Journal of Plant Physiology 221 (2018) 107–118

Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: www.elsevier.com/locate/jplph

De novo transcriptome assembly and comparative analysis between male and benzyladenine-induced female inflorescence buds of *Plukenetia volubilis*

Qiantang Fu, Longjian Niu, Mao-Sheng Chen, Yan-Bin Tao, Xiulan Wang, Huiying He, Bang-Zhen Pan, Zeng-Fu Xu*

Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Yunnan, 666303, China

ARTICLE INFO

Keywords:
Sacha inchi
Transcriptome
Cytokinin
Floral sex differentiation
Flower development
Monoecious plants

ABSTRACT

Plukenetia volubilis is a promising oilseed crop due to its seeds being rich in unsaturated fatty acids, especially alpha-linolenic acid. *P. volubilis* is monoecious, with separate male and female flowers on the same inflorescence. We previously reported that male flowers were converted to female flowers by exogenous cytokinin (6-benzyladenine, 6-BA) treatment in *P. volubilis*. To identify candidate genes associated with floral sex differentiation of *P. volubilis*, we performed *de novo* transcriptome assembly and comparative analysis on control male inflorescence buds (MIB) and female inflorescence buds (FIB) induced by 6-BA using Illumina sequencing technology. A total of 57,664 unigenes with an average length of 979 bp were assembled from 104.1 million clean reads, and 45,235 (78.45%) unigenes were successfully annotated in the public databases. Notably, Gene Ontology analyses revealed that 4193 and 3880 unigenes were enriched in the categories of reproduction and reproductive processes, respectively. Differential expression analysis identified 1385 differentially expressed unigenes between MIB and FIB, of which six unigenes related to cytokinin and auxin signaling pathways and 16 important transcription factor (TF) genes including MADS-box family members were identified. In particular, several unigenes encoding important TFs, such as homologs of CRABS CLAW, RADIALIS-like 1, RADIALIS-like 2, HECATE 2, WUSCHEL-related homeobox 9, and SUPERMAN, were expressed at higher levels in FIB than in MIB. The expression patterns of the 36 selected unigenes revealed by transcriptome analysis were successfully validated by quantitative real-time PCR. This study not only provides comprehensive gene expression profiles of *P. volubilis* inflorescence buds, but also lays the foundation for research on the molecular mechanism of floral sex determination in *P. volubilis* and other monoecious plants.

1. Introduction

Floral sex differentiation plays an important role in fruit set and seed production, and this biological process is regulated by both genetic factors and the external and internal environmental conditions (Tanurdzic and Banks, 2004). Monoecious species provide a comprehensive system for studying the developmental programs underlying the establishment of male and female organs in unisexual flowers (Rocheta et al., 2014). Many studies have reported floral sex determination in monoecious species, such as cucumber (*Cucumis sativus*), bitter gourd (*Momordica charantia*), maize (*Zea mays*), *Jatropha curcas*, and *Quercus suber* (Bai and Xu, 2013; Chen et al., 2016; Rocheta et al., 2014; Shukla et al., 2015; Thompson, 2014; Wu et al., 2010). However, molecular data resources are limited for most monoecious non-model species, especially those without a reference genome, and the mechanisms that control floral sex in different plant species are far from conclusive.

Plukenetia volubilis, a species of the Euphorbiaceae family, also known as Sacha inchi, or Inca peanut, is a perennial oilseed vine and native to the rainforests of South America. *P. volubilis* seeds contain 25–27% protein and 41–54% oil, which comprises approximately 90% unsaturated fatty acids (oleic, linoleic, linolenic) and is rich in vitamins E and A (Chirinos et al., 2013; Gutierrez et al., 2011; Niu et al., 2014). *P. volubilis* oil has great potential economic value in cosmetic, pharmaceutical, and food industries (Chirinos et al., 2013; Hanssen and Schmitz-Huebsch, 2011). Wang et al. (2012) reported the transcriptome analysis of Sacha inchi seeds at the initial and fast oil accumulation stages, and identified 397 unigenes associated with the biosynthesis of fatty acids. Recently, an *in vitro* regeneration system for *P. volubilis* has been successfully established using cotyledons as explants (Dong et al., 2016). These studies lay the foundation for further gene function studies and utilization of *P. volubilis*. However, studies on flower

* Corresponding author.
E-mail addresses: zfxu@xtbg.ac.cn, zengfu.xu@gmail.com (Z.-F. Xu).

https://doi.org/10.1016/j.jplph.2017.12.006
Received 22 July 2017; Received in revised form 4 December 2017; Accepted 4 December 2017
Available online 09 December 2017

development, especially floral sex determination, in *P. volubilis* remain limited. *P. volubilis* is monoecious, with separate female and male flowers on the same inflorescence. Only one or two female flowers are at the basal-most node, and numerous male flowers are radially set around the narrow raceme-like inflorescence axis (Gillespie, 1993). We previously developed an efficient method to convert most of the male flower buds to female flower buds by exogenous 6-benzyladenine (6-BA) treatment (Fu et al., 2014). Therefore, the molecular mechanisms of floral sex determination in *P. volubilis* are worthy of further research.

Many studies have shown that phytohormones are significantly involved in modulating floral sex-developmental processes in a variety of species. In general, ethylene, cytokinin, and auxin promote female sex development, while gibberellic acid promotes male sex development (Chailakhyan and Khryanin, 1978; Heslopharrison, 1956; Orozco-Arroyo et al., 2012; Pan et al., 2014; Yin and Quinn, 1995). Some phytohormone biosynthesis and perception genes that affect sex determination have been discovered, such as the key regulatory genes of ethylene biosynthesis, the 1-aminocyclopropane-1-carboxylic acid synthase (ACS) genes *CmACS-7* and *CmACS-11*, which regulate the sex determination in melon (Boualem et al., 2008). Some transcription factors (TFs) play important roles in regulating floral sex development, such as SUPERMAN (SUP) (Kazama et al., 2009), MYB (Song et al., 2013; Wu et al., 2010), and basic helix-loop-helix (bHLH) (Gremski et al., 2007), in different plant species. In addition, MADS-box TF genes are also implicated in sex differentiation of male and female flowers in unisexual plants. Chawla et al. (2015) demonstrated that *HrAP1* showed female-specific expression, while *HrAP2* was expressed particularly in male flowers of seabuckthorn (*Hippophae rhamnoides*). *SpAP*3 of spinach (*Spinacia oleracea*) is strongly expressed in male flowers and weakly expressed in female flowers, and *SpPI* is expressed early in male flower development (Pfent et al., 2005). These results indicate that floral sex determination is governed by many signals and complex transcriptional regulatory networks.

In recent years, transcriptome-sequencing technology has become increasingly common for unravelling the genetic networks that regulate floral sex determination and development in many species. Rocheta et al. (2014) analyzed the transcriptome of male and female flowers in monoecious *Quercus suber* using 454 pyrosequencing technology, and identified some genes involved in pollen development and ovule formation. In the transcriptome analysis of flower buds at different developmental stages in *Jatropha curcas*, 15 sex-related genes contributing to stamen differentiation and embryo sac development were obtained (Xu et al., 2016). Other studies have also been conducted to detect the candidate genes underlying sex determination in a variety of plant species, such as wild grapevine (*Vitis vinifera*) (Ramos et al., 2014, 2017), bitter gourd (Shukla et al., 2015), and castor bean (*Ricinus communis*) (Tan et al., 2016). Although many genes involved in regulating floral sex have been identified in different species, the gene expression information in florescence buds and the factors regulating floral sex in *P. volubilis* remain poorly understood.

To reveal sex differences in *P. volubilis* at the transcription level, we performed *de novo* transcriptome assembly and comparative analysis on control male inflorescence buds (MIB) and female inflorescence buds (FIB) induced by 6-BA. More than 104.1 million clean reads and 57,664 unigenes were obtained from the two transcriptome libraries. We identified multiple phytohormone-related genes and some important transcription factors (TFs) genes including MADS-box family members, some of which might play important roles in flower development and floral sex determination in *P. volubilis*.

2. Materials and methods

2.1. Plant material

Two-year-old *P. volubilis* trees were grown at Xishuangbanna Tropical Botanical Garden (21°54′ N, 101°46′ E, 580 m asl), Chinese Academy of Sciences, located in Mengla County, Yunnan Province, China, under natural conditions. Because only one to two female flowers are located near the base of the inflorescences, it is difficult to obtain female flower buds in the early developmental stage of inflorescence. To obtain female flower buds, *P. volubilis* trees were treated with 20 mg/L 6-BA in July–August 2014, as previously reported (Fu et al., 2014); control trees were sprayed with distilled water containing 0.05% (v/v) Tween-20. When the first poly-female inflorescence (> 2 female flowers) was observed, the fourth to sixth inflorescences of the branch were selected as FIB. The selected inflorescences were less than 0.5 cm in length. MIB were selected from control inflorescences during a similar developmental stage. Each sample consisted of more than 20 male or female inflorescence buds, respectively. The collected samples were immediately frozen in liquid nitrogen and stored at −80 °C for RNA extraction.

2.2. RNA isolation, library construction, and transcriptome sequencing

Total RNA of each sample was isolated using Trizol reagent (Invitrogen, USA) and purified using an RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturers' protocols. The quality and quantity of total RNA were assessed using a 1.0% agarose gel and a Nano-Drop ND1000 spectrophotometer, and RNA integrity was evaluated using an Agilent 2100 Bioanalyzer. Then, the isolated RNA samples were prepared for construction of transcriptome libraries and quantitative real-time PCR (qRT-PCR) validation.

Poly (A) mRNA was isolated and enriched from total RNA using oligo (dT)-attached magnetic beads according to the manufacturer's (Illumina) instructions and then chemically broken into short fragments in fragmentation buffer. Using these short mRNA fragments as templates, random hexamers were used as primers to synthesize first-strand cDNA. Subsequently, second-strand cDNA was synthesized in a buffer containing dNTPs, DNA polymerase I, and RNaseH. Short cDNA fragments were purified with a QIAquick PCR extraction kit and resolved in extraction buffer (10 mmol/L Tris-HCl, pH 8.4). After the purified cDNA ends were repaired and single nucleotide adenine (A) was added, the fragments were ligated to sequencing adapters. The suitable fragments were screened and purified as templates for PCR amplification to enrich the cDNA. The cDNA libraries were sequenced using an Illumina HiSeq™ 2000 sequencing platform at the Beijing Genomics Institute (BGI, Shenzhen, China), and the raw reads were generated in a 90-bp paired-end format.

