# An ecological transcriptome approach to capture the molecular and physiological mechanisms of mass flowering in *Shorea curtisii*

**Ahmad Husaini Suhaimi** [1], **Masaki J. Kobayashi** [2], **Akiko Satake** [3], **Ching Ching Ng** [1], **Soon Leong Lee** [4], **Norwati Muhammad** [4], **Shinya Numata** [5], **Tatsuya Otani** [6], **Toshiaki Kondo** [7], **Naoki Tani** [Corresp., 2, 8], **Suat Hui Yeoh** [Corresp. 1]

[1] Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur, Malaysia

[2] Forestry Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki, Japan

[3] Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan

[4] Forestry Biotechnology Division, Forest Research Institute Malaysia, Selangor, Malaysia

[5] Department of Tourism Science, Tokyo Metropolitan University, Tokyo, Japan

[6] Shikoku Research Center, Forestry Research and Management Organization, Kochi, Japan

[7] Bio-Resources and Utilization Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki, Japan

[8] Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

Corresponding Authors: Naoki Tani, Suat Hui Yeoh
Email address: ntani@affrc.go.jp, suathui_yeoh@um.edu.my

Climatic factors have commonly been attributed as the trigger of general flowering, a unique community-level mass flowering phenomenon involving most dipterocarp species that forms the foundation of Southeast Asian tropical rainforests. This intriguing flowering event is often succeeded by mast fruiting, which provides a temporary yet substantial burst of food resources for animals, particularly frugivores. However, the physiological mechanism that triggers GF is not well understood largely due to its irregular and unpredictable occurrences in the tall and dense forests. To shed light on this mechanism, we employed ecological transcriptomic analyses on an RNA-seq dataset of a GF species, *Shorea curtisii*, sequenced from leaves and buds collected at multiple vegetative and flowering phenological stages. We assembled 64,219 unigenes from the transcriptome of which 1,730 and 3,559 were differentially expressed in the leaf and the bud, respectively. Differentially expressed unigene clusters were found to be enriched with *Arabidopsis thaliana* gene sets that are associated with response to biotic and abiotic stresses, nutrient level, and hormonal treatments. When combined with rainfall data, our transcriptome data reveals that the trees were responding to a brief period of drought prior to the elevated expression of key floral promoters and followed by differential expression of unigenes that indicates physiological changes associated with the transition from vegetative to reproductive stages. Our study is timely for a representative general flowering species that occurs in forests that are under the constant threat of deforestation and climate change as

it pinpoints important climate sensitive and flowering-related homologs and offers a glimpse into the cascade of gene expression before and after the onset of floral initiation.

# An Ecological Transcriptome Approach to Capture the Molecular and Physiological Mechanisms of Mass Flowering in *Shorea curtisii*

Ahmad Husaini Suhaimi[1], Masaki J. Kobayashi[2], Akiko Satake[3], Ching Ching Ng[1], Soon Leong Lee[4], Norwati Muhammad[4], Shinya Numata[5], Tatsuya Otani[6], Toshiaki Kondo[7], Naoki Tani[2,8], Suat Hui Yeoh[1]

[1] Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur, Malaysia
[2] Forestry Division, Japan International Research Center for Agricultural Sciences (JIRCAS), Tsukuba, Ibaraki, Japan
[3] Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan
[4] Forestry Biotechnology Division, Forest Research Institute Malaysia (FRIM), Selangor, Malaysia
[5] Department of Tourism Science, Tokyo Metropolitan University, Tokyo, Japan
[6] Shikoku Research Center, Forestry Research and Management Organization (FRMO), Kochi, Japan
[7] Bio-Resources and Utilization Division, Japan International Research Center for Agricultural Sciences (JIRCAS), Tsukuba, Ibaraki, Japan
[8] Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

Corresponding Author:
Suat Hui Yeoh[1]
Institute of Biological Sciences, Lingkungan Budi, Lembah Pantai, Kuala Lumpur, 50603, Malaysia
Email address: suathui_yeoh@um.edu.my

Naoki Tani[2,8]
Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8577, Japan
Email address: ntani@affrc.go.jp

## Abstract

Climatic factors have commonly been attributed as the trigger of general flowering, a unique community-level mass flowering phenomenon involving most dipterocarp species that forms the foundation of Southeast Asian tropical rainforests. This intriguing flowering event is often succeeded by mast fruiting, which provides a temporary yet substantial burst of food resources for animals, particularly frugivores. However, the physiological mechanism that triggers general flowering is not well understood largely due to its irregular and unpredictable occurrences in the tall and dense forests. To shed light on this mechanism, we employed ecological transcriptomic analyses on an RNA-seq dataset of a general flowering species, *Shorea curtisii*, sequenced from leaves and buds collected at multiple vegetative and flowering phenological stages. We assembled 64,219 unigenes from the transcriptome of which 1,730 and 3,559 were differentially expressed in the leaf and the bud, respectively. Differentially expressed unigene clusters were found to be enriched with *Arabidopsis thaliana* gene sets that are associated with response to biotic and abiotic stresses, nutrient level, and hormonal treatments. When combined with rainfall data, our transcriptome data reveals that the trees were responding to a brief period of drought prior to the elevated expression of key floral promoters and followed by differential expression of unigenes that indicates physiological changes associated with the transition from vegetative to reproductive stages. Our study is timely for a representative general flowering species that occurs in forests that are under the constant threat of deforestation and climate change as it pinpoints important climate sensitive and flowering-related homologs and offers a glimpse into the cascade of gene expression before and after the onset of floral initiation.


## Introduction

General flowering (GF) is a unique intermittent flowering phenomenon that occurs at irregular intervals of 1–10 years in Southeast Asian mixed-lowland dipterocarp forests (Ashton et al., 1988; Chen et al., 2018). During a GF event, dozens of plant families, including Dipterocarpaceae which are among the most abundant trees in the canopy and emergent layers of the tropical rainforests (Ashton 1988), exhibit synchronous flowering (Appanah, 1993; Ashton et al., 1988). Each GF episode lasts for about four weeks and is typically followed by mast fruiting (Appanah, 1985; Sakai et al., 1999), resulting in a massive pulse of resources in the form of seeds and fruits that are consumed by nectarivorous and frugivorous fauna (Ashton et al., 1988; Janzen, 1974). The interval and magnitude of the GF are influenced by climate extremes such as those caused by ENSO in Southeast Asia (Sakai et al., 2006), which also affects the fauna through fluctuations in floral and fruit resources (Butt et al., 2015), thereby enhancing the ecological services they provide. The unpredictable nature of GF presents challenges for the conservation of GF species, especially in terms of seed collection (Kettle, 2010). Determining the cues that dictate the timing of GF is crucial for conservation management of dipterocarp species in a region that is constantly under threat of deforestation.

The absence of clear seasonality in the aseasonal tropics poses difficulty in identifying proximate cues of GF. In temperate plants, photoperiod or day length is a key factor in reproductive initiation (Andres & Coupland, 2012). Given the minimal day length variation in the tropics (Janzen, 1967), this may not be the case in tropical plants. Instead, several studies have suggested that other climatic factors, such as cool temperatures and drought, signal flowering time in GF species (Chen et al., 2018; Ushio et al., 2020; Yeoh et al., 2017). However, little is known about how climate change can alter the flowering phenology in tropical trees (Satake et al., 2022). To fill this knowledge gap, the information available on the model species, *Arabidopsis thaliana*, coupled with the ecological transcriptome approach that combines transcriptome and meteorological data (Richards et al., 2009), offer a means to elucidate the cues and molecular mechanism of flowering in GF species, thus enabling the prediction of flowering time and the effect of changing climate on GF.

Progress to understand the molecular mechanism of GF has been slow mainly due to limited sample availability for frequent sample collection and difficulties in accessing the dense and tall canopies of tropical forests. Despite the sampling challenges, several studies have made use of gene expression and ecological data to infer the environmental factors regulating GF. For example, comparisons between transcriptome sequencing data of bud samples from *Shorea beccariana* and rainfall data supported drought as a significant floral trigger (Kobayashi et al., 2013). A separate study conducted on the expression of two key flowering homologs, *FLOWERING LOCUS T* (*FT*) and *LEAFY* (*LFY*), in two *Shorea* species, *S. curtisii* and *Shorea leprosula,* using real-time PCR found that the expression levels of these two genes were strongly correlated with floral initiation and that the synergistic effect of drought and low temperature was a reliable predictor of GF (Yeoh et al., 2017). This report was followed by our recent research using RNA-seq data from *S. curtisii* leaf samples collected at three time points including before and after floral initiation, where homologs of *A. thaliana* genes that are associated with drought response were found to have a significantly higher representation among genes that were differentially expressed during flowering time (Suhaimi et al., 2023). Both the transcriptome studies on the members of red meranti group (Dipterocarpaceae), *S. curtisii* (Suhaimi et al., 2023) and *S. beccariana* (Kobayashi et al., 2013) have reported that majority of *A. thaliana* flowering-time genes, including key floral promoters, are conserved in sequence and function in the dipterocarp species. All these studies have shown that despite differences in the flowering physiology and phenology between *A. thaliana* and *Shorea* spp., the former is still a valuable model for understanding the molecular mechanisms in dipterocarps.

The studies thus far, nevertheless important, were conducted on a single tissue (Kobayashi et al., 2013; Suhaimi et al., 2023) or a small number of key flowering genes (Yeoh et al., 2017). However, physiological changes during flowering span across multiple organs (Parcy, 2005) with both leaves and buds playing direct roles in the production and transportation of proteins that regulate floral initiation and as the primary site of flower development in the trees. Concurrent study of both the organs is needed to capture the missing link and gain a

comprehensive view of the gene expression dynamics that occur between the two organs. In our current study, we utilized the transcriptome of both leaf and bud tissues, and more than double the number of time points compared to our previous study (Suhaimi et al., 2023) to identify genes that are important during GF and gain a greater resolution of the molecular events leading to GF. We further characterized differentially expressed unigenes by comparing them to *A. thaliana* gene sets that associate with responses to biotic and abiotic stresses, nutrient levels, and hormonal treatments. We also integrated meteorological records to uncover the molecular pathways interacting with the environmental signals and finally discussed the possible mechanisms involved in the floral regulation of this GF species.

## Materials & Methods

### Study species and sample collection

The species selected for this study, *Shorea curtisii* (Dipterocarpaceae), and other trees from the genus *Shorea* section *Mutica*, have been documented as reliable indicators of GF (Ashton et al., 1988). *Shorea* spp. can be found occupying ridges of dipterocarp forests in the Malay Peninsula ranging from 300–800 m above sea level (Appanah & Chan, 1981).

The samples in this study were collected from two *S. curtisii* trees, designated as C1 and C2, with diameter at breast height of 72 cm and 89 cm, respectively (Yeoh et al., 2017). The trees are located at Semangkok Forest Reserve (2°58'N, 102°18'E, 340–450 m above sea level) in the state of Selangor Darul Ehsan, Peninsular Malaysia (Fig. 1A). The permission to conduct this research in the forest reserve was granted by the Selangor State Forest Department. The forest reserve consists of a 6-ha (200 m × 300 m) hill dipterocarp forest conservation plot, and another 5.4-ha (140 m × 400 m) plot established on a selectively logged area (Tani et al., 2015; Yagihashi et al., 2010). Daily temperature and precipitation data were collected from Hospital Kuala Kubu Bharu meteorological station and Kampung Pertak hydrological station, respectively (Fig. 1B), and were previously reported in Yeoh et al. (2017). The cumulative rainfall data (defined as 30-d moving rainfall total) was calculated from the daily rainfall record. Drought is defined as < 40 mm cumulative rainfall (Sakai et al., 2006).

