


Article

Development of SSR Molecular Markers and Genetic Diversity Analysis of TPS Gene Family in *Chimonanthus praecox*

Xuemei Fu ^{1,†} , Nan Yang ^{1,†}, Yongqin Du ², Hafiz Muhammad Kamran ¹, Huabo Wang ¹, Shaoyuan Chen ¹ and Longqing Chen ^{1,*}

¹ Yunnan Province Engineering Research Center for Functional Flower Resources and Industrialization, College of Landscape Architecture and Horticulture Sciences, Southwest Forestry University, Kunming 650224, China; fuxm@swfu.edu.cn (X.F.); yangnan@swfu.edu.cn (N.Y.)

² Agro-Technical Popularization Service Centre of Jiading District, Shanghai 201800, China

* Correspondence: chenlq@swfu.edu.cn; Tel.: +86-0871-6386-2842

† These authors contributed equally to this work.

Abstract: Terpene synthase (TPS) plays a key role in the biosynthesis of terpenoids, which are the most important components of the volatile compounds of wintersweet (*Chimonanthus praecox*). In this study, 52 *CpTPS* genes were found in wintersweet which were divided into 5 subfamilies. We identified 146 SSRs in the *CpTPS* genes, and obtained 33 pairs of SSR primers with good polymorphism through amplification in 6 wintersweet samples. Then, these primers were amplified in 69 samples from China's main wintersweet production areas. Through structural analysis, 69 samples were divided into 2 clusters, and were divided into 4 groups in a genetic cluster analysis, of which SH-33 and SW were separate groups. Through AMOVA analysis, it was found that the variation mainly occurred in the population, and that the gene flow between populations was $N_m > 1$, so it might lead to population differentiation. In other words, these findings provided useful information for the biosynthesis of terpenoids, the construction of a genetic linkage map, the detection of quantitative trait loci, marker-assisted selection and other aspects of wintersweet.

Keywords: terpene synthase; terpenoids; SSR; *Chimonanthus praecox*



Citation: Fu, X.; Yang, N.; Du, Y.; Kamran, H.M.; Wang, H.; Chen, S.; Chen, L. Development of SSR Molecular Markers and Genetic Diversity Analysis of TPS Gene Family in *Chimonanthus praecox*. *Agriculture* **2023**, *13*, 893. <https://doi.org/10.3390/agriculture13040893>

Academic Editor: Rodomiro Ortiz

Received: 14 March 2023

Revised: 12 April 2023

Accepted: 13 April 2023

Published: 18 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Plants are pioneers in producing different terpenoids [1]. So far, more than 80,000 terpenoids, including monoterpenes, sesquiterpenes, diterpenes, triterpenes and tetraterpenes have been found [2,3]. Terpenoids are utilized to mediate interactions in environmental defense and adaptability [4–7], play a role in plant allelopathy and abiotic stress responses [8–10], and participate in the synergistic interaction between plants and herbivores and other soil organisms [11–14]. Terpene synthase (TPS) genes catalyze a stereospecific carbon cation cascade reaction that transforms a few common glutaryl diphosphate substrates into a variety of terpenoids [15–19]. Terpene synthases (TPSs) are the primary enzyme that catalyze the conversion of GPP, FPP, or GGPP into monoterpene (C10), sesquiterpene (C15), or diterpene (C20) [20].

Wintersweet (*Chimonanthus praecox*) is a plant of *Chimonanthus* in the family Calycanthaceae, which is a traditional fragrant flower plant and an important garden plant in China and has been introduced to the United States, Japan, South Korea and other countries. Wintersweet has a unique fragrance and winter flowering characteristics, and it is a plant with aesthetic value and economic value. In addition to garden use, wintersweet also has many commercial uses, such as in essential oils, cut flowers, cosmetics, etc. At present, the research on the fragrance components of wintersweet has proven that a large number of volatile terpenes (mainly monoterpenes and sesquiterpenes) and a small amount of benzene ring compounds are the sources of the strong aroma of wintersweet [21,22]. Different wintersweet genotypes might contain different floral components. For example,

linalool was the most abundant volatile compound in H29, and the level of it is low in SW001, but the level of trans- β -ocimene of SW001 was significantly higher than that of H29, and there were 16 differentially expressed *CpTPS* genes in SW001 and H29 [23].

CpTPS genes play an important role in the synthesis of terpene volatiles in wintersweet. The biosynthesis of terpene volatiles has seen some progress. The functions of some *CpTPS* genes have been identified. For example, *CpTPS04* is the main gene that regulates the synthesis of linalool. At the same time, it was found that *CpTPS17*, *18* and *19* are three copies of *CpTPS04*, which might be the reason for the large degree of linalool synthesis in wintersweet [24]. *CpTPS1* could regulate the synthesis of ocimene through prokaryotic expression [25].

Nowadays, molecular markers are extensively used in the construction of genetic maps, the identification of plant germplasm resources, the division of heterosis groups, molecular marker-assisted breeding, the location of important trait genes and functional genes, or quantitative trait locus (QTL) analysis in species and much more. Simple sequence repeats (SSRs) have the benefits of codominance, good repeatability, easy operation, wide-coverage, and good polymorphism [26,27]. Other markers, such as RAPD, AFLP, ISSR, EST-SSR, and SRAP have previously been used to categorize wintersweet, discover new species, classify variations, and analyze the genetic structure of populations [28–32]. The genetic diversity and genetic relationship of 201 individuals from seven populations of *C. niten* complex including *C. niten*, *C. zhejiangensis* and *C. grammatus*, were investigated with ISSR and RAPD markers [28]. The development and primer screening of EST-SSR markers were carried out using the transcriptome data of wintersweet [30]. In addition, SSR sequences might lead to different functions and genes. In *Salvia guaranitica*, several SSR motifs were associated with many unique sequences that encoded enzymes involved in terpenoid biosynthesis [33]. The allelic rearrangement in the *MpMYB10* promoter produces an auto-regulated SSR site, which leads to increased *MpMYB10* transcription levels and the ectopic accumulation of anthocyanins in apple [34]. A total of 35 TabZIPMSs (*Triticum aestivum* bZIP MicroSatellites) with high polymorphism and interspecific metastasis were identified in wheat [35]. The SSRs in *CpMYB2* may be related to the flower color of wintersweet [36]. SSR markers of regulatory flavonoid biosynthesis genes were found to be highly polymorphic in strawberry and raspberry, indicating that their allelic variability may be related to differences in flavonoid accumulation and composition [37]. Based on the studies of the function of *CpTPS* genes, the SSR markers of the *CpTPS* gene family were analyzed, which might lead to the development of functional markers related to the terpene synthesis of wintersweet. These functional markers would be helpful in distinguishing the differences in terpene compounds between different wintersweet plants.

Genome sequencing of wintersweet has given us the opportunity to study the genomic SSR of the *TPS* gene family. In this study, we focused on the terpene synthetase gene family for the development of polymorphic SSR markers. These SSR markers were used to assess the population structure and genetic diversity of wintersweet genotypes and cultivars originating from the main wintersweet-producing areas of China. These findings provide valuable information and tools for genetic resources identification and conservation measures, as well as for molecular marker-assisted (MAS) breeding strategies.

2. Materials and Method

2.1. Plant Materials

The 69 wintersweet cultivars and genotypes were collected from Shanghai, Nanjing, Chengdu, Wuhan, and Kunming (Figure 1; Table 1), which were the main production areas of wintersweet in China. These DNA were extracted by the CTAB method [38]. The DNA was sent to Tianyi Huayu, China (Wuhan Tianyi Huayu Gene Technology Co., Ltd., Wuhan, China), in dry ice for SSR primer amplification.



Figure 1. Distribution of sampling sites. Purple font represents sampling sites, and number of samples are indicated in parentheses. Black font represents the name of the provincial capital.

