diversity

MDPI

Article

Development of Polymorphic Microsatellite Markers and Identification of Applications for Wild Walnut (*Juglans regia* L.) in Middle Asia

Xuerong Li [1,2], Xiyong Wang [2], Zhijun Cui [3], Wei Shi [1,2,*], Junhua Huang [1,*] and Jiancheng Wang [2,*]

1 College of Forestry and Landscape Architecture, Xinjiang Agricultural University, Urumqi 830052, China; li_xuerong_123@163.com
2 State Key Laboratory of Desert and Oasis Ecology, Key Laboratory of Ecological Safety and Sustainable Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China; wxy.701@163.com
3 Gongliu County Forestry and Steppe Bureau, Gongliu 835400, China; cuizhijunabc123@163.com
* Correspondence: shiwei@ms.xjb.ac.cn (W.S.); huangjunhua-7311@163.com (J.H.); www-1256@ms.xjb.ac.cn (J.W.)

Abstract: The common walnut (*Juglans regia* L.), a species of significant economic and ecological importance, along with its variant, holds developmental value. Simple sequence repeat (SSR) markers are used as an effective method for material identification due to their co-dominant inheritance. In this work, a total of 357629 SSR loci were searched based on previously published walnut genome sequences, with a density of 662.28 (SSRs/Mb). The majority of all these loci were found to be single nucleotide A/T (58.49%), followed by the dinucleotide AT (19.48%). Based on the results of genome-wide SSR site design, 22 effective primer pairs were designed and screened to analyze the genetic diversity of 48 wild walnut samples from three countries. Each locus had an average of 5.17 effective alleles (Ne), with an average polymorphism information content (PIC) of 0.71, average heterozygosity (He) of 0.52, and average genetic differentiation index (Fst) of 0.09. The analysis of 48 wild walnut samples from three countries (Kyrgyzstan, Tajikistan, and China) using the unweighted pairwise clustering method and arithmetic mean (UPGMA) along with principal coordinate analysis revealed the division of all samples into three groups. In summary, these novel SSR markers serve as a reference for the identification of wild walnut germplasm in Central Asia and for the development and utilization of wild walnut seed resources.

Keywords: *Juglans regia*; SSR markers; species identification; whole genome; genetic diversity

Citation: Li, X.; Wang, X.; Cui, Z.; Shi, W.; Huang, J.; Wang, J. Development of Polymorphic Microsatellite Markers and Identification of Applications for Wild Walnut (*Juglans regia* L.) in Middle Asia. *Diversity* **2023**, *15*, 1073. https://doi.org/10.3390/d15101073

Academic Editors: Petras Prakas and Mario A. Pagnotta

Received: 17 August 2023
Revised: 30 September 2023
Accepted: 7 October 2023
Published: 10 October 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

Walnut (*Juglans regia* L.) is a monoecious and dichogamous tree species that was first domesticated in Central Asia [1,2]. The genus Juglans (order Fagales, family *Juglandaceae*) includes more than 20 diploid species, with $2n = 2x = 32$ chromosomes [3]. Today, walnuts are farmed around the world, from Europe to South Africa. Globally, around 1.3 million hectares of walnut are planted every year [4]. The United States, China, and Turkey account for 69.22% of the world's walnut acreage, and these three countries also contribute to 81.32% of the world's walnut production [5]. It is generally accepted that *J. regia* is native to the mountain ranges of southeastern Europe (Carpathian Mountains) and west–central Asia [4,6,7]. The natural range of walnut extends from Xinjiang in China in the east to the Caucasus in the west, and is mainly found in mountainous areas [7,8]. Cultivated varieties of walnuts are distinct from wild varieties [9,10], and it has been speculated that the types of walnuts cultivated around the globe were selected as seedlings from these Eurasian natural populations thousands of years ago [10,11]. Walnut has been studied in various ways, including cultivar development and nutrient extraction [12,13].

Molecular markers such as simple sequence repeats (SSRs), random amplified polymorphic DNA (RAPD) [14], amplified fragment length polymorphisms (AFLPs), and inter simple sequence repeats (ISSRs) [15] markers have extensively been used in conservation genetics to guide conservation [16,17]. Knowledge of the genetic diversity of a species is crucial for the effective management and utilization of its germplasm. In particular, RAPD- [18], SSR- [19], and AFLP- [20] based approaches are powerful tools for studying genetic relationships and identity due to the innate consideration of complicating factors such as codominance, super mutation, and multiple alleles [21].

SSRs produce different types based on different classification criteria, and different types of SSRs bring out different messages. SSRs are classified into three major categories according to the different base-sequence arrangements. Accordingly, the categories are: complete microsatellites, incomplete microsatellites, and complex microsatellites. Complete microsatellites are sequences composed of bases linked without interruption, and interrupted microsatellites are sequences with fewer than three non-repetitive bases between bases and more than three repeating bases at either end of the sequence. A complex microsatellite is a sequence composed of no fewer than two or more tandem bases separated by no fewer than three non-repetitive bases, and there are always at least five base repeats within the sequence [22]. The application of SSR in different sequences derived from simple sequence repeat interval amplification polymorphism-Inter-Simple Sequence Repeat (ISSR), expressed sequence tag microsatellites-Expressed Sequenced Tags-Simple Sequence Repeat (EST-SSR), whole-genome SSR, and other molecular markers, and the above SSR molecular marker technology in the nuclear genome and chloroplast genome has been widely used. Evidence suggests that genomic SSRs exhibit greater levels of polymorphism compared with the genic SSR markers [23,24]. This may be due to a reduction in genic regions sequence polymorphism caused by the selective implication effect during species evolution [25], or the presence of introns in the genomic sequence itself, but not in the EST sequence. Alternatively, high levels of genomic-SSR polymorphism may be caused by intron polymorphism [26]. As a consequence, attempting to use genomic-SSR to identify closely related individuals will be a more sensitive form of analysis than using other genic SSRs [23]. However, it is worth emphasizing that other genic SSRs (such as EST-SSR from DNA-transcribed regions) may be better indicators of phenotypic traits as well as physiological and biochemical characteristics [27].

Cultivated walnuts are domesticated from wild walnuts [28]. Cultivated walnuts have a high economic value, and considerable effort in studying the species using molecular methods has resulted in the identification of 18 core loci that may be used to distinguish cultivars from wild varieties [29]. These techniques have been proven effective, and researchers have been able to identify multiple cultivars by solely using four pairs of SSR primers [13]. Similar work has been able to distinguish the genetic diversity of walnuts living at different altitudes in Morocco [30]. In the same vein, these molecular techniques (namely, 13 SSR markers) have been used to identify 22 hybrid genotypes selected from a study plot [31]. These same approaches have been used to identify wild walnuts available in China [32]. Clearly, SSR molecular markers can be effective tools to determine the genetic relationship between walnut species. However, despite the sequence and assembly of the walnut genome and the fairly robust history of using molecular methods to study walnuts, few have focused on the economically important Central Asian population.

In our study, the assembled genome sequence of walnut was used to excavate SSR loci, design and screen primers, and analyze the genetic diversity of wild walnut to complete the identification of wild walnut. Our goals were the following: (1) develop and identify SSRs in the whole genome sequence of walnut; (2) select primers with better polymorphism from wild walnut materials; and (3) identify relationships among different populations of walnut in Central Asia.