To obtain gene expression profiles of the study species over a flowering period, we collected leaves and buds sampled from the top canopies of the two *S. curtisii* trees around noon (Table S1). These samples, which are part of the collection described in Yeoh et al. (2017), were chosen from eight time points (TPs): 08 July 2013 (TP-A), 17 October 2013 (TP-B), 18 December 2013 (TP-C), 14 February 2014 (TP-D), 21 March 2014 (TP-E), 02 April 2014 (TP-F), 14 May 2014 (TP-G), and 17 June 2014 (TP-H). The TPs were selected to represent different flowering stages of the trees, namely vegetative, inflorescence, flowering, and abortion/fruiting stages (Table S1). After collection, the samples were immediately submerged in RNAlater solution (Ambion, Austin, Texas, USA) and stored at –80°C until RNA extraction. Using RNASeqPower package

(Hart et al., 2013) in R v4.0.3, the statistical power of the experimental design was calculated to be 0.38 for a sample size of two (the number of biological replicates used in the study).

## Transcriptome sequencing and annotation

The workflow of the current study is outlined in Fig. 2. Total RNA was extracted from leaf and bud samples using CTAB method described by Kobayashi et al. (2013). The extracted RNA was used to create paired-end cDNA libraries as per manufacturer's protocol (Illumina, USA) followed by RNA-seq using Illumina HiSeq 4000 Sequencer (Illumina, USA). Low quality bases and traces of Illumina adapters were removed using Trimmomatic v0.36 (Bolger et al., 2014) prior to *de novo* assembly using Trinity v2.8.5 (Grabherr et al., 2011). To minimize transcript redundancy, only transcripts with complete open reading frames (ORFs) as predicted by TransDecoder v5.5.0 (http://transdecoder.github.io) were retained. The number of transcripts was further reduced by clustering highly similar protein sequences using CD-HIT v4.8.1 (Fu et al., 2012) (parameters: -n 5 -c 1). The completeness of the assembly was assessed using BUSCO v5.2.1 (Manni et al., 2021). Unless otherwise stated, all bioinformatic analyses in this study were conducted using default parameters.

The non-redundant transcripts, henceforth referred to as unigenes, were annotated by querying against the proteome of *A. thaliana* (Cheng et al., 2017) using BLASTx v2.10 (Camacho et al., 2009) with E-value cut-off: 1E–10. The annotated unigenes were compared to the *A. thaliana* flowering genes database (Bouché et al., 2016) to identify their homologs in *S. curtisii*. Translated unigenes were searched against UniProt database using BLASTp v2.10 (Camacho et al., 2009) with E-value cut-off: 1E–05, as well as Pfam v32.0 (El-Gebali et al., 2018) and PANTHER v14.1 (Thomas et al., 2003) databases using InterProScan v5.36 (Jones et al., 2014). The unigenes were assigned Gene Ontology (GO) terms (Ashburner et al., 2000) associated with the earlier BLASTx matches and visualized using WEGO online tool (Ye et al., 2018). The unigenes were also queried against KEGG database (Kanehisa & Goto, 2000) using BlastKOALA online tool (Kanehisa et al., 2016).

## Differential expression analysis and enrichment tests

Transcript quantification was conducted using Salmon v1.5.1 (Patro et al., 2016). To summarize the expression data, we performed a principal component analysis (PCA) using *pcaExplorer* package v2.22.0 (Marini & Binder, 2019) in R, which normalized the expression values using DESeq method (Anders & Huber, 2010) prior to PCA. Subsequent bioinformatic analyses were performed on leaf and bud samples separately. We conducted differential expression analysis using DESeq2 software v1.32.0 (Love et al., 2014) by comparing pairwise TPs to identify differentially expressed unigenes (DEUs; defined as transcripts with absolute $\log_2$ fold change ≥ 1 and false discovery rate (FDR) < 1E–03). To identify significantly enriched GO terms (P <

0.05) in the DEUs, enrichment analysis was performed using topGO package v2.40.0 (Alexa & Rahnenfuhrer, 2020) in R. Enriched KEGG pathways (FDR < 0.05) were identified using an online tool, KO-Based Annotation System (KOBAS) v3.0 (Bu et al., 2021).
To validate the expression profiles of the unigenes, the expression of key flowering gene, *FLOWERING LOCUS T* (*FT*), homologs was compared to relative expression values measured by qRT-PCR in a previous study in *S. curtisii* (Yeoh et al., 2017). We utilized *ggpubr* package in R to perform Spearman's correlation test and to plot linear regression scatter plots.

## Clustering and characterization of differentially expressed unigenes

To summarize the expression profiles of the DEUs, the unigenes were clustered based on their expression patterns using *coseq* package (Rau & Maugis-Rabusseau, 2018) in R. The expression values were normalized using Trimmed Mean of M-values (TMM) method (Robinson & Oshlack, 2010), transformed using centered log ratio (Aitchison, 1982), and clustered using K-mean algorithm (MacQueen, 1967).

For the characterization of the DEUs, we followed the method used in the transcriptome study of *S. beccariana* (Kobayashi et al., 2013), which compared DEUs to *A. thaliana* transcriptome data and assumed that the regulation of homologs in both species is similar. This approach enabled functional annotations of DEUs with categories that are not covered by the GO database, such as "response to prolonged moderate drought" (Kobayashi et al., 2013). We compared the DEUs in *S. curtisii* to 35 gene sets of up- and downregulated genes in *A. thaliana* that were exposed to various external and endogenous factors (Scheible et al., 2004; Misson et al., 2005; Cao et al., 2006; Gould et al., 2006; Nemhauser et al., 2006; Ma and Bohnert, 2007; Osuna et al., 2007; Harb et al., 2010; see Table S2 for details). Fisher's exact tests with Bonferroni multiple testing correction were performed to determine the significance in the number of overlapping genes between DEUs and the *A. thaliana* gene sets (Fig. S1). When the number of DEUs that overlaps with the genes in the gene set is significant, we assumed that the *S. curtisii* trees used in this study experienced similar conditions as the *A. thaliana* from which the gene set was obtained. To determine if the significantly tested *A. thaliana* gene sets showed specific expression patterns, we further compared the gene sets to DEU clusters using gene enrichment tests (Fig. S2; Tables S14−S19, S21−S25). The enrichment tests were performed using Fisher's exact tests with Bonferroni corrections. We assumed enriched gene sets had the same expression pattern as their significantly associated DEU clusters. Finally, the DEU clusters with significantly enriched gene sets were compared to the temperature and rainfall records to determine potential association between the transcriptome and meteorological data.

# Results

## RNA sequencing and transcriptome assembly

In this study, the transcriptome profile of *S. curtisii* at different stages of flowering (vegetative, inflorescence, flowering, and fruiting/abortion) was captured by sequencing total RNA from leaves and buds at eight time points (TPs). The sequencing generated 1,595,323,292 reads with an average length of 150 bp which were assembled into 664,053 transcripts. After removing redundant transcripts, the resulting transcriptome assembly comprises 64,219 unigenes with an average length of 2,148 bp and N50 of 2,624 bp (Fig. S3). The GC content of the assembled transcriptome was 43.02%. The BUSCO analysis indicated an almost complete transcriptome assembly in which 95.5% of highly conserved proteins in the Embryophyta lineage were recovered (Table S3). Of the total unigenes, 58,722 (91.44%) were annotated with at least one of the public protein databases (Table 1; Tables S4 and S5; Fig. S4).

We conducted PCA on normalized unigene expression data to obtain an overview of the variance in gene expression among samples. The analysis clustered the samples based on tissue type, individual tree, and TPs (Fig. S5A). Most of the variance (PC1: 34.71%) can be attributed to transcriptional differences between buds and leaves (Fig. S5A). Variation in transcriptional profiles between the two biological replicates was also high (PC2: 18.63%; Fig. S5A). In bud, 11.29% of the variation (PC3) distinguished majority of the vegetative samples (TP-A–D) from the inflorescence and reproductive samples (TP-E–H; Fig. S5B). The clustering of TP-E with TPs at the inflorescence and reproductive stages suggests that the switch to inflorescence may have already begun although morphologically the buds at TP-E are still in vegetative stage. Interestingly, a number of the top contributing unigenes for PC3 (Table S8) are homologs of stress-responsive genes such as *EXPANSIN A15* (Wieczorek et al., 2006) and *WRKY33* (Zheng et al., 2006).

## Differential gene expression and clusters of differentially expressed unigenes

To identify candidate unigenes that are involved in floral regulation of *S. curtisii*, we conducted a differential expression analysis for pairwise comparison of all TPs. A total of 1,730 and 3,559 DEUs were identified in leaf and bud samples, respectively, with 789 unigenes shared between the two tissues (Tables S9 and S10). The differences in the number of DEUs between the tissues may be due to the regulation of organ-specific unigenes that function in floral morphogenesis, as previously reported in *A. thaliana* (Smaczniak et al., 2017). To summarize the expression profiles of the DEUs, K-mean clustering was performed, resulting in the identification of six DEU clusters with distinctive expression profiles in leaf transcriptomes (Fig. 3) and five DEU clusters in bud transcriptomes (Fig. 4).

## Differential expression of putative flowering-time unigenes

To gain insight into the floral regulatory pathways of *S. curtisii*, we first identified homologs of *A. thaliana* flowering-time genes from the annotated unigenes. Of the 51,654 unigenes that could be mapped to *A. thaliana* genes, 918 unigenes were homologous to 228 non-redundant flowering-time genes (Table 2 and Table S11). This indicates that >75% (228/306) of *A. thaliana* flowering-time genes have homologs in *S. curtisii,* similar to the findings in a closely related species, *S. beccariana* (Kobayashi et al., 2013). Out of the total *A. thaliana* flowering-time homologs, 103 were differentially expressed throughout the study period (Table S11).

There are 274 flowering-time homologs associated with the photoperiod and circadian clock pathway (Table 2 and Table S11). Photoperiod-associated homologs accounted for over half (15/26 DEUs) of flowering-time DEUs in leaf, while only one third (37/96 DEUs) in bud were from this pathway (Table S11). The number of homologs related to the circadian clock and photoperiod pathway in this study is consistent with our earlier study in *S. curtisii* leaves despite using fewer time points in the previous study (Suhaimi et al., 2023). We also identified 99 homologs belonging to the vernalization pathway, 34 homologs from the autonomous pathway, and 16 ambient temperature-responsive homologs in the transcriptome. Of the homologs from the vernalization pathway, eight were differentially expressed only in bud, while two homologs of *VERNALIZATION 5* (*VRN5*) were differentially regulated in both leaf and bud (Table 2). Only two homologs from the autonomous pathway were differentially expressed in bud (Table 2). Of the homologs associated with ambient temperature, two homologs of *RELATED TO ABI3/VP1 2* (*RAV2*) were differentially expressed in both leaf and bud (Table S11). Additionally, we identified 41 homologs of integrators in this study, ten of which were differentially expressed only in bud. Two homologs of *FT*, namely *ScFT1* and *ScFT2*, were differentially regulated in both tissues (Table S11). The expression of these *FT* homologs is congruent with qRT-PCR data from a previous study (Yeoh et al., 2017; Fig. S6), thus, validating the RNA-Seq results.

