Electronic Journal of Biotechnology 29 (2017) 39–46

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Research article

Transcriptome analysis of female and male flower buds of *Idesia polycarpa* Maxim. var. *vestita* Diels

Lanju Mei [1], Na Dong [1], Fosheng Li, Na Li, Min Yao, Fang Chen *, Lin Tang *

Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, College of Life Science, Sichuan University, Chengdu 610000, Sichuan, China

ARTICLE INFO

Article history:
Received 12 May 2017
Accepted 6 June 2017
Available online 13 July 2017

Keywords:
Dioecious
Floral homeotic genes
Flower development
Functional annotation
Next generation sequencing technology
RNA sequencing (RNA-seq)
Sex determination
Sex differentiation
Simple sequence repeats (SSR)
Transcription factors
Transcriptome profiling

ABSTRACT

Background: Idesia polycarpa Maxim. var. *vestita* Diels, a dioecious plant, is widely used for biodiesel due to the high oil content of its fruits. However, it is hard to distinguish its sex in the seedling stage, which makes breeding and production problematic as only the female tree can produce fruits, and the mechanisms underlying sex determination and differentiation remain unknown due to the lack of available genomic and transcriptomic information. To begin addressing this issue, we performed the transcriptome analysis of its female and male flower.

Results: 28,668,977 and 22,227,992 clean reads were obtained from the female and male cDNA libraries, respectively. After quality checks and *de novo* assembly, a total of 84,213 unigenes with an average length of 1179 bp were generated and 65,972 unigenes (78.34%) could be matched in at least one of the NR, NT, Swiss-Prot, COG, KEGG and GO databases. Functional annotation of the unigenes uncovered diverse biological functions and processes, including reproduction and developmental process, which may play roles in sex determination and differentiation. The Kyoto Encyclopedia of Genes and Genomes pathway analysis showed many unigenes annotated as metabolic pathways, biosynthesis of secondary metabolites pathways, plant–pathogen interaction, and plant hormone signal transduction. Moreover, 29,953 simple sequence repeats were identified using the microsatellite software.

Conclusion: This work provides the first detailed transcriptome analysis of female and male flower of *I. polycarpa* and lays foundations for future studies on the molecular mechanisms underlying flower bud development of *I. polycarpa*.

1. Introduction

Flower development is a complex process, including flowering induction, development of floral meristem identity, and floral organ development and is initiated when the plant meristem changes its identity from vegetative to reproductive growth, this transition which is triggered by five genetically defined pathways in *Arabidopsis thaliana* [1]. Floral repressor can maintain a vegetative state in the center of the shoot apical meristem [2]. However, floral integrating proteins, such as FLOWERING LOCUS T (FT) and SUPPRESSOR OF OVEREXPRESSION OF CO1 (SOC1) can accept signals from the flowering control genetic pathway and activate the floral meristem identity genes *LEAFY* (*LFY*) and *APETALA1* (*AP1*) [3]. LFY, in turn, activates the floral organ identity genes, including a set of floral homeotic genes [4]. Development of a functional androecium and gynoecium involves a large number of genes and mutations in any of the many regulatory genes have the potential to trigger abortion or loss of function of male and/or female organs [5]. In dioecious plants, two broad categories of unisexual flowers have been recognized: the 'type I' category includes flowers that are unisexual through abortion of reproductive organs, and the 'type II' category includes flowers that are unisexual from inception [6]. Transcription factors (TFs), which are usually classified by their DNA-binding domains, play crucial roles in regulating flower development. For example, the MADS-box family represents the best studied floral gene family, of which multiple members are crucial for floral development [7], and the MYB family has been reported to be involved in stamen and ovule development [8]. Phytohormones are non-genic elements important to plasticity in floral developmental pathways and they also function in the maturation of unisexual inflorescences [9]. Previous research has reported that ethylene promotes female sex expression in cucumber [10]. Chen et al. reported that application of exogenous cytokinin (6-benzyladenine, BA) on inflorescence buds of *Jatropha curcas* can significantly increase the number of female flowers [11]. In *Populus tomentosa* [12], endogenous gibberellins (GA) and auxin (IAA)

* Corresponding authors.
E-mail addresses: fangchenscu@163.com (F. Chen), tanglin@scu.edu.cn (L. Tang).
[1] These authors contributed equally to this work.
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

http://dx.doi.org/10.1016/j.ejbt.2017.07.002

contents were higher in male flowers than in female flowers during development. In general, the expression of TFs and plant hormones can affect a mechanism that may be important for flower development and sex determination in plants.

With the advent of next generation sequencing technology, a novel approach to transcriptome profiling, called RNA sequencing (RNA-seq) has emerged; advantages over traditional techniques include rapid processing, high throughput, and greater cost effectiveness [13]. For plants, particularly those without a whole-genome database, transcriptome-wide RNA-seq analysis is an efficient method for obtaining information on unigenes involved in developmental processes, such as flower development [14]. Recently, deep sequencing has been widely used to study the transcriptome dynamics of flower development, both in herbaceous plants, such as cucumber [15] and wheat [16], and in woody trees, such as Populus [12], *Quercus suber* [17] and Metasequoia [18].

Idesia polycarpa Maxim. var. *vestita* Diels is a dioecious tree of the Flacourtiaceae family, which is native to eastern Asia including China, Japan, and Korea [19]. This tree can be 8–21 m high and the fruit has a high oil content, confirmed to be edible, and has the potency to be useful in preparation of biodiesel [20]. However, it's difficult to identify its sex during the long juvenile stage, and the reproductive maturity of seedlings takes 4 or 5 years, before the small flower buds begin to appear in the panicle rachis in early April and bloom three weeks later [16]. However, because only female plants produce flowers and subsequently fruit, knowing the sexual identity of the trees in the seedling stage would be advantageous for optimal production, and also for breeding purposes. Flowers of *I. polycarpa*, lack of petals, are imperfect; the early development of female flowers includes bisexual tissues, with male sexual degradation occurring at later developmental phases; and early development of male flowers similarly includes bisexual tissues, with the presence of a rudimentary pistillode in male flowers at later developmental phases. However, currently available genomic and genetic information for *I. polycarpa* are limited, and genes implicated in flower development and organ abortion in male and female flowers of *I. polycarpa* have not yet been reported. In the present work, flower buds of both sexes of *I. polycarpa* were collected, and RNA-seq was performed on the Illumina HiSeq 2000 platform. Many candidate genes for flower development were identified. This represents a valuable resource for future investigating the molecular mechanisms involved in female and male floral development in *I. polycarpa*.