2. Materials and Methods

2.1. Excavation and Primer Design of SSRs

We used MISA to identify microsatellite loci in the assembled *J. regia* L. nuclear genome http://pgrc.ipk-gatersleben.de/misa/misa.html (accessed on 15 January 2023). The parameters set for this analysis included nucleotide repeats (MNRs), dinucleotide repeats (DNRs), trinucleotide repeats (TNRs), tetranucleotide repeats (TTRs), pentenucleotide repeats (PNRs), and hexonucleotide repeats (HNRs). Their minimum repeat units were 10, 6, 5, 5, 5, and 5, respectively. Any repeated sequence longer than 100 bp is regarded as an SSR locus. We then input these SSR loci into Primer3 http://primer3.sourceforge.net/releases.php (accessed on 15 January 2023) to design our primers. In this effort, the primer design looked to identify loci wherein the optimal nucleotide length was 21 bp (18 Bp–23 bp), the size of the amplified product was 80–280 bp, the optimal annealing temperature was 58 °C, and the optimal GC content was set as 50% (50–60%). To ease the scoring procedure in the genotyping process, we tagged forward primer sequences with a 5′-end fluorescent dye (FAM Blue) (Shanghai General Biotechnology Co., Ltd., Shanghai, China).

2.2. DNA Extraction and SSR-PCR Reaction

Wild walnuts were collected from wild walnut populations in several Central Asian countries. In all, we collected 23 wild walnuts (Y) from Ili, Xinjiang, China, 14 wild walnuts (D) from Dushanbe, Tajikistan, and 9 wild walnuts (J) from Wild Fruit Forest Reserve, Kyrgyzstan (Table 1, Figure 1). All leaf material was taken from fresh and healthy leaves. Upon collection, we ensured that each plant from which we collected walnuts was more than fifteen meters apart from another collected sample. The collected leaves were dried with silica gel and numbered. Using the DNA safe Plant Kit (Tiangen Biotechnology, Beijing, China), genomic DNA was extracted from the leaves of *J. regia* L. in accordance with the manufacturer's instructions. Before polymerase chain reaction (PCR), the DNA samples were diluted to approximately 20 ng/mL.

Table 1. Walnut samples from three regions of Central Asia.

Serial Number	ID	Location	Number	Latitude	Longitude
1	J1-J9	Kyrgyzstan	8	40.853421	73.66128
2	D1-D14	Tajikistan	14	38.584371	68.768641
3	Y1-Y24	China	23	43.217701	82.152833

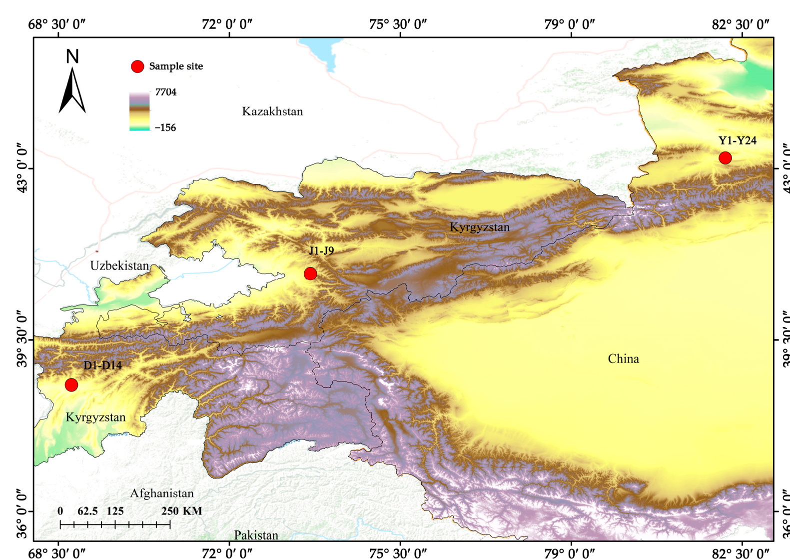

Figure 1. Map of walnut collection areas.

All samples were amplified by PCR using a 25 µL reaction system. Template DNA 1 µL, forward primer 1 µL (10 µmol/L), reverse primer 1 µL (10 µmol/L), 2 × Easy Taq RPCR Super Mix (Trans Gen Biotech, Beijing, China) 12.5 µL, and ddH$_2$O 9.5 µL. The amplification reaction was divided into three stages. The first stage was pre-denaturation at 94 °C for 5 min, while the second stage had 35 cycles: denaturation at 94 °C for 30 s, annealing (Table 2) for 1 min, and extension at 72 °C for 40 s. The last cycle was extended for 10 min at 72 °C. Capillary electrophoresis and fluorescent labeling were used to detect PCR products. The samples were screened for alleles (FSA) generated by the samples using GeneMapper Software (Applied Biosystems™, Waltham, MA, USA), which is genotyping software with independent readings of the chromatograms (Table S4).

Table 2. The characteristics of the 22 polymorphic SSR primer pairs of *Juglans regia*.

Name	Primer Sequences (5′–3′)	Repeat Motif	Ta (°C)	Fluorescent Dye	Size (bp)
WJR001	WJR001F: GGTGTAGGTTTGGAAGGCCA WJR001R: ATTGAGGCAACGGGAAGAGG	CT (8)	60	5′-FAM	139
WJR029	WJR029F: TTCTTGGCCGCAGAGCATTA WJR029R: TGTGCGTGCTAGATGGATGT	TC (7)	58	5′-FAM	148
WJR105	WJR105F: CACACACACACACACACACA WJR105R: CGTCTCACTCTCACTTCCAGG	GA (7)	60	5′-FAM	114
WJR157	WJR157F: GGTCGAGATCACCAAATGGC WJR157R: GCCGGCAGCTTTACTACTCA	AG (9)	60	5′-FAM	150
WJR621	WJR621F: TGCATGCTGTCAAAGGTGTC WJR621R: CGAGCTAGTGAACATTTGCAGT	AT (15)	58	5′-FAM	155
WJR679	WJR679F: TTTTCTCGCAAAGCAGCTGG WJR679R: TGGTCATCGTCTGGTTGCAA	AT (9)	58	5′-FAM	81
WJR1022	WJR1022F: AACTGGACAACCTTGCCCAA WJR1022R: CAGCTCAATGGCTTCTTGGC	AT (16)	58	5′-FAM	166
WJR1193	WJR1193F: GGGCGCCGTTGAACAAATAT WJR1193R: CGGCCATCAGAGAGGGATTC	TC (12)	60	5′-FAM	134
WJR19057	WJR19057F: CCGTGGCACCTAATCCTTGT WJR19057R: AGAGAGAGAGAGAGAGAGAGAGAGA	AC (10)	61	5′-FAM	129
WJR18870	WJR18870F: TCACCTCTCTCTCACTCTCTCA WJR18870R: GCGCGCAACAGAAAGAGAAA	CT (6)	59	5′-FAM	151
WJR18123	WJR18123F: GCATTTTTGCCACCCCACCTT WJR18123R: TGCCAAGTGGTACAAGTGGA	AC (8)	58	5′-FAM	129
WJR27081	WJR27081F: ACAAACAACACCGACGAGGA WJR27081R: TTGACGTTGTTAGTGTGCCC	GA (9)	58	5′-FAM	112
WJR20785	WJR20785F: TCAGGACGGTATGCTTGACT WJR20785R: TGTGTGTTGGTGTGTGTGGA	TC (14)	58	5′-FAM	182
WJR19806	WJR19806F: AGAGTGGTGTGTGTGAGTGC WJR19806R: TCCCACCCTTCTCCTTCCTT	AG (6)	60	5′-FAM	138
WJR28944	WJR28944F: AAAAGACCTTCGATCGAGCC WJR28944R: AAAAGACCTTCGATCGAGCC	TA (7)	60	5′-FAM	161
WJR30675	WJR30675F: CGCTGGTTTCTGGCATGAAA WJR30675R: GAAGAGATCAGAACCGGCCA	AT (8)	59	5′-FAM	183
WJR32508	WJR32508F: AGCAGAGCGAAAGAGAGCAG WJR32508R: AGACGCAACCCTCAAACCAT	GA (21)	59	5′-FAM	183
WJR37937	WJR37937F: TGCATTCAGAACACGGGTGA WJR37937R: AAAGCATGAGTTATCCTTGCAAAA	TCT (7)	56	5′-FAM	150
WJR41945	WJR41945F: TGGTTAATCAGGCCATGGCT WJR41945R: GCAGTTGCCGCAAACTTGTA	AT (15)	58	5′-FAM	178
WJR17197	WJR17197F: TGCCATCACCATGTTCACCA WJR17197R: TTGCGGCAACCCTAGTTCTT	AT (9)	58	5′-FAM	165
WJR20924	WJR20924F: AGAGATGTGCGTGTGTGTGT WJR20924R: AAGTGACGGTGTCCCACAAG	TA (9)	59	5′-FAM	116
WJR19609	WJR19609F: GAAGCATGTGTGTGTGTGTGT WJR19609R: AGGCTCGTTCGTTTATGCCC	TA (10)	59	5′-FAM	141