Table 1. Summary of wintersweet cultivars and genotypes used in the study.

Sampling Site	Sample Name	Quantity
Shanghai	SH-1, SH-2, SH-3, SH-4, SH-5, SH-6, SH-7, SH-8, SH-9, SH-10, SH-11, SH-12, SH-13, SH-14, SH-15, SH-16, SH-17, SH-18, SH-19, SH-20, SH-21, SH-22, SH-23, SH-24, SH-25, SH-26, SH-27, SH-28, SH-29, SH-30, SH-31, SH-32, SH-33, SH-34, SH-35, SH-36, SH-37, SH-38, SH-39, SH-40, HHT, JZ, WG, YZH	44
Nanjing	ZSB, XRDY, QYJT, JLHZ	4
Chengdu	XGCL, HQ, ZH, CTYX, BC, XY, SN, XFHX, JSYG, HTXD	10
Kunming	HLT012, HLT015, HLT016, HLT040, HLT053, XL003	6
Wuhan	H29, H45, H64, H93, SW	5

2.2. Phylogeny Analysis of CpTPS Proteins

Based on the gene ID of CpTPS in Shang's paper [24], we identified the ORFs of 52 CpTPS genes from the wintersweet genome (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA600650> accessed on 16 April 2022) by the Gtf/Gff3 Sequences Extract tool of TBtools [39]. The protein sequences of the CpTPSs were obtained through the Batch Translate CDS to Protein tool of TBtools. We compared the CpTPS proteins in multiple series by using MEGA 7.0 [40]. To create a phylogenetic tree, we utilized the maximum likelihood approach and the Jones–Taylor–Thornton with gamma distribution (JTT + G) model, and the bootstrap value of the phylogenetic tree was 1000 repetitions. The phylogenetic tree was beautified through iTOL [41].

2.3. SSR Mining and Primer Design of CpTPS Genes

By the Gtf/Gff3 Sequences Extract tool of TBtools [39], the 52 CpTPS gene sequences were excavated from the wintersweet genome and the base pair values were set to 2000 bp before the start codon and 2000 bp after the stop codon. These sequences were used for SSRs mining and primer design. The SSRs in the CpTPS genes were mined by MISA [42]. The minimum standard length for dinucleotide repeats was 6 repeats, the minimum

standard length standard for trinucleotide repeats, tetranucleotide repeats, hexanucleotide repeats and pentanucleotide repeats was 5 repeats. Composite SSRs were defined as a repeating sequence with a non-repetitive sequence interval of less than 100 bp. Primers were created utilizing the flanking sequences of SSR loci by the online program Primer3 (<https://primer3.ut.ee/> (16 April 2022)).

2.4. Amplification Verification and Evaluation of CpTPS SSR

The SSR primers was synthesized by Tianyi Huayu, China (Wuhan Tianyi Huayu Gene Technology Co., Ltd., Wuhan, China). SSR primers were amplified and verified in wintersweet cultivars and genotypes, and PCR products were separated by capillary electrophoresis (CE). The number of observed and expected alleles, as well as the observed and expected heterozygosity for each marker, were calculated by the Popgen32 program [43]. We calculated the polymorphic information content (PIC) of each SSR marker by the PIC calculator (<https://www.liverpool.ac.uk/kempsj/pic.html> accessed on 16 April 2022) and used the Gene Location Visualize tool from GTF/GFF of TBtools [39] to locate SSR loci on the wintersweet chromosomes.

2.5. Clustering and Genetic Structure Analysis

The Bayesian algorithm-based STRUCTURE software version 2.3.4 [44] was used to estimate the genetic structure of 69 wintersweet cultivars and genotypes. The population numbers (K) were calculated by running 20 iterations of K from 1 to 10 with a burn-in period of 20,000 and a total of number of 20,000 Markov Chain Monte Carlo (MCMC) repeats after burn-in. We used structure harvester v0.6.94 [45] to determine K via posterior probability using a hierarchical island model utilizing the log-likelihood value (LnPd value) [46]. NTSYS-pc ver. 2.10 [47] was used for the clustering analysis, and the tree cluster diagram was obtained by the unweighted matching method (UPGMA). We performed principal component analysis (PCA) and an analysis of molecular variance (AMOVA) of the population structure by using GenAlEx [48].

3. Result

3.1. Phylogenetic Analysis of CpTPS Proteins

In our preliminary study, 52 complete CpTPSs were identified in the wintersweet assembly genome and the TPS genes were revealed to be significantly expanded by comparative genomic analysis [24]. The CpTPS gene family could be divided into five subfamilies based on phylogenetic analysis: TPS-a, TPS-b, TPS-c, TPS-g and TPS-e/f (Figure 1). The TPS-b subfamily had the largest number of genes (24), followed by the TPS-a subfamily (18), while the TPS-c subfamily had only one gene (Figure 2). The TPS-b subfamily participated in the synthesis of monoterpenes [49], so the expansion of the TPS-b subfamily might lead to a large accumulation of monoterpenes in the volatile organic compounds (VOCs) of wintersweet flower.

3.2. Identification and Characteristic Analysis of SSR Loci of CpTPS Family

The 146 SSRs were discovered from 44 CpTPS sequences (each sequence included a 2000 bp promoter gene and 2000 bp sequence of the 3' UTR area) (Supplementary Table S1), accounting for 84.62% of the overall sequences. In total, 35 of the 44 sequences had multiple SSRs. Also, 17 SSRs present in compound formation were detected (Table 2). The 2000 bp promoter regions contained 32 SSRs, the gene regions contained 82 SSRs, the post-2000 bp sequences contained 15 SSRs, and CpTPS23 contained the most SSR sites, of which only one SSR (19-CDs) in CpTPS19 was located in the CDs region (Supplementary Table S1). Among all the SSR types, dinucleotide repeats were the most abundant (67.12%), followed by trinucleotide repeats (32.19%), while there was only one SSR with a pentanucleotide repeat. SSRs with 5 (23.97%) and 6 (26.03%) sequence repetitions were the most abundant. The number of sequence repetitions was 41 (Table 3). The AT/AT type was the most common (67.35%) among the dinucleotide repeated sequences, followed by AG/CT (32.65%).

AAT/ATT was the most frequent trinucleotide repeat (80.85%), followed by AAG/CTT (12.77%) and ATC/ATG (6.38%) (Figure 3).

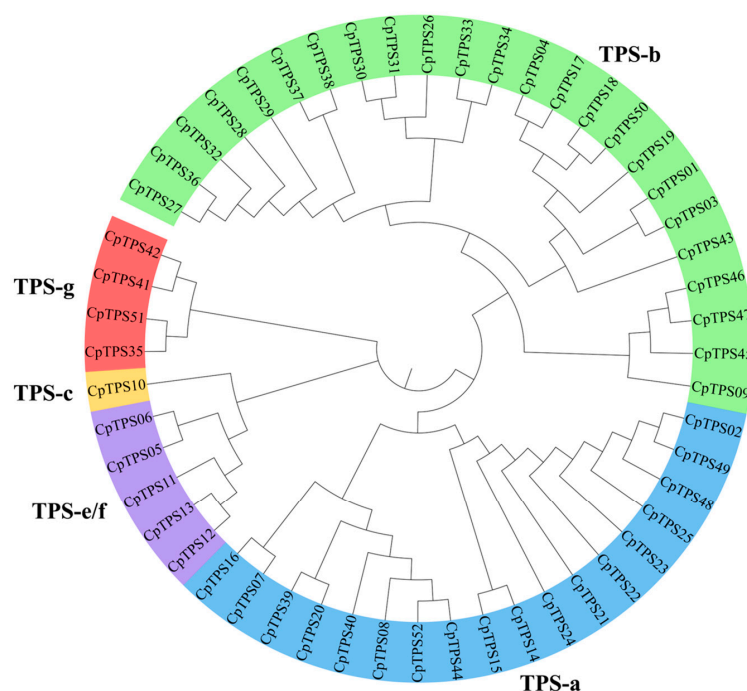


Figure 2. Phylogenetic relationship of *CpTPS* genes. MEGA 7.0 software was used to generate a phylogenetic tree using ML method. The blue, green, yellow, purple and orange branches represent the TPS-a, TPS-b, TPS-c, TPS-e/f and TPS-g subfamilies.