2. Materials and methods

2.1. Plant materials

I. polycarpa flower buds were collected from female and male *Idesia polycarpa* Maxim. var. *vestita* Diels trees of the same age and the height of 5 m in Ziyang, Sichuan, Southwest China (Fig. 1a). The female and male panicles were both nearly 20 cm (Fig. 1b and c) with flower buds about 5 mm in diameter at the late developmental stage before blooming (Fig. 1d). After the calyx of those flower buds removed and observed under a stereo microscope (Olympus, Tokyo, Japan), the abortive stamens were obviously found in female flower buds (Fig. 1e) and the arrested pistil in male flower buds (Fig. 1f). Compared with normal stamens in male buds, the anthers in the female buds were degenerated and whitish in color; similarly, the pistil in the male buds also stopped developing with thin and small in size. Samples were immediately frozen in liquid nitrogen and then stored at -80°C until use.

2.2. RNA extraction, cDNA library preparation and Illumina sequencing

Total RNA of each sample was isolated from male and female buds using the TRIzol Reagent (Invitrogen, USA) according to the manufacturer's protocol. The quality and quantity of RNA were evaluated by 1% agarose gel electrophoresis and a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA).

To eliminate the individual differences, equal amounts (2 μg) of high quality total RNA of three replicates of each sample were mixed to get pooled samples. A total amount of 2 μg RNA of each pooled sample was used as input material for generating sequencing librariesas using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA)

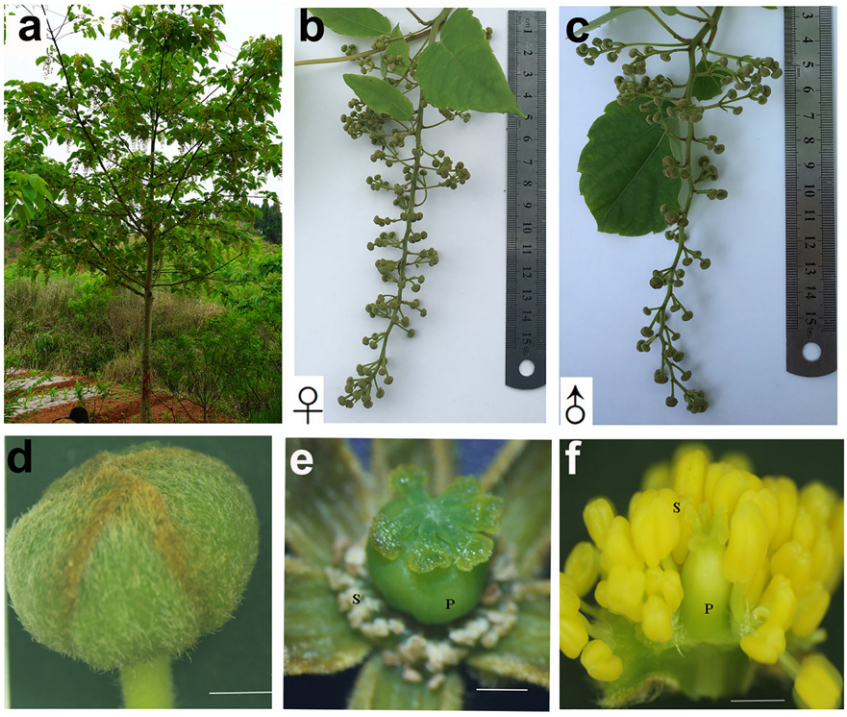

Fig. 1. Morphological structures of female and male flower buds of *I. polycarpa*. (a) *I. polycarpa* tree with the height of 5 m. (b and c) The female and male panicle at the late developmental stage with the length of nearly 20 cm, respectively. (d) Flower bud with 5 mm in diameter. (e) Female bud. (f) Male bud. S: stamen; P: Pistil. Bars = 1 mm.

following manufacturer's instruction. The two (male and female) cDNA libraries were sequenced using the Illumina HiSeq™ 2000 platform (BGI, Shenzhen, China).

2.3. Data filtering and de novo assembly

Raw reads were generated by the Illumina instrument software, and then were processed through in-house Perl scripts to remove reads containing adapter and low-quality reads to obtain clean reads. The Q20, GC-content, and sequence duplication level in the clean reads were calculated. All downstream analyses were based on clean reads with high quality. Transcriptome *de novo* assembly of the clean reads was carried out with Trinity software (release-20130225) [21]. The reads with overlapping segments were assembled into contigs. Then the contigs were assembled into unigenes by paired-end assembly and gap filling. The distribution of contigs and unigenes lengths was also calculated as an evaluation standard for assembly quality.

2.4. Functional annotation and single sequence repeat (SSR) analysis

To identify unigene putative functions, all unigenes were aligned with the NCBI non-redundant protein database (NR, release-20130408), Swiss-Prot protein database (release-2013_03), the Cluster of Orthologous Groups of proteins database (COG, release-20090331) and Kyoto Encyclopedia of Genes and Genomes database (KEGG, release 63.0) using BLASTX with the E-value cutoff <1e-5. Unigenes were also aligned with non-redundant nucleotide sequence database (NT, release-20130408) in GenBank by BLASTn (E-value <1e-5). On the basis of NR annotations, the Blast2GO program (version 2.5.0, http://www.blast2go.com/b2ghome) was used to obtain the gene ontology (GO) annotations of unigenes [22]. Subsequently, the WEGO software program was used to perform GO functional classification (biological process, molecular function and cellular component) for all unigenes and to understand the distribution of gene functions of *I. polycarpa* from the macro level [23]. In addition, KEGG, a major public pathway-related database [24], was used to further study the complex biological behaviors of unigenes related to flower development in *I. polycarpa*. Moreover, MicroSAtellite software (MISA) was employed to identify SSR in the assembled unigenes of *I. polycarpa*, the repeat thresholds for Mono-, Di-, Tri-, Quad-, Penta- and Hexa-nucleotide motifs with a minimum of 12, 6, 5, 4, 3 and 3, respectively.

3. Results and discussion

3.1. Illumina sequencing and de novo assembly

A total of 55,100,394 raw reads were acquired by Illumina paired-end sequencing and submitted to the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra/) under the accession number SRP065835. After stringent quality assessment and data filtering, 28,668,977 and 22,227,992 clean reads of female and male buds were harvested, respectively, indicating that the sequencing was qualified for subsequent assembly (Table 1). Because no reference genome exists for *I. polycarpa*, the high-quality reads from two cDNA libraries were combined and *de novo* assembled into a reference transcriptome using Trinity software [21]. This assembly yielded 405,212 contigs with a maximum length of 13,746 bp, a minimum length of 101 bp, an average length of 247 bp and an N50 of 332 bp and 84,213 unigenes with a maximum length of 13,746 bp, a minimum length of 300 bp, an average length of 1179 bp and an N50 of 1767 bp (Table 2). In addition, 57,837 unigenes (68.7%) are longer than 500 nt and 36,119 unigenes (42.9%) are longer than 1000 nt (Table 2). These quality statistics indicated the sequencing and assembly results were qualified for further analysis.

Table 1
Summary of Illumina transcriptome sequencing for *I. Polycarpa.*

Sample	Total reads	Total bases	Q20%	GC %	N %
Female	28,668,977	5,791,133,354	98.73%	43.03%	0
Male	22,227,992	4,490,054,384	98.73%	43.11%	0

3.2. Functional annotation of the unigenes

3.2.1. Unigene sequence similarity analysis

To provide putative annotations for the unigenes, all of the assembled unigenes were aligned to the protein database NR, Swiss-Prot, KEGG and COG using BLASTx and the nucleotide sequence database NT by BLASTn. A total of 65,972 unigenes (78.34% of all 84,213 unigenes) were annotated, including 55,539 in NR, 64,763 in NT, 33,864 in Swiss-Prot, 31,387 in KEGG, 20,214 in COG, and 45,298 in GO (Table 3). However, the lack of a complete genomic or transcriptomic reference sequence for *I. polycarpa* increased the difficulty of unigene annotation. Thus, it is acceptable that some unigenes had no annotation in those databases and some unmatched unigenes may represent lineage-specific genes that have not been previously characterized.

The E-value and similarity distribution of the best hits against the NR database were shown in Fig. 2a and Fig. 2b, respectively. For species distribution, *Populus balsamifera subsp. trichocarpa* provided the best BLASTx matches with 44,087 unigenes (79.38%). The second closest species was *Ricinus communis* with 5104 unigenes (9.19%), followed by *Vitis vinifera* (3.95%), *Amygdalus persica* (1.56%) and other species (5.92%) (Fig. 2c).

To classify the function of the *I. polycarpa* unigenes, GO terms were assigned using the Blast2go tool based on the NR protein sequence database annotation of the unigenes. This analysis provided hierarchical relationships that represent information on molecular functions, cellular components and biological processes and 45,298 unigenes was obtained and then classified into 55 GO terms (Fig. 3, Table S1). The GO terms with the most unigenes were 'cell' and 'cell part' both belonged to the cellular component category, with 33,882 (74.80%) and 33,880 (74.79%) unigenes respectively. In the biological process category, 'cellular process' and 'metabolic process' covered 28,706 (63.37%) and 28,202 (62.26%) unigenes, respectively. In the molecular function category, 'binding' (23,465 unigenes, 51.80%) and 'catalytic activity' (22,295 unigenes, 49.22%) were the most abundant. Furthermore, a number of unigenes associated with 'developmental process' (8403 unigenes, 18.55%) and 'reproduction' (5662 unigenes, 12.50%) may play important roles in sex determination and differentiation of *I. polycarpa.* These relative percentages were similar to those observed for transcriptome analysis of flowers in *Jatropha curcas* [25] and *Quercus suber* [17].

Table 2
Assembly quality and length distribution of assembled contigs and unigenes.

Length	Contigs	Unigenes
Total number	405,212	84,213
Total length(bp)	100,097,601	99,348,277
Max length(bp)	13,746	13,746
Min length (bp)	101	300
Mean length (bp)	247	1179
N50 length (bp)	332	1767
0–200	29,771 (21.7%)	–
200–300	45,916 (33.5%)	–
300–500	29,116 (21.2%)	26,376 (31.3%)
500–1000	16,995 (12.4%)	21,718 (25.8%)
1000–2000	10,208 (7.4%)	21,723 (25.8%)
>2000	5206 (3.8%)	14,396 (17.1%)

Table 3
Summary for annotation results of *I. Polycarpa* unigenes.

	NR	NT	Swiss-Port	KEGG	COG	GO	All annotated unigenes	All assembled unigenes
Number of unigenes	55,539	64,763	33,864	31,387	20,214	45,298	65,972	84,213
Percentage (%)	65.95	76.90	40.21	37.27	24.01	53.79	78.34	100

3.2.2. COG annotation and classification

Subsequently, for deeper understanding functional classification, 20,214 unigenes mapped to COG database were divided into 25 functional categories, covering most of the life processes. The top category was 'General function prediction only' (n = 6462), followed by 'Replication, recombination and repair' (n = 3402) and 'Transcription' (n = 3116); while there was only one unigene in the 'Extracellular structures' and five unigenes in the 'Nuclear structure' (Fig. 4).