2.3. De novo transcriptome assembly and unigene functional annotation

Before data analysis, the raw reads were filtered using the Filter_fq (an internal program of BGI) by removing adaptor reads, low-quality reads (containing more than 20% bases with Q-value ≤ 10), and reads containing more than 5% unknown nucleotides. The clean reads obtained from each library were assembled separately, and pooled reads of both libraries were also assembled *de novo* into unigenes using Trinity program (Grabherr et al., 2011). The longest non-redundant unigenes were acquired by removing sequence splicing and redundancy using TGI Clustering tools (TGICL) (Pertea et al., 2003). Then the unigenes were divided into two classes: "clusters", with the prefix "CL", and "singletons", with the prefix "unigene". The transcriptome data have been deposited in the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) database under the accession number SRP101414 (ftp://ftp-trace.ncbi.nlm.nih.gov/sra/review/SRP101414_20171019_083700_a3f2a910685f5b07f5f45a5fc1fdb389).

For further annotation of unigenes, all the assembly unigenes were first searched against the NCBI non-redundant protein database (Nr) and Swiss-Prot protein database using BLASTx alignment (E-value < 1.0E-05), and the NCBI non-redundant nucleotide sequence (Nt) database using BLASTn (E-value < 1.0E-05). The best aligning

results were used to decide sequence direction of unigenes. Then the Blast2GO program (Conesa et al., 2005) was used to obtain GO annotation of unigenes based on the Nr annotation. The GO functional classification was then performed using WEGO software (Ye et al., 2006) to deduce the distribution of each unigene catalogued as biological processes, cellular components, and molecular functions ontologies. The unigene sequences were also aligned to the Clusters of Orthologous Groups of proteins (COG) database to predict and classify their functions. Pathway assignments were carried out according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Kanehisa et al., 2012) using BLASTx (E-value < 1.0E-05).

2.4. Gene expression analysis and screening of differentially expressed unigenes

To quantify transcript abundances and identify significant changes in transcript expression levels, all read counts were calculated using the fragments per kilobase of transcript per million fragments mapped (FPKM) method (Trapnell et al., 2010). The method described by Audic and Claverie (1997) was used to determine the differences in gene expression analysis. The false discovery rate (FDR) ≤ 0.001 and $|\log_2 \text{ratio}| \geq 1$ were used as thresholds to judge the differentially expressed unigenes (DEGs) between the two samples.

2.5. Quantitative real-time PCR analysis

cDNA was synthesized from the same RNA samples used for transcriptome sequencing with the PrimeScript® RT reagent kit (Takara) following the manufacturer's protocols. qRT-PCR was performed on a Roche LightCycler 480 Real-Time PCR instrument (Roche Diagnostics) using a LightCycler® 480 SYBR Green I Master (Roche). The cycling procedure was initiated at 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s, 59 °C for 10 s, and 72 °C for 30 s. Three biological replicates for each sample and three technical replicates were performed. The relative expression levels of the candidate genes were calculated according to the $2_T^{-\Delta\Delta C}$ method (Livak and Schmittgen, 2001) using the *PvoGAPDH* gene as a reference gene (Niu et al., 2015). Gene-specific primers were designed based on the candidate unigene sequences (Supplemental Table 6).

3. Results

3.1. Sequencing and de novo assembly of male and female inflorescence bud transcriptome of P. volubilis

To identify candidate genes involved in flower development and sex differentiation of *P. volubilis*, female inflorescence buds (FIB) induced by 6-BA and control male inflorescence buds (MIB) at early developmental stages (Fig. 1A and B), which may further develop into female inflorescences and male inflorescences, respectively (Fig. 1C and D), were collected for transcriptome sequencing in this study.

The two cDNA libraries constructed from MIB and FIB were sequenced on an Illumina high-throughput sequencing platform. After filtering the adaptors and low-quality sequences, 51,201,412 and 52,876,984 high-quality clean reads (90 bp) were generated from the male and female flower buds libraries, respectively (Supplemental Table 1). The Q20 percentage reached more than 96%, and the GC contents of the two libraries were 43.21% (MIB) and 43.36% (FIB), respectively (Supplemental Table 1). A total of 67,019 and 66,723 unigenes with mean sizes of 750 bp and 775 bp were generated from the MIB and FIB libraries, respectively (Table 1). Altogether, 57,664 unigenes with a mean size of 979 bp and an N50 value of 1525 bp were obtained (Table 1). The size distribution of the assembled contigs and unigenes are presented in Supplemental Fig. S1.

3.2. Functional annotation of unigenes from inflorescence buds of P. volubilis

To identify the putative functions of the assembled unigenes, these unigene sequences were aligned to the Nr, Swiss-Prot, KEGG, COG, and GO (E-value < 1.0E-05) protein databases by BLASTx and to the Nt (E-value < 1.0E-05) nucleotide database by BLASTn. The results showed that 43,690 (75.77%), 43,049 (74.65%), 28,182 (48.87%), 24,888 (43.16%) 15,919 (27.61%), and 36,387 (63.10%) of the 57,664 unigenes were significantly matched with sequences in the Nr, Nt, Swiss-Prot, KEGG, COG, and GO databases, respectively. In total, 45,235 (78.45%) unigenes were successfully annotated in the public databases (Table 2).

The E-value distribution of the top hits showed that 56.0% of the unigenes had high homologies to previously deposited sequences (< 1.0E-45) and 34.0% had E-values in the range of 1.0E-05 to 1.0E-45 (Fig. 2A). Likewise, the similarity distribution of the top BLAST hits revealed that 50.4% of the matched sequences had reached higher than 80% similarity, and 45.8% had similarities ranging from 40 to 80% (Fig. 2B). Further analysis of homologies among different species showed that 76.5% of the annotated sequences matched sequences from castor bean, 9.7% to poplar (*Populus trichocarpa*), 4.2% to grape, 2.2% to peach (*Amygdalus persica*), and 1.1% to woodland strawberry (*Fragaria vesca*) (Fig. 2C).

3.3. Functional classification of unigenes by clusters of orthologous groups, gene ontology, and Kyoto encyclopedia of genes and genomes

The annotated unigenes were aligned to the Clusters of Orthologous Groups of proteins (COG) database for functional prediction and classification. In total, 15,919 unigenes (27.61% of all unigenes) were assigned a Cluster of Orthologous Groups classification (Table 2). Among the 25 categories, transcripts associated with "transcription" (2651; 16.65%), "replication, recombination and repair" (2554; 16.04%), "signal transduction mechanisms" (2135; 13.41%), and "post-translational modification, protein turnover, chaperones" (2039; 12.81%) were dominant. The categories "nuclear structure" (1), "extracellular structures" (5) and "cell motility" (177) represented the smallest groups. However, categories with no concrete assignment, such as "function unknown" (1396; 8.77%) and "general function prediction only" (5232; 32.87%) accounted for a large fraction of the transcripts (Supplemental Table 2; Supplemental Fig. S2).

The Gene Ontology (GO) analysis categorized 36,387 unigenes into 41 functional groups under three main categories (biological processes, cellular components, and molecular functions) (Supplemental Table 3; Fig. 3). Among the biological processes category, "cellular processes" (22,788; 62.63%), "metabolic processes" (22,371; 61.48%), and "single-organism processes" (15,941; 43.81%) were the predominant groups. Notably, unigenes categorized into "reproduction" and "reproductive processes" were also significantly enriched and reached 4193 (11.52%) and 3880 (10.66%) unigenes, respectively. Within the cellular components category, "cell" (27,335; 75.12%), "cell part" (27,335; 75.12%), and "organelle" (21,919; 60.24%) were the most overrepresented groups. For the molecular function category, a large proportion of unigenes were clustered into "binding" (18,630; 51.20%), "catalytic activity" (18,014; 49.51%), and "transporter activity" (2464; 6.78%) groups. Additionally, 1338 (3.68%) unigenes categorized into "nucleic acid binding transcription factors" were identified (Supplemental Table 3; Fig. 3).

To further understand the biological functions of unigenes, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database was used. A total of 24,888 matched sequences were assigned to 128 KEGG pathways (Supplemental Table 4). The most represented pathways included "metabolic pathways" (5252; 21.10%), "biosynthesis of secondary metabolite pathways" (2688; 10.80%), "plant-pathogen interaction" (1623; 6.52%), and "plant hormone signal transduction

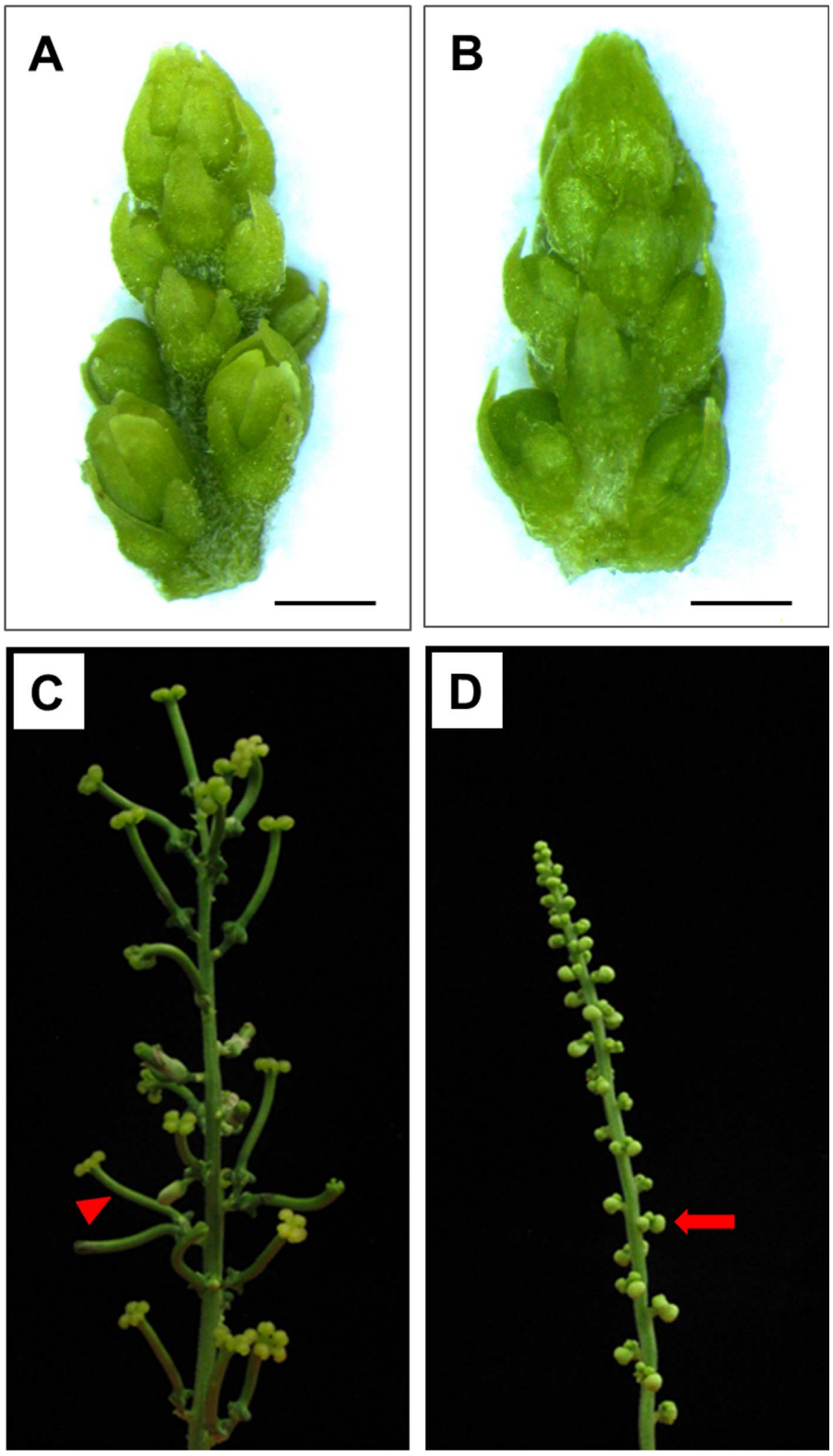


Fig. 1. 6-BA converts male flower buds to female flower buds and induces poly-female inflorescence in *P. volubilis*. A. Female inflorescence buds induced by exogenous 6-BA; B. Control male inflorescence buds; C. Inflorescence from 6-BA-treated plants; D. Inflorescence from control plants. Bar = 1 mm. Red arrow indicates a male flower; red arrowhead indicates a female flower. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pathways" (1515; 6.09%). Notably, unigenes involved in "zeatin biosynthesis" (152; 0.61%) and "indole alkaloid biosynthesis" (23; 0.09%) were also identified (Supplemental Table 4).