## Characterization of the differentially expressed unigenes through *A. thaliana* gene sets enrichment analyses

Only 2.7% of the total DEUs (122/4,500) are homologs of *A. thaliana* flowering-time genes (Tables S9–11), which suggests that most of them have functions that are not directly related to known flowering mechanisms in the model species. To characterize these DEUs, enrichment analyses were performed on 35 gene sets of *A. thaliana* (Table 3) and the DEUs separately for leaf and bud samples. We found 25 *A. thaliana* gene sets to be significantly enriched in the DEUs for both *S. curtisii* leaf and bud (Table 3; Tables S12 and S13). These gene sets include responses to drought and temperature, as well as changes in level of hormones such as jasmonic acid (JA) and gibberellin (GA). We assumed that the trees experienced similar conditions as the *A. thaliana* from which the gene sets were obtained when the tests were significant.

## Expression profiles of the enriched gene sets

To determine if the enriched gene sets were expressed in a specific pattern across the study period, enrichment analysis was conducted on gene sets that were found to be significant (Table 3) and the DEU clusters in leaf and bud of *S. curtisii* (Figs. 3 and 4). We assumed that the expression pattern of gene sets enriched in a particular DEU cluster is represented by that of the cluster. Overall, 12 gene sets were enriched in at least one DEU cluster in leaf (Fig. 3 and Table S20) while 13 gene sets were overrepresented in at least one DEU cluster in bud (Fig. 4 and Table S26). Notably, a number of gene sets related to nutrient levels, biotic and abiotic stresses, and hormonal treatments were enriched in DEU clusters that showed similar expression in both trees (Figs. 3 and 4).

Our analysis of DEU clusters in leaf has found that the homologs of *A. thaliana* genes that are upregulated under medium-term phosphate-limited condition were enriched in DEU cluster 1 (Fig. 3A). The expression profile of this cluster indicates a decrease in phosphorus level between TP-B and TP-D, followed by another decrease prior to flowering. DEU cluster 3 in leaf was overrepresented with homologs of genes commonly upregulated in response to biotic and abiotic stresses, as well as genes upregulated with hormonal treatment, specifically abscisic acid (ABA), indole-3-acetic acid (IAA, or commonly known as auxin), and JA (Fig. 3C). These phytohormones are also involved in plant stress response (Yang et al., 2019). Furthermore, GO terms related to stress response such as “response to reactive oxygen species” and “MAPK cascade”, were enriched in this cluster (Table S30). The expression of the DEUs was elevated from TP-D to TP-F (Fig. 3C), suggesting that the plants were under stress during this period. DEU cluster 4 in leaf was enriched in genes that are upregulated under carbon-limited condition and downregulated with addition of sucrose (Fig. 3D). The expression profile of this cluster suggests that the concentration of carbon in the leaf was decreasing at the onset of flowering, whilst sucrose level was increasing prior to floral induction, specifically at TP-E in C1 and TP-D in C2. Additionally, the KEGG pathway analysis identified the “carbon metabolism” pathway as being enriched in this cluster (Table S30). Finally, DEU cluster 6 in leaf was overrepresented with gene sets responding to drought and changes in the levels of GA and carbon (Fig. 3F). However, the expression patterns of the DEUs in the cluster varied between the two individuals, implying that the response observed was specific to each individual. As a result, the enriched gene sets in the cluster were not further analyzed.

Examination of DEU clusters in bud identified that the homologs of *A. thaliana* genes upregulated in response to carbon-limited condition were enriched in DEU cluster 1 (Fig. 4A). The expression pattern of the cluster suggests that the level of carbon in bud was depleted as flowering approached (Fig. 4A), which is consistent with the results of enrichment tests in leaf (Fig. 3D). GO terms related to catabolism such as “polysaccharide catabolic process”, “cellular amide catabolic process”, and “allantoin catabolic process”, were also found to be overrepresented in the cluster (Table S31). This aligns with previous research in *A. thaliana*

which demonstrated that carbon starvation leads to an upregulation of genes involved in catabolism (Osuna et al., 2007). Among the DEU clusters in bud, clusters 2 and 4 showed inverse expression patterns in both individuals (Figs. 4B and 4D). Cluster 2 was enriched with homologs of genes that are downregulated under moderate drought, while cluster 4 was not. The expression profile of DEU cluster 2 suggests that the *Shorea* individuals were experiencing mild drought between TP-D and TP-F. Additionally, both clusters 2 and 4 were enriched with homologs of genes that respond to elevated level of GA. The expression profiles of these two clusters indicate an increase in GA level between TP-D and TP-F. Both clusters were also enriched with GO terms related to stress, signaling, growth, and floral regulation, such as "response to reactive oxygen species", "vasculature development", "MAPK cascade", and "signal transduction" in cluster 2, as well as "response to hypoxia", "response to light stimulus", "microtubule cytoskeleton organization", and "cytokinesis" in cluster 4 (Table S31). Our PCA also revealed that the majority of top contributing unigenes to PC3 belong to these two clusters (Table S8), highlighting the importance of the unigenes in these clusters in the transition from vegetative to reproductive stages. In DEU cluster 3, we found homologs of *A. thaliana* genes that are typically upregulated in response to diverse abiotic and biotic stresses, as well as genes that are affected by cytokinin and JA treatment (Fig. 4C). However, this cluster showed inconsistent expression profiles between the two *S. curtisii* individuals (Fig. 4C), hence, the enriched gene sets in this cluster were not examined further. Lastly, DEU cluster 5 in bud was overrepresented with homologs of genes upregulated under moderate drought and downregulated with increasing temperature (Fig. 4E). The expression profile of this cluster indicates an increase in temperature between TP-D and TP-E (Fig. 4E), which is in line with the enrichment of relevant GO terms such as "response to stress" and "response to stimulus" (Table S31).

## Association between meteorological data and transcriptome profiles

Prolonged moderate drought and drop in temperature have been hypothesized as proximate cues for GF (Ashton et al., 1988; Brearley et al., 2007; Chen et al., 2018; Yeoh et al., 2017). Therefore, we examined meteorological data to look for these signals prior to flowering. These data were then compared to the expression patterns of DEU clusters that were significantly enriched with *A. thaliana* gene sets related to temperature or drought (Figs. 3 and 4) to infer the regulatory pathways that respond to these climatic cues.

The daily mean temperature at the nearest meteorological station situated 11 km from the study site was 26.8°C, while the total annual precipitation at the nearest hydrological station 6 km from the study site was 2,272 mm. The lowest minimum daily temperature (18.3°C) was recorded in early February around TP-D (Fig. 5B). We also calculated 30-day cumulative rainfall to identify drought periods, which are defined as 30-day cumulative rainfall of below 40 mm (Sakai et al., 2006). The cumulative rainfall data showed a period of drought beginning before TP-D and lasting until TP-E (Fig. 5C). Therefore, the expression patterns of leaf DEU cluster 3 (Fig. 3C) as well as bud DEU clusters 2, 4, and 5 (Figs. 4B, 4D and 4E) were examined.

Among these DEU clusters, only DEU cluster 5 in bud was enriched with homologs of genes downregulated with increased temperature (Fig. 4E) and its expression pattern implies that the plants were responding to rising temperature around TP-D. This pattern concurs with the daily mean temperature data, which showed a slight increase within the same period (Fig. 5A). Although the increase in temperature was subtle, it was not unexpected for the plants to be able to capture the slight change in temperature stimuli as it has been suggested that trees in tropical regions are more sensitive to change in their surroundings (Singh & Kushwaha, 2016). Of the DEU clusters overrepresented with homologs of genes responding to drought, only DEU cluster 3 in leaf (Fig. 3C) and cluster 2 in bud (Fig. 4B) displayed expression patterns consistent with the drought period (Fig. 5C). This indicates that the trees were experiencing a moderate drought between TP-D and TP-E. The expression pattern of DEU cluster 5 in bud (Fig. 4E), on the other hand, did not match the cumulative rainfall data.

## Discussion

Our study found high transcriptional variation between leaf and bud of *S. curtisii* during GF, including the expression of *A. thaliana* homologs involved in flowering. Despite the inter-individual differences in the expression profiles, we were able to identify and characterize transcriptional profiles of unigenes that are differentially expressed in the vegetative and reproductive stages of the trees. We interpreted the outcome of our transcriptome analysis along with available meteorological data and proposed a preliminary framework of floral initiation and development in this GF species.

### Proximate environmental cues for general flowering

The increasing rate of climate change, characterized by higher temperature fluctuations (IPCC, 2022) and more frequent and severe droughts (Dai, 2013; Trenberth et al., 2014), among others, is expected to have a disproportionate impact on the phenology of plants and animals in tropical regions (Santer et al., 2018). The fitness and abundance of tropical communities, including GF species, can be adversely affected by a slight temperature increase (Lister & Garcia, 2018; Numata et al., 2022). As GF species is a valuable source of sustenance for a variety of seed and fruit predators including wild pigs and rodents (Ickes, 2001; Miura et al., 1997), alteration in their phenology could have potential cascading effects on the tropical forest communities.

Long term phenological records of tropical rainforests in the Malay Peninsula have shown that flowering events are always preceded by a drop in daily minimum temperature after a brief period of drought (Ashton et al., 1988; Sakai et al., 2006). In fact, mathematical models that take into account the synergistic effects of both drought and a drop in daily minimum temperature have been found to be reliable at predicting GF occurrences (Chen et al., 2018; Yeoh et al., 2017). The cumulative rainfall data (Fig. 5C), along with the expression profile of DEU clusters that were enriched with homologs of drought-responsive genes (Figs. 3C and 4B) indicate that

the *S. curtisii* in our study experienced water stress due to a moderate drought lasting from before TP-D until TP-E. The upregulation of *ScFT1* (Unigene: TRINITY_DN6921_c0_g2_i2) in leaf shortly after drought began (TP-D; Table S9) and the upregulation of *ScFT1* in bud post-drought (TP-E; Table S10) further corroborated the hypothesis that drought is one of the floral triggers in dipterocarps (Kobayashi et al., 2013; Sakai et al., 2006; Yeoh et al., 2017). Moreover, elevation in the levels of ABA, IAA, and JA hormones (Fig. 3C) that coincides with the occurrence of drought also supported the aforementioned hypothesis. Roles of these hormones in the floral regulation and signaling pathways of plants responding to stress, including drought, have been established in many other species such as *A. thaliana*, *rice*, and *Citrus* spp. (reviewed in Singh & Laxmi, 2015; Yang et al., 2019). Among these hormones, ABA is particularly important in the hormonal response to drought stress (Gupta et al., 2020) as it regulates stomatal opening to reduce water loss (Okamoto et al., 2013). Interestingly, ABA has also been reported to function in drought-induced flowering in *A. thaliana* (Riboni et al., 2016) and *Citrus* spp. (Khan et al., 2022; Li et al., 2017). Auxin (IAA) plays a role in modulating leaf water uptake and regulating the expression of antioxidant enzymes to detoxify reactive oxygen species (Shi et al., 2014). Similarly, JA contributes to abiotic stress responses by increasing resistance to oxidative stress (Wu et al., 2012) and maintaining downstream drought-responsive pathways such as those related to modulation of shoot biomass and photosynthetic rate (Wang et al., 2021).

Despite a sharp drop in daily minimum temperature observed prior to TP-D (Fig. 5B), we did not find any gene set associated with differential regulation in response to decrease in temperature to be overrepresented in any of the DEU clusters (Figs. 3 and 4). However, several studies have shown that temperature drops always precede GF (Ashton et al., 1988; Numata et al., 2003; Yasuda et al., 1999), and that the occurrence of drought alone does not always result in GF (Chen et al., 2018; Yeoh et al., 2017). Therefore, it is possible that *S. curtisii* employs different temperature-responsive mechanisms than *A. thaliana*. This scenario is not unlikely, as similar divergence has been reported in other floral regulatory pathways. For example, *FLOWERING LOCUS C* (*FLC*) encodes a key floral repressor in the vernalization pathway of *A. thaliana*, grasses have developed vernalization response that is completely independent of FLC, relying instead on four central genes: *VERNALIZATION 1–3* (*VRN1–3*) and *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT2*) (reviewed in Trevaskis et al., 2007). Further studies using transcriptome profiles of GF trees treated with low temperature and drought under controlled environments could help to verify the potential of low temperature as a flowering cue and further prove that a short period of drought can induce flowering in the species. However, this would be a practically challenging and costly endeavor for tall tropical trees that flower irregularly.

