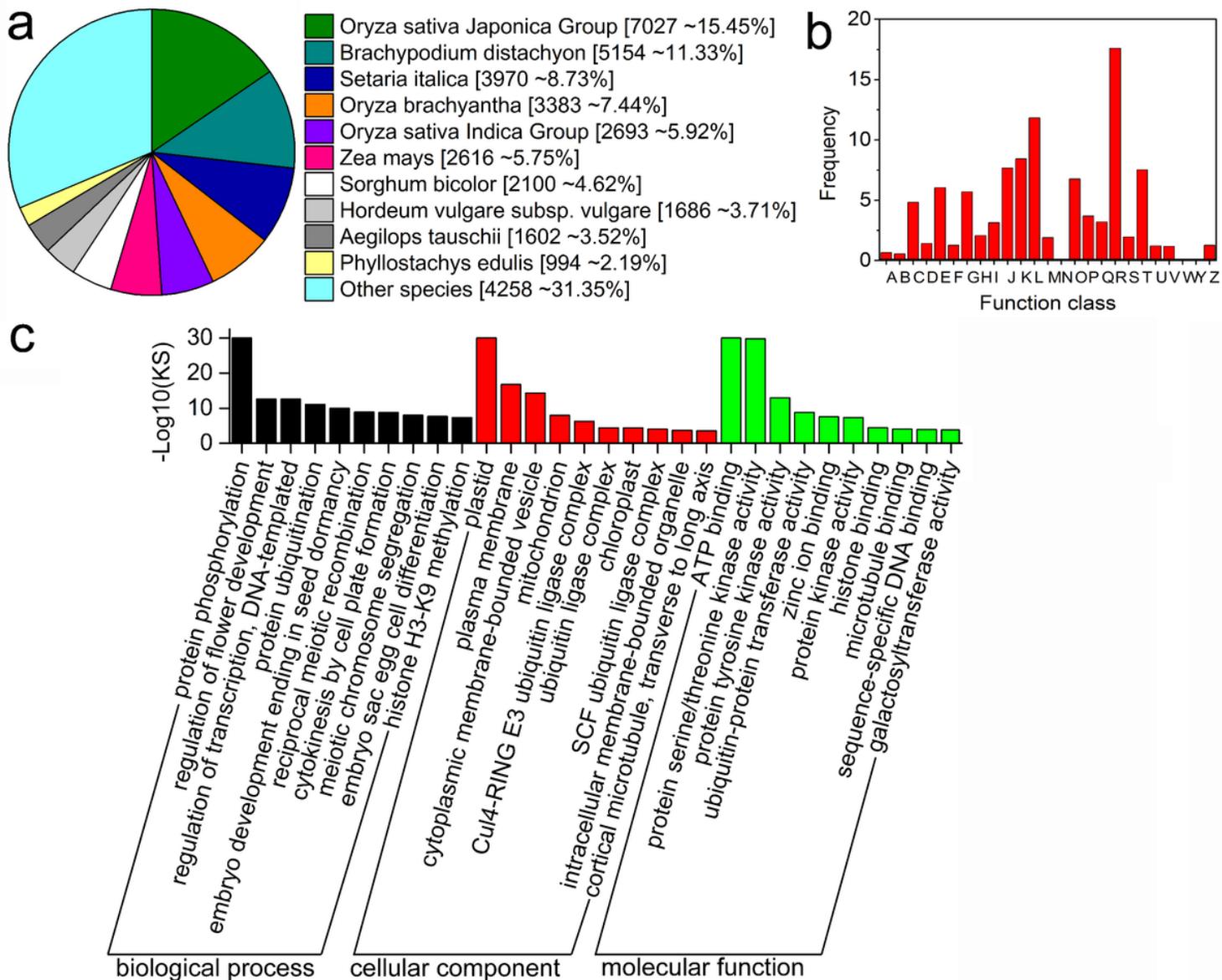
**Additional file 13: Table S9.** Primer pairs used for RT-qPCR.

