A peer-reviewed version of this preprint was published in PeerJ on 28 February 2018.

View the peer-reviewed version (peerj.com/articles/4448), which is the preferred citable publication unless you specifically need to cite this preprint.

https://doi.org/10.7717/peerj.4448
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 heterophyllly 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;
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 heterophylous 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 CO₂ 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. arcuata*, 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\textit{HOMEOBOS (KNOX)} gene (Smith et al., 1992; Jackson, Veit & Hake, 1994), and down-regulation of the \textit{KNOX} gene in regions where the leaf primordia will initiate is one of the earliest indications of leaf development\textsuperscript{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 \textit{ABRH} for ABA-responsive heterophyllly, in the aquatic fern \textit{M. quadrifolia}. The \textit{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-\textit{cis}-epoxycarotenoid dioxygenase 3 (\textit{NCED3}) plays a key role in regulating the ABA-mediated heterophyllly.

Nakayama \textit{et al.} (2014) investigated the mechanism underlying heterophyllly in \textit{R. aquatica}, and their results suggested that regulating heterophyllly in this species mainly through the regulation of GA level via \textit{KNOXI} genes. Because heterophyllly 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 heterophyllly
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 ≤ 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 $\leq 0.01$. Absolute values of $\log_2$ (foldchange) $\geq 1$ 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
enrichment analysis were adjusted with Bonferroni’s correction. A corrected $P$-value $\leq 0.05$ was selected as the threshold for significantly enriched GO terms. All DEGs were mapped to terms in the KEGG database to obtain enriched pathway terms (Kanehisa & Goto, 2000). The pathways with an FDR value of $\leq 0.05$ were considered as significant DEGs.

Quantitative real-time PCR (qPCR) analysis

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

RESULTS

Illumina sequencing and de novo assembly

After removal of duplicate sequences, adaptor sequences, low quality reads, and ambiguous reads, a total of 135.83 Gb clean data with Q30 $\geq 96.0\%$ were obtained from the 15 cDNA libraries. The sequencing raw data was deposited to the NCBI Short Reads Archive (SRA) with the accession number SRR3184674. All clean reads were pooled and *de novo* assembled into 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
overlapping and unique sets (Fig. 3a, b). Among these comparisons, shoot/JFL contained the most DEGs (3,706), whereas the smallest number of DEGs (64) was detected in JSL/ASL. Among the DEGs, 2,186, 1,621 and 78 were up-regulated, while 1,835, 1,679 and 332 were down-regulated between comparisons of shoot/JFL, shoot/AFL, and JFL/AFL during the development of floating leaves, respectively. During the development of submerged leaves, 2,413, 2,242 and 45 genes were found to be up-regulated, and 1,458, 1,284 and 33 were down-regulated upon analysis of shoot/JSL, shoot/ASL, and JSL/ASL, respectively (Fig. 3c). These results indicated that the greatest differential expression occurred in the early stages of development for both floating and submerged leaves (shoot/JFL and shoot/JSL). In addition, the number of up-regulated DEGs was more than the number of down-regulated DEGs during the development of submerged leaves. However, the number of up-regulated and down-regulated DEGs was basically the same during the development of floating leaves.

Based on the functional annotations of the DEGs, the numbers of DEGs annotated in shoot/JSL and shoot/ASL were greater than those in JSL/ASL upon comparisons of the development of floating leaves (Fig. 4). A similar pattern was found with comparisons of the development of the submerged leaves (Fig. 5). The highest number of DEGs for each comparison was found for “biological process”, while the most abundant terms were “cellular process”, “metabolic process”, and “single-organism” for different stages of submerged and floating leaf development. The enriched GO terms during the early leaf developmental stage (JFL and JSL) included “translation”, “regulation of transcription, DNA-templated”, “RNA methylation” and “photosynthesis”. However, the GO terms enriched during the later leaf developmental stage were assigned into “response to salt stress” (AFL and ASL), “salicylic acid biosynthetic process” (AFL) and “negative regulation of programmed cell death” (ASL). In the
cellular component category, “cell” and “cell part” were the two most highly represented components throughout the development of all leaves. The GO terms “ribosome” and “cytosolic small ribosomal subunit” were enriched in different leaf developmental stages, while the GO terms “cytoplasmic membrane-bounded vesicle” (JFL) and “chloroplast envelope” (JSL) were enriched in the early leaf developmental stages. In the molecular function category, the DEGs mapped to “catalytic activity” and “binding” were present in high proportion in all developmental stages.

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

Analysis of transcription factors associated with heterophyllous leaf types

We queried *P. octandrus* TF genes in the Plant Transcription Factor Database (http://planttfdb.cbi.pku.edu.cn/) to uncover the function of transcription factors (TFs) during the development of heterophyllous leaves. We identified 1,681 putative TF genes that could be 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
the signaling of abscisic acid (ABA), auxin, cytokinin (CTK), ethylene, jasmonic acid (JA),

gibberellin (GA), brassinosteroid and salicylic acid (SA).