Additionally, several homologs of flowering-time genes were found to be differentially expressed after the short dry spell (Figs. 4 and 5; Tables S9–11). Given the accumulating evidence of organisms and populations adapting to climate change through phenotypic plasticity and genotypic evolution (reviewed in Peñuelas et al., 2013), further research focusing on these

homologs could help us to understand how these rapid environmental changes will impact the reproductive phenology of GF species and how this phenomenon will evolve in the future.

## Resource dynamics in *Shorea* during flowering time

Accumulated nutrients have been proposed as a limiting factor in GF, which could explain why some mature trees do not flower during GF (Ichie et al., 2013; Kelly & Sork, 2002). In the present study, we found molecular evidence showing a decrease in phosphorus level prior to flowering (Fig. 3A). This finding is consistent with a previous report that investigated the mineral nutrient storage dynamics prior to mast reproduction in another dipterocarp species, *Dryobalapnos aromatica* (Ichie & Nakagawa, 2013). Although we identified homologs of *A. thaliana* genes responsive to nitrogen to be overrepresented in the DEUs of *S. curtisii* (Table 3), these unigenes did not show any specific expression pattern (Figs. 3 and 4). Besides phosphorus and nitrogen, carbohydrate has been demonstrated to be utilized for floral initiation and subsequent development of floral organs (Peng & Iwahori, 1994). The results of our cluster analysis indicate that sucrose level increased in leaves during the brief period of drought (Fig. 3D), similar to observations in other trees such as East Asian white birch (Bhusal et al., 2021) and tropical starfruit (Pingping et al., 2017). Furthermore, the cluster analysis suggests a reduction in carbon concentrations in bud prior to flowering time (Fig. 4A). Previous studies have suggested that flower bud act as carbohydrate 'utilizing sinks', metabolizing rather than storing carbohydrate (Eshghi et al., 2007; Ho, 1988).

The integration of nutritional content information such as nitrogen, phosphorus, and carbon concentrations, in future studies would be useful to understand the roles and dynamics of plant resources during flowering time. However, collecting sufficient tissues from carbohydrate sinks such as stems and trunks (Kozlowski, 1992) without damaging the branches was not feasible in this study due to the need for frequent repeated sampling to capture the floral initiation time points of these trees with unpredictable flowering intervals. Besides that, canopy access for tall tropical trees has thus far been limited, hence observations and sample collections in our study were done by climbing approximately 40 m above the ground canopy. Despite these challenges, we were able to conduct the present study on two individual trees and obtain clear and consistent results.

Based on the findings in the current study and earlier reports (Chen et al., 2018; Ichie et al., 2013; Kobayashi et al., 2013; Yang et al., 2019; Yeoh et al., 2017), we propose a preliminary framework that depicts the interactions between external environmental cues, phytohormones, nutrient resources and expression of florigen, *FT* in *Shorea* (Fig. 6) during GF. According to this framework, flowering signals such as drought (and potentially, drops in daily minimum temperature) accumulate over a period of 2–3 months prior to floral initiation. During this period, the levels of stress-associated phytohormones, including IAA, ABA, and JA, increase, leading to the upregulation of the key flowering gene, *FT* in the leaf. Subsequently, *FT*

expression in the bud is elevated, thus initiating the transition to the reproductive stage. Further research is needed to develop and complete this proposed framework and determine the extent of its applicability to other GF species.

## Conclusions

To date, this is the first genome-wide transcriptome sequencing study conducted concurrently on both leaf and bud of a dipterocarp species over a flowering season. The characterization of *S. curtisii* transcriptome in this study has led to the identification of numerous unigenes that are homologous to *A. thaliana* flowering-time genes, including unigenes that were differentially expressed during the flowering season. We have also identified unigenes that are important during the switch from vegetative to reproductive stages but are not known to be involved in the *A. thaliana* flowering mechanism, hence warrant further investigation. Our ecological transcriptome approach, which involved cluster analysis of DEUs in *S. curtisii* and enrichment analysis with *A. thaliana* gene sets and climatic data, suggests that the trees perceived proximate flowering cues, such as drought, approximately three months prior to floral initiation. This was followed by changes in hormone levels, as well as phosphorus and carbohydrate contents, as the trees entered reproductive stages. The outcomes of this study provide insight into the molecular and physiological mechanisms underlying floral regulation in a GF species and highlight the need for further research in this area to fully elucidate these complex processes.

## Acknowledgements

We acknowledge the Selangor State Forest Department for granting the permission to conduct this research in Semangkok Forest Reserve. The authors would also like to thank B Yasri and P Ramli for their assistance in the fieldwork, and Dr. KKS Ng for his assistance in RNA extraction. This research was supported in part through computational resources provided by the Data Intensive Computing Centre, Universiti Malaya.

## References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)* 44(2), 139-160.

Alexa, A., & Rahnenfuhrer, J. (2020). topGO: Enrichment analysis for Gene Ontology.

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Nature Precedings* 11(10).

Andres, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics* 13(9), 627-639. doi:10.1038/nrg3291

Appanah, S. (1985). General flowering in the climax rain forests of South-east Asia. *Journal of Tropical Ecology* 1(3), 225-240.

Appanah, S. (1993). Mass flowering of dipterocarp forests in the aseasonal tropics. *Journal of Biosciences* 18(4), 457-474. doi:Doi 10.1007/Bf02703079

Appanah, S., & Chan, H. T. (1981). Thrips: The pollinators of some dipterocarps. *Malaysian Forester* 44(2/3), 234-252.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Consortium, G. O. (2000). Gene Ontology: Tool for the unification of biology. *Nature Genetics* 25(1), 25-29. doi:Doi 10.1038/75556

Ashton, P. S. (1988). Dipterocarp biology as a window to the understanding of tropical forest structure. *Annual Review of Ecology and Systematics* 19**,** 347-370. doi: DOI 10.1146/annurev.es.19.110188.002023

Ashton, P. S., Givnish, T. J., & Appanah, S. (1988). Staggered flowering in the Dipterocarpaceae - New insights into floral induction and the evolution of mast fruiting in the aseasonal tropics. *American Naturalist* 132(1), 44-66. doi:Doi 10.1086/284837

Bhusal, N., Lee, M., Lee, H., Adhikari, A., Han, A. R., Han, A., & Kim, H. S. (2021). Evaluation of morphological, physiological, and biochemical traits for assessing drought resistance in eleven tree species. *Science of The Total Environment* 779, 146466.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Bouché, F., Lobet, G., Tocquin, P., & Périlleux, C. (2016). FLOR-ID: An interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Research* 44(D1), D1167-D1171.

Brearley, F. Q., Proctor, J., Suriantata, Nagy, L., Dalrymple, G., & Voysey, B. C. (2007). Reproductive phenology over a 10-year period in a lowland evergreen rain forest of central Borneo. *Journal of Ecology* 95(4), 828-839. doi:10.1111/j.1365-2745.2007.01258.x

Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., . . . Kong, L. (2021). KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research* 49(W1), W317-W325. doi:10.1093/nar/gkab447

Butt, N., Seabrook, L., Maron, M., Law, B. S., Dawson, T. P., Syktus, J., & McAlpine, C. A. (2015). Cascading effects of climate extremes on vertebrate fauna through changes to low-latitude tree flowering and fruiting phenology. *Global Change Biology* 21(9), 3267-3277.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST plus: Architecture and applications. *BMC Bioinformatics* 10. doi:Artn 42110.1186/1471-2105-10-421

Cao, D. N., Cheng, H., Wu, W., Soo, H. M., and Peng, J.R. (2006). Gibberellin mobilizes distinct DELLA-dependent transcriptomes to regulate seed germination and floral development in arabidopsis. *Plant Physiology* 142(2)**,** 509-525. doi: 10.1104/pp.106.082289.

Chen, Y. Y., Satake, A., Sun, I. F., Kosugi, Y., Tani, M., Numata, S., . . . Wright, S. J. (2018). Species-specific flowering cues among general flowering *Shorea* species at the Pasoh

Research Forest, Malaysia. *Journal of Ecology* 106(2), 586-598. doi:10.1111/1365-2745.12836

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2017). Araport11: A complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal* 89(4), 789-804. doi:10.1111/tpj.13415

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change* 3(1), 52-58.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., . . . Finn, R. D. (2018). The Pfam protein families database in 2019. *Nucleic Acids Research* 47(D1), D427-D432. doi:10.1093/nar/gky995

Eshghi, S., Tafazoli, E., Dokhani, S., Rahemi, M., & Emam, Y. (2007). Changes in carbohydrate contents in shoot tips, leaves and roots of strawberry (*Fragaria×ananassa* Duch.) during flower-bud differentiation. *Scientia Horticulturae* 113(3), 255-260.

Fu, L. M., Niu, B. F., Zhu, Z. W., Wu, S. T., & Li, W. Z. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23), 3150-3152. doi:10.1093/bioinformatics/bts565

Gould, P. D., Locke, J. C. W., Larue, C., Southern, M. M., Davis, S. J., Hanano, S., et al. (2006). The molecular basis of temperature compensation in the Arabidopsis circadian clock. *Plant Cell* 18(5)**,** 1177-1187. doi: 10.1105/tpc.105.039990.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., . . . Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7), 644-U130. doi:10.1038/nbt.1883

Gupta, A., Rico-Medina, A., & Caño-Delgado, A. I. (2020). The physiology of plant responses to drought. *Science* 368(6488), 266-269.

Harb, A., Krishnan, A., Ambavaram, M. M. R., and Pereira, A. (2010). Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiology* 154(3)**,** 1254-1271. doi: 10.1104/pp.110.161752.

Hart, S. N., Therneau, T.M., Zhang, Y., Poland, G. A., Kocher, J. P. (2013). Calculating sample size estimates for RNA sequencing data. *Journal of Computational Biology* 20(12), 970-978. doi: 10.1089/cmb.2012.0283.

Ho, L. C. (1988). Metabolism and compartmentation of imported sugars in sink organs in relation to sink strength. *Annual Review of Plant Physiology and Plant Molecular Biology* 39(1), 355-378.

Ichie, T., Igarashi, S., Yoshida, S., Kenzo, T., Masaki, T., & Tayasu, I. (2013). Are stored carbohydrates necessary for seed production in temperate deciduous trees? *Journal of Ecology* 101(2), 525-531.

Ichie, T., & Nakagawa, M. (2013). Dynamics of mineral nutrient storage for mast reproduction in the tropical emergent tree *Dryobalanops aromatica*. *Ecological Research 28*(2), 151-158. doi:10.1007/s11284-011-0836-1

Ickes, K. (2001) Hyper-abundance of native wild pigs (*Sus scrofa*) in a lowland dipterocarp rain forest of Peninsular Malaysia. *Biotropica* 33, 682–690.

IPCC. (2022). *Summary for Policymakers. Climate Change 2022: Impacts, Adaptation, and Vulnerability*. In (pp. 1-3676): Cambridge University Press. In Press Cambridge.

Janzen, D. H. (1967). Why mountain passes are higher in the tropics. *The American Naturalist* 101(919), 233-249.