## Figures



### Figure 1

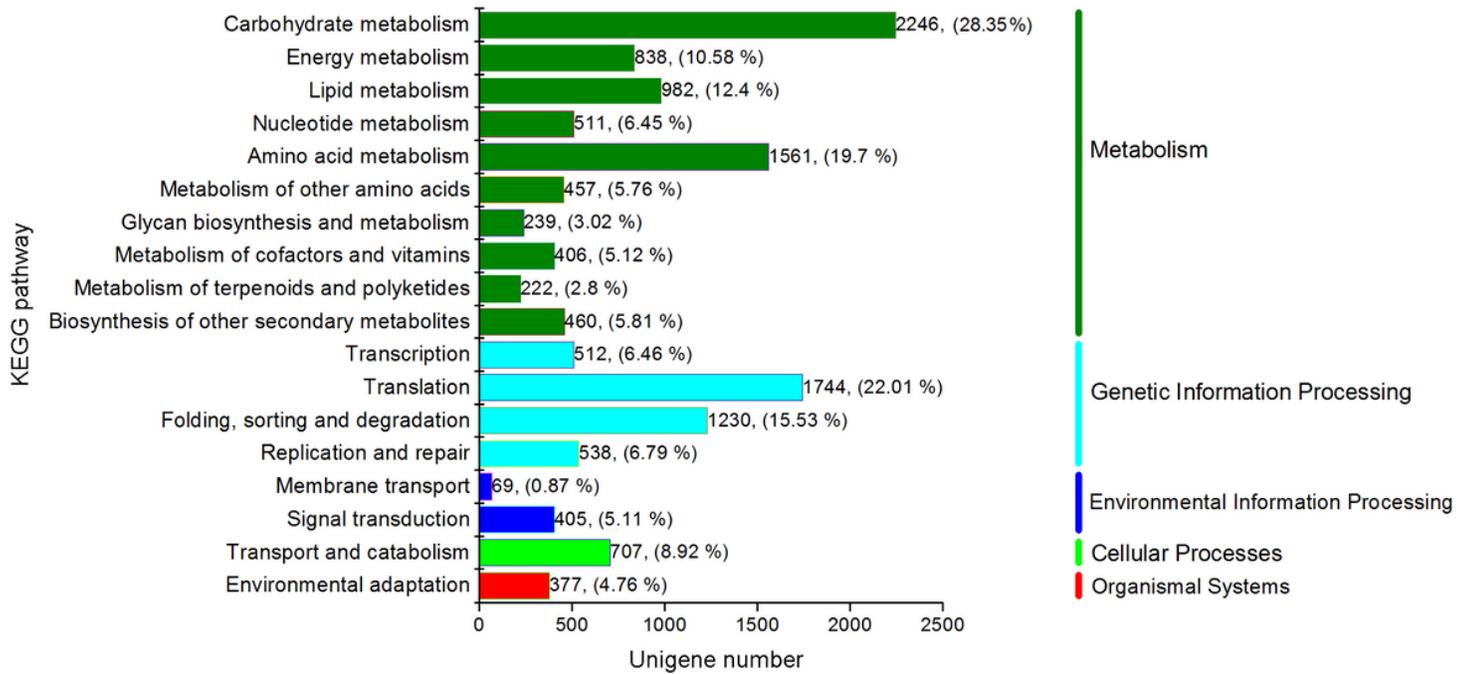
Seedlings of *Fargesia macclureana* flowered shortly after being transferred from the Qinghai–Tibet Plateau (QTP) approximately 2,000 ~ 3,800 m above sea level to a low altitude lab. (a-b) Floret and spikelet of a flowering seedling maintained at the low altitude lab outside the QTP. (c-d) The seedling and shoot of plants growing on the QTP. (e) The original growing environment of *F. macclureana*.