2.3. Data Collection and Analysis

Using POPGENE32, we collected data on the following: the number of alleles (Na), effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), unbiased expected heterozygosity (uHe), polymorphism index (I), gene flow (Nm), polymorphism information content (PIC), inbreeding coefficient (Fis), proportion of differentiation (Fst), and estimations of the fixation index (F) 1.32 [33]. NTSYS-PC 2.10e software [34] was used to construct a dendrogram based on cluster analysis using the unweighted pair group method with the arithmetic mean average (UPGMA) based on a similarity matrix. Significant differences between the groups and samples were tested by an analysis of molecular variance in

GenAlEx v6.5 [35,36]. The genetic distances between walnut samples were calculated using the PowerMarker v3.25 [37] program. A principal coordinate analysis (PCoA) was conducted based on Nei's genetic distance using GenALEx v6.5 [36] software. The probability of identity (PID) and probability of identity in siblings (PIDsib) at microsatellite loci were calculated by CervusV3.0 software [38]. The conversion of data types was conducted by DataFormater [39]. The model-based Bayesian clustering method implemented in Structure 2.3 [40] derives the proportion of affiliation and overall structure of each sample in the predefined totals in each inferred cluster (q-estimation). In the absence of information about the origin of the populations, a total of 10,000 MCMC iterations per run were performed after a 10,000-step break-in period using a mixture of models with correlated allele frequencies [41,42]. The number of clusters (K) and statistics ΔK estimated by Structure were analyzed in the web-based StructureE Harvester program [43,44]. Population size (K) was determined for each K value by 10 independent runs ranging from K = 1 to K = 10.

3. Results

3.1. Genome-Wide Identification and Characterization of SSR Loci

Screening the whole genome of *J. regia* at 540,793,986 bp, 357,629 SSR markers were identified with a total length of 7,641,320 bp, and the number of SSRs present in the compound formation was 54,620. The frequency and density of SSRs in the whole genome were 662.28 and 14,150.59 bp/Mb, respectively, and accordingly we estimated SSRs accounting for 1.41% of the whole genome sequence (Table 3).

Table 3. SSRs in the *Juglans regia* genome.

Items	Numbers
Total size of genome (Mb)	540
Total number of identified SSRs	357,629
Total length of SSRs (bp)	7,641,320
Frequency (SSRs/Mb)	662.28
Density (bp/Mb)	14,150.59
Total content of genome SSRs (%)	1.41

In the *J. regia* genome, the length of SSRs varied from 10 bp to 1252 bp, with an average length of 25 bp. The most common (49,317 SSRs) repeat length was 10 bp and the number of repeats was 49,317, followed by 12, 11, and 14 bp, and the number was 33,779, 26,475, and 21,459, respectively (Figure 2a). Statistical analysis of the number of motif repeats of all SSR locus showed that the highest number was 134 (Figure 2b). Among them, the proportion of ten tandem repeats of the SSR locus was the highest at 21.16%, followed by four eleven repeats of the SSR locus at 12.17% (Figure 2b). Due to the different sizes of the repeated base sequences, all SSRs were divided into six SSR types: nucleotide, dinucleotide, trinucleotide, tetraconucleotide, pentanucleotide, and hexanucleotide. Among all SSRs in the whole genome of *J. regia*, single nucleotide repeats were the most abundant (58.49%), followed by dinucleotide (33.95%) repeats (Table S1). The total genome length of SSRs was 5,766,962 bp, and the total genome length of SSRs with Mono-, Di-, Tri-, Tetra-, Penta-, and Hexa- nucleotide repeats was 2,758,042 bp, 1,502,169 bp, 18,670 bp, 3920 bp, 2534 bp, and 624 bp, respectively. The main types were Di- and Tri- repeats.

The chromosome set of walnut has a total of sixteen chromosomes, from chromosome 1 (chr1) to chromosome 16 (chr16). According to the statistics, SSR loci are most distributed on chromosome 1, followed by chromosome 3, with both exceeding 30,000 loci. The number of SSR loci varied on each chromosome, with chr1 having the most SSR loci. We found chromosome 1 to contain most of the SSRs identified in this study, followed by chromosome 3 (chr3), with more than 30,000 loci on both. On each chromosome, it is the single nucleotide repeat that has the most SSR sites, followed by the dinucleotide repeat, trinucleotide repeat, quadruple nucleotide repeat, pentanucleotide repeat, and hexanucleotide repeat, with the exception of the chromosome chr4, which has the most dinucleotide repeats, followed by the

single nucleotide repeat, trinucleotide repeat, quadruple nucleotide repeat, pentanucleotide repeat, and hexanucleotide repeat (Table S2).

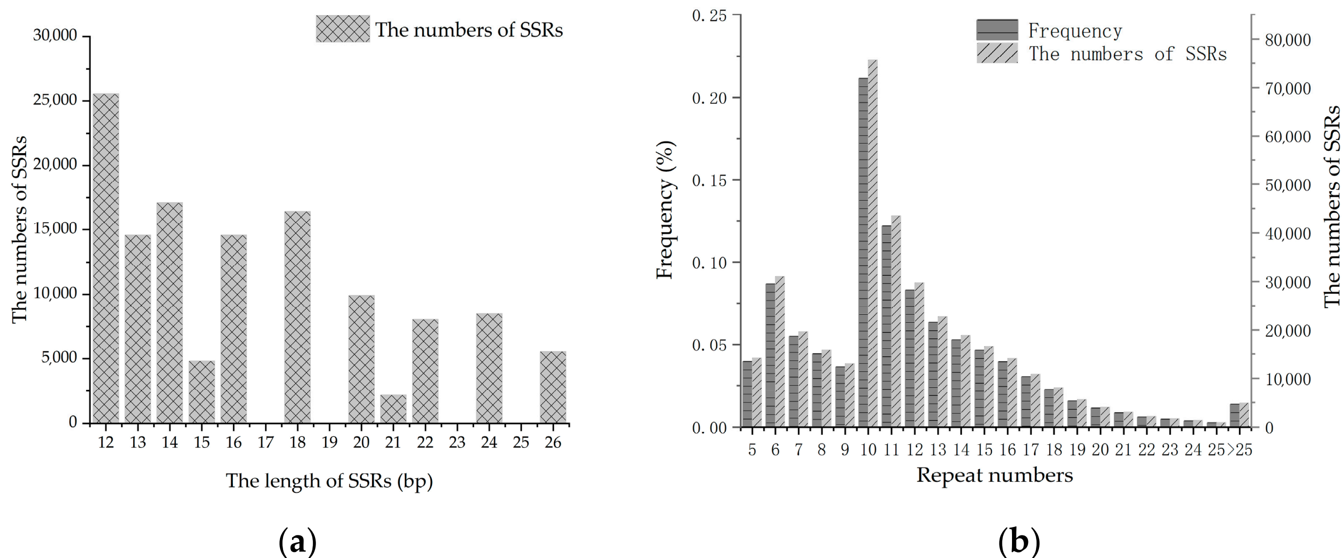

Figure 2. Characterization of SSR loci. (**a**) Number of SSRs of different lengths. (**b**) Frequency and number of SSRs with different number of replicates.

3.2. Development and Characterization of SSR Markers

Ultimately, we identified 316 number of alleles in 48 walnut samples, with an average of 14.36 alleles per SSR marker (Table 4). The number of alleles (Na) varied from 7 (WJR001, WJR27081, WJR29601, and WJR19609) to 27 (WJR105). The average number of Ne for the 22 SSR loci was 5.17, with a maximum value of 19.51 from WJR105 and a minimum of 1.47 from WJR41945. The maximum value of observed heterozygosity (Ho) was 1 (WJR37937), the minimum value was 0.11 (WJR001), and the mean value was 0.52. The expected heterozygosity (He) had a maximum value of 0.95 (WJR105), a minimum value of 0.32, and a mean value of 0.72. The Ne, Ho, and He values indicate good polymorphism in the 22 primer pairs screened. The polymorphism information content (PIC) ranged from 0.39 to 0.93 with an average of 0.71. There were up to 20 pairs of SSRs with PIC values higher than 0.5, indicating a high level of polymorphism in these markers. Shannon's information index (I) had a maximum value of 3.12 (WJR105), a minimum value of 0.89, and a mean value of 1.82. Gene flow (Nm) has a maximum value of 8.35 (WJR17197), a minimum value of 1.6, and a mean value of 2.68. The inbreeding coefficient (Fis) ranged from a maximum value of 0.81 to a minimum value of -0.4, and it had a mean value of 0.19. The proportion of differentiation (Fst) had a maximum value of 0.13, a minimum value of 0.03, and a mean value of 0.09 (Table 4). The results showed that the PID was 7.744×10^{-25}, and the PIDsib was 3.043×10^{-9}. In the individual identification analysis, the PIDsib value of the cumulative microsatellite loci should be less than 0.01 [45] to ensure the accuracy of individual identification, confirming that all samples are from single individuals.

Table 4. Summary of 22 polymorphic SSR locus developed for *Juglans regia*.

Locus	Na	Ne	Ho	He	uHe	I	Nm	PIC	Fis	Fst	F (Null)
WJR001	7	2.23	0.11	0.55	0.55	1.08	5.23	0.52	0.81	0.05	0.68
WJR029	11	2.16	0.17	0.49	0.54	1.27	3.05	0.53	0.55	0.08	0.54
WJR105	27	19.51	0.55	0.89	0.95	3.12	3.70	0.93	0.25	0.06	0.26
WJR157	10	3.04	0.49	0.64	0.67	1.48	3.76	0.65	0.28	0.06	0.17

Table 4. *Cont.*

Locus	Na	Ne	Ho	He	uHe	I	Nm	PIC	Fis	Fst	F (Null)
WJR621	23	11.05	0.92	0.84	0.91	2.67	2.91	0.90	−0.12	0.08	−0.003
WJR679	19	5.66	0.53	0.82	0.82	2.22	4.27	0.82	0.32	0.06	0.23
WJR1022	23	8.23	0.62	0.82	0.88	2.54	2.30	0.87	0.21	0.10	0.18
WJR1193	12	3.16	0.48	0.65	0.68	1.62	1.62	0.66	0.09	0.13	0.21
WJR19057	15	5.10	0.38	0.75	0.80	1.99	2.30	0.82	0.45	0.10	0.38
WJR18870	15	5.20	0.64	0.71	0.81	2.01	1.63	0.79	0.13	0.13	0.12
WJR18123	15	3.99	0.40	0.72	0.75	1.90	1.9	0.73	0.32	0.12	0.31
WJR17197	17	4.46	0.63	0.76	0.78	1.95	8.35	0.75	0.24	0.03	0.13
WJR27081	7	2.62	0.29	0.57	0.62	1.33	7.53	0.59	0.40	0.03	0.37
WJR20924	21	7.68	0.46	0.79	0.87	2.54	1.6	0.87	0.27	0.13	0.32
WJR20785	25	7.53	0.96	0.82	0.87	2.55	2.80	0.86	−0.13	0.08	−0.06
WJR19806	9	2.75	0.20	0.58	0.64	1.29	1.96	0.61	0.51	0.11	0.55
WJR28944	12	1.88	0.27	0.51	0.47	1.16	2.70	0.45	0.34	0.08	0.32
WJR29601	7	2.81	0.39	0.61	0.64	1.37	1.77	0.64	0.30	0.12	0.26
WJR32508	12	6.27	0.83	0.77	0.84	2.10	1.92	0.83	−0.12	0.12	0.008
WJR37937	9	3.44	1.00	0.71	0.71	1.52	6.82	0.66	−0.40	0.04	−0.19
WJR41945	13	1.47	0.31	0.36	0.32	0.89	2.75	0.39	−0.04	0.08	0.03
WJR19609	7	3.56	0.78	0.68	0.72	1.41	3.84	0.70	−0.08	0.06	−0.06
Total	316										
Mean	14.36	5.17	0.52	0.68	0.72	1.82	2.68	0.71	0.19	0.09	0.22

Na = number of alleles, Ne = effective number of alleles, Ho = observed heterozygosity, He = expected heterozygosity, uHe = unbiased expected heterozygosity, I = Shannon's information index, Nm = gene flow, PIC = polymorphism information content, Fis = inbreeding coefficient, Fst = proportion of differentiation, F(Null) = null allele.

3.3. Analysis of Genetic Diversity of Walnut

Based on the information presented in Table 5, it can be seen that populations from three regions of Kyrgyzstan (J), Tajikistan (D), and China (Y) were genetically characterized at 22 SSRs. The number of alleles (Na) had a mean value of 8.27 in the Kyrgyzstan region, 14.68 in the Tajikistan region, and 23.18 in China. The effective number of alleles (Ne) had a mean value of 7.64 in the Kyrgyzstan region, 7.86 in the Tajikistan region, and 7.73 in the China. The mean value of Shannon's information index (I) for the three regions was the highest at 1.69 in the Kyrgyzstan region and the lowest at 1.33 in China. The observed heterozygosity (Ho) results for the three regions show that Kyrgyzstan, Tajikistan, and China have values of 0.64, 0.59, and 0.43, respectively. The expected heterozygosity (He) for Kyrgyzstan and Tajikistan are not much different from each other at 0.74 and 0.71, respectively, while it is lower for China at 0.60. The unbiased expected heterozygosity (uHe) of the three regions also shows that Kyrgyzstan is the highest, followed by Tajikistan and then China. The fixation index (F) of the three regions is different, with the lowest being 0.14 in Kyrgyzstan, followed by 0.18 in Tajikistan, and the highest being 0.31 in China (Table 5).

Table 5. Population genetic characteristics based on 22 SSR locus data in three *Juglans regia* populations.

Pop		N	Na	Ne	I	Ho	He	uHe	F
J	Mean	8.27	7.64	5.06	1.69	0.64	0.74	0.79	0.14
	SE	0.19	0.58	0.57	0.10	0.05	0.03	0.03	0.06
D	Mean	14.68	7.86	4.29	1.59	0.59	0.71	0.73	0.18
	SE	0.17	0.63	0.43	0.10	0.06	0.03	0.03	0.07
Y	Mean	23.18	7.73	3.32	1.33	0.43	0.60	0.62	0.31
	SE	0.31	0.75	0.54	0.11	0.07	0.04	0.04	0.09

N = number of individuals sampled, Na = number of alleles, Ne = effective number of alleles, I = Shannon's information index, Ho = observed heterozygosity, He = expected heterozygosity, uHe = unbiased expected heterozygosity, F = fixation index, SE = standard error of the mean (S.E. Mean).

Bayesian clustering methods revealed a stronger genetic structure among wild walnut populations. The provisional statistic ΔK shows that the maximum likelihood value of K = 3 distinguishes wild walnuts from the three countries (Supplement Figure S1). The results of the population structure analysis plots showing K = 2 to K = 5 in Figure 3 indicate that walnuts from the three regions can be roughly distinguished when K = 3, and that individuals in each cluster have higher membership (q-value) in their clusters (Figure 3).

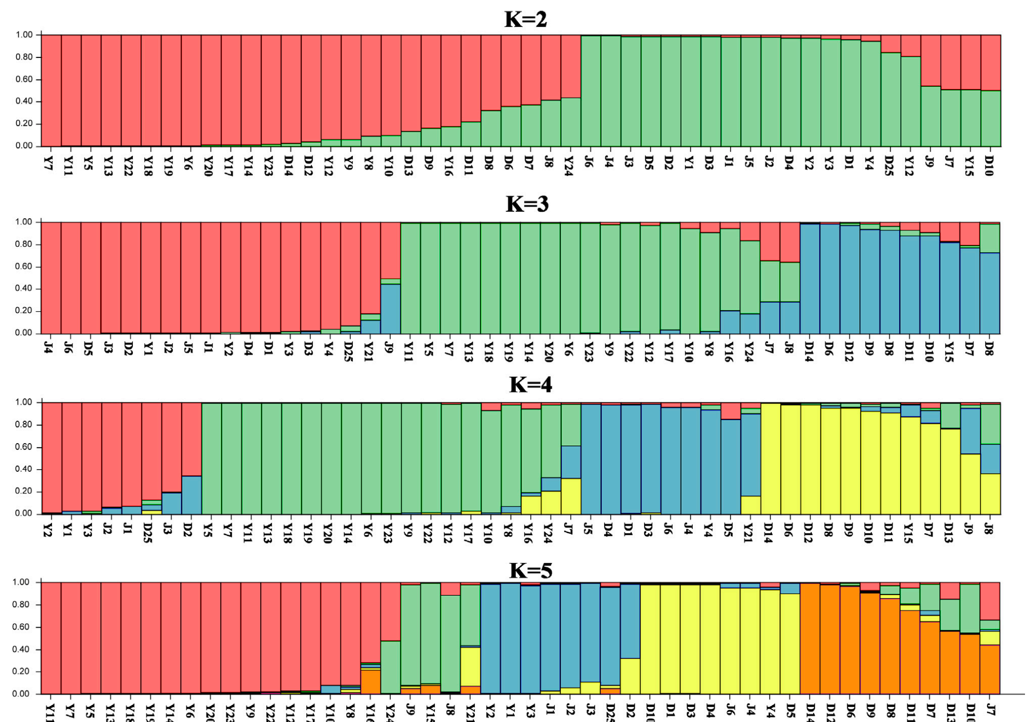

Figure 3. Bar plot showing the genetic structure of the 48 walnuts between K = 2 and K = 5. The vertical bars represent the membership coefficients (qi) of each individual.

The genetic similarity coefficient among the 48 samples of walnut accessions ranged from a maximum of 0.877 between Y18 and Y19, a minimum of 0 between J6 and Y21, with an overall average of 0.328 (Table S3). A dendrogram drawn from the genetic distance based on an unweighted pair group method with arithmetic means (UPGMA) classified the 48 samples into three clusters. Cluster I is composed of sixteen samples of which six samples are from Kyrgyzstan, six samples are from Tajikistan, and four samples are from Ili. Cluster II is composed of nineteen samples of which eighteen samples are from Ili and one sample is from Kyrgyzstan. Cluster III is composed of thirteen samples of which nine samples are from Tajikistan, two samples are from Kyrgyzstan, and two samples are from Ili (Figure 4).

The PCoA of 48 wild walnut materials is shown in Figure 5. The results showed that 48 wild walnut materials were divided into three groups: the first group comprised five samples of walnuts from Tajikistan (D), five samples of walnuts from Kyrgyzstan (J), and three samples of walnuts from Ili (Y); the second group consisted of nine samples of wild walnuts from Kyrgyzstan (J), two samples of walnuts from Kyrgyzstan (J), and two samples of wild walnuts from Ili (Y); and the third group consisted of 18 samples of wild walnuts from Ili (Y) and one sample of walnuts from Kyrgyzstan (J). The first principal coordinate and the second principal coordinate accounted for 24.13% and 11.48% of the total variation of the total coordinates, respectively, explaining the total genetic variation of the 22 SSR loci.

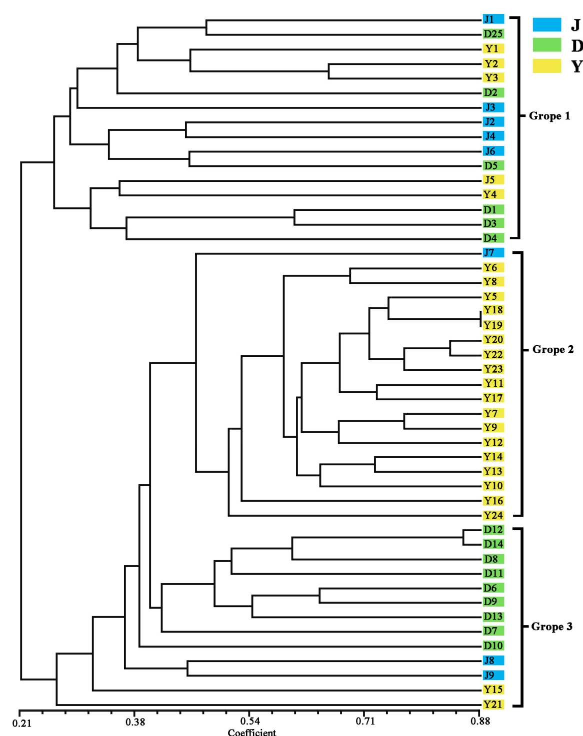

Figure 4. A dendrogram of genetic relationship among 48 walnuts on 21 SSR markers.

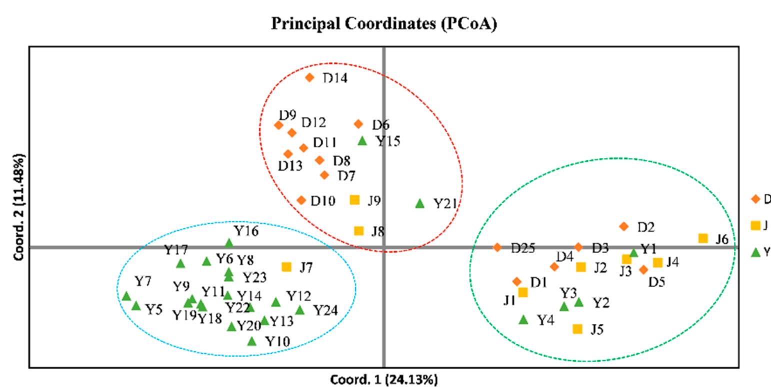

Figure 5. Principal coordinate analysis (PCoA) of 48 walnuts based on 21 SSR markers. Use red, blue and green circles to circle the three clusters.

4. Discussion

With the development of sequencing technology, more and more plant genomes have been sequenced, and screening SSR loci in genomes has been applied to a variety of plants, from cotton [24] to watermelon [12]. In this study, 357,629 SSR loci with a density of 662.28 (SSRs/Mb) were obtained by a reference genome, and the number and density are much higher than that of transcriptome data and watermelon [46], but slightly higher than cucumber [47]. In the previous selection of SSR loci in walnut, Wu et al. [48] obtained 31.2 Mb BAC end sequences for walnut, and found an average frequency of one SSR per 2.3 kb. Topcu et al. [49] obtained only 516 points due to limitations in genome sequencing. The percentages of various bases in the results are comparable to the transcriptome data [50]. However, there were changes when compared to the SSR loci derived from the transcriptome of *Juglans cathayensis* [51]. One reason for observing a higher percentage of A/T than C/G in the data is because the A/T slippage rate is greater than the G/C slippage rate; hence, A/T is seen more frequently in SSR motifs [52]. This conclusion is comparable to the fact that the AT or AAT type is more prevalent among dicots [46]. Although GC, TC,

and GA types have very stable structures, their frequency is low in the results, which are similar to those of other genomes [53].

Due to the remarkable conservation of flanking areas, microsatellite markers may be employed for genetic investigations, species, and genus levels [54]. These markers can also be utilized for other purposes, such as genetic diversity analysis, linkage analysis, population structure determination, evolutionary analysis, and so on. Therefore, we used SSR markers to analyze 48 walnut samples. When PIC $\geq$ 0.5, the gene seat is highly polymorphic. When 0.25 $\leq$ PIC < 0.5, it is moderately polymorphic, and when PIC < 0.25, it is lowly polymorphic [55]. The average PIC index of the 22 pairs of SSRs selected in this study was higher than 0.5. We observed high rates of polymorphism across the 22 genomic-SSR markers identified in this study. Our results were consistent with other studies focused on the species [56–59]. The small discrepancies between our results and those of other studies are likely due to variations in sampling size and the different genetic motifs studied [60]. The heterozygosity index can effectively assess the genetic diversity of species [61]. Ho and He were higher compared to Magige et al. (Ho = 0.47, He = 0.56) [57], but similar to Anthony et al., Torokeldiev et al. (Ho = 0.573, He = 0.545) [62], and Wang et al. (Ho = 0.5116, He = 0.5864) [63]. The Ho degree of sunflower was 0.56. The Ho and He of non-heading Chinese cabbage were 0.530 and 0.726, respectively. Both of them were higher than the present study, however they were artificially selected plant species [64,65]. Genetic diversity is influenced by many factors, including reproductive systems, natural selection, bottleneck effect, evolution, life history, habitat fragmentation, anthropogenic disturbance, and many others [60]. A comparison of He and Ho among walnuts from the three regions showed that walnuts from China had lower values in both He and Ho than those from the other two regions. In general, there was higher genetic variation among populations in areas with lower genetic differentiation among populations, which is consistent with other existing work on heterotic woody plant species [66]. The comparison of the genetic diversity of different populations also shows that the conservation of wild walnut germplasm resources in Xinjiang, China, urgently needs to consider the possibility of utilizing germplasm resources from other countries to improve the genetic diversity of wild walnut.

SSRs are widely used as genotyping markers that are experimentally transferable between related species [67]. For wild species, SSRs are useful for diversity studies that may be used to infer rates of gene flow and hybridization and, consequently, estimate the relationship between and amongst subspecies [68]. Based on the results shown by principal coordinate analysis (PCoA) and previous studies, SSR markers can be used to identify the level of genetic variation in *J. regia* [69,70]. According to the clustering results, we found that the separation of walnuts is slightly different from their geographical origin. Although the walnuts come from different regions, the clustering diagram in the results shows three groups. In the results of PCoA, the 48 walnut samples were also divided into three groups. The clustering results show that all individuals are classified into three categories. It is worth noting that one category of walnuts from China absolutely dominated in the PCoA diagram, indicating that the wild walnuts from China are related to the wild walnuts from Kyrgyzstan and Tajikistan, but the relationship is more distant. The relatively large seed mass of walnut plants means that they cannot be spread over long distances by wind, birds, etc. [71]. Therefore, we consider the behavior of human activities that interfere with wild walnuts, especially walnuts as a plant with economic value. For example, based on preliminary genetic analyses in Yunnan (China), rural networks and family relationships were found to be associated with the genetic structure of native populations of walnut [70]. Understanding the geographic distribution of common walnuts needs to be based on a clear historical record of the human use of walnuts as well as their dissemination over the past thousand years [72,73]. Historical sources indicate that the common walnut has been used for nut production and silviculture in Central Asia since at least the 5th century AD [74]. Similar to walnuts, almonds and pomegranates are perennial food plants that were spread throughout Asia by humans [75], but walnuts have

only been widely propagated by grafting in the last century, and are still propagated by seed in Asia [76]. In particular, the emergence of the Silk Road, human activities (cultural barriers represented by language differences), and evolutionary processes such as selection and drift have sometimes influenced the genetic spatial structure of walnuts in ancient Asian commerce [77]. The present study was able to use these 22 SSR pairs to broadly distinguish walnuts from different geographical locations. Normally, the domestication of perennials produces dramatic changes in plant reproduction, inflorescence, and fruiting characteristics [78,79], but wild walnut does not meet the criteria for domestication when compared to cultivated walnut, which is not different from the cultivated form in the first place [9]. Therefore, cultivated walnuts may have been derived from seedlings selected from geographically distinct natural populations over the course of thousands of years [79]. We need to first clarify the distinction between the geographical origins of wild walnuts for the subsequent use of more informative molecular markers or sequencing for the study of the genetic structure of wild walnuts in Central Asia, the walnut identification system, and even to help in the selection of the core germplasm of walnuts.

5. Conclusions

In this study, the published SSR loci with 1–6 nucleotide repeats in the chromosomal sequence of the walnut genome was mined by using published whole-base sequencing of walnut, and the sequence features were analyzed. Meanwhile, synthetic primers were designed for polymorphism detection based on the desired type of SSR loci, and 22 pairs of primers with high polymorphism were obtained by screening. These 22 pairs of primers were used to analyze and identify the genetic diversity of 48 wild walnut samples collected from Central Asia. It was possible to use the 22 primer pairs to differentiate wild walnut. The results are also informative for other species. The results show that the genetic diversity of wild walnut in Xinjiang, China, is low, and the use of more genetically diverse walnut populations from other countries may be considered for conservation. The affinities of wild nuclei in different parts of Central Asia may be affected by anthropogenic factors. The study showed that the 22 SSR markers could be used for the identification and genetic diversity analysis of wild walnut, and contribute to the subsequent selection of core germplasm and the conservation of the genetic diversity of wild walnuts in Central Asia.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d15101073/s1, Table S1: The main motif of *Juglans regia* genome; Table S2: The distribution of SSR repeat types on each chromosome; Table S3: The genetic similarity coefficient among the 48 accessions based on 22 SSR markers; Table S4: Genotype multilocus of 48 wild walnut samples. Figure S1 Second order of change of the log-likelihood of the data (ΔK) as a function of K.

Author Contributions: Conceptualization, W.S. and J.H.; methodology, X.L. and W.S.; software, X.L.; validation, W.S. and X.L.; formal analysis, X.L. and X.W.; investigation, Z.C. and X.W.; resources, W.S.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, W.S., J.W. and X.L.; visualization, X.L.; supervision, W.S.; project administration, J.W.; funding acquisition, W.S. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategic Biological Resources Capacity Building Project, CAS (grant number KFJ-BRP-017-72), and the Third Xinjiang Scientific Expedition Program (grant number 2022xjkk1503).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

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

References

1. McGranahan, G.; Leslie, C. Walnuts (*Juglans*). *Acta Hortic.* **1991**, *290*, 907–974. [CrossRef]
2. Zeven, A.C.; Zhukovsky, P.M. *Dictionary of Cultivated Plants and Their Centres of Diversity: Excluding Ornamentals, Forest Trees and Lower Plants*; Centre for Agricultural Publishing and Documentation: Wageningen, The Netherlands, 1975.
3. Woodworth, R.H. Meiosis of microsporogenesis in the Juglandaceae. *Am. J. Bot.* **1930**, *17*, 863–869. [CrossRef]
4. Aradhya, M.K.; Potter, D.; Simon, C.J. *Darwin's Harvest: New Approaches to the Origins, Evolution, and Conservation of Crops*; Columbia University Press: New York, NY, USA, 2006; pp. 143–170.
5. Wang, P.; Zhong, L.; Yang, H.; Zhu, F.; Hou, X.; Wu, C.; Zhang, R.; Cheng, Y. Comparative analysis of antioxidant activities between dried and fresh walnut kernels by metabolomic approaches. *LWT-Food Sci. Technol.* **2022**, *155*, 112875. [CrossRef]
6. Hemery, G.E.; Savill, P.S.; Thakur, A. Height growth and flushing in common walnut (*Juglans regia* L.): 5-year results from provenance trials in Great Britain. *Forestry* **2005**, *78*, 121–133. [CrossRef]
7. Aradhya, M.; Woeste, K.; Velasco, D. Genetic Diversity, Structure and Differentiation in Cultivated Walnut (*Juglans regia* L.). In Proceedings of the 6th International Walnut Symposium, Melbourne, Australia, 25–27 February 2010.
8. Akça, Y.; Mehmet Sen, S. Studies on selection of walnut (*Juglans regia* L.) in Gürün. In *Progress in Temperate Fruit Breeding*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 179–181.
9. Carrion, J.S.; Sanchezgomez, P. Palynological data in support of the survival of walnut (*Juglans regia* L.) in the western mediterranean area during last glacial times. *J. Biogeogr.* **1992**, *19*, 623–630. [CrossRef]
10. Squires, V.R. *The Role of Food, Agriculture, Forestry and Fisheries in Human Nutrition-Volume III*; EOLSS Publications: Abu Dhabi, United Arab Emirates, 2011.
11. Molnar, T.; Zaurov, D.; Capik, J.; Eisenman, S.; Ford, T.; Nikolyi, L.; Funk, C. Persian walnuts (*Juglans regia* L.) in Central Asia. *Annu. Rep. North. Nut Grow. Assoc.* **2011**, *101*, 56–69.
12. Zhu, Z.; Zhu, W.; Yi, J.; Liu, N.; Cao, Y.; Lu, J.; Decker, E.A.; McClements, D.J. Effects of sonication on the physicochemical and functional properties of walnut protein isolate. *Food Res. Int.* **2018**, *106*, 853–861. [CrossRef]
13. Chen, L.N.; Ma, Q.G.; Chen, Y.K.; Wang, B.Q.; Pei, D. Identification of major walnut cultivars grown in China based on nut phenotypes and SSR markers. *Sci. Hortic.* **2014**, *168*, 240–248. [CrossRef]
14. Nicese, F.P.; Hormaza, J.I.; McGranahan, G.H. Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. *Euphytica* **1998**, *101*, 199–206. [CrossRef]
15. Potter, D.; Gao, F.Y.; Aiello, G.; Leslie, C.; McGranahan, G. Intersimple sequence repeat markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia*) cultivars. *J. Am. Soc. Hort. Sci.* **2002**, *127*, 75–81. [CrossRef]
16. Kantety, R.V.; La Rota, M.; Matthews, D.E.; Sorrells, M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* **2002**, *48*, 501–510. [CrossRef]
17. Zha, H.G.; Milne, R.I.; Sun, H. Morphological and molecular evidence of natural hybridization between two distantly related Rhododendron species from the Sino-Himalaya. *Bot. J. Linn. Soc.* **2008**, *156*, 119–129. [CrossRef]
18. Fatahi, R.; Ebrahimi, A.; Zamani, Z. Characterization of Some Iranians and Foreign Walnut Genotypes Using Morphological Traits and RAPD Markers. *Hortic. Environ. Biote.* **2010**, *51*, 51–60.
19. Christopoulos, M.V.; Rouskas, D.; Tsantili, E.; Bebeli, P.J. Germplasm diversity and genetic relationships among walnut (*Juglans regia* L.) cultivars and Greek local selections revealed by Inter-Simple Sequence Repeat (ISSR) markers. *Sci. Hortic.* **2010**, *125*, 584–592. [CrossRef]
20. Bayazit, S.; Kazan, K.; Gulbitti, S.; Cevik, V.; Ayanoglu, H.; Ergul, A. AFLP analysis of genetic diversity in low chill requiring walnut (*Juglans regia* L.) genotypes from Hatay, Turkey. *Sci. Hortic.* **2007**, *111*, 394–398. [CrossRef]
21. Foroni, I.; Rao, R.; Woeste, K. Molecular characterization of *Juglans regia* L. cultivars with SSR markers. In Proceedings of the 5th International Walnut Symposium, Sorrento, Italy, 9–13 November 2006.
22. Weber, J.L. Informativeness of human $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphisms. *Genomics* **1990**, *7*, 524–530. [CrossRef]
23. Parthiban, S.; Govindaraj, P.; Senthilkumar, S. Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. *3 Biotech* **2018**, *8*, 144. [CrossRef]
24. Lu, M.; Zhang, H.; An, H.; Zhou, W. Genetic variation and population structure of rosa roxburghii by est-based and genomic SSR markers. *Pak. J. Bot.* **2020**, *52*, 1315–1322. [CrossRef]
25. Maynard Smith, J.; Haigh, J. Hitch-hiking effect of a favorable gene. *Genet. Res.* **1974**, *23*, 23–35. [CrossRef]
26. Van, K.; Hwang, E.Y.; Kim, M.Y.; Kim, Y.H.; Cho, Y.I.; Cregan, P.B.; Lee, S.H. Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. *Euphytica* **2004**, *139*, 147–157. [CrossRef]
27. Casu, R.; Dimmock, C.; Thomas, M.; Bower, N.; Knight, D.; Grof, C.; McIntyre, L.; Jackson, P.; Jordan, D.; Whan, V.; et al. Genetic and expression profiling in sugarcane. In Proceedings of the 24th Congress of the International Society of Sugar Cane-Technologists, Brisbane, Australia, 17–21 September 2001.
28. Hsin-shi, C. On the eco-geographical characters and the problems of classification of the wild fruit-tree forest in the Ili Valley of Sinkiang. *J. Integr. Plant Biol.* **1973**, *2*, 239–253.
29. Wambulwa, M.C.; Fan, P.-Z.; Milne, R.; Wu, Z.Y.; Luo, Y.H.; Wang, Y.H.; Wang, H.; Gao, L.M.; Xiahou, Z.Y.; Jin, Y.C.; et al. Genetic analysis of walnut cultivars from southwest China: Implications for germplasm improvement. *Plant Divers.* **2022**, *44*, 530–541. [CrossRef] [PubMed]

30. Kabiri, G.; Bouda, S.; Haddioui, A. Evaluation of genetic diversity and structuration across altitude of walnut (*Juglans regia* L.) accessions from Morocco using SSR markers. *Bulg. J. Agric. Sci.* **2022**, *28*, 451–458.
31. Cseke, K.; Bujdoso, G.; Bader, M.; Mertl, T.; Benke, A.; Kampel, J.D. Genetic Identification of Hybrid Walnuts (*Juglans* × *intermedia* Carr.) in Hungary, the Hidden Potential for Future Breeding. *Sustainability* **2022**, *14*, 4782. [CrossRef]
32. Feng, X.; Zhou, H.; Zulfiqar, S.; Luo, X.; Hu, Y.; Feng, L.; Malvolti, M.E.; Woeste, K.; Zhao, P. The Phytogeographic History of Common Walnut in China. *Front. Plant. Sci.* **2018**, *9*, 1399. [CrossRef] [PubMed]
33. Yeh, F. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J. Bot.* **1997**, *129*, 157.
34. Rohlf, F.J. *NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System*, version 2.1; Exeter Publication Ltd. Setauket: New York, NY, USA, 2000.
35. Peakall, R.; Smouse, P.E. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **2006**, *6*, 288–295. [CrossRef]
36. 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]
37. Liu, K.; Muse, S.V. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* **2005**, *21*, 2128–2129. [CrossRef]
38. Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising how the computer program Cervus accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [CrossRef]
39. Fan, W.; Ge, H.; Sun, X.; Yang, A.; Zhang, Z.; Ren, M. DataFormater, a Software for SSR Data Formatting to Develop Population Genetics Analysis. *Mol. Plant Breed.* **2016**, *14*, 265–270.
40. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef] [PubMed]
41. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol. Ecol. Notes* **2007**, *7*, 574–578. [CrossRef]
42. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **2003**, *164*, 1567–1587. [CrossRef] [PubMed]
43. 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]
44. 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]
45. Waits, J.L.; Leberg, P.L. Biases associated with population estimation using molecular tagging. *Anim. Conserv.* **2000**, *3*, 191–199. [CrossRef]
46. Zhu, H.; Song, P.; Koo, D.H.; Guo, L.; Li, Y.; Sun, S.; Weng, Y.; Yang, L. Genome wide characterization of simple sequence repeats in watermelon genome and their application in comparative mapping and genetic diversity analysis. *BMC Genom.* **2016**, *17*, 557. [CrossRef] [PubMed]
47. Cavagnaro, P.F.; Senalik, D.A.; Yang, L.; Simon, P.W.; Harkins, T.T.; Kodira, C.D.; Huang, S.; Weng, Y. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics.* **2010**, *11*, 569. [CrossRef] [PubMed]
48. Wu, J.; Gu, Y.Q.; Hu, Y.; You, F.M.; Dandekar, A.M.; Leslie, C.A.; Aradhya, M.; Dvorak, J.; Luo, M.C. Characterizing the walnut genome through analyses of BAC end sequences. *Plant Mol. Biol.* **2012**, *78*, 95–107. [CrossRef]
49. Topcu, H.; Ikhsan, A.S.; Sutyemez, M.; Coban, N.; Guney, M.; Kafkas, S. Development of 185 polymorphic simple sequence repeat (SSR) markers from walnut (*Juglans regia* L.). *Sci. Hortic.* **2015**, *194*, 160–167. [CrossRef]
50. Yang, G.; Chen, S.; Chen, L.; Sun, K.; Huang, C.; Zhou, D.; Huang, Y.; Wang, J.; Liu, Y.; Wang, H.; et al. Development of a core SNP arrays based on the KASP method for molecular breeding of rice. *Rice* **2019**, *12*, 21. [CrossRef] [PubMed]
51. Dang, M.; Liu, Z.X.; Chen, X.; Zhang, T.; Zhou, H.J.; Hu, Y.H.; Zhao, P. Identification, development, and application of 12 polymorphic EST-SSR markers for an endemic Chinese walnut (*Juglans cathayensis* L.) using next-generation sequencing technology. *Biochem. Syst. Ecol.* **2015**, *60*, 74–80. [CrossRef]
52. Ashworth, V.; Kobayashi, M.C.; De La Cruz, M.; Clegg, M.T. Microsatellite markers in avocado (*Persea americana* Mill.): Development of dinucleotide and trinucleotide markers. *Sci. Hortic.* **2004**, *101*, 255–267. [CrossRef]
53. Portis, E.; Lanteri, S.; Barchi, L.; Portis, F.; Valente, L.; Toppino, L.; Rotino, G.L.; Acquadro, A. Comprehensive Characterization of Simple Sequence Repeats in Eggplant (*Solanum melongena* L.) Genome and Construction of a Web Resource. *Front. Plant. Sci.* **2018**, *9*, 401. [CrossRef] [PubMed]
54. Hamza, S.; Ben Hamida, W.; Rebai, A.; Harrabi, M. SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. *Euphytica* **2004**, *135*, 107–118. [CrossRef]
55. 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. [PubMed]
56. Shah, U.N.; Mir, J.; Ahmed, N.; Fazili, K.M. Assessment of germplasm diversity and genetic relationships among walnut (*Juglans regia* L.) genotypes through microsatellite markers. *J. Saudi Soc. Agric. Sci.* **2018**, *17*, 339–350.
57. Magige, E.A.; Fan, P.Z.; Wambulwa, M.C.; Milne, R.; Wu, Z.Y.; Luo, Y.H.; Khan, R.; Wu, H.Y.; Qi, H.L.; Zhu, G.F.; et al. Genetic Diversity and Structure of Persian Walnut (*Juglans regia* L.) in Pakistan: Implications for Conservation. *Plants* **2022**, *11*, 1652. [CrossRef]

58. Nagano, Y.; Tashiro, H.; Nishi, S.; Hiehata, N.; Nagano, A.J.; Fukuda, S. Genetic diversity of loquat (*Eriobotrya japonica*) revealed using RAD-Seq SNP markers. *Sci. Rep.* **2022**, *12*, 10200. [CrossRef]
59. Bernard, A.; Barreneche, T.; Lheureux, F.; Dirlewanger, E. Analysis of genetic diversity and structure in a worldwide walnut (*Juglans regia* L.) germplasm using SSR markers. *PLoS ONE* **2018**, *13*, e0208021. [CrossRef]
60. Wen, Y.; Han, W.; Wu, S. Plant genetic diversity and its influencing factors. *J. Cent. South Univ. For. Technol.* **2010**, *30*, 80–87.
61. Slatkin, M.; Barton, N.H. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **1989**, *43*, 1349–1368. [CrossRef] [PubMed]
62. Torokeldiev, N.; Ziehe, M.; Gailing, O.; Finkeldey, R. Genetic diversity and structure of natural *Juglans regia* L. populations in the southern Kyrgyz Republic revealed by nuclear SSR and EST-SSR markers. *Tree Genet. Genom.* **2019**, *15*, 5. [CrossRef]
63. Wang, H.; Pan, G.; Ma, Q.G.; Zhang, J.P.; Pei, D. The genetic diversity and introgression of Juglans regia and *Juglans sigillata* in Tibet as revealed by SSR markers. *Tree Genet. Genom.* **2015**, *11*, 1. [CrossRef]
64. Filippi, C.V.; Aguirre, N.; Rivas, J.G.; Zubrzycki, J.; Puebla, A.; Cordes, D.; Moreno, M.V.; Fusari, C.M.; Alvarez, D.; Heinz, R.A.; et al. Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. *BMC Plant Biol.* **2015**, *15*, 52. [CrossRef]
65. Yin, J.; Zhao, H.; Wu, X.; Ma, Y.; Zhang, J.; Li, Y.; Shao, G.; Chen, H.; Han, R.; Xu, Z. SSR marker based analysis for identification and of genetic diversity of non-heading Chinese cabbage varieties. *Front. Plant. Sci.* **2023**, *14*, 1112748. [CrossRef]
66. Hamrick, J.L.; Godt, M.J.W.; Shermanbroyles, S.L. Factors influencing levels of genetic diversity in woody plant-species. In Proceedings of the International Symp on Population Genetics of Forest Trees, Corvallis, OR, USA, 31 July–2 August 1992.
67. Mason, A.S. SSR genotyping. *Methods Mol. Biol.* **2015**, *1245*, 77–89.
68. Carneiro Vieira, M.L.; Santini, L.; Diniz, A.L.; Munhoz, C.d.F. Microsatellite markers: What they mean and why they are so useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [CrossRef]
69. Pollegioni, P.; Woeste, K.; Mugnozza, G.S.; Malvolti, M.E. Retrospective identification of hybridogenic walnut plants by SSR fingerprinting and parentage analysis. *Mol. Breed.* **2009**, *24*, 321–335. [CrossRef]
70. Gunn, B.F.; Aradhya, M.; Salick, J.M.; Miller, A.J.; Yang, Y.; Liu, L.; Hai, X. Gentic variation in walnuts (*Juglans regia* and *J.* sigillata; Juglandaceae): Species distinctions, human impacts, and the conservation of agrobiodiversity in Yunnan, China. *Am. J. Bot.* **2010**, *97*, 660–671. [CrossRef]
71. Victory, E.R.; Glaubitz, J.C.; Rhodes, O.E.; Woeste, K.E. Genetic homogeneity in Juglans nigra (*Juglandaceae*) at nuclear microsatellites. *Am. J. Bot.* **2006**, *93*, 118–126. [CrossRef]
72. Vahdati, K. Traditions and Folks for Walnut Growing around the Silk Road. In Proceedings of the 1st International Symposium on Fruit Culture and Its Traditional Knowledge along Silk Road Countries, Yerevan, Armenia, 4–8 November 2014.
73. French, D. Pre and early Roman roads of Asia Minor: The Persian royal road. *Iran* **1998**, *36*, 15–43. [CrossRef]
74. