Comparative transcriptomic analysis of heterophylly of the aquatic plant *Potamogeton octandrus* (Potamogetonaceae)

Dingxuan He [1], **Pin Guo** [2], **Paul F Gugger** [3], **Youhao Guo** [1], **Xing Liu** Corresp., [1], **Jinming Chen** Corresp. [4]

[1] College of life Sciences, Wuhan University, Wuhan, Hubei, China

[2] Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan

[3] Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, Maryland, USA

[4] Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

Corresponding Authors: Xing Liu, Jinming Chen
Email address: xingliu@whu.edu.cn, jmchen@wbgcas.cn

Many plant species exhibit heterophylly, displaying different leaves upon a single plant. The molecular mechanisms regulating this phenomenon, however, have remained elusive. In this study, the transcriptomes of submerged and floating leaves of an aquatic heterophyllous plant, *Potamogeton octandrus* Poir, were sequenced using a high-throughput sequencing technique (RNA-Seq), which aims to assist with the gene discovery and functional studies of genes involved in heterophyllous leaf development. A total of 81,103 unigenes were identified from the submerged and floating leaves, and a total of 6,822 differentially expressed genes (DEGs) were identified by comparing the samples from each developmental stage. KEGG pathway enrichment analysis categorized these unigenes into 128 pathways (p-value < 10^{-5}). A total of 24,025 differentially expressed genes were involved in the carbon metabolic pathway, biosynthesis of amino acids, ribosomes, and plant-pathogen interaction. KEGG pathway enrichment analysis categorized a total of 70 DEGs into plant hormone signal transduction pathways. This study describes the initial results of the high-throughput transcriptome sequencing of heterophylly. Understanding the transcriptomes of floating and submerged leaves of the aquatic plant *P. octandrus* will assist with gene cloning and functional studies of genes involved in leaf development. This is especially the case with those involved in heterophyllous leaf development.

Comparative transcriptomic analysis of heterophylly of the aquatic plant *Potamogeton octandrus* (Potamogetonaceae)

Dingxuan He[1], Pin Guo[2], Paul F. Gugger[3], Youhao Guo[1], Xing Liu[1*], Jinming Chen[4*]

1 Laboratory of Plant Systematics and Evolutionary Biology, College of Life Science, Wuhan University, Wuhan 430072, China

2 Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan

3 University of Maryland Center for Environmental Science, Appalachian Laboratory, 301 Braddock Road, Frostburg, Maryland 21532, USA

4 Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China

*** Corresponding authors:**

Xing Liu:

Laboratory of Plant Systematics and Evolutionary Biology, College of Life Science, Wuhan University, Wuhan 430072, China;

Tel.: +86-27-68756612; Fax: +86-27-68752560; E-mail: xingliu@whu.edu.cn

Jin-Ming Chen:

Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China;

Tel: 86-27-87617211; Fax: 86-27-87510526; E-mail: jmchen@wbgcas.cn

ABSTRACT

Many plant species exhibit heterophylly, displaying different leaves upon a single plant. The molecular mechanisms regulating this phenomenon, however, have remained elusive. In this study, the transcriptomes of submerged and floating leaves of an aquatic heterophyllous plant, *Potamogeton octandrus* Poir, were sequenced using a high-throughput sequencing technique (RNA-Seq), which aims to assist with the gene discovery and functional studies of genes involved in heterophyllous leaf development. A total of 81,103 unigenes were identified from the submerged and floating leaves, and a total of 6,822 differentially expressed genes (DEGs) were identified by comparing the samples from each developmental stage. KEGG pathway enrichment analysis categorized these unigenes into 128 pathways (p-value < $^{10-5}$). A total of 24,025 differentially expressed genes were involved in the carbon metabolic pathway, biosynthesis of amino acids, ribosomes, and plant-pathogen interaction. KEGG pathway enrichment analysis categorized a total of 70 DEGs into plant hormone signal transduction pathways. This study describes the initial results of the high-throughput transcriptome sequencing of heterophylly. Understanding the transcriptomes of floating and submerged leaves of the aquatic plant *P. octandrus* will assist with gene cloning and functional studies of genes involved in leaf development. This is especially the case with those involved in heterophyllous leaf development.

Keywords: Gene expression, Heterophyllous leaves, *Potamogeton octandrus*, Transcriptome

INTRODUCTION

Heterophylly, displaying two or more leaf forms upon a single plant, is widely observed across diverse plant species, most notably in aquatic plants (Minorsky, 2003; Zotz, Wilhelm & Becker, 2011). In some cases, heterophylly is believed to be an adaptive response to the environment, and it has been linked to an increase in fitness (Cook & Johnson, 1968; Wells & Pigliucci, 2000; Minorsky, 2003). For example, the heterophylly of aquatic plants may increase their fitness by decreasing leaf damage, decreasing water loss, enhancing photosynthesis, or promoting sexual reproductive success (Winn, 1999a,b; Wells & Pigliucci, 2000; Minorsky, 2003; Zhang et al., 2009; Zotz, Wilhelm & Becker, 2011). Accordingly, heterophylly has been used as a model system for studying gene-environment interactions (Pigliucci, 2010; Nakayama et al., 2014).

In the past century, numerous studies have been conducted to describe morphological changes in heterophyllous plants in response to environmental factors such as CO_2 concentration, oxygen capacity, salt concentration, temperature, water level, seasonal change, and light intensity and quality (McCallum, 1902; Arber, 1920; Fassett, 1930; Sculthorpe, 1967; Cook & Johnson, 1968; Bodkin, Spence & Weeks, 1980; Deschamp & Cooke, 1984; Titus & Sullivan, 2001). In addition to such environmental factors, more recent studies have revealed that plant hormones, including ethylene, abscisic acid (ABA), and gibberellin (GA), could affect heterophyllous leaf formation in many plant species, such as *Potamogeton nodosus* (Anderson, 1978), *Hippuris vulgaris* (Kane & Albert, 1987), *Marsilea quadrifolia* (Liu, 1984), *Callitriche heterophylla* (Deschamp & Cooke, 1985), *Ranunculus flaellaris* (Young & Horton, 1985; Young, Dengler & Horton, 1987), *Ludwigia arcuata* (Kuwabara, 2003), and *Rorippa aquatica* (Nakayama et al., 2014). Studies with *L. arcuate*, for example, have suggested that ethylene gas induces the development of submerged-type leaves on terrestrial shoots. ABA, however, resulted in the

formation of terrestrial-type leaves on submerged shoots (Kuwabara, 2003). However, the molecular mechanisms regulating these alterations in leaf form in heterophyllous plant species remain largely unclear.

Plant leaves come from a group of initial cells that are called shoot apical meristems (SAMs). In recent years, the molecular mechanisms responsible for the initiation and maintenance of the SAM and polar processes of leaf expansion have been studied in detail (Hay & Tsiantis, 2006; Tsukaya, 2006; Uchida et al., 2007, 2010; Shani et al., 2010; Moon & Hake, 2011). For example, the SAM is characterized by expression of the Class I *KNOTTED1-LIKE HOMEOBOS* (*KNOX*) gene (Smith et al., 1992; Jackson, Veit & Hake, 1994), and down-regulation of the *KNOX* gene in regions where the leaf primordia will initiate is one of the earliest indications of leaf development[29]. However, few studies have investigated the molecular biological changes that occur during the course of the heterophyllous switches to date. Hsu et al. (2001) identified several ABA-regulated early genes, designated *ABRH* for ABA-responsive heterophylly, in the aquatic fern *M. quadrifolia*. The ABRHs contain encoding transcription factors, protein kinases, membrane transporters, metabolic enzymes and structural proteins. Chen et al. (2011) studied the effects of endogenous ABA on heterophyllous alternating and its molecular mechanism in two different lily varieties, demonstrating that 9-*cis*-epoxycarotenoid dioxygenase 3 (*NCED3*) plays a key role in regulating the ABA-mediated heterophylly. Nakayama *et al*. (2014) investigated the mechanism underlying heterophylly in *R. aquatica*, and their results suggested that regulating heterophylly in this species mainly through the regulation of GA level via *KNOX1* genes. Because heterophylly has been found to occur across diverse taxa and may have arisen from convergent evolution (Minorsky, 2003), different developmental processes and molecular mechanisms may exist in different species; in addition, the heterophylly

on a single plant is controlled through multiple signalling pathways (Lin & Yang, 1999; Hsu et al., 2001). Thus, when attempting to elucidate the complex molecular mechanisms that regulate heterophylly, more study systems in diverse heterophyllous plants and large datasets generated at the whole genomic or transcriptomic levels would be helpful.

To elucidate the mechanisms underlying heterophylly, we used *Potamogeton octandrus* Poir (Potamogetonaceae), a heterophyllous pondweed with two different leaf forms (submerged and floating), as a study system. We sequenced the transcriptomes of the submerged and floating leaf materials at different developmental stages using the Illumina RNA-Seq method. Next, we performed *de novo* assembly of the DNA reads generated from all submerged and floating leaf materials as the reference transcriptome, and then mapped the short sequence reads generated from each developmental stage of the submerged and floating leaf onto the assembled transcriptome and identified the genes showing different expression between leaf morphologies. Our study provides a reference transcriptome for investigating the regulatory mechanisms of each leaf form and a list of candidate genes likely to be involved in heterophylly development. Overall, this is a first step toward elucidating the regulatory mechanisms of heterophylly in some heterophyllous plants.

MATERIALS AND METHODS

Plant materials

Potamogeton octandrus is a perennial aquatic herb that is self-compatible and can reproduce vegetatively through rhizomes or sexually by selfing and outcrossing seeds. This species can produce many floating leaves that are flat and ovate with a sharp leaf tip and submerged leaves that are linear in shape (Fig. 1). During the initial development stage, all seedlings are submerged

under water and the stem apex can produce sessile, linear and entire submerged leaves arranged in a decussate phyllotaxy (Fig. 1a). When submerged stems reach the surface of the water, the stems begin plagiotropic growth, and as the stem elongates the stem apex can produce both floating and submerged leaves (Fig.1b).