**Figure 2**

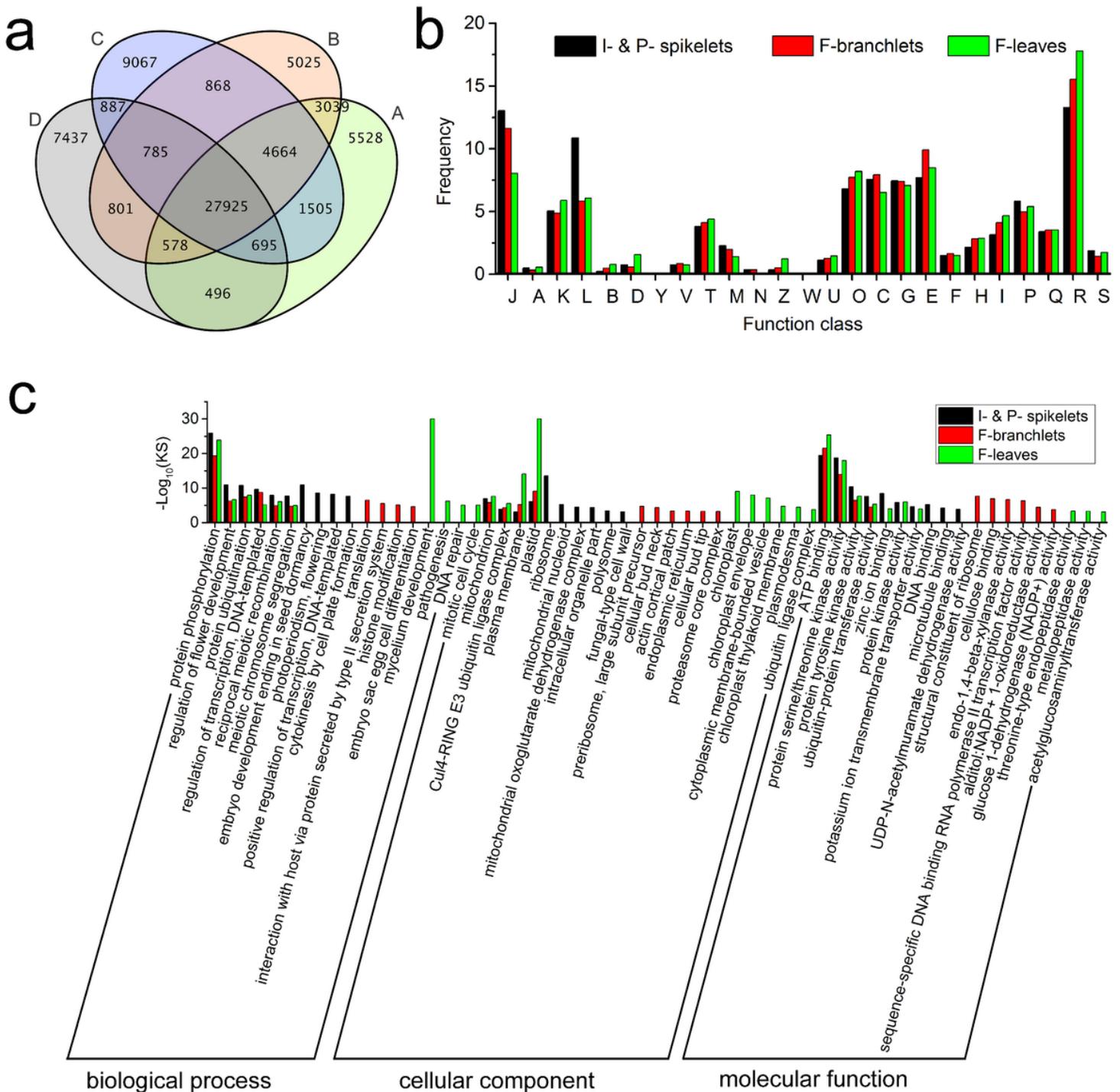
Function annotation and classification of unigenes identified from the transcriptome of *F. macclureana*. (a) Nr annotation. (b) Clusters of orthologous groups (COG) annotation. Out of 45,516 Nr hits, 13,128 unigenes had a COG classification. A: RNA processing and modification B: Chromatin structure and dynamics C: Energy production and conversion D: Cell cycle control, cell division, chromosome partitioning E: Amino acid transport and metabolism F: Nucleotide transport and metabolism G: Carbohydrate transport and metabolism H: Coenzyme transport and metabolism I: Lipid transport and metabolism J: Translation, ribosomal structure and biogenesis K: Transcription L: Replication, recombination and repair M: Cell wall/membrane/envelope biogenesis N: Cell mobility O: Posttranslational modification, protein turnover, chaperones P: Inorganic ion transport and metabolism Q: Secondary metabolites biosynthesis, transport and metabolism R: General function prediction only S: Function unknown T: Signal transduction mechanism U: Intracellular trafficking, secretion, and vesicular transport V: Defense mechanisms W: Extracellular structures Y: Nuclear structure Z: Cytoskeleton. (c) GO annotation. Results were summarized in three main categories: biological process, cellular