Janzen, D. H. (1974). Tropical blackwater rivers, animals, and mast fruiting by the Dipterocarpaceae. *Biotropica***,** 69-103.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., . . . Nuka, G. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30(9), 1236-1240.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 28(1), 27-30.

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology* 428(4), 726-731.

Kelly, D., & Sork, V. L. (2002). Mast seeding in perennial plants: Why, how, where? *Annual Review of Ecology and Systematics* 33, 427-447. doi:10.1146/annurev.ecolsys.33.020602.095433

Kettle, C. J. (2010). Ecological considerations for using dipterocarps for restoration of lowland rainforest in Southeast Asia. *Biodiversity and Conservation* 19(4), 1137-1151. doi:10.1007/s10531-009-9772-6

Khan, F. S., Gan, Z.-M., Li, E.-Q., Ren, M.-K., Hu, C.-G., and Zhang, J.-Z. (2022). Transcriptomic and physiological analysis reveals interplay between salicylic acid and drought stress in citrus tree floral initiation. *Planta* 255(1)**,** 1-22.

Kobayashi, M. J., Takeuchi, Y., Kenta, T., Kume, T., Diway, B., & Shimizu, K. K. (2013). Mass flowering of the tropical tree *Shorea beccariana* was preceded by expression changes in flowering and drought-responsive genes. *Molecular Ecology* 22(18), 4767-4782. doi:10.1111/mec.12344

Kozlowski, T. (1992). Carbohydrate sources and sinks in woody plants. *The Botanical Review* 58(2), 107-222.

Li, J.-X., Hou, X.-J., Zhu, J., Zhou, J.-J., Huang, H.-B., Yue, J.-Q., et al. (2017). Identification of genes associated with lemon floral transition and flower development during floral inductive water deficits: a hypothetical model. *Frontiers in Plant Science* 8**,** 1013.

Lister, B. C., & Garcia, A. (2018). Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences* 115(44), E10397-E10406.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12). doi:ARTN 55010.1186/s13059-014-0550-8

637 Ma, S. S., & Bohnert, H. J. (2007). Integration of *Arabidopsis thaliana* stress-related transcript
638 profiles, promoter structures, and cell-specific expression. *Genome Biology* 8(4). doi:
639 ARTN R4910.1186/gb-2007-8-4-r49.

640 MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations.*
641 Paper presented at the Proceedings of the fifth Berkeley symposium on mathematical
642 statistics and probability.

643 Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO
644 update: Novel and streamlined workflows along with broader and deeper phylogenetic
645 coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology*
646 *and Evolution* 38(10), 4647-4654.

647 Marini, F., & Binder, H. (2019). pcaExplorer: An R/Bioconductor package for interacting with
648 RNA-seq principal components. *BMC Bioinformatics* 20(1), 1-8.

649 Misson, J., Raghothama, K. G., Jain, A., Jouhet, J., Block, M. A., Bligny, R., et al. (2005). A
650 genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips
651 determined plant responses to phosphate deprivation. *Proceedings of the National*
652 *Academy of Sciences of the United States of America* 102(33)**,** 11934-11939. doi:
653 10.1073/pnas.0505266102

654 Miura, S., Yasuda, M. & Ratnam, L. (1997) Who steals the fruits? Monitoring frugivory of
655 mammals in a tropical rain forest. *Malayan Nature Journal* 50, 183–193.

656 Nemhauser, J. L., Hong, F. X., and Chory, J. (2006). Different plant hormones regulate similar
657 processes through largely nonoverlapping transcriptional responses. *Cell* 126(3)**,** 467-
658 475. doi: 10.1016/j.cell.2006.05.050

659 Numata, S., Yasuda, M., Okuda, T., Kachi, N., & Noor, N. S. M. (2003). Temporal and spatial
660 patterns of mass flowerings on the Malay Peninsula. *American Journal of Botany* 90(7),
661 1025-1031. doi:DOI 10.3732/ajb.90.7.1025

662 Numata, S., Yamaguchi, K., Shimizu, M., Sakurai, G., Morimoto, A., Alias, N., . . . Satake, A.
663 (2022). Impacts of climate change on reproductive phenology in tropical rainforests of
664 Southeast Asia. *Communications Biology* 5(1), 1-10.

665 Okamoto, M., Peterson, F. C., Defries, A., Park, S.-Y., Endo, A., Nambara, E., . . . Cutler, S. R.
666 (2013). Activation of dimeric ABA receptors elicits guard cell closure, ABA-regulated
667 gene expression, and drought tolerance. *Proceedings of the National Academy of*
668 *Sciences* 110(29), 12132-12137.

669 Osuna, D., Usadel, B., Morcuende, R., Gibon, Y., Blasing, O. E., Hohne, M., . . . Stitt, M.
670 (2007). Temporal responses of transcripts, enzyme activities and metabolites after adding
671 sucrose to carbon-deprived Arabidopsis seedlings. *Plant Journal* 49(3), 463-491.
672 doi:10.1111/j.1365-313X.2006.02979.x

673 Parcy, F. (2005). Flowering: a time for integration. *International Journal of Developmental*
674 *Biology* 49(5-6), 585-593. doi:10.1387/ijdb.041930fp

675 Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2016). Salmon provides
676 accurate, fast, and bias-aware transcript expression estimates using dual-phase inference.
677 *BioRxiv*, 021592.

678 Peng, S.-A., & Iwahori, S. (1994). Morphological and cytological changes in apical meristem
679 during flower bud differentiation of Japanese pear, *Pyrus pyrifolia* Nakai. *Journal of the*
680 *Japanese Society for Horticultural Science* 63(2), 313-321.

681 Peñuelas, J., Sardans, J., Estiarte, M., Ogaya, R., Carnicer, J., Coll, M., . . . Garbulsky, M.
682 (2013). Evidence of current impact of climate change on life: A walk from genes to the
683 biosphere. *Global Change Biology* 19(8), 2303-2338.

684 Pingping, W., Chubin, W., & Biyan, Z. (2017). Drought stress induces flowering and enhances
685 carbohydrate accumulation in *Averrhoa carambola*. *Horticultural Plant Journal* 3(2), 60-
686 66.

687 Rau, A., & Maugis-Rabusseau, C. (2018). Transformation and model choice for RNA-seq co-
688 expression analysis. *Briefings in Bioinformatics* 19(3), 425-436. doi:10.1093/bib/bbw128

689 Riboni, M., Robustelli Test, A., Galbiati, M., Tonelli, C., and Conti, L. (2016). ABA-dependent
690 control of GIGANTEA signalling enables drought escape via up-regulation of
691 *FLOWERING LOCUS T* in *Arabidopsis thaliana*. *Journal of Experimental Botany*
692 67(22), 6309-6322.

693 Richards, C. L., Hanzawa, Y., Katari, M. S., Ehrenreich, I. M., Engelmann, K. E., &
694 Purugganan, M. D. (2009). Perspectives on ecological and evolutionary systems biology.
695 In: *Annual Plant Reviews* (eds Coruzzi, G. M. & Gutiérrez, R. A.), pp. 331-351. Wiley-
696 Blackwell, Oxford.

697 Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential
698 expression analysis of RNA-seq data. *Genome Biology* 11(3). doi:ARTN R2510.1186/gb-
699 2010-11-3-r25

700 Santer, B. D., Po-Chedley, S., Zelinka, M. D., Cvijanovic, I., Bonfils, C., Durack, P. J., . . .
701 Painter, J. (2018). Human influence on the seasonal cycle of tropospheric temperature.
702 *Science* 361(6399), eaas8806.

703 Sakai, S., Harrison, R. D., Momose, K., Kuraji, K., Nagamasu, H., Yasunari, T., . . .
704 Nakashizuka, T. (2006). Irregular droughts trigger mass flowering in aseasonal tropical
705 forests in Asia. *American Journal of Botany* 93(8), 1134-1139. doi:DOI
706 10.3732/ajb.93.8.1134

707 Sakai, S., Momose, K., Yumoto, T., Nagamitsu, T., Nagamasu, H., Hamid, A. A., Nakashizuka,
708 T. (1999). Plant reproductive phenology over four years including an episode of general
709 flowering in a lowland dipterocarp forest, Sarawak, Malaysia. *American Journal of*
710 *Botany* 86(10), 1414-1436. doi: Doi 10.2307/2656924.

711 Satake, A., Nagahama, A., & Sasaki, E. (2022). A cross-scale approach to unravel the molecular
712 basis of plant phenology in temperate and tropical climates. *New Phytologist* 233(6),
713 2340-2353.

714 Scheible, W. R., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D., Palacios-Rojas, N., et al.
715 (2004). Genome-wide reprogramming of primary and secondary metabolism, protein
716 synthesis, cellular growth processes, and the regulatory infrastructure of Arabidopsis in
717 response to nitrogen. *Plant Physiology* 136(1)**,** 2483-2499. doi: 10.1104/pp.104.047019.

718 Shi, H., Chen, L., Ye, T., Liu, X., Ding, K., & Chan, Z. (2014). Modulation of auxin content in
719 Arabidopsis confers improved drought stress resistance. *Plant Physiology and*
720 *Biochemistry* 82, 209-217.

721 Singh, D., & Laxmi, A. (2015). Transcriptional regulation of drought response: A tortuous
722 network of transcriptional factors. *Frontiers in Plant Science* 6. doi:ARTN
723 89510.3389/fpls.2015.00895

724 Singh, K., & Kushwaha, C. (2016). Deciduousness in tropical trees and its potential as indicator
725 of climate change: A review. *Ecological Indicators* 69, 699-706.

726 Smaczniak, C., Muiño, J. M., Chen, D., Angenent, G. C., & Kaufmann, K. (2017). Differences in
727 DNA binding specificity of floral homeotic protein complexes predict organ-specific
728 target genes. *The Plant Cell* 29(8), 1822-1835.

729 Suhaimi, A. H., Kobayashi, M. J., Satake, A., Lee, S. L., Muhammad, N., Otani, T., … Yeoh, S.
730 H. (2023). Characterization of leaf transcriptome in a tropical tree species, *Shorea curtisii*
731 over a flowering season. *Japan Agricultural Research Quarterly* 57(2).

732 Tani, N., Tsumura, Y., Fukasawa, K., Kado, T., Taguchi, Y., Lee, S. L., . . . Kassim, A. R.
733 (2015). Mixed mating system are regulated by fecundity in *Shorea curtisii*
734 (Dipterocarpaceae) as revealed by comparison under different pollen limited conditions.
735 *Plos One* 10(5). doi:ARTN e012344510.1371/journal.pone.0123445

736 Team, R. C. (2018). R: A language and environment for statistical computing: R Foundation for
737 Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

738 Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., . . .
739 Narechania, A. (2003). PANTHER: A library of protein families and subfamilies indexed
740 by function. *Genome Research* 13(9), 2129-2141.

741 Trenberth, K. E., Dai, A., Van Der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., &
742 Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*
743 4(1), 17-22.

744 Trevaskis, B., Hemming, M. N., Dennis, E. S., & Peacock, W. J. (2007). The molecular basis of
745 vernalization-induced flowering in cereals. *Trends in Plant Science* 12(8), 352-357.

746 Ushio, M., Osada, Y., Kumagai, T., Kume, T., Pungga, R. A. S., Nakashizuka, T., . . . Sakai, S.
747 (2020). Dynamic and synergistic influences of air temperature and rainfall on general
748 flowering in a Bornean lowland tropical forest. *Ecological Research* 35(1), 17-29.

749 Wang, X., Li, Q., Xie, J., Huang, M., Cai, J., Zhou, Q., . . . Jiang, D. (2021). Abscisic acid and
750 jasmonic acid are involved in drought priming-induced tolerance to drought in wheat.
751 *The Crop Journal* 9(1), 120-132.

