

Article

Development and Characterization of High-Throughput EST-Based SSR Markers for *Pogostemon cablin* Using Transcriptome Sequencing

Puyue Ouyang^{1,2,†}, Dali Kang^{3,†}, Xiaolu Mo², Enwei Tian⁴ , Yanyu Hu¹ and Rongshao Huang^{1,*}

¹ Agricultural College, Guangxi University, Nanning 530005, China; ouyangpy@gdyzy.edu.cn (P.O.); striverhu@163.com (Y.H.)

² College of Chinese Materia Medica, Guangdong Food and Drug Vocational College, Guangzhou 510520, China; moxl@gdyzy.edu.cn

³ College of Medical Technology, Ningbo College of Health Sciences, Ningbo 315100, China; 202099@nchs.edu.cn

⁴ School of Traditional Chinese Medicine, Southern Medical University, Guangzhou 510515, China; tianenwei@126.com

* Correspondence: hrs17252@gxu.edu.cn; Tel.: +86-771-323-5612

† These authors contributed equally to this work.

Received: 21 June 2018; Accepted: 5 August 2018; Published: 9 August 2018



Abstract: Simple sequence repeats (SSRs) or microsatellite markers derived from expressed sequence tags (ESTs) are routinely used for molecular assisted-selection breeding, comparative genomic analysis, and genetic diversity studies. In this study, we investigated 54,546 ESTs for the identification and development of SSR markers in *Pogostemon cablin* (Patchouli). In total, 1219 SSRs were identified from 1144 SSR-containing ESTs. Trinucleotides (80.8%) were the most abundant SSRs, followed by di- (10.8%), mono- (7.1%), and hexa-nucleotides (1.3%). The top six motifs were CCG/CGG (15.3%), AAG/CTT (15.0%), ACC/GGT (13.5%), AGG/CCT (12.4%), ATC/ATG (9.9%), and AG/CT (9.8%). On the basis of these SSR-containing ESTs, a total of 192 primer pairs were randomly designed and used for polymorphism analysis in 38 accessions collected from different geographical regions of Guangdong, China. Of the SSR markers, 45 were polymorphic and had allele variations from two to four. Furthermore, a transferability analysis of these primer pairs revealed a 10–40% cross-species transferability in 10 related species. This report is the first comprehensive study on the development and analysis of a large set of SSR markers in *P. cablin*. These markers have the potential to be used in quantitative trait loci mapping, genetic diversity studies, and the fingerprinting of cultivars of *P. cablin*.

Keywords: *Pogostemon cablin*; expressed sequence tag (EST); simple sequence repeat (SSR); polymorphism; transferability

1. Introduction

Pogostemon cablin (Blanco) Benth., Lamiaceae (Patchouli) is a hardy, perennial, and aromatic herb that adapts to hot and humid climatic surroundings. It is widely distributed in the wilds of Indonesia, Philippines, Malaysia, and other countries of Southeast Asia, and is extensively cultivated in South China [1]. Patchouli essential oil is one of the most important materials used in the perfume industry [2]. In China, *P. cablin* is also used as a medicine for the treatment of the common cold, headache, fever, vomiting, indigestion, and diarrhea [3,4]. Patchouli oil is also known for its protection against gastric ulcers [5] and its antiseptic, anti-inflammatory, and antifungal properties [6,7]. Nowadays, growing

interest in its fragrance and medicinal applications has promoted patchouli's cultivation in South China. Nevertheless, the therapeutic properties of *P. cablin* are directly correlated with the qualitative and quantitative composition of its components, such as patchouli oil, which are obviously different depending on the cultivar region of origin [8].

In this genomic era, deciphering the genomes of medicinal plants will enable insights into the biochemistry of the genes responsible for secondary metabolism biosynthetic pathways, and then improve the contents of the active ingredients. Recent advances in genome sequencing technologies and assembly methods have resulted in the generation of genome sequences for a wide range of plant species. However, despite the prominent roles of *P. cablin* in both traditional Chinese medicine and the perfume industry, very few studies have examined the genomes of *P. cablin*. He et al. [9] reported the first de novo assembled 1.15-Gb draft genome sequence for *P. cablin* from next-generation sequencing technology in 2016, which is also the first draft of the genome of Lamiaceae.

In modern phylogenetic studies and genetic breeding studies of plant cultivars, molecular markers have become important and efficient tools for genetic diversity assessment, quantitative trait loci (QTL) analysis, and gene mapping. The genetic improvement strategy mainly relies on access to genetic variation in the gene pool [10]. Diverse genetic materials need to be assessed for genetic diversity among the accessions collected from widely separated geographical locations. Molecular phylogenetic studies have clarified many relationships within Lamiaceae, inferred from cpDNA sequences, nuclear DNA, and plastid DNA [11–17], but the genus *Pogostemon* has received relatively scant attention. Some efforts have been made to assess the genetic variability in *P. cablin*; however, only a few molecular marker-based studies have been carried out to achieve this goal, employing, for instance, the random amplification of polymorphic DNA (RAPD) [18,19], inter-simple sequence repeat (ISSR), and sequence-related amplified polymorphism (SRAP) [20]. These studies showed different levels of polymorphism in *P. cablin*, suggesting the necessity to develop a large number of markers that can be utilized for germplasm characterization and QTL analysis for *P. cablin*.

Simple sequence repeats (SSRs) or microsatellites consist of short tandem repeated motifs of 1 to 6 bp units, and occur in both the coding and noncoding regions of the genome [21,22]. SSR markers are a good choice for a broad application in basic and applied plant biology, because of their high level of polymorphism, high abundance, co-dominance, reproducibility, and high rates of transferability across species and genera [21,23,24]. Expressed sequence tag-simple sequence repeats (EST-SSRs) are markers mined from EST sequence collections. As markers, they differ from traditional genomic SSR (gSSR) markers, as they are more likely embedded in functional gene sequences, less costly to identify, and may provide more abundant information. EST-SSR markers are present in more conserved regions in comparison to markers that are generated from genomic sequences, therefore, they show more transferability between species [24,25]. In this study, we examined 54,546 EST sequences obtained from the transcriptome sequencing of *P. cablin* in order to (1) develop EST-derived SSRs for *P. cablin*; (2) analyze the frequency and distribution of EST-SSRs in the expressed portion of the genome; (3) construct a functional annotation and prediction of amino acids from SSR loci; and (4) evaluate the polymorphism and cross-species and genera transferability of EST-SSR markers.

2. Results

2.1. Characterization of EST-SSRs

A total of 54,546 patchouli EST sequences were screened for SSRs. The SSR screening resulted in the identification of 1114 sequences containing 1219 SSRs (Table 1). A total of 1047 sequences contained single SSRs of different motif types, and 67 sequences contained more than two SSRs.

Of the total 1219 SSRs, 1182 (97.0%) were simple repeat motifs and 37 (3.0%) were in compound formation. The highest proportion of repeats (80.8%) was represented by trinucleotide repeats, followed by dinucleotide (10.8%) and mononucleotide (7.1%) repeats, and only 1.3% hexanucleotide SSRs were found (Figure 1a). The trinucleotide and dinucleotide repeats thus represented the majority of EST-SSRs in *P. cablin*. The number of repeats of the different SSR motifs varied from 5 to 24, with a maximum

frequency for five repeats, followed by six and seven repeats. The frequency of the trinucleotide motifs with five repeats was more common (55.1%), followed by six-repeat trinucleotides (19.2%) and six-repeat dinucleotides (7.6%). The frequency of different SSR motifs for each repeat number is presented in Figure 1b. The length of the SSR region (motif length \times repeat numbers) varied from 10 to 57 bases, with a maximum frequency for 15 bases (55.1%), followed by 18 bases (19.2%) and 12 bases (7.6%) (Figure 1c).

Table 1. Summary of the in silico search for simple sequence repeats (SSRs) in *Pogostemon cablin*.

Parameters Used in Screening	Data Generated by MISA
Total number of sequences examined	54,546
Total number of identified SSRs	1219
Number of SSR-containing sequences	1144
Number of sequences containing more than 1 SSR	67
Number of SSRs present in compound formation	37
Total size of examined sequences (bp)	36,417,906

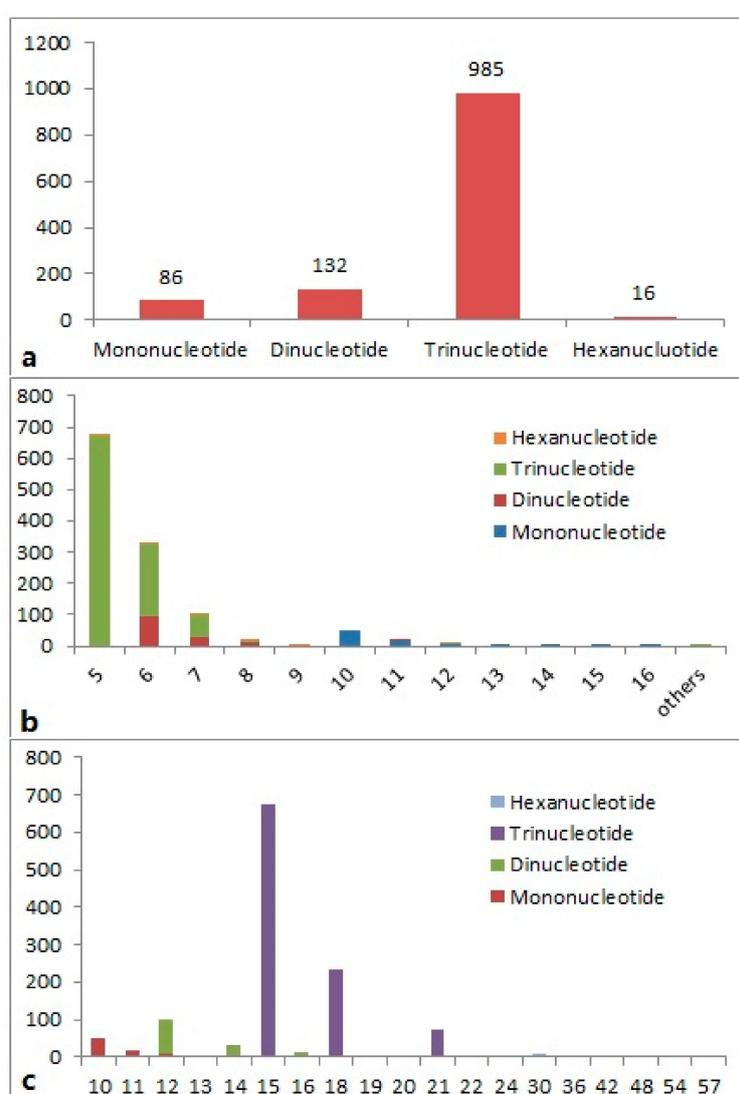


Figure 1. Frequency distribution of the putative expressed sequence tag-simple sequence repeats (EST-SSRs) from patchouli expressed sequence tags (ESTs) based on (a) the distribution of EST-SSR types, (b) the EST-SSRs repeat number, and (c) the EST-SSRs' repeat lengths (bp).

Among the dinucleotide repeats, AG/CT was the most common (90.9%) motif, followed by AC/GT (8.3%) and AT/AT (0.8%). Among the trinucleotide repeats, CCG/CGG was the most common, accounting for 18.9% of the total trinucleotide repeats, followed by the AAG/CTT (18.6%) and ACC/GGT (16.8%) motifs (Table 2).

Table 2. The number of non-redundant SSRs in a set of 1219 patchouli expressed sequence tags (ESTs).

SSR Motif	Number of Repeats												Total
	5	6	7	8	9	10	11	12	13	14	15	>15	
A/T	-	-	-	-	-	47	14	6	2	5	1	2	77
C/G	-	-	-	-	-	5	2	1		1			9
AC/GT	-	10	1										11
AG/CT	-	83	24	10	1		2						120
AT/AT	-											1	1
AAC/GTT	20	6	5										31
AAG/CTT	116	58	9										183
AAT/ATT	7	4											11
ACC/GGT	114	35	14	1								1	165
ACG/CGT	17	6	2										25
ACT/AGT	5	2											7
AGC/CTG	66	20	16	3									105
AGG/CCT	105	36	10										151
ATC/ATG	95	21	5										121
CCG/CGG	127	46	12					1					186
AAAGAT/ATCTTT	2												2
AAAGGC/CCTTTG	2												2
AACAGC/CTGTTG		2											2
AACCCT/AGGGTT				1									1
AATCCC/ATTGGG		1		1									2
AATCTG/AGATTC	1												1
AATGGT/ACCATT	1												1
ACAGCC/CTGTGG					2								2
ACCTCC/AGGTGG	1												1
ACTCCG/AGTCGG			1										1
AGCCCT/AGGGCT	1												1
													Total
													1219

2.2. Categories and Annotations of Unigenes

Based on a Basic Local Alignment Search Tool X (BLASTX) analysis, the EST-SSR-containing unigenes were divided into three classifications according to their functions, namely (a) a biological process; (b) a cellular component; and (c) a molecular function. The unigenes classified as being involved in a biological process were shown to have a putative function in the metabolic process, primary metabolism process, response to stimuli, biological regulation, cellular process, biosynthetic process, and so on (Figure 2a). The unigenes classified as cellular components are shown in Figure 2b. In the cellular component category, the greatest numbers of genes were found in the terms cellular component, cell, and cell part. The unigenes classified in the molecular function category were shown to have a putative function in binding, catalytic activity, DNA binding transcription factor activity, transporter activity, nucleic acid binding transcription factor activity, and so on (Figure 2c).

Among the patchouli active components, the sesquiterpenoids and flavonoids are the most abundant. However, although some of the genes involved in the terpenoid and flavonoid biosynthesis have been isolated and functionally characterized, the global terpenoid and flavonoid biosynthesis pathways in *P. cablin* remain yet to be fully characterized. The gene ontology (GO) analysis showed that the functions of the identified genes are involved various biological processes. A large number of catalysts and transferases were annotated, which suggests that our study may allow for the identification of novel genes involved in the sesquiterpenoids and flavonoids synthesis pathways.

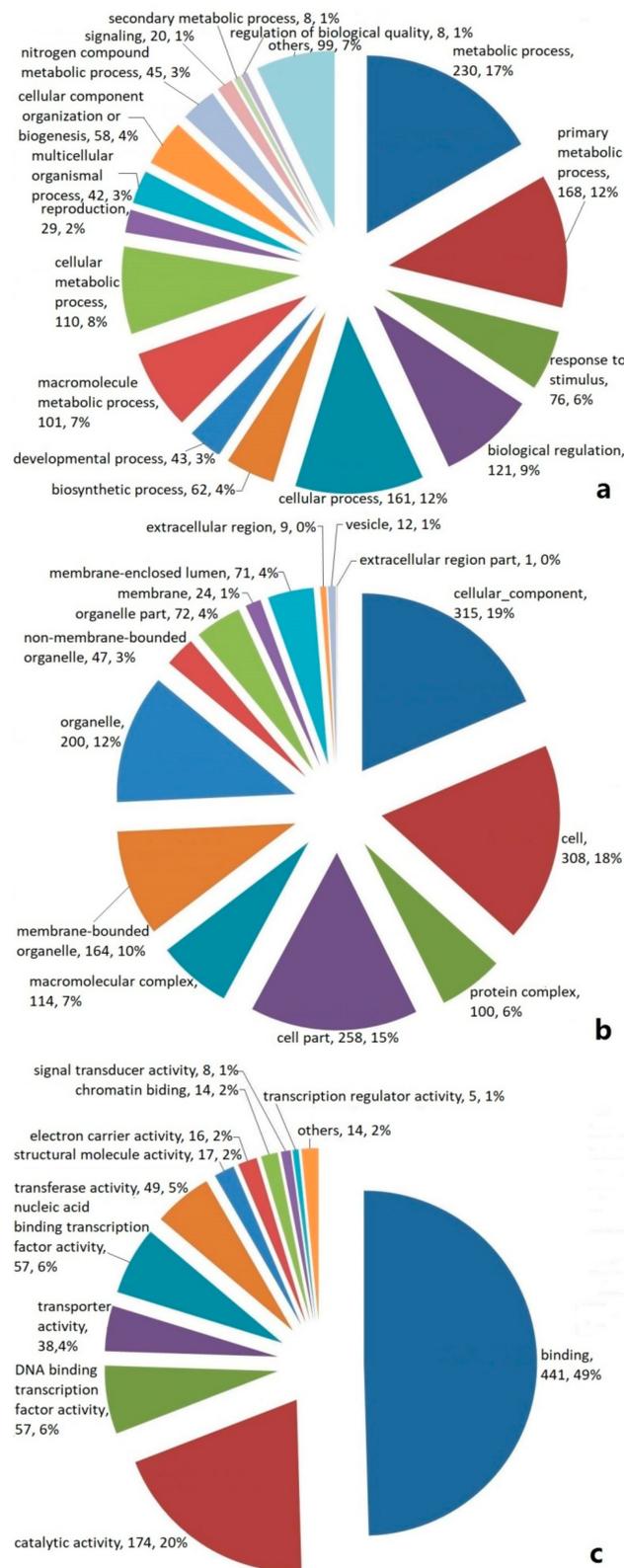


Figure 2. Distribution of EST-SSR-encoded unigenes in *P. cablin* classified according to their biological function, namely, (a) biological process, (b) cellular component, and (c) molecular function.

2.3. Polymorphism and Cross-Species Transferability

In order to validate the EST-SSRs and test their transferability across related species of *P. cablin*, 192 primer pairs were designed randomly and 48 accessions were selected, including 10 from related species that are widely used in traditional Chinese medicine and 38 *P. cablin* accessions. Out of these 192 primer pairs, 45 SSR markers were polymorphic (23.4%) and with allele variations from two to four. Furthermore, only 12 SSR markers were found to be interspecifically amplifiable, and 10–40% cross-species transferability in 10 related species were revealed. The characteristics of 45 EST-derived SSRs for *P. cablin* are given in the Supplementary Materials.

3. Discussion

In this study, EST sequences were obtained and used for the development of a novel set of functional markers. Among all of the EST-SSRs developed, the trinucleotide motifs were predominant, corresponding to 80.8% of the EST-SSRs, which was 7.5 times higher than the amount of the second abundant motif (dinucleotide, 10.8%). This may simply be reflecting the fact that trinucleotide SSRs do not cause frame shifts when present in ESTs, because they are multiples of three (i.e., of the number of nucleotides in a codon). This is not the case for di-, tetra-, and penta-nucleotide motifs. The most common SSR lengths were 12, 15, 18 and 21, which were all multiples of three, supporting the idea that there is a positive selection pressure for the in-frame SSRs in genic sequences [26].

EST-SSRs are derived from the expressed portion of the genome, therefore, they can be assayed as gene-based functional markers for diversity assessment, gene mapping, and marker assisted selection. The gene ontology (GO) classifies genes according to their function in three categories (i.e., molecular function, biological process, and cellular component) [27]. In our study, most of the identified unigenes matched unique known proteins that are present in the public databases, implying that the transcriptome sequencing yielded a great number of unique genes in *P. cablin*. A large number of unigenes were assigned to a wide range of GO functions, which indicated that the transcriptome data represented a broad diversity of transcripts in *P. cablin*. Similar results were also reported in other species [28–32]. The unigenes related to a biological process, such as the response to a stimulus and a secondary metabolic process, should be paid more attention when studying the medicinal plants whose activities rely on the content of secondary functional components generated by the response against a stressor. By employing transcriptome sequencing and gene annotation, such a large number of transcriptome sequences will provide an excellent resource for gene isolation and gene expression profile analysis in *P. cablin*.

In China, cultivars of *P. cablin* in different regions have different morphological characters and essential oil constituents. It was reported that *P. cablin* in China could be classified into two chemotypes, the pogostone type and patchoulol type, whose essential oils contain extremely high amounts of pogostone and patchoulol, respectively [8,33]. The pogostone type patchouli plant is mainly cultivated in the areas of Guangzhou and Zhaoqing in Guangdong Province, and is called ‘Paixiang’ and ‘Zhaoxiang’, respectively, in Chinese, whereas the patchoulol type *P. cablin* is cultivated in the areas of Zhanjiang in Guangdong Province and Wanning in Hainan Province, and is called ‘Zhanxiang’ and ‘Nanxiang’, respectively. Traditionally, Paixiang is considered an authentic herb, according to the traditional Chinese medicine theory. Additionally, the patchoulol type is mainly used in the perfume industry because of its higher yield of essential oil compared with the pogostone type *P. cablin* [33]. Unfortunately, Paixiang, which is considered a genuine herbal material, is becoming rare, because its cultivation has been greatly decreased as a consequence of the expansion of the urban area of Guangzhou city. In order to explore the genetic features of the authentic cultivar and to accelerate the development of genetically improved cultivars, it is desirable to understand the inheritance pattern of complex quantitative traits like yield and adaptability. This can be facilitated by developing DNA-based markers, such as SSRs, and applying them to estimate the level of genetic diversity and to construct a framework linkage map and QTL mapping for tagging the target traits [24,34,35]. Molecular markers are widely distributed in various regions of the genome. By comparing the polymorphism

of molecular markers randomly-distributed in the genome, the diversity of plants can be evaluated and used to cluster the species, and to understand their phylogeny and phylogenetic relationships. In this study, the cross-species transferability of EST-SSRs could help understand the evolution of microsatellites in a phylogenetic context, and EST-SSRs on a genic loci might play an important role in speciation or gene functionality diversification during the evolutionary process. The development of patchouli EST-SSR markers provides a powerful tool for the study of phylogenetic *Pogostemon* genus and Labiatae family. More importantly, the EST-SSR markers developed in this study are a good resource for molecular marker-assisted breeding. Consequently, these EST-SSR markers are useful tools to map the genes related to bioactive components. This study, for the first time, reports the development of high-throughput EST-SSR markers in *P. cablin*, and assesses their transferability across a wide spectrum of related species and genera. The patchouli EST-SSR markers that have been developed are valuable genetic resources for investigating species relationships and for comparative mapping in the Labiatae family.

4. Materials and Methods

4.1. Plant Materials and DNA Isolation

A total of 38 patchouli accessions and one accession each of 10 related plants (*Mesona chinensis* Benth., *Prunella vulgaris* L., *Salvia bowleyana* Dunn, *Stachys geobombycis* C. Y. Wu, *Teucrium viscidum*, *Dysophylla aurl* Cularia [L.] Blume, *Mentha haplocalyx* Briq., *Mentha spicata* Linn., *Leonurus artemisia* [Laur.] S. Y. Hu F, and *Salvia plebeia* R. Br.) were used for the examination of polymorphism and for the transferability of patchouli EST-derived SSR-markers (Table 3).

Genomic DNA was isolated from fresh, young leaf tissues of patchouli accessions using the Cetyltrimethylammonium bromide (CTAB) method [36]. The genomic DNA was purified using the GenElute genomic DNA miniprep kit (SIGMA Aldrich, St. Louis, MO, USA). The DNA quality and final concentration were estimated by agarose gel electrophoresis, using a known concentration of uncut λ DNA as a standard.

Table 3. List of 38 patchouli accessions used in the analysis of genetic diversity.

Resource	No. of Accessions Sampled	Origin	Latitude (N)	Longitude (E)
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	3	Longdong, Guangzhou city	23°12'20''	113°22'36''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Shipaiensis	3	Longdong, Guangzhou city	23°12'20''	113°22'36''
<i>Pogostemon cablin</i> (Blanco) Benth	3	Longdong, Guangzhou city	23°12'20''	113°22'36''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	6	Didou Town, Zhaoqing City	23°33'58''	112°43'03''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	4	Liantang Town, Zhaoqing City	22°57'04''	112°27'54''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Gaoyaoensis	4	Liantang Town, Zhaoqing City	22°57'04''	112°27'54''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	5	Yingli Town, Zhanjiang City	20°29'51''	109°58'10''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	5	Yingli Town, Zhanjiang City	20°33'51''	110°04'03''
<i>Pogostemon cablin</i> (Blanco) Benth cv. Hainangensis	5	Tanshui Town, Yangjiang City	22°04'25''	111°30'19''

4.2. Transcriptome Sequencing of *P. cablin*

For the transcriptome sequencing, the young leaf tissues of three patchouli cultivars were collected and used to build up RNA libraries. The samples were immersed in liquid nitrogen immediately after collection and then stored at $-80\text{ }^{\circ}\text{C}$. The total RNA was isolated using the Trizol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. The quality of the total RNA was checked on a spectrophotometer, the NanoDrop 2000c (Thermo Scientific, Waltham, MA, USA), and a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). The RNA samples were stored at $-80\text{ }^{\circ}\text{C}$ until further use. Poly(A) mRNA was isolated with oligo(dT) beads. The first-strand cDNA was synthesized using a random hexamer-primer and reverse transcriptase (Invitrogen, Carlsbad, CA, USA). The second-strand cDNA was synthesized using RNase H (Invitrogen, Carlsbad, CA, USA) and DNA polymerase I (New England BioLabs, Ipswich, MA, USA). Then, the cDNA libraries were prepared according to Illumina's protocols (San Diego, CA, USA), and were subjected to the Illumina Genome Analyzer's proprietary sequencing-by-synthesis method. The sequencing was carried out at the Beijing Genomics Institute (BGI, Shenzhen, China). In total, three single-end cDNA libraries were constructed for each patchouli cultivar. The raw sequencing dataset for the *P. cablin* was submitted to the NCBI/SRA database under accession number SRP150583.

4.3. De Novo Assembly and Mining for EST-SSRs

The raw sequencing reads were first filtered by removing the invalid reads, and then including the reads with an adaptor contamination, with ambiguous 'N' bases at a ratio greater than 5% and reads with a more than 50% base with a quality lower than 20 in one sequence. The short clean reads were assembled using Trinity software to construct contigs. These contigs were used for the further process of sequence clustering, with the software TGICL (TIGR Gene Indices clustering tools) to form unigenes [37]. The 54,546 ESTs were searched for microsatellites using MicroSatellite (MISA) scripts (<http://pgrc.ipk-gatersleben.de/misa/>). The criteria for the SSR search by the MISA were repeat stretches having a minimum of six repeat units for dinucleotide SSRs, and five repeat units in case of tri-, tetra-, penta-, and hexa-nucleotide SSRs. The identified EST-SSR sequences were deposited in the Genebank so as to obtain the accession numbers. The putative function of the developed functional markers-encoding unigenes was found using the BLASTX tool of NCBI, assuming an *e*-value $< 1 \times 10^{-10}$ as a significant homology.

4.4. Designing EST-SSR Primers

The primer sets were designed using Primer3 [38] and the Primer3 interface modules scripts by the MISA developers at <http://pgrc.ipk-gatersleben.de/misa/primer3.html>. The parameters used for primer design were (1) a primer length of 18–24 bp, with an optimum of 20 bp; (2) an annealing temperature of 50–60 $^{\circ}\text{C}$, with an optimum of 55 $^{\circ}\text{C}$; (3) a percentage GC in the range of 40–50; and (4) a product size in the range of 100–400 bp.

4.5. Polymerase Chain Reaction (PCR) Amplification and Polymorphism Validation

A total of 38 wild patchouli accessions and one accession each of 10 related species were used for the examination of the polymorphism and transferability of the patchouli EST-derived SSR-markers. The PCR amplifications of the genomic DNA were performed in a 10 μL solution containing the following components: 1 \times PCR buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl_2), 0.2 mM of each dNTP, 0.2 μM of each locus-specific primer, 0.075 μM M13 (–21) primer labelled at the 5' end with 6-carboxy-fluoresceine (FAM) and hexachloro-6-carboxy-fluoresceine (HEX), 0.25 U of Taq polymerase (Fermentas, Vilnius, Lithuania), and 20 ng of genomic DNA. The amplification was performed in PE9700 (Applied Biosystems, Waltham, MA, USA) or T-gradient (Biometra, Jena Germany) thermalcyclers, using a two-step PCR protocol, with an initial touchdown cycle of 94 $^{\circ}\text{C}$ for 5 min, followed by five cycles of 30 s at 94 $^{\circ}\text{C}$, 30 s at 60 $^{\circ}\text{C}$ (which was lowered by 1 $^{\circ}\text{C}$ for each

cycle), and 30 cycles at 55 °C, and ending with a 20 min extension step at 72 °C. The PCR products were detected by gel electrophoresis in a 2% agarose-1 × Tris-borate-EDTA (TBE) buffer. The sequence analysis was performed on an ABI 3730 DNA Sequence Analyzer.

5. Conclusions

Currently, the development of functional molecular markers such as EST-SSRs is becoming a valuable objective with regards to plants, especially in marker-assisted breeding programs. In this research article, a set of novel EST-SSRs developed from cDNA libraries were identified and characterized. A total of 45 polymorphic EST-SSR markers were ultimately screened and confirmed across 38 patchouli populations and, interestingly, showed a 10–40% cross-transferability, in 10 related species of Lamiaceae. There is no doubt that these novel EST-SSRs will be helpful for future research on genetic diversity, population structure, evolutionary processes, linkage map construction, and QTL mapping for *P. cablin*. In addition, as a codominant marker, EST-SSRs are abundant in the genome, and can be used as an ideal tool for maker-assisted breeding, constructing genetic maps, and for the cloning of key genes for *P. cablin*.

Supplementary Materials: The supplementary materials are available online.

Author Contributions: R.H. conceived and designed the experiments; P.O. performed the experiments; D.K. and Y.H. analyzed the data; D.K. and P.O. wrote the paper; and E.T. reviewed the paper. All of the authors read and approved the final manuscript.

Funding: This research was funded by [Science and Technology Program of Guangzhou City] grant number [201604020041]; [Science and Technology Program of Guangdong Province] grant number [2015A040404029]; [Natural Science Foundation of Guangdong Province] grant number [2018A030313255]; and [Natural Science Project of the Guangdong Food and Drug Vocational College] grant number [301009A1807].

Acknowledgments: We gratefully acknowledge Shaohua Zeng from the South China Botanical Garden in the Chinese Academy of Sciences for his valuable support in the laboratory tasks and in reviewing the paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

EST	expressed sequence tag
SSR	simple sequence repeat
RAPD	random amplification of polymorphic DNA
ISSR	inter-simple sequence repeat
SRAP	sequence-related amplified polymorphism
QTL	quantitative trait loci
BLASTX	Basic Local Alignment Search Tool X
PCR	polymerase chain reaction
GO	gene ontology

References

1. Miyazawa, M.; Okuno, Y.; Nakamura, S.; Kosaka, H. Antimutagenic activity of flavonoids from *Pogostemon cablin*. *J. Agric. Food Chem.* **2000**, *48*, 642–647. [[CrossRef](#)] [[PubMed](#)]
2. Hasegawa, Y.; Tajima, K.; Toi, N.; Sugimura, Y. An additional constituent occurring in the oil from a patchouli cultivar. *Flavour Fragr. J.* **1992**, *7*, 333–335. [[CrossRef](#)]
3. Chinese Pharmacopoeia Commission. *Chinese Pharmacopoeia Part I*, 2015 ed.; Chemical Industry Press: Beijing, China, 2015; p. 45.
4. Wu, Y.G.; Guo, Q.S.; Zheng, H.Q. Textual research on history of introduction and herbal medicine of *Pogostemon cablin*. *Chin. J. Chin. Mater. Med.* **2007**, *32*, 2114–2117.

5. Chen, H.M.; Liao, H.J.; Liu, Y.H.; Zheng, Y.F.; Wu, X.L.; Su, Z.Q.; Zhang, X.; Lai, Z.Q.; Lai, X.P.; Lin, Z.X.; et al. Protective effects of pogostone from *Pogostemonis herba* against ethanol-induced gastric ulcer in rats. *Fitoterapia* **2015**, *100*, 110–117. [[CrossRef](#)] [[PubMed](#)]
6. Liu, H.M.; Li, Y.L.; Yang, J.Z. Review of pharmacological effects of patchouli oil. *Chin. Pharm.* **2012**, *23*, 4506–4508.
7. Yi, Y.Y.; He, J.J.; Su, J.Q.; Kong, S.Z.; Su, J.Y.; Li, Y.C.; Huang, S.H.; Li, C.W.; Lai, X.P.; Su, Z.R. Synthesis and antimicrobial evaluation of pogostone and its analogues. *Fitoterapia* **2013**, *84*, 135–139. [[CrossRef](#)] [[PubMed](#)]
8. Luo, J.P.; Liu, Y.P.; Feng, Y.F.; Guo, X.L.; Cao, H. Two chemotypes of *Pogostemon cablin* and influence of region of cultivation and harvesting time on volatile oil composition. *Acta Pharm. Sin.* **2003**, *38*, 307–310.
9. He, Y.; Xiao, H.T.; Deng, C.; Xiong, L.; Nie, H.; Peng, C. Survey of the genome of *Pogostemon cablin* provides insights into its evolutionary history and sesquiterpenoid biosynthesis. *Sci. Rep.* **2016**, *6*, 26405. [[CrossRef](#)] [[PubMed](#)]
10. Yadav, H.K.; Ranjan, A.; Asif, M.H.; Mantri, S.; Sawant, S.V.; Tuli, R. EST-derived SSR markers in *Jatropha curcas* L.: Development, characterization, polymorphism, and transferability across the species/genera. *Tree Genet. Genomes* **2011**, *7*, 207–219. [[CrossRef](#)]
11. Trusty, J.L.; Olmstead, R.; Bogler, D.J.; Santos-Guerra, A.; Francisco-Ortega, J. Using molecular data to test a biogeographic connection of the Macaronesian genus *Bystropogon* (Lamiaceae) to the New World: A case of conflicting phylogenies. *Syst. Bot.* **2004**, *29*, 702–715. [[CrossRef](#)]
12. Bräuchler, C.; Meimberg, H.; Heubl, G. Molecular phylogeny of Menthinae (Lamiaceae, Nepetoideae, Mentheae)—Taxonomy, biogeography and conflicts. *Mol. Phylogenet. Evol.* **2010**, *55*, 501–523. [[CrossRef](#)] [[PubMed](#)]
13. Bendiksby, M.; Thorbek, L.; Scheen, A.C.; Lindqvist, C.; Ryding, O. An updated phylogeny and classification of Lamiaceae subfamily Lamioideae. *Taxon* **2011**, *60*, 471–484.
14. Drew, B.; Sytsma, K.J. Phylogenetics, biogeography, and staminal evolution in the tribe Mentheae (Lamiaceae). *Am. J. Bot.* **2012**, *99*, 933–953. [[CrossRef](#)] [[PubMed](#)]
15. Xiang, C.L.; Zhang, Q.; Scheen, A.C.; Cantino, P.; Funamoto, T.; Peng, H. Molecular phylogenetics of *Chelonopsis* (Lamiaceae: Gomphostemmatae) as inferred from nuclear and plastid DNA and morphology. *Taxon* **2013**, *62*, 375–386. [[CrossRef](#)]
16. Roy, T.; Chang, T.H.; Lan, T.; Lindqvist, C. Phylogeny and biogeography of New World Stachydeae (Lamiaceae) with emphasis on the origin and diversification of Hawaiian and South American taxa. *Mol. Phylogenet. Evol.* **2013**, *69*, 218–238. [[CrossRef](#)] [[PubMed](#)]
17. Roy, T.; Cole, L.W.; Chang, T.H.; Lindqvist, C. Untangling reticulate evolutionary relationships among New World and Hawaiian mints (Stachydeae, Lamiaceae). *Mol. Phylogenet. Evol.* **2015**, *89*, 46–62. [[CrossRef](#)] [[PubMed](#)]
18. Pan, C.M.; Li, W.; He, H.; Deng, W.Q.; Li, T.H.; Xu, H.H. Study on intraspecific genetic diversity in different plant populations of *Pogostemon cablin*. *China J. Chin. Mater. Med.* **2006**, *31*, 723–726.
19. Zhang, Y. Study on Guangdong Authentic and Superior Medical Material *Pogostemon cablin* by GC–MS Fingerprinting and DNA Molecular Markers Analysis. Ph.D. Thesis, Beijing University of Chinese Medicine, Beijing, China, 2007.
20. Wu, Y.G.; Wu, L.H.; He, J.C. Progress in Research of the Hereditary Basis and Biotechnology of *Pogostemon cablin* (Blanco) Benth. *J. Trop. Org.* **2010**, *1*, 288–292.
21. Powell, W.; Machray, G.C.; Provan, J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1996**, *1*, 215–222. [[CrossRef](#)]
22. Toth, G.; Gaspari, Z.; Jurka, J. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* **2000**, *10*, 967–981. [[CrossRef](#)] [[PubMed](#)]
23. Gupta, P.K.; Rustgi, S.; Sharma, S.; Singh, R.; Kumar, N.; Balyan, H.S. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Gen. Genom.* **2003**, *270*, 315–323. [[CrossRef](#)] [[PubMed](#)]
24. Varshney, R.K.; Sigmund, R.; Borner, A.; Korzun, V.; Stein, N.; Sorrels, M.E.; Langridge, P.; Graner, A. Interspecific Transferability and Comparative Mapping of Barley EST-SSR Markers in Wheat, Rye and Rice. *Plant Sci.* **2005**, *168*, 195–202. [[CrossRef](#)]

25. Portis, E.; Nagy, I.; Sasvari, Z.; Stigel, A.; Barchi, L.; Lanteri, S. The design of Capsicum spp. SSR assays via analysis of in silico DNA sequence, and their potential utility for genetic mapping. *Plant Sci.* **2007**, *172*, 640–648. [[CrossRef](#)]
26. Cloutier, S.; Niu, Z.X.; Datla, R.; Duguid, S. Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). *Theor. Appl. Genet.* **2009**, *119*, 53–63. [[CrossRef](#)] [[PubMed](#)]
27. DeJongh, M.; Van Dort, P.; Ramsay, B. Linking molecular function and biological process terms in the ontology for gene expression data analysis. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2004**, *4*, 2984–2986. [[PubMed](#)]
28. Guo, S.G.; Liu, J.G.; Zheng, Y.; Huang, M.Y.; Zhang, H.Y.; Gong, G.Y.; He, H.J.; Ren, Y.; Zhong, S.L.; Fei, Z.J.; et al. Characterization of transcriptome dynamics during watermelon fruit development: Sequencing, assembly, annotation and gene expression profiles. *BMC Genom.* **2011**, *12*, 454. [[CrossRef](#)] [[PubMed](#)]
29. Zhang, J.N.; Liang, S.; Duan, J.L.; Wang, J.; Chen, S.L.; Cheng, Z.S.; Zhang, Q.; Liang, X.Q.; Li, Y.R. De novo assembly and characterization of the transcriptome during seed development, and generation of genic-SSR markers in Peanut (*Arachis hypogaea* L.). *BMC Genom.* **2012**, *13*, 90. [[CrossRef](#)] [[PubMed](#)]
30. Huang, Q.S.; Dong, S.J.; Fang, C.; Wu, X.L.; Ye, T.; Lin, Y. Deep sequencing-based transcriptome profiling analysis of *Oryzias melastigma* exposed to PFOS. *Aquat. Toxicol.* **2012**, *120–121*, 54–58. [[CrossRef](#)] [[PubMed](#)]
31. Huang, L.L.; Yang, X.; Sun, P.; Tong, W.; Hu, S.Q. The first illumina-based de novo transcriptome sequencing and analysis of safflower flowers. *PLoS ONE* **2012**, *7*, e38653.
32. Gao, X.G.; Han, J.B.; Lu, Z.C.; Li, Y.F.; He, C.B. Characterization of the spotted seal *Phoca largha* transcriptome using Illumina paired-end sequencing and development of SSR markers. *Comp. Biochem. Physiol. D* **2012**, *7*, 277–284. [[CrossRef](#)] [[PubMed](#)]
33. Guo, X.L.; Feng, Y.F.; Luo, J.P. Re-study on characteristic fingerprint of volatile oil from herba Pogostemonis by GC. *J. Chin. Med. Mater.* **2004**, *27*, 903–908.
34. Gupta, M.N.; Roy, I. Applied biocatalysis: An overview. *Indian J. Biochem. Biophys.* **2002**, *39*, 220–228. [[PubMed](#)]
35. Collard, B.C.Y.; Jahufer, M.Z.Z.; Brouwer, J.B.; Pang, E.C.K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **2005**, *142*, 169–196. [[CrossRef](#)]
36. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
37. Zhou, X.J.; Wang, Y.Y.; Xu, Y.N.; Yan, R.S.; Zhao, P.; Liu, W.Z. 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*, 12855–12870. [[CrossRef](#)] [[PubMed](#)]
38. Rozen, S.; Skaletsky, H.J. Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*; Krawetz, S., Misener, S., Eds.; Humana Press: Totowa, NJ, USA, 2000; pp. 365–386.

Sample Availability: Samples of the compounds are not available from the authors.



© 2018 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 (<http://creativecommons.org/licenses/by/4.0/>).