component and molecular function. The right and left y-axes indicated the number and percentage of unigenes in a certain category, respectively.



**Figure 3**

KEGG annotation of unigenes in the transcriptome of *F. macclureana*. The x-axis indicated the number of unigenes in a certain category. The right y-axis showed the main clusters of KEGG pathways.



**Figure 4**

Unigenes that were specifically expressed in different tissues collected from flowering plants of *F. macclureana*. (a) Venn diagram of unigenes expressed in spikelets in the initial flower stage (I-spikelets, A) and the peak flower stage (P-spikelets, B), branchlets (F-branchlets, C) and leaves (F-leaves, D) of flowering plants. (b) COG annotation of unigenes that were specifically expressed in I-spikelets, P-spikelets, F-branchlets and F-leaves. (c) GO enrichment of unigenes that were specifically expressed in I- & P-spikelets, F-branchlets and F-leaves. BP: biological process; CC: cellular component; MF: molecular function.

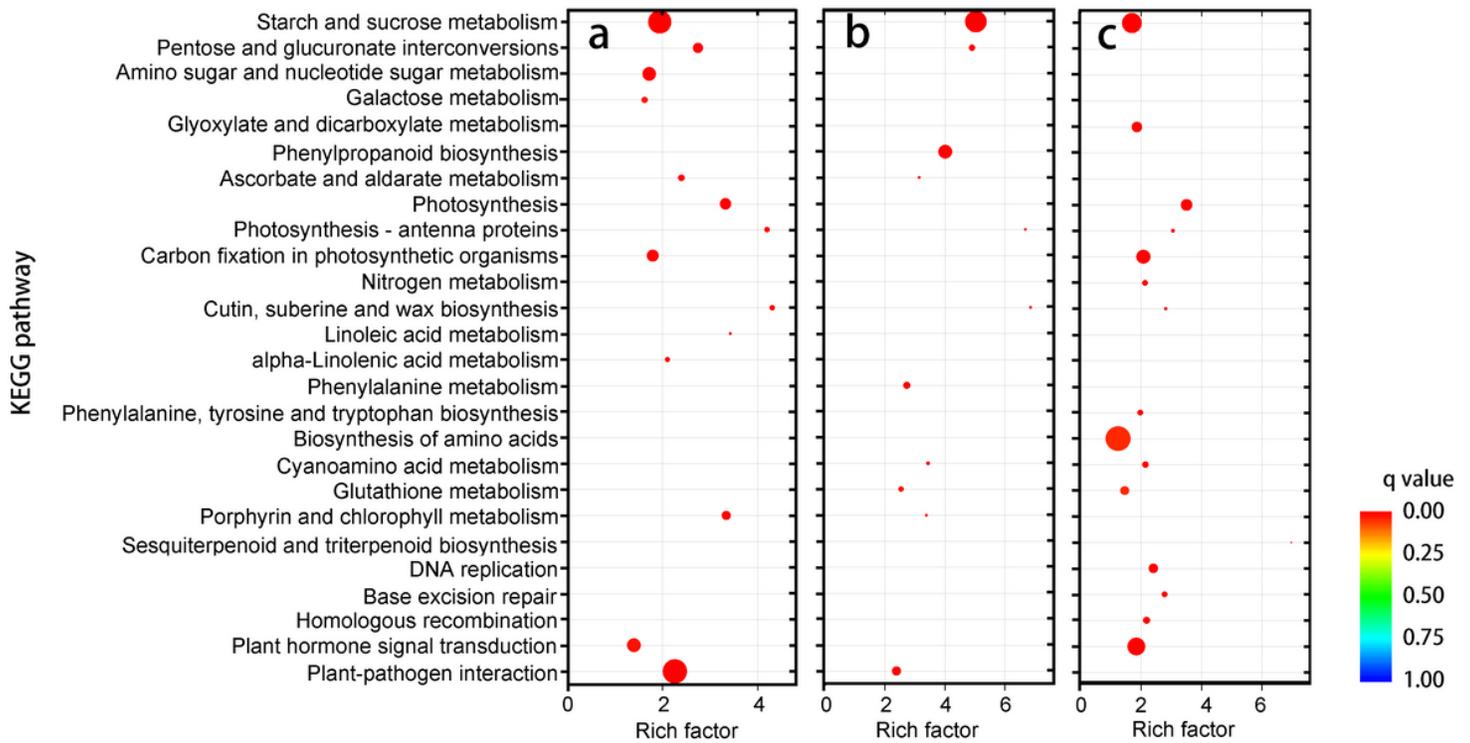


Figure 5

KEGG annotation of unigenes that were specifically expressed in P-spikelets (a), F-branchlets (b) and F-leaves (c) of arrow bamboo flowering plants. The size of dots is proportional to the number of unigenes.