752 Wieczorek, K., Golecki, B., Gerdes, L., Heinen, P., Szakasits, D., Durachko, D. M., . . .
753 Bohlmann, H. (2006). Expansins are involved in the formation of nematode-induced
754 syncytia in roots of *Arabidopsis thaliana*. *The Plant Journal* 48(1), 98-112.

755 Wu, H., Wu, X., Li, Z., Duan, L., & Zhang, M. (2012). Physiological evaluation of drought
756 stress tolerance and recovery in cauliflower (*Brassica oleracea* L.) seedlings treated with
757 methyl jasmonate and coronatine. *Journal of Plant Growth Regulation* 31(1), 113-123.

Yagihashi, T., Otani, T., Tani, N., Nakaya, T., Rahman, K. A., Matsui, T., & Tanouchi, H. (2010). Habitats suitable for the establishment of *Shorea curtisii* seedlings in a hill forest in Peninsular Malaysia. *Journal of Tropical Ecology* 26, 551-554. doi:10.1017/S026646741000026x

Yang, J., Duan, G., Li, C., Liu, L., Han, G., Zhang, Y., & Wang, C. (2019). The crosstalks between jasmonic acid and other plant hormone signaling highlight the involvement of jasmonic acid as a core component in plant response to biotic and abiotic stresses. *Frontiers in Plant Science* 10, 1349.

Yasuda, M., Matsumoto, J., Osada, N., Ichikawa, S., Kachi, N., Tani, M., . . . Manokaran, N. (1999). The mechanism of general flowering in Dipterocarpaceae in the Malay Peninsula. *Journal of Tropical Ecology* 15, 437-449. doi:Doi 10.1017/S0266467499000930

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., . . . Zhou, A. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research* 46(W1), W71-W75.

Yeoh, S. H., Satake, A., Numata, S., Ichie, T., Lee, S. L., Basherudin, N., . . . Tani, N. (2017). Unravelling proximate cues of mass flowering in the tropical forests of South-East Asia from gene expression analyses. *Molecular Ecology* 26(19), 5074-5085. doi:10.1111/mec.14257

Zheng, Z., Qamar, S. A., Chen, Z., & Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *The Plant Journal* 48(4), 592-605.

## Table 1(on next page)

The number of unigenes in the leaf and bud transcriptome of *Shorea curtisii* annotated using public protein databases.

**Table 1:**
**The number of unigenes in the leaf and bud transcriptome of *Shorea curtisii* annotated using public protein databases.**

| Database | No. of unigenes | Percentage (%)* |
|---|---|---|
| Araport | 51,654 | 80.43 |
| UniProt | 50,381 | 78.45 |
| Pfam | 47,610 | 74.14 |
| PANTHER | 55,929 | 87.09 |
| GO | 51,503 | 80.20 |
| KEGG | 25,589 | 39.85 |
| Annotated in at least one database | 58,722 | 91.44 |
| Annotated in all the databases | 22,508 | 35.05 |

* Percentage of annotated unigenes over total unigenes.

## Table 2(on next page)

Summary of *Arabidopsis thaliana* flowering-time homologs and differentially expressed unigenes (DEUs) identified in the leaf and bud transcriptomes of *Shorea curtisii*.

**Table 2:**
**Summary of *Arabidopsis thaliana* flowering-time homologs and differentially expressed unigenes (DEUs) identified in the leaf and bud transcriptomes of *Shorea curtisii*.**

| Floral regulatory pathway | No. of *A. thaliana* genes | No. of *S. curtisii* homologs | No. of DEUs | |
|---|---|---|---|---|
| | | | Leaf | Bud |
| General | 117 | 402 (100) | 2 (2) | 22 (12) |
| Photoperiod and circadian clock | 103 | 274 (70) | 15 (9) | 37 (20) |
| Hormones | 28 | 62 (21) | 3 (3) | 5 (4) |
| Vernalization | 28 | 99 (21) | 2 (1) | 10 (4) |
| Aging | 22 | 23 (6) | 0 (0) | 8 (4) |
| Floral meristem identity | 9 | 26 (6) | 0 (0) | 7 (4) |
| Sugar | 9 | 46 (9) | 2 (1) | 9 (3) |
| Integrator | 8 | 41 (7) | 2 (2) | 12 (4) |
| Ambient temperature | 7 | 16 (5) | 2 (1) | 2 (1) |
| Autonomous | 7 | 34 (7) | 0 (0) | 2 (1) |

Number in parentheses indicates the number of non-redundant *A. thaliana* genes that corresponds to *S. curtisii* homologs.

## Table 3(on next page)

Overrepresentation of *Arabidopsis thaliana* gene sets in the differentially expressed unigenes of *Shorea curtisii*.

**Table 3:**
**Overrepresentation of *Arabidopsis thaliana* gene sets in the differentially expressed unigenes of *Shorea curtisii*.**

| No. | Description of *A. thaliana* gene set | Leaf | Bud |
|---|---|---|---|
| 1. | Upregulated under severe drought | 4.2E–11 | NS |
| 2. | Downregulated under severe drought | 1.3E–05 | 7.7E–15 |
| 3. | Upregulated under prolonged moderate drought | 8.0E–06 | 7.7E–15 |
| 4. | Downregulated under prolonged moderate drought | 7.7E–15 | 7.7E–15 |
| 5. | Commonly upregulated under 41 abiotic and biotic stress conditions | 7.7E–15 | 7.7E–15 |
| 6. | Upregulated with temperature increase | 3.3E–12 | 1.3E–06 |
| 7. | Downregulated with temperature increase | 7.7E–15 | 7.7E–15 |
| 8. | Upregulated with temperature decrease | NS | NS |
| 9. | Downregulated with temperature decrease | 1.8E–04 | NS |
| 10. | Upregulated with abscisic acid treatment | 7.7E–15 | 7.7E–15 |
| 11. | Downregulated with abscisic acid treatment | NS | 1.1E–04 |
| 12. | Upregulated with brassinosteroid treatment | 2.3E–12 | 7.7E–15 |
| 13. | Downregulated with brassinosteroid treatment | NS | NS |
| 14. | Upregulated with cytokinin treatment | NS | NS |
| 15. | Downregulated with cytokinin treatment | 1.3E–08 | 5.0E–07 |
| 16. | Upregulated with ethylene treatment | NS | 5.3E–06 |
| 17. | Downregulated with ethylene treatment | 3.6E–06 | 7.7E–15 |
| 18. | Upregulated with indole-3-acetic treatment | 7.7E–15 | 7.7E–15 |
| 19. | Downregulated with indole-3-acetic treatment | NS | 7.4E–06 |
| 20. | Upregulated with jasmonic acid treatment | 7.7E–15 | 7.7E–15 |
| 21. | Downregulated with jasmonic acid treatment | 1.1E–12 | 7.7E–15 |
| 22. | Upregulated with gibberellin in young flower buds | 7.7E–15 | 7.7E–15 |
| 23. | Downregulated with gibberellin in young flower buds | 1.2E–11 | 7.7E–15 |
| 24. | Upregulated under carbon-limited condition | 7.7E–15 | 2.5E–09 |
| **No.** | **Description of *A. thaliana* gene set** | **Leaf** | **Bud** |

| | | | |
|---|---|---|---|
| 25. | Downregulated under carbon-limited condition | 7.7E–15 | 7.7E–15 |
| 26. | Upregulated after the addition of sucrose under carbon-limited condition | 7.7E–15 | 7.7E–15 |
| 27. | Downregulated after the addition of sucrose under carbon-limited condition | 1.7E–11 | NS |
| 28. | Upregulated under long-term phosphate-limited condition | 7.7E–15 | 7.7E–15 |
| 29. | Downregulated under long-term phosphate-limited condition | NS | NS |
| 30. | Upregulated under medium-term phosphate-limited condition | 1.6E–08 | 6.2E–04 |
| 31. | Downregulated under medium-term phosphate-limited condition | NS | NS |
| 32. | Upregulated under shot-term phosphate-limited condition | 2.0E–10 | NS |
| 33. | Downregulated genes under short-term phosphate-limited condition | NS | NS |
| 34. | Upregulated genes under nitrogen-limited condition | 1.1E–09 | 1.4E–08 |
| 35. | Downregulated genes under nitrogen-limited condition | NS | 2.5E–08 |

Fisher's exact test *P*-value cut-off: 1E–03 after Bonferroni correction was used.
NS indicates not significant.

## Table 4(on next page)

Clusters of differentially expressed unigenes in *Shorea curtisii* leaf with significant enrichment of *Arabidopsis thaliana* gene sets.

**Table 4:**
**Clusters of differentially expressed unigenes in *Shorea curtisii* leaf with significant enrichment of *Arabidopsis thaliana* gene sets.**

| Description of *Arabidopsis thaliana* gene set | Cluster | | | |
|---|---|---|---|---|
| | 1 | 3 | 4 | 6 |
| Upregulated under severe drought condition | NS | 4.2E–09 | NS | NS |
| Downregulated under severe drought condition | NS | NS | NS | 1.3E–15 |
| Downregulated under prolonged moderate drought condition | NS | 8.1E–12 | NS | NS |
| Commonly upregulated under 41 abiotic and biotic stress conditions | NS | 1.3E–15 | NS | NS |
| Upregulated with abscisic acid treatment | NS | 2.3E–07 | NS | NS |
| Upregulated with indole-3-acetic acid treatment | NS | 1.8E–09 | NS | NS |
| Upregulated with jasmonic acid treatment | NS | 6.5E–09 | NS | NS |
| Upregulated with gibberellin treatment | NS | NS | NS | 1.6E–07 |
| Upregulated under carbon-limited condition | NS | NS | 4.5E–13 | NS |
| Downregulated under carbon-limited condition | NS | NS | NS | 5.7E–09 |
| Downregulated after the addition of sucrose | NS | NS | 2.1E–14 | NS |
| Upregulated under medium-term phosphate-limited condition | 1.1E–04 | NS | NS | NS |

Fisher's exact test *P*-value cut-off: 1E–03 after Bonferroni correction.
NS indicates not significant.

## Table 5(on next page)

Clusters of differentially expressed unigenes in *Shorea curtisii* bud with significant enrichment of *Arabidopsis thaliana* gene sets.