3.4. Identification of differentially expressed unigenes between MIB and FIB

Analysis of differentially expressed genes (DEGs) between MIB and FIB could reveal potential key genes underlying floral sex determination in *P. volubilis*. The differences of gene expression levels are plotted with differently colored dots. Red and green dots represent up-regulated and down-regulated DEGs in FIB compared with MIB, respectively. Blue dots represent genes without significant expression changes (Fig. 4A). A total of 1385 unigenes exhibited the threshold ≥two-fold difference in expression levels between MIB and FIB. Among these

Table 1
Summary of MIB and FIB transcriptome assembly.

	Sample	Total Number	Total Length (nt)	Mean Length (nt)	N50	Total Consensus Sequence
Contig	MIB	88,567	41,884,082	473	1059	–
	FIB	91,384	42,311,215	463	1064	–
Unigene	MIB	67,019	50,258,475	750	1226	67,019
	FIB	66,723	51,721,957	775	1260	66,723
	All	57,664	56,464,422	979	1525	57,664

Table 2
Statistics on the number of all the assembled unigenes in two libraries annotated with the National Center for Biotechnology Information (NCBI) non-redundant nucleotide sequence (Nt) database, the NCBI non-redundant protein (Nr) database, Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups of proteins (COG), and Gene Ontology (GO) databases.

Sequences	Nt	Nr	Swiss-Prot	KEGG	COG	GO	All
All Unigenes	43,049	43,690	28,182	24,888	15,919	36,387	45,235
Percent (%)	74.65	75.77	48.87	43.16	27.61	63.10	78.45

DEGs, 451 unigenes were up-regulated, and 934 unigenes were down-regulated in FIB compared with MIB (Fig. 4B), from which candidate genes that might be involved in floral sex determination were selected and are shown below.

3.4.1. Phytohormone biosynthesis-, metabolism-, and signaling-related genes

Phytohormones are essential factors for floral sex determination in various plant species (Golenberg and West, 2013). Among DEGs, two transcripts (Unigene12983 and Unigene21253) encoding ethylene-responsive transcription factor (ERF) homolog (PvoERF109 and PvoERF017) were down-regulated by 3.0- and 2.8-fold, respectively, while one transcript (CL3383.contig1) encoding PvoERF034 was up-regulated by 9.9-fold in FIB compared with MIB (Table 3). Two transcripts (CL1556.contig1 and CL4094.contig1) encoding cytokinin oxidase1 (CKX1) homolog (PvoCKX1) and adenine phosphoribosyltransferase 1 (APT1) homolog (PvoAPT1) were down-regulated by 1.1- and 1.0-fold (Table 3), respectively, and unigene16829 encoding a two-component *Arabidopsis* response regulator 9 (ARR9) homolog protein (PvoARR9) was up-regulated by 4.2-fold in FIB (Table 3). In particular, auxin-related genes, such as auxin-conjugating Gretchen Hagen3.1 homolog gene (*PvoGH3.1*; unigene30072), were differentially up-regulated, reaching 9540.5-fold in FIB compared with MIB (Table 3; Fig. 5). *PvoGH3.3* (unigene31691) and indole-3-acetic acid gene (*IAA*) homolog *PvoIAA14* (CL2060.contig1) were also up-regulated by 6.7- and 4.7-fold, respectively, in FIB compared with MIB (Table 3; Fig. 5). These results indicated that these phytohormone-related genes might be involved in male and female flower bud development in *P. volubilis*.

3.4.2. Transcription factor genes

Transcription factors (TFs) play important roles in the regulation of sex differentiation in plants (Harkess and Leebens-Mack, 2017). Some

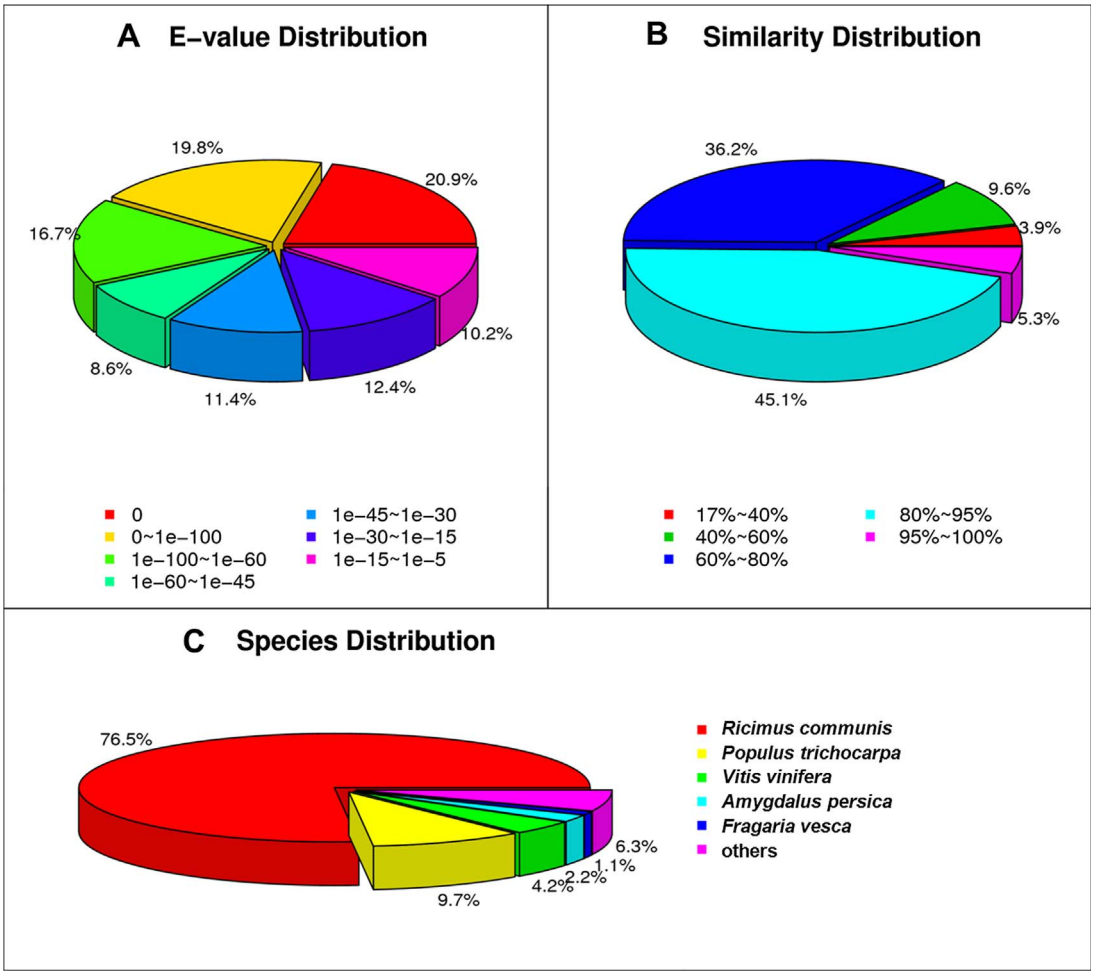


Fig. 2. The annotation and classification of the assembled unigenes in the non-redundant protein databases. E-value distribution (A), similarity distribution (B), and species distribution (C) of the result of non-redundant protein annotation.

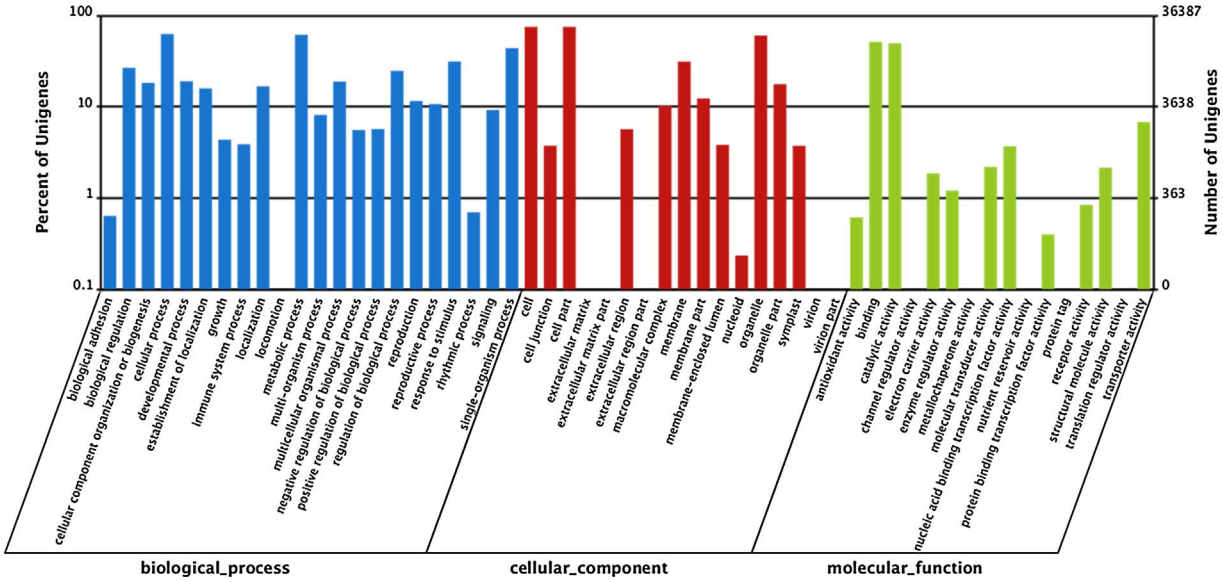


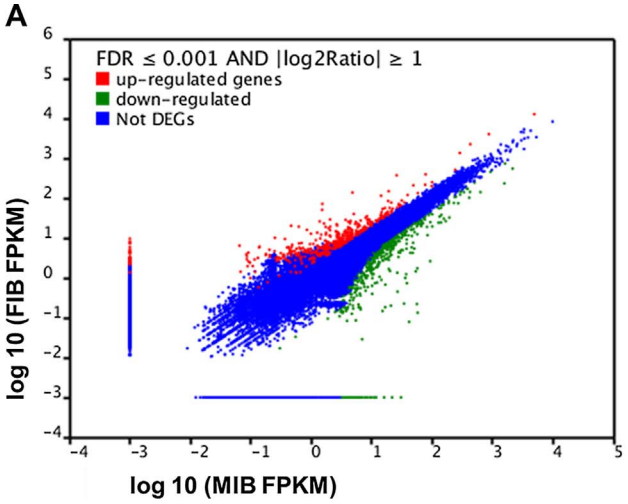
Fig. 3. Gene Ontology categories of the assembled unigenes. Unigenes were assigned to three categories: biological processes, cellular components, or molecular functions.