3.2.3. KEGG pathway annotation

To further illustrate the biochemical pathways and transduction processes used during *I. polycarpa*, 31,387 unigenes (47.58%) were mapped to 128 KEGG pathways (Table S2). Among these, 'metabolic pathways [ko01100]' (6392; 20.37%), were the most represented groups, followed by 'biosynthesis of secondary metabolites pathways [ko01110]' (2996; 9.55%), 'plant-pathogen interaction [ko04626]' (1831; 5.83%), and 'plant hormone signal transduction [ko04075]' (1532; 4.88%). Additionally, some unigenes also were mapped to several pathways related with reproduction and development such as plant hormone signal transduction, circadian rhythm and flavonoid biosynthesis. These pathways probably played significant roles in investigation of specific biochemical and development processes.

3.3. Identification of I. polycarpa SSRs based on the de novo assembled transcriptome

SSRs are widely distributed throughout eukaryote genomes and transcriptomes, with high variability, co-dominant inheritance and detection convenience, and very useful as molecular markers for genetics and biology researches [26]. In 84,213 Unigene of *I. polycarpa*, we detected 29,953 SSRs, of which Mono-nucleotides comprised the largest fraction (9782, 32.66%) followed by Di-nucleotide (6921, 23.11%) and Tri-nucleotide (5420, 18.10%). Moreover, a total of 1862, 3100 and 2868 SSRs were detected by Quad-nucleotide, Penta-nucleotide and Hexa-nucleotide, respectively. AG/CT (4679) represented the largest fraction of Di-nucleotide, whereas AAG/CTT (1220) was the most abundant Tri-nucleotide (Fig. 5), which was also coincident with results in other angiosperms [27,28]. SSRs can be used not only for genetic diversity analysis, characters identification and

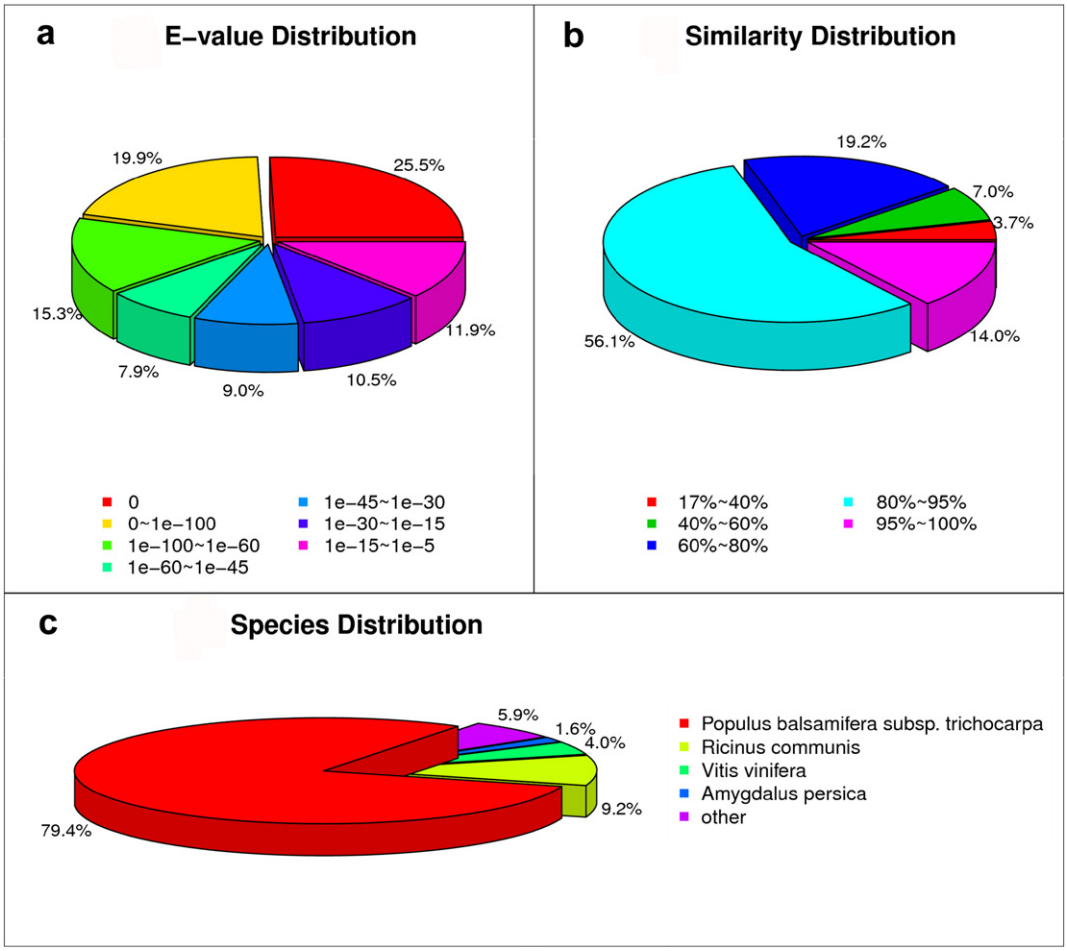

Fig. 2. Statistics of NR annotation e-value, similarity and species distribution of the annotated All-unigenes.

