Investigating the molecular basis for heterophylly in the aquatic plant Potamogeton octandrus (Potamogetonaceae) with comparative transcriptomics

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

[1] Laboratory of Plant Systematics and Evolutionary Biology, 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 different leaf morphologies within a single plant, or heterophylly. 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, at different developmental stages, were sequenced using high-throughput sequencing (RNA-Seq), in order to aid gene discovery and functional studies of genes involved in heterophyllous leaf development. A total of 81,103 unigenes were identified in submerged and floating leaves and 6,822 differentially expressed genes (DEGs) were identified by comparing samples from different developmental stages. KEGG pathway enrichment analysis categorized these unigenes into 128 pathways. A total of 24,025 differentially expressed genes were involved in carbon metabolic pathways, biosynthesis of amino acids, ribosomal processes, and plant-pathogen interactions. In particular, KEGG pathway enrichment analysis categorized a total of 70 DEGs into plant hormone signal transduction pathways. The high-throughput transcriptomic results presented here highlight the potential for understanding the molecular mechanisms underlying heterophylly, which is still poorly understood. Further, these data provide a framework to better understand heterophyllous leaf development in P. octandrus via targeted studies utilizing gene cloning and functional analyses.

Investigating the molecular basis for heterophylly in the aquatic plant *Potamogeton octandrus* (Potamogetonaceae) with comparative transcriptomics

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, 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:

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 different leaf morphologies within a single plant, or heterophylly. 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, at different developmental stages, were sequenced using high-throughput sequencing (RNA-Seq), in order to aid gene discovery and functional studies of genes involved in heterophyllous leaf development. A total of 81,103 unigenes were identified in submerged and floating leaves and 6,822 differentially expressed genes (DEGs) were identified by comparing samples from different developmental stages. KEGG pathway enrichment analysis categorized these unigenes into 128 pathways. A total of 24,025 differentially expressed genes were involved in carbon metabolic pathways, biosynthesis of amino acids, ribosomal processes, and plant-pathogen interactions. In particular, KEGG pathway enrichment analysis categorized a total of 70 DEGs into plant hormone signal transduction pathways. The high-throughput transcriptomic results presented here highlight the potential for understanding the molecular mechanisms underlying heterophylly, which is still poorly understood. Further, these data provide a framework to better understand heterophyllous leaf development in *P. octandrus* via targeted studies utilizing gene cloning and functional analyses.

INTRODUCTION

The presence of two or more leaf forms on a single plant, or heterophylly, is widely observed across diverse plant species, and 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

45 their fitness by decreasing leaf damage, decreasing water loss, enhancing photosynthesis, or
46 promoting sexual reproductive success (Winn, 1999a, b; Wells & Pigliucci, 2000; Minorsky,
47 2003; Zhang et al., 2009; Zotz, Wilhelm & Becker, 2011). Accordingly, heterophylly has been
48 used as a model system for studying gene-environment interactions (Pigliucci, 2010; Nakayama
49 et al., 2014).

50 In the past century, numerous studies have described morphological changes in
51 heterophyllous plants in response to environmental factors such as CO_2 concentration, oxygen
52 capacity, salt concentration, temperature, water levels, seasonal changes, and light intensity and
53 quality (McCallum, 1902; Arber, 1920; Fassett, 1930; Sculthorpe, 1967; Cook & Johnson, 1968;
54 Bodkin, Spence & Weeks, 1980; Deschamp & Cooke, 1984; Titus & Sullivan, 2001). More
55 recent studies have also revealed that plant hormones, including ethylene, abscisic acid (ABA),
56 and gibberellin (GA) can affect heterophyllous leaf formation in many plant species, including
57 *Potamogeton nodosus* (Anderson, 1978), *Hippuris vulgaris* (Kane & Albert, 1987), *Marsilea*
58 *quadrifolia* (Liu, 1984), *Callitriche heterophylla* (Deschamp & Cooke, 1985), *Ranunculus*
59 *flaellaris* (Young & Horton, 1985; Young, Dengler & Horton, 1987), *Ludwigia arcuata*
60 (Kuwabara, 2003), and *Rorippa aquatica* (Nakayama et al., 2014). For example, ethylene gas has
61 been shown to induce the development of submerged-type leaves on terrestrial shoots of *L.*
62 *arcuate*. ABA, however, induces the formation of terrestrial-type leaves on submerged shoots
63 (Kuwabara, 2003). However, the molecular mechanisms regulating the alterations of leaf forms
64 in heterophyllous plant species remain largely unclear.

65 Plant leaves arise from a group of initial cells termed shoot apical meristems (SAMs). The
66 molecular mechanisms responsible for the initiation and maintenance of SAMs and polar
67 processes of leaf expansion have been recently studied in detail (Hay & Tsiantis, 2006; Tsukaya,
68 2006; Uchida et al., 2007, 2010; Shani et al., 2010; Moon & Hake, 2011). For example, SAMs
69 are 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 (Moon & Hake, 2011). However, few studies have investigated the molecular biological changes that occur during the course of heterophyllous transitions. Hsu *et al.* (2001) identified several early-response ABA-regulated genes, designated *ABRH* (ABA-responsive heterophylly), in the aquatic fern *M*. *quadrifolia*. *ABRH* genes encode transcription factors, protein kinases, membrane transporters, metabolic enzymes, and structural proteins. In addition, Chen *et al.* (2011) demonstrated that 9-*cis*-epoxycarotenoid dioxygenase 3 (*NCED3*) plays a key role in regulating ABA-mediated heterophylly via endogenous ABA in two different lily varieties. Nakayama *et al*. (2014) also showed that regulation of GA levels via *KNOX1* genes is primarily responsible for modulating heterophylly in *R*. *aquatica*. Heterophylly occurs across diverse taxa and may have evolved through convergent evolution (Minorsky, 2003), and thus different developmental processes and molecular mechanisms may exist among different species. Further, heterophylly within a single plant is controlled through multiple signalling pathways (Lin & Yang, 1999; Hsu et al., 2001). Thus, additional study systems in diverse heterophyllous plants and the use of large comparative datasets generated at the whole genomic or transcriptomic levels would aid in elucidating the complex molecular mechanisms that regulate heterophylly.

Here, we investigated the molecular mechanisms underlying heterophylly in *Potamogeton* (Potamogetonaceae), a genus comprises more than eight species being heterophyllous (Guo et al., 2010). *P*. *octandrus* Poir (2n = 28), a heterophyllous pondweed with submerged and floating leaf forms, is a perennial aquatic herb that is self-compatible and can reproduce vegetatively through rhizomes, or sexually by selfing and outcrossing seeds. *P. octandrus* produces many floating leaves that are flat and ovate with sharp leaf tips 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 the stem apex can produce both floating and submerged leaves as the stem elongates (Fig.1b).

The transcriptomes of submerged and floating leaf materials of *P. octandrus* at different developmental stages were sequenced using the Illumina RNA-Seq platform. *De novo* assembly of the DNA reads generated from submerged and floating leaf materials were then used as the reference transcriptome, and short sequence reads that were generated from each developmental stage of the submerged and floating leaf were mapped to the assembled transcriptome in order to identify genes exhibiting differential expression between leaf morphologies. These results provide a crucial reference transcriptome for investigating the regulatory mechanisms underlying each leaf form, in addition to a list of candidate genes that are likely to be involved in heterophylly development. Further, these results add to the growing body of data aimed at understanding the regulatory mechanisms of heterophylly among heterophyllous plants.

MATERIALS AND METHODS

Plant materials

A *P. octandrus* plant was collected from a Tongcheng population (29°16’05.6” N, 113°48’46.9” E) in the Hubei Province, China, and used for transcriptomic analysis. The plant was transplanted to a greenhouse at Wuhan University in April 2015 and the seeds were harvested in August 2015. Seedlings were cultured in a pool at Wuhan University and floating and submerged leaf forms were sampled at the following developmental stages that we have defined as follows: (1) juvenile floating leaves with length less than 0.5 cm (JFL) (72 hours); (2) adult floating leaves with length more than 1 cm (AFL) (84 hours); (3) juvenile submerged leaves with length less than 1.5 cm (JSL) (96 hours); (4) adult submerged leaves with length more than 3 cm (ASL) (120 hours). An additional developmental stage was also included: (5) leaf-shoots of

plants with stems that have an apex just reaching the water surface (shoot) (24 hours), where identification of which leaf form it will develop into is not possible at this stage. The "shoot" was considered as the initial stage of either floating or submerged leaves. Fifteen samples with three biological replicates for each developmental stage were collected in total. Sampled tissues were immediately frozen in liquid nitrogen and stored at −80°C until further analysis.

RNA extraction, cDNA library construction and sequencing

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

cDNA libraries were then assembled for the 15 RNA samples (T01-T15). cDNA was synthesized using a SuperScript VILO cDNA Synthesis Kit (Invitrogen., Carlsbad, CA, USA) according to the manufacturer's protocol. A total of 10 μg of RNA was purified using oligo (dT) magnetic beads in order to enrich for poly (A) mRNA. Fragmentation buffer was then added to split mRNA into short fragments. Fragments were used as templates to synthesize first-strand cDNA using random hexamer-primers (Invitrogen, Carlsbad, CA, USA) and 1000 U SuperScriptTM II reverse transcriptase (Invitrogen, Carlsbad, CA, USA). RNase H and DNA polymerase I were used to synthesize second-strand cDNA. The short fragments were then amended with adapters and end repair ligation. The products were first purified using a QIAquick® Polymerase Chain Reaction (PCR) Extraction Kit (Qiagen, Hilden, Germany), and then enriched with PCR (15 cycles) to create the final cDNA libraries. The 15 cDNA libraries were then sequenced to obtain 150 bp paired-end reads using the Illumina HiSeq™ 2500

143 sequencing platform (Illumina Inc., San Diego, CA, USA). The raw sequencing data was
144 deposited into the NCBI Short Reads Archive (SRA) under the accession number SRR3184674.

