

## Article

# Assessment of the Genetic Relationship and Population Structure in Oil-Tea Camellia Species Using Simple Sequence Repeat (SSR) Markers

Heqin Yan <sup>1,2</sup>, Huasha Qi <sup>3</sup>, Yang Li <sup>1,2</sup>, Yougen Wu <sup>1,4</sup>, Yong Wang <sup>5</sup>, Jianmiao Chen <sup>1,4,\*</sup> and Jing Yu <sup>1,2,\*</sup> 

<sup>1</sup> Sanya Nanfan Research Institute of Hainan University, Hainan Yazhou Bay Seed Laboratory, Sanya 572025, China

<sup>2</sup> Key Laboratory for Quality Regulation of Tropical Horticultural Plants of Hainan Province, College of Horticulture, Hainan University, Haikou 570228, China

<sup>3</sup> Institute of Tropical Horticulture Research, Hainan Academy of Agricultural Sciences, Haikou 571100, China

<sup>4</sup> Engineering Research Center for the Selection and Breeding of New Tropical Crop Varieties of Ministry of Education, College of Tropical Crops, Hainan University, Haikou 570228, China

<sup>5</sup> Ministry of Education Key Laboratory for Ecology of Tropical Islands, College of Life Sciences, Hainan Normal University, Haikou 571158, China

\* Correspondence: cjm@hainanu.edu.cn (J.C.); yujinghxy@163.com (J.Y.)

**Abstract:** Oil-tea camellia trees, the collective term for a class of economically valuable woody oil crops in China, have attracted extensive attention because of their rich nutritional and pharmaceutical value. This study aimed to analyze the genetic relationship and genetic diversity of oil-tea camellia species using polymorphic SSR markers. One-hundred and forty samples of five species were tested for genetic diversity using twenty-four SSR markers. In this study, a total of 385 alleles were identified using 24 SSR markers, and the average number of alleles per locus was 16.0417. The average Shannon's information index (I) was 0.1890, and the percentages of polymorphic loci (P) of oil-tea camellia trees were 7.79–79.48%, indicating that oil-tea camellia trees have low diversity. Analysis of molecular variance (AMOVA) showed that the majority of genetic variation (77%) was within populations, and a small fraction (23%) occurred among populations. Principal coordinate analysis (PCoA) results indicated that the first two principal axes explained 7.30% (PC1) and 6.68% (PC2) of the total variance, respectively. Both UPGMA and PCoA divided the 140 accessions into three groups. *Camellia oleifera* clustered into one class, *Camellia vietnamensis* and *Camellia gauchowensis* clustered into one class, and *Camellia crapnelliana* and *Camellia chekiangoleosa* clustered into another class. It could be speculated that the genetic relationship of *C. vietnamensis* and *C. gauchowensis* is quite close. SSR markers could reflect the genetic relationship among oil-tea camellia germplasm resources, and the results of this study could provide comprehensive information on the conservation, collection, and breeding of oil-tea camellia germplasms.

**Keywords:** oil-tea camellia; SSR; genetic relationship; genetic diversity; population structure



**Citation:** Yan, H.; Qi, H.; Li, Y.; Wu, Y.; Wang, Y.; Chen, J.; Yu, J. Assessment of the Genetic Relationship and Population Structure in Oil-Tea Camellia Species Using Simple Sequence Repeat (SSR) Markers. *Genes* **2022**, *13*, 2162. <https://doi.org/10.3390/genes13112162>

Academic Editor: Pietro Gramazio

Received: 28 September 2022

Accepted: 17 November 2022

Published: 19 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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

Oil-tea camellia trees is the collective term for a group of plants of high economic value. There are approximately 50 species of these trees, and they belong to the family Theaceae [1]. Oil-tea camellia trees have high value and a wide range of uses. They can be used as chemical bioenergetics, chemical feedstock, and a nutrient source [2,3]. Oil-tea camellia trees have a long history of cultivation in China and are mainly distributed in areas south of the Yangtze River Valley [4,5]. The main cultivated species are *Camellia chekiangoleosa*, *Camellia oleifera*, *Camellia crapnelliana*, *Camellia vietnamensis* Huang, etc. [5,6]. Nevertheless, the quality and oil yield of oil-tea camellia trees may vary depending on the species [3]. Thus, it is essential to form molecular markers for identification of populations

or species to support breeding improvement and promote the development of genetic resources for oil-tea camellia trees.

Due to its extensive planting under different ambient conditions in China, oil-tea camellia trees have formed species with different growth habits, morphological characteristics, and degrees of oil quality [7]. *C. vietnamensis* is a species of oil-tea tree from Hainan Island, the southernmost city in China with a unique geographical location and superior climate [8], and some other tropical countries, such as Thailand and Vietnam [9]. *C. vietnamensis* from Hainan Island, which is considered an independent and traditional plant resource according to the long-term isolation from the mainland [7,10], is somewhat different from *C. oleifera*, which is widely grown in mainland China. It is more suitable for a tropical climate, has a large amount of genetic variation, and has higher contents of active ingredients in the oil [1,8].

Simple sequence repeats (SSRs), also called short tandem repeats (STRs) or microsatellites, are widely distributed in the genomes of animals and plants [11]. The random distribution of SSRs in the genome, together with the high level of allelic variation in microsatellite loci, makes them an ideal marker for studying population structure and genetic relationships [12,13]. Designing suitable genetic markers using SSRs allows detailed understanding of the composition and regulatory mechanisms of loci controlling quantitative or disease-resistance traits, allowing one to construct genetic linkage groups with genetic markers [13]. Further manipulation of these genes, the identification and cloning of QTLs affecting target traits, and studying the diversity of population genetics will assist with reaching the goal of marker-assisted selection of an improved population or genotypic selection of an improved population [14]. Several polymorphic SSR markers have been built and used to analyze population structure and genetic relationship in *Camellia* species [15], such as *Camellia sinensis* [11], *C. chekiangoleosa* [16], *C. oleifera* [17], *Camellia japonica* [15], and *Camellia fascicularis* [18]. Huang for the first time analyzed the inter-species hybrid introgression and genetic structure between *Camellia meiocarpa* and *C. oleifera* by SSR markers [19]. Combining morphological traits and SSR markers analysis, He et al. found that *C. oleifera* had abundant genetic variation [17]. An unidentified oil-tea *Camellia* species from Hainan was identified by the chloroplast genome sequences and SSR analysis [1]. As a consequence, it was feasible to study the population structure and genetic relationships of oil-tea camellia species using SSRs.

To date, molecular marker studies in *Camellia* species have mainly involved SSRs [11,17], RAPD [20,21]), ISSR [22,23], and so on. However, few SSRs studies have compared *C. vietnamensis* and other *Camellia* species. Therefore, in this study, we collected SSR molecular markers from 140 oil-tea camellia samples, followed by the non-hierarchical analysis of molecular variance (AMOVA), the unweighted pair group method with arithmetic (UPGMA), principal coordinates analysis (PCoA), and population structure analysis, in the hope of providing some data basis and theoretical basis for the delineation of the relatives, resource system, and population structure of oil-tea camellia species.

## 2. Materials and Methods

### 2.1. Materials

A collection of 140 oil-tea camellia accessions was used in this study, which were divided into 5 groups (Table 1), including 114 oil-tea camellia leaves and 26 oil-tea camellia seeds. The samples in this study were identified by Prof. Kaibing Zhou in 2018. Among them, the leaves included 95 *C. oleifera*, 17 *C. vietnamensis* species, and 2 *C. chekiangoleosa*. The seeds included 13 *C. oleifera*, 3 *C. chekiangoleosa*, 5 *C. crapnelliana*, and 5 *Camellia gauchowensis* specimens. Details of the samples are shown in Table 1 and Figure S1.