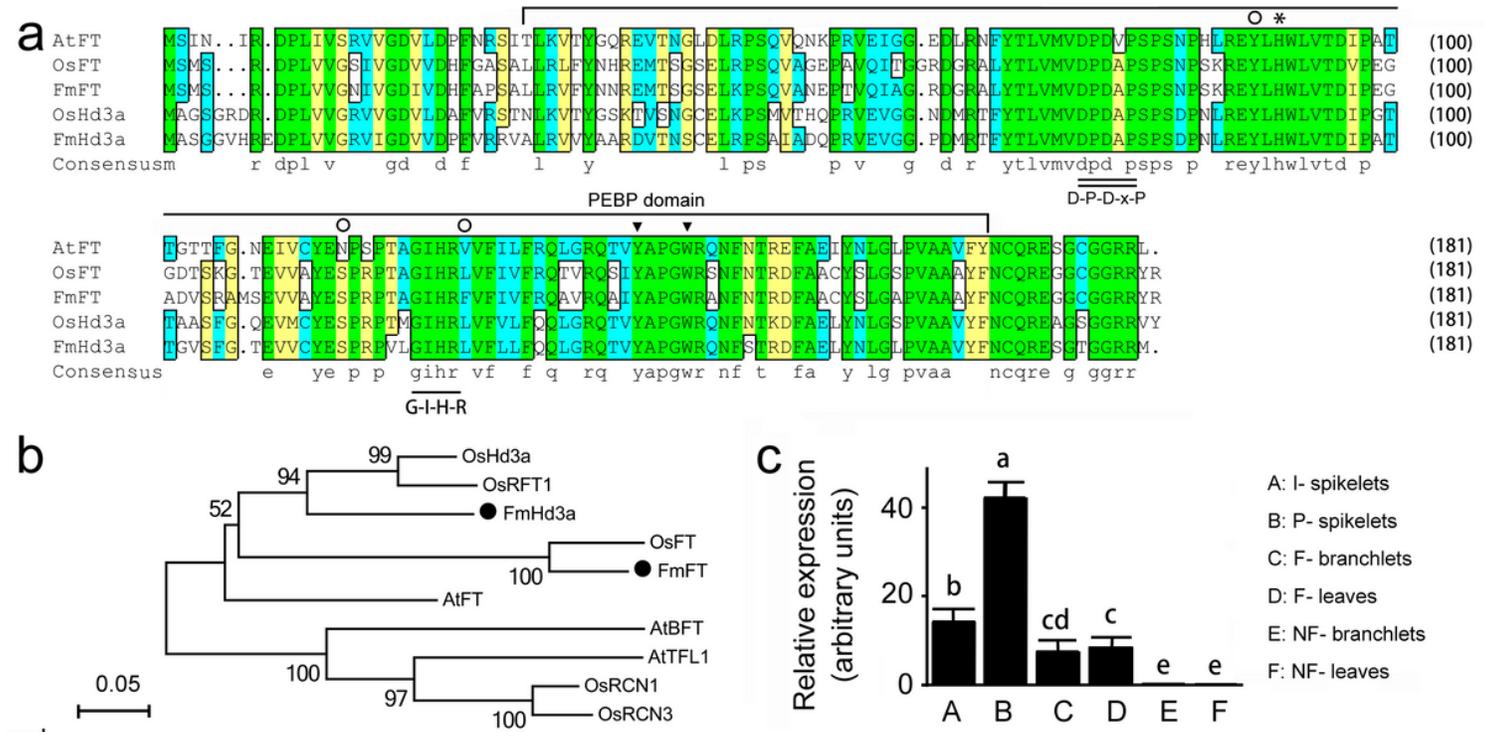