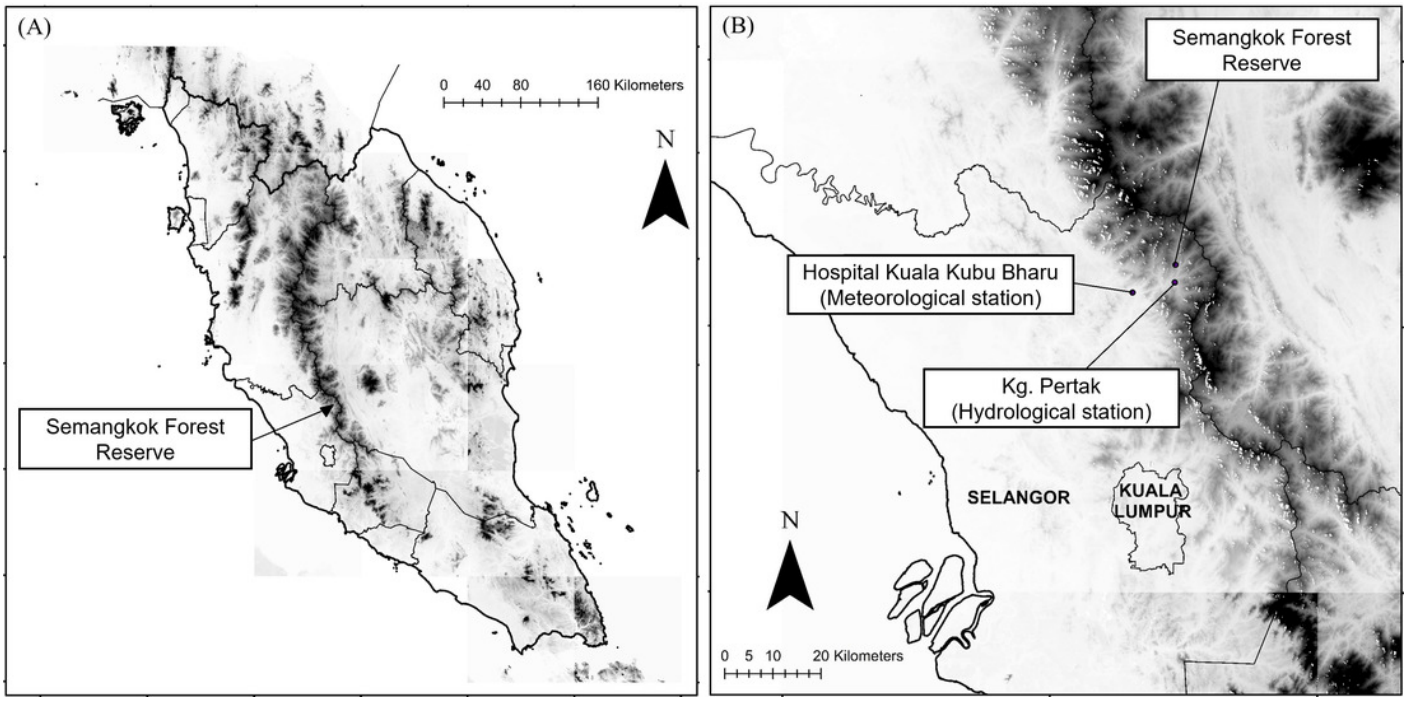
**Table 5:**
**Clusters of differentially expressed unigenes in *Shorea curtisii* bud with significant enrichment of *Arabidopsis thaliana* gene sets.**

| Description of *Arabidopsis thaliana* gene set | Cluster | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Downregulated under severe drought condition | NS | NS | NS | 1.1E–15 | NS |
| Upregulated under prolonged moderate drought condition | NS | NS | NS | NS | 3.8E–05 |
| Downregulated under prolonged moderate drought condition | NS | 1.2E–06 | NS | NS | NS |
| Commonly upregulated under 41 abiotic and biotic stress conditions | NS | 1.3E–15 | 5.2E–05 | NS | NS |
| Downregulated with increase in temperature | NS | NS | NS | NS | 8.2E–05 |
| Downregulated with cytokinin treatment | NS | NS | 2.8E–04 | NS | NS |
| Downregulated with ethylene treatment | NS | NS | NS | 4.1E–08 | NS |
| Upregulated with indole-3-acetic acid treatment | NS | 1.7E–05 | NS | NS | NS |
| Upregulated with jasmonic acid treatment | NS | 8.8E–10 | 3.4E–04 | NS | NS |
| Upregulated with gibberellin treatment | NS | NS | NS | 3.1E–04 | NS |
| Downregulated with gibberellin treatment | NS | 5.2E–07 | NS | NS | NS |
| Upregulated under carbon-limited condition | 9.7E–04 | NS | NS | NS | NS |
| Downregulated under carbon-limited condition | NS | NS | NS | 1.4E–12 | NS |

Fisher's exact test *P*-value cut-off: 1E–03 after Bonferroni correction.
NS indicates not significant.

# Figure 1

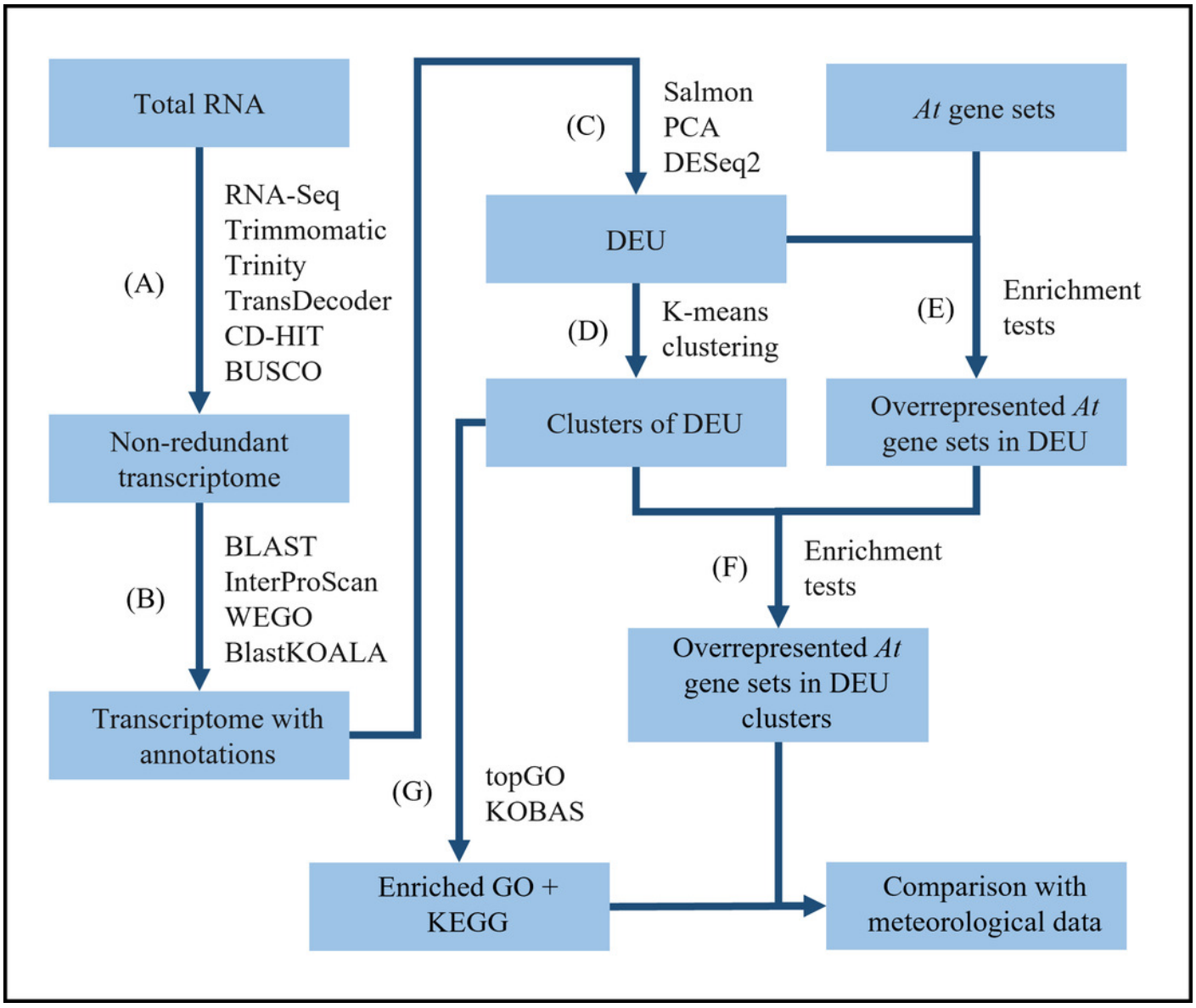
Map of (A) sample collection site (Semangkok Forest Reserve) and (B) the meteorological and hydrological stations nearest to the study site.



# Figure 2

The research strategy employed to study the leaf and the bud transcriptome of *Shorea curtisii* during a flowering season.

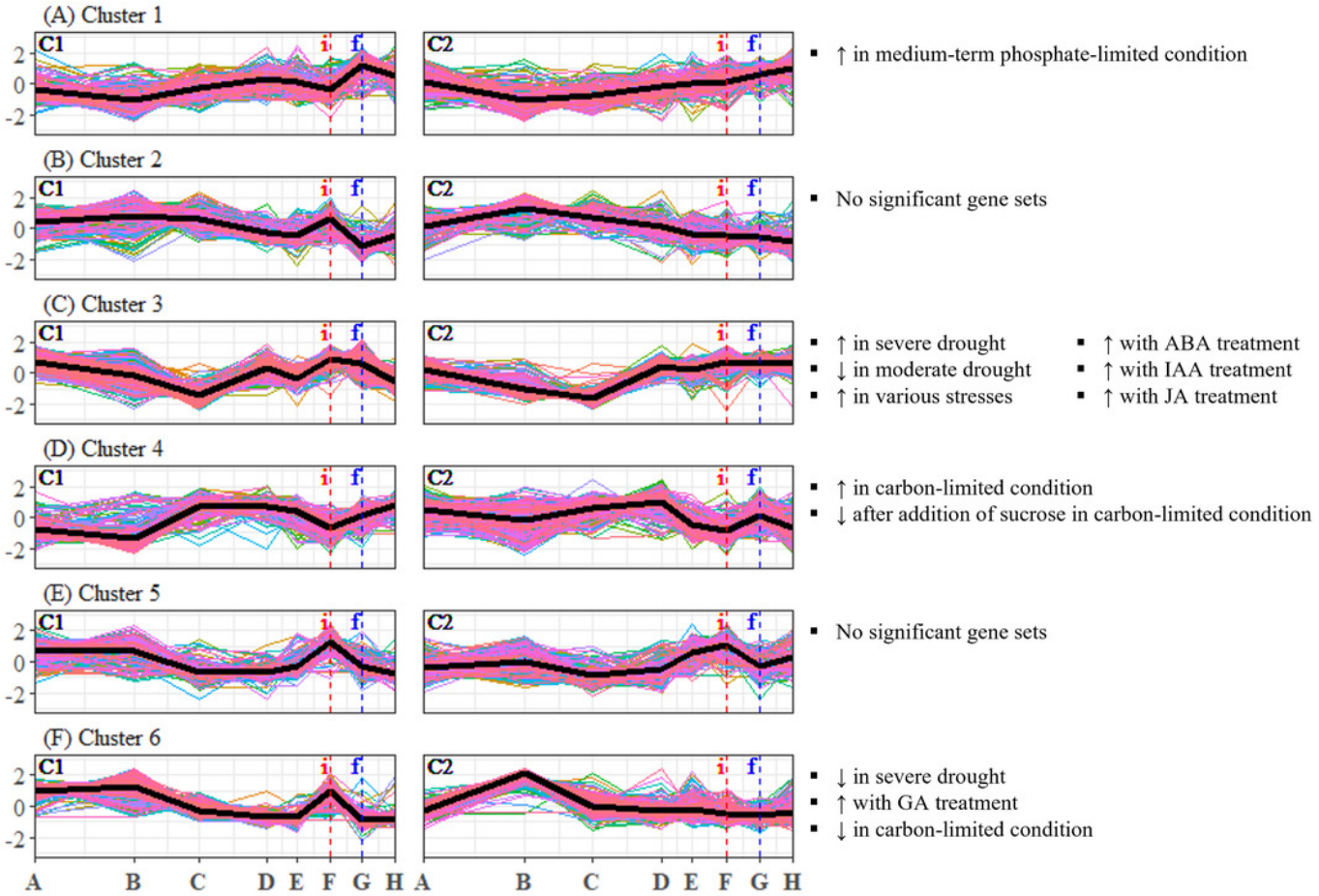
We pooled and analyzed all the samples together until transcript quantifications (A–C) and performed the rest of the bioinformatic pipeline (D–G) on the leaf and bud samples separately. For more detailed description of the analyses, refer to the main text. **(A)** Removal of low quality bases and traces of adapters prior to *de novo* assembly. Coding regions within each transcript were predicted and only transcripts with complete open reading frame were retained. Highly similar sequences were clustered to reduce the number of transcripts. The completeness of the assembly was also assessed. **(B)** Annotation of the transcriptome using publicly available protein databases. **(C)** Transcript quantification was performed prior to principal component analysis. Differential expression analysis was performed to identify differentially expressed unigenes. Samples at every TP were compared in a pairwise manner. **(D)** Analysis of the expression profiles of the DEUs. The DEUs were clustered based on their expression patterns. **(E)** Identification of *A. thaliana* gene sets that were overrepresented in the DEUs. Gene enrichment tests were performed between each *A. thaliana* gene set and the DEUs in *S. curtisii.* **(F)** Identification of enriched *A. thaliana* gene sets from the previous step that show specific expression pattern. Gene enrichment tests were conducted between each significantly enriched *A. thaliana* gene set and all DEU clusters to identify *A. thaliana* gene sets that were enriched in any specific DEU cluster. **(G)** Identification of GO terms and KEGG pathways that were enriched in the DEUs. **(G)** Comparison between the results of the enrichment analyses with the meteorological data (temperature records and cumulative rainfall data).


Total RNA
(A)
RNA-Seq
Trimmomatic
Trinity
TransDecoder
CD-HIT
BUSCO
Non-redundant transcriptome
(B)
BLAST
InterProScan
WEGO
BlastKOALA
Transcriptome with annotations
(C)
Salmon
PCA
DESeq2
At gene sets
DEU
(D)
K-means clustering
(E)
Enrichment tests
Clusters of DEU
Overrepresented At gene sets in DEU
(F)
Enrichment tests
Overrepresented At gene sets in DEU clusters
(G)
topGO
KOBAS
Enriched GO + KEGG
Comparison with meteorological data

# Figure 3

The expression profiles of differentially expressed unigene clusters in *S. curtisii* leaf and the corresponding significantly enriched *A. thaliana* gene sets.

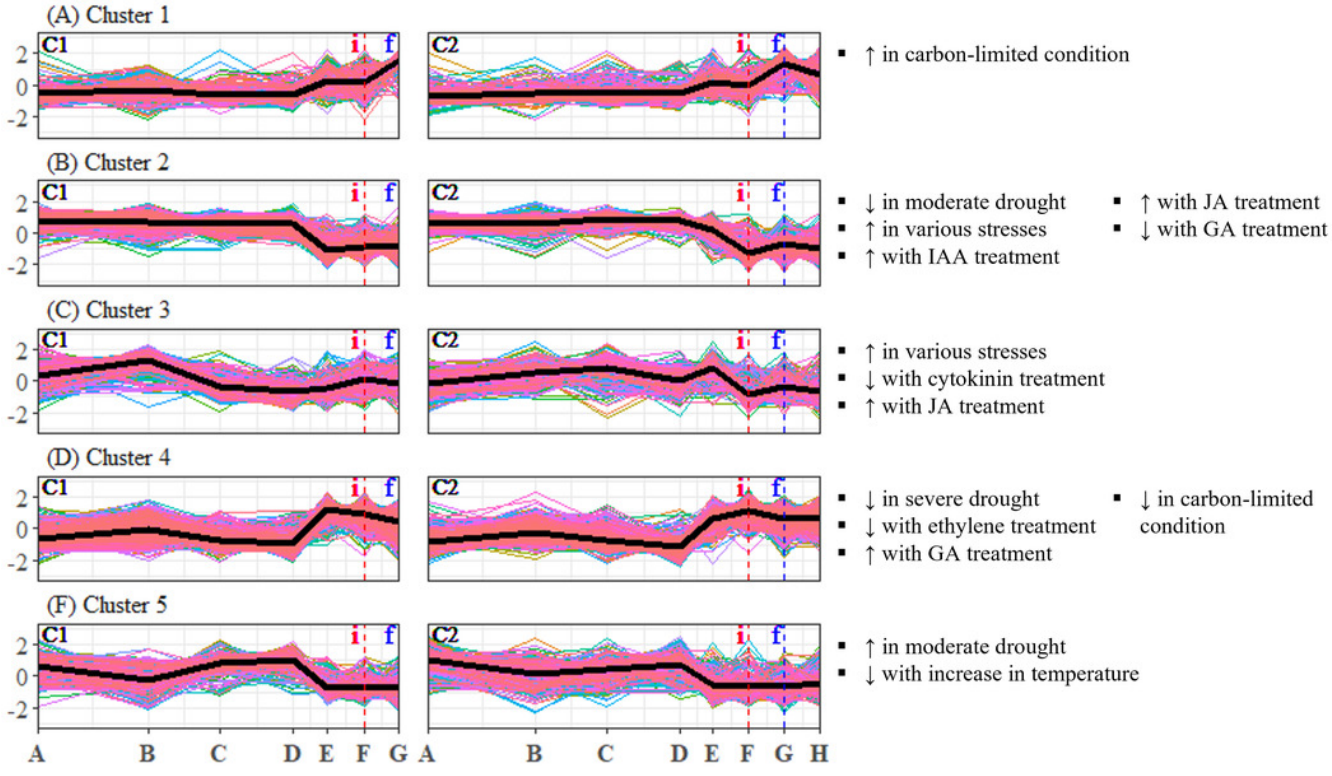
The expression of unigenes (*y*-axis) was scaled to a mean of 0 and a standard deviation of 1 across the samples. The time points are shown on *x*-axis. Black lines represent the mean expression pattern of the respective cluster. Significantly enriched *A. thaliana* gene sets in each cluster are shown on the right-hand side of each profile (see Supplementary Table S20 for details). i: Inflorescence stage; f: Flowering stage; C1: Individual 1, C2: Individual 2, ↑: Upregulated genes, ↓: Downregulated genes.



# Figure 4

The expression profiles of differentially expressed unigene clusters in *Shorea curtisii* bud and the corresponding significantly enriched *Arabidopsis thaliana* gene sets.

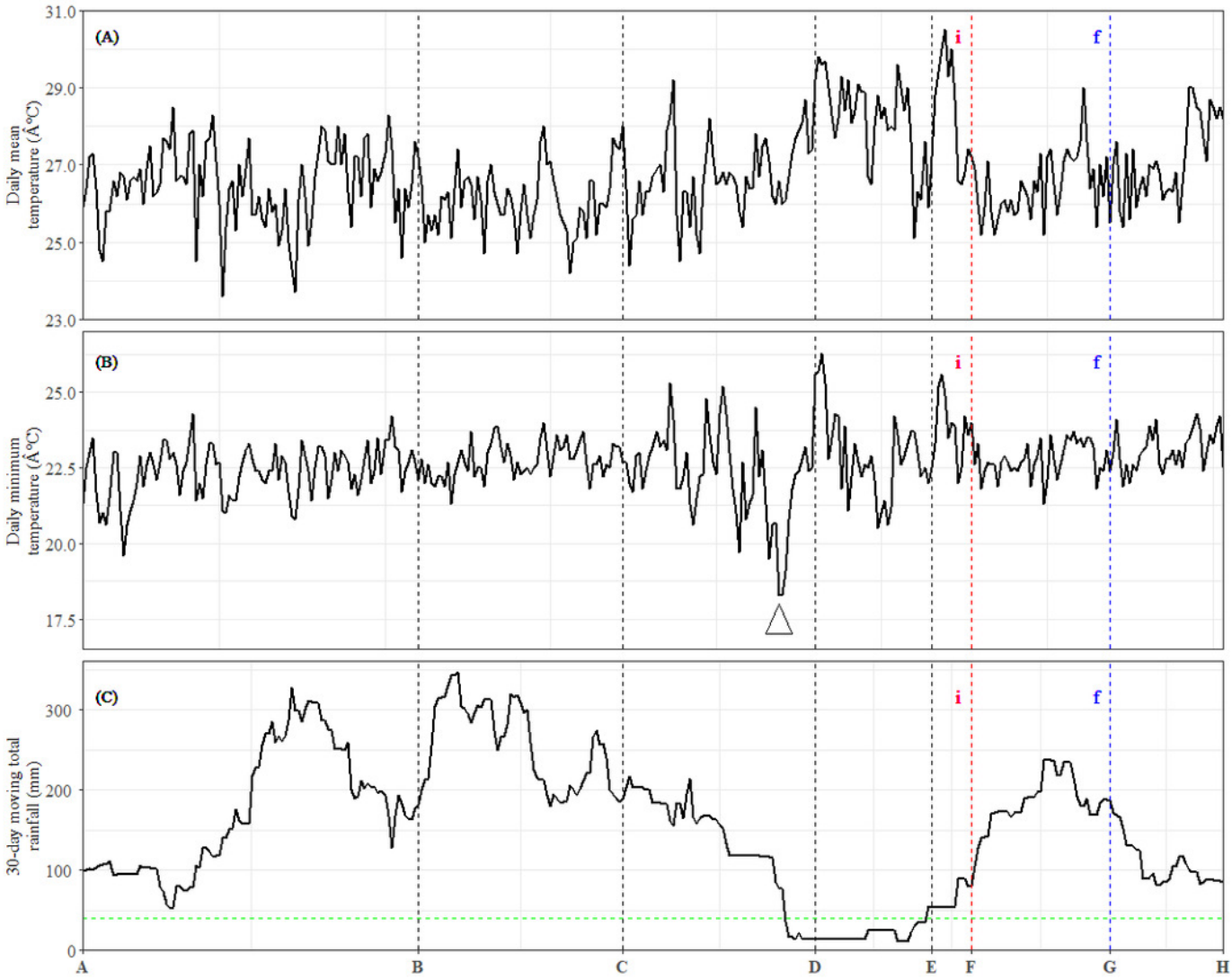
The expression of unigenes (*y*-axis) was scaled to a mean of 0 and a standard deviation of 1 across the samples. The time points are shown on *x*-axis. Black lines represent the mean expression pattern of the respective cluster. Significantly enriched *A. thaliana* gene sets in each cluster are shown on the right-hand side of each profile (see Supplementary Table S20 for details). i: Inflorescence stage; f: Flowering stage; C1: Individual 1, C2: Individual 2, ↑: Upregulated genes, ↓: Downregulated genes.



# Figure 5

Meteorological data and occurrences of general flowering. Data from July 2013 to June 2014. (A) Daily mean temperature. (B) Daily minimum temperature.

The arrow indicates the sharpest drop in daily minimum temperature within the study period. **(C)** Total rainfall over the preceding 30-day period. The dashed horizontal line indicates 40-mm drought threshold. Dashed vertical lines indicate time points (*x*-axis). i: Inflorescence stage; f: Flowering stage.



# Figure 6

Dynamics of phytohormones, carbohydrate resources, and *FT* expression from signal accumulation to onset of flowering in *Shorea curtisii*.

Flowering signal accumulation e.g., drought signal begins 2–3 months prior to floral initiation. In response to the stress signal, levels of phytohormones such as IAA, ABA, and JA are elevated. During this period, sucrose level in leaf increases whilst carbon concentration in bud decreases, presumably in preparation for floral initiation. Expression of *FT* in leaf gradually increases and peaks at floral initiation which is followed by appearance of inflorescence buds. Expression of *FT* in bud remains high towards flowering time. Carbohydrate content in bud decreases as the plants readying itself for anthesis. I: Inflorescence/floral initiation; F: Flowering; ↑: Upregulation, ↓: Downregulation; IAA: Auxin; ABA: Abscisic acid; JA: Jasmonic acid; GA: Gibberellic acid.