Mining Simple Sequence Repeat and Single Nucleotide Polymorphism Markers in a Transcriptomic Database of Wintersweet (*Chimonanthus praecox*)

Daofeng Liu, Jing Ma, and Jianfeng Yang

Chongqing Engineering Research Center for Floriculture, Key Laboratory of Horticulture Science for Southern Mountainous Regions of Ministry of Education, College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400715, China

Tien V. Nguyen

Fruit and Vegetable Research Institute, Vietnam Academy of Agricultural Science, Hanoi, Vietnam

Huamin Liu, Renwei Huang, Shunzhao Sui¹, and Mingyang Li¹

Chongqing Engineering Research Center for Floriculture, Key Laboratory of Horticulture Science for Southern Mountainous Regions of Ministry of Education, College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400715, China

Additional index words. *Chimonanthus praecox*, molecular marker, RNA-Seq, simple sequence repeat, single nucleotide polymorphism

Abstract. Wintersweet is a woody ornamental plant and has a long history of human cultivation. Few molecular markers have been characterized and remain scant in wintersweet. This study aimed to mine simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) from the transcriptomic database of wintersweet. A total of 3972 SSRs and 97,060 putative SNPs/indels (92,307 SNPs and 4753 indels) were identified in this data set. This study marks the highest number of SSR and SNP markers discovered to date from wintersweet by using transcriptome sequencing data. These identified markers will provide a useful source for molecular genetic studies such as genetic diversity and characterization, association mapping, and map-based gene cloning in wintersweet.

The small, evolutionarily ancient Calycanthaceae family comprises four genera, namely *Calycanthus* L. in North America, *Idiospermum* S. T. Blake in Australia, and *Sinocalycanthus* W. C. Cheng & S. Y. Chang, and *Chimonanthus* L. in China (Li and Li, 2000). Wintersweet is one of the most economically important species in the Calycanthaceae family. Wintersweet (*Chimonanthus praecox*) is a hardy, fast-growing perennial shrub native to China; it is dichogamous and a diploid (2n = 22) (Zhang and Liu, 1998). As its name indicates, it blooms particularly in winter, from late November to March in south-central and

southwest China. Its unique flowering time and strong fragrance make it one of the most popular ornamental plants in China; it is appreciated as a common garden plant, a pot plant, or as cut flowers, and it has a high ornamental and economic value.

Wintersweet already has over 1000 years of cultivation history as well as many varieties and cultivars that were developed mainly using vegetative propagation and seedling selection. Some cultivars have been named and identified based on morphological traits such as the color or morphology of petals, and correspondingly several morphological classification criteria have been suggested (Chen et al., 2004; Zhao et al., 2004). Although morphological traits can be used to investigate genetic relationships and parentage information, these traits are often limited in numbers and influenced by the environment (Zhao and Zhang, 2008). Diversity analysis and measurement of genetic similarity or differences among varieties or cultivars provide important information about wintersweet conservation and breeding development. Moreover, a major focus in wintersweet breeding

programs has been to develop wintersweet hybrid varieties with improved ornamental traits such as flower size, color, or fragrance. Breeding new hybrids by conventional practices mostly is slow and uncertain. The application of molecular markers can improve efficiency of plant selection, saving time and providing accuracy in a breeding program (Snowdon and Friedt, 2004). The molecular markers have been used for a variety of applications including examination of genetic relationships between individuals, mapping of useful genes, construction of linkage maps, marker-assisted selections and backcrosses, population genetics, and phylogenetic studies (Kalia et al., 2011) such as marker-assisted selection in rice (Jena and Mackill, 2008) and identify quantitative trait loci (QTL) associated with seed protein content in soybean (Jun et al., 2008).