One plant of *P. octandrus* collected from the Tongcheng population (29°16'05.6"N, 113°48'46.9"E) in Hubei Province, China was used for our transcriptome analysis. The plant was transplanted to the greenhouse of Wuhan University in April, 2015 and the seeds were harvested in August, 2015. The seedlings were cultured in a pool at Wuhan University and leaves of floating and submerged forms were sampled at the following developmental stages: (1) juvenile floating leaf with length less than 0.5 cm (JFL); (2) adult floating leaf with length more than 1 cm (AFL); (3) juvenile submerged leaf with length less than 1.5 cm (JSL); (4) adult submerged leaf with length more than 3 cm (ASL). Our samples also included one developmental stage: (5) leaf-shoot of plants with stems that have an apex just reaching the water surface (shoot). At this stage, one cannot identify which leaf form it will develop into. In this study, we regarded the "shoot" as the initial stage of either floating or submerged leaves. In total, 15 samples with three biological replicates for each developmental stage were collected. Sampled tissues were immediately frozen in liquid nitrogen and then stored at −80°C until use.

RNA extraction, cDNA library construction and sequencing

Total RNA from each sample was extracted with TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. It was then treated with RNase-free DNase I (Fermentas, ThermoFisher, Waltham, MA, USA). The quantity and quality of RNA was assessed using 1% agarose gels with an RNA 6000 Nano Assay Kit and an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The total RNA was stored at −80°C.

CDNA libraries (T01-T15) were assembled for the 15 RNA samples. The cDNA was synthesized with a cDNA Synthesis Kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer's protocol. A total of 10 μg of total RNA was purified using oligo (dT) magnetic beads to isolate poly (A) mRNA. Fragmentation buffer was added to split the mRNA into short fragments. These fragments were used as templates to synthesize the first-strand cDNA using random hexamer-primers (Invitrogen, Carlsbad, CA, USA) and reverse transcriptase. RNase H and DNA polymerase I were used to synthesize second-strand cDNA. The short fragments were amended with adapter and end repair ligation. These products were first purified and then enriched with PCR (15 cycles) to create the final library of cDNA. The 15 cDNA libraries were then sequenced to obtain 150 bp paired-end short reads using an Illumina HiSeq™ 2500 sequencing platform, separately.

Sequence assembly and gene annotation

The raw reads were cleaned by removing the low quality reads, adapter reads, and reads with 5% or more unknown nucleotides. Transcriptome *de novo* assembly was performed with the clean reads that resulted from all 15 materials using the Trinity program (Grabherr et al., 2011) with min_kmer_cov set to 2 by default. Trinity first combined the overlapping sequences with the short reads to form contigs without gaps. The reads were then reverse mapped to the contigs. The paired-end reads allowed the detection of contigs from the same transcript and the distances among these contigs was determined. Trinity connected the contigs and then obtained the sequences that could not be extended on either end. Unigenes were then formed. As a final step, the putative functions of the unigene sequences were annotated using BLASTx (E-value $\leq 10^{-5}$) with several protein databases (NCBI non-redundant protein (Nr), Swiss-Prot, Cluster of Orthologous Groups (COG), euKaryotic Orthologou Groups (KOG), eggNOG, Protein family

(Pfam), Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). The best alignments determined the sequencing direction of unigenes. ESTScan was used to determine its sequencing direction when a unigene could not be aligned to any of the above databases. The program Blast2GO 2.5.0 (Conesa et al., 2005) was employed to compare and determine the unigene Gene Ontology (GO) annotations. The GO functional classifications for all annotated genes were obtained with the program WEGO (Ye et al., 2006), which was also used to plot the distribution of gene functions.

Analysis of differentially expressed genes (DEGs)

All clean reads from each sample were mapped onto the assembled reference transcriptome using the program Bowtie2 (Langmead & Salzberg, 2012) to quantify the abundance of transcripts. A likelihood ratio test was used to calculate gene expression levels and comparisons of the gene fragments per kb per million (FPKM) fragments values (Marioni et al., 2008) were performed. To determine the differentially expressed genes (DEGs) between samples from different developmental stages (control/experiment: shoot/JFL, shoot/AFL, JFL/AFL, shoot/JSL, shoot/ASL, JSL/ASL), we applied a previously developed algorithm (Audic and Claverie, 1997). False discovery rate (FDR) control was used for multiple hypotheses to correct for *P* values. Genes with changes in expression that were two-fold or greater were used in our study coupled with an FDR ≤ 0.01. Absolute values of $\log_2$ (foldchange) ≥ 1 were regarded as DEGs.

GO functional analysis provided both GO functional enrichment analysis and GO functional classification annotation for the DEGs. We mapped all DEGs to each term of the GO database. We then calculated the gene numbers for each GO term. We acquired a gene number and gene list for each GO term. Then a hypergeometric test was applied to identify the significantly enriched GO terms in DEGs relative to the genome background. All *P*-values from the GO

179 enrichment analysis were adjusted with Bonferroni's correction. A corrected *P*-value ≤ 0.05 was
180 selected as the threshold for significantly enriched GO terms. All DEGs were mapped to terms in
181 the KEGG database to obtain enriched pathway terms (Kanehisa & Goto, 2000). The pathways
182 with an FDR value of ≤ 0.05 were considered as significant DEGs.

183 **Quantitative real-time PCR (qPCR) analysis**

184 To confirm the transcriptome data, 16 DEGs were randomly selected and their expression
185 profiles were investigated by qPCRs. Total RNA was extracted from *P. octandrus* leaves at five
186 development stages and cDNA libraries were constructed using the same methods as described
187 above. The BioRad Real-Time thermal cycler system (BioRad, Hercules, CA, USA) with a
188 SYBR Premix Ex Taq^{TM} Kit (TaKaRa) were used to perform qPCR. The gene-specific primers
189 were designed using Primer Premier5, and the primer sequences are listed in File S5. The PCR
190 cycle was developed according to the manufacturer's protocol (95°C for 30s, 40 cycles of 95°C
191 for 5 s and 58°C for 30 s). Each reaction was performed three times. *β*-actin was used for a
192 reference gene. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative gene expression levels
193 (Livak & Schmittgen, 2001).

194 **RESULTS**

195 **Illumina sequencing and *de novo* assembly**

196 After removal of duplicate sequences, adaptor sequences, low quality reads, and ambiguous
197 reads, a total of 135.83 Gb clean data with Q30 ≥ 96.0% were obtained from the 15 cDNA
198 libraries. The sequencing raw data was deposited to the NCBI Short Reads Archive (SRA) with
199 the accession number SRR3184674. All clean reads were pooled and *de novo* assembled into
200 81,103 unigenes, which had an average length of 841 bp (N50=1713 bp) (Table 1). Unigenes in

the most abundant group were 201–300 bp, followed by 500–1,000 bp and 300–500 bp (Table 1). All clean data were mapped to the unigenes to analyze sequencing randomness. The sequencing randomness of each sample was shown to be sufficient (Additional file 1). The statistics of mapped reads in each sample are shown in Additional file 2.

Functional annotation and classification of unigenes

A total of 48,235 (56% of all unigenes) consensus sequences were annotated using the program BLASTx (E-value $\leq 1 \times 10^{-5}$) against public protein databases including Nr, Swiss-Prot, COG, KOG, eggNOG, Pfam, GO and KEGG. The overall functional annotation for *P. octandrus* is listed in Additional file 3. Similarity searches with known proteins in publicly available databases resulted in 42,539 unigenes being annotated to 3,711 (8.73%) proteins of *Elaeis*. Among the annotated unigenes, 24,025 were categorised into 52 functional subcategories under three main GO categories: cellular component, molecular function and biological processes (Fig. 2), among which the biological processes represented the largest number of GO annotations, with metabolic process, cellular process and single-organism process being the three top-level subcategories. In the cellular component category, the “cell” and “cell part” were dominant, while in molecular function category, the “catalytic activity” was prominent, followed by “binding”. In the COG and KOG functional classifications, the cluster of “translation, ribosomal structure and biogenesis” represented the largest group (Additional files 4 and 5). Overall, 22,346 unigenes were assigned to 128 KEGG pathways.

DEGs identification and functional analysis

A total of 6,822 DEGs were identified by comparing samples from each developmental stage (control/experiment: shoot/JFL, shoot/AFL, JFL/AFL, shoot/JSL, shoot/ASL, JSL/ASL) (Additional file 6). A Venn diagram was used to represent the numbers of DEGs in both

224 overlapping and unique sets (Fig. 3a, b). Among these comparisons, shoot/JFL contained the
225 most DEGs (3,706), whereas the smallest number of DEGs (64) was detected in JSL/ASL.
226 Among the DEGs, 2,186, 1,621 and 78 were up-regulated, while 1,835, 1,679 and 332 were
227 down-regulated between comparisons of shoot/JFL, shoot/AFL, and JFL/AFL during the
228 development of floating leaves, respectively. During the development of submerged leaves,
229 2,413, 2,242 and 45 genes were found to be up-regulated, and 1,458, 1,284 and 33 were down-
230 regulated upon analysis of shoot/JSL, shoot/ASL, and JSL/ASL, respectively (Fig. 3c). These
231 results indicated that the greatest differential expression occurred in the early stages of
232 development for both floating and submerged leaves (shoot/JFL and shoot/JSL). In addition, the
233 number of up-regulated DEGs was more than the number of down-regulated DEGs during the
234 development of submerged leaves. However, the number of up-regulated and down-regulated
235 DEGs was basically the same during the development of floating leaves.

236 Based on the functional annotations of the DEGs, the numbers of DEGs annotated in
237 shoot/JSL and shoot/ASL were greater than those in JSL/ASL upon comparisons of the
238 development of floating leaves (Fig. 4). A similar pattern was found with comparisons of the
239 development of the submerged leaves (Fig. 5). The highest number of DEGs for each
240 comparison was found for “biological process”, while the most abundant terms were “cellular
241 process”, “metabolic process”, and “single-organism” for different stages of submerged and
242 floating leaf development. The enriched GO terms during the early leaf developmental stage
243 (JFL and JSL) included “translation”, “regulation of transcription, DNA-templated”, “RNA
244 methylation” and “photosynthesis”. However, the GO terms enriched during the later leaf
245 developmental stage were assigned into “response to salt stress” (AFL and ASL), “salicylic acid
246 biosynthetic process” (AFL) and “negative regulation of programmed cell death” (ASL). In the

247 cellular component category, “cell” and “cell part” were the two most highly represented
248 components throughout the development of all leaves. The GO terms “ribosome” and “cytosolic
249 small ribosomal subunit” were enriched in different leaf developmental stages, while the GO
250 terms “cytoplasmic membrane-bounded vesicle” (JFL) and “chloroplast envelope” (JSL) were
251 enriched in the early leaf developmental stages. In the molecular function category, the DEGs
252 mapped to “catalytic activity” and “binding” were present in high proportion in all
253 developmental stages.

254 The annotated sequences were searched against the KEGG database. Among the
255 annotated DEGs, 1,490, 1,141, 135, 1,425, 1,362 and 30 were assigned to 114, 114, 57, 114, 111
256 and seven pathways in shoot/JFL, shoot/AFL, JFL/AFL, shoot/JSL, shoot/ASL, and JSL/ASL
257 comparisons, respectively (Additional file 7). The pathways with the largest proportions of
258 DEGs were “ribosome”, “biosynthesis of amino acids”, “carbon metabolism”, and “plant-
259 pathogen interaction” in shoot/JFL, shoot/AFL, shoot/JSL, and shoot/ASL, respectively
260 (Additional file 7). However, these four pathways accounted for less proportion in JFL/AFL and
261 JSL/ASL, which may be important in the early stages of development of floating and submerged
262 leaves. “Plant hormone signal transduction” was the most enriched cluster among the DEGs.
263 This indicated that this category may be essential for differences that occur in morphology and
264 physiology during the early stages of leaf development.

265 **Analysis of transcription factors associated with heterophyllous leaf types**

266 We queried *P. octandrus* TF genes in the Plant Transcription Factor Database
267 (http://planttfdb.cbi.pku.edu.cn/) to uncover the function of transcription factors (TFs) during the
268 development of heterophyllous leaves. We identified 1,681 putative TF genes that could be
269 categorized into 48 families. Of the 1,681 putative TF genes, 469 from 42 families showed

significant differential expression between developmental stages (Additional file 8 and Fig. 6). Most of these genes (53%) were expressed at the highest levels in AFL, ASL, JFL, and JSL (G2). However, only 37% were expressed at the highest levels in the “shoot” (G1). An additional 10% showed peak expression in JFL (G3). We also identified family-specific expression trends (Fig. 7). We found that the WRKY, C3H and AP2 families of TFs, which play roles in plant processes including stresses response, leaf senescence and leaf epidermal cell identity, were highly expressed in G2 (Additional file 9). Many TFs that mediate hormone signaling, such as ethylene signaling (ERF family) and auxin signaling (NAC family), were also highly expressed during the development of floating and submerged leaves development (G2). The HD-Zip family members, which have been proposed to be regulators of vascular development, stomatal complex morphogenesis, leaf polarity, and epidermal cell differentiation, were preferentially expressed in G2. MYB, ARF and B3 TFs are involved in light and hormone signaling pathways. These genes accumulated to the highest levels in the “shoot” (G1), where cell differentiation and cellular morphogenesis are modulated. Homeobox genes participating in a number of developmental events were also highly expressed in G1. The Knotted related homeobox (KNOX) was found to be related to maintenance and initiation of the shoot apical meristem and leaf morphogenesis. Several FAR1 and bHLH TFs, which have been reported to function in light signaling and stomata development, were enriched in juvenile floating leaves (G3). Our results suggested that transcriptional regulatory genes are necessary for a wide variety of developmental processes in the leaf transcriptome.

Hormone metabolism and signalling pathway among heterophyllous leaf samples

KEGG annotation revealed that most DEGs were enriched to “plant hormone signal transduction”. A total of 70 DEGs were annotated in diverse hormone-related genes, including

293 the signaling of abscisic acid (ABA), auxin, cytokinin (CTK), ethylene, jasmonic acid (JA),
294 gibberellin (GA), brassinosteroid and salicylic acid (SA).

295 Among the genes involved in hormone signaling, most were responsive to auxin during the
296 development of both floating and submerged leaves. Genes encoding auxin response factor (ARF)
297 were down-regulated, while genes encoding SAUR family members were found to be
298 differentially expressed, with two DEGs up-regulated and two down-regulated. Genes encoding
299 auxin influx transport protein (AUX1) and auxin-responsive protein AUX/IAA were up-
300 regulated in floating leaves, whereas those genes were down-regulated in submerged leaves.

301 Eight genes associated with ABA and GA were highly expressed in floating leaves,
302 including those encoding PYR (pyrabactin resistance)/PYL (PYR1-like), SNF1 related protein
303 kinase 2 (SnRK2), ABA responsive element binding factor (ABF), and phytochrome-interacting
304 factor (PIF). Genes related to GA and ethylene were highly expressed in submerged leaves,
305 including one that encoded gibberellin receptor (GID1) and two encoding ethylene-responsive
306 transcription factor 1 (ERF1).

307 Among the 31 hormone-related genes that were up-regulated during development of
308 floating leaves, four were involved in the ABA signal transduction pathway, four were
309 associated with the GA signal transduction pathway, five were responsive to auxin stimulus,
310 three were associated with ethylene-mediated signaling, two were responsive to CTK stimulus,
311 and three were related to brassinosteroid signaling. Most of these genes were up-regulated
312 significantly in the early stage of floating leaves. Additionally, 34 genes involved in the hormone
313 signal transduction pathway were up-regulated in the development of submerged leaves. When
314 compared with floating leaves, more genes were involved in GA, ethylene and CTK signal
315 transduction pathways during the development of submerged leaves (Additional file 10).

316 **Expression patterns of ABA- and GA-biosynthesis genes**

317 In the ABA biosynthesis pathway, seven DEGs, which were annotated to encode zeaxanthin
318 epoxidase (ZEP), 9-*cis*-epoxycarotenoid dioxygenase (NCED) and abscisic-aldehyde oxidase
319 (AAO), showed different levels of expression. Genes encoding ZEP and NCED were highly
320 expressed during the development of floating leaves and submerged leaves, especially in floating
321 leaves. Moreover, the expression of genes encoding AAO was down-regulated in the later stages
322 of development of submerged leaves.

323 In the GA biosynthesis pathway, the expression of one DEG encoding gibberellin 20
324 oxidase (GA20ox), a key enzyme in GA biosynthesis, was down-regulated throughout the leaf
325 development, except during the later stage of development of floating leaves. Additionally, two
326 genes encoding gibberellin 2beta-dioxygenase (GA2ox), which has an inactive effect on GAs,
327 were down-regulated. These results suggested that GA is important in early stages of leaf
328 development.

329 **Expression patterns of genes mapped to "stomatal complex morphogenesis" and "cuticle**
330 **development"**

331 In the functional categories, some genes were clustered to the terms "stomatal complex
332 morphogenesis" and "cuticle development" throughout the leaf development process.
333 Specifically, 23 DEGs were mapped to "stomatal complex morphogenesis", including 12 that
334 were up-regulated during the development of floating leaves, and especially highly expressed in
335 juvenile floating leaves, and three that were down-regulated in submerged leaves. The up-
336 regulated DEGs were categorized as phosphoribulokinase, chloroplastic, auxin-binding protein
337 ABP19a, leaf isozyme and phototropins.

338 A total of 11 DEGs were mapped to "cuticle development", four of which were up-

regulated during the development of floating leaves and three that were down-regulated. The up-regulated DEGs include orthologs of 3-ketoacyl-CoA synthase 10 and beta-ketoacyl-CoA synthase like protein. The large number of up-regulated stomatal morphogensis and cuticle development genes may regulate leaf morphology and structure.

Validation of DEGs expression by qPCR

A total of 16 candidate DEGs were selected for qPCR validation. These genes were suggested to be related to transcription factors, biological processes and hormone signaling. Five transcription factors were tested, including *NF-YB, MYB, GRF, NAC* and *NF-YA* (c260025.graph_c1, c265373.graph_c0, c267010.graph_c0, c268319.graph_c0, c269000.graph_c0). Additionally, six genes were involved in biological processes, including guard cell differentiation (c260078.graph_c1), stomatal complex morphogenesis (c260159.graph_c0, c262330.graph_c0, c267324.graph_c0), wax biosynthetic processes (c271037.graph_c1) and cuticle development (c271065.graph_c0). In hormone signaling, *AUX/IAA* (c264173.graph_c0), *AUX1* (c264621.graph_c1), *ARF* (c268917.graph_c0), *GA2ox* (c266852.graph_c0), and *GA20ox* (c268628.graph_c0) were selected. The gene annotations of these candidate DEGs are presented in Additional file 11. The correlation between the RNA-Seq results (fold change) and qPCR results ($2^{-\Delta\Delta CT}$) was measured by scatter plotting the log2 fold changes (Additional file 12a). The results revealed that the qPCR data had significant similarity ($r^2 = 0.65$) to the RNA-Seq data (Additional file 12b).

DISCUSSION

Comparative transcriptomic analysis is an efficient method for discovering genes and investigating biochemical pathways involved in physiological processes (Varshney et al., 2009; Ozsolak & Milos, 2011; Shi et al., 2011; Mutasa-Goettgen et al., 2012; Yang et al., 2014). In our

study, leaves of the aquatic heterophyllous plant *P. octandrus*, which has two leaf forms (e.g., submerged and floating), were sampled and used for transcriptome sequencing. Overall, 81,103 unigenes were assembled and 48,235 unigenes were annotated in public protein databases. A total of 6,822 DEGs between each comparison of developmental stages were identified. KEGG pathway enrichment analysis was used to sort a number of DEGs into plant hormone signal transduction pathways, including the Cytokinin, Auxin, Abscisic Acid, Gibberellin, Brassinosteroid, Ethylene, and Jasmonic Acid pathways.

The initial stage of leaf shoot development is enriched in basic biological processes, such as metabolic processes, cellular processes and responses to stimuli. The expression level of genes related to leaf morphogenesis, photomorphogenesis, and hormone signaling increased during development of submerged and floating leaves. This was especially evident in juvenile floating leaves, in which genes involved in establishing stomata and leaf petioles were more highly expressed than in other tissues. In adult submerged leaves, genes associated with leaf senescence were more highly expressed than other tissues. These gene transcriptional regulations are coincident with the developmental dynamics of heterophyllous leaves.

When grouped into the KEGG pathway, most DEGs were assumed to be connected to hormone signaling. Both exogenous and endogenous ABA and GA could regulate the leaf form alteration in heterophyllous plant species (Allsopp, 1962; Deschamp & Cooke, 1984; Gee & Anderson, 1996; Kuwabara, 2003), and the ABA-mediated regulation of morphological changes of heterophylly has been studied intensively (Wanke, 2011; Nakayama et al., 2012). In our study, a large number of annotated hormone-related DEGs belonged to ABA and GA signal transduction pathways, suggesting that genes responsive to ABA and GA might have played an important role in heterophyllous leaf formation in *P. octandrus*. The results of our study

385 indicated that in the ABA signaling pathway, the genes encoding the PYR/PYL ABA receptor,
386 which interacts with PP2C phosphatases, were up-regulated in floating leaves. The genes
387 encoding SnRF2 proteins were also up-regulated, and the activated SnRF2 proteins can
388 phosphorylate downstream targets, such as AREB/ABF transcription factors (Cutler et al., 2010).
389 Since endogenous levels of ABA have been shown to increase in the leaves of water-stressed
390 terrestrial plants and ABA plays a pivotal role in drought stress in terrestrial plant species
391 (Walton & Li, 2013), ABA might have a similar function in the initiation of heterophyllous
392 leaves in response to the transition from submerged to aerial conditions in *P. octandrus* (Goliber
393 & Feldman, 1989). Moreover, plant cells respond to environmental stimuli through a series of
394 intracellular signals. To minimize transpirational water loss (Hirayama & Shinozaki, 2010) ABA
395 controls stomatal closure. This is brought about by transcriptional reprogramming via the ABA
396 signalosome complex (PYP/PYL-PP2C-SnRK2). With the increasing of the concentration of
397 ABA, the ABA signaling complex (PYP/PYL-PP2C-SnRK2) can cause the stomatal closure in
398 guard cells through mediating the calcium-independent manner and some key biochemical
399 messengers (Geiger et al., 2011).

400 Many enzymes in ABA biosynthesis are also induced during the drought stress response.
401 Among these, ZEP, NCED and AAO are regarded as key enzymes (Iuchi et al., 2001; Chen et
402 al., 2011; Wanke, 2011). Our results revealed that 19 genes encoding ZEP, NCED and AAO
403 showed different expression levels. Six of the 19 genes were highly expressed in floating leaves
404 and up-regulated when compared to the leaf shoot and submerged leaves, while only one gene
405 was down-regulated, suggesting that differential expressions of genes that encode key enzymes
406 in ABA biosynthesis might control ABA function at specific stages during the initiation of

heterophyllous leaves. Thus, the ABA might play a complex role in signalling transduction in heterophyllous development in *P. octandrus*.

Recently, ethylene has been suggested to influence the formation of heterophyllous leaves of *L. arcuata* in an opposite way of ABA and to be an endogenous factor inducing the formation of submerged leaves (Kuwabara, Tsukaya & Nagata, 2001). Our results show that the expressions of three genes encoding 1-aminocyclopropane-1-carboxylate synthase (ACS), a rate-limiting enzyme in ethylene biosynthesis, were highly expressed in floating leaves relative to the leaf shoots and submerged leaves. GA likely has an antagonistic effect on heterophylly in aquatic plants and it also is involved in the formation of submerged leaves; however, it induces heterophylly only indirectly through ethylene (Kuwabara, Tsukaya & Nagata, 2001). Therefore, changes in endogenous ABA concentrations can influence the formation of aerial leaves and antagonistically apply feedback on ethylene and GA (Wanke, 2011). Our results also showed that two genes encoding GA2ox, which have an inactive effect on GA, were down-regulated in submerged leaves. Taken together, these findings suggest that ABA signalling was enhanced in the aerial leaves of heterophyllous plant species.

Transcription factors (TFs) are regulators that modulate the concentrations of local proteins as limiting factors at target promoters. They also play an important role in responses to environmental stress (Yuan & Perry, 2011) and plant development. Previous studies have reported that many genes in the AP2/EREBP family participate in the transcriptional regulation of biological processes related to growth and development. For example, one member of the ERF gene family, *ESR1*, regulates the shoot regeneration (Banno et al., 2001), while overexpression of the *SHN* genes, which are AP2/EREBP transcription factors, increases cuticular wax (Aharoni et al., 2004), the gene *Glossy15* from maize regulates leaf epidermal cell identity (Moose &

Sisco, 1996), and the *LEAFY PETIOLE* (*LEP*) gene influences leaf petiole development in *Arabidopsis thaliana* (van der Graaff et al., 2000). In this study, we found that 48 TFs are predominantly expressed in *P. octandrus* and 42 TFs are up- or down-regulated in each comparison. Among these detected TFs, several ERF homologs to SHN and LEAFY PETIOLE showed up-regulation during floating leaves development, which is consistent with the finding that floating leaves of *P. octandrus* have thicker waxy cuticles and extended leaf petioles. In addition, we detected the closest homolog of maize *Glossy15*, which is highly expressed in floating leaves and submerged leaves, suggesting the same function of ERF in leaves of *P. octandrus* and maize.

KNOX homeobox protein (Knotted1-like homebox, KNOX) is a homeodomain transcription factor that maintains cell pluripotency in plant shoot apical meristems (SAM) (Vollbrecht et al., 1991). Three *KNOTTED1-like HOMEOBOX* (*KNOX*) homologs were detected as a result of our DEG analysis, and all genes were highly expressed in leaf shoots. KNOX proteins regulate the homeostasis of CTK and GA to maintain meristematic cells in an undifferentiated state (Shani, Yanai & Ori, 2006). CTK is a plant hormone involved in cell proliferation and GA controls leaf morphogenesis (Hooley, 1994; Mok & Mok, 2001). Therefore, these results implied that regulation of GA levels by *KNOX1* genes is involved in regulating heterophylly in *P. octandrus*.

bZIP TFs regulate a variety of plant development and abiotic resistance processes. *AtbZIP1* from *Arabidopsis* regulates ABA signal transduction by binding to the ABA-responsive elements (ABREs) and alters the expression of ABA-responsive genes (Sun et al., 2011). In this study, we revealed three ABREs binding factors (AREB/ABF) that were differentially expressed between leaf shoots, floating leaves and submerged leaves. These findings suggested that

453 morphological differences between heterophyllous leaves may be directed by genes of multiple
454 functional groups, such as bZIP genes.

455 Several other TFs involved in ABA singling were identified in this study. For example,
456 the NAC TF, ANAC072, responds to exogenous ABA and may regulate ABA-responsive genes
457 as target genes (Tran et al., 2004). Moreover, ABA regulates gene expression through additional
458 TFs such as *MYB*, *HD-ZF*, *B3* and *bHLH* (Fujita et al., 2011). Our data indicate that diverse TFs
459 may be involved in heterophyllous leaf development. The genes are both down-regulated and up-
460 regulated, which suggests that TFs may be involved in different processes of heterophyllous leaf
461 development.

462 **Conclusions**

463 Our study describes the generation, assembly and annotation of the transcriptomes of submerged
464 and floating leaves of an aquatic plant, *P. octandrus*. We identified a number of genes that
465 showed differential expression at different developmental stages during the development of
466 heterophylly in this species. Our results aid gene cloning and functional studies of genes that are
467 involved in leaf development, especially those genes that are involved in the process of
468 heterophyllous leaf development.

469 **REFERENCES**

470 Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A. 2004. The SHINE clade of AP2
471 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers
472 drought tolerance when overexpressed in *Arabidopsis*. *Plant Cell* 16: 2463-2480.
473 Allsopp A. 1962. The effects of gibberellic acid on morphogenesis in *Marsilea drummondii*.
474 *Phytomorphology* 12: 1-10.

475 Anderson LW. 1978. Abscisic acid induces formation of floating leaves in the heterophyllous
476 aquatic angiosperm *Potamogeton nodosus*. *Science* 201: 1135-1138.
477 Arber A. 1920. Water Plants: A Study of Aquatic Angiosperms. Cambridge University Press.
478 Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome*
479 *Research* 7: 986-995.
480 Banno H, Ikeda Y, Niu QW, Chua NH. 2001. Overexpression of *Arabidopsis* ESR1 induces
481 initiation of shoot regeneration. *Plant Cell* 13: 2609-2618.
482 Bodkin PC, Spence D, Weeks DC. 1980. Photoreversible control of heterophylly in *Hippuris*
483 *vulgaris* L. *New Phytologist* 84: 533.
484 Chen H, Hwang S, Chen S, Shii C, Cheng W. 2011. ABA-mediated heterophylly is regulated by
485 differential expression of 9-*cis*-epoxycarotenoid dioxygenase 3 in lilies. *Plant and Cell*
486 *Physiology* 52: 1806-1821.
487 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal
488 tool for annotation, visualization and analysis in functional genomics research.
489 *Bioinformatics* 21: 3674-3676.
490 Cook SA, Johnson MP. 1968. Adaptation to heterogenous environments I. Variation in
491 heterophylly in *Ranunculus flammula* L. *Evolution* 22: 496-516.
492 Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic acid: emergence of a core
493 signaling network. *Annual Review of Plant Biology* 61: 651-679.
494 Deschamp PA, Cooke TJ. 1984. Causal mechanisms of leaf dimorphism in the aquatic
495 angiosperm *Callitriche heterophylla*. *American Journal of Botany* 71: 319-329.
496 Deschamp PA, Cooke TJ. 1985. Leaf dimorphism in the aquatic angiosperm *Callitriche*
497 *heterophylla*. *American Journal of Botany* 72: 1377-1387.

498 Fassett NC. 1930. A Manual of Aquatic Plants. The University of Wisconsin Press.
499 Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K. 2011. ABA-mediated transcriptional
500 regulation in response to osmotic stress in plants. *Journal of Plant Research* 124: 509-525.
501 Gee D, Anderson L. 1996. ABA induced differences during leaf development in the aquatic
502 angiosperm, *Potamogeton nodosus*, are detected with differential display. *Plant Physiology*
503 111: 110.
504 Geiger D, Maierhofer T, Al-Rasheid KA, Scherzer S, Mumm P, Liese A, et al. 2011. Stomatal
505 closure by fast abscisic acid signaling is mediated by the guard cell anion channel SLAH3
506 and the receptor RCAR1. *Science Signaling* 4 (173): ra32.
507 Goliber T, Feldman L. 1989. Osmotic stress, endogenous abscisic acid and the control of leaf
508 morphology in *Hippuris vulgaris* L. *Plant Cell and Environment* 12: 163-171.
509 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. 2011. Full-length
510 transcriptome assembly from RNA-Seq data without a reference genome. *Nature*
511 *Biotechnology* 29: 130-644.
512 Hay A, Tsiantis M. 2006. The genetic basis for differences in leaf form between *Arabidopsis*
513 *thaliana* and its wild relative *Cardamine hirsuta*. *Nature Genetics* 38: 942-947.
514 Hirayama T, Shinozaki K. 2010. Research on plant abiotic stress responses in the post-genome
515 era: past, present and future. *The Plant Journal* 61:1041-1052.
516 Hooley R. 1994. Gibberellins: perception, transduction and responses. *Plant Molecular Biology*
517 26: 1529-1555.
518 Hsu TC, Liu HC, Wang JS, Chen RW, Wang YC, Lin BL. 2001. Early genes responsive to
519 abscisic acid during heterophyllous induction in *Marsilea quadrifolia*. *Plant Molecular*
520 *Biology* 47: 703-715.

521 Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, et al. 2001. Regulation of drought
522 tolerance by gene manipulation of 9-*cis*-epoxycarotenoid dioxygenase, a key enzyme in
523 abscisic acid biosynthesis in *Arabidopsis*. *The Plant Journal* 27: 325-333.
524 Jackson D, Veit B, Hake S. 1994. Expression of maize *KNOTTED1* related homeobox genes in
525 the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot.
526 *Development* 120: 405-413.
527 Kane ME, Albert LS. 1987. Abscisic acid induces aerial leaf morphology and vasculature in
528 submerged *Hippuris vulgaris* L. *Aquatic Botany* 28: 81-88.
529 Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids*
530 *Research* 28: 27-30.
531 Kuwabara A, Ikegami K, Koshiba T, Nagata T. 2003. Effects of ethylene and abscisic acid upon
532 heterophylly in *Ludwigia arcuata* (Onagraceae). *Planta* 217: 880-887.
533 Kuwabara A, Tsukaya H, Nagata T. 2001. Identification of factors that cause heterophylly in
534 *Ludwigia arcuata* Walt.(Onagraceae). *Plant Biology* 3: 98-105.
535 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:
536 357-359.
537 Lin BL, Yang WJ. 1999. Blue light and abscisic acid independently induce heterophyllous
538 switch in *Marsilea quadrifolia*. *Plant Physiology* 119: 429-434.
539 Liu BL. 1984. Abscisic acid induces land form characteristics in *Marsilea quadrifolia* L.
540 *American Journal of Botany* 71: 638-644.
541 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
542 quantitative PCR and the 2-ΔΔCt method. *Methods* 25: 402-408.

543 Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. 2008. RNA-seq: an assessment of
544 technical reproducibility and comparison with gene expression arrays. *Genome Research* 18:
545 1509-1517.
546 McCallum WB. 1902. On the nature of the stimulus causing the change of form and structure in
547 *Prosperpinaca palustris*. *Botanical Gazette* 34:93-108.
548 Minorsky PV. 2003. Heterophylly in aquatic plants. Plant Physiology 133: 1671-1672.
549 Mok D, Mok MC. 2001. Cytokinin metabolism and action. Annual Review of Plant Physiology
550 and Plant Molecular Biology 52: 89-118.
551 Moon J, Hake S. 2011. How a leaf gets its shape. Curr. Opin. Plant Biol. 14: 24-30.
552 Moose SP, Sisco PH. 1996. Glossy15, an *APETALA2*-like gene from maize that regulates leaf
553 epidermal cell identity. Genes & Development 10: 3018-3027.
554 Mutasa-Goettgen, ES, Joshi A, Holmes HF, Hedden P, Goettgens B. 2012. A new RNASeq-
555 based reference transcriptome for sugar beet and its application in transcriptome-scale
556 analysis of vernalization and gibberellin responses. BMC Genomics 13: 99.
557 Nakayama H, Nakayama N, Nakamasu A, Sinha N, Kimura S. 2012. Toward elucidating the
558 mechanisms that regulate heterophylly. Plant Morphology. 24: 57-63.
559 Nakayama H, Nakayama N, Seiki S, Kojima M, Sakakibara H, Sinha N, et al. 2014. Regulation
560 of the *KNOX-GA* gene module induces heterophyllic alteration in North American Lake Cress.
561 *Plant Cell* 26: 4733-4748.
562 Ozsolak F, Milos PM. 2011. RNA sequencing: advances, challenges and opportunities. Nature
563 Reviews Genetics 2: 87-98.
564 Pigliucci M. 2010. Phenotypic plasticity. In: Pigliucci M, Muller GM, editors. Evolution: the
565 extended synthesis. MIT Press; p. 355-378.

566 Sculthorpe CD. 1967. The Biology of Aquatic Vascular Plants. Arnold E.
567 Shani E, Ben-Gera H, Shleizer-Burko S, Burko Y, Weiss D, Ori N. 2010. Cytokinin regulates
568 compound leaf development in Tomato. Plant Cell 22: 3206-3217.
569 Shani E, Yanai O, Ori N. 2006. The role of hormones in shoot apical meristem function. Current
570 Opinion in Plant Biology 9: 484-489.
571 Shi C, Yang H, Wei CL, Yu O, Zhang ZZ, Jiang CJ, et al. 2011. Deep sequencing of the
572 *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of
573 tea-specific compounds. BMC Genomics 12: 131.
574 Smith LG, Greene B, Veit B, Hake S. 1992, A dominant mutation in the maize homeobox gene,
575 *Knotted-1*, causes its ectopic expression in leaf cells with altered fates. Development 116: 21-
576 30.
577 Sun X L, Li Y, Cai H, Bai X, Ji W, Ji ZJ, et al. 2011. Arabidopsis bZIP1 transcription factor
578 binding to ABRE cis-element regulates abscisic acid signal transduction. Acta Agronomica
579 Sinica. 37: 612-619.
580 Titus JE, Sullivan PG. Heterophylly in the yellow waterlily, *Nuphar variegata* (Nymphaeaceae):
581 effects of [CO_2] natural sediment type, and water depth. Amer J Bot. 2001; 88: 1469-1478.
582 Tran L, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, et al. Isolation and
583 functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a
584 drought-responsive *cis*-element in the early responsive to dehydration stress 1 promoter.
585 Plant Cell. 2004; 16: 2481-2498.
586 Tsukaya H. Mechanism of leaf-shape determination. Annu Rev Plant Biol. 2006; 57: 477-496.
587 Uchida N, Kimura S, Koenig D, Sinha N. Coordination of leaf development via regulation of
588 *KNOX1*genes. J Plant Res. 2010; 123:7-14.

589 Uchida N, Townsley B, Chung KH, Sinha N. Regulation of *SHOOT MERISTEMLESS* genes via
590 an upstream-conserved noncoding sequence coordinates leaf development. Proc Natl Acad
591 Sci. USA. 2007; 104: 15953-15958.
592 van der Graaff E, Den Dulk-Ras A, Hooykaas P, Keller B. Activation tagging of the LEAFY
593 PETIOLE gene affects leaf petiole development in Arabidopsis thaliana. Development. 2000;
594 127: 4971-4980.
595 Varshney R, Nayak S, May G, Jackson S. Next-generation sequencing technologies and their
596 implications for crop genetics and breeding. Trends Biotechnol. 2009; 27: 522-530.
597 Vollbrecht E, Veit B, Sinha N, Hake S. The developmental gene Knotted-1 is a member of a
598 maize homeobox gene family. Nature. 1991; 350: 241-243.
599 Walton DC, Li Y. Abscisic acid biosynthesis and metabolism. In: Davies P, editor. Plant
600 hormones: Physiology, Biochemistry and Molecular Biology. Springer Science & Business
601 Media; 2013. p.140-157.
602 Wanke D. The ABA-mediated switch between submersed and emersed life-styles in aquatic
603 macrophytes. J Plant Res. 2011; 124: 467-475.
604 Wells CL, Pigliucci M. Adaptive phenotypic plasticity: the case of heterophylly in aquatic plants.
605 Perspect Plant Ecol Evol Syst. 2000; 3: 1-18.
606 Winn AA. Is seasonal variation in leaf traits adaptive for the annual plant *Dicerandra linearifolia*?
607 J Evol Biol. 1999; 12: 306-313.
608 Winn AA. The functional significance and fitness consequences of heterophylly. Int J Plant Sci.
609 1999; 160: S113-S121.

610 Yang M, Zhu L, Xu L, Pan C, Liu Y. Comparative transcriptomic analysis of the regulation of
611 flowering in temperate and tropical lotus (*Nelumbo nucifera*) by RNA-Seq. Annals Applied
612 Biol. 2014; 165: 73-95.
613 Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting GO
614 annotations. Nucleic Acids Res. 2006; 34: W293-W297.
615 Young JP, Dengler NG, Horton RF. Heterophylly in *Ranunculus flabellaris*: The effect of
616 abscisic acid on leaf anatomy. Ann Bot. 1987; 60: 117-125.
617 Young JP, Horton RF. Heterophylly in *Ranunculus flabellaris*: the effect of abscisic acid. Ann
618 Bot. 1985; 55: 899-902.
619 Yuan L, Perry SE. Plant transcription factors. Meth Protocol Mol Biol. 2011; 754: 347.
620 Zhang X, Gituru RW, Yang C, Guo YH. Variations of floral traits among different life forms
621 illustrate the evolution of pollination systems in *Potamogeton* species from China. Aquat Bot.
622 2009; 90: 124-128.
623 Zotz G, Wilhelm K, Becker A. Heteroblasty-a review. Bot Rev. 2011; 77: 109-151.
624

625 **ADDITIONAL INFORMATION AND DECLARATIONS**

626 **Funding**

627 This study was supported by the Specialized Research Fund for the Doctoral Program of Higher
628 Education (20110141110019), the National Natural Science Foundation of China (Nos.
629 31270278 and 31570220), the Strategic Priority Research Program of the Chinese Academy of
630 Sciences (No. XDPB02), and the Wuhan Botanical Garden, Chinese Academy of Sciences (No.
631 Y655261W03).

Grant Disclosures

The following grant information was disclosed by the authors:

Specialized Research Fund for the Doctoral Program of Higher Education: 20110141110019;

National Natural Science Foundation of China: 31270278, 31570220;

Strategic Priority Research Program of the Chinese Academy of Sciences: XDPB02;

Wuhan Botanical Garden, Chinese Academy of Sciences: Y655261W03.

Competing financial interests

The authors declare no competing financial interests.

Data Availability

All the sequencing raw data was deposited to the NCBI Short Reads Archive (SRA) with the accession number SRR3184674. The other supporting results from this study can be found within both the article and Additional files 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

Author Contributions

Dingxuan He, Pin Guo, and Xing Liu conceived and designed the experiments. Dingxuan He and Ping Guo performed the experiments.

Dingxuan He analyzed the data.

Youhao Guo and Dingxuan He contributed reagents/materials/analysis tools.

Dingxuan He, Jinming Chen and Paul F. Gugger wrote the paper.

Supplemental Information

Supplemental information for this article can be found online.

Figure legends:

Fig. 1 Morphological features of *P. octandrus*. (A) The initial development stage of the plant that produces only submerged leaves. (B) The later development stage of the plant that produces both floating and submerged leaves.

Fig. 2 Gene ontology (GO) annotations of all detected genes. Three main categories including biological process, cellular component and molecular function were summarized.

Fig. 3 The DEGs in different comparisons during *P. octandrus* leaves development. "T01, T04, T12", "T02, T03, T05", "T06, T08, T13", "T07, T10, T14", and "T08, T11, T15" indicate three biological replicates of shoot, juvenile floating leaves, adult floating leaves, juvenile submerged leaves and adult submerged leaves, respectively. (A) Venn diagram showing the unique and common DEGs among different comparisons in floating leaves. (B) Venn diagram showing the unique and common DEGs among different comparisons in submerged leaves. (C) The expression patterns of DEGs among different comparisons.

Fig.4 GO function classifications of DEGs among different comparisons during floating leaves development. T01, T04, T12, T02, T03, T05, T06, T08 and T13 indicate three biological replicates of shoot, juvenile floating leaves and adult floating leaves, respectively.

Fig. 5 GO function classifications of DEGs among different comparisons during submerged leaves development. "T01, T04, T12", "T07, T10, T14", and "T08, T11, T15" indicate three biological replicates of shoot, juvenile submerged leaves and adult submerged leaves, respectively.

Fig. 6 Dendrogram showing the transcription factors. Three clusters (G1, G2 and G3) resulted from the 469 significantly differentially expressed transcription factors from leaves in different stages of development.

675 **Fig. 7** Distribution of transcription factor families among the three clusters (G1, G2 and G3).

676

677 **Supplemental files**

678 **Additional file 1:** The distribution map of randomness test of the sequence reads in *P.*

679 *Octandrus*.

680 **Additional file 2:** Aligning statistics of clean reads with assembled unigenes. Q30 percentage

681 is proportion of nucleotides with quality value larger than 30 in reads; GC percentage is

682 proportion of guanine and cytosine nucleotides among total nucleotides.

683 **Additional file 3:** Functional annotation of the *P. octandrus*.

684

685 **Additional file 4:** Clusters of orthologous groups (COG) classifications for *P. octandrus*. In

686 total, 20,856 of the 48,235 sequences with Nr hits were groups into 25 classifications.

687 **Additional file 5:** Clusters of euKaryotic orthologous groups (KOG) classifications for *P.*

688 *octandrus*. In total, 30,177 of the 48,235 sequences with Nr hits were groups into 25

689 classifications

690 **Additional file 6:** A list of 6822 differentially expressed genes (DEGs) at each development

691 stage. We defined genes with expression change of $|\log_2 FC| \geq 1$ and FDR < 0.01 as DEGs. The

692 FPKM of two stages, value of log2FC and expression pattern of DEGs are presented in the table.

693 **Additional file 7:** A list of KEGG pathways mapped by 1,490, 1,141, 135, 1,425, 1,362 and

694 30 DEGs in all comparisons, respectively.

695 **Additional file 8:** A list of putative transcription factors (TF) differentially expressed during
696 heterophyllous leaves development. 469 putative TF genes are identified into 42 families.
697 **Additional file 9:** Representative functions of TF genes showing expression gradients
698 between floating leaf and submerged leaf tissue samples.
699 **Additional file 10:** A list of 70 DEGs involved in the plant hormone signal transduction
700 pathway during heterophyllous leaves development. We defined genes with expression
701 change of $|\log2FC| \geq 1$ and FDR < 0.01 as DEGs. The values of log2FC and
702 expression pattern of DEGs are presented in the table.
703 **Additional file 11:** The candidate DEGs IDs, annotations and the primers used for qPCR
704 analysis.
705 **Additional file 12:** Quantitative real-time polymerase chain reaction (qPCR) validation of
706 RNA-seq data. (a) Expression profiles of the candidate genes by qPCR. The value of relative
707 expression level on the y-axis calculated according to $\log(2^{-\Delta\Delta CT})$. The expression level of
708 each gene in shoot was arbitrarily set as 1, the value of the y-axis was 0. (b) Correlation of
709 gene expression results. The x-axis represents the value of Log2 fold change by RNA-seq and
710 the y-axis represents the value of Log2 relative expression level by qPCR.

Figure 1(on next page)

Fig. 1 Morphological features of P. octandrus. (A) The initial development stage of the plant that produces only submerged leaves. (B) The later development stage of the plant that produces both float

Fig. 1 Morphological features of P. octandrus. (A) The initial development stage of the plant that produces only submerged leaves. (B) The later development stage of the plant that produces both floating and submerged leaves.

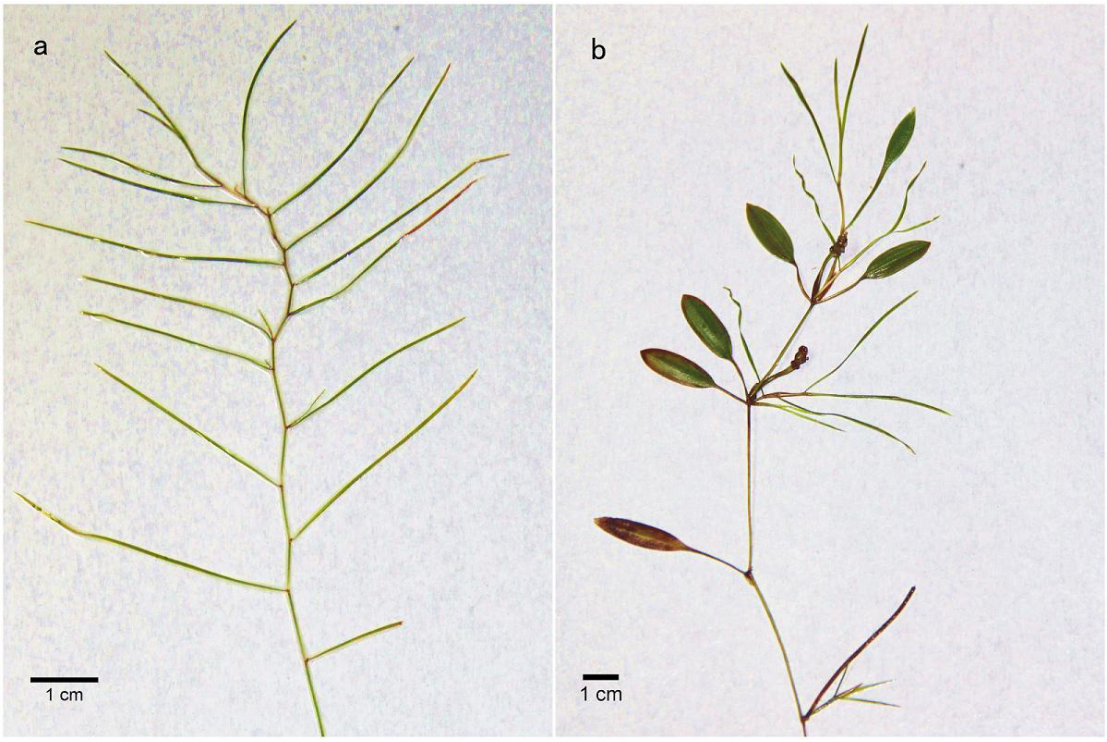

Fig. 1

Figure 2(on next page)

Fig. 2 Gene ontology (GO) annotations of all detected genes. Three main categories including biological process, cellular component and molecular function were summarized.

Fig. 2 Gene ontology (GO) annotations of all detected genes. Three main categories including biological process, cellular component and molecular function were summarized.

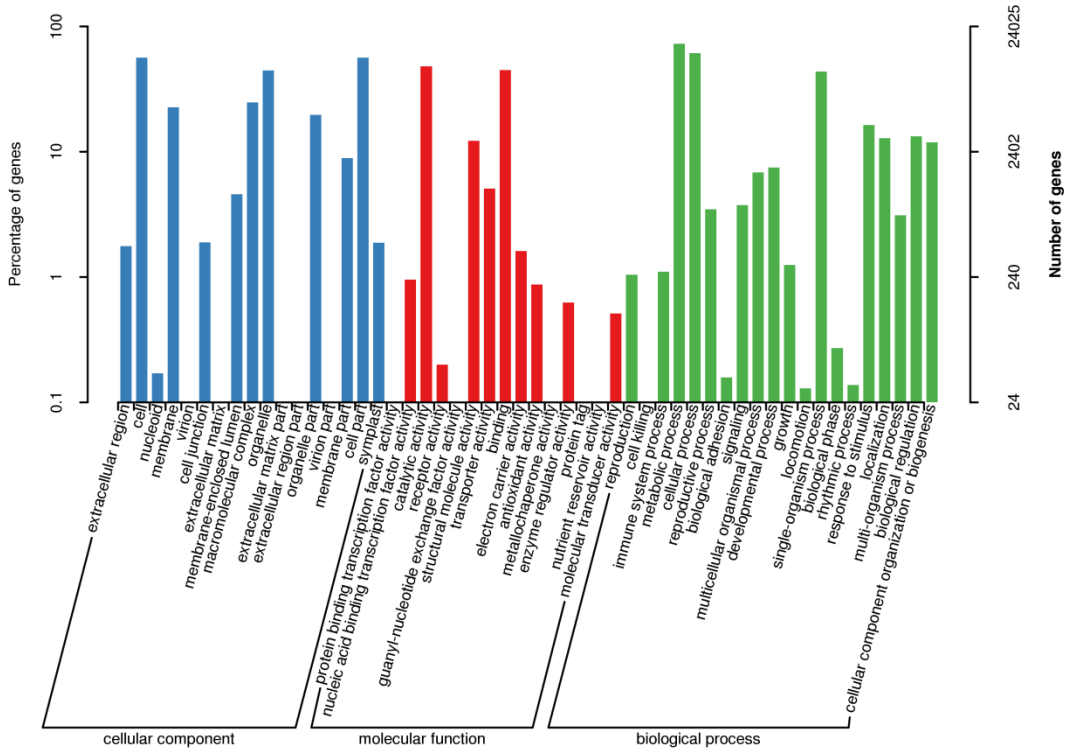
Percentage of genes
100
10
1
0.1
Number of genes
24025
2402
240
24
extracellular region
cell
nucleoid
membrane
virion
cell junction
extracellular matrix
membrane-enclosed lumen
macromolecular complex
organelle
extracellular matrix part
extracellular region part
organelle part
virion part
membrane part
cell part
symplast
protein binding transcription factor activity
nucleic acid binding transcription factor activity
catalytic activity
receptor activity
guanyl-nucleotide exchange factor activity
structural molecule activity
transporter activity
binding
electron carrier activity
antioxidant activity
metallochaperone activity
enzyme regulator activity
protein tag
nutrient reservoir activity
molecular transducer activity
reproduction
cell killing
immune system process
metabolic process
cellular process
reproductive process
biological adhesion
multicellular organismal process
signaling
developmental process
growth
locomotion
single-organism process
biological phase
rhythmic process
response to stimulus
localization
multi-organism process
biological regulation
cellular component organization or biogenesis
cellular component
molecular function
biological process

Fig. 2

Figure 3(on next page)

Fig. 3 The DEGs in different comparisons during P. octandrus leaves development.

Fig. 3 The DEGs in different comparisons during P. octandrus leaves development. “T01, T04, T12”, “T02, T03, T05”, “T06, T08, T13”, “T07, T10, T14”, and “T08, T11, T15” indicate three biological replicates of shoot, juvenile floating leaves, adult floating leaves, juvenile submerged leaves and adult submerged leaves, respectively. (A) Venn diagram showing the unique and common DEGs among different comparisons in floating leaves. (B) Venn diagram showing the unique and common DEGs among different comparisons in submerged leaves. (C) The expression patterns of DEGs among different comparisons.

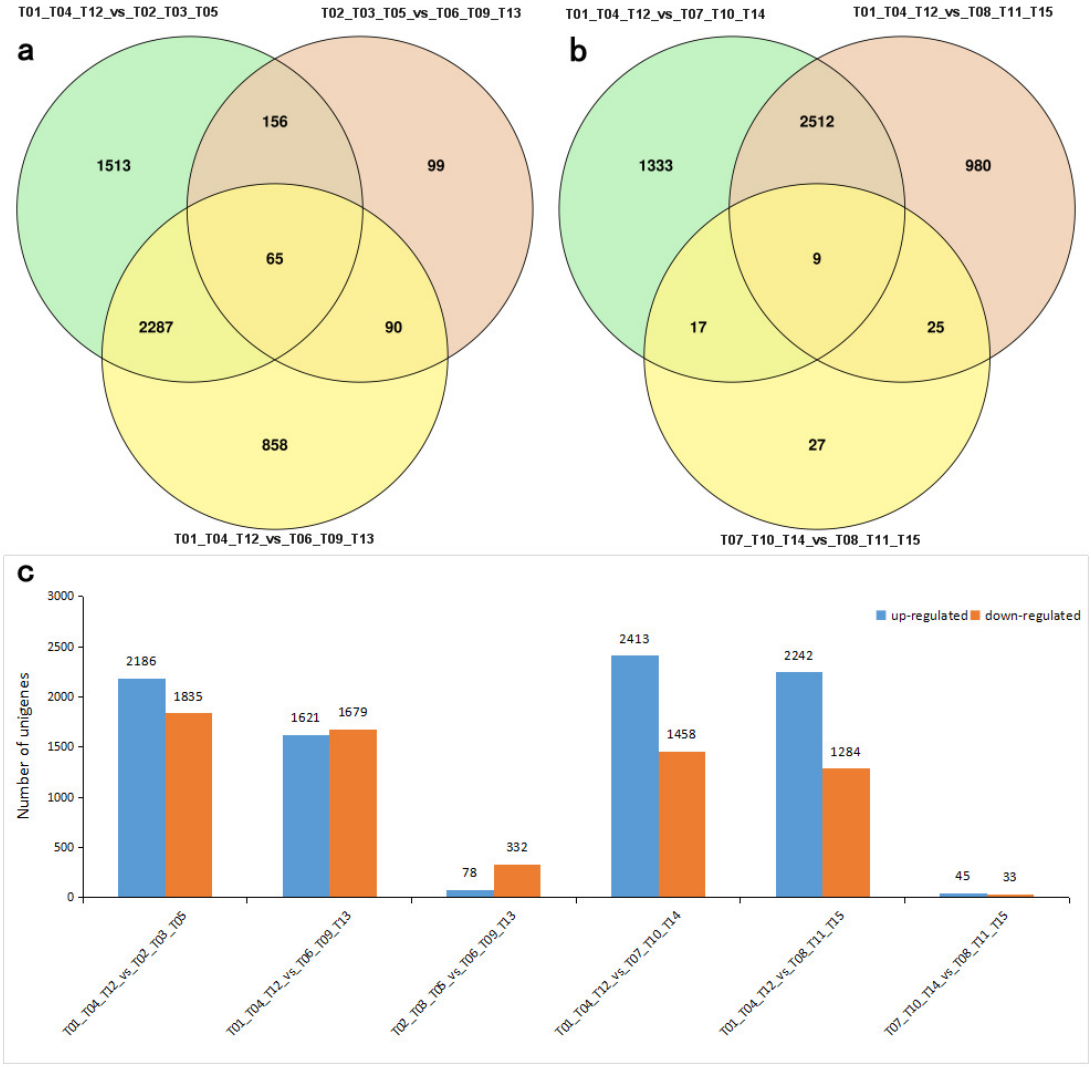
a
T01_T04_T12_vs_T02_T03_T05
T02_T03_T05_vs_T06_T09_T13
T01_T04_T12_vs_T06_T09_T13
1513
156
99
65
2287
90
858
b
T01_T04_T12_vs_T07_T10_T14
T01_T04_T12_vs_T08_T11_T15
T07_T10_T14_vs_T08_T11_T15
1333
2512
980
9
17
25
27
c
Number of unigenes
up-regulated
down-regulated
2186
1835
1621
1679
78
332
2413
1458
2242
1284
45
33
T01_T04_T12_vs_T02_T03_T05
T01_T04_T12_vs_T06_T09_T13
T02_T03_T05_vs_T06_T09_T13
T01_T04_T12_vs_T07_T10_T14
T01_T04_T12_vs_T08_T11_T15
T07_T10_T14_vs_T08_T11_T15

Fig. 3

Figure 4(on next page)

Fig.4 GO function classifications of DEGs among different comparisons during floating leaves development.

Fig.4 GO function classifications of DEGs among different comparisons during floating leaves development. T01, T04, T12, T02, T03, T05, T06, T08 and T13 indicate three biological replicates of shoot, juvenile floating leaves and adult floating leaves, respectively.

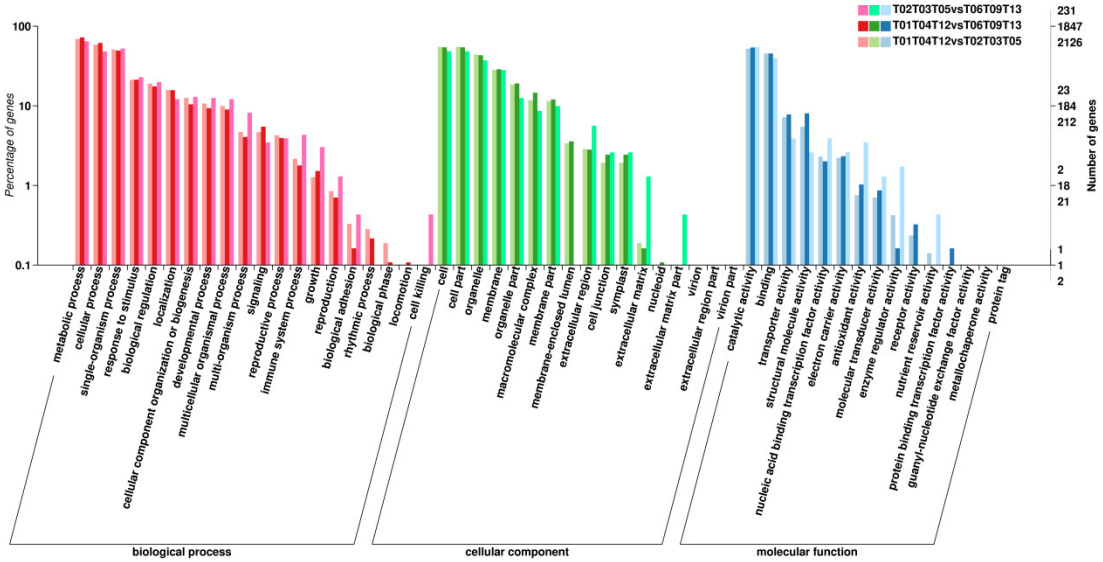
T02T03T05vsT06T09T13
T01T04T12vsT06T09T13
T01T04T12vsT02T03T05
Percentage of genes
Number of genes
100
10
1
0.1
231
1847
2126
23
184
212
2
18
21
1
1
2
metabolic process
cellular process
single-organism process
response to stimulus
biological regulation
cellular component organization or biogenesis
localization
developmental process
multicellular organismal process
multi-organism process
signaling
reproductive process
immune system process
growth
reproduction
biological adhesion
rhythmic process
biological phase
locomotion
cell killing
cell
cell part
organelle
membrane
organelle part
macromolecular complex
membrane part
membrane-enclosed lumen
extracellular region
cell junction
symplast
extracellular matrix
nucleoid
extracellular matrix part
virion
extracellular region part
virion part
catalytic activity
binding
transporter activity
structural molecule activity
nucleic acid binding transcription factor activity
electron carrier activity
antioxidant activity
molecular transducer activity
enzyme regulator activity
receptor activity
nutrient reservoir activity
protein binding transcription factor activity
guanyl-nucleotide exchange factor activity
metallochaperone activity
protein tag
biological process
cellular component
molecular function

Fig. 4

Figure 5(on next page)

Fig. 5 GO function classifications of DEGs among different comparisons during submerged leaves development.

Fig. 5 GO function classifications of DEGs among different comparisons during submerged leaves development. “T01, T04, T12”, “T07, T10, T14”, and “T08, T11, T15” indicate three biological replicates of shoot, juvenile submerged leaves and adult submerged leaves, respectively.