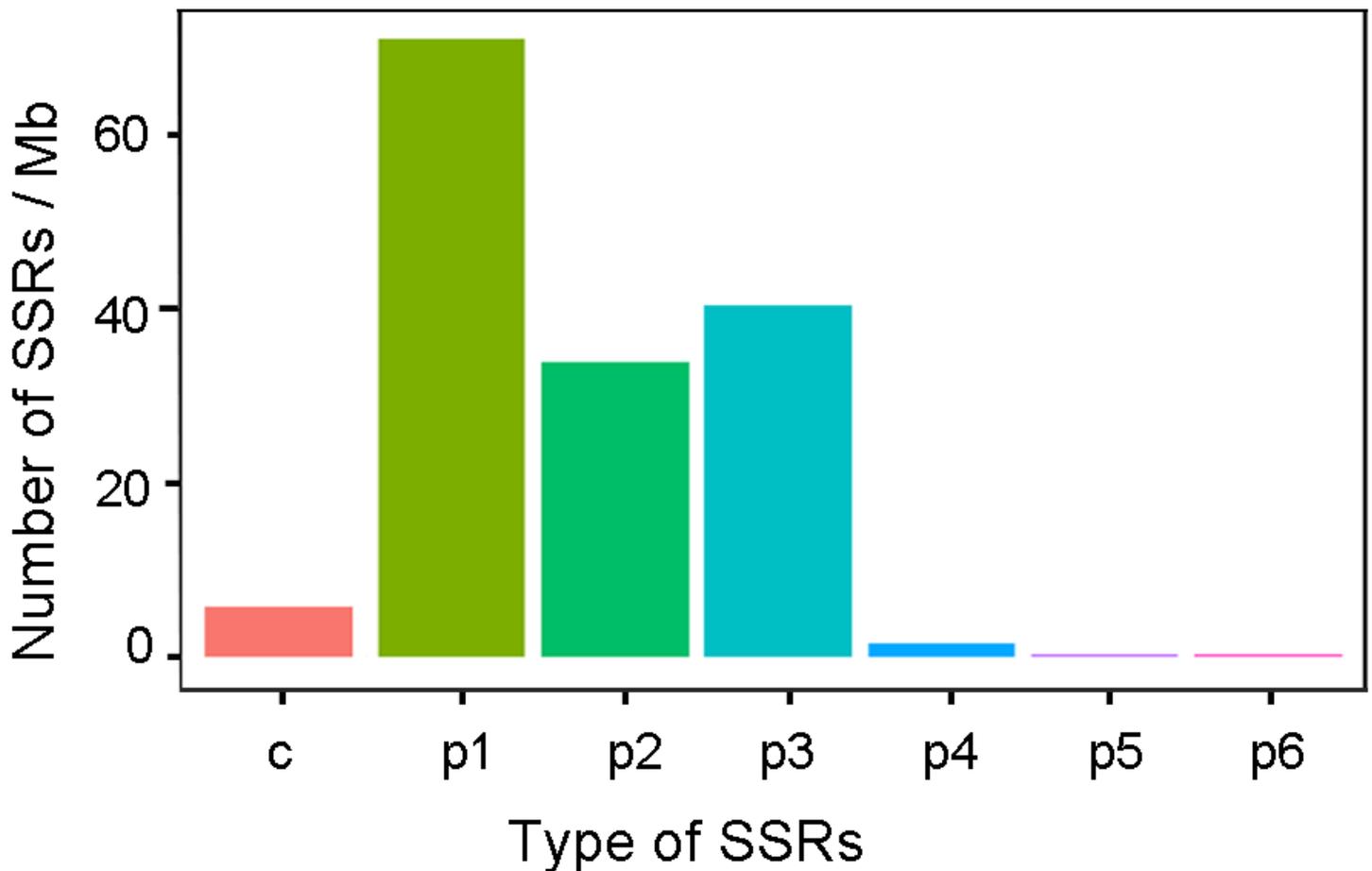