**Table 1.** Detailed information for the 140 oil-tea camellia accessions used in this study.

No.	Name	Tissue	Species	Group	Origin	Location
1	1_11		<i>C. oleifera</i>	S1		
2	15_15		<i>C. oleifera</i>	S1		
3	15_2		<i>C. oleifera</i>	S1		
4	15_3		<i>C. oleifera</i>	S1		
5	2_18		<i>C. oleifera</i>	S1		
6	3_18		<i>C. oleifera</i>	S1		
7	3_8		<i>C. oleifera</i>	S1		
8	4_13		<i>C. oleifera</i>	S1		
9	4_18		<i>C. oleifera</i>	S1		
10	5_3		<i>C. oleifera</i>	S1		
11	A29		<i>C. oleifera</i>	S1		
12	A7		<i>C. oleifera</i>	S1		
13	A8		<i>C. oleifera</i>	S1		
14	B18		<i>C. oleifera</i>	S1		
15	B19		<i>C. oleifera</i>	S1		
16	B22		<i>C. oleifera</i>	S1		
17	B26		<i>C. oleifera</i>	S1		
18	B30		<i>C. oleifera</i>	S1		
19	B33		<i>C. oleifera</i>	S1		
20	B34		<i>C. oleifera</i>	S1		
21	B6		<i>C. oleifera</i>	S1		
22	CL18		<i>C. oleifera</i>	S1		
23	CL23		<i>C. oleifera</i>	S1		
24	CL3		<i>C. oleifera</i>	S1		
25	CL40		<i>C. oleifera</i>	S1		
26	CL4		<i>C. oleifera</i>	S1		
27	CL53		<i>C. oleifera</i>	S1		
28	CR11	Leaves	<i>C. oleifera</i>	S1	Nursery of oil-tea camellia germplasm resources, Danzhou campus, Hainan University, China	109°29'45" E, 19°30'28" N
29	D2		<i>C. oleifera</i>	S1		
30	D5		<i>C. oleifera</i>	S1		
31	D6		<i>C. oleifera</i>	S1		
32	E10		<i>C. oleifera</i>	S1		
33	GC11		<i>C. oleifera</i>	S1		
34	GC5		<i>C. oleifera</i>	S1		
35	GC7		<i>C. oleifera</i>	S1		
36	GC8		<i>C. oleifera</i>	S1		
37	H12		<i>C. oleifera</i>	S1		
38	H13	<i>C. oleifera</i>	S1			
39	H15	<i>C. oleifera</i>	S1			
40	H17	<i>C. oleifera</i>	S1			
41	H18	<i>C. oleifera</i>	S1			
42	H19	<i>C. oleifera</i>	S1			
43	H1	<i>C. oleifera</i>	S1			
44	H20	<i>C. oleifera</i>	S1			
45	H3	<i>C. oleifera</i>	S1			
46	H5	<i>C. oleifera</i>	S1			
47	H6	<i>C. oleifera</i>	S1			
48	H7	<i>C. oleifera</i>	S1			
49	H9	<i>C. oleifera</i>	S1			
50	HUA16	<i>C. oleifera</i>	S1			
51	HUA23	<i>C. oleifera</i>	S1			
52	HUA32	<i>C. oleifera</i>	S1			
53	HUA42	<i>C. oleifera</i>	S1			
54	HUA44	<i>C. oleifera</i>	S1			
55	HUA88	<i>C. oleifera</i>	S1			
56	HUA97	<i>C. oleifera</i>	S1			
57	HY14	<i>C. oleifera</i>	S1			
58	HY22	<i>C. oleifera</i>	S1			
59	HY29	<i>C. oleifera</i>	S1			

Table 1. Cont.

No.	Name	Tissue	Species	Group	Origin	Location
60	HY52		<i>C. oleifera</i>	S1		
61	HY8		<i>C. oleifera</i>	S1		
62	K6		<i>C. oleifera</i>	S1		
63	KE13		<i>C. oleifera</i>	S1		
64	KE2		<i>C. oleifera</i>	S1		
65	KE5		<i>C. oleifera</i>	S1		
66	MQ150		<i>C. oleifera</i>	S1		
67	N1		<i>C. oleifera</i>	S1		
68	N26		<i>C. oleifera</i>	S1		
69	N3		<i>C. oleifera</i>	S1		
70	N41		<i>C. oleifera</i>	S1		
71	N8		<i>C. oleifera</i>	S1		
72	N9		<i>C. oleifera</i>	S1		
73	SHI11		<i>C. oleifera</i>	S1		
74	SK15-2		<i>C. oleifera</i>	S1		
75	SK15-5		<i>C. oleifera</i>	S1		
76	SK9-1		<i>C. oleifera</i>	S1		
77	SK		<i>C. oleifera</i>	S1		
78	XIAN207		<i>C. oleifera</i>	S1		
79	XIAN3		<i>C. oleifera</i>	S1		
80	XIAN41		<i>C. oleifera</i>	S1		
81	XIAN46		<i>C. oleifera</i>	S1		
82	XIAN67		<i>C. oleifera</i>	S1		
83	XIAN72		<i>C. oleifera</i>	S1		
84	XIAN82		<i>C. oleifera</i>	S1		
85	XIAN87		<i>C. oleifera</i>	S1		
86	XIAN89		<i>C. oleifera</i>	S1		
87	XIAN93		<i>C. oleifera</i>	S1		
88	ZA12		<i>C. oleifera</i>	S1		
89	ZA16		<i>C. oleifera</i>	S1		
90	ZA1		<i>C. oleifera</i>	S1		
91	ZA22		<i>C. oleifera</i>	S1		
92	ZA24		<i>C. oleifera</i>	S1		
93	ZA27		<i>C. oleifera</i>	S1		
94	ZA31		<i>C. oleifera</i>	S1		
95	ZA3		<i>C. oleifera</i>	S1		
96	FS1		<i>C. vietnamensis</i>	S2	Fansai Village, Wuzhishan City,	109°32'24" E,
97	FS2		<i>C. vietnamensis</i>	S2	Hainan Province	18°50'37" N
98	HC1H		<i>C. vietnamensis</i>	S2	Fushan Town, Chengmai County,	109°54'55" E,
99	HC2H		<i>C. vietnamensis</i>	S2	Hainan Province	19°52'20" N
100	HD2H		<i>C. vietnamensis</i>	S2	Shangke Town, Qionghai City,	110°20'39" E,
101	HD4H		<i>C. vietnamensis</i>	S2	Hainan Province	19°04'20" N
102	HL1H		<i>C. vietnamensis</i>	S2	Qionghai Area, Haikou City,	110°21'54" E,
103	HL2H		<i>C. vietnamensis</i>	S2	Hainan Province	19°59'25" N
104	HS1		<i>C. vietnamensis</i>	S2		
105	HS3		<i>C. vietnamensis</i>	S2	Hongshan Village, Wuzhishan City,	109°30'56" E,
106	HS4		<i>C. vietnamensis</i>	S2	Hainan Province	18°51'35" N
107	RY1H		<i>C. vietnamensis</i>	S2	Wencheng Town, Wenchang City,	110°47'38" E,
108	RY2H		<i>C. vietnamensis</i>	S2	Hainan Province	19°33'13" N
109	WH1H		<i>C. vietnamensis</i>	S2		
110	WH2H		<i>C. vietnamensis</i>	S2	Wanling Town, Qiongzong County,	109°53'48" E,
111	WH3H		<i>C. vietnamensis</i>	S2	Hainan Province	19°08'35" N
112	WH4H		<i>C. vietnamensis</i>	S2		
113	HONG3		<i>C. chekiangoleosa</i>	S3		
114	HONG4		<i>C. chekiangoleosa</i>	S3	Wuzhishan City, Hainan Province	109°30'57" E, 18°46'29" N

Table 1. Cont.

No.	Name	Tissue	Species	Group	Origin	Location
115	CL		<i>C. oleifera</i>	S1		
116	CL-1		<i>C. oleifera</i>	S1		
117	CL-2		<i>C. oleifera</i>	S1		
118	DZ		<i>C. oleifera</i>	S1		
119	DZ-1		<i>C. oleifera</i>	S1		
120	DZ-2		<i>C. oleifera</i>	S1		
121	DZ-3		<i>C. oleifera</i>	S1		
122	DZ-4		<i>C. oleifera</i>	S1		
123	XL		<i>C. oleifera</i>	S1		
124	XL-1		<i>C. oleifera</i>	S1		
125	XL-2		<i>C. oleifera</i>	S1		
126	XL-3		<i>C. oleifera</i>	S1		
127	XL-4		<i>C. oleifera</i>	S1		
128	GN	Seeds	<i>C. chekiangoleosa</i>	S3	Xixiangtang Area, Nanning City, Guangxi Zhuang Autonomous Region	108°21'7" E, 22°55'6" N
129	GN-1		<i>C. chekiangoleosa</i>	S3		
130	GN-2		<i>C. chekiangoleosa</i>	S3		
131	BB		<i>C. crapnelliana</i>	S4		
132	BB-1		<i>C. crapnelliana</i>	S4		
133	BB-2		<i>C. crapnelliana</i>	S4		
134	BB-3		<i>C. crapnelliana</i>	S4		
135	BB-4		<i>C. crapnelliana</i>	S4		
136	LC		<i>C. gauchowensis</i>	S5		
137	LC-1		<i>C. gauchowensis</i>	S5		
138	LC-2		<i>C. gauchowensis</i>	S5		
139	LC-3		<i>C. gauchowensis</i>	S5		
140	LC-4		<i>C. gauchowensis</i>	S5		

## 2.2. DNA Extraction

Sample DNA was extracted by the TIANGEN genomic DNA extraction kit (Beijing, China). DNA quality and concentration were then checked by 1% (*w/v*) agarose gel electrophoresis and the Agilent 2100 Bioanalyzer (USA). Good quality DNA was used directly for SSR analysis or stored at  $-20\text{ }^{\circ}\text{C}$  for further use.

## 2.3. SSR Analysis

Ninety-six pairs of SSR primers were selected for pre-screening based on the transcriptome data of *C. vietnamensis* (NCBI accession number: PRJNA825399) [24], in which 15 fluorescently labeled SSR primers were selected for further research (Table 2). In addition, nine pairs of primers with good polymorphism were screened, referring to Song's study (Table 2) [25]. The 5' end of each forward primer for this analysis was labelled with FAM fluorescent dye (Applied Biosystems, USA). The M13 universal linker sequence (TGTAACAACGACGGCCAGT) was used to add to the 5' direction of the forward primer of each pair of primers, and M13 linker sequences with different fluorescent groups were synthesized. Following the method of Gu [26] with minor modification, the SSR-PCR amplification was performed in a 15  $\mu\text{L}$  total reaction volume, including 1.0  $\mu\text{L}$  ( $5\text{ pmol}\cdot\mu\text{L}^{-1}$ ) of forward and reverse primers, 7.5  $\mu\text{L}$  of  $2\times$  Taq PCR master mix (Gene tech, Shanghai, China), 1  $\mu\text{L}$  ( $50\text{ ng}\cdot\mu\text{L}^{-1}$ ) of template DNA, and 4.5  $\mu\text{L}$  of ddH<sub>2</sub>O. The PCR program was as follows:  $96\text{ }^{\circ}\text{C}$ , 3 min;  $96\text{ }^{\circ}\text{C}$  for 30 s,  $50\text{--}60\text{ }^{\circ}\text{C}$  for 30 s, and  $72\text{ }^{\circ}\text{C}$  for 1 min, and these three procedures were cycled 30 times;  $72\text{ }^{\circ}\text{C}$ , 10 min. Two microliters of amplified PCR products were used in 2% (*w/v*) agarose gel electrophoresis to check whether the amplified fragment size and concentration were in the normal ranges at each locus with reference to the DNA marker alignment. Then, 1.0  $\mu\text{L}$  of the fluorescent PCR product was diluted 30 fold with ultrapure water and prepared for machine detection. The diluted PCR products were separated by capillary electrophoresis by the ABI 3730XL DNA Analyzer (Applied

Biosystems, Foster City, CA, USA), and data were handled by Gene Marker v.2.2.0 software (Soft Genetics, State College, PA, USA).

**Table 2.** Detailed information for 24 pairs of primers in the study.

No.	Locus	Repeat Unit	Forward Sequence	Reverse Sequence	Pre Experiment Size (bp)	Fluorescent Dyes
1	CoA007	(TCT)6	CCAATCTCCAAACGCAACTT	CAGAGGAAATCGAGAGGCAG	245	FAM
2	CoA008	(ATAG)6	CCAGCCAGCTAAGAGGTTG	CAGGTCATAGCTACCACGGA	188	FAM
3	CoA011	(CTT)5	TGGGTGGCTCAATATCATCA	ACCGGCCATTTATATGGGTT	200	FAM
4	CoA016	(ATC)6	GTAAGTCTCTGCACCCGCTC	TCGATTTCGTCCAATCCTTC	211	FAM
5	CoA020	(AGG)6	AGGGCATAAGAGGGAGTGGT	CGACCTCGACCTTCAAGAAC	207	FAM
6	CoA022	(GA)12	TAGCCAATAACATGCCACACA	AGTTGTCCAACCCCTTCCTCA	147	FAM
7	CoA032	(GCG)5	TTATTCTTCGGGAACAACGG	ACACATGAAACAACGGCAAA	170	FAM
8	CoA038	(GTG)7	GAGATCGGCCAGAGTTTGAG	CATCAAAGCCACACTCGCTA	202	FAM
9	CoA039	(TTA)6	GCAAGAGGTCTCTTTGGGTG	AACCTCATGAGCTAAAGCCG	113	FAM
10	CoA045	(ACC)5	TCCAAACAGGCCAACTAAGC	GCTTGAGAAACCCAAAGCAG	244	FAM
11	CoA046	(TAAC)4	AACCAGAGGAACATCCAACG	TATCCTTGCCGCTTTGAATC	196	FAM
12	CoA055	(CAT)6	TCTGGTGTGCTTCAAACCTGC	GTCCAGCAAATATTCAGGC	265	FAM
13	CoA069	(TGC)6	CATGGCTTGGCTTCAATCTT	CAATGTTCCCAAGCGATTCT	224	FAM
14	CoA081	(CAA)5	ATATGAATCGGCCAATCGAC	AGATGACGCCTTTCGAAGAA	154	FAM
15	CoA084	(GTG)6	GACGGCTTAAACATGGAGGA	TTCATTTAATGGCAGGAGGC	110	FAM
16	SJMCoa003	(CAA)7	ACGAAACATGTCCGACGTGA	GGAATGGACGAGACTTGGG	120	FAM
17	SJMCoa007	(TTC)6	GCAGCAGCGAGAGTAACAGT	GTGGGACGATTGAGCTTCTT	149	FAM
18	SJMCoa030	(CCT)10	GGTGTGGTGTGAAGCAGTA	TTGTCTGGATCCATAGCCGC	248	FAM
19	SJMCoa038	(TTAT)5	TGCTTGGTCACTACCCAGTC	TGACACCTTGGTGCCAAAGA	266	FAM
20	SJMCoa045	(AAT)5	TTTGGGCGGGCAAAGATTG	ACTCAAGCATGGACATCGGG	276	FAM
21	SJMCoa049	(AAT)5	AAGACCCAAACTGGACTGCA	ACCTTGCACCATAATGGGTT	254	FAM
22	SJMCoa050	(AAT)7	TGGAGCGTTAGTCTGGAGTC	GGCCTTCATCCATGTCAGG	249	FAM
23	SJMCoa058	(CCA)9	GTGCCCTGTGACACCAAGTA	CGACGGTGGAGATTGGTGA	245	FAM
24	SJMCoa090	(TCA)9	ACAGAAGGCGTTGAGTCAA	GGCTTCTTCTCGGAACCCA	165	FAM

#### 2.4. Data Acquisition and Analysis

According to the PCR results, a binary matrix was formed in which the presence of the product was marked as 1 and the absence of the product as 0. The results of the 1/0 data matrix were utilized to analyze the genetic diversity of oil-tea camellia trees. Based on the number of alleles, the level of discrimination of each SSR marker was assessed by calculating the percentage of polymorphic loci (P), Nei's genetic diversity (h), Shannon diversity index (I) [27], gene differentiation coefficient (Gst) [26], and gene flow from Gst (Nm). Nei's genetic diversity (h) and Shannon diversity index (I) were calculated using the POPGENE software [28].

According to the DICE coefficient [29], Nei's genetic distance (D) and genetic identity between different groups were further calculated using GenAlex software [30]. The degrees of genetic variation among and within groups were analyzed by the non-hierarchical analysis of molecular variance (AMOVA) method, with 9999 random permutations [28]. Then, the unweighted pair group method with arithmetic (UPGMA) and the principal coordinates analysis (PCoA) were performed [31]. PCoA analysis was performed with GenAlex software. Linkage disequilibrium was analyzed using the pair.ia method of the R package poppr, and plots were drawn in R. In addition, the genetic structure of oil-tea camellia samples was analyzed by STRUCTURE [32], which is a model-based Bayesian clustering program with a range of genetic clusters from K = 3 to 10. Twenty independent runs were evaluated for each fixed K, and the best potential clusters (K value) were checked by the  $\Delta K$  method on the STRUCTURE Harvester program [32]. The running results were integrated by CLUMPP software [33].

### 3. Results

#### 3.1. Assessment of SSR Marker Diversity Levels

The 140 accessions from five oil-tea camellia species were analyzed by SSR markers. The alleles detected by 24 pairs of primers at the polymorphic sites ranged from 6 to 31. A total of 385 alleles were generated by amplification, resulting in an average of 16.0417 alleles per locus (Table 3). The mean of Nei's gene diversity (h) and Shannon's information index (I) were 0.1104 and 0.1890, which indicate that the genetic diversity was

not very rich. It can be seen in Table 3 and Table S1 that the range of total genetic variation  $H_t$  was 0.0019–0.5000; the average value was 0.1153. The range of genetic variation within population  $H_s$  was 0–0.4601; the average value was 0.0698. The gene differentiation coefficient  $G_{st}$  value ranged from 0.0037 to 1.0000, and the average was 0.3948, indicating 39.48% genetic variation among individuals and a high degree of genetic differentiation. The range of gene flow ( $N_m$ ) values of the whole population was 0–134.0618, and the average value was 0.7666, indicating that there was little gene exchange among the oil-tea camellia group.

**Table 3.** Genetic parameters of the SSR locus analysis.

Locus	Product Size (bp)	Number of Alleles	$N_e$	$h$	$I$	$G_{st}$	$N_m$
CoA007	176–257	16	1.0072–1.6880	0.0071–0.4076	0.0237–0.5976	0.0037–0.7133	0.2010–134.0618
CoA008	138–230	26	1.0072–1.7521	0.0071–0.4293	0.0237–0.6206	0.0037–0.7981	0.0029–134.0618
CoA011	167–206	12	1.0072–1.5542	0.0071–0.3566	0.0237–0.5393	0.0037–1.0000	0–134.0618
CoA016	208–367	17	1.0072–1.7183	0.0071–0.3833	0.0237–0.6088	0.0037–0.7853	0.1367–134.0618
CoA020	164–256	21	1.0072–2.0000	0.0071–0.5000	0.0237–0.6931	0.0037–0.6424	0.2783–134.0618
CoA022	127–184	31	1.0072–1.9468	0.0071–0.4863	0.0237–0.6794	0.0037–0.7032	0.2110–134.0618
CoA032	130–204	22	1.0072–1.9872	0.0071–0.4968	0.0237–0.6899	0.0037–0.6555	0.2628–134.0618
CoA038	192–221	13	1.0072–1.9619	0.0071–0.4903	0.0237–0.6834	0.0037–0.6220	0.3038–134.0618
CoA039	103–126	13	1.0073–1.5290	0.0072–0.3460	0.0240–0.4973	0.0162–0.6118	0.3172–30.4262
CoA045	237–288	14	1.0072–1.9993	0.0071–0.4998	0.0237–0.6930	0.0037–0.7650	0.1536–134.0618
CoA046	169–208	17	1.0072–1.9983	0.0071–0.4996	0.0237–0.6927	0.0037–0.5994	0.3126–134.0618
CoA055	151–313	22	1.0073–1.2904	0.0072–0.2250	0.0240–0.3849	0.0112–1.0000	0–44.0603
CoA069	211–266	18	1.0072–1.9155	0.0071–0.4779	0.0237–0.6709	0.0037–0.8486	0.1211–134.0618
CoA081	150–184	7	1.0073–1.9835	0.0072–0.4959	0.0240–0.6890	0.0112–0.5857	0.3537–44.0603
CoA084	106–119	6	1.0072–1.6058	0.0071–0.3773	0.0237–0.5648	0.0037–0.4816	0.5381–134.0618
SJMCoa003	126–167	12	1.0072–1.9971	0.0071–0.4993	0.0237–0.6924	0.0037–0.6596	0.2581–134.0618
SJMCoa007	224–311	23	1.0072–1.9989	0.0071–0.4997	0.0237–0.6929	0.0037–0.4972	1.0755–134.0618
SJMCoa030	238–277	14	1.0072–1.7639	0.0071–0.4331	0.0237–0.6246	0.0037–1	0–134.0618
SJMCoa038	273–304	18	1.0072–1.8695	0.0071–0.4651	0.0237–0.6578	0.0037–0.2993	1.1705–134.0618
SJMCoa045	291–317	19	1.0073–1.9215	0.0072–0.4796	0.0240–0.6726	0.0075–0.3244	1.0415–66.5610
SJMCoa049	270–286	7	1.0072–1.9989	0.0071–0.4997	0.0237–0.6929	0.0037–0.3559	0.9048–134.0618
SJMCoa050	253–273	10	1.0072–1.3412	0.0071–0.2544	0.0237–0.4219	0.0037–1.0000	0–134.0618
SJMCoa058	197–266	16	1.0072–1.5438	0.0071–0.3522	0.0237–0.5371	0.0037–0.2238	1.7344–134.0618
SJMCoa090	173–203	11	1.0218–1.9829	0.0213–0.4957	0.0596–0.6888	0.0112–0.5934	0.3426–44.0603
Mean		16.0417	1.1676	0.1104	0.1890	0.3948	0.7666

Note:  $N_e$ , Number of Effective Alleles;  $h$ , Nei's gene diversity;  $I$ , Shannon's Information Index;  $G_{st}$ , Gene differentiation coefficient;  $N_m$ , estimate of gene flow from  $G_{st}$ .

The alleles at each locus in each sample were coded into a fingerprint in the form of a 0/1 matrix based on bands amplified using 24 pairs of primers. Fingerprinting gives a visual representation of the differences for each sample (Figure 1). As could be found from the fingerprinting of 140 oil-tea camellia accessions, these 24 pairs of primers could discriminate some of the 140 accessions.

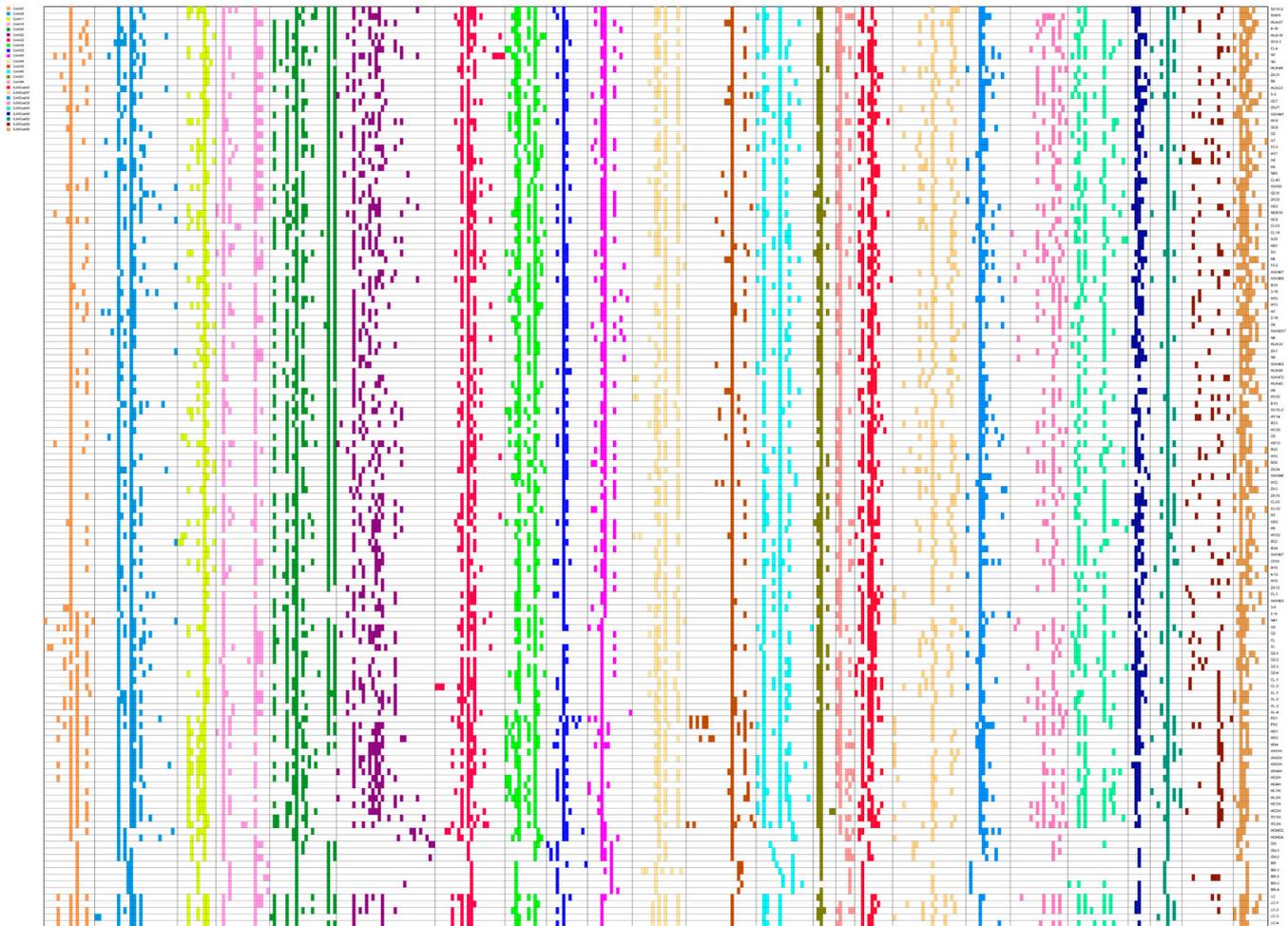
### 3.2. Genetic Diversity of Oil-Tea Camellia Species Based on SSR Analysis

The detailed information of each genetic locus of each species is shown in Table S2. The average sample size was 28 for each species (Table 4). The mean  $N_a$  was 0.735 (range: 0.200–1.605). The average  $N_e$  was 1.138 (range: 1.041–1.197). The mean  $h$  was 0.086 (range: 0.027–0.128), and the mean  $u_h$  was 0.096. The average  $I_s$  within species reached 0.134. S1 had the highest genetic variability (0.214), and S4 had the lowest value (0.041). When computed at the individual level, the mean  $I$  was 0.1890. The results indicate that the genetic differences among different groups were small and the genetic diversity was not very rich.

**Table 4.** The population average diversity index.

Group	N	$N_a$	$N_e$	$I_s$	$h$	$u_h$	P (%)
S1	108	1.605	1.197	0.214	0.128	0.130	79.48
S2	17	0.969	1.195	0.193	0.121	0.122	45.97
S3	5	0.434	1.140	0.119	0.081	0.101	20.52
S4	5	0.200	1.041	0.041	0.027	0.033	7.79
S5	5	0.468	1.116	0.104	0.070	0.087	18.70
Mean	28	0.735	1.138	0.134	0.086	0.096	34.49

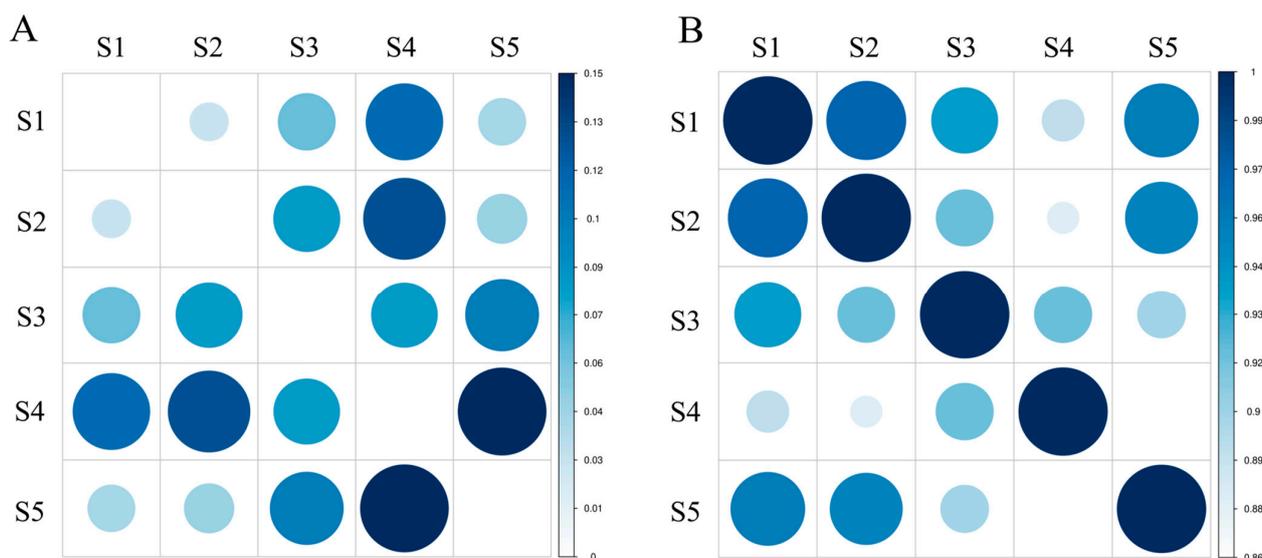
Note: N, Sample size;  $N_a$ , Number of different alleles;  $N_e$ , Number of effective alleles;  $I_p$ , intra-specie diversity;  $h$ , Nei's gene diversity;  $u_h$ , Unbiased diversity; P, Percentage of Polymorphic Loci.



**Figure 1.** The fingerprinting of each allele in 140 samples.

### 3.3. Analysis of Nei's Genetic Distance between Species

Nei's genetic distance ( $D$ ) is a measure of genetic difference among biological populations and can be measured in terms of quality traits and also with quantitative traits. The estimation of genetic distance is important for exploring the origins of cultivars, analyzing the relationships among populations, mapping phylogenetic trees and predicting heterosis, and guiding parental selection. The range of genetic identity among species was 0.8616–0.9719, calculated from 285 amplified fragments. As shown in Figure 2, S1 and S2 had the smallest genetic distance (0.0285) and the largest genetic identity (0.9719) with the closest relatives, followed by S1 and S5. S5 and S4 had the largest genetic distance (0.1490), shared the least genetic identity (0.8616), and were the most distantly related, followed by S4 and S2. The results of AMOVA indicated that most of the genetic variation (77%) occurred within species and only a small fraction (23%) occurred among species (Table 5). In addition, there were significant differences within and among groups. The mean fixation index ( $F_{st}$ ) among five groups showed moderate genetic differentiation ( $F_{st} = 0.231$ ).



**Figure 2.** Genetic distance and genetic identity among groups. (A): Nei's genetic distance; (B): Nei's genetic identity.

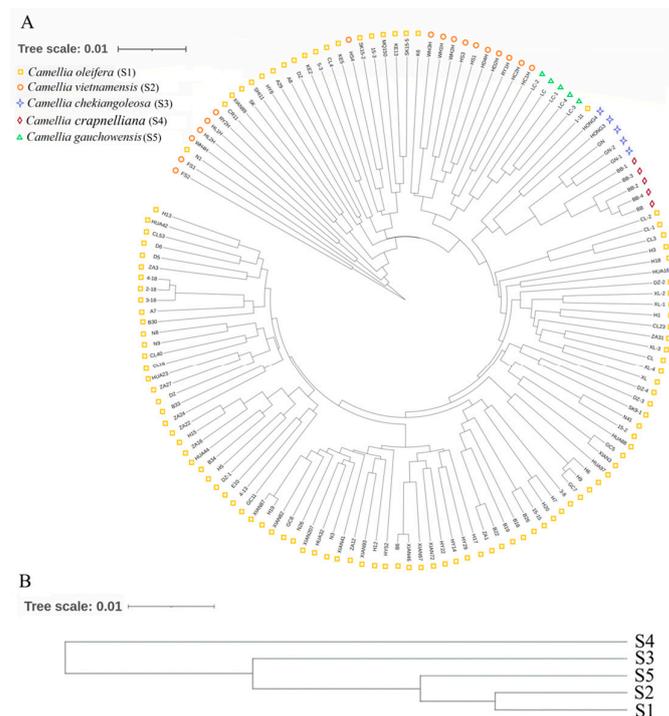
**Table 5.** An analysis of molecular variance among and within *Camellia* species.

Variation Source	df	SS	MS	Est. Var.	PMV (%)	$F_{st}$	$p$ value
Among Pops	4	485.260	121.315	7.197	23	0.231	0.001
Within Pops	135	3240.862	24.006	24.006	77		
Total	139	3726.121		31.203	100		

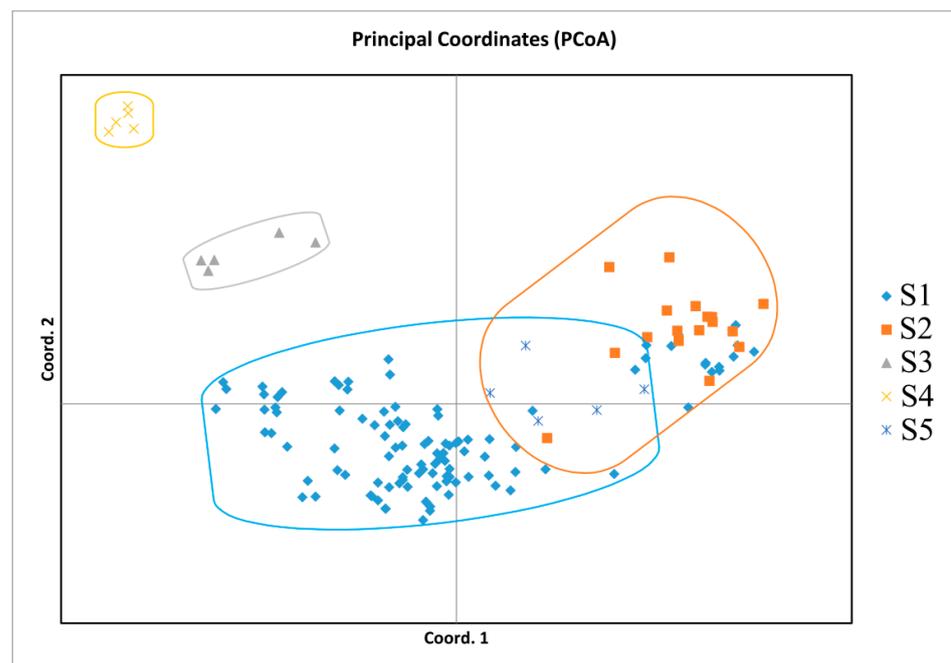
Note: df, degree of freedom; SS, Square deviation; MS, Mean square deviation; Est. Var., Exist variance; PMV, Percentages of molecular variance;  $F_{st}$ , coefficient of genetic differentiation.  $p$  value indicated significant differences of  $p \leq 0.001$ .

### 3.4. UPGMA and PCoA Analysis

Based on Nei's genetic distances among individuals and groups, the clustering analysis among individuals was accomplished using the boot method of the R package poppr, by selecting Nei's distance and bootstrapping 1000 times. Cluster analysis among populations was subjected to UPGMA trees drawn using the phylip software. According to the genetic distances, a phylogenetic tree was built (Figure 3). As can be seen in Figure 3A, most individuals from S1 grouped together; S2, S5, and a small part of S1 were clustered together; individuals of S3 and S4 grouped together. The phylogenetic tree obtained with Nei's genetic distance classified the species into three main clades (Figure 3B). The first clades included S1, S2, and S5; the other two were S3 and S4. Among them, *C. oleifera* was clustered into one subclade, *C. vietnamensis* and *C. crapnelliana* were clustered into one subclade, and *C. chekiangoleosa* and *C. gauchowensis* were clustered into one subclade. In addition, some *C. oleifera* and *C. vietnamensis* were clustered into one subclade. Furthermore, two-dimensional PCoA revealed four distinct clusters on the basis of Nei's genetic distance among individuals (Figure 4). PCoA analysis reflects the variability between two samples or two groups by an intuitive comparison of the straight-line distances between samples in the coordinate axis, which indicates whether the two samples or two groups of samples are notably divergent. PCoA of the first three axes explained 17.61% of the total variation (7.30%, 6.68%, and 3.63%, respectively). The results of PCoA were relatively similar to the individual-based phylogenetic tree. S1 (*C. oleifera*) samples were clustered together, S2 (*C. vietnamensis*) and S5 (*C. gauchowensis*) samples were clustered together, S3 (*C. chekiangoleosa*) samples were clustered together, and S4 (*C. crapnelliana*) samples were clustered together.



**Figure 3.** The UPGMA phylogenetic tree of oil-tea camellia samples based on SSR data. (A): The phylogenetic tree of 140 samples; (B): The phylogenetic tree of five *Camellia* species.

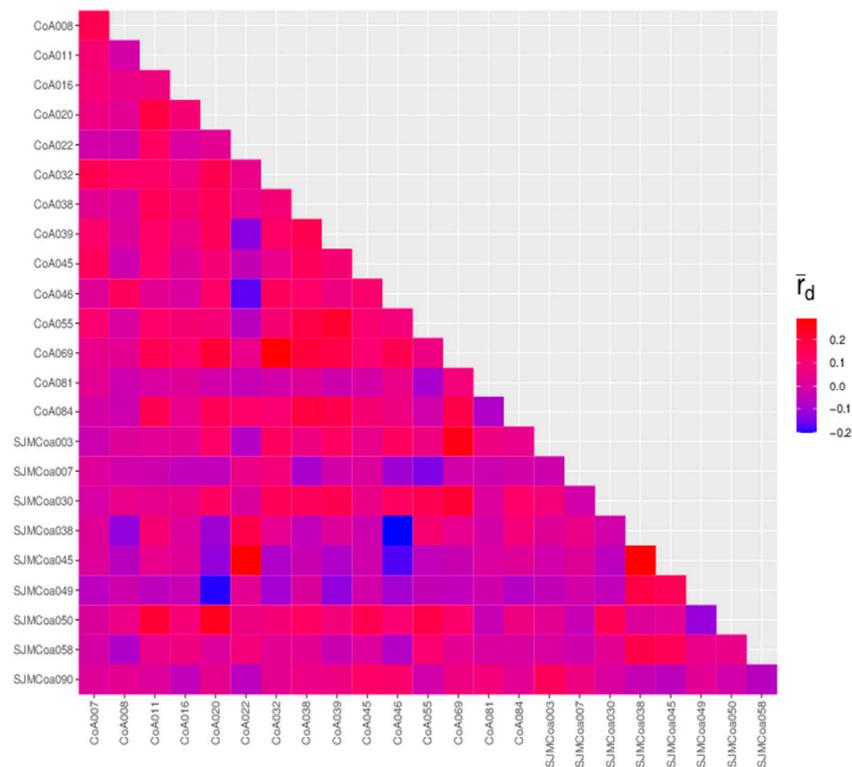


**Figure 4.** PCoA of oil-tea camellia samples based on SSR data.

### 3.5. Linkage Disequilibrium Analysis and Population Structure

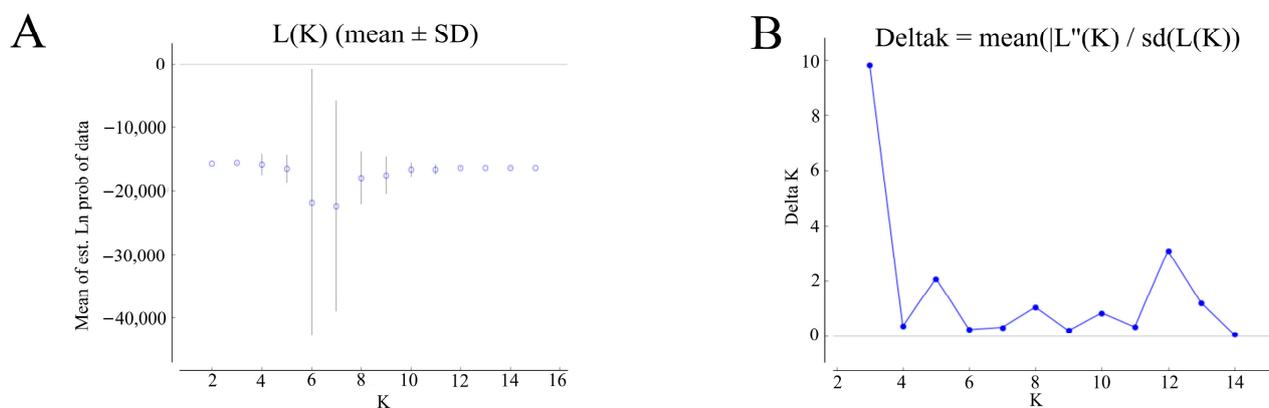
In linkage disequilibrium, there is a shift between the probability that a haplotype will appear and the probability that it will be randomly combined. The extent of this offset determines the extent of linkage disequilibrium. The degree of linkage disequilibrium was characterized by the square of the  $R$  value, which, when equal to 0, indicates complete linkage equilibrium—independent inheritance. When the  $R$ -squared equals 1, it indicates complete

linkage disequilibrium. All 24 SSR loci were in linkage disequilibrium with each other; the a maximum R-squared was 1, and a minimum R-squared was 0.0099 (Figure 5 and Table S3).

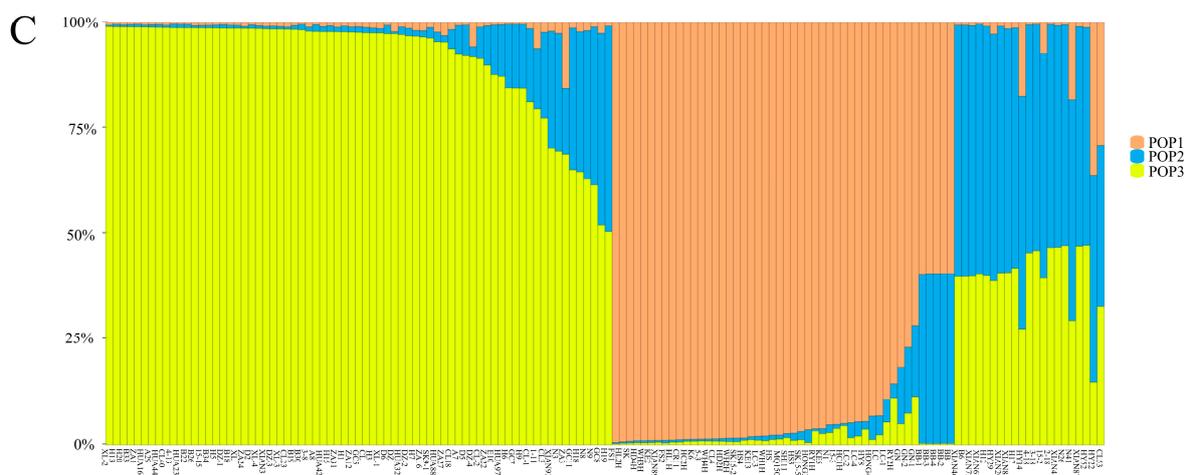


**Figure 5.** The linkage disequilibrium of 24 SSR loci.

The results of STRUCTURE showed a clear maximum for Ln(PD)-derived delta K ( $\Delta K$ ) at  $K = 3$  (Figure 6A,B), and this was considered as a possible number for the population of oil-tea camellia. Therefore, it indicated that the studied accessions belonged to three different clusters (Figure 6C). Among them, most of *C. oleifera* samples were clustered in one population; a small proportion of *C. oleifera* samples were clustered separately; *C. vietnamensis*, *C. gauchowensis*, *C. crapnelliana*, and *C. chekiangoleosa* were another cluster. The results of population structure analysis were similar to those of UPGMA analysis (Figure 3A).



**Figure 6.** Cont.



**Figure 6.** A structure analysis of 140 oil-tea camellia samples. (A): Estimated  $\ln P(D)$  of  $K$  from 2 to 16. (B):  $\Delta K$  according to the rate of change of  $\ln P(D)$  between successive  $K$ . (C): Genetic structure of oil-tea camellia population.

#### 4. Discussion

In this research, the genetic diversity of five oil-tea camellia species was analyzed by using SSR markers. The range of alleles in SSR was 6–31. A total of 385 alleles were found. An average of 16.0417 alleles were found for each SSR-primer pair. At the species level, the range of  $I_p$  of oil-tea camellia was 0.041–0.214, and the range of  $p$  was 7.79–79.48% (the mean value was 34.49%), showing moderate genetic diversity. When the polymorphism information content (PIC) was less than 0.25, SSR primers showed little polymorphism; when  $0.25 < \text{PIC} < 0.5$ , moderate polymorphism; when  $\text{PIC} > 0.5$ , high polymorphism [34]. The results from the amplification of 345 pairs of SSR primers by Shi et al. [16] indicated that the proportion of polymorphic sites (31.9%) was relatively high. Chai et al. analyzed six natural populations of *C. pubipetala*, and the results showed that the  $I$  value was 0.4100; the PPB (percentage of polymorphic bands) was 80.43% [35]. Although the six populations' distribution was narrow, the genetic diversity was high. A total of 495 alleles were identified by 111 SSR loci in *C. japonica*, and the range of alleles was 1–12. The mean was 4.46 alleles per locus. The range of PIC was 0.15–0.86, and the average was 0.59 [15]. The mean of  $p$  in this study was 34.49%, which is similar to the above results. The ranges of  $N_e$ ,  $h$ , and  $I_p$  of 24 markers in this study were 1.041–1.197, 0.027–0.128, and 0.041–0.214, respectively. The differences are larger when compared with the results of Dong et al., who used 16 SSR marker pairs for 54 oil-tea trees (including *C. polyodonta*, *C. oleifera*, *C. gauchowensis*, and *C. semiserrata*) for genetic diversity analysis. The ranges of  $N_e$ ,  $h$ , and  $I$  were 1.17–1.70, 0.14–0.40, and 0.26–0.59, respectively [3]. In conclusion, the SSR primers in this study showed moderate to high levels of polymorphism, which indicates that they were suitable for genetic diversity analysis of oil-tea camellia trees.

Accurate genetic relationships among germplasm accessions are important for variety development, evolutionary studies, and resource conservation [31,36]. Three main clusters were determined by the UPGMA method on 140 samples. *C. vietnamensis* was clustered with *C. gauchowensis*, which is similar to the findings of Qi et al. [37] and Chen et al. [1]. They found that various indexes of leaf, flower, fruit, and seed morphologies of *C. vietnamensis* collected from Hainan Province showed high similarity to those of *C. gauchowensis*, whose provenance was Gaozhou in Guangdong Province [37], and *C. vietnamensis* and *C. gauchowensis* were found to be clustered together by cpDNA sequences and SSR marker analysis [1], so it could be speculated that the relative proximity of *C. vietnamensis* and *C. gauchowensis* to each other was quite near. Dai et al. analyzed the chloroplast genome *trnH-psbA* and *matK* sequences of 101 different kinds of oil-tea camellia seedlings by DNA barcoding technology [38]. They found that *C. vietnamensis* was clustered into one branch and *C. chekiangoleosa* was clustered into another, and the

clustering results in this study agree strongly with these results. The findings suggest that *C. vietnamensis* in Hainan has a relatively close relative. Additionally, the phenomenon of self-incompatibility might occur in close relatives, which might be one of the reasons for the low seed-setting rate of *C. vietnamensis* in Hainan.

For a more accurate analysis of the genetic structure of oil-tea camellia, STRUCTURE was used for further analysis, and the results indicated that the 140 accessions were classified into three clusters. Among them, most of *C. oleifera* samples were clustered one population; a small proportion of *C. oleifera* were in another cluster; and *C. vietnamensis*, *C. gauchowensis*, *C. crapnelliana*, and *C. chekiangoleosa* were one cluster. There were small fractions of *C. oleifera* samples that clustered with other *Camellia* species. It was indicated that plants of the same group came not only from the same region but also from different regions. Possibly, species with different genetic backgrounds may cluster together. This indicates that the kinship of germplasm is extremely complex. The reasons for this phenomenon might be as follows. First, occasional genetic mutations and long-term natural selection have made finding the relatives of oil-tea camellia more complicated. Second, the effect of genetic drift was greater during the natural differentiation of oil-tea camellia than those of natural environmental factors, leading to the failure to divide by geographic region when clustering. Using indirect measures, the gene flow among populations was estimated by the value of  $Nm$  [28]. The  $Nm$  value (0.7666) indicated low gene flow among species and might promote population differentiation. When  $Nm < 1$ , genetic drift is thought to be a major contributor to population differentiation [26,28]. Third, the uneven number of selected samples makes the clustering result not accurate enough, and so on.

The results of genetic structure analysis indicate that the genetic variation of oil-tea camellia samples mainly appeared within species, accounting for 77% of the total variation, leaving only a small portion (23%) occurring among species. That might result from habitat fragmentation and geographical barriers. Some experts have also obtained similar results with other *Camellia* plants. He et al. used nine pairs of SSR primers to analyze 150 accessions of *C. oleifera*, and the results indicated that the genetic diversity level of *C. oleifera* is high [17]. In addition, Li et al. also analyzed 84 accessions of eight natural populations of *C. fascicularis* with fourteen pairs of primers for SSR markers [18]. The results indicated that the eight populations of *C. fascicularis* were roughly divided into three clusters, and the genetic variation within populations accounted for 49.95% of variation. To sum up, the results of this study indicate that the genetic variation of oil-tea camellia samples was mainly found within populations, and inbreeding occurred within the population, such as with *C. vietnamensis*. The degree of gene exchange among species was low.

## 5. Conclusions

In this research, 24 pairs of SSR primers were selected to analyze the genetic relationship and population structure of 140 oil-tea camellia accessions using fluorescence detection by capillary electrophoresis. The results indicate that genetic diversity was abundant among the 140 *Camellia* accessions. Based on genetic distances and clustering by UPGMA, the 140 accessions could be classified into three clusters. Most individuals from S1 grouped together, samples from S2 and S5 grouped together, and samples from S3 and S4 formed the same branch. In addition, some individuals from S1 and S2 were clustered together, which relates to the results of the PCoA. The Bayesian-model-based genetic structure analysis indicated that the studied accessions belonged to three populations. Among them, most of the *C. oleifera* samples were clustered into one population; a small proportion of *C. oleifera* were in another cluster; *C. vietnamensis*, *C. gauchowensis*, *C. crapnelliana*, and *C. chekiangoleosa* were one cluster. Taken together, the findings should be instructive for oil-tea camellia species' introduction, breeding, germplasm preservation, and new-variety development, and provide a theoretical foundation for the classification and identification of oil-tea camellia species in southern China and the research on relatives.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13112162/s1>, Table S1: Genetic diversity parameters of each SSR locus; Table S2: Genetic diversity parameters of each SSR locus of each specie; Table S3: Linkage disequilibrium of 24 SSR loci.

**Author Contributions:** Conceptualization, J.Y. and Y.L.; methodology, Y.W. (Yongen Wu); software, H.Q.; validation, Y.W. (Yong Wang), and J.Y.; formal analysis, Y.L.; investigation, H.Y.; resources, J.C.; data curation, H.Y.; writing—original draft preparation, H.Y.; writing—review and editing, H.Q.; visualization, Y.L.; supervision, J.C.; project administration, Y.W. (Yong Wang); funding acquisition, J.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the Open Project of Ministry of Education Key Laboratory for Ecology of Tropical Islands, Hainan Normal University, China (No. HNSF-OP-2021-2), the High-level Talents Project of Hainan Natural Science Foundation (2019RC173), the Key R&D Program of Hainan Province, China (ZDYF2022SHFZ020), and the High-level talent project of Hainan Natural Science Foundation (No. 820RC585).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank Jiaming Song (Hainan Academy of Forestry) for providing nine primer pairs.

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

## References

1. Chen, J.; Guo, Y.J.; Hu, X.W.; Zhou, K.B. Comparison of the chloroplast genome sequences of 13 oil-tea *Camellia* samples and identification of an undetermined oil-tea *Camellia* species from Hainan province. *Front. Plant Sci.* **2022**, *12*, 798581. [[CrossRef](#)] [[PubMed](#)]
2. Jia, B.G.; Lin, Q.; Feng, Y.Z.; Hu, X.Y.; Tan, X.F.; Shao, F.G.; Zhang, L. Development and cross-species transferability of unigene-derived microsatellite markers in an edible oil woody plant, *Camellia oleifera* (Theaceae). *Genet. Mol. Res.* **2015**, *14*, 6906–6916. [[CrossRef](#)] [[PubMed](#)]
3. Dong, B.; Deng, Z.; Liu, W.; Rehman, F.; Yang, T.J.; Huang, Y.F.; Gong, H.G. Development of expressed sequence tag simple sequence repeat (EST-SSR) markers and genetic resource analysis of tea oil plants (*Camellia* spp.). *Conserv. Genet. Resour.* **2022**, *14*, 41–45. [[CrossRef](#)]
4. Feng, J.L.; Yang, Z.J.; Chen, S.P.; El-Kassaby, Y.A.; Chen, H. High throughput sequencing of small RNAs reveals dynamic microRNAs expression of lipid metabolism during *Camellia oleifera* and *C. meiocarpa* seed natural drying. *BMC Genom.* **2017**, *18*, 546. [[CrossRef](#)] [[PubMed](#)]
5. Yu, J.; Yan, H.; Wu, Y.; Wang, Y.; Xia, P. Quality Evaluation of the Oil of *Camellia* spp. *Foods* **2022**, *11*, 2221. [[CrossRef](#)]
6. Yang, C.Y.; Liu, X.M.; Chen, Z.Y.; Lin, Y.S.; Wang, S.Y. Comparison of Oil Content and Fatty Acid Profile of Ten New *Camellia oleifera* Cultivars. *J. Lipids* **2016**, *2016*, 3982486. [[CrossRef](#)]
7. Ye, Z.C.; Wu, Y.G.; Ul Haq Muhammad, Z.; Yan, W.P.; Yu, J.; Zhang, J.F.; Yao, G.L.; Hu, X.W. Complementary transcriptome and proteome profiling in the mature seeds of *Camellia oleifera* from Hainan Island. *PLoS ONE* **2020**, *15*, e0226888. [[CrossRef](#)]
8. Dai, J.N.; Zheng, W.; Yu, J.; Yan, H.Q.; Wang, Y.; Wu, Y.G.; Hu, X.W.; Lai, H.G. cDNA cloning, prokaryotic expression, and functional analysis of squalene synthase (SQS) in *Camellia vietnamensis* Huang. *Protein Expres. Purif.* **2022**, *194*, 106078. [[CrossRef](#)]
9. Zhu, G.F.; Liu, H.; Xie, Y.C.; Liao, Q.; Lin, Y.W.; Liu, Y.H.; Liu, Y.H.; Xiao, H.W.; Cao, Z.J.; Hu, S.Z. Postharvest Processing and Storage Methods for *Camellia oleifera* Seeds. *Food Rev. Int.* **2020**, *36*, 319–339. [[CrossRef](#)]
10. Zhang, W.; Zhao, Y.; Yang, G.; Peng, J.; Chen, S.; Xu, Z. Determination of the evolutionary pressure on *Camellia oleifera* on Hainan Island using the complete chloroplast genome sequence. *PeerJ* **2019**, *7*, e7210. [[CrossRef](#)]
11. Ma, J.Q.; Ma, C.L.; Yao, M.Z.; Jin, J.Q.; Wang, Z.L.; Wang, X.C.; Chen, L. Microsatellite markers from tea plant expressed sequence tags (ESTs) and their applicability for cross-species/genera amplification and genetic mapping. *Sci. Hortic. Amst.* **2012**, *134*, 167–175. [[CrossRef](#)]
12. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2010**, *177*, 309–334. [[CrossRef](#)]
13. Hao, L.; Zhang, G.; Lu, D.; Hu, J.; Jia, H. Analysis of the genetic diversity and population structure of *Salix psammophila* based on phenotypic traits and simple sequence repeat markers. *PeerJ* **2019**, *7*, e6419. [[CrossRef](#)]
14. Taniguchi, F.; Fukuoka, H.; Tanaka, J. Expressed sequence tags from organ-specific cDNA libraries of tea (*Camellia sinensis*) and polymorphisms and transferability of EST-SSRs across *Camellia* species. *Breed. Sci.* **2012**, *62*, 186–195. [[CrossRef](#)]

15. Li, Q.Y.; Su, X.J.; Ma, H.H.; Du, K.B.; Yang, M.; Chen, B.L.; Fu, S.; Fu, T.J.; Ciang, C.L.; Zhao, Q.; et al. Development of genic SSR marker resources from RNA-seq data in *Camellia japonica* and their application in the genus *Camellia*. *Sci. Rep.* **2021**, *11*, 9919. [CrossRef]
16. Shi, J.; Dai, X.G.; Chen, Y.N.; Chen, J.H.; Shi, J.S.; Yin, T.M. Discovery and experimental analysis of microsatellites in an oil woody plant *Camellia chekiangoleosa*. *Plant Syst. Evol.* **2013**, *299*, 1387–1393. [CrossRef]
17. He, Z.L.; Liu, C.X.; Wang, X.N.; Wang, R.; Chen, Y.Z.; Tian, Y. Assessment of genetic diversity in *Camellia oleifera* Abel. accessions using morphological traits and simple sequence repeat (SSR) markers. *Breed. Sci.* **2020**, *70*, 586–593. [CrossRef]
18. Li, B.; Liu, C.; Tang, J.R.; Xin, Y.X.; Dong, Z.H.; Bai, B.; Xin, P.Y. Genetic diversity analysis of *Camellia fascicularis* H. T. Chang based on SSR markers. *J. Appl. Res. Med. Aroma.* **2022**, *31*, 100404. [CrossRef]
19. Huang, Y. Population genetic structure and interspecific introgressive hybridization between *Camellia meiocarpa* and *C. oleifera*. *Chin. J. Appl. Ecol.* **2013**, *24*, 2345–2352. [CrossRef]
20. Tang, S.; Bin, X.; Wang, L.; Zhong, Y. Genetic Diversity and Population Structure of Yellow Camellia (*Camellia nitidissima*) in China as Revealed by RAPD and AFLP Markers. *Biochem. Genet.* **2006**, *44*, 444. [CrossRef]
21. Huseynov, M.; Suleymanova, Z.; Ojaghi, J.; Mammadov, A. Characterization and Phylogeny Analysis of Azerbaijan Tea (*Camellia sinensis* L.) Genotypes by Molecular Markers. *Cytol. Genet.* **2022**, *56*, 285–291. [CrossRef]
22. Zhou, L.Y.; Wang, X.N.; Wang, L.P.; Chen, Y.Z.; Jiang, X.C. Genetic diversity of oil-tea camellia germplasms revealed by ISSR analysis. *Int. J. Biomath.* **2015**, *08*, 1550070. [CrossRef]
23. Xiao, P.; Liu, H.; Wang, D.; Tang, W.; Yang, H.; Wang, C.; He, Z.; Wang, R.; Wang, X.; Lu, X.; et al. Assessment of genetic diversity in *Camellia oleifera* Abel. accessions using inter-simple sequence repeat (ISSR) and start codon targeted (SCoT) polymorphic markers. *Genet. Resour. Crop Evol.* **2020**, *67*, 1115–1124. [CrossRef]
24. Yan, H.; Zheng, W.; Wang, Y.; Wu, Y.; Yu, J.; Xia, P. Integrative Metabolome and Transcriptome Analysis Reveals the Regulatory Network of Flavonoid Biosynthesis in Response to MeJA in *Camellia vietnamensis* Huang. *Int. J. Mol. Sci.* **2022**, *23*, 9370. [CrossRef] [PubMed]
25. Song, J.M.; Li, X.Y.; Zhang, S.H.; Lai, H.G.; Zhou, Y.; Chen, J.M.; Wang, J.; Pang, Z.Z. Development and evaluation of Hainan *Camellia* SSR molecular marker basrd on transcriptome data. *Mol. Plant Breed.* **2022**, 1–19. Available online: <http://kns.cnki.net/kcms/detail/46.1068.S.20220316.2203.008.html> (accessed on 27 September 2022).
26. Gu, X.Y.; Guo, Z.H.; Ma, X.; Bai, S.Q.; Zhang, X.Q.; Zhang, C.B.; Chen, S.Y.; Peng, Y.; Yan, Y.H.; Huang, L.K.; et al. Population genetic variability and structure of *Elymus breviaristatus* (Poaceae: Triticeae) endemic to Qinghai–Tibetan Plateau inferred from SSR markers. *Biochem. Syst. Ecol.* **2015**, *58*, 247–256. [CrossRef]
27. Wu, W.D.; Liu, W.H.; Sun, M.; Zhou, J.Q.; Liu, W.; Zhang, C.L.; Zhang, X.Q.; Peng, Y.; Huang, L.K.; Ma, X. Genetic diversity and structure of *Elymus tangutorum* accessions from Western China as unrevealed by AFLP markers. *Hereditas* **2019**, *156*, 8. [CrossRef]
28. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
29. Dice, L.R. Measures of the amount of ecologic association between species. *J. Ecol.* **1945**, *26*, 297–302. [CrossRef]
30. Nei, M. Genetic Distance between Populations. *Am. Nat.* **1972**, *106*, 283–292. [CrossRef]
31. Xiong, Y.; Liu, W.; Xiong, Y.; Yu, Q.; Ma, X.; Lei, X.; Zhang, X.; Li, D. Revelation of genetic diversity and structure of wild *Elymus excelsus* (Poaceae: Triticeae) collection from western China by SSR markers. *PeerJ* **2019**, *7*, e8038. [CrossRef]
32. Evanno, G.S.; Regnaut, S.J.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef]
33. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [CrossRef]
34. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314–331. [CrossRef]
35. Chai, S.F.; Zhuang, X.Y.; Zou, R.; Chen, Z.; Jiang, Y.; Wei, X. Genetic diversity analysis of endangered plant *Camellia pubipetala* detected by ISSR. *Acta Bot. Boreali-Occident. Sin.* **2014**, *34*, 93–98. [CrossRef]
36. Najaphy, A.; Parchin, R.A.; Farshadfar, E. Evaluation of genetic diversity in wheat cultivars and breeding lines using inter simple sequence repeat markers. *Biotechnol. Biotec. Eq.* **2011**, *25*, 2634–2638. [CrossRef]
37. Qi, H.S.; Dai, J.N.; Yang, Y.N.; Li, S.Y.; Wang, J.; Shi, L.C.; Wu, Y.G.; Lai, H.G.; Hu, X.W.; Yu, J. Identification of *Camellia oleifera* seed DNA barcode based on trnH-psbA and matK sequence. *Mol. Plant Breed.* **2019**, *17*, 5057–5065. [CrossRef]
38. Dai, J.N.; Yu, J.; Qi, H.S.; Zheng, W.; Wang, J.; Wu, Y.G.; Lai, H.G.; Hu, X.W. DNA Barcoding Identification of Different Species in *Camellia* Based on trnH-psbA and matK Sequences. *Chin. J. Trop. Crops* **2021**, *42*, 611–619.