145 **Sequence assembly and gene annotation**

146 The raw reads were cleaned by removing the low quality reads, adapter reads, and reads
147 with 5% or more unknown nucleotides. Transcriptome *de novo* assembly was performed with the
148 clean reads that resulted from all 15 samples using the Trinity program (Grabherr et al., 2011)
149 with the min_kmer_cov parameter set to 2 by default. Overlapping sequence reads were first
150 combined to form contigs without gaps, and reads were then reverse mapped to the contigs.
151 Paired-end sequencing allowed the inference of contigs arising from the same transcript, and
152 distances between these contigs were determined. Contigs were connected in Trinity and
153 sequences that could not be extended on either end were identified, followed by construction of
154 'unigenes' that represented transcripts from the same locus. The completeness of *P. octandrus*
155 transcriptome assembly was estimated by reciprocal BLAST search against the annotated
156 proteins from the seagrass genome *Zostera marina* (a closely related species of *P. octandrus*)
157 (Olsen et al., 2016) with E-value threshold of 1e-5.
158 The putative functions of unigenes were then assessed using BLASTx (E-value $\leq 10^{-5}$) and
159 several protein databases including the: NCBI non-redundant protein (nr), Swiss-Prot, Clusters of
160 Orthologous Groups (COG), euKaryotic Orthologous Groups (KOG), eggNOG, Protein family
161 (Pfam), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG)
162 databases. The sequencing direction of unigenes was determined based on the best alignments.
163 ESTScan program (version 3.0.1, http://www.ch.embnet.org/software/ESTScan.html) was used to
164 determine sequencing direction when a unigene could not be aligned with any of the above
165 databases. Blast2GO 2.5.0 (Conesa et al., 2005) was used to compare and determine unigene
166 Gene Ontology (GO) annotations. 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)

To quantify the abundance of transcripts, clean reads from each sample were then mapped onto the assembled reference transcriptome using the Bowtie2 read-mapping, alignment software (Langmead & Salzberg, 2012). Gene expression levels and comparisons of gene fragments per kb per million (FPKM) fragments values (Marioni et al., 2008) were then assessed using likelihood ratio tests in order to statistically identify differentially expressed genes (DEGs). DEGs were assessed between samples from different developmental stages (control/experimental: shoot/JFL, shoot/AFL, JFL/AFL, shoot/JSL, shoot/ASL, JSL/ASL), a previously developed algorithm was applied (Audic and Claverie, 1997). False discovery rate (FDR) control was used to correct for *P* values for multiple hypotheses. Genes with changes in expression that were two-fold or greater and had FDR values ≤ 0.01 were identified as differentially expressed and then absolute values of $\log_2$ (foldchange) ≥ 1 were regarded as DEGs.

GO functional analysis was used to assess GO functional enrichment and GO functional annotation for the DEGs. All DEGs were mapped to each term of the GO database. We then calculated the gene abundances for each GO term. A hypergeometric test was then applied to identify significantly enriched GO terms in DEGs relative to the genome background. *P*-values from the GO enrichment analysis were adjusted using Bonferroni's correction. A corrected *P*-value ≤ 0.05 was selected as the threshold for significantly enriched GO terms. DEGs were then mapped to terms in the KEGG database to obtain enriched pathway terms (Kanehisa & Goto, 2000). Pathways with an FDR value of ≤ 0.05 were considered significant.

Quantitative real-time PCR (qPCR) analysis

To validate the transcriptomic data, 16 DEGs were randomly selected and their expression profiles were investigated by qPCR. Total RNA was extracted from *P. octandrus* leaves at five development stages and cDNA libraries were constructed as described above. A SYBR Premix Ex Taq™ Kit (TaKaRa) was used for the qPCR reactions in a BioRad Real-Time thermal cycler system (BioRad, Hercules, CA, USA). Gene-specific primers were designed using Primer Premier5, and the primer sequences are provided in File S5. The PCR cycling regimen was developed according to the manufacturer's protocol and consisted of 95°C for 30s, 40 cycles of 95°C for 5 s and 58°C for 30 s. Each reaction was performed in triplicate and *β*-actin was used as a reference gene. The $2^{-\Delta\Delta Ct}$ method was used to calculate 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 reads with ambiguous bases, a total of 135.83 Gbp clean sequence data with Q30 ≥ 96.0% were obtained from the 15 cDNA libraries. 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). The most abundant length fraction of Unigenes were 201–300 bp, followed by 500–1,000 bp and 300–500 bp sequences (Table 1). Cleaned reads were then mapped to the unigenes to assess sequencing randomness, which was found to be sufficient (Additional file 1). Further statistics of read mapping for each sample are provided in Additional file 2. In the test of completeness of the assembly of the transcriptome, by homologous search, about 91.1% (18,738/20,559) of all annotated seagrass proteins could be identified in our transcriptome. This indicates that our transcriptome assembly may cover majority of the *P. octandrus* transcriptome. Meanwhile, only 60.5% (101,218/ 167,270) unique *P. octandrus* transcripts could be annotated by seagress

proteins, which imply that the rest transcripts are from lineage-specific genes or long non-coding RNAs.

Functional annotation and classification of unigenes

A total of 48,235 (56% of all unigenes) consensus sequences were annotated using BLASTx (*E*-value ≤ 1×10^{-5}) against several public protein databases including NCBI-nr, Swiss-Prot, COG, KOG, eggNOG, Pfam, GO, and KEGG (Additional file 3). Among the annotated unigenes, 24,025 were categorized into 52 functional subcategories under three main GO categories: cellular components, molecular functions, and biological processes (Fig. 2). The GO category "biological processes" represented the largest number of GO annotations, with metabolic processes, cellular processes, and single-organism processes comprising the three most abundant top-level subcategories. In the cellular components category, the "cell" and "cell part" categories were most represented, while in the molecular functions category, the "catalytic activity" subgroup was prominent, followed by the "binding" group. In the COG and KOG functional classifications, "translation, ribosomal structure and biogenesis" associated proteins represented the largest group (Additional files 4 and 5). Lastly, a total of 22,346 unigenes were assigned to 128 KEGG pathways.

DEG identification and functional analysis

A total of 6,822 differentially expressed genes (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 visualize the numbers of overlapping and unique DEGs among and within groups (Fig. 3a, b). The shoot/JFL comparison contained the most DEGs (3,706), whereas the smallest number of DEGs (64) was detected in the JSL/ASL comparison. Among the DEGs associated with the development of floating leaves, 2,186, 1,621 and 78 were up-regulated, while 1,835, 1,679 and 332 were down-

238 regulated in comparisons of shoot/JFL, shoot/AFL, and JFL/AFL, respectively. During the
239 development of submerged leaves, 2,413, 2,242 and 45 DEGs were found to be up-regulated,
240 while 1,458, 1,284 and 33 were down-regulated in the shoot/JSL, shoot/ASL, and JSL/ASL
241 comparisons, respectively (Fig. 3c). These results indicated that the greatest differential
242 expression occurred in the early stages of development for both floating and submerged leaves
243 (shoot/JFL and shoot/JSL). Further, the number of up-regulated DEGs was higher than the
244 number of down-regulated DEGs during the development of submerged leaves. However, the
245 number of up-regulated and down-regulated DEGs was essentially equivalent during the
246 development of floating leaves.

247 The number of DEGs that were annotated in the shoot/JSL and shoot/ASL comparisons
248 during floating leaf development was greater than those in the JSL/ASL comparison (Fig. 4). A
249 similar pattern was found in comparisons for the development of submerged leaves (Fig. 5). The
250 highest number of DEGs for each comparison was identified in the “biological process” category,
251 while the most abundant sub-group terms were in the “cellular process”, “metabolic process” and
252 “single-organism” categories for different stages of submerged and floating leaf development.
253 Enriched GO terms during the early leaf developmental stages (JFL and JSL) included
254 “translation”, “regulation of transcription, DNA-templated”, “RNA methylation”, and
255 “photosynthesis”. However, the GO terms enriched during the later developmental stages were
256 assigned to “response to salt stress” (AFL and ASL), “salicylic acid biosynthetic process” (AFL),
257 and “negative regulation of programmed cell death” (ASL). In the cellular component category,
258 “cell” and “cell part” were the two most highly represented components throughout the
259 development of all leaves. The GO terms “ribosome” and “cytosolic small ribosomal subunit”
260 were enriched in different leaf developmental stages, while the GO terms “cytoplasmic
261 membrane-bounded vesicle” (JFL) and “chloroplast envelope” (JSL) were enriched in the early
262 developmental stages of leaves. In the molecular function category, DEGs that mapped to

263 "catalytic activity" and "binding" were present in high proportions through all developmental
264 stages.

265 Annotated sequences were then searched against the KEGG pathways database. Among the
266 annotated DEGs, 1,490, 1,141, 135, 1,425, 1,362, and 30 were assigned to 114, 114, 57, 114, 111,
267 and seven pathways in the shoot/JFL, shoot/AFL, JFL/AFL, shoot/JSL, shoot/ASL, and JSL/ASL
268 comparisons, respectively (Additional file 7). The pathways with the largest proportions of DEGs
269 were "ribosome", "biosynthesis of amino acids", "carbon metabolism", and "plant-pathogen
270 interaction" in the shoot/JFL, shoot/AFL, shoot/JSL, and shoot/ASL comparisons, respectively
271 (Additional file 7). However, these four pathways accounted for a lower percentage of the
272 proportion in the JFL/AFL and JSL/ASL groups, which may be important in the early stages of
273 development of floating and submerged leaves. "Plant hormone signal transduction" was the
274 most enriched cluster among the DEGs, indicating that proteins in this category may be important
275 in the morphological and physiological differentiation that occurs in the early stages of leaf
276 development.

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

278 To assess the function of transcription factors (TFs) during the development of
279 heterophyllous leaves, we identified *P. octandrus* TF genes using the Plant Transcription Factor
280 Database (http://planttfdb.cbi.pku.edu.cn/). In total, we identified 1,681 putative TF genes that
281 were categorized into 48 families. Four-hundred and sixty-nine of the 1,681 putative TF genes
282 that were from 42 families showed significant differential expression between developmental
283 stages (Fig. 6 and Additional file 8). Most of these genes (53%) were expressed at the highest
284 levels in the AFL, ASL, JFL, and JSL groups (Group 2; G2). Only 37% were expressed at the
285 highest levels in the "shoot" group (Group 1; G1). Lastly, an additional 10% exhibited highest
286 expression in the JFL group (Group 3; G3). We further investigated differential expression at the

family-specific level (Fig. 7). The WRKY, C3H, and AP2 families of TFs, which play roles in various processes including stress response, leaf senescence, and leaf epidermal cell identity, were highly expressed in G2 (Additional file 9). Many TFs that mediate hormone signaling, including ethylene signaling (ERF family) and auxin signaling (NAC family), were highly expressed during the development of floating and submerged leaves (G2). The HD-Zip family members that have been proposed as regulators of vascular development, stomatal complex morphogenesis, leaf polarity, and epidermal cell differentiation, were most highly expressed in G2. MYB, ARF, and B3 TFs are involved in light and hormone signaling pathways. These genes were most expressed in the "shoot" (G1) group, where cell differentiation and cellular morphogenesis are modulated. Homeobox genes, which participate in a number of developmental events, were also highly expressed in G1. Several FAR1 and bHLH TFs, which have been reported to function in light signaling and stomatal development, were enriched in juvenile floating leaves (G3). These results suggest that transcriptional regulatory genes are critical for a wide variety of developmental processes in the leaf transcriptome.

Hormone metabolism and signalling pathways among heterophyllous leaf samples

KEGG annotation revealed that most DEGs were classified in the "plant hormone signal transduction" category. A total of 70 DEGs were annotated as diverse hormone-related genes, including those involved in signalling with abscisic acid (ABA), auxin, cytokinin (CTK), ethylene, jasmonic acid (JA), gibberellin (GA), brassinosteroid, and salicylic acid (SA).

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

Eight genes that were associated with ABA and GA pathways 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 that were related to GA and ethylene pathways were also highly expressed in submerged leaves, including one that encoded gibberellin receptor (GID1) and two encoding the ethylene-responsive transcription factor 1 (ERF1).

Among the 31 hormone-related genes that were up-regulated during the 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 stimulation, 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 significantly up-regulated in the early developmental 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 that were annotated to encode zeaxanthin epoxidase (ZEP), 9-*cis*-epoxycarotenoid dioxygenase (NCED), and abscisic-aldehyde oxidase (AAO), exhibited different levels of expression. Genes encoding ZEP and NCED were highly expressed during the development of submerged leaves and particularly highly expressed in floating leaves. Further, 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), which is a key enzyme in GA biosynthesis, was down-regulated throughout leaf

development, except during the later stage of floating leaf development. Additionally, two genes that encode gibberellin 2beta-dioxygenase (GA2ox), which has an inactive effect on GAs, were down-regulated. These results suggested that GA signaling pathways are important in early stages of leaf development.

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

Of the genes that were functionally annotated, some were associated with the functional groups "stomatal complex morphogenesis" and "cuticle development" throughout the leaf development process. Twenty-three DEGs were associated with "stomatal complex morphogenesis", including 12 that were up-regulated during the development of floating leaves, and particularly highly expressed in juvenile floating leaves. Three DEGs were also down-regulated in submerged leaves. The up-regulated DEGs were annotated 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-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 indicate regulation of leaf morphology and structure.

Validation of DEGs expression by qPCR

Sixteen candidate DEGs were selected for qPCR validation of expression levels. These genes were categorized into groups that were 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 that were involved in biological processes were analyzed, 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 provided in Additional file 11. The correlation between RNA-Seq results (fold change) and qPCR results ($2^{-\Delta\Delta CT}$) was measured by plotting the $\log_2$ fold changes (Additional file 12a). The results suggested that the qPCR abundances of genes were significant similar ($r^2 = 0.65$, $P < 0.01$) to the RNA-Seq data (Additional file 12b).

DISCUSSION

A growing body of research has indicated that heterophylly is a common, adaptive response by plants to increase their fitness through responses to environmental changes. However, the molecular mechanisms underlying these processes have remained unclear. Comparative transcriptomic analysis is an efficient method for discovering genes and investigating biochemical pathways that are involved in physiological processes, and has recently become an increasingly tractable methodology (Varshney et al., 2009; Ozsolak & Milos, 2011; Shi et al., 2011; Mutasa-Goettgen et al., 2012; Yang et al., 2014). Here, leaves of the aquatic heterophyllous plant *P. octandrus* that has two leaf forms (submerged and floating), were subjected to transcriptomic sequencing to identify differentially expressed genes among the two leaf morphologies in different developmental phases. Overall, 81,103 unigenes were assembled and 48,235 unigenes were annotated against public protein databases. A total of 6,822 differentially expressed genes (DEGs) were identified in comparisons between developmental stages. KEGG pathway enrichment analysis was used to sort a number of DEGs into plant hormone signal transduction pathways, including Cytokinin, Auxin, Abscisic Acid, Gibberellin, Brassinosteroid, Ethylene, and Jasmonic Acid related pathways.

The initial stage of leaf shoot development is enriched in metabolic processes, cellular processes, and responses to stimuli. The expression levels of genes related to leaf morphogenesis, photomorphogenesis, and hormone signaling increased during the development of submerged and floating leaves. This was especially evident in juvenile floating leaves, where 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 in other tissues. These gene transcriptional regulations were coincident with the developmental dynamics of heterophyllous leaves.

When grouped into KEGG pathways, most DEGs were associated with hormone signaling. Exogenous and endogenous ABA and GA can regulate leaf form alteration in heterophyllous plant species (Allsopp, 1962; Deschamp & Cooke, 1984; Gee & Anderson, 1996; Kuwabara, 2003). The ABA-mediated regulation of heterophyllous morphological changes has been studied intensively (Wanke, 2011; Nakayama et al., 2012). Here, a large number of hormone-related DEGs were associated with ABA and GA signal transduction pathways, suggesting that genes responsive to ABA and GA might play an important role in heterophyllous leaf formation in *P. octandrus*. The genes encoding the PYR/PYL ABA receptor, which interacts with PP2C phosphatases and is a component of the ABA signaling pathway, were up-regulated in floating leaves. Further, genes encoding SnRF2 proteins were also up-regulated. Activated SnRF2 proteins can phosphorylate downstream targets, such as AREB/ABF transcription factors and contribute to those signaling pathways (Cutler et al., 2010). Endogenous levels of ABA have been shown to increase in leaves of water-stressed terrestrial plants and ABA plays a pivotal role in drought stress in terrestrial plant species (Walton & Li, 2013). Thus, ABA may 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). Plant cells respond to environmental stimuli through a series of intracellular signals. To minimize transpirational water

loss, ABA controls stomatal closure (Hirayama & Shinozaki, 2010). This change is induced by transcriptional reprogramming via the ABA signalosome complex (PYP/PYL-PP2C-SnRK2). With increasing concentrations of ABA, the ABA signaling complex (PYP/PYL-PP2C-SnRK2) can cause stomatal closure in guard cells in a calcium-independent manner and through key biochemical messengers (Geiger et al., 2011).

Many enzymes in ABA biosynthesis are also induced during the drought stress response, including the key enzymes ZEP, NCED, and AAO (Iuchi et al., 2001; Chen et al., 2011; Wanke, 2011). Our results indicated that 19 genes encoding ZEP, NCED, and AAO exhibited different expression levels among leaf morphological groups. Six of the 19 genes were highly expressed in floating leaves and up-regulated when compared to leaf shoots and submerged leaves, while only one gene was down-regulated. These results suggest that differential expression of genes that encode key enzymes in ABA biosynthesis might control ABA function at specific stages during the initiation of heterophyllous leaf changes. Thus, ABA may play a complex role in signalling transduction during heterophylly development in *P. octandrus*.

Recently, ethylene has been suggested to influence the formation of heterophyllous leaves in *L. arcuata* in a manner that is opposite of ABA, and also has been suggested to be an endogenous factor that induces the formation of submerged leaves (Kuwabara, Tsukaya & Nagata, 2001). Our results indicated that the expression 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 leaf shoots and submerged leaves. GA likely has an antagonistic effect on heterophylly in aquatic plants and is also 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 feedback antagonistically on ethylene and GA (Wanke, 2011). Our results also indicated that two genes encoding GA2ox, which has an inactive effect on GA, were

435 down-regulated in submerged leaves. Taken together, these findings suggest that ABA signalling
436 was enhanced in the aerial leaves of heterophyllous plant species.

437 Transcription factors (TFs) that target promoter regions regulate the concentrations of
438 proteins as limiting factors. They also play an important role in responses to environmental stress
439 (Yuan & Perry, 2011) and plant development. Previous studies have demonstrated that many
440 genes in the AP2/EREBP family participate in the transcriptional regulation of processes related
441 to growth and development. For example, one member of the ERF gene family, *ESR1*, regulates
442 shoot regeneration (Banno et al., 2001), while overexpression of *SHN* gene that are AP2/EREBP
443 transcription factors, increases cuticular wax (Aharoni et al., 2004). Further, the *Glossy15* gene
444 from maize regulates leaf epidermal cell identity (Moose & Sisco, 1996), and the *LEAFY*
445 *PETIOLE* (*LEP*) gene influences leaf petiole development in *Arabidopsis thaliana* (van der
446 Graaff et al., 2000). Here, we found that 48 TFs are predominantly expressed in *P. octandrus* and
447 42 TFs are up- or down-regulated among morphological comparisons. Among these TFs, several
448 ERF homologs to SHN and LEAFY PETIOLE exhibited up-regulation during floating leaf
449 development, which is consistent with thicker waxy cuticles and extended leaf petioles on
450 floating leaves of *P. octandrus*. In addition, we detected a close homolog of maize *Glossy15*,
451 which was highly expressed in floating and submerged leaves, suggesting a similar function of
452 ERF in leaves of *P. octandrus* and maize.

453 The KNOX homeobox protein (Knotted1-like homebox, KNOX) is a homeodomain
454 transcription factor that maintains cell pluripotency in plant shoot apical meristems (SAM)
455 (Vollbrecht et al., 1991). Three *KNOTTED1-like HOMEOBOX* (*KNOX*) homologs were detected
456 in our DEG analysis, and all of these were highly expressed in leaf shoots. KNOX proteins
457 regulate the homeostasis of CTK and GA in order to maintain meristematic cells in an
458 undifferentiated state (Shani, Yanai & Ori, 2006). CTK is a plant hormone involved in cell
459 proliferation while GA controls leaf morphogenesis (Hooley, 1994; Mok & Mok, 2001). These

460 results imply that regulation of GA levels by *KNOX1* genes is involved in regulating heterophylly
461 in *P. octandrus*.

462 bZIP TFs regulate a variety of plant development and abiotic resistance processes. For
463 example, *AtbZIP1* from *Arabidopsis* regulates ABA signal transduction by binding ABA-
464 responsive elements (ABREs) and altering the expression of ABA-responsive genes (Sun et al.,
465 2011). Here, three ABRE binding factors (AREB/ABF) were identified that were differentially
466 expressed between leaf shoots, floating leaves and submerged leaves. These findings suggest that
467 morphological differences between heterophyllous leaves may be directed by genes from
468 multiple functional groups, such as bZIP genes.

469 We also identified several other TFs involved in ABA signalling. For example, the NAC TF,
470 ANAC072, was detected and responds to exogenous ABA and may regulate ABA-responsive
471 genes (Tran et al., 2004). Moreover, ABA regulates gene expression through additional TFs such
472 as *MYB*, *HD-ZF*, *B3* and *bHLH* (Fujita et al., 2011). Our data indicate that diverse TFs may be
473 involved in heterophyllous leaf development and that the genes are both down-regulated and up-
474 regulated, suggesting that TFs may be involved in different processes during heterophyllous leaf
475 development.

476 **Conclusions**

477 Here, we describe the production, assembly, and annotation of transcriptomes from
478 submerged and floating leaves of a heterophyllous aquatic plant, *P. octandrus*. We identified
479 several genes that showed differential expression at different developmental stages during the
480 development of heterophylly in this species. These results inform on a more complete
481 understanding of the molecular mechanisms underlying heterophylly in plants. In particular, these
482 results can be used for cloning and functional studies of genes that are involved in leaf
483 development and heterophylly development in order to further probe the mechanistic basis for

484 these important developmental phenomena.