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

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

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

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

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

Expression patterns of genes mapped to “stomatal complex morphogenesis” and “cuticle development”

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

A total of 11 DEGs were mapped to “cuticle development”, four of which were up-

https://doi.org/10.7287/peerj.preprints.3336v1 | CC BY 4.0 Open Access | rec: 12 Oct 2017, publ: 12 Oct 2017
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 morphogenesis 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 heterophyll 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
indicated that in the ABA signaling pathway, the genes encoding the PYR/PYL ABA receptor, which interacts with PP2C phosphatases, were up-regulated in floating leaves. The genes encoding SnRF2 proteins were also up-regulated, and the activated SnRF2 proteins can phosphorylate downstream targets, such as AREB/ABF transcription factors (Cutler et al., 2010). Since endogenous levels of ABA have been shown to increase in the leaves of water-stressed terrestrial plants and ABA plays a pivotal role in drought stress in terrestrial plant species (Walton & Li, 2013), ABA might have a similar function in the initiation of heterophyllous leaves in response to the transition from submerged to aerial conditions in *P. octandrus* (Goliber & Feldman, 1989). Moreover, plant cells respond to environmental stimuli through a series of intracellular signals. To minimize transpirational water loss (Hirayama & Shinozaki, 2010) ABA controls stomatal closure. This is brought about by transcriptional reprogramming via the ABA signalosome complex (PYP/PYL-PP2C-SnRK2). With the increasing of the concentration of ABA, the ABA signaling complex (PYP/PYL-PP2C-SnRK2) can cause the stomatal closure in guard cells through mediating the calcium-independent manner and some key biochemical messengers (Geiger et al., 2011).

Many enzymes in ABA biosynthesis are also induced during the drought stress response. Among these, ZEP, NCED and AAO are regarded as key enzymes (Iuchi et al., 2001; Chen et al., 2011; Wanke, 2011). Our results revealed that 19 genes encoding ZEP, NCED and AAO showed different expression levels. Six of the 19 genes were highly expressed in floating leaves and up-regulated when compared to the leaf shoot and submerged leaves, while only one gene was down-regulated, suggesting that differential expressions of genes that encode key enzymes 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 heterophyllly in aquatic
plants and it also is involved in the formation of submerged leaves; however, it induces
heterophyllly 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.
ocandrus 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 HOMEBOX (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 KNOXI genes is involved in
regulating heterophylly in P. octandrus.

bZIP TFs regulate a variety of plant development and abiotic resistance processes.
AthZIP1 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
morphological differences between heterophyllous leaves may be directed by genes of multiple functional groups, such as bZIP genes.

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

Conclusions

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

REFERENCES


Allsopp A. 1962. The effects of gibberellic acid on morphogenesis in *Marsilea drummondii*.

*Phytomorphology* 12: 1-10.


Sculthorpe CD. 1967. The Biology of Aquatic Vascular Plants. Arnold E.


**ADDITIONAL INFORMATION AND DECLARATIONS**

**Funding**

This study was supported by the Specialized Research Fund for the Doctoral Program of Higher Education (20110141110019), the National Natural Science Foundation of China (Nos. 31270278 and 31570220), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDPB02), and the Wuhan Botanical Garden, Chinese Academy of Sciences (No. 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.
Fig. 7 Distribution of transcription factor families among the three clusters (G1, G2 and G3).

Supplemental files

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

Additional file 2: Aligning statistics of clean reads with assembled unigenes. Q30 percentage is proportion of nucleotides with quality value larger than 30 in reads; GC percentage is proportion of guanine and cytosine nucleotides among total nucleotides.

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

Additional file 4: Clusters of orthologous groups (COG) classifications for *P. octandrus*. In total, 20,856 of the 48,235 sequences with Nr hits were groups into 25 classifications.

Additional file 5: Clusters of euKaryotic orthologous groups (KOG) classifications for *P. octandrus*. In total, 30,177 of the 48,235 sequences with Nr hits were groups into 25 classifications.

Additional file 6: A list of 6822 differentially expressed genes (DEGs) at each development stage. We defined genes with expression change of $|\log_2 \text{FC}| \geq 1$ and FDR < 0.01 as DEGs. The FPKM of two stages, value of log2FC and expression pattern of DEGs are presented in the table.

Additional file 7: A list of KEGG pathways mapped by 1,490, 1,141, 135, 1,425, 1,362 and 30 DEGs in all comparisons, respectively.
Additional file 8: A list of putative transcription factors (TF) differentially expressed during heterophyllous leaves development. 469 putative TF genes are identified into 42 families.

Additional file 9: Representative functions of TF genes showing expression gradients between floating leaf and submerged leaf tissue samples.

Additional file 10: A list of 70 DEGs involved in the plant hormone signal transduction pathway during heterophyllous leaves development. We defined genes with expression change of $|\log_{2}FC| \geq 1$ and FDR $< 0.01$ as DEGs. The values of log2FC and expression pattern of DEGs are presented in the table.

Additional file 11: The candidate DEGs IDs, annotations and the primers used for qPCR analysis.

Additional file 12: Quantitative real-time polymerase chain reaction (qPCR) validation of RNA-seq data. (a) Expression profiles of the candidate genes by qPCR. The value of relative expression level on the y-axis calculated according to $\log(2^{\Delta\Delta CT})$. The expression level of each gene in shoot was arbitrarily set as 1, the value of the y-axis was 0. (b) Correlation of gene expression results. The x-axis represents the value of Log2 fold change by RNA-seq and the y-axis represents the value of Log2 relative expression level by qPCR.
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
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
**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.
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.
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.
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.
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).
Fig. 7
Table 1 Overview of the sequencing and assembly from RNA-Seq data.

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

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