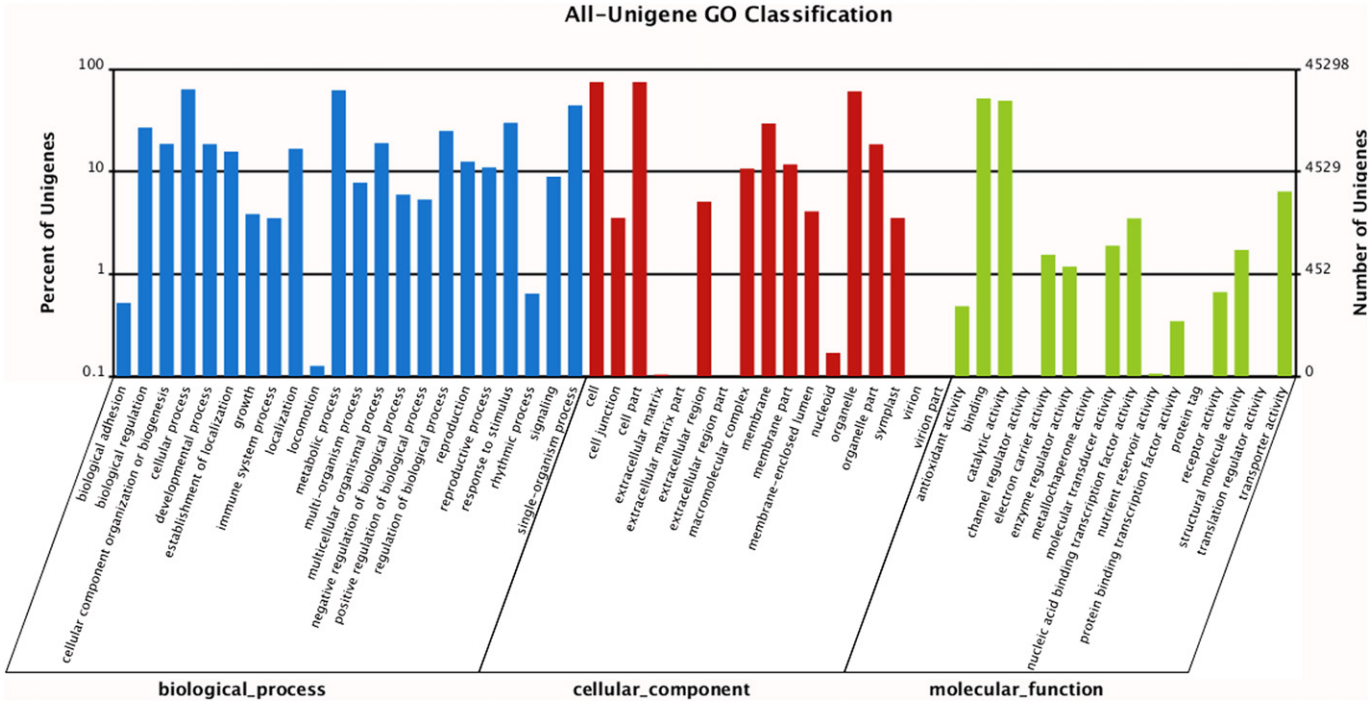

Fig. 3. GO classification of the annotated All-unigenes.

germplasm resources identification, but also for gender identification, which may provide insight into genetic diversity and a set of sex-linked SSR markers in *I. polycarpa*. Moreover, SSR markers based on transcriptome have been widely developed and also linked to sex determination in many plants, such as *Phoenix dactylifera* [29], kiwifruit [30], *Tapiscia sinensis* [31] and *Myrica rubra* [32].

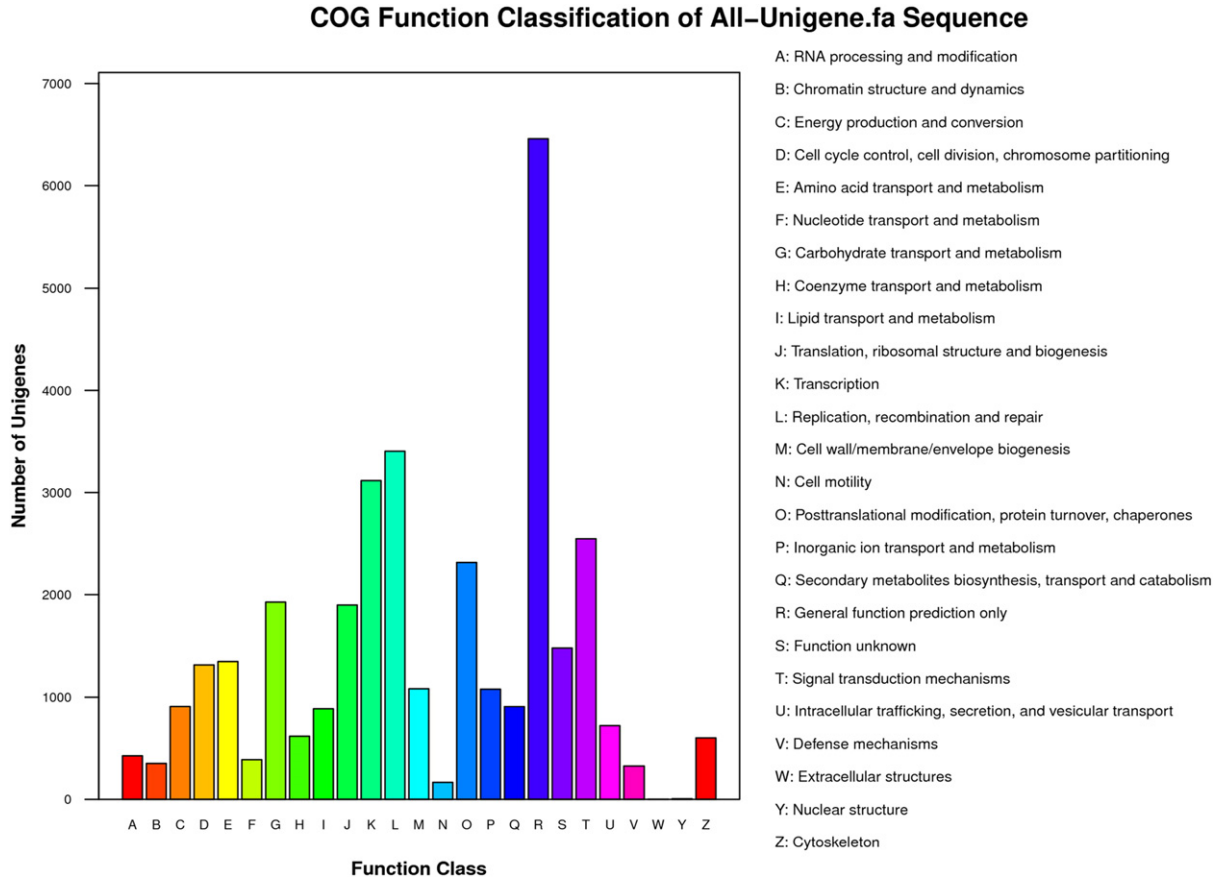

Fig. 4. COG categories of the annotated All-unigenes.

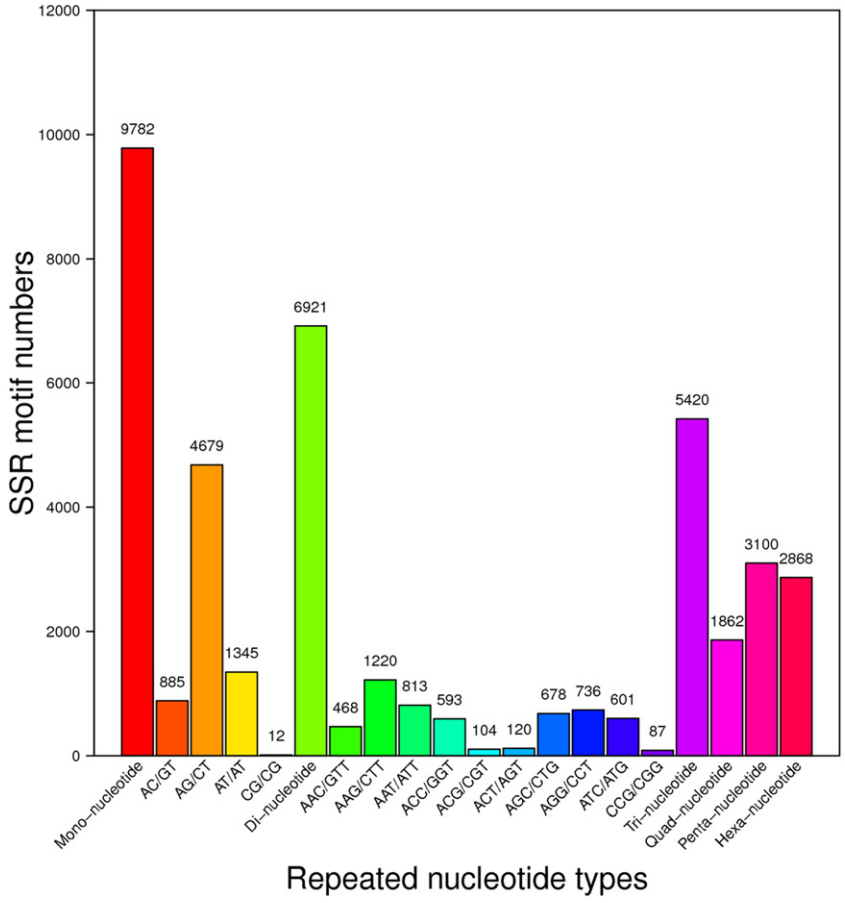

Fig. 5. SSRs identified in the transcriptome of *I. polycarpa*. X-axis: motif types. Y-axis: numbers of SSRs matched to different motif types.

Table 4
A list of putative unigenes related to flower development in *I. Polycarpa.*

Function	Gene name	Unigene umber	KO Name/Gene ID
Photoreceptor gene	*PHYA*(*phytochrome A*)	6	K12120
	PHYB(*phytochrome B*)	9	K12121
	PHYE(*phytochrome E*)	1	K12123
	CRY1(*cryptochrome 1*)	2	K12118
	CRY2(*cryptochrome 2*)	3	K12119
Circadian rhythm import genes	*FKF1*(*flavin-binding kelch repeat F-box protein 1*)	5	K12116
	PIF3 (*phytochrome-interacting factor 3*)	56	K12126
Circadian rhythm clock genes	*TOC1*(*APRR1,pseudo-response regulator 1*)	21	K12127
	CCA1(*circadian clock associated 1*)	22	K12134
	LHY(*late elongated hypocotyl*)	24	K12133
Circadian rhythm export genes	*CO* (*zinc finger protein CONSTANS*)	50	K12135
	GI (*GIGANTEA*)	2	K12124
Flowering integrators	*FLOWERING LOCUS T (FT)*	5	K16223
	SUPPRESSOR OF OVEREXPRESSION OF CO1 (SOC1)	7	no KO assigned
Flowering repressors	*TERMINAL FLOWER1 (TFL1)*	2	no KO assigned
	SHORT VEGETATIVE PHASE (SVP)	2	no KO assigned
Floral meristem identity gene	*LEAFY (LFY)*	2	no KO assigned
Homologous gene of *FT* and *TFL1*	*MOTHER OF FT AND TFL1 (MFT)*	2	no KO assigned
A-class genes	*APETALA1* (*AP1*)	3	K09264
	APETALA2 (*AP2*)	5	K09284
B-class genes	*GLOBOSA*(*GLO*)	2	K09264
	APETALA 3 (*AP3*)	3	K09264
	PMADS 1	2	K09264
	PMADS 2	2	K09264
C-class genes	*Agamous* (*AG*)	2	no KO assigned
E-class genes	*SEPALLATA1* (*SEP1*)	2	K09264
	SEPALLATA3 (*SEP3*)	2	no KO assigned
D-class genes	*Agamous-like MADS-box protein AGL11*	1	K09264

3.4. Unigenes putatively related to flower development of I. polycarpa

Flower development is a complex process regulated by the elaborate coordination of many genes. Functional annotation revealed that a number of unigenes with similarity to genes from other systems were involved in flower development. For example, many unigenes were similar to photoreceptor and circadian rhythm genes, flowering integration and repressor genes and floral meristem and organ identity genes (Table 4). Notably, two putative *MOTHER OF FT AND TFL1 (MFT)*, homologous to *FT* and *TFL1*, which were reported to promote floral induction in *A. thaliana* [33], were also found in *I. polycarpa* with higher expression in the male buds compared to the female bud, which together with *SOC1* and *LFY* were identified to play roles in the late stages of *Metasequoia glyptostroboides* buds [34].

Moreover, some unigenes involved in floral organ determination in other systems were identified in *I. polycarpa*, which were also reported to be crucial for flower organ formation in *Taihangia rupestris* [35]. According to the 'ABCDE' model, B-class is related to the production of petals, and stamens are regulated by the activity of the B-class and C-class genes, while the same C-class genes alone are responsible for carpel development [36]. Among them, some identified unigenes, homologs for A-class genes, *AP2*, and B-class genes showed lower expression levels in the *I. polycarpa* female library, and may account for the lack of petals in *I. polycarpa* flower buds. In addition, two homologs of the C-class gene *AGAMOUS* (*AG*) and four homologs of E-class genes showed similar levels of accumulation in both female and male libraries. Additionally, it's reported that the D-class gene *SEEDSTICK* (*STK*), a homolog of *Agamous-like gene 11* (*AGL11*), plays an important role in carpel and ovule development [37] and one homolog for *AGL11* was more highly expressed in female buds, indicating it may be responsive to signals involved in carpel and ovule development. These results provide reliable information for further studies on the expression of unigenes related to flower development in *I. polycarpa*.

4. Conclusion

In this study, the transcriptomes of *Idesia polycarpa* Maxim. var. *vestita* Diels female and male flower buds were sequenced using Illumina sequencing technology for the first time. *De novo* assembly generated 84,213 unigenes, 78.34% of which were mapped to major public databases, enriching the functional genomic resources available for *I. polycarpa*. Those unmatched unigenes may be species specific genes in *I. polycarpa*. *Idesia polycarpa* unigenes uncovered diverse biological functions and processes, such as reproduction and developmental process, and some unigenes putatively related to flower development were detected, and may play important roles in sex determination and differentiation. In addition, we also identified 29,953 SSRs based on the *de novo* assembled transcriptome of *I. polycarpa.* Taken together, these data provide insights into the mechanisms of female and male flower development in *I. polycarpa* and important candidate genes for further functional studies to analyze differences in sexual reproduction between female and male flower buds of *I. polycarpa*, all of which will prove valuable for further research into the reproductive biology and functional genomics of *I. polycarpa*. Further analysis of these SSRs will provide useful resources for conservation genetics and gender identification research on *I. polycarpa* in the future.

Financial support

This work was funded by Grants from Science and Technology support projects of Sichuan Province (2016NZYZF0102).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary data

http://dx.doi.org/10.1016/j.ejbt.2017.07.002

References

[1] Srikanth A, Schmid M. Regulation of flowering time: All roads lead to Rome. Cell Mol Life Sci 2011;68(12):2013–37. http://dx.doi.org/10.1007/s00018-011-0673-y.
[2] Baumann K, Venail J, Berbel A, et al. Changing the spatial pattern of *TFL1* expression reveals its key role in the shoot meristem in controlling *Arabidopsis* flowering architecture. J Exp Bot 2015;66(15):4769–80. http://dx.doi.org/10.1093/jxb/erv247.
[3] Corbesier L, Vincent C, Jang S, et al. FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. Science 2007;316(5827):1030–3. http://dx.doi.org/10.1126/science.1141752.
[4] Busch MA, Bomblies K, Weigel D. Activation of a floral homeotic gene in *Arabidopsis*. Science 1999;285(5427):585–7. http://dx.doi.org/10.1126/science.285.5427.585.
[5] Ming R, Bendahmane A, Renner SS. Sex chromosomes in land plants. Annu Rev Plant Biol 2011;62:485–514. http://dx.doi.org/10.1146/annurev-arplant-042110-103914.
[6] Mitchell CH, Diggle PK. The evolution of unisexual flowers: Morphological and functional convergence results from diverse developmental transitions. Am J Bot 2005; 92(7):1068–76. http://dx.doi.org/10.3732/ajb.92.7.1068.
[7] Alvarez-Buylla ER, Liljegren SJ, Pelaz S, et al. MADS-box gene evolution beyond flowers: Expression in pollen, endosperm, guard cells, roots and trichomes. Plant J 2000;24(4):457–66. http://dx.doi.org/10.1111/j.1365-313X.2000.00891.x.
[8] Dubos C, Stracke R, Grotewold E, et al. MYB transcription factors in *Arabidopsis*. Trends Plant Sci 2010;15(10):573–81. http://dx.doi.org/10.1016/j.tplants.2010.06.005.
[9] Golenberg EM, West NW. Hormonal interactions and gene regulation can link monoecy and environmental plasticity to the evolution of dioecy in plants. Am J Bot 2013;100(6):1022–37. http://dx.doi.org/10.3732/ajb.1200544.
[10] Wu T, Qin Z, Zhou X, et al. Transcriptome profile analysis of floral sex determination in cucumber. J Plant Physiol 2010;167(11):905–13. http://dx.doi.org/10.1016/j.jplph.2010.02.004.
[11] Chen M-S, Pan B-Z, Wang G-J, et al. Analysis of the transcriptional responses in inflorescence buds of *Jatropha curcas* exposed to cytokinin treatment. BMC Plant Biol 2014;14:318. http://dx.doi.org/10.1186/s12870-014-0318-z.
[12] Song Y, Ma K, Ci D, et al. Sexual dimorphic floral development in dioecious plants revealed by transcriptome, phytohormone, and DNA methylation analysis in *Populus tomentosa*. Plant Mol Biol 2013;83:559–76. http://dx.doi.org/10.1007/s11103-013-0108-2.
[13] Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63. http://dx.doi.org/10.1038/nrg2484.
[14] Bräutigam A, Gowik U. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. Plant Biol 2010;12:831–41. http://dx.doi.org/10.1111/j.1438-8677.2010.00373.x.
[15] Guo S, Zheng Y, Joung J-G, et al. Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. BMC Genomics 2010;11:384. http://dx.doi.org/10.1186/1471-2164-11-384.
[16] Wang S, Li Y, Li Z, et al. Identification of an SCAR marker related to female phenotype in *Idesia polycarpa* Maxim. Genet Mol Res 2015;14. http://dx.doi.org/10.4238/2015.March.20.11.
[17] Rocheta M, Sobral R, Magalhães J, et al. Comparative transcriptomic analysis of male and female flowers of monoecious *Quercus suber*. Front Plant Sci 2014;5:599. http://dx.doi.org/10.3389/fpls.2014.00599.
[18] Zhao Y, Liang H, Li L, et al. Digital gene expression analysis of male and female bud transition in *Metasequoia* reveals high activity of MADS-box transcription factors and hormone-mediated sugar pathways. Front Plant Sci 2015;6:467. http://dx.doi.org/10.3389/fpls.2015.00467.
[19] Chou C-J, Lin L-C, Tsai W-J, et al. Phenyl β-d-glucopyranoside derivatives from the fruits of *Idesia polycarpa*. J Nat Prod 1997;60(4):375–7. http://dx.doi.org/10.1021/np960335n.
[20] Ye Y, Tang X-S, Chen F, et al. Optimization of phenolics extracted from *Idesia polycarpa* defatted fruit residue and its antioxidant and depigmenting activity *In Vitro* and *In Vivo*. Evid Based Complement Alternat Med 2014;931269. http://dx.doi.org/10.1155/2014/931269.
[21] Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011;29:644–52. http://dx.doi.org/10.1038/nbt.1883.
[22] Conesa A, Götz S, García-Gómez JM, et al. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005; 21(18):3674–6. http://dx.doi.org/10.1093/bioinformatics/bti610.
[23] Ye J, Fang L, Zheng H, et al. WEGO: A web tool for plotting GO annotations. Nucleic Acids Res 2006;34:W293–7. http://dx.doi.org/10.1093/nar/gkl031.
[24] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008;36:D480–4. http://dx.doi.org/10.1093/nar/gkm882.
[25] Xu G, Huang J, Yang Y, et al. Transcriptome analysis of flower sex differentiation in *Jatropha curcas* L. using RNA sequencing. PLoS One 2016;11(2):e0145613. http://dx.doi.org/10.1371/journal.pone.0145613.

[26] Faircloth BC. Msatcommander: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 2008;8(1):92–4. http://dx.doi.org/10.1111/j.1471-8286.2007.01884.x.
[27] Wu T, Luo S, Wang R, et al. The first Illumina-based de novo transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. Mol Breed 2014;34(3):1437–47. http://dx.doi.org/10.1007/s11032-014-0128-x.
[28] Li D, Deng Z, Qin B, et al. *De novo* assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). BMC Genomics 2012;13:192. http://dx.doi.org/10.1186/1471-2164-13-192.
[29] Cherif E, Zehdi S, Castillo K, et al. Male-specific DNA markers provide genetic evidence of an XY chromosome system, a recombination arrest and allow the tracing of paternal lineages in date palm. New Phytol 2013;197(2):409–15. http://dx.doi.org/10.1111/nph.12069.
[30] Zhang Q, Liu CY, Liu YF, et al. High-density interspecific genetic maps of kiwifruit and the identification of sex-specific markers. DNA Res 2015;22(5):367–75. http://dx.doi.org/10.1093/dnares/dsv019.
[31] Zhou XJ, Wang YY, Xu YN, et al. *De novo* characterization of flower bud transcriptomes and the development of EST-SSR markers for the endangered tree *Tapiscia sinensis*. Int J Mol Sci 2015;16(6):12855–70. http://dx.doi.org/10.3390/ijms160612855.
[32] Jia HM, Jiao Y, Wang GY, et al. Genetic diversity of male and female Chinese bayberry (*Myrica rubra*) populations and identification of sex-associated markers. BMC Genomics 2015;16:394. http://dx.doi.org/10.1186/S12864-015-1602-5.
[33] Yoo SY, Kardailsky I, Lee JS, et al. Acceleration of flowering by overexpression of MFT. Mol Cells 2004;17(1):95–101.
[34] Zhao Y, Thammannagowda S, Staton M, et al. An EST dataset for *Metasequoia glyptostroboides* buds: The first EST resource for molecular genomics studies in *Metasequoia*. Planta 2013;237(3):755–70. http://dx.doi.org/10.1007/s00425-012-1783-y.
[35] Li W, Zhang L, Ding Z, et al. *De novo* sequencing and comparative transcriptome analysis of the male and hermaphroditic flowers provide insights into the regulation of flower formation in andromonoecious *Taihangia rupestris*. BMC Plant Biol 2017; 17:54. http://dx.doi.org/10.1186/s12870-017-0990-x.
[36] Wellmer F, Graciet E, Riechmann JL. Specification of floral organs in *Arabidopsis*. J Exp Bot 2014;65(1):1–9. http://dx.doi.org/10.1093/jxb/ert385.
[37] Pinyopich A, Ditta GS, Savidge B, et al. Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 2003;424:85–8. http://dx.doi.org/10.1038/nature01741.