Figure 6

Alignment (a) and phylogenetic analyses (b) of FT/TFL1 proteins from *F. macclureana* and other plants, as well as the expression pattern of *FmHd3a* across six different tissues collected by RT-qPCR (c). The involving sequences and their accession numbers are: three FT/TERMINAL FLOWER 1 (TFL1) proteins from *A. thaliana*: AtFT (AT1G65480.1), AtTFL1 (NP\_196004.1) and BROTHER OF FT AND TFL1 (AtBFT, AT5G62040.1); and four FT/TFL1 proteins from *Oryza sativa*: OsFT (XP\_015619436.1), OsRFT1 (rice RICE FLOWERING LOCUS T 1 (RFT1), BAO03220.1), OsHd3a (XP\_015611892.1) and two RCNs (putative TFL1/CENTRORADIALIS (CEN) orthologs): OsRCN1 (AAD42895.1) and OsRCN3 (AAD42896.1). Amino acid sequences with double underline and single underline indicate the critical motifs in PEBP proteins [27]; asterisk and triangles indicate the invariant histidine in the PEBP domain [28] and two conserved sites in inducer FTs [29], respectively. Hollow circles indicate the position where the critical amino acid variation occurred between *Arabidopsis* TFL1 and FT, an activator and a repressor of flowering [36]. Relative expression levels were calculated using the  $2^{-\Delta\Delta CT}$  method to reflect expressions relatively more veritably. The statistical significance was tested by one-way ANOVA, considering  $P < 0.05$  and  $P < 0.01$  as statistically significant and extremely significant, respectively. Significant differences in transcript abundance between different tissues were then compared by Duncan's multiple range test using SPSS 17.0 (SPSS Inc., Chicago, USA).



**Figure 7**

Densities of different SSR types. c and p1-6 represent mono-, di-, tri-, tetrad-, penta- and hexa-nucleotide repeats, respectively.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS5.xlsx](#)
- [TableS6.xlsx](#)
- [FigureS6.docx](#)
- [FigureS1.tif](#)
- [FigureS2.tif](#)
- [FigureS4.tif](#)
- [TableS4.xlsx](#)
- [TableS3.xlsx](#)
- [TableS7.xlsx](#)
- [FigureS3.tif](#)
- [TableS2.xlsx](#)
- [TableS1.docx](#)
- [TableS8.xlsx](#)
- [FigureS5.docx](#)
- [TableS9.xlsx](#)