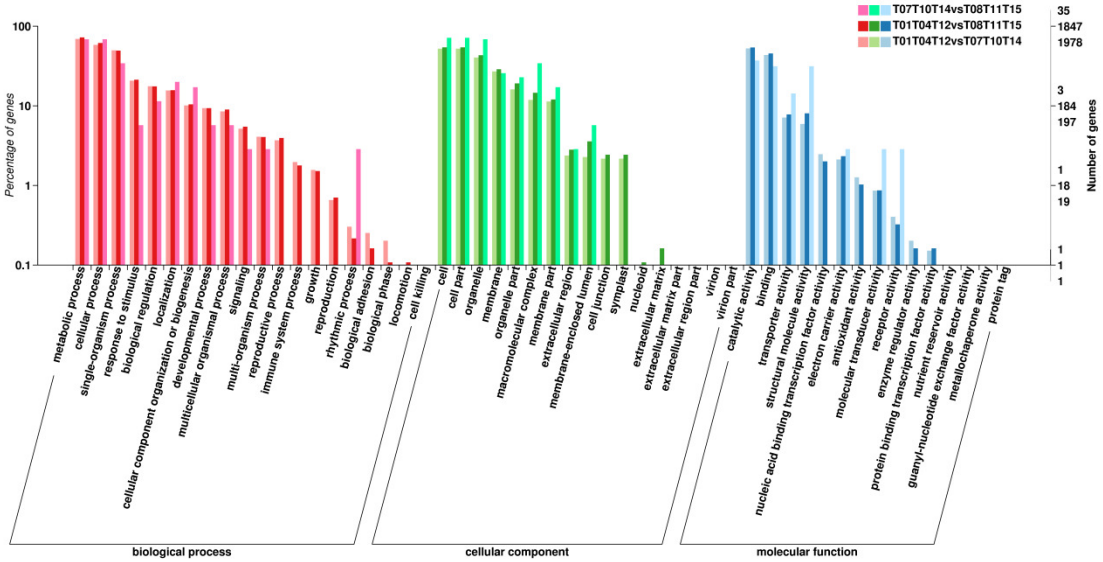
T07T10T14vsT08T11T15
T01T04T12vsT08T11T15
T01T04T12vsT07T10T14
Percentage of genes
Number of genes
100
10
1
0.1
35
1847
1978
3
184
197
1
18
19
1
1
1
metabolic process
cellular process
single-organism process
response to stimulus
biological regulation
cellular component organization or biogenesis
localization
developmental process
multicellular organismal process
signaling
multi-organism process
reproductive process
immune system process
growth
reproduction
rhythmic process
biological adhesion
biological phase
locomotion
cell killing
cell
cell part
organelle
membrane
organelle part
macromolecular complex
membrane part
extracellular region
membrane-enclosed lumen
cell junction
symplast
nucleoid
extracellular matrix
extracellular matrix part
extracellular region part
virion
virion part
catalytic activity
binding
transporter activity
structural molecule activity
nucleic acid binding transcription factor activity
electron carrier activity
antioxidant activity
molecular transducer activity
receptor activity
enzyme regulator activity
protein binding transcription factor activity
nutrient reservoir activity
guanyl-nucleotide exchange factor activity
metallochaperone activity
protein tag
biological process
cellular component
molecular function

Fig. 5

Figure 6(on next page)

Fig. 6 Dendrogram showing the transcription factors.

Fig. 6 Dendrogram showing the transcription factors. Three clusters (G1, G2 and G3) resulted from the 469 significantly differentially expressed transcription factors from leaves in different stages of development.

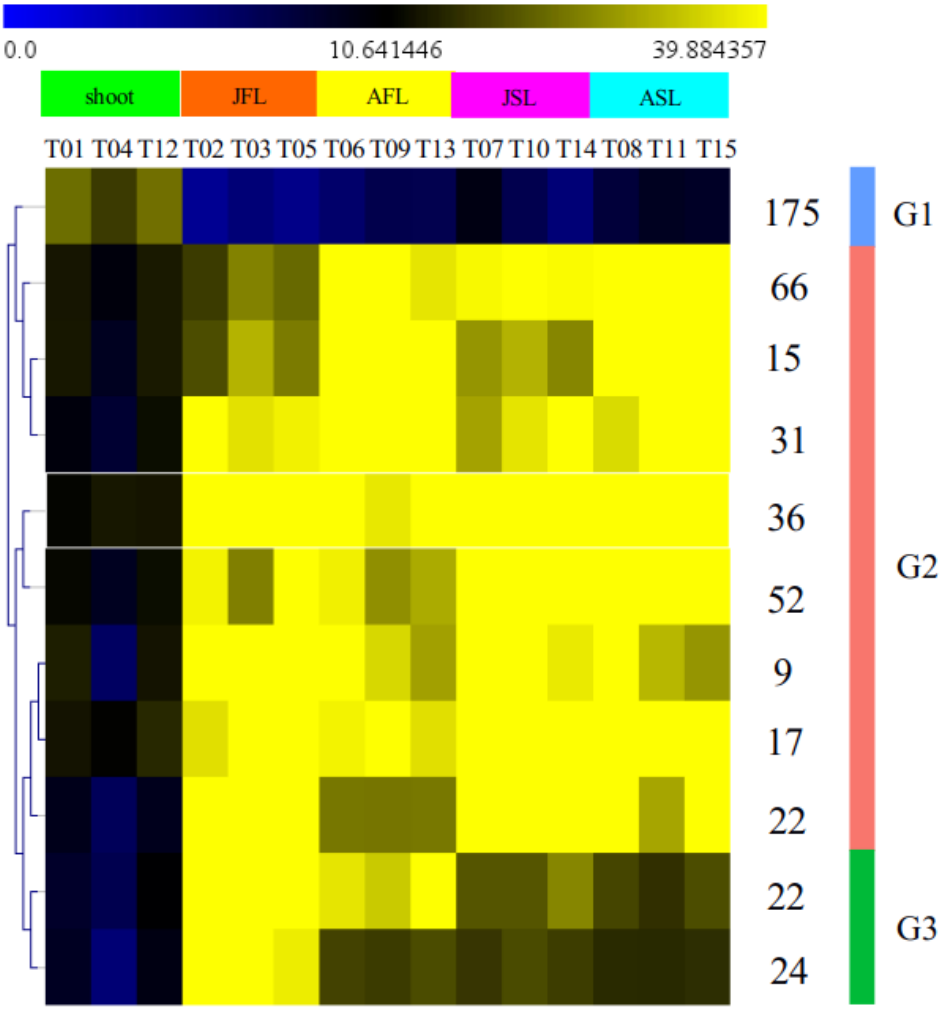
0.0
10.641446
39.884357
shoot
JFL
AFL
JSL
ASL
T01 T04 T12 T02 T03 T05 T06 T09 T13 T07 T10 T14 T08 T11 T15
175
66
15
31
36
52
9
17
22
22
24
G1
G2
G3

Fig. 6

Figure 7 (on next page)

Fig. 7 Distribution of transcription factor families among the three clusters (G1, G2 and G3).

Fig. 7 Distribution of transcription factor families among the three clusters (G1, G2 and G3).

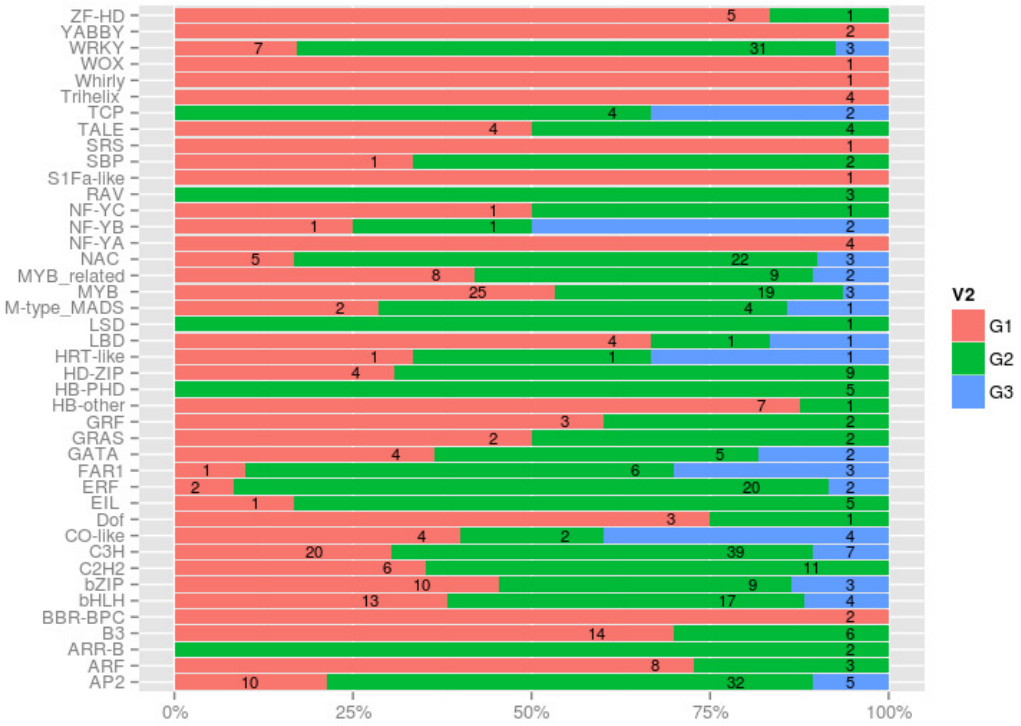
ZF-HD
YABBY
WRKY
WOX
Whirly
Trihelix
TCP
TALE
SRS
SBP
S1Fa-like
RAV
NF-YC
NF-YB
NF-YA
NAC
MYB_related
MYB
M-type_MADS
LSD
LBD
HRT-like
HD-ZIP
HB-PHD
HB-other
GRF
GRAS
GATA
FAR1
ERF
EIL
Dof
CO-like
C3H
C2H2
bZIP
bHLH
BBR-BPC
B3
ARR-B
ARF
AP2
0%
25%
50%
75%
100%
V2
G1
G2
G3

Fig. 7

Table 1(on next page)

Table 1 Overview of the sequencing and assembly from RNA-Seq data.

Table 1 Overview of the sequencing and assembly from RNA-Seq data.

1 **Table 1** Overview of the sequencing and assembly from RNA-Seq data.

Length	Transcript	Unigene
200-300	40,102(23.97%)	38,142(44.30%)
300-500	17,132(10.24%)	11,774(13.67%)
500-1000	30,450(18.20%)	15,041(17.47%)
1000-2000	39,384(23.55%)	11,399(13.24%)
2000+	40,202(24.03%)	9,747(11.32%)
Total Number	167,270	86,103
Total Length	224,870,533	72,422,038
N50 Length	2,273	1,713
Mean Length	1344.36	841.11

2

3

4