important TFs were found differentially expressed between MIB and FIB in *P. volubilis*. The expression levels of two transcripts (unigene20388 and CL6356.contig1) encoding an ABORTED MICROSPORES (AMS) homolog (PvoAMS) and a bHLH homolog (PvobHLH) were down-regulated by 5.4- and 5.2-fold, respectively, in FIB compared with MIB (Table 3; Fig. 5). In FIB, the expression levels of two transcripts (unigene30446 and unigene31003) encoding HECATE (HEC) homolog (PvoHEC) were up-regulated by 8191.0- and 42.7-fold, respectively (Table 3; Fig. 5), and the expression levels of two transcripts (CL4156.contig1 and unigene12502) encoding RADIALIS-like (RL) homolog protein (PvoRL) were up-regulated by 46.2- and 27.8-fold, respectively (Table 3). CL5764.contig2 encoding CRABS CLAW (CRC) homolog (PvoCRC) had a 13.5-fold higher expression in FIB compared with MIB (Table 3). Furthermore, CL2866.contig3 and unigene18781 encoding WUSCHEL-related homeobox 9 (WOX9) homolog (PvoWOX9) and SUP homolog (PvoSUP) were also identified as highly expressed in FIB (Table 3; Fig. 5).

The expression analysis of the ABC floral organ identity genes showed that A-class gene *APETALA1* (*AP1*) homolog (*PvoAP1*, CL5426.contig2) was slightly up-regulated in MIB (Table 3; Fig. 5), and B-class gene *APETALA3* (*AP3*) homolog (*PvoAP3*, CL1230.contig3), *PISTILLATA* (*PI*) homolog (*PvoPI*, unigene22152), *PI*-like homolog (*PvoPIL*, unigene14817), and *TOMATO MADS BOX 6* (*TM6*) homolog (*PvoTM6*, unigene12072) were highly expressed in MIB (Table 3; Fig. 5). In contrast, the expression levels of the C-class gene *AGAMOUS* (*AG*) homolog (*PvoAG*, CL7379.contig2) were up-regulated by 0.9-fold in FIB compared with MIB (Table 3; Fig. 5). The expression levels of *AG*-like homolog (*PvoAGL11*, unigene30578) and MADS-box JOINTLESS-like homolog gene (*PvoJOL*, CL5107.contig1) were also up-regulated by 97.4- and 1.1-fold, respectively, in FIB (Table 3; Fig. 5). These MADS-box genes might be responsible for the differentiation of MIB and FIB.

3.4.3. Energy transfer-, storage-, and metabolism-related genes

We also identified some genes that have not previously been characterized in flower development among DEGs. Genes encoding proteins involved in energy transfer, storage, and metabolism were also identified from DEGs between the MIB and FIB. Unigene26431 encoding non-specific lipid-transfer homolog protein (PvoNSLTP) was down-regulated by 526.4-fold in FIB compared with MIB (Table 3). Two transcripts (unigene26980 and CL395.contig3) encoding ABC transporter G family homolog protein (PvoABCG) were down-regulated by 14.9-fold and up-regulated 2.6-fold, respectively, in FIB compared with MIB. *PvoTKPR1* (unigene5320) encoding tetraketide alpha-pyrone reductase 1 homolog was also down-regulated by 10.3- fold in FIB (Table 3; Fig. 5). *PvoUGT85A2* (unigene5320) and *PvoSUC2* (unigene690) encoding UDP-glycosyltransferase 85A2 homolog and sucrose transport



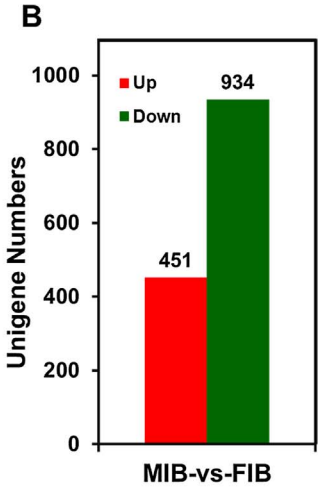


Fig. 4. Comparison of gene expression levels between the MIB library and the FIB library using the FPKM method. A. The difference of gene expression levels was plotted for the MIB library versus the FIB library. The DEGs with FDR ≤0.001 and $|\log_2 \text{ratio}|$ ≥1 are shown in red dots and green dots, respectively. Genes without significant expression changes are shown in blue dots. B. Number of DEGs. Red and green columns represent the number of significantly up-regulated and down-regulated expressed genes in the FIB library compared with the MIB library, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Candidate genes related to flower development and sex differentiation of *P. volubilis*.

Functional group	Gene designation	Unigene ID	MIB (FPKM)	FIB (FPKM)	Log_2 ratio	P-value	Annotation
Phytohormones biosynthesis, metabolism, and signaling related genes	*PvoERF109*	Unigene12983	22.8	5.7	−2.00	6.83E-54	Ethylene-responsive transcription factor ERF109
	PvoERF017	Unigene21253	132.6	34.6	−1.94	4.81E-95	Ethylene-responsive transcription factor ERF017
	PvoCKX1	CL1556.contig1	6.0	2.9	−1.04	1.53E-07	Cytokinin oxidase 1
	PvoAPT1	CL4094.contig1	121.0	60.4	−1.00	5.51E-107	Adenine phosphoribosyltransferase 1
	PvoGH3.1	Unigene30072	0	9.6	13.22	2.02E-39	Auxin conjugating Gretchen Hagen3.1
	PvoERF034	CL3383.contig1	1.3	13.6	3.44	4.81E-46	Ethylene-responsive transcription factor ERF034
	PvoGH3.3	Unigene31691	0.7	5.6	2.95	4.74E-43	Auxin conjugating Gretchen Hagen3.3
	PvoIAA14	CL2060.contig1	2.1	11.5	2.51	1.09E-33	Auxin-responsive protein IAA14
	PvoARR9	Unigene16829	2.6	13.8	2.39	1.36E-17	Two-component response regulator ARR9
Transcription factors genes	*PvoAMS*	Unigene20388	48.2	7.6	−2.67	8.51E-113	ABORTED MICROSPORES
	PvobHLH	CL6356.contig1	30.6	4.9	−2.63	4.46E-130	Basic helix-loop-helix protein
	PvoHEC1	Unigene30446	0	8.2	13.00	7.01E-24	Transcription factor HECATE 1
	PvoRL2	CL4156.contig1	1.5	71.6	5.56	8.17E-264	RADIALIS-like 2
	PvoHEC2	Unigene31003	0.5	20.3	5.45	2.50E-47	Transcription factor HECATE 2
	PvoRL1	Unigene12502	4.8	139.1	4.85	0	RADIALIS-like 1
	PvoCRC	CL5764.contig2	25.8	375.5	3.86	0	CRABS CLAW protein
	PvoWOX9	CL2866.contig3	5.1	40.2	2.99	1.53E-108	WUSCHEL-related homeobox 9
	PvoSUP	Unigene18781	2.1	15.5	2.92	4.97E-47	Transcriptional regulator SUPERMAN
Floral organ identity MADS- box genes	*PvoAP3*	CL1230.contig3	268.9	74.3	−1.86	0	APETALA3
	PvoPIL	Unigene14817	227.1	75.7	−1.59	1.18E-186	PISTILLATA-like
	PvoPI	Unigene22152	420.2	143.1	−1.55	0	PISTILLATA
	PvoTM6	Unigene12072	186.7	74.7	−1.32	1.32E-251	TOMATO MADS BOX 6
	PvoAP1	CL5426.contig2	49.3	28.1	−0.81	1.89E-38	APETALA1
	PvoAGL11	Unigene30578	0.07	6.4	6.62	6.52E-29	AGAMOUS-like 11
	PvoJOL	CL5107.contig1	21.0	43.2	1.04	1.33E-55	MADS-box protein JOINTLESS-like
	PvoAG	CL7379.contig2	76.8	145.8	0.92	1.60E-108	Floral homeotic protein AGAMOUS
Energy transfer, storage and metabolism related genes	*PvoNSLTP*	Unigene26431	41.4	0.1	−9.04	5.47E-155	Non-specific lipid-transfer protein
	PvoABCG26	Unigene26980	10.1	0.6	−3.99	1.06E-79	ABC transporter G family member 26
	PvoTKPR1	Unigene5320	47.5	4.2	−3.50	3.35E-212	Tetraketide alpha-pyrone reductase 1
	PvoPER9	Unigene4280	139.7	24.3	−2.52	0	Peroxidase 9 precursor
	PvoUGT85A2	Unigene31590	1.6	23.1	3.83	3.95E-83	UDP-glycosyltransferase 85A2
	PvoFAD	Unigene17998	2.1	28.9	3.82	4.77E-181	Omega-6 fatty acid desaturase
	PvoBSPA4	Unigene9493	19.0	149.9	2.98	0	Bark storage protein A
	PvoBSPA2	Unigene15165	863.6	4076.1	2.24	0	Bark storage protein A
	PvoABCG40	CL395.contig3	15.1	53.6	1.83	3.23E-155	ABC transporter G family member 40
	PvoSUC2	Unigene690	6.7	15.7	1.24	3.35E-27	Sucrose transport protein SUC2
Other functional genes	*PvoSSG*	Unigene26964	30.4	0	−14.89	8.31E-72	Sulfated surface glycoprotein
	PvoCYS	Unigene27104	16.0	0	−14.0	2.64E-44	Cysteine proteinase inhibitor
	Unknown	CL8116.contig2	11.7	0	−13.52	1.13E-94	Conserved hypothetical protein
	PvoAGG	Unigene26506	11.5	0	−13.49	1.58E-96	Agglutinin family protein
	PvoEXOL	Unigene26711	10.0	0	−13.29	3.41E-54	Exopolygalacturonase-like
	PvoAPG	CL5036.contig1	57.4	0.1	−8.75	1.20E-126	Anther-specific proline-rich protein APG
	PvoBCP	CL5129.contig2	55.7	0.2	−7.97	7.76E-218	Blue copper protein
	PvoCHS1	Unigene27397	34.2	0.2	−7.24	3.25E-301	Chalcone synthase
	PvoCHS2	Unigene5446	53.4	1.5	−5.20	0	Chalcone synthase
	Unknown	Unigene24560	92.2	2.6	−5.14	1.26E-160	Probable F-box protein
	PvoMen-8.2	Unigene12710	83.3	11.8	−2.82	9.00E-184	Men-8 protein precursor
	PvoMen-8.1	Unigene17527	1556.8	234.9	−2.73	0	Men-8 protein precursor
	Unknown	Unigene30610	0.6	19.3	4.95	3.14E-48	Unknown
	PvoCLE43	Unigene19410	0.8	16.0	4.25	9.59E-80	CLAVATA3/ESR (CLE)-related protein 43
	Unknown	Unigene24633	14.2	114.6	3.01	8.10E-89	Unknown
	pvoCYCD1-1	Unigene7730	31.3	87.1	1.48	3.22E-254	Cyclin-D1-1

protein SUC2 homolog were up-regulated by 13.2- and 1.4-fold, respectively, in FIB (Table 3). Two transcripts (unigene9493 and unigene15165) encoding Bark Storage Protein A (BSPA) homolog (PvoBSPA) were also up-regulated by 6.9- and 3.7-fold, respectively, in FIB (Table 3; Fig. 6).

3.4.4. Other functional genes

Some other functional and unknown genes were identified among the DEGs. *PvoSSG* (unigene26964), *PvoCYS* (unigene27104), *PvoAGG* (unigene26506), *PvoEXOL* (unigene26711), and an unknown gene (CL8116.contig2) were specifically and highly expressed in MIB (Table 3). *PvoAPG* (CL5036.contig1) and *PvoBCP* (CL5129.contig2) encoding anther-specific proline-rich homolog protein APG and blue copper homolog protein, respectively, were down-regulated by 429.5- and 249.7-fold in FIB compared with MIB (Table 3). Two *PvoCHS* genes (unigene27397 and unigene5446) encoding chalcone synthase homolog were down-regulated by 150.2- and 35.8-fold (Table 3; Fig. 5), respectively, and two transcripts (unigene12710 and unigene17527) encoding male-enhanced homolog protein (PvoMen-8) were also down-regulated by 6.1- and 5.6-fold, respectively, in FIB (Table 3). *PvoCLE43* (unigene19410) encoding CLAVATA3/ESR (CLE)-related 43 (CLE43) homolog protein was highly up-regulated, and *PvoCYCD1-1*

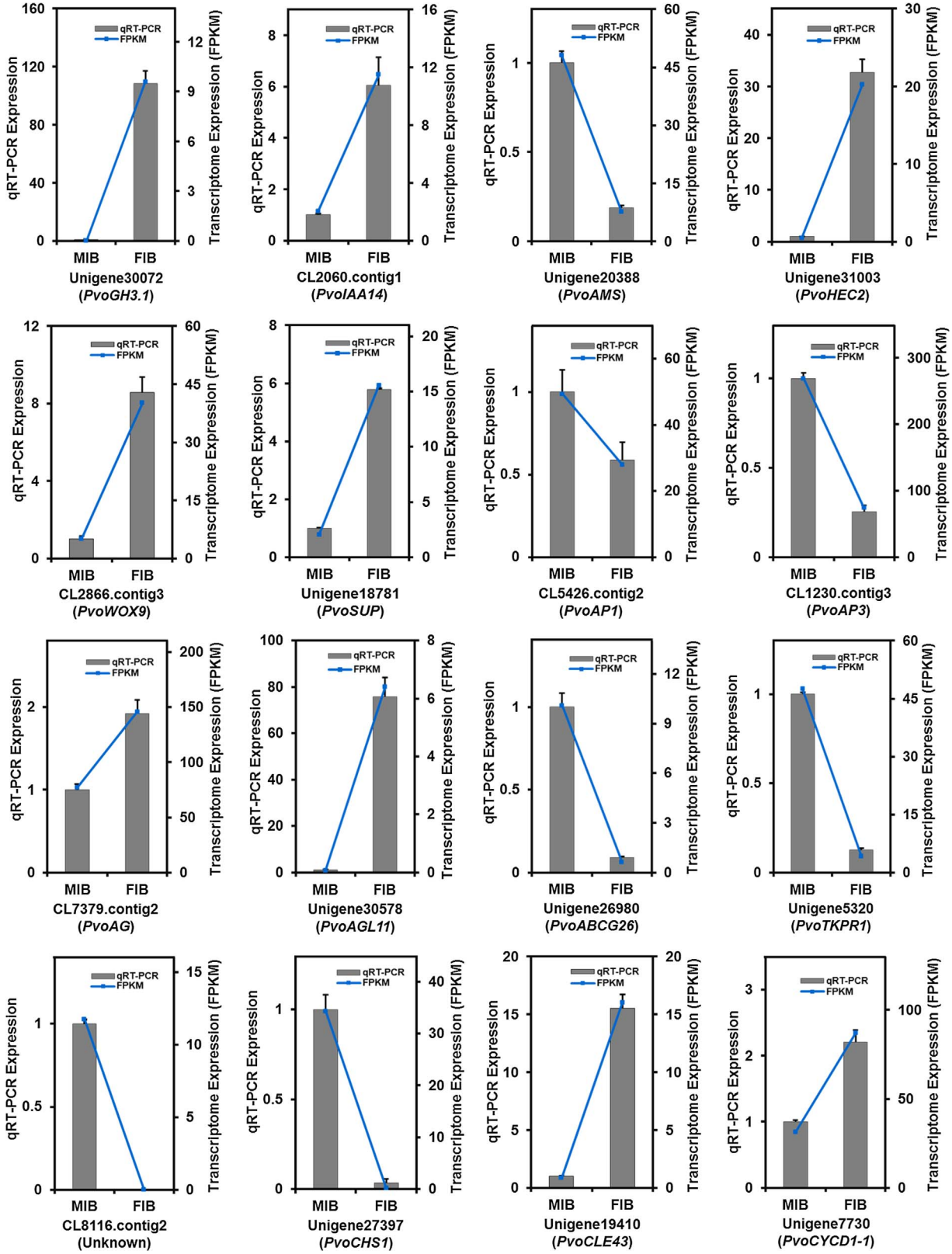


Fig. 5. qRT-PCR validation of differentially expressed unigenes. MIB means male inflorescence buds and FIB means female inflorescence buds. These differentially expressed genes (DEGs) include two auxin-related genes (Unigene30072: *PvoGH3.1* and CL2060.contig1: *PvoIAA14*), four TF genes (unigene20388: *PvoAMS*, unigene31003: *PvoHEC*, CL2866.contig3: *PvoWOX9*, and unigene18781: *PvoSUP*), four MADS-Box genes (CL5426.contig2: *PvoAP1*, CL1230.contig3: *PvoAP3*, CL7379.contig2: *PvoAG*, and unigene30578: *PvoAGL11*), and six other genes (unigene26980: *PvoABCG26*, unigene5320: *PvoTKPR1*, CL8116.contig2: Unknown, Unigene27397: *PvoCHS*, unigene19410: *PvoCLE43*, and unigene7730: *PvoCYCD1-1*).

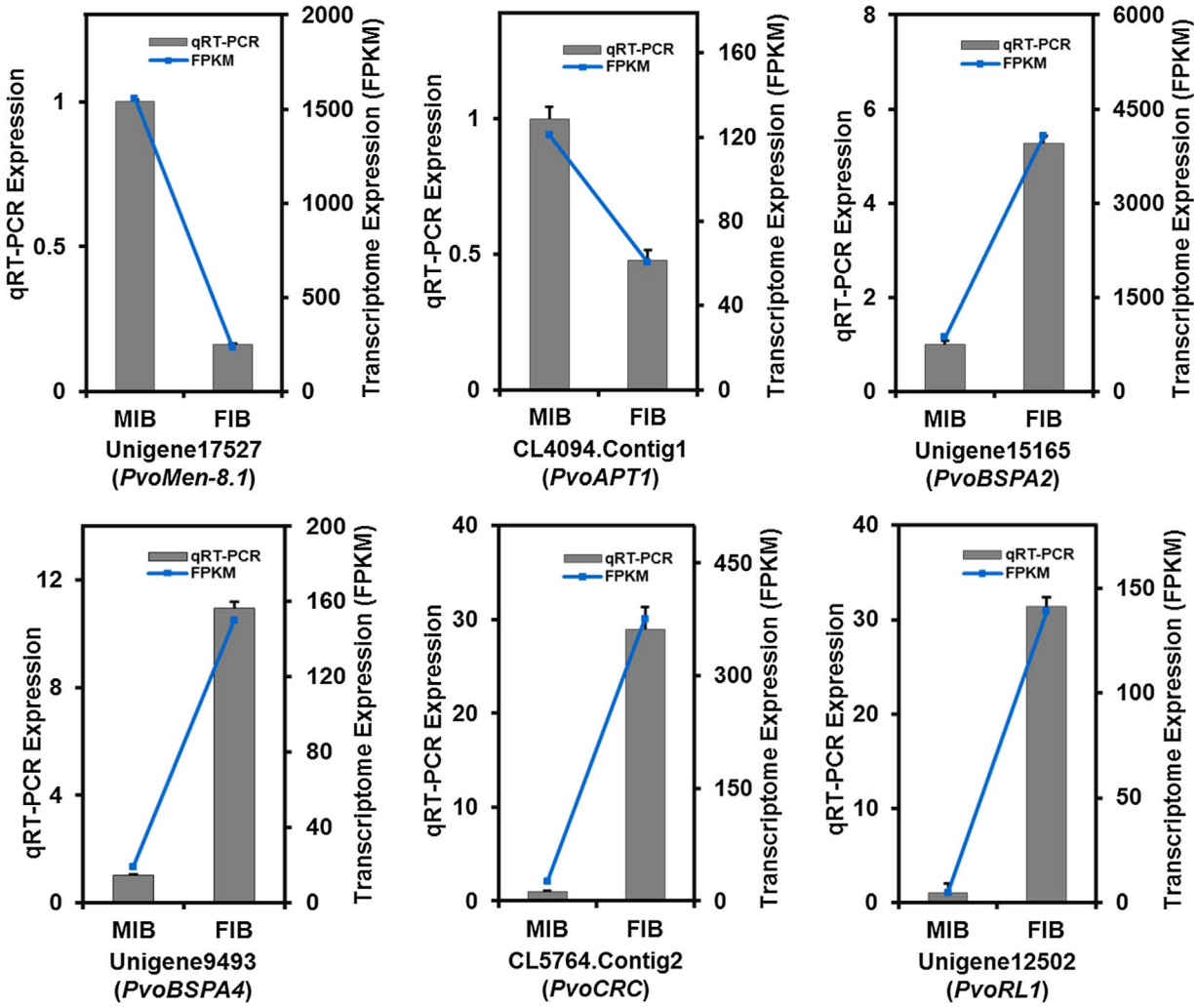


Fig. 6. qRT-PCR validation of highly expressed DEGs (FPKM values ≥ 100). MIB means male inflorescence buds and FIB means female inflorescence buds. These highly expressed DEGs include unigene17527: *PvoMen-8.1*; CL4094.contig1: *PvoAPT1*; unigene15165: *PvoBSPA2*; unigene9493: *PvoBSPA4*; CL5764.contig2: *PvoCRC*; and unigene12502: *PvoRL1*.

(unigene7730) encoding cyclin-D1-1 homolog was slightly up-regulated in FIB (Table 3; Fig. 5).

3.5. Highly expressed DEGs in MIB and FIB

To further select highly expressed DEGs, we used FPKM values ≥100 as the threshold (Sablok et al., 2014). Among the 43 unigenes with FPKM values ≥100 in MIB, 36 unigenes were down-regulated in FIB compared with MIB, such as unigene17527 encoding PvoMen-8.1 (Supplemental Table 5; Fig. 6) and three transcripts (CL3645.contig2, unigene2246, and CL5998.contig1) encoding EXTENSIN (EXT) homolog (PvoEXT) (Supplemental Table 5) required for male fertility. *PvoERF017* (unigene21253) and *PvoAPT1* (CL4094.contig1) were abundant in MIB (Supplemental Table 5; Fig. 6). Consistent with responding to the development of floral buds, four unigenes (unigene22152; CL1230.contig3; unigene14817; and unigene12072) encoding PvoPI, PvoAP3, PvoPIL, and PvoTM6, respectively, were also down-regulated in FIB compared with MIB (Supplemental Table 5). Sixteen unigenes with FPKM values ≥100 were up-regulated in FIB compared with MIB, and four transcripts (CL5018.contig1; unigene15165; unigene12423; and unigene9493) encoding PvoBSPA, CL5764.contig2 encoding PvoCRC, unigene12502 encoding PvoRL1, and CL2088.contig3 encoding a GRF1-interacting factor 1 (GIF1) homolog protein (PvoGIF1) were identified (Supplemental Table 5; Fig. 6).

3.6. Validation of the differentially expressed candidate genes by qRT-PCR analysis

To validate the reliability of the transcriptome sequencing results, a total of 36 genes (Supplemental Table 7) with high or low expression levels were selected for qRT-PCR analysis. The results of 22 genes are shown in Fig. 5 and 6, and the expression patterns of the tested genes were all consistent with the transcriptome data. Meanwhile, a highly significant correlation (Pearson correlation coefficient r = 0.93) of the 36 genes was found between the qRT-PCR results and the transcriptome data (Fig. 7), suggesting that our transcriptome data are reliable and do, in fact, reflect the expression levels of the transcripts.

4. Discussion

Transcriptome sequencing is widely applied in revealing the key genes regulating growth and development of many species. Genetic information on *P. volubilis* is limited; only transcriptome information of its seeds has been reported (Wang et al., 2012). In this study, a comparative transcriptome analysis between MIB and FIB in *P. volubilis* was performed. A total of 104.1 million 90-bp paired-end clean reads were generated and assembled into 57,664 unigenes with a mean size of 979 bp (Table 1), and 45,235 unigenes were successfully annotated in the public databases (Table 2). The species distribution of the annotated unigenes showed that 76.5% of unigenes had the highest similarity with

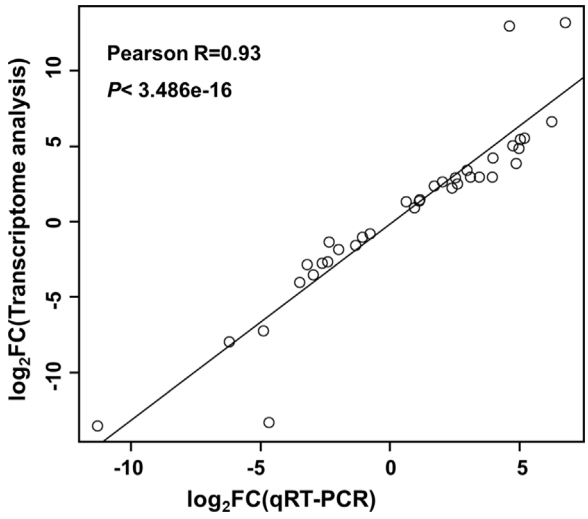


Fig. 7. Correlation between qRT-PCR and the transcriptome data. The values of $\log_2$(relative expression ratios) in qRT-PCR (x-axis) were plotted against the values of $\log_2$(FPKM ratios) in transcriptome data (y-axis) for the 36 selected genes.

sequences from castor bean (Fig. 2C), a species that belongs to the same family (Euphorbiaceae) as *P. volubilis*. These results suggested that the assembly and annotation of the transcriptome were reliable, and the genome of castor bean (Chan et al., 2010) can be used as a reference for the transcriptome analysis of inflorescence buds in *P. volubilis*. The other top-hit species, such as poplar (9.7%), grape (4.2%), peach (2.2%), and woodland strawberry (1.1%) (Fig. 2C), are perennial woody plants in line with *P. volubilis*. These results provided valuable information for selecting candidate genes involved in flower development and sex determination in *P. volubilis*.

Cytokinin is a positive regulator exerting a feminizing effect in *Luffa cylindrica* (Takahashi et al., 1980) and spinach (Chailakhyan and Khryanin, 1978). Treatment with exogenous cytokinin was also found to induce floral feminization in *P. volubilis* (Fu et al., 2014). The breakdown and metabolism of cytokinin is catalyzed by CKXs and APTs (Allen et al., 2002; Bartrina et al., 2011), and *ARRs* are rapidly induced by cytokinin treatment (To and Kieber, 2008). Our data showed that the expression levels of *PvoCKX1* and *PvoAPT1* were down-regulated, whereas *PvoARR9* was up-regulated in FIB (Table 3; Fig. 6), suggesting that cytokinin might play a positive role in female floral bud development, which occurs prior to the development of male floral buds in *P. volubilis*. Auxin also exerts a feminizing effect in *Cannabis sativus* (Heslopharrison, 1956) and *Opuntia stenopetala* (Orozco-Arroyo et al., 2012). Tan et al. (2016) reported that auxin signaling was up-regulated during the transformation from leaf apices to floral buds in the pistillate line compared with the wild monoecious line of *Ricinus communis*. *GH3* genes encode IAA-amido synthetases, which help to maintain auxin homeostasis (Staswick et al., 2005). In this study, several genes, such as *PvoGH3.1*, *PvoGH3.3*, and *PvoIAA14*, were up-regulated in FIB (Table 3; Fig. 5). Therefore, an abundance of auxin signaling in FIB may potentially lead to female tissue development of *P. volubilis*. In addition, ethylene has been proven to have a strong feminizing effect in cucumber (Yamasaki et al., 2001) and melon (Boualem et al., 2008). We found that *PvoERF109* and *PvoERF017* were significantly down-regulated; however, *PvoERF034* was up-regulated in FIB compared with MIB (Table 3). These results indicated that the mechanism of sex differentiation in *P. volubilis* might be different from that of cucumber, melon, or *Opuntia stenopetala*. The crosstalk and balance of multiple phytohormones might affect sex differentiation in *P. volubilis*.

Many TFs exert pivotal roles in flower sex differentiation, such as *QsAMS* in *Quercus suber* (Rocheta et al., 2014), *JcCRC* in *Jatropha curcas* (Xu et al., 2016), and *SlSUP* in *Silene latifolia* (Kazama et al., 2009). AMS is a master regulator that coordinates pollen wall development and sporopollenin biosynthesis in *Arabidopsis* (Xu et al., 2014). The sex-specific candidate gene *PvoAMS*, a homolog of *AMS*, exhibited high expression levels in MIB (Table 3; Fig. 5). The conserved function of *CRC* and its' orthologs have been reported in carpel fusion and style and nectary development (Bowman and Smyth, 1999; Fourquin et al., 2014). We found that *PvoCRC* had a higher expression in FIB compared with MIB (Table 3; Fig. 6). Therefore, *PvoCRC* is expected to exert similar functions in the development of gynoecium in *P. volubilis*. *AtRL2* is expressed in the funiculus of ovules and in embryos (Baxter et al., 2007). *PvoRL1* and *PvoRL2* showed up-regulated expression in FIB (Table 3; Fig. 6). HEC controls gynoecium development by balancing auxin and cytokinin signaling in *Arabidopsis* (Schuster et al., 2015). *PvoHEC1* and *PvoHEC2* were significantly more highly expressed in FIB than in MIB (Table 3; Fig. 5). Two *PvoHECs* might also be involved in mediating auxin and cytokinin signaling pathways during FIB development, in agreement with notably up-regulated expression of the two phytohormone-related genes in FIB, as mentioned above. *WOX9* is required for shoot apical meristem growth and maintenance and embryonic development in *Arabidopsis* (Wu et al., 2005). *PaWOX9* is highly expressed in early embryo development (Palovaara et al., 2010) and is an important regulator of apical-basal embryo pattern establishment in Norway spruce (*Picea abies*) (Zhu et al., 2014). Our results showed that *PvoWOX9* was up-regulated in FIB (Table 3; Fig. 5), suggesting its involvement in female reproductive development in *P. volubilis*. SUP has been reported as an important regulator of floral organ boundaries between whorl 3 and whorl 4 in *Arabidopsis* (Sakai et al., 1995) and increased feminization in monoecious flowers (Nibau et al., 2011). Similarly to *SlSUP* in *S. latifolia* (Kazama et al., 2009), the up-regulated *PvoSUP* in FIB might also play a positive role in female flower development of *P. volubilis*. Moreover, similarly to *HEC*, *SUP* also stimulated auxin- and cytokinin-regulated processes (Nibau et al., 2011). Therefore, *PvoSUP* might regulate FIB development by balancing auxin and cytokinin signaling together with *PvoHEC*.

Generally, differential expression of the ABC floral organ-identity genes, specifically the B-class genes *AP3* and *PI* and the C-class gene *AG*, has been implicated in regulating the development of male and female flowers (Ainsworth et al., 1995; Pfent et al., 2005), as well as in specifying the identity of the stamen and carpel whorls (Coen and Meyerowitz, 1991). Comparing female-specific expression of the *AP1* homolog in *Quercus suber* (Rocheta et al., 2014) and seabuckthorn (Chawla et al., 2015), the expression of *PvoAP1* was slightly lower in FIB than in MIB (Table 3; Fig. 5). As the expression levels of *HoAP3_1*, *HoAP3_2*, and *HoPI_2* from *Hedyosmum orientale* with unisexual inflorescence were higher in male flowers than in female flowers (Liu et al., 2013), the expression of B-class homolog genes, such as *PvoTM6*, *PvoAP3*, *PvoPI*, and *PvoPIL*, was also highly expressed in MIB (Table 3; Fig. 5). These results showed that these B-class homolog genes might be involved in the development of male flowers. In contrast, the expression of C-class homolog gene *PvoAG* was up-regulated in FIB (Table 3; Fig. 5), which is consistent with the significant expression of *HrAG* in seabuckthorn female flower buds at stage II (Chawla et al., 2015). *AGL11* (now named SEEDSTICK) is specifically expressed in developing ovules and associated placental tissues (Rounsley et al., 1995) and is required for ovule identity and normal funiculus development in *Arabidopsis* (Pinyopich et al., 2003). Similarly, *PvoAGL11* had higher expression levels in FIB than in MIB (Table 3; Fig. 5). Thus, these floral organ-identity genes may have crucial roles in establishing floral organ identity and sexual differentiation in male and female flower buds of *P. volubilis*, respectively.

CHS is known to be involved in the metabolic pathway of flavonoids and anthocyanin pigments in several species (Winkel-Shirley, 2002). *QsCHSA* is uniquely and highly expressed in male flowers of *Quercus suber* (Rocheta et al., 2014). Similarly, we found that two *PvoCHS* genes were highly up-regulated in MIB compared with FIB (Table 3; Fig. 5). *Men* is specifically expressed in male flower buds of *Silene latifolia* (Scutt et al., 1997), and, interestingly, we found that two *PvoMen-8* members

were also primarily expressed in MIB (Table 3; Fig. 6). It has been commonly proposed that females expend more resources on seed production than males, and an abundance of energy could promote plants to produce more female flowers (Bickel and Freeman, 1993; Torices and Mendez, 2011). BSPA is involved in nitrogen storage and cycling pathways (Wildhagen et al., 2010; Zhu and Coleman, 2001), and SUC2 serves sucrose transport in phloem (Gould et al., 2012). Correlated with the additional allocation of resources to FIB, we found that four *PvoBSPA* genes and a *PvoSUC2* gene were markedly up-regulated in FIB compared with MIB (Table 3 and Supplemental Table 5). UGT85A2 is mainly expressed in areas of active cell division in *Arabidopsis* (Koroleva et al., 2004), and *CYCD1-1* can accelerate the cell cycle in cultured tobacco BY-2 cells (Woo et al., 2007). In our study, *PvoUGT85A2* and *PvoCYCD1-1* were up-regulated in FIB compared with MIB (Table 3; Fig. 5) and possibly involved in the regulation of cell cycle and division in FIB. Furthermore, we also obtained some genes of unknown function among the DEGs. These genes might play specific roles during the differentiation and development of MIB and/or FIB in *P. volubilis*.

5. Conclusions

Our study provides comprehensive transcripts of inflorescence buds in monoecious species *P. volubilis*. Transcriptome analysis revealed many important genes with known roles in flower development, such as genes involved in phytohormone signaling pathways, transcription regulation, and floral organ identity. Some genes that have not previously been characterized in flower development or in other species were identified, such as unigenes encoding homologs of bark storage protein A, non-specific lipid-transfer protein, CLAVATA3/ESR (CLE)-related protein 43, Cyclin-D1-1, and some proteins of unknown function. In addition, some genes expressed exclusively in each type of MIB and FIB were also identified. These candidate genes provide the foundation for further studies on the molecular mechanisms that regulate flower development and sex differentiation in *P. volubilis*. Further research is needed to investigate the biological function of these genes and the application of these genes in the modification of flower sex.

Author contribution statement

QF and ZFX designed the entire research and analyzed data. NL, CMS, WX, TYB, HH, and PBZ performed experiments and helped to analyze data. QF and ZFX wrote and edited the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

This work was supported by the Natural Science Foundation of China (31300568, 31370595, and 31670612), the Natural Science Foundation of Yunnan Province (2014FB186 and 2016FB051), and the CAS 135 program (2017XTBG-T02). The authors gratefully acknowledge the Central Laboratory of the Xishuangbanna Tropical Botanical Garden for providing the research facilities.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jplph.2017.12.006.

References

Ainsworth, C., Crossley, S., Buchananwollaston, V., Thangavelu, M., Parker, J., 1995. Male and female flowers of the dioecious plant sorrel show different patterns of MADS box gene expression. Plant Cell 7 (10), 1583–1598.
Allen, M., Qin, W.S., Moreau, F., Moffatt, B., 2002. Adenine phosphoribosyltransferase isoforms of *Arabidopsis* and their potential contributions to adenine and cytokinin metabolism. Physiol. Plant. 115 (1), 56–68.
Audic, S., Claverie, J.M., 1997. The significance of digital gene expression profiles. Genome Res. 7 (10), 986–995.
Bai, S.-N., Xu, Z.-H., 2013. Unisexual cucumber flowers, sex and sex differentiation. In: In: Jeon, K.W. (Ed.), International Review of Cell and Molecular Biology 304. pp. 1–55.
Bartrina, I., Otto, E., Strnad, M., Werner, T., Schmulling, T., 2011. Cytokinin regulates the activity of reproductive meristems, flower organsize, ovule formation, and thus seed yield in *Arabidopsis thaliana*. Plant Cell 23 (1), 69–80.
Baxter, C.E.L., Costa, M.M.R., Coen, E.S., 2007. Diversification and co-option of RAD-like genes in the evolution of floral asymmetry. Plant J. 52 (1), 105–113.
Bickel, A.M., Freeman, D.C., 1993. Effects of pollen vector and plant geometry on floral sex ratio in monoecious plants. Am. Midl. Nat. 130 (2), 239–247.
Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A., Morin, H., Sari, M.A., Collin, F., Flowers, J.M., Pitrat, M., Purugganan, M.D., Dogimont, C., Bendahmane, A., 2008. A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science 321 (5890), 836–838.
Bowman, J.L., Smyth, D.R., 1999. CRABS *CLAW*, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. Development 126 (11), 2387–2396.
Chailakhyan, M.K., Khryanin, V.N., 1978. Effect of growth regulators and role of roots in sex expression in spinach. Planta 142 (2), 207–210.
Chan, A.P., Crabtree, J., Zhao, Q., Lorenzi, H., Orvis, J., Puiu, D., Melake-Berhan, A., Jones, K.M., Redman, J., Chen, G., 2010. Draft genome sequence of the oilseed species *Ricinus communis*. Nat. Biotechnol. 28 (9), 951–956.
Chawla, A., Stobdan, T., Srivastava, R.B., Jaiswal, V., Chauhan, R.S., Kant, A., 2015. Sex-biased temporal gene expression in male and female floral buds of seabuckthorn (*Hippophae rhamnoides*). PLoS One 10 (4), e0124890.
Chen, M.-S., Pan, B.-Z., Fu, Q., Tao, Y.-B., Martinez Herrera, J., Niu, L., Ni, J., Dong, Y., Zhao, M.-L., Xu, Z.-F., 2016. Comparative transcriptome analysis between gynoecious and monoecious plants identifies regulatory networks controlling sex determination in *Jatropha curcas*. Front. Plant Sci. 7, 1953.
Chirinos, R., Zuloeta, G., Pedreschi, R., Mignolet, E., Larondelle, Y., Campos, D., 2013. Sacha inchi (*Plukenetia volubilis*): a seed source of polyunsaturated fatty acids, tocopherols, phytosterols, phenolic compounds and antioxidant capacity. Food Chem. 141 (3), 1732–1739.
Coen, E.S., Meyerowitz, E.M., 1991. The war of the whorls: genetic interactions controlling flower development. Nature 353 (6339), 31–37.
Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21 (18), 3674–3676.
Dong, Y., Chen, M., Wang, X., Niu, L., Fu, Q., Xu, Z.-F., 2016. Establishment of *in vitro* regeneration system of woody oil crop *Plukenetia volubilis*. Mol. Plant Breed. 14 (2), 462–470.
Fourquin, C., Primo, A., Martinez-Fernandez, I., Huet-Trujillo, E., Ferrandiz, C., 2014. The CRC orthologue from *Pisum sativum* shows conserved functions in carpel morphogenesis and vascular development. Ann. Bot. 114 (7), 1535–1544.
Fu, Q., Niu, L., Zhang, Q., Pan, B.-Z., He, H., Xu, Z.-F., 2014. Benzyladenine treatment promotes floral feminization and fruiting in a promising oilseed crop *Plukenetia volubilis*. Ind. Crop. Prod. 59, 295–298.
Gillespie, L.J., 1993. A synopsis of neotropical *Plukenetia* (Euphorbiaceae) including two new species. Syst. Bot. 18 (4), 575–592.
Golenberg, E.M., West, N.W., 2013. Hormonal interactions and gene regulation can link monoecy and environmental plasticity to the evolution of dioecy in plants. Am. J. Bot. 100 (6), 1022–1037.
Gould, N., Thorpe, M.R., Pritchard, J., Christeller, J.T., Williams, L.E., Roeb, G., Schurr, U., Minchin, P.E.H., 2012. AtSUC2 has a role for sucrose retrieval along the phloem pathway: evidence from carbon-11 tracer studies. Plant Sci. 188, 97–101.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29 (7), 644–652.
Gremski, K., Ditta, G., Yanofsky, M.F., 2007. The HECATE genes regulate female reproductive tract development in *Arabidopsis thaliana*. Development 134 (20), 3593–3601.
Gutierrez, L.-F., Rosada, L.-M., Jimenez, A., 2011. Chemical composition of sacha inchi (*Plukenetia volubilis* L.) seeds and characteristics of their lipid fraction. Grasas Aceites 62 (1), 76–83.
Hanssen, H.-P., Schmitz-Huebsch, M., 2011. Sacha inchi (*Plukenetia volubilis* L.) nut oil and its therapeutic and nutritional uses. In: Preedy, V.R., Watson, R.R., Patel, V.B. (Eds.), Nuts and Seeds in Health and Disease Prevention. Elsevier Academic Press Inc., San Diego, USA, pp. 991–994.
Harkess, A., Leebens-Mack, J., 2017. A century of sex determination in flowering plants. J. Hered. 108 (1), 69–77.
Heslopharrison, J., 1956. Auxin and sexuality in *Cannabis sativa*. Physiol. Plant. 9 (4), 588–597.
Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 40 (D1), D109–D114.
Kazama, Y., Fujiwara, M.T., Koizumi, A., Nishihara, K., Nishiyama, R., Kifune, E., Abe, T., Kawano, S., 2009. A SUPERMAN-like gene is exclusively expressed in female flowers of the dioecious plant *Silene latifolia*. Plant Cell Physiol. 50 (6), 1127–1141.
Koroleva, O.A., Tomlinson, M., Parinyapong, P., Sakvarelidze, L., Leader, D., Shaw, P., Doonan, J.H., 2004. CycD1, a putative G1 cyclin from *Antirrhinum majus*, accelerates the cell cycle in cultured tobacco BY-2 cells by enhancing both G1/S entry and progression through S and G2 phases. Plant Cell 16 (9), 2364–2379.
Liu, S.J., Sun, Y.H., Du, X.Q., Xu, Q.J., Wu, F., Meng, Z., 2013. Analysis of the APETALA3- and PISTILLATA-like genes in *Hedyosmum orientale* (Chloranthaceae) provides insight

into the evolution of the floral homeotic B-function in angiosperms. Ann. Bot. 112 (7), 1239–1251.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\triangle\triangle CT}$ method. Methods 25 (4), 402–408.
Nibau, C., Di Stilio, V.S., Wu, H.M., Cheung, A.Y., 2011. *Arabidopsis* and tobacco *SUPERMAN* regulate hormone signalling and mediate cell proliferation and differentiation. J. Exp. Bot. 62 (3), 949–961.
Niu, L.J., Li, J.L., Chen, M.S., Xu, Z.F., 2014. Determination of oil contents in sacha inchi (*Plukenetia volubilis*) seeds at different developmental stages by two methods: soxhlet extraction and time-domain nuclear magnetic resonance. Ind. Crop. Prod. 56, 187–190.
Niu, L.J., Tao, Y.B., Chen, M.S., Fu, Q.T., Li, C.Q., Dong, Y.L., Wang, X.L., He, H.Y., Xu, Z.F., 2015. Selection of reliable reference genes for gene expression studies of a promising oilseedcrop, *Plukenetia volubilis*, by real-time quantitative PCR. Int. J. Mol. Sci. 16 (6), 12513–12530.
Orozco-Arroyo, G., Vazquez-Santana, S., Camacho, A., Dubrovsky, J.G., Cruz-Garcia, F., 2012. Inception of maleness: auxin contribution to flower masculinization in the dioecious cactus *Opuntia stenopetala*. Planta 236 (1), 225–238.
Palovaara, J., Hallberg, H., Stasolla, C., Hakman, I., 2010. Comparative expression pattern analysis of *WUSCHEL-related homeobox* 2 (WOX2) and WOX8/9 in developing seeds and somatic embryos of the gymnosperm Picea abies. New Phytol. 188 (1), 122–135.
Pan, B.-Z., Chen, M.-S., Ni, J., Xu, Z.-F., 2014. Transcriptome of the inflorescence meristems of the biofuel plant *Jatropha curcas* treated with cytokinin. BMC Genomics 15, 974.
Pertea, G., Huang, X.Q., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J., Quackenbush, J., 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19 (5), 651–652.
Pfent, C., Pobursky, K.J., Sather, D.N., Golenberg, E.M., 2005. Characterization of *SpAPETALA3* and *SpPISTILLATA*, B class floral identity genes in Spinacia oleracea, and their relationship to sexual dimorphism. Dev. Genes Evol. 215 (3), 132–142.
Pinyopich, A., Ditta, G.S., Baumann, E., Wisman, E., Yanofsky, M.F., 2003. Unraveling the redundant roles of MADS-box genes during carpel and fruit development. Nature 424 (6944), 85–88.
Ramos, M.J.N., Coito, J.L., Silva, H.G., Cunha, J., Costa, M.M.R., Rocheta, M., 2014. Flower development and sex specification in wild grapevine. BMC Genomics 15 (1), 1095.
Ramos, M.J.N., Coito, J.L., Fino, J., Cunha, J., Silva, H., de Almeida, P.G., Costa, M.M.R., Amâncio, S., Paulo, O.S., Rocheta, M., 2017. Deep analysis of wild Vitis flower transcriptome reveals unexplored genome regions associated with sex specification. Plant Mol. Biol. 93 (1–2), 151–170.
Rocheta, M., Sobral, R., Magalhaes, J., Amorim, M.I., Ribeiro, T., Pinheiro, M., Egas, C., Morais-Cecilio, L., Costa, M.M.R., 2014. Comparative transcriptonnic analysis of male and female flowers of monoecious *Quercus suber*. Front. Plant Sci. 5, 599.
Rounsley, S.D., Ditta, G.S., Yanofsky, M.F., 1995. Diverse roles for MADS box genes in *Arabidopsis* development. Plant Cell 7 (8), 1259–1269.
Sablok, G., Fu, Y., Bobbio, V., Laura, M., Rotino, G.L., Bagnaresi, P., Allavena, A., Velikova, V., Viola, R., Loreto, F., Li, M., Varotto, C., 2014. Fuelling genetic and metabolic exploration of C3 bioenergy crops through the first reference transcriptome of *Arundo donax* L. Plant Biotechnol. J. 12 (5), 554–567.
Sakai, H., Medrano, L.J., Meyerowitz, E.M., 1995. Role of SUPERMAN in maintaining *Arabidopsis* floral whorl boundaries. Nature 378 (6553), 199–203.
Schuster, C., Gaillochet, C., Lohmann, J.U., 2015. *Arabidopsis HECATE* genes function in phytohormone control during gynoecium development. Development 142 (19), 3343–3350.
Scutt, C.P., Li, Y., Robertson, S.E., Willis, M.E., Gilmartin, P.M., 1997. Sex determination in dioecious *Silene latifolia* (Effects of the Y chromosome and the parasitic smut fungus (*Ustilago violacea*) on gene expression during flower development). Plant Physiol. 114 (3), 969–979.
Shukla, A., Singh, V.K., Bharadwaj, D.R., Kumar, R., Rai, A., Rai, A.K., Mugasimangalam, R., Parameswaran, S., Singh, M., Naik, P.S., 2015. *De novo* assembly of bitter gourd transcriptomes: gene expression and sequence variations in gynoecious and monoecious lines. PLoS One 10 (6), e0128331.
Song, Y., Ma, K., Ci, D., Chen, Q., Tian, J., Zhang, D., 2013. Sexual dimorphic floral development in dioecious plants revealed by transcriptome, phytohormone, and DNA methylation analysis in *Populus tomentosa*. Plant Mol. Biol. 83 (6), 559–576.
Staswick, P.E., Serban, B., Rowe, M., Tiryaki, I., Maldonado, M.T., Maldonado, M.C., Suza, W., 2005. Characterization of an *Arabidopsis* enzyme family that conjugates amino acids to indole-3-acetic acid. Plant Cell. 17 (2), 616–627.
Takahashi, H., Suge, H., Saito, T., 1980. Sex expression as effected by N^6-benzylaminopurine in staminate inflorescence of *Luffa cylindrica*. Plant Cell Physiol. 21 (4), 525–536.
Tan, M.L., Xue, J.F., Wang, L., Huang, J.X., Fu, C.L., Yan, X.C., 2016. Transcriptomic analysis for different sextypes of *Ricinus communis* L. during development from apical buds to inflorescences by digital gene expression profiling. Front. Plant Sci. 6, 1208.
Tanurdzic, M., Banks, J.A., 2004. Sex-determining mechanisms in land plants. Plant Cell 16, S61–S71.
Thompson, B., 2014. Genetic and hormonal regulation of maize inflorescence development. In: Fornara, F. (Ed.), Molecular Genetics of Floral Transition and Flower Development. Academic Press Ltd-Elsevier Science Ltd, London, pp. 263–296.
To, J.P.C., Kieber, J.J., 2008. Cytokinin signaling: two-components and more. Trends Plant Sci. 13 (2), 85–92.
Torices, R., Mendez, M., 2011. Influence of inflorescence size on sexual expression and female reproductive success in a monoecious species. Plant Biol. 13 (s1), 78–85.
Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28 (5), 511–515.
Wang, X., Xu, R., Wang, R., Liu, A., 2012. Transcriptome analysis of sacha inchi (*Plukenetia volubilis* L.) seeds at two developmental stages. BMC Genomics 13, 716.
Wildhagen, H., Duerr, J., Ehlting, B., Rennenberg, H., 2010. Seasonal nitrogen cycling in the bark of field-grown Grey poplar is correlated with meteorological factors and gene expression of bark storage proteins. Tree Physiol. 30 (9), 1096–1110.
Winkel-Shirley, B., 2002. Biosynthesis of flavonoids and effects of stress. Curr. Opin. Plant Biol. 5 (3), 218–223.
Woo, H.H., Byeong, R.J.B., Hirsch, A.M., Hawes, M.C., 2007. Characterization of *Arabidopsis* AtUGT85A and AtGUS gene families and their expression in rapidly dividing tissues. Genomics 90 (1), 143–153.
Wu, T., Qin, Z., Zhou, X., Feng, Z., Du, Y., 2010. Transcriptome profile analysis of floral sex determination in cucumber. J. Plant Physiol. 167 (11), 905–913.
Wu, X.L., Dabi, T., Weigel, D., 2005. Requirement of homeobox gene *STIMPY/WOX9* for *Arabidopsis* meristem growth and maintenance. Curr. Biol. 15 (5), 436–440.
Xu, G., Huang, J., Yang, Y., Yao, Y.A., 2016. Transcriptome analysis of flower sex differentiation in *Jatropha curcas* L. using RNA sequencing. PLoS One 11 (2), e0145613.
Xu, J., Ding, Z.W., Vizcay-Barrena, G., Shi, J.X., Liang, W.Q., Yuan, Z., Werck-Reichhart, D., Schreiber, L., Wilson, Z.A., Zhang, D.B., 2014. ABORTED MICROSPORES acts as a master regulator of pollen wall formation in *Arabidopsis*. Plant Cell 26 (4), 1544–1556.
Yamasaki, S., Fujii, N., Matsuura, S., Mizusawa, H., Takahashi, H., 2001. The *M* locus and ethylene-controlled sex determination in andromonoecious cucumber plants. Plant Cell Physiol. 42 (6), 608–619.
Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., Wang, J., Li, S., Li, R., Bolund, L., Wang, J., 2006. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 34 (s2), W293–W297.
Yin, T.J., Quinn, J.A., 1995. Tests of a mechanistic model of one hormone regulating both sexes in *Cucumis sativus* (Cucurbitaceae). Am. J. Bot. 82 (12), 1537–1546.
Zhu, B.L., Coleman, G.D., 2001. The poplar bark storage protein gene (*Bspa*) promoter is responsive to photoperiod and nitrogen in transgenic poplar and active in floral tissues, immature seeds and germinating seeds of transgenic tobacco. Plant Mol. Biol. 46 (4), 383–394.
Zhu, T.Q., Moschou, P.N., Alvarez, J.M., Sohlberg, J.J., von Arnold, S., 2014. *WUSCHEL-RELATED HOMEOBOX 8/9* is important for proper embryo patterning in the gymnosperm Norway spruce. J. Exp. Bot. 65 (22), 6543–6552.