Among the many types of molecular markers, SSRs and SNPs are the preferred marker types for many genetic applications. SSRs are efficient codominant anchor markers with high levels of polymorphism and can easily be amplified by polymerase chain reaction (PCR) using primers designed from flanking sequences of the SSR motifs. SSR markers have been useful for integrating the genetic, physical, and sequence-based physical maps in plant species, and they are an efficient tool to link phenotypic and genotypic variation (Varshney et al., 2005). Meanwhile, SNPs are single-base differences between DNA sequences of individuals or lines. SNPs are ideal molecular markers for conducting genetic studies because they are codominant markers, abundant in both coding and non-coding regions, and are increasingly being used in phylogenetic and population genetic studies (Garvin et al., 2010). However, the traditional process of developing SSR markers is time-consuming and expensive because of the preparation of genomic libraries and sequencing of a large number of clones containing SSR regions in most species (Squirrel et al., 2003), and the traditional method for SNP discovery (cloning and comparative sequencing), which has been used for decades, is also expensive and time-consuming. In recent years, high-throughput sequencing technology has greatly decreased the inherent cost and time required for obtaining genomic and transcriptomic data and has become a more powerful resource than traditional methods in the discovery of SSR markers and SNP identification (Barbazuk et al., 2007; Metzker, 2010; Parchman et al., 2010). Moreover, the high-throughput capacity of the latest generation of RNA sequencing (RNA-Seq) technology provides a unique opportunity for the development of SSR and SNP markers in coding genes with a high degree of accuracy, both in model and non-model plant species, even in those for which little or no genome sequence information is available (Osman et al., 2003). Thousands of SSRs and SNPs have been detected within transcriptomic data for different species, especially in non-model species such as *Cucurbita pepo* (Blanca et al., 2011), sabaigrass (Zou et al.,

Received for publication 13 May 2014. Accepted for publication 22 Aug. 2014.

This work was supported by National Natural Science Foundation of China (Grant no. 31370698), the Fundamental Research Funds for the Central Universities (XDJK2013A004), the postdoctoral research foundation of Chongqing (Xm201345), and the “111” Project (B12006).

¹To whom reprint requests should be addressed; e-mail sszcq@swu.edu.cn; limy@swu.edu.cn.

2013), and *Lycoris aurea* (Wang et al., 2013).

Although entire genomes of model plant species such as *Arabidopsis* (Weigel and Mott, 2009), rice (Goff et al., 2002), and *Medicago* (Bell et al., 2001) have been sequenced in the past decade, the genome of wintersweet is still unavailable. As a result, research on wintersweet breeding and its application has been progressing slowly. We consider that only a very limited number of molecular markers has been developed in wintersweet (Chen and Chen, 2010; Dai et al., 2012; Yang et al., 2013; Zhao et al., 2007). For example, 31 expressed sequence tags (ESTs)-SSR markers have been developed from wintersweet EST sequences, and only eight polymorphic EST-SSR markers have been used to analyze the genetic diversity and structure of 10 natural populations (Yang et al., 2013). Although the whole genome of wintersweet has not been sequenced yet, a transcriptomic database of wintersweet has been established using the Illumina RNA-Seq method and a total of 10,699 transcripts assembled (Liu et al., 2014). This study aimed to mine SSRs and SNPs within the transcriptomic database of wintersweet and identify a substantial set of genetic markers for use in genetics, evaluation, or QTL identification. This work will facilitate and accelerate wintersweet breeding by marker-assisted selection.

Materials and Methods

Sequence data sources. We collected wintersweet [*Chimonanthus praecox* (L.) Link] flowers from three distinct wintersweet plants in the nursery at Southwest University, Chongqing, China. Total RNA was extracted for use in cDNA library construction; then the libraries were sequenced respectively with an Illumina HiSeq™ 2000 sequencing system (2 × 100-bp read length) at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China) as described by Liu et al. (2014). We assembled all the reads using Trinity software (Grabherr et al., 2011), and from the transcriptome database, we obtained 106,995 transcripts with total residues of 127,396,344 bp (Liu et al., 2014). We used nucleotide sequences of raw reads from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA, <<http://www.ncbi.nlm.nih.gov/Traces/sra>>) with accession number SRA106143.

SSR marker discovery. We subjected all transcripts to preprocessing to remove poly A/T tails using the program Trimest (<<http://mobyle.pasteur.fr/cgi-bin/portal.py?form=trimest>>) and removed any contaminations using an NCBI VecScreen system. The SSRs were identified by a msatcommander program (Faircloth, 2008). We adjusted the parameters to enable identification of perfect dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide motifs with a minimum of eight, five, four, four, and four repeats, respectively. Next, we designed the primer pairs to flank the detected SSR sites using a Primer3 tool (Rychlik, 1995). The major parameters for primer design were set as

follows: PCR products ranging from 150 to 450 bp; primer lengths ranging from 11 to 27 bp with an optimum of 19 bp; 60 °C optimal annealing temperature; and guanine-cytosine content (or GC content) from 35% to 75% with an optimum of 50%.

Identification of SNPs. We used the assembled transcripts as a reference sequence (Liu et al., 2014). To detect SNPs and indels, we mapped the raw reads against the transcripts using VarScan (<<http://varscan.sourceforge.net>>) and screened the SNPs using SAMtools (<<http://samtools.sourceforge.net/>>). We aligned each read to only a single best homologous site in the reference sequence. Reads aligning equally well in more than one location in the reference were discarded. Finally, we calculated the transition to transversion ratio and each type of substitution.

Results and Discussion

SSR. SSRs are one of the most popular marker systems and provide consistency by varying numbers of tandemly repeated di-, tri-, or tetra-nucleotide DNA motifs. Using the wintersweet transcriptomic data, we explored the msatcommander 0.8.1 program (Faircloth, 2008) to mine potential SSRs in wintersweet and analyzed their nature and frequency. A search for di-, tri-, and tetra-nucleotide repeats yielded 3972 potential SSRs. Of these, the most frequent repeat motifs were dinucleotides, which accounted for 64.65% of all SSRs, followed by trinucleotide (19.79%), tetranu-

cleotide (10.35%), pentanucleotide (3.25%), and hexanucleotide (3.25%) repeats (Table 1). Based on the distribution of SSR motifs, (AG/CT)_n was the most predominant type among the dinucleotide repeat motifs with a frequency of 86.33%. Among the 10 types of trinucleotide repeats, (AAG/CTT)_n was the most common motif with a frequency of 40.08% followed by (ATC/GAT)_n at 19.21% and (AGC/GCT)_n at 9.29%. Among the 28 types of tetranucleotide repeats, (AAAT/ATTT)_n and (AAAG/CTTT)_n were the most predominant motifs with frequencies of 31.87% and 17.76%, respectively (Table 1). Earlier studies showed that AG/CT and AAG/CTT were the predominant di- and trinucleotide SSR motifs, respectively, in dicotyledonous plants such as *Arabidopsis* (Lawson and Zhang, 2006) and *Cucurbita pepo* (Blanca et al., 2011). Thus, AG and AAG motifs may be common features of SSRs in dicotyledonous plants.

Although SSRs (48.77%) were found to be located in undetermined regions, more SSRs (21.85%) were located in the 5' untranslated regions (UTRs), and the number of SSRs located in open reading frames (ORFs) was similar to the number located in the 3'-UTRs (Table 2). Analysis of the localization of di-, tri-, and tetra-repeats showed that trinucleotides localized preferentially in ORFs, consistent with maintenance of the ORF coding capacity, whereas di- and tetra-nucleotides were more frequent in UTRs (Table 2). It has been reported that ≈3% to 7% of expressed genes contain putative SSR motifs, mainly

Table 1. Summary of simple sequence repeat types in the transcriptome of *Chimonanthus praecox*.

Repeat motif	Number	Percentage (%)	Total percentage (%)
Dinucleotide	2,568	100	64.65
AC/GT	228	8.88	
AG/CT	2,217	86.33	
AT/TA	121	4.71	
CG/GC	2	0.08	
Trinucleotide	786	100	19.79
AAC/GTT	28	3.56	
AAG/CTT	315	40.08	
AAT/ATT	53	6.74	
ACC/GGT	71	9.03	
ACG/CGT	10	1.27	
ACT/AGT	8	1.02	
AGC/GCT	73	9.29	
AGG/CCT	65	8.27	
ATC/GAT	151	19.21	
CCG/CGG	12	1.53	
Tetranucleotide	411	100	10.35
AAAT/ATTT	131	31.87	
AAAG/CTTT	73	17.76	
ATCC/GGAT	29	7.06	
AGGG/CCCT	27	6.57	
ACAT/ATGT	24	5.84	
AGAT/ATCT	20	4.87	
Other tetranucleotide repeats (less than 4% of each)	107	26.03	
Pentanucleotide	129	100	3.25
AGAGG/CCTCT	15	11.63	
AAAAT/ATTTT	14	10.85	
AAAAG/CTTTT	13	10.08	
AAGAG/CTCTT	13	10.08	
Other pentanucleotide repeats (less than 4% of each)	74	57.36	
Hexanucleotide	78		1.96
Total	3,972		