485 **REFERENCES**

486 Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A. 2004. The SHINE clade of AP2
487 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers
488 drought tolerance when overexpressed in *Arabidopsis*. *Plant Cell* 16: 2463-2480.

489 Allsopp A. 1962. The effects of gibberellic acid on morphogenesis in *Marsilea drummondii*.
490 *Phytomorphology* 12: 1-10.

491 Anderson LW. 1978. Abscisic acid induces formation of floating leaves in the heterophyllous
492 aquatic angiosperm *Potamogeton nodosus*. *Science* 201: 1135-1138.

493 Arber A. 1920. Water Plants: A Study of Aquatic Angiosperms. Cambridge University Press.

494 Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome*
495 *Research* 7: 986-995.

496 Banno H, Ikeda Y, Niu QW, Chua NH. 2001. Overexpression of *Arabidopsis* ESR1 induces
497 initiation of shoot regeneration. *Plant Cell* 13: 2609-2618.

498 Bodkin PC, Spence D, Weeks DC. 1980. Photoreversible control of heterophylly in *Hippuris*
499 *vulgaris* L. *New Phytologist* 84: 533.

500 Chen H, Hwang S, Chen S, Shii C, Cheng W. 2011. ABA-mediated heterophylly is regulated by
501 differential expression of 9-*cis*-epoxycarotenoid dioxygenase 3 in lilies. *Plant and Cell*
502 *Physiology* 52: 1806-1821.

503 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal
504 tool for annotation, visualization and analysis in functional genomics research.
505 *Bioinformatics* 21: 3674-3676.

506 Cook SA, Johnson MP. 1968. Adaptation to heterogenous environments I. Variation in
507 heterophylly in *Ranunculus flammula* L. *Evolution* 22: 496-516.

508 Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic acid: emergence of a core
509 signaling network. *Annual Review of Plant Biology* 61: 651-679.
510 Deschamp PA, Cooke TJ. 1984. Causal mechanisms of leaf dimorphism in the aquatic
511 angiosperm *Callitriche heterophylla*. *American Journal of Botany* 71: 319-329.
512 Deschamp PA, Cooke TJ. 1985. Leaf dimorphism in the aquatic angiosperm *Callitriche*
513 *heterophylla*. *American Journal of Botany* 72: 1377-1387.
514 Fassett NC. 1930. A Manual of Aquatic Plants. The University of Wisconsin Press.
515 Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K. 2011. ABA-mediated transcriptional
516 regulation in response to osmotic stress in plants. *Journal of Plant Research* 124: 509-525.
517 Gee D, Anderson L. 1996. ABA induced differences during leaf development in the aquatic
518 angiosperm, *Potamogeton nodosus*, are detected with differential display. *Plant Physiology*
519 111: 110.
520 Geiger D, Maierhofer T, Al-Rasheid KA, Scherzer S, Mumm P, Liese A, Ache P, Wellmann C,
521 Marten I, Grill E, Romeis T, Hedrich R. 2011. Stomatal closure by fast abscisic acid signaling
522 is mediated by the guard cell anion channel SLAH3 and the receptor RCAR1. *Science*
523 *Signaling* 4 (173): ra32.
524 Goliber T, Feldman L. 1989. Osmotic stress, endogenous abscisic acid and the control of leaf
525 morphology in *Hippuris vulgaris* L. *Plant Cell and Environment* 12: 163-171.
526 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
527 Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F,
528 Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length
529 transcriptome assembly from RNA-Seq data without a reference genome. *Nature*
530 *Biotechnology* 29: 130-644.

Guo YH., Haynes RR., Hellquist CB., Kaplan Z. 2010. Potamogetonaceae. In: Wu ZY, Peter HR. (Eds.) *Flora of China*. Volume 23. Science Press and Missouri Botanical Garden Press, Beijing and St. Louis, 91–102.

Hay A, Tsiantis M. 2006. The genetic basis for differences in leaf form between *Arabidopsis thaliana* and its wild relative *Cardamine hirsuta*. *Nature Genetics* 38: 942-947.

Hirayama T, Shinozaki K. 2010. Research on plant abiotic stress responses in the post-genome era: past, present and future. *The Plant Journal* 61:1041-1052.

Hooley R. 1994. Gibberellins: perception, transduction and responses. *Plant Molecular Biology* 26: 1529-1555.

Hsu TC, Liu HC, Wang JS, Chen RW, Wang YC, Lin BL. 2001. Early genes responsive to abscisic acid during heterophyllous induction in *Marsilea quadrifolia*. *Plant Molecular Biology* 47: 703-715.

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K. 2001. Regulation of drought tolerance by gene manipulation of 9-*cis*-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *The Plant Journal* 27: 325-333.

Jackson D, Veit B, Hake S. 1994. Expression of maize *KNOTTED1* related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development* 120: 405-413.

Kane ME, Albert LS. 1987. Abscisic acid induces aerial leaf morphology and vasculature in submerged *Hippuris vulgaris* L. *Aquatic Botany* 28: 81-88.

Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28: 27-30.

Kuwabara A, Ikegami K, Koshiba T, Nagata T. 2003. Effects of ethylene and abscisic acid upon heterophylly in *Ludwigia arcuata* (Onagraceae). *Planta* 217: 880-887.

556 Kuwabara A, Tsukaya H, Nagata T. 2001. Identification of factors that cause heterophylly in
557 *Ludwigia arcuata* Walt.(Onagraceae). *Plant Biology* 3: 98-105.
558 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:
559 357-359.
560 Lin BL, Yang WJ. 1999. Blue light and abscisic acid independently induce heterophyllous switch
561 in *Marsilea quadrifolia*. *Plant Physiology* 119: 429-434.
562 Liu BL. 1984. Abscisic acid induces land form characteristics in *Marsilea quadrifolia* L.
563 *American Journal of Botany* 71: 638-644.
564 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
565 quantitative PCR and the 2-△△Ct method. *Methods* 25: 402-408.
566 Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. 2008. RNA-seq: an assessment of
567 technical reproducibility and comparison with gene expression arrays. *Genome Research* 18:
568 1509-1517.
569 McCallum WB. 1902. On the nature of the stimulus causing the change of form and structure in
570 *Prosperpinaca palustris*. *Botanical Gazette* 34:93-108.
571 Minorsky PV. 2003. Heterophylly in aquatic plants. Plant Physiology 133: 1671-1672.
572 Mok D, Mok MC. 2001. Cytokinin metabolism and action. *Annual Review of Plant Physiology*
573 *and Plant Molecular Biology* 52: 89-118.
574 Moon J, Hake S. 2011. How a leaf gets its shape. Curr. Opin. Plant Biol. 14: 24-30.
575 Moose SP, Sisco PH. 1996. Glossy15, an *APETALA2*-like gene from maize that regulates leaf
576 epidermal cell identity. *Genes & Development* 10: 3018-3027.
577 Mutasa-Goettgen, ES, Joshi A, Holmes HF, Hedden P, Goettgens B. 2012. A new RNASeq-based
578 reference transcriptome for sugar beet and its application in transcriptome-scale analysis of
579 vernalization and gibberellin responses. *BMC Genomics* 13: 99.

580 Nakayama H, Nakayama N, Nakamasu A, Sinha N, Kimura S. 2012. Toward elucidating the
581 mechanisms that regulate heterophylly. *Plant Morphology*. 24: 57-63.
582 Nakayama H, Nakayama N, Seiki S, Kojima M, Sakakibara H, Sinha N, Kimura S. 2014.
583 Regulation of the *KNOX-GA* gene module induces heterophyllic alteration in North American
584 Lake Cress. *Plant Cell* 26: 4733-4748.
585 Olsen J L, Rouzé P, Verhelst B, Lin YC, Bayer T, Collen J, Dattolo E, De Paoli E, Dittami S,
586 Maumus F, Michel G, Kersting A, Lauritano C, Lohaus R, Töpel M, Tonon T, Vanneste K,
587 Amirebrahimi M, Brakel J, Boström C, Chovatia M, Grimwood J, Jenkins JW, Jueterbock A,
588 Mraz A, Stam WT, Tice H, Bornberg-Bauer E, Green PJ, Pearson GA, Procaccini G, Duarte
589 CM, Schmutz J, Reusch TB, Van de Peer Y. 2016. The genome of the seagrass *Zostera*
590 *marina* reveals angiosperm adaptation to the sea. *Nature*, 530: 331-335.
591 Ozsolak F, Milos PM. 2011. RNA sequencing: advances, challenges and opportunities. *Nature*
592 *Reviews Genetics* 2: 87-98.
593 Pigliucci M. 2010. Phenotypic plasticity. In: Pigliucci M, Muller GM, editors. Evolution: the
594 extended synthesis. MIT Press; p. 355-378.
595 Sculthorpe CD. 1967. The Biology of Aquatic Vascular Plants. Arnold E.
596 Shani E, Ben-Gera H, Shleizer-Burko S, Burko Y, Weiss D, Ori N. 2010. Cytokinin regulates
597 compound leaf development in Tomato. *Plant Cell* 22: 3206-3217.
598 Shani E, Yanai O, Ori N. 2006. The role of hormones in shoot apical meristem function. *Current*
599 *Opinion in Plant Biology* 9: 484-489.
600 Shi C, Yang H, Wei CL, Yu O, Zhang ZZ, Jiang CJ, Sun J, Li YY, Chen Q, Xia T, Wan XC. 2011.
601 Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major
602 metabolic pathways of tea-specific compounds. *BMC Genomics* 12: 131.

603 Smith LG, Greene B, Veit B, Hake S. 1992, A dominant mutation in the maize homeobox gene,
604 *Knotted-1*, causes its ectopic expression in leaf cells with altered fates. *Development* 116: 21-
605 30.
606 Sun X L, Li Y, Cai H, Bai X, Ji W, Ji ZJ, Zhu YM. 2011. Arabidopsis bZIP1 transcription factor
607 binding to ABRE cis-element regulates abscisic acid signal transduction. *Acta Agronomica*
608 *Sinica*. 37: 612-619.
609 Titus JE, Sullivan PG. 2001. Heterophylly in the yellow waterlily, *Nuphar variegata*
610 (Nymphaeaceae): effects of [CO_2] natural sediment type, and water depth. *American Journal*
611 *of Botany* 88: 1469-1478.
612 Tran L, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M,
613 Shinozaki K, Yamaguchi-Shinozaki K. 2004. Isolation and functional analysis of *Arabidopsis*
614 stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in
615 the early responsive to dehydration stress 1 promoter. *Plant Cell*. 16: 2481-2498.
616 Tsukaya H. Mechanism of leaf-shape determination. *Annual Review of Plant Biology* 2006; 57:
617 477-496.
618 Uchida N, Kimura S, Koenig D, Sinha N. 2010. Coordination of leaf development via regulation
619 of *KNOX1*genes. *Journal of Plant Research* 123:7-14.
620 Uchida N, Townsley B, Chung KH, Sinha N. 2007. Regulation of *SHOOT MERISTEMLESS*
621 genes via an upstream-conserved noncoding sequence coordinates leaf development.
622 *Proceedings of the National Academy of Sciences of America*. 104: 15953-15958.
623 van der Graaff E, Den Dulk-Ras A, Hooykaas P, Keller B. 2000. Activation tagging of the
624 LEAFY PETIOLE gene affects leaf petiole development in Arabidopsis thaliana.
625 *Development*. 127: 4971-4980.
626 Varshney R, Nayak S, May G, Jackson S. 2009. Next-generation sequencing technologies and
627 their implications for crop genetics and breeding. *Trends in Biotechnology* 27: 522-530.

628 Vollbrecht E, Veit B, Sinha N, Hake S. 1991. The developmental gene Knotted-1 is a member of
629 a maize homeobox gene family. *Nature* 350: 241-243.
630 Walton DC, Li Y. 2013. Abscisic acid biosynthesis and metabolism. In: Davies P, editor. Plant
631 hormones: Physiology, Biochemistry and Molecular Biology. Springer Science & Business
632 Media; p.140-157.
633 Wanke D. 2011. The ABA-mediated switch between submersed and emersed life-styles in aquatic
634 macrophytes. *Journal of Plant Research* 124: 467-475.
635 Wells CL, Pigliucci M. 2000. Adaptive phenotypic plasticity: the case of heterophylly in aquatic
636 plants. P*erspectives in Plant Ecology Evolution and Systematics* 3: 1-18.
637 Winn AA. 1999. Is seasonal variation in leaf traits adaptive for the annual plant *Dicerandra*
638 *linearifolia*? Journal of Evolutionary Biology 1999; 12: 306-313.
639 Winn AA. The functional significance and fitness consequences of heterophylly. *International*
640 *Journal of Plant Sciences* 160: S113-S121.
641 Yang M, Zhu L, Xu L, Pan C, Liu Y. 2014. Comparative transcriptomic analysis of the regulation
642 of flowering in temperate and tropical lotus (*Nelumbo nucifera*) by RNA-Seq. *Annals of*
643 *Applied Biology* 165: 73-95.
644 Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li ST, Li RQ, Bolund L, Wang J.
645 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research* 34: W293-
646 W297.
647 Young JP, Dengler NG, Horton RF. 1987. Heterophylly in *Ranunculus flabellaris*: The effect of
648 abscisic acid on leaf anatomy. *Annals of Botany* 60: 117-125.
649 Young JP, Horton RF. 1985. Heterophylly in *Ranunculus flabellaris*: the effect of abscisic acid.
650 *Annals of Botany* 55: 899-902.
651 Yuan L, Perry SE. 2011. Plant transcription factors. Methods and Protocols. *Methods in*
652 *Molecular Biology* 754: 347.

653 Zhang X, Gituru RW, Yang C, Guo YH. 2009. Variations of floral traits among different life
654 forms illustrate the evolution of pollination systems in *Potamogeton* species from China.
655 *Aquatic Botany* 90: 124-128.
656 Zotz G, Wilhelm K, Becker A. 2011. Heteroblasty-a review. *Botanical Review* 77: 109-151.

Figure legends:

Fig. 1 Morphological features of *P*. *octandrus*. (A) The initial developmental stage of the plant that produces only submerged leaves. (B) The later developmental 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 DEGs identified in comparisons among different *P. octandrus* leaf development stages. "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 common and unique DEGs among different comparisons of floating leaves. (B) Venn diagram showing common and unique DEGs among different comparisons of submerged leaves. (C) Expression patterns of DEGs among different comparisons. The numbers above each bar indicate the total number of genes in each group.

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

Fig. 5 GO functional classifications of DEGs identified from comparisons among different groups during submerged leaf 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 similarities in transcription factor expression profiles among samples. Three transcription factor cluster groups (G1, G2 and G3) resulted from the 469

681 significantly differentially expressed transcription factors in leaves from different stages of
682 development.

683 **Fig. 7** Distribution of transcription factor families among the three transcription factor cluster
684 groups (G1, G2 and G3).

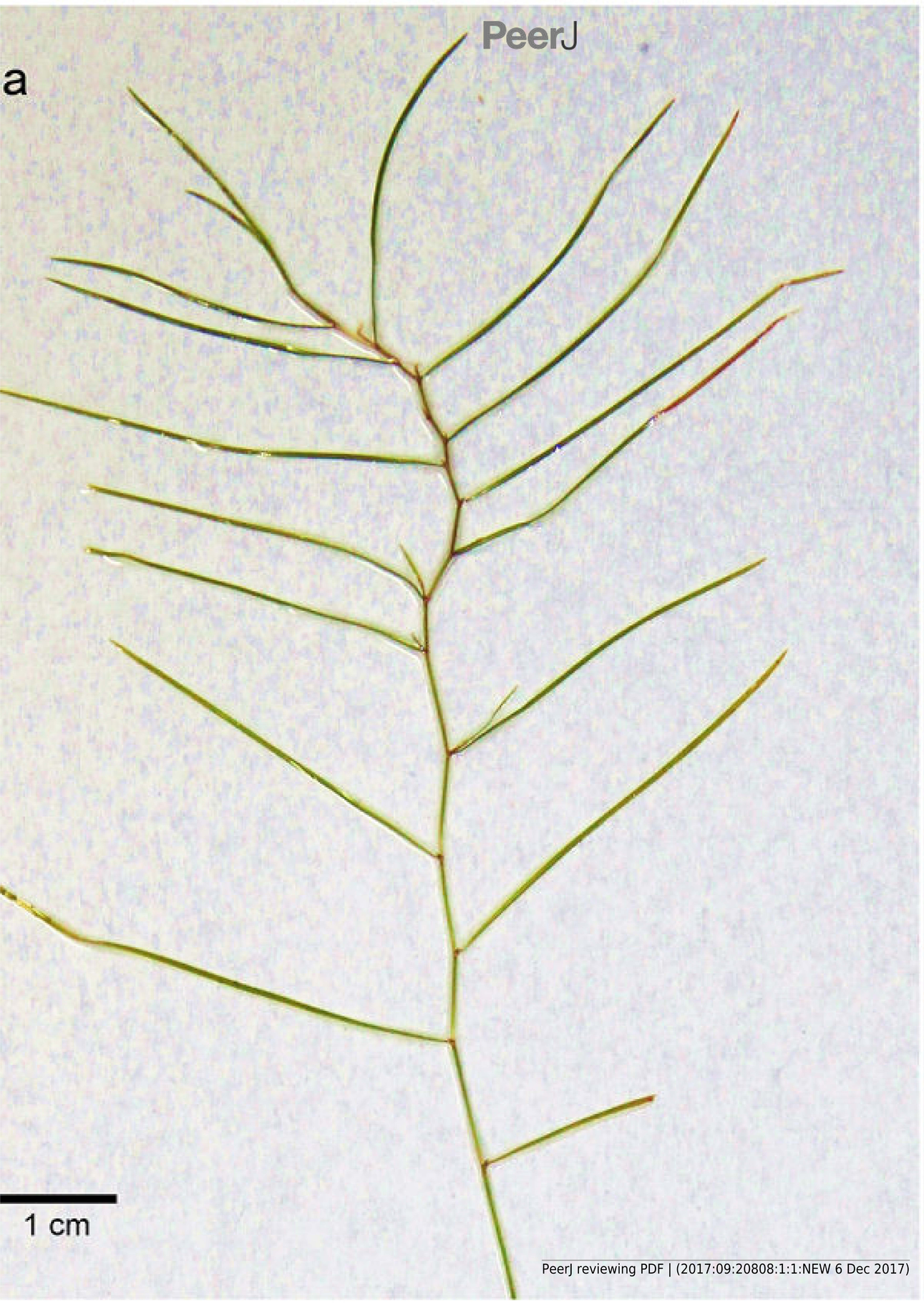
Figure 1(on next page)

Morphological features of P. octandrus.

(A) The initial developmental stage of the plant that produces only submerged leaves. (B) The later developmental stage of the plant that produces both floating

Manuscript to be reviewed

a



b

Figure 2(on next page)

Gene ontology (GO) annotations of all detected genes.

Three main categories including “biological process”, “cellular component”, and “molecular function” were summarized.

PeerJ reviewing PDF | (2017:09:20808:1:1:NEW 6 Dec 2017)