Table 2. Mining results of SSR in the *CpTPS* genes.

Mining Results of SSR in the <i>CpTPS</i> Genes		
Total number of checked sequences		52
The total size of checked sequences		977,634
Total number of identified SSRs		146
Number of sequences with SSRs		44
Number of sequences with more than one SSR		35
Number of SSRs present in compound formation		17
Number of designed SSR primers		104
Number of wintersweet cultivars and genotypes used for genotyping		69
Number of polymorphic primers found		39

Table 3. The relationship between the frequency of different SSR repeat motif types and the number of repeat motifs in the *CpTPSs*.

Number of Repetitions	Di	Tri	Penta	Total	Percentage (%)
5	0	35	0	35	23.97
6	31	6	1	38	26.03
7	24	0	0	24	16.44
8	5	2	0	7	4.79
9	5	4	0	9	6.16
10	4	0	0	4	2.74
11	2	0	0	2	1.37
12	6	0	0	6	4.11
14	1	0	0	1	0.68
15	3	0	0	3	2.05
>15	17	0	0	17	11.64
Total	98	47	1	146	
Percentage (%)	67.12	32.19	0.68		

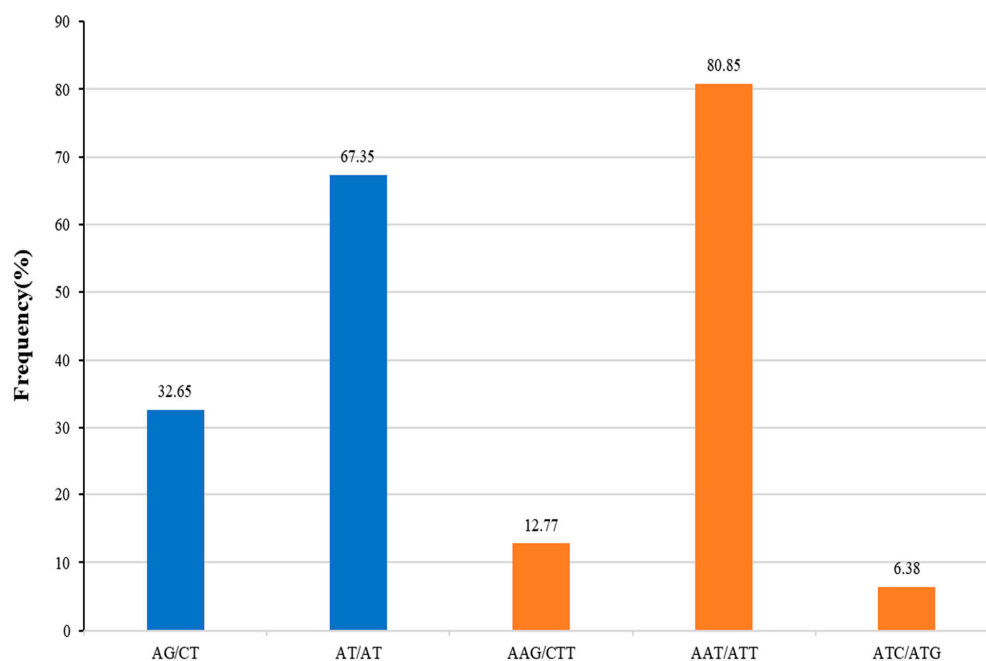


Figure 3. Ratio of different repeat motifs in SSR loci. Blue indicates a dinucleotide repeat, and orange indicates a trinucleotide repeat.

3.3. Chromosome Location of SSR Loci

In order to understand the distribution of CpTPSMSs (microsatellites of the TPS gene family in wintersweet) on wintersweet chromosomes, we located the 146 CpTPSMSs on the wintersweet chromosome. Based on the wintersweet genome, these CpTPSMSs were mapped onto the physical map of the wintersweet chromosomes (Figure 4), and it was discovered that CpTPSMSs were dispersed on 8 chromosomes and 5 scaffolds. The majority of CpTPSMSs were found on Chr5 (contained 31 SSRs) and Chr6 (contained 47 SSRs). The CpTPS genes of the TPS-b subfamily contained 66 SSRs, followed by the TPS-a subfamily, which contained 43 SSRs. According to the ratio of the number of SSRs to the number of genes in each subfamily, the average SSR number of the CpTPS-c subfamily genes was the highest, at an average of 4.00 SSRs. The average number of CpTPS-a, CpTPS-b and CpTPS-e/f subfamily genes was 3.44, 2.79 and 3.40, respectively. The lowest average number was of CpTPS-g subfamily genes, at an average of 0.50 SSRs. It was indicated that the distribution of CpTPSMSs among different CpTPS subfamilies was quite different.

3.4. Development of CpTPSMS Primers

We screened 103 SSRs from 146 CpTPSMSs and designed primers. These primers were amplified in six wintersweet samples (H29, HLT040, HLT053, SW, XL003 and YZH). The results of capillary electrophoresis showed that most SSR primers had a poor amplification effect (Figure 5). In total, 13 pairs of primers had no amplification products in three or more samples, 35 pairs of primers had no polymorphism or poor polymorphism, 22 pairs of the primer amplification products could not be determined, and 33 SSR primers had obvious amplification bands and good polymorphism. The 33 polymorphic markers generated 108 SSR variants, and each range was 2–6, with an average of 2.91 SSR variations per marker. Meanwhile, the effective allele range was 1.41–6, with an average of 2.36 per locus (Table 4 and Table S2). The expected heterozygosity was 0.32–0.91 (an average of 0.58), the observed heterozygosity was 0–0.67 (an average of 0.20), and the expected heterozygosity was higher than that which was observed. The range of the polymorphism information content (PIC) value was 0.27–0.81, with an average of 0.46. In total, 13 CpTPSMSs were highly polymorphic (PIC > 0.5), CpTPSMS03-2, CpTPSMS43-1 and CpTPSMS47-1 were from CpTPS genes of the TPS-b subfamily, CpTPSMS06-3, CpTPSMS06-4, CpTPSMS12-2

were from *CpTPS* genes of the TPS-e/f subfamily, and *CpTPSMS10-2* was from the *CpTPS* gene of the TPS-c subfamily (Figure 4).