Table 2. Localization of simple sequence repeats (SSRs) with respect to putative initiation and termination codons in the transcriptomic database of wintersweet.^a

	Dinucleotide		Trinucleotide		Tetranucleotide		Pentanucleotide		Hexanucleotide		All SSRs	
	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
5'-UTR	586	22.82	143	18.19	85	20.68	34	26.36	20	25.64	868	21.85
ORF	302	11.76	269	34.22	35	8.52	9	6.98	35	44.87	650	16.36
3'-UTR	359	13.98	65	8.27	66	16.06	22	17.05	5	6.41	517	13.02
Undetermined	1,321	51.44	309	39.31	225	54.74	64	49.61	18	23.08	1,937	48.77
Total	2,568	100	786	100	411	100	129	100	78	100	3,972	100

^aUnigenes were checked for the presence of the start and stop codons. "Undetermined" means imprecise localization of the SSRs with respect to putative initiation or termination codons.

UTR = untranslated region; ORF = open reading frame.

within the untranslated regions of the mRNA (Thiel et al., 2003). SSRs within gene sequences may have different putative functions. For example, SSR variations in 5'-UTRs may regulate gene expression by affecting gene transcription or regulation; SSRs found in 3'-UTRs are involved in gene silencing and transcription slippage; intronic SSRs can affect gene transcription, mRNA splicing, or export to cytoplasm; and SSRs within genes are likely to be subjected to stronger selective pressure than other genomic regions (Li et al., 2004). Moreover, many of these SSRs occur in the protein coding sequences of annotated isogenes, representing genes of known or predicted identity and function.

To date, only a few microsatellites have been available for wintersweet. Thus, the development of SSRs for this species is highly desirable. Therefore, the SSR markers obtained in this work have some intrinsic advantages over genomic SSRs in the non-transcribed regions such as relative ease of generation, inexpensiveness, stability, and cross-species transferability (Varshney et al., 2005). These SSR markers will contribute to the construction of genetic linkage maps, genetic identification, and molecular marker-assistant breeding in *Chimonanthus* species. SSRs derived from unigenes are tightly linked with functional genes that may control useful characteristics, which is a good resource for further research. To facilitate the designing of SSR markers, we used the Primer3 software to design primer pairs for each SSR under a series of primer-designing parameters (see "Materials and Methods"). Thus, we designed 3687 primer pairs in 3972 SSRs (Supplemental Table 1). The primers designed from our transcriptome sequences with SSR locus will be very useful resources for SSR marker development in wintersweet; however, these primers need further experimental validation.

SNP/indel identification. Because the transcriptome sequences were obtained from a mixture of distinct wintersweet plants (Liu et al., 2014), our investigation was also very useful for detecting SNP polymorphism. In this study, we identified more than 97,060 putative SNPs/indels (92,307 SNPs and 4753 indels) from the transcriptomic database of wintersweet. Of the SNPs, we identified 54,484 transitions and 37,823 transversions from the isogenes (Fig. 1). Moreover, transitions (59%) were more frequent than transversions (41%). Regarding transition-type SNPs, the frequency of the A/G type (27,361; 29.64%) was found to be similar to that of the

Fig. 1. Summary of putative single nucleotide polymorphisms (SNPs) and insertion/deletions (indels) in the transcriptome of wintersweet.

C/T type (27,123; 29.38%). Meanwhile, for transversion-type SNPs, the A/T type (10,163; 11.01%) was the most common followed by the C/G type (9,270; 10.04%) (Fig. 1). In addition, ~60% of the isogenes contained only one or two SNPs/indels and those with no more than 10 SNPs/indels constituted 95% of the total isogenes. We observed 1453 isogenes containing more than 10 SNPs/indels. Figure 2A shows the detailed SNPs/indel distribution among these isogenes. To investigate the indel types among the isogenes, we calculated the insertion and deletion frequency within the isogenes (Fig. 2B). For the indels, deletions (73%) were more frequent than insertions (27%). A and T constituted the most frequent insertions (62.8%) and deletions (62.5%) among the indels (Fig. 2B).

Currently, no SNPs/indels have been identified in wintersweet. In this study, large numbers of SNPs/indels were identified from the transcriptome of wintersweet, providing a wealth of markers that will be potentially useful to applications ranging from population genetics and linkage mapping to comparative genomics. SNP discovery using transcriptomic data is advantageous in identifying SNPs that are directly associated with useful characteristics such as morphological traits or growth advantages. However, the disadvantage with SNP discovery using the transcriptome is that genes are more conserved than non-coding

DNA, which will lead to the discovery of fewer SNPs (Hyten et al., 2010). The more conserved sequence will also lead to primers or probes hybridizing to both the gene sequence that contains the SNP as well as to any conserved paralogous sequence, thereby decreasing the success rate of assays for such an SNP. In addition, without a genomic reference sequence, the proportion of successful SNP assays designed to cDNA sequence will also be reduced owing to introns interfering with oligo hybridization.

Conclusions

This study has revealed the highest number of SSR and SNP markers to date from wintersweet based on transcriptome sequencing data. A total of 3972 potential SSRs and 97,060 putative SNP/indels were identified from the transcriptomic database. Next-generation transcriptome sequencing leads to superior resources for the development of such markers not only because of the vast amount of sequence information from which markers can be identified, but also because newly discovered markers will be gene-based with associated studies aimed at understanding the genetic control of adaptive traits. The molecular markers identified in this study will accelerate our understanding of genetic variation in populations and the genetic control of important traits in

Fig. 2. Statistics of single nucleotide polymorphisms (SNPs) and insertion/deletions (indels) in the transcriptome of wintersweet. (A) Numbers of SNPs and indels detected among the isogenes; the x-axis represents the numbers of SNPs and indels detected per isogenes; the y-axis represents the numbers of isogenes. (B) Frequencies of different insertion/deletion types of indels; the x-axis represents the type of insertion or deletion; the y-axis represents the numbers of different insertion or deletion types of indels.

wintersweet. Additionally, our investigation will provide a material basis for future genetic linkage and QTL analyses and will provide useful information for functional genomic research in the future.

Literature Cited

Barbazuk, W.B., S.J. Emrich, H.D. Chen, L. Li, and P.S. Schnable. 2007. SNP discovery via 454 transcriptome sequencing. *Plant J.* 51:910–918.

Bell, C.J., R.A. Dixon, A.D. Farmer, R. Flores, J. Inman, R.A. Gonzales, M.J. Harrison, N.L. Paiva, A.D. Scott, J.W. Weller, and G.D. May. 2001. The *Medicago* Genome Initiative: A model legume database. *Nucleic Acids Res.* 29:114–117.

Blanca, J., J. Cañizares, C. Roig, P. Ziarso, F. Nuez, and B. Picó. 2011. Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics* 12:104.

Chen, D.W. and L.Q. Chen. 2010. The first intraspecific genetic linkage maps of wintersweet [*Chimonanthus praecox* (L.) Link] based on AFLP and ISSR markers. *Sci. Hort.* 124:88–94.

Chen, L.Q., K.G. Zhao, and M.Q. Zhou. 2004. Cultivar classification system of *Chimonanthus*. *J. Beijing Forestry Univ.* 26(suppl.):88–90.

Dai, P.F., J. Yang, T.H. Zhou, Z.H. Huang, L. Feng, H.L. Su, Z.L. Liu, and G.F. Zhao. 2012. Genetic diversity and differentiation in *Chimonanthus praecox* and *Ch. salicifolius* (Calycanthaceae) as revealed by inter-simple sequence repeat (ISSR) markers. *Biochem. Syst. Ecol.* 44:149–156.

Faircloth, B.C. 2008. msatcommander: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8:92–94.

Garvin, M.R., K. Saitoh, and A.J. Gharrett. 2010. Application of single nucleotide polymorphisms to non-model species: A technical review. *Mol. Ecol. Resour.* 10:915–934.

Goff, S.A., D. Ricke, T.H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, and H. nVarma, et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296:92–100.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29(7):644–652.

Hyten, D.L., Q. Song, E.W. Fickus, C.V. Quigley, J.S. Lim, I.Y. Choi, E.Y. Hwang, M. Pastor-Corrales, and P.B. Cregan. 2010. High-throughput SNP discovery and assay development in common bean. *BMC Genomics* 11:475.

Jena, K.K. and D.J. Mackill. 2008. Molecular markers and their use in marker-assisted selection in rice. *Crop Sci.* 48:1266–1276.

Jun, T.H., K. Van, M.Y. Kim, S.H. Lee, and D.R. Walker. 2008. Association analysis using SSR markers to find QTL for seed protein content in soybean. *Euphytica* 162:179–191.

Kalia, R.K., M.K. Rai, S. Kalia, R. Singh, and A.K. Dhawan. 2011. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* 177:309–334.

Lawson, M.J. and L. Zhang. 2006. Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biol.* 7:R14.

Li, Y. and P.T. Li. 2000. Origin, evolution and distribution of the Calycanthaceae. *Guizhou Sci.* 20:295–300.

Li, Y.C., A.B. Korol, T. Fahima, and E. Nevo. 2004. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* 6:991–1007.

Liu, D., S. Sui, J. Ma, Z. Li, Y. Guo, D. Luo, J. Yang, and M. Li. 2014. Transcriptomic analysis of flower development in wintersweet (*Chimonanthus praecox*). *PLoS One* 9:E86976.

Metzker, M.L. 2010. Sequencing technologies—The next generation. *Nat. Rev. Genet.* 11:31–46.

Osman, A., B. Jordan, P.A. Lessard, N. Muhammad, M.R. Haron, N.M. Riffin, A.J. Sinskey, C. Rha, and D.E. Housman. 2003. Genetic diversity of *Eurycoma longifolia* inferred from single nucleotide polymorphisms. *Plant Physiol.* 131:1294–1301.

Parchman, T.L., K.S. Geist, J.A. Grahnen, C.W. Benkaman, and C.A. Buerkle. 2010. Transcriptome sequencing in an ecologically important tree species: Assembly, annotation, and marker discovery. *BMC Genomics* 11:180.

Rychlik, W. 1995. Selection of primers for polymerase chain reaction. *Mol. Biotechnol.* 3:129–134.

Snowdon, R.J. and W. Friedt. 2004. Molecular markers in *Brassica* oilseed breeding: Current status and future possibilities. *Plant Breed.* 123:1–8.

Squirrel, J., P.M. Hollingsworth, M. Woodhead, J. Russell, A.J. Lowe, M. Gibby, and W. Powell. 2003. How much effort is required to isolate nuclear microsatellites from plants? *Mol. Ecol.* 12:1339–1348.

Thiel, T., W. Michalek, R.K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106:411–422.

Varshney, R.K., A. Graner, and M.E. Sorrells. 2005. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* 23:48–55.

- Wang, R., S. Xu, Y. Jiang, J. Jiang, X. Li, L. Liang, J. He, F. Peng, and B. Xia. 2013. *De novo* sequence assembly and characterization of *Lycoris aurea* transcriptome using GS FLX titanium platform of 454 pyrosequencing. PLoS One 8:e60449.
- Weigel, D. and R. Mott. 2009. The 1001 genomes project for *Arabidopsis thaliana*. Genome Biol. 10:107.
- Yang, J., P. Dai, T. Zhou, Z. Huang, L. Feng, H. Su, Z. Liu, and G. Zhao. 2013. Genetic diversity and structure of wintersweet (*Chimonanthus praecox*) revealed by EST-SSR markers. Sci. Hort. 150:1–10.
- Zhang, R.H. and H.E. Liu. 1998. Wax Shrubs in World (Calycanthaceae). China Science and Technology Press, Beijing, China. p. 24–25.
- Zhao, B. and Q.X. Zhang. 2008. Analysis of flower character variation of *Chimonanthus praecox* in China. Acta. Horticult. Sin. 35:383–388.
- Zhao, K.G., J.F. Yujiang, and L.Q. Chen. 2004. Numerical classification and principal component analysis of wintersweet cultivars. J. Beijing Forestry Univ. 26(suppl):79–83.
- Zhao, K.G., M.Q. Zhou, L.Q. Chen, D.L. Zhang, and W.R. Gituru. 2007. Genetic diversity and discrimination of *Chimonanthus praecox* (L.) Link germplasm using ISSR and RAPD markers. HortScience 42:1144–1148.
- Zou, D., X. Chen, and D. Zou. 2013. Sequencing, de novo assembly, annotation and SSR and SNP detection of sabaigrass (*Eulaliopsis binata*) transcriptome. Genomics 102:57–62.