PeerJ

Manuscript to be reviewed

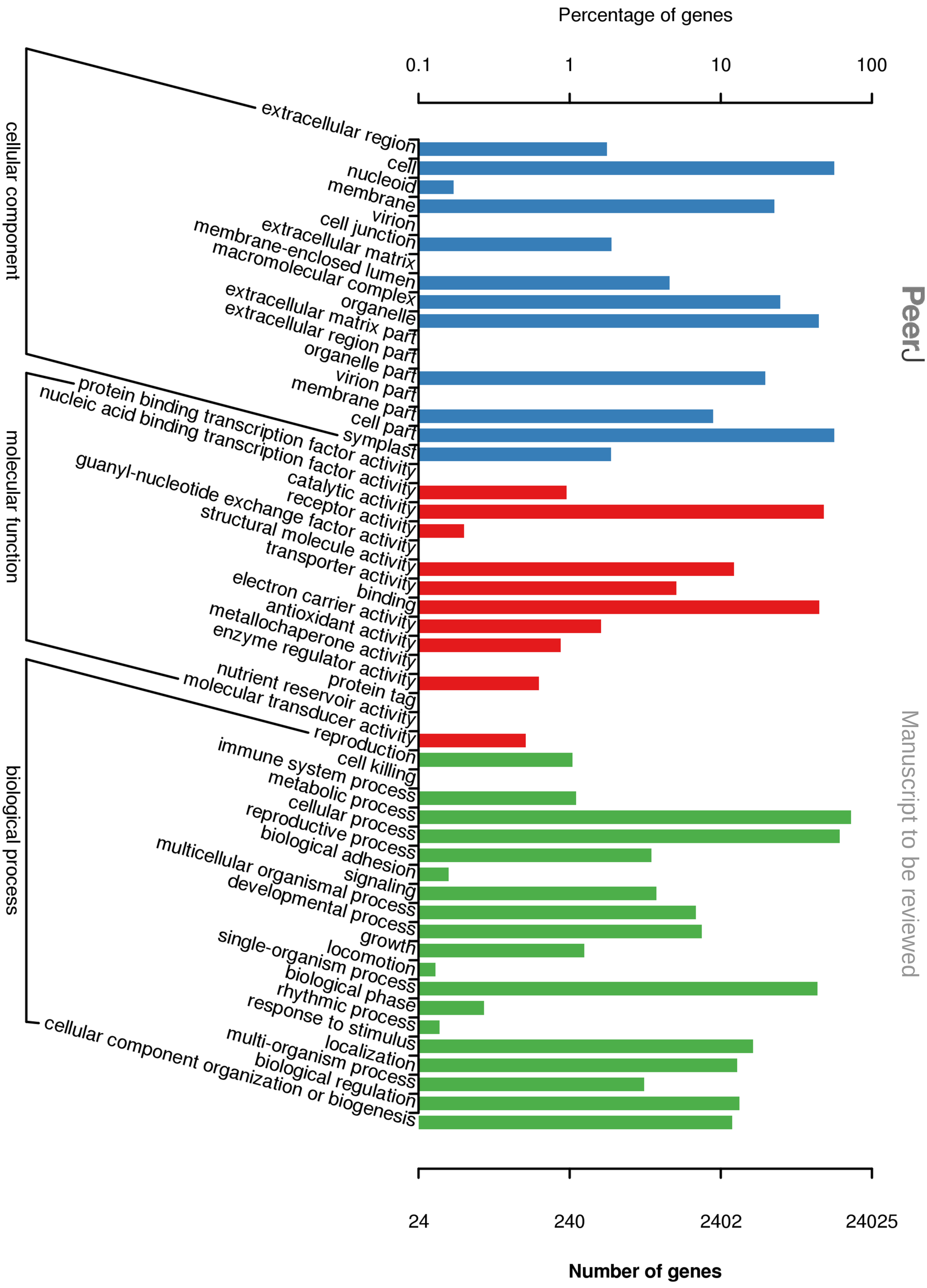
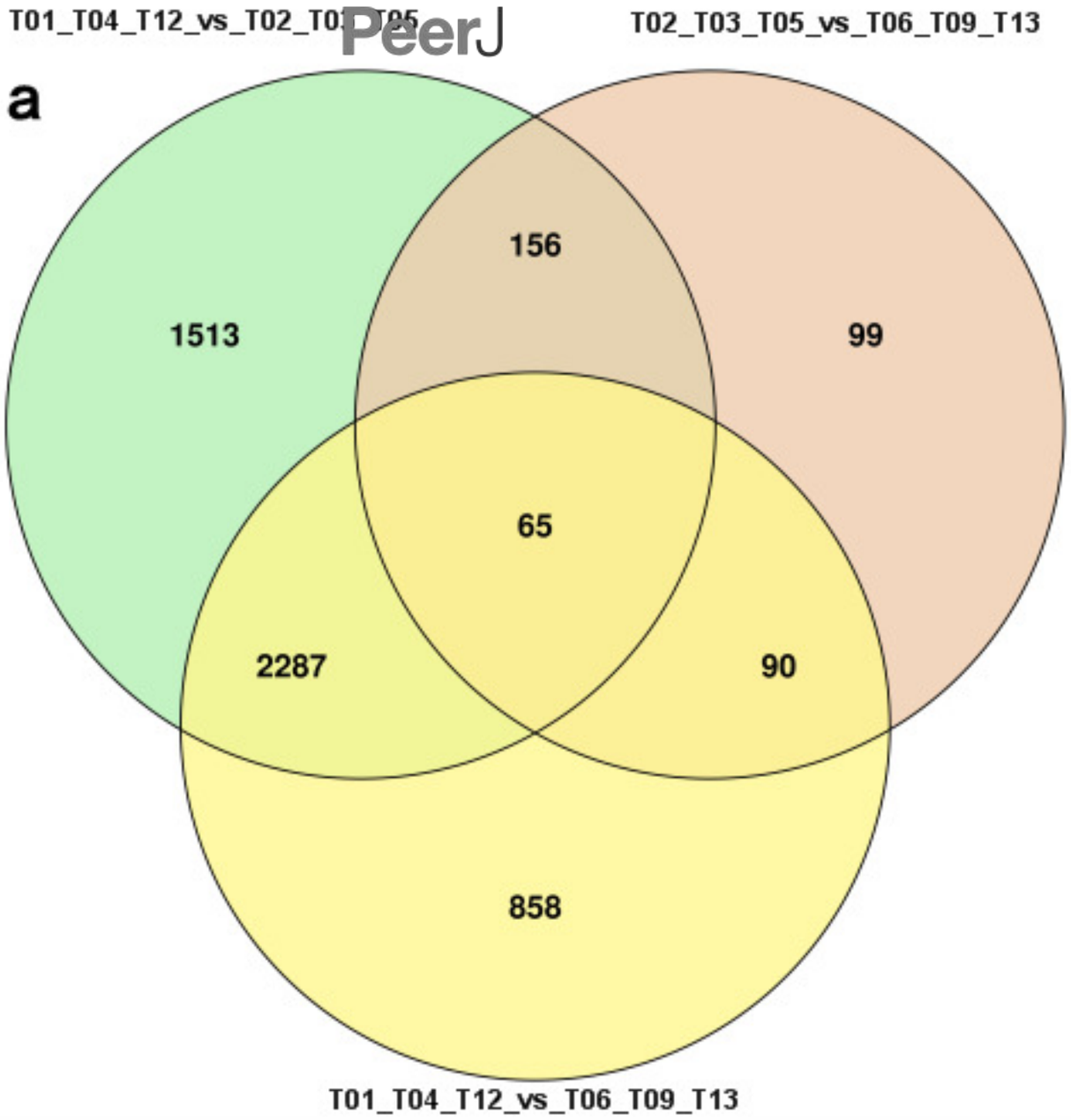
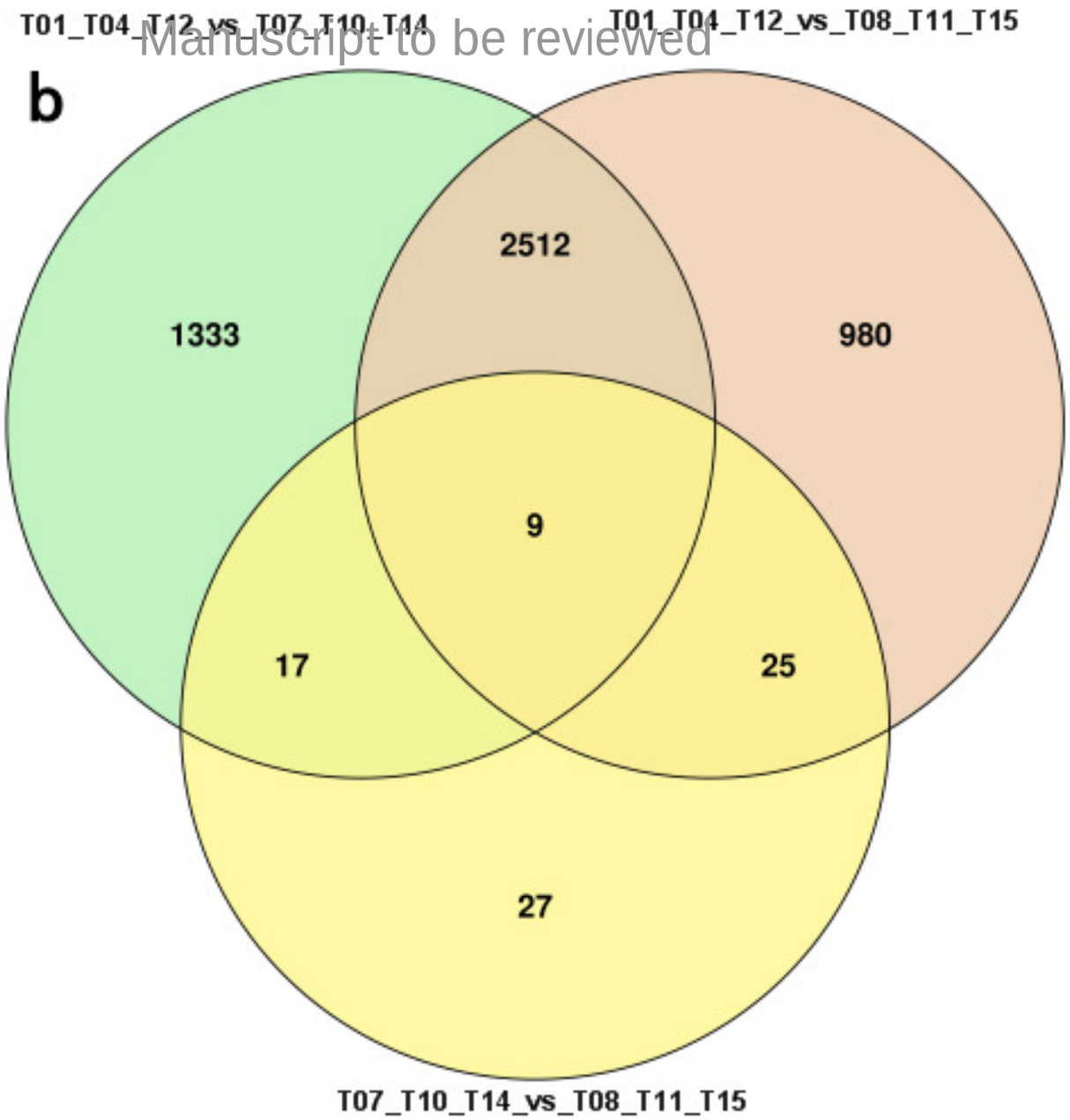


Figure 3(on next page)

DEGs identified in comparisons among different P. octandrus leaf development stages.

“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 common and unique DEGs among different comparisons of floating leaves. (B) Venn diagram showing common and unique DEGs among different comparisons of submerged leaves. (C) Expression patterns of DEGs among different comparisons. The numbers above each bar indicate the total number of genes in each group.


a
T01_T04_T12_vs_T02_T03_T05
T02_T03_T05_vs_T06_T09_T13
1513
156
99
65
2287
90
858
T01_T04_T12_vs_T06_T09_T13


b
T01_T04_T12_vs_T07_T10_T14
T01_T04_T12_vs_T08_T11_T15
1333
2512
980
9
17
25
27
T07_T10_T14_vs_T08_T11_T15

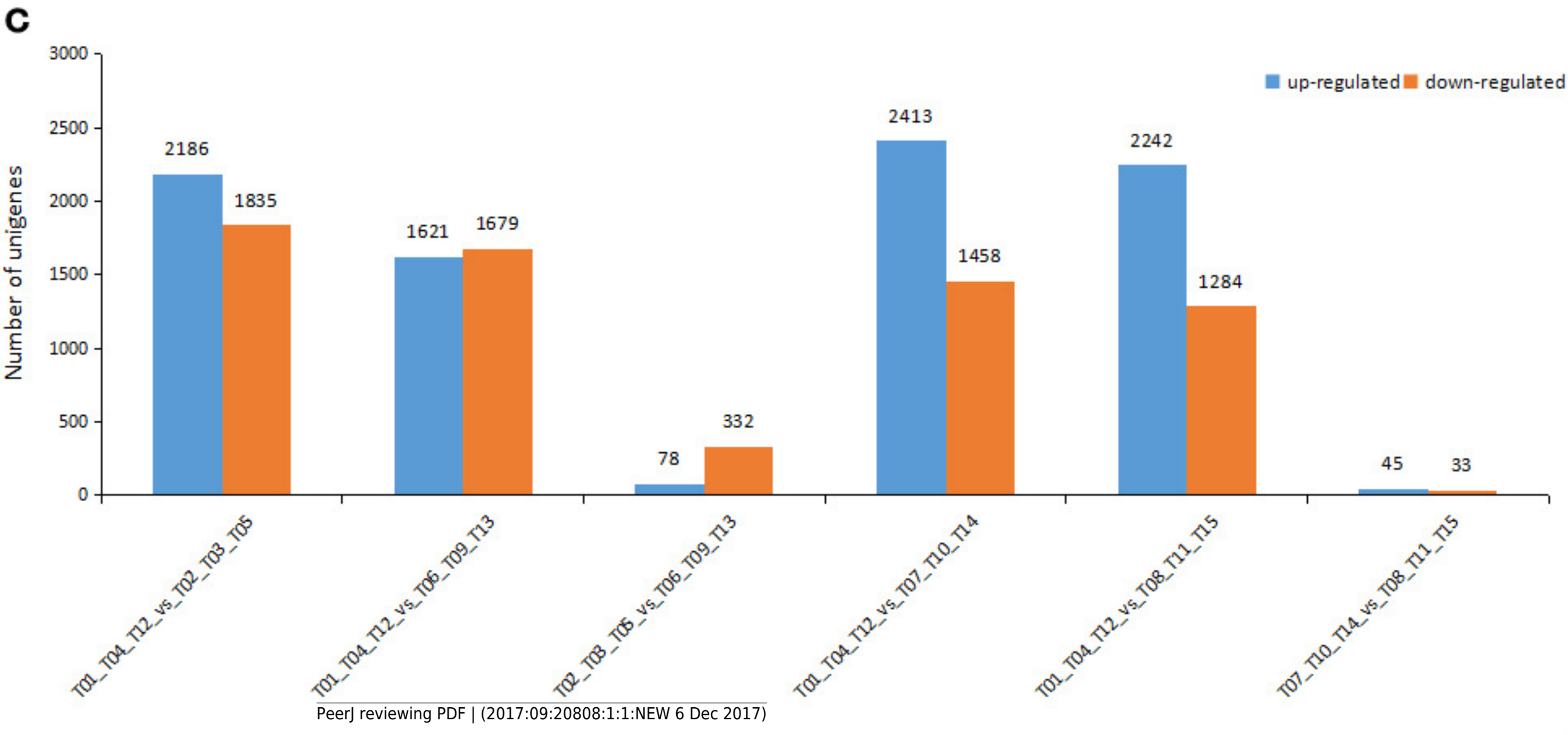

c
Number of unigenes
3000
2500
2000
1500
1000
500
0
up-regulated
down-regulated
2186
1835
1621
1679
78
332
2413
1458
2242
1284
45
33
T01_T04_T12_vs_T02_T03_T05
T01_T04_T12_vs_T06_T09_T13
T02_T03_T05_vs_T06_T09_T13
T01_T04_T12_vs_T07_T10_T14
T01_T04_T12_vs_T08_T11_T15
T07_T10_T14_vs_T08_T11_T15

Figure 4(on next page)

GO functional classifications of DEGs identified from comparisons among different groups during floating leaf development.

“T01, T04, T12”, “T02, T03, T05”, and “T06, T08, T13” indicate three biological replicates of shoot, juvenile floating leaves, and adult floating leaves, respectively.

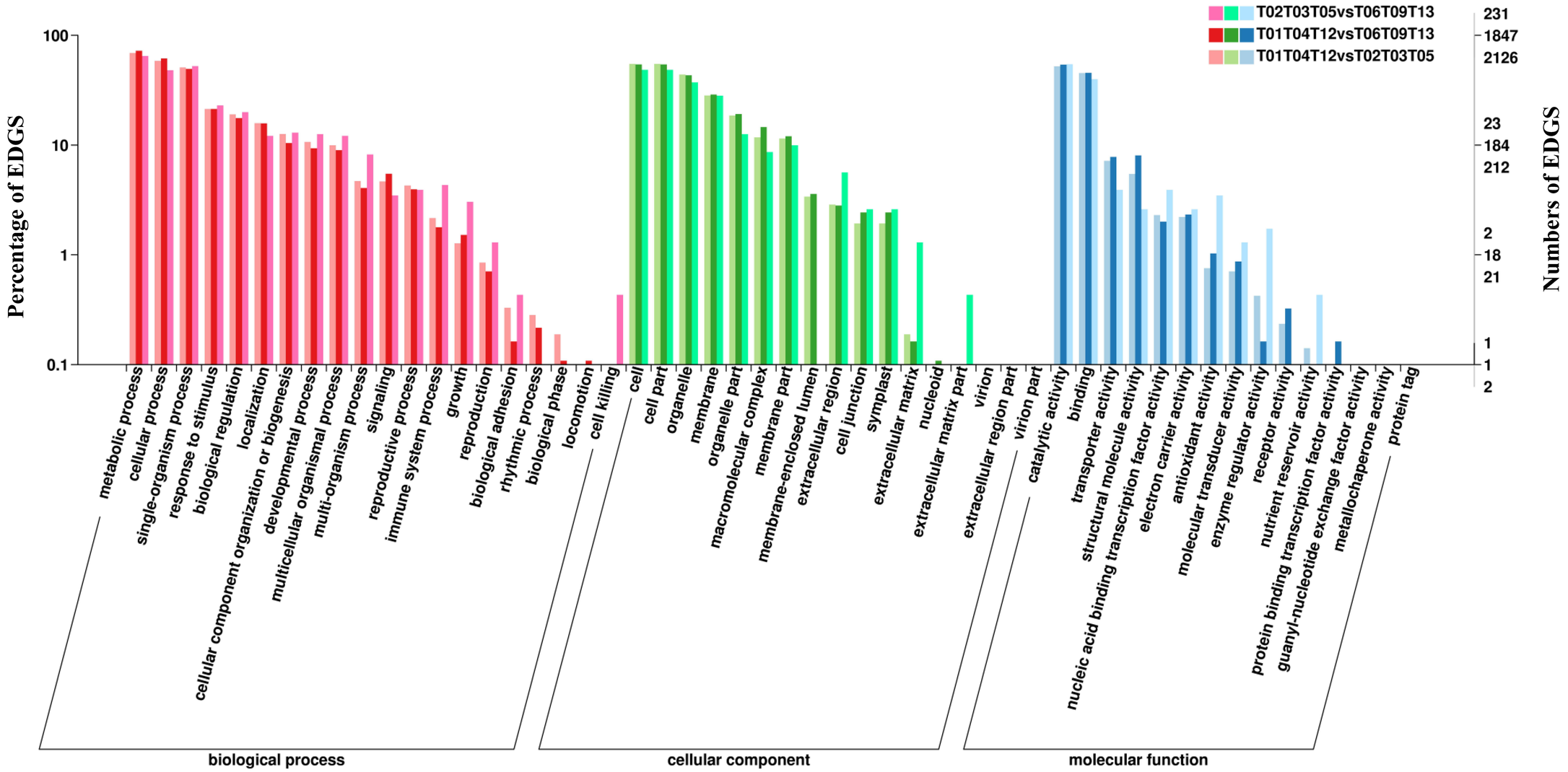

T02T03T05vsT06T09T13
T01T04T12vsT06T09T13
T01T04T12vsT02T03T05
Percentage of EDGS
Numbers of EDGS
100
10
1
0.1
231
1847
2126
23
184
212
2
18
21
1
1
2
metabolic process
cellular process
single-organism process
response to stimulus
biological regulation
localization
cellular component organization or biogenesis
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
biological process
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
cellular component
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
molecular function

Figure 5(on next page)

GO functional classifications of DEGs identified from comparisons among different groups during submerged leaf 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.

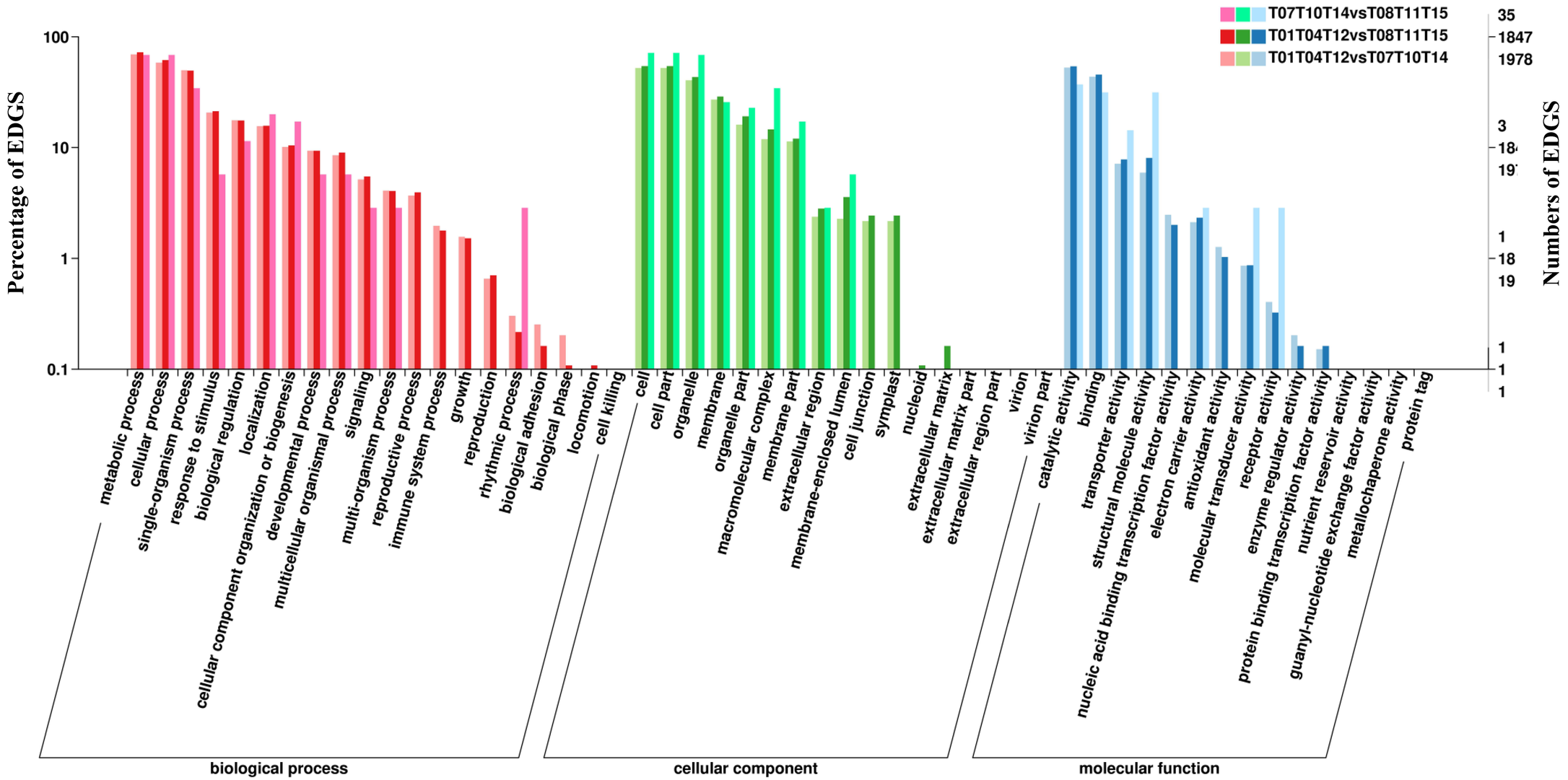

T07T10T14vsT08T11T15
T01T04T12vsT08T11T15
T01T04T12vsT07T10T14
Percentage of EDGS
100
10
1
0.1
Numbers of EDGS
35
1847
1978
3
1
18
19
metabolic process
cellular process
single-organism process
response to stimulus
biological regulation
localization
cellular component organization or biogenesis
developmental process
multicellular organismal process
signaling
multi-organism process
reproductive process
immune system process
growth
reproduction
rhythmic process
biological adhesion
biological phase
locomotion
cell killing
biological process
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
cellular component
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
molecular function

Figure 6(on next page)

Dendrogram showing similarities in transcription factor expression profiles among samples.

Three transcription factor cluster groups (G1, G2 and G3) resulted from the 469 significantly differentially expressed transcription factors in leaves from different stages of development.

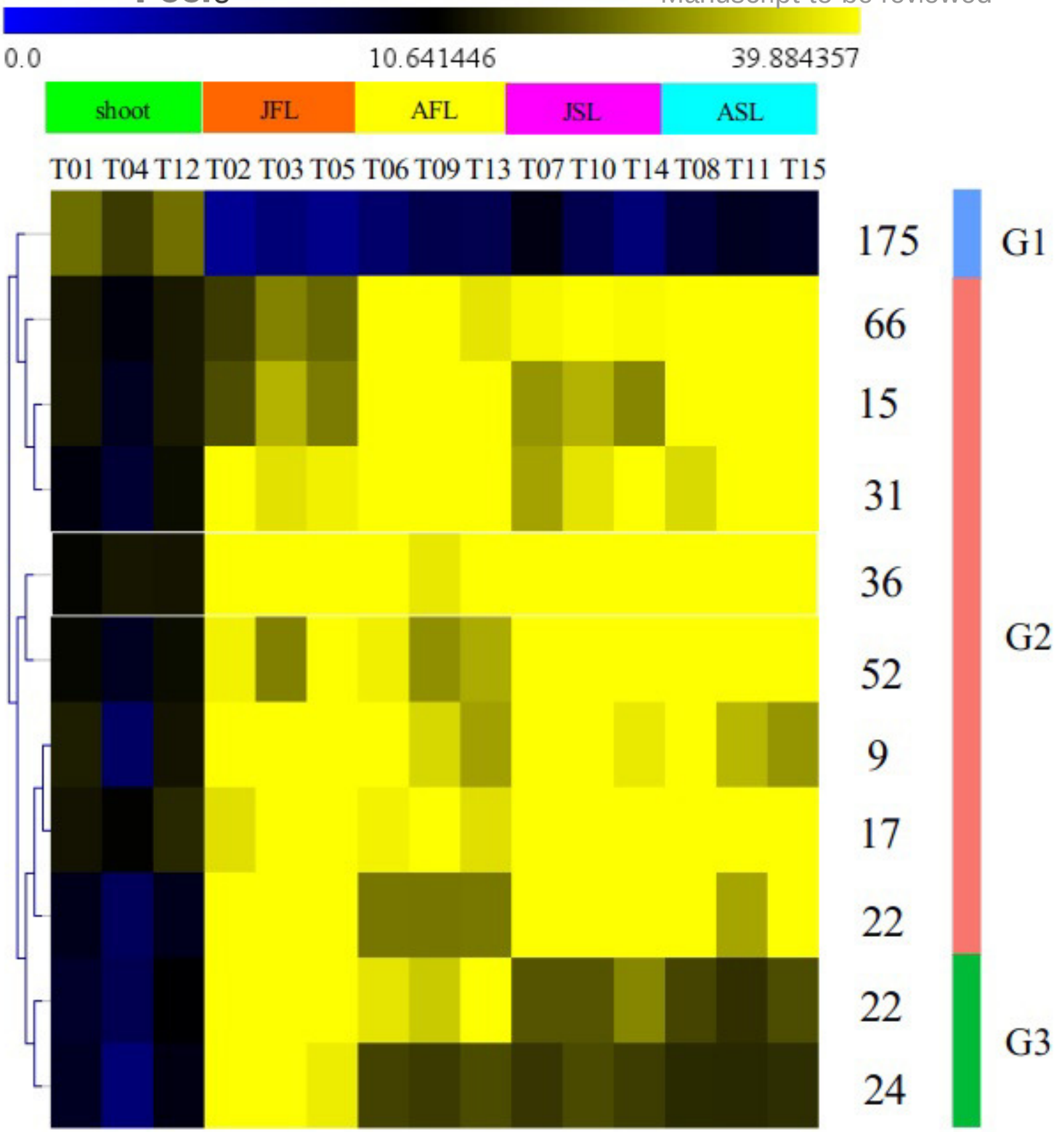

0.0
10.641446
39.884357
shoot
JFL
AFL
JSL
ASL
T01 T04 T12 T02 T03 T05 T06 T09 T13 T07 T10 T14 T08 T11 T15
175
66
15
31
36
52
9
17
22
22
24
G1
G2
G3

Figure 7 (on next page)

Distribution of transcription factor families among the three transcription factor cluster groups (G1, G2 and G3).

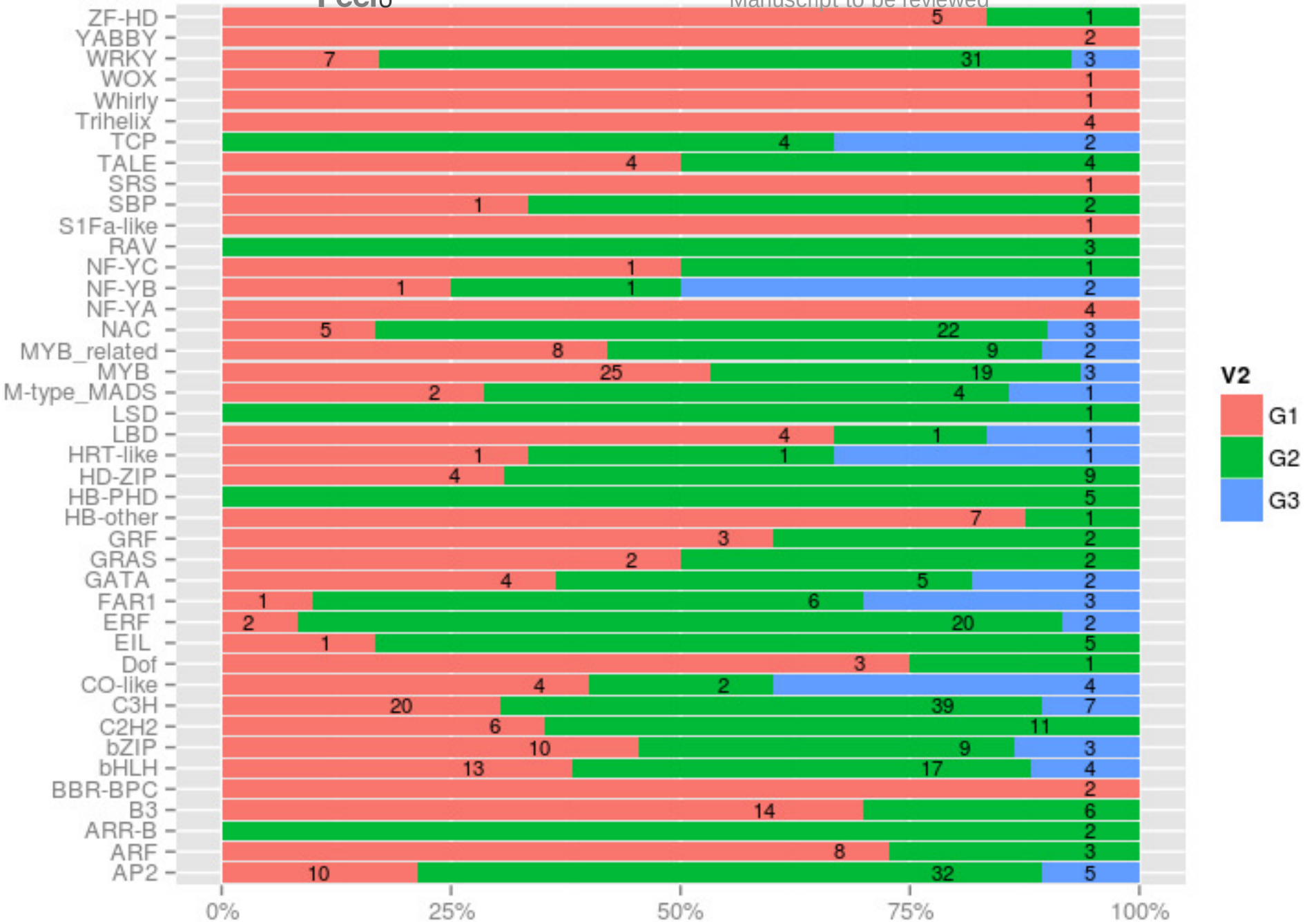

ZF-HD 5 1
YABBY 2
WRKY 7 31 3
WOX 1
Whirly 1
Trihelix 4
TCP 4 2
TALE 4 4
SRS 1
SBP 1 2
S1Fa-like 1
RAV 3
NF-YC 1 1
NF-YB 1 1 2
NF-YA 4
NAC 5 22 3
MYB_related 8 9 2
MYB 25 19 3
M-type_MADS 2 4 1
LSD 1
LBD 4 1 1
HRT-like 1 1 1
HD-ZIP 4 9
HB-PHD 5
HB-other 7 1
GRF 3 2
GRAS 2 2
GATA 4 5 2
FAR1 1 6 3
ERF 2 20 2
EIL 1 5
Dof 3 1
CO-like 4 2 4
C3H 20 39 7
C2H2 6 11
bZIP 10 9 3
bHLH 13 17 4
BBR-BPC 2
B3 14 6
ARR-B 2
ARF 8 3
AP2 10 32 5
0% 25% 50% 75% 100%
V2
G1
G2
G3

Table 1(on next page)

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

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