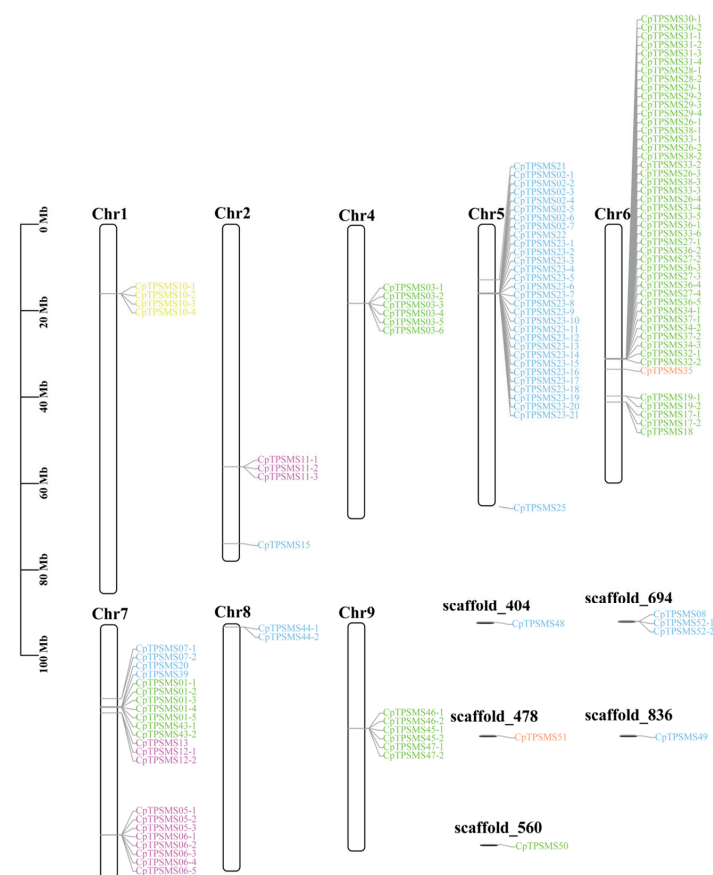


Figure 4. Chromosomal location of *CpTPSMSs*. Names of makers are shown on the right side of the chromosome. Blue, green, yellow, purple and orange indicate that *CpTPSMSs* were from the TPS-a, TPS-b, TPS-c, TPS-e/f and TPS-g subfamilies, respectively.

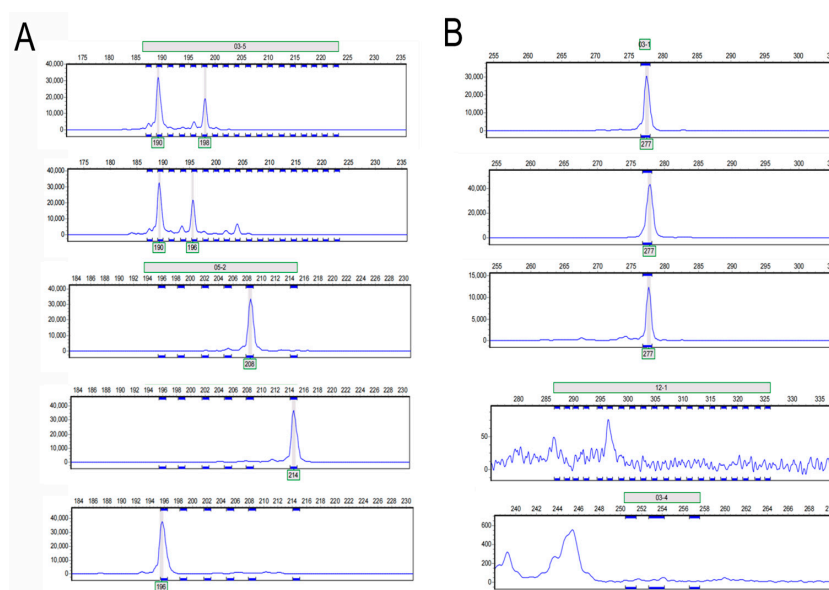


Figure 5. Examples of SSR primer amplification peaks and percentages of different types of polymorphic SSRs. (A) Examples of SSR amplification peaks with good polymorphism. (B) Examples of SSR peak graphs for SSR amplification without polymorphism and failed SSR amplification.

Table 4. Amplification characteristics of 33 CpTPSMSs.

Locus Name	na	ne	Obs_Hom	Exp_Hom	Obs_Het	Exp_Het	PIC
CpTPSMS01-4	2	1.60	0.83	0.59	0.17	0.41	0.30
CpTPSMS01-5	3	2.94	0.80	0.27	0.20	0.73	0.59
CpTPSMS02-7	3	2.63	0.80	0.31	0.20	0.69	0.55
CpTPSMS03-1	3	1.67	0.83	0.56	0.17	0.44	0.36
CpTPSMS03-2	6	6.00	0.67	0.09	0.33	0.91	0.81
CpTPSMS03-4	3	2.67	1.00	0.29	0.00	0.71	0.55
CpTPSMS03-5	2	1.47	1.00	0.64	0.00	0.36	0.27
CpTPSMS03-6	3	2.32	0.50	0.38	0.50	0.62	0.50
CpTPSMS05-1	2	1.80	0.33	0.52	0.67	0.48	0.35
CpTPSMS05-2	2	1.60	0.83	0.59	0.17	0.41	0.30
CpTPSMS05-3	3	2.32	0.67	0.38	0.33	0.62	0.50
CpTPSMS06-1	2	1.60	0.83	0.59	0.17	0.41	0.30
CpTPSMS06-3	2	1.80	1.00	0.52	0.00	0.48	0.35
CpTPSMS06-4	3	2.57	1.00	0.33	0.00	0.67	0.54
CpTPSMS06-5	3	3.00	0.33	0.27	0.67	0.73	0.59
CpTPSMS10-2	5	4.50	1.00	0.15	0.00	0.85	0.74
CpTPSMS10-3	3	3.00	1.00	0.20	0.00	0.80	0.59
CpTPSMS11-1	2	1.60	0.83	0.59	0.17	0.41	0.30
CpTPSMS11-2	3	2.27	1.00	0.38	0.00	0.62	0.50
CpTPSMS12-1	4	3.57	1.00	0.20	0.00	0.80	0.67
CpTPSMS12-2	4	3.57	1.00	0.20	0.00	0.80	0.67
CpTPSMS13	3	2.00	1.00	0.45	0.00	0.55	0.45
CpTPSMS19-CDs	2	2.00	0.80	0.44	0.20	0.56	0.38
CpTPSMS19	4	2.94	0.40	0.27	0.60	0.73	0.61
CpTPSMS23-6	3	1.67	0.50	0.56	0.50	0.44	0.36
CpTPSMS23-21	2	1.80	0.67	0.52	0.33	0.48	0.35
CpTPSMS26-4	3	1.41	0.83	0.68	0.17	0.32	0.27
CpTPSMS33-6	3	1.85	0.80	0.49	0.20	0.51	0.41
CpTPSMS38-2	2	1.80	0.67	0.52	0.33	0.48	0.35
CpTPSMS43-1	3	2.18	0.67	0.41	0.33	0.59	0.46
CpTPSMS47-1	3	1.67	0.83	0.56	0.17	0.44	0.36
CpTPSMS47-2	3	2.00	1.00	0.45	0.00	0.55	0.45
CpTPSMS52-1	2	1.95	0.83	0.47	0.17	0.53	0.37
Average	2.91	2.36	0.80	0.42	0.20	0.58	0.46

na: observed number of alleles; ne: effective number of alleles; Obs_Hom: observed homozygosity; Exp_Hom: expected homozygosity; Obs_Het: observed heterozygosity; Exp_Het: expected heterozygosity; PIC: polymorphism information content.

3.5. Structural and Genetic Clustering Analysis

We collected 69 wintersweet cultivars and genotypes from Wuhan, Kunming, Nanjing, Chengdu and Shanghai, which were the main places of origin of wintersweet in China. The DNA of these samples was amplified using 33 SSR primers with good polymorphism. The polymorphism of the amplified products was detected by capillary electrophoresis, which found that 26 pairs of SSR primers had a good amplification effect in 69 samples. The structure of the 69 samples was analyzed by the STRUCTURE software. The hierarchical island model revealed that the ΔK reached the highest peak at $K = 2$, and the 69 wintersweet cultivars and genotypes were grouped into two groups (Figure 6). Cluster I and Cluster II had Q values ranging from 0.56 to 0.99 and 0.50 to 0.99, respectively. There were 50 cultivars and genotypes (72.46%) with a Q score greater than 0.8, suggesting that their genotypes are pure lines, while suggesting that the other 19 cultivars and genotypes are mixed lines. The expected heterozygosity between individuals in Cluster I was 0.3844, and that between individuals in Cluster II was 0.5972. The paired F_{st} values for Cluster I and Cluster II were 0.1096, showing the substantial difference between the populations, but that the differences were not associated with geographical distances between the cultivars and genotypes. Cluster I contained 49 wintersweet cultivars and genotypes, all of which were from Nanjing

and Sichuan. Cluster II has 20 wintersweet cultivars and genotypes, isome of which were from Wuhan and four of which were from Kunming.

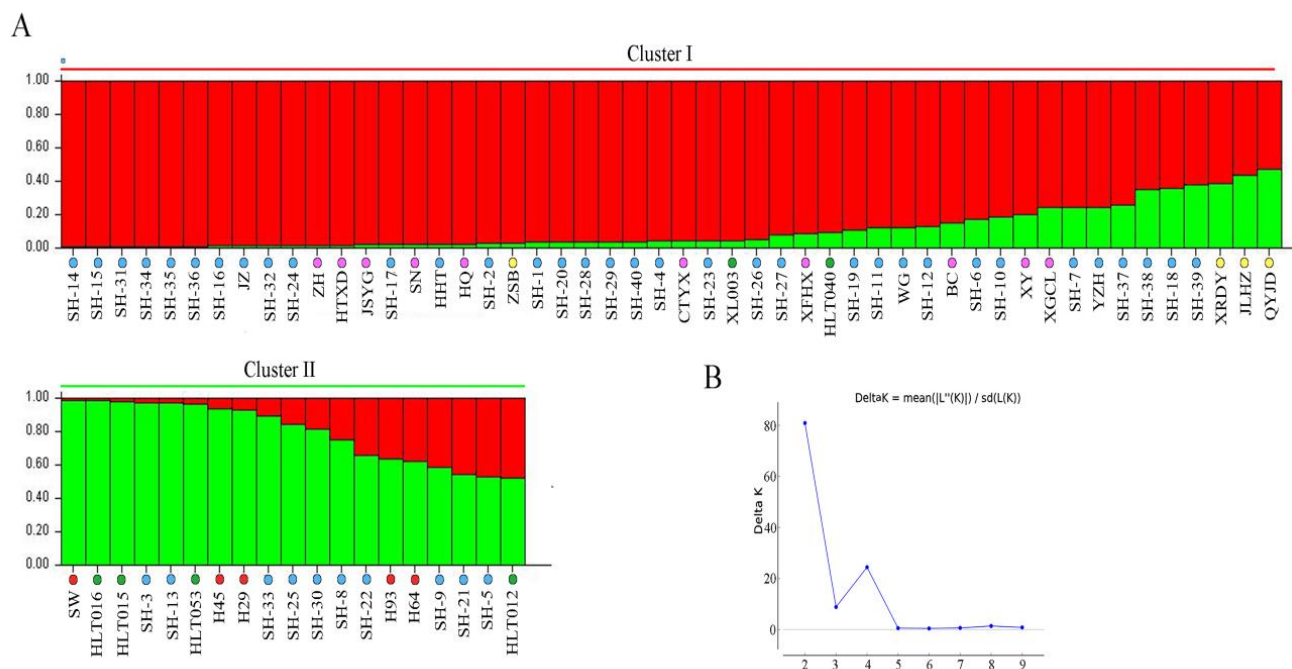
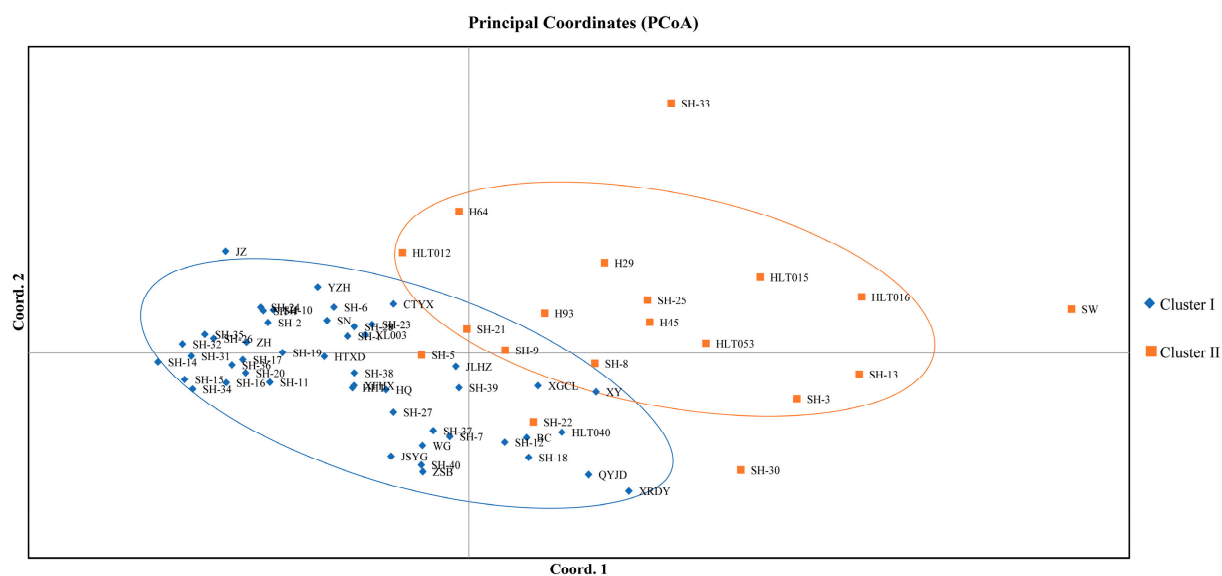
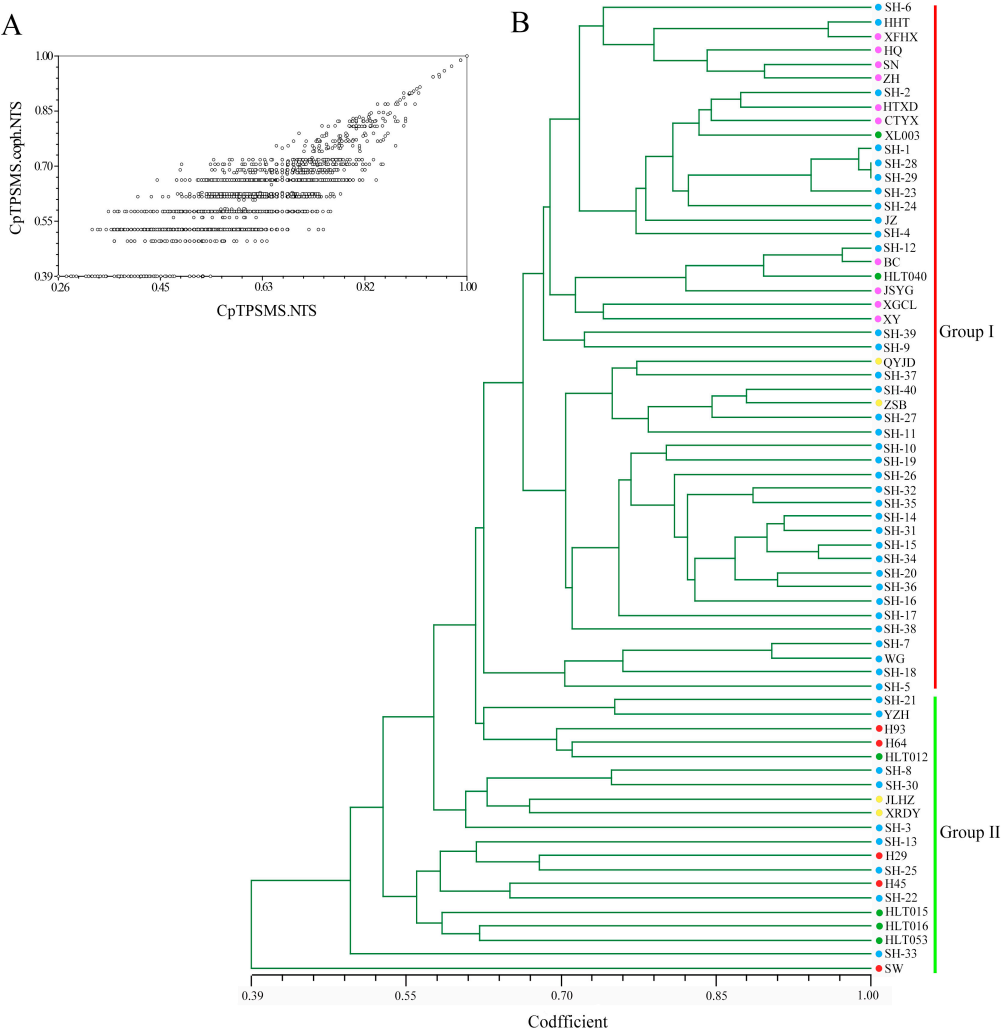


Figure 6. Structure analysis of 69 wintersweet cultivars and genotypes based on the Bayesian model. (A) Group structure was obtained by structure analysis. Red and green represent Cluster I and Cluster II, respectively. Blue, purple, yellow, green and red dots represent wintersweet samples from Shanghai, Sichuan, Nanjing, Kunming and Wuhan, respectively. (B) Hierarchical island model was obtained through the online software Structure Harvester.

At the same time, according to the amplification results of 26 SSR primers in the 69 samples, after calculating the SM genetic similarity coefficient between the samples, a cluster map was obtained using the unweighted pair-group method with arithmetical means (UPGMA). In order to test the reliability of the clustering results, the Mantel test was carried out through the molecular marker clustering results and the similarity coefficient matrix, and the correlation coefficient $r = 0.80$ (Figure 7A) was obtained, indicating that the molecular marker clustering results accurately clustered 69 samples with a close genetic relationship into a group. When the similarity coefficient was 0.55, 69 samples were divided into 4 groups (Figure 7B). Group I contained 59 samples. The samples from Sichuan and Nanjing and most of the samples from Shanghai belonged to Group I. Group II contained 8 samples, including 3 samples from Kunming, 3 samples from Shanghai and 2 samples from Wuhan. Groups III and IV 4 contained only one sample, SH-33 and SW, respectively.

3.6. Analysis Principal Component and Molecular Variance

According to the principal component analysis, among the top three principal components, the interpretation rate of principal component 1 (PC1) was 22.66%, while that of principal component 2 (PC2) was 12.75% and that of principal component 3 (PC3) was 11.10% (Supplementary Table S3). All of cultivars and genotypes of wintersweet were split into two clusters (Figure 8), confirming the grouping pattern of the group structure analysis (Figure 6). The distribution of Cluster II was relatively scattered, with SH-33 and SW being far from the main distribution area of the cluster. There was some overlap between the two clusters, indicating that gene flow occurred between the two clusters.



The molecular variance analysis was utilized to estimate the evolutionary distance between alleles. Based on the AMOVA the variance between wintersweet cultivars and genotypes was estimated to construct the genetic distance. The AMOVA results revealed that the variance index among groups of 69 wintersweet cultivars and genotype was 11%, the intra-group variation index was 62%, the inter-individual variation index was 27%. Besides that, the F_{st} among the three groups was 0.112 ($p \geq 0.001$) (Table 5). The gene flow (N_m) calculated from the F_{st} value was 1.976, which showed a high level of gene exchange between the two populations. When $N_m > 1$, gene flow was sufficient to overcome genetic drift within the population and caused population differentiation [50].

Table 5. Summary AMOVA Table.

Source	df	SS	MS	Est. Var.	Percentage Variation	F_{st}
Among Pops	1	51.459	51.459	0.787	11%	0.112 ($p \geq 0.001$)
Among Indiv	67	542.281	8.094	1.869	27%	
Within Indiv	69	300.500	4.355	4.355	62%	
Total	137	894.239		7.012	100%	

4. Discussion

Wintersweet, a Chinese traditional flower, is famous for its strong floral scent, which is one of its most prominent ornamental features. Terpene volatiles are the most significant component of wintersweet volatile compounds, and they play a key role in the scent of the wintersweet [21,22]. Plant terpenoids are not only necessary for floral fragrance but are also required to cope with biotic and abiotic stress [6–8,10,51,52]. Currently, the research on the TPS genes related to the wintersweet flower's scent has focused mostly on identification, and some important genes have been verified [23,24]. Shang found 52 *CpTPS* genes from the wintersweet genome and identified their subfamilies [24].

Through the screen of *CpTPSMSs* by MISA, we discovered that the dinucleotide repeated sequence was the most prevalent (67.12%), followed by the trinucleotide repetitive sequence (32.19%) (Table 3, Figure 3), which was consistent with the repetitive motif of SSR from the wintersweet transcriptome [30]. The AT/AT type was the most abundant dinucleotide repetitive sequence (67.35%), and the AAT/ATT type was the most abundant trinucleotide repetitive sequence (80.85%), which was not the same as the EST-SSR repeat motifs of the wintersweet transcriptome [30,53]. In this study, only one SSR locus (*CpTPSMS19-CDs*) existed in the exon, and the other SSRs came from introns and promoters. Most polymorphic *CpTPSMSs* were found on Chr7 and 6, which were linked to the presence of more *CpTPS* genes on the two chromosomes (Figure 5). The TPS-b and TPS-e/f subfamilies contained most of the *CpTPSMSs*. The TPS-b subfamily was the largest subfamily of *CpTPS* genes, which was closely related to the synthesis of a large number of terpene compounds in wintersweet, and the content of terpene compounds in different wintersweet cultivars varied greatly [21,22]. The TPS-e/f subfamily, as a more conservative subfamily, allowed a certain degree of conservation among different cultivars of wintersweet. The identification and analysis of the polymorphism SSR loci on these genes could help to distinguish the function of TPS genes and further develop functional SSR markers in the future.

We amplified 103 *CpTPSMSs* at random in six wintersweet genotypes and discovered 33 polymorphic SSR markers (32.04%) (Table 4). The effective amplification rate of primers in the study was much greater than that of SSR primers from *Calycanthaceae* in earlier investigations [30,53,54]. In these 33 *CpTPSMSs*, there was an average of 2.91 SSR variants per locus (Table 4), which was larger than the number of amplified loci in *Sinocalycanthus chinensis* [53,54]. If the PIC value was larger than 0.5, the marker was more informative [55]. The PIC value of the *CpTPSMS* was 0.27–0.81, with an average value of 0.46, showing only moderate polymorphism. However, the PIC of 13 SSR loci was more than 0.5, and these SSR loci were highly polymorphic (Table 4). We discovered that SSRs with a trinucleotide

repetitive sequence were more likely to have high polymorphism, based on the ratio of SSRs with high polymorphism in their corresponding types.

The genetic differences in the population could be inferred via population structure analysis. It is easier to understand the evolutionary relationship between groups and individuals if they are divided into subgroups. Because quantitative characteristics are impacted by complex factors such as genetic variations across populations and individuals, we employed a combination of NJ-tree-building and structural analysis to validate the genetic clustering and population structure of 69 wintersweet cultivars and genotypes (Figures 6 and 7). These wintersweet samples came from the main production areas and germplasm resource nurseries of wintersweet in China. In the structural analysis, 69 wintersweet samples were divided into two clusters. In the phylogenetic analysis, when the similarity coefficient was 0.55, the samples were divided into four groups, of which SH-33 and SW were separated into one group each (Figure 7), which was similar to the PCA analysis results. Meanwhile, we discovered that the two groups intersected each other in the PCA analysis (Figure 8). The result of PCA analysis made up for the defects of the genetic clustering using the genetic similarity coefficient and the population structuring using Bayesian algorithm. Our results indicate that even when genotypes came from the same area, they could be from different original production areas, because they were not specifically clustered into one group. For example, in the structural analysis, two samples from Kunming were in Cluster I, and four samples were in Cluster II (Figure 6); in the cluster analysis, three samples from Kunming belonged to Group I, and the other three samples belonged to Group II (Figure 7). This might have led to the deviation in the results of the genetic clustering and structural analysis. Although the clustering results matched the geographic origin found in an earlier research [56], some people believe that microsatellite analysis cannot be used to cluster species based on their geographic origin [55].

The 19 wintersweet samples belonged to mixed genotypes, such as JYJD, H93 and H64, indicating that they had mixed ancestors and lost their original gene pool. While the remaining 50 wintersweet samples were homozygous, such as H29, SW and HLT015, their *Q* values were greater than 0.8, and their genetic structure was maintained during evolution (Figure 6). The majority of the variance, according to the AMOVA analysis, originated from within the population (Table 5). Similar results have been obtained from a genetic diversity analysis based on ISSR and RAPD markers [31]. This might also have been due to the complexity of the sample source. On the other hand, wintersweet is a cross-pollinated plant, which means that there could be a lot of hybridization in the population.

5. Conclusions

In total, 52 terpene synthetase genes were found in the genome of wintersweet. The polymorphic SSR loci in *CpTPS* genes were screened in six wintersweet samples by capillary electrophoresis, and 33 SSR loci with good polymorphism were found. The 33 pairs of SSR primers were amplified in 69 wintersweet samples and the genetic diversity and structure of these samples were analyzed. Through these research results, 26 SSR loci with good polymorphism in the *CpTPS* gene family were discovered, which expanded the marker library of wintersweet. SSR markers with high polymorphism can be used for germplasm identification, population structure research, genetic linkage analysis, etc.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/agriculture13040893/s1>, Table S1: Distribution of SSR loci in *CpTPS* genes, Table S2: Basic information of 33 pairs of SSR primers, Table S3: Percentage of variation explained by the first 3 axes.

Author Contributions: X.F. performed most of the experiments, analyzed the data and wrote the manuscript. N.Y. designed the experiments and revised the manuscript. Y.D. contributed to the plant material. H.M.K. provided scientific advice. H.W. and S.C. performed some experiments. L.C. designed and oversaw the project. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by National Natural Science Foundation of China (grant No. 32002073); Natural Science Foundation of Yunnan Province (grant No. 202101AU070145); The Talents Introduction Plan of Yunnan Province—"High-End Foreign Experts" Program (grant No. YNQR-GDWG-2018-018).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the support provided by their respective institutions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Tholl, D. Biosynthesis and biological functions of terpenoids in plants. In *Biotechnology of Isoprenoids*; Springer International: Cham, Switzerland, 2015; pp. 63–106.
- Christianson, D.W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **2017**, *117*, 11570–11648. [[CrossRef](#)] [[PubMed](#)]
- Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-Ya, K.; Omura, S.; Cane, D.E.; Ikeda, H. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 857–862. [[CrossRef](#)] [[PubMed](#)]
- Gershenzon, J.; Dudareva, N. The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **2007**, *3*, 408–414. [[CrossRef](#)] [[PubMed](#)]
- Keeling, C.I.; Bohlmann, J. Diterpene resin acids in conifers. *Phytochemistry* **2006**, *67*, 2415–2423. [[CrossRef](#)] [[PubMed](#)]
- Schmelz, E.A.; Huffaker, A.; Sims, J.W.; Christensen, S.A.; Lu, X.; Okada, K.; Peters, R.J. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.* **2014**, *79*, 659–678. [[CrossRef](#)]
- Vaughan, M.M.; Wang, Q.; Webster, F.X.; Kiemle, D.; Hong, Y.J.; Tantillo, D.J.; Coates, R.M.; Wray, A.T.; Askew, W.; O'donnell, C. Formation of the unusual semivolatile diterpene rhizathalene by the Arabidopsis class I terpene synthase *TPS08* in the root stele is involved in defense against belowground herbivory. *Plant Cell* **2013**, *25*, 1108–1125. [[CrossRef](#)] [[PubMed](#)]
- Kato-Noguchi, H.; Peters, R.J. The role of momilactones in rice allelopathy. *J. Chem. Ecol.* **2013**, *39*, 175–185. [[CrossRef](#)]
- López, M.L.; Bonzani, N.E.; Zygadlo, J.A. Allelopathic potential of *Tagetes minuta* terpenes by a chemical, anatomical and phytotoxic approach. *Biochem. Sys. Ecol.* **2008**, *36*, 882–890. [[CrossRef](#)]
- Vaughan, M.M.; Christensen, S.; Schmelz, E.A.; Huffaker, A.; Mcauslane, H.J.; Alborn, H.T.; Romero, M.; Allen, L.H.; Teal, P.E. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell Environ.* **2015**, *38*, 2195–2207. [[CrossRef](#)]
- Agrawal, A.A.; Heil, M. Synthesizing specificity: Multiple approaches to understanding the attack and defense of plants. *Trends Plant Sci.* **2012**, *17*, 239–242. [[CrossRef](#)]
- Dudareva, N.; Pichersky, E. Biochemical and molecular genetic aspects of floral scents. *Plant Physiol.* **2000**, *122*, 627–634. [[CrossRef](#)] [[PubMed](#)]
- Heil, M.; Ton, J. Long-distance signalling in plant defence. *Trends Plant Sci.* **2008**, *13*, 264–272. [[CrossRef](#)] [[PubMed](#)]
- Pichersky, E.; Gershenzon, J. The formation and function of plant volatiles: Perfumes for pollinator attraction and defense. *Curr. Opin. Plant Biol.* **2002**, *5*, 237–243. [[CrossRef](#)]
- Banerjee, A.; Hamberger, B. P450s controlling metabolic bifurcations in plant terpene specialized metabolism. *Phytochem. Rev.* **2018**, *17*, 81–111. [[CrossRef](#)]
- Bathe, U.; Tissier, A. Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry* **2019**, *161*, 149–162. [[CrossRef](#)]
- Karunanithi, P.S.; Zerbe, P. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Front. Plant Sci.* **2019**, *10*, 1166. [[CrossRef](#)] [[PubMed](#)]
- Peters, R.J. Two rings in them all: The labdane-related diterpenoids. *Nat. Prod. Rep.* **2010**, *27*, 1521–1530. [[CrossRef](#)] [[PubMed](#)]
- Zerbe, P.; Bohlmann, J. Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.* **2015**, *33*, 419–428. [[CrossRef](#)]
- Trapp, S.C.; Croteau, R.B. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **2001**, *158*, 811–832. [[CrossRef](#)]
- Du, Y.Q.; Tian, X.L.; Gan, J.Z.; Pan, C.; Zhang, W.Q.; Xia, W. Essential oils extracted from different varieties of wintersweet (*Chimonanthus praecox*). *J. Beijing For. Univ.* **2013**, *35*, 81–85.
- Zhou, M.Q.; Xiang, L.; Chen, L.Q. Preliminary studies on the components of volatile floral flavor and flower pigments of *Chimonanthus praecox*. *J. Beijing For. Univ.* **2007**, *29*, 22–25.
- Tian, J.; Ma, Z.; Zhao, K.; Zhang, J.; Xiang, L.; Chen, L. Transcriptomic and proteomic approaches to explore the differences in monoterpene and benzenoid biosynthesis between scented and unscented genotypes of wintersweet. *Physiol. Plant.* **2019**, *166*, 478–493. [[CrossRef](#)] [[PubMed](#)]

24. Shang, J.; Tian, J.; Cheng, H.; Yan, Q.; Li, L.; Jamal, A.; Xu, Z.; Xiang, L.; Saski, C.A.; Jin, S.; et al. The chromosome-level wintersweet (*Chimonanthus praecox*) genome provides insights into floral scent biosynthesis and flowering in winter. *Genome Biol.* **2020**, *21*, 200. [[CrossRef](#)] [[PubMed](#)]
25. Tian, J. Analysis of Floral Volatile Biosynthetic Pathways and Functional Characterization of Mono-TPSs Genes from *Chimonanthus praecox*. Ph.D. Thesis, Huazhong Agricultural University, Wuhan, China, 2019.
26. Varshney, R.K.; Chabane, K.; Hendre, P.S.; Aggarwal, R.K.; Graner, A. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Sci.* **2007**, *173*, 638–649. [[CrossRef](#)]
27. Russell, J.; Fuller, J.; Macaulay, M.; Hatz, B.; Jahoor, A.; Powell, W.; Waugh, R. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor. Appl. Genet.* **1997**, *95*, 714–722. [[CrossRef](#)]
28. Li, X.; Zhou, M.Q.; Zhao, K.G.; Chen, L.Q. Genetic diversity and differentiation of *Chimonanthus nitens* complex populations. *J. Beijing For. Univ.* **2012**, *136*, 111–117.
29. Chen, D.W.; Chen, L.Q. The first intraspecific genetic linkage maps of wintersweet [*Chimonanthus praecox* (L.) Link] based on AFLP and ISSR markers. *Sci. Hortic.* **2010**, *124*, 88–94. [[CrossRef](#)]
30. Li, X.; Yang, N.; Zhao, K.G.; Chen, Y.X.; Tang, R.J.; Chen, L.Q. Development and primer selection of EST-SSR molecular markers based on transcriptome sequencing of *Chimonanthus praecox*. *J. Beijing For. Univ.* **2013**, *35*, 25–32.
31. Zhao, K.G.; Zhou, M.Q.; Chen, L.Q.; Zhang, D.; Robert, G.W. Genetic diversity and discrimination of *Chimonanthus praecox* (L.) link germplasm using ISSR and RAPD markers. *HortScience* **2007**, *42*, 1144–1148. [[CrossRef](#)]
32. Zhou, M.Q.; Chen, L.Q.; Ruan, R. Genetic diversity of *Chimonanthus nitens* Oliv. complex revealed using inter-simple sequence repeat markers. *Sci. Hortic.* **2012**, *136*, 38–42. [[CrossRef](#)]
33. Ali, M.; Hussain, R.M.; Rehman, N.U.; She, G.; Li, P.; Wan, X.; Guo, L.; Zhao, J. De novo transcriptome sequencing and metabolite profiling analyses reveal the complex metabolic genes involved in the terpenoid biosynthesis in Blue Anise Sage (*Salvia guaranitica* L.). *DNA Res.* **2018**, *25*, 597–617. [[CrossRef](#)] [[PubMed](#)]
34. Espley, R.V.; Brendolise, C.; Chagne, D.; Kutty-Amma, S.; Green, S.; Volz, R.; Putterill, J.; Schouten, H.J.; Gardiner, S.E.; Hellens, R.P. Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. *Plant Cell* **2009**, *21*, 168–183. [[CrossRef](#)] [[PubMed](#)]
35. Sharma, H.; Bhandawat, A.; Kumar, P.; Rahim, M.S.; Parveen, A.; Kumar, P.; Madhawan, A.; Rishi, V.; Roy, J. Development and characterization of bZIP transcription factor based SSRs in wheat. *Gene* **2020**, *756*, 144912. [[CrossRef](#)] [[PubMed](#)]
36. Yang, N. Functional Characterization of Anthocyanin Synthase *CpANS1* and Transcription Factor *CpMYB2* in *Chimonanthus praecox* Tepals. Ph.D. Thesis, Huazhong Agricultural University, Wuhan, China, 2019.
37. Lebedev, V.G.; Subbotina, N.M.; Maluchenko, O.P.; Lebedeva, T.N.; Krutovsky, K.V.; Shestibratov, K.A.G. Transferability and polymorphism of SSR markers located in flavonoid pathway genes in *Fragaria* and *Rubus* species. *Genes* **2019**, *11*, 11. [[CrossRef](#)] [[PubMed](#)]
38. Doyle, J.J.; Doyle, J.L. *Isolation of Plant DNA from Fresh Tissue*; Focus: San Francisco, CA, USA, 1990; Volume12, pp. 13–15.
39. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)] [[PubMed](#)]
40. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
41. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
42. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
43. Yeh, F.C. *POPGENE*, version 1.3.1. Microsoft Window-Bases Freeware for Population Genetic Analysis. University of Alberta: Edmonton, AB, Canada, 1999.
44. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **2009**, *9*, 1322–1332. [[CrossRef](#)]
45. Earl, D.A.; Vonholdt, B.M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
46. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)] [[PubMed](#)]
47. Rohlf, F. *NTSYSpc: Numerical Taxonomy System*, version 2.1; Exeter Publishing, Ltd.: Setauket, NY, USA, 2002.
48. Smouse RP, P.; Peakall, R. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539.
49. Chen, F.; Tholl, D.; Bohlmann, J.; Pichersky, E. The family of terpene synthases in plants: A mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* **2011**, *66*, 212–229. [[CrossRef](#)]
50. Slatkin, M. Gene flow and the geographic structure of natural populations. *Science* **1987**, *236*, 787–792. [[CrossRef](#)]
51. Li, J.J.; Hu, H.; Mao, J.; Yu, L.; Stoop, G.; Wang, M.; Mumm, R.; De Ruijter, N.C.; Dicke, M.; Jongsma, M.A. Defense of pyrethrum flowers: Repelling herbivores and recruiting carnivores by producing aphid alarm pheromone. *New Phytol.* **2019**, *223*, 1607–1620. [[CrossRef](#)]

52. Huang, M.; Sanchez-Moreiras, A.M.; Abel, C.; Sohrabi, R.; Lee, S.; Gershenzon, J.; Tholl, D. The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (E)- β -caryophyllene, is a defense against a bacterial pathogen. *New Phytol.* **2012**, *193*, 997–1008. [[CrossRef](#)]
53. Huang, Y.H.; Zhang, C.; Zhou, L.H.; Zhao, H.B. Development and primer screening of SSR markers based on transcriptome sequences in *Sinocalycanthus chinensis*. *J. Zhejiang Agric. For. Univ.* **2017**, *34*, 589–596.
54. Wu, J.; Huang, Y.H.; Dong, B.; Zhang, C.; Fu, J.X.; Zhou, L.H.; Zhao, H.B. Suitability test of SSR primer of *Sinocalycanthus chinensis* and its application in genetic diversity. *J. Nanjing For. Univ. Nat. Sci. Ed.* **2018**, *42*, 58–66.
55. Bandelj, D.; Jakše, J.; Javornik, B. Assessment of genetic variability of olive varieties by microsatellite and AFLP markers. *Euphytica* **2004**, *136*, 93–102. [[CrossRef](#)]
56. Jiang, Y.M.; Wang, J.F.; Liu, Y.X.; Zhang, Z.B.; Cao, G.P.; Lu, S.B.; Zhu, D. Genetic patterns investigation of wild *Chimonanthus grammatus* MC Liu by using SSR markers. *Acta Ecol. Sin.* **2015**, *35*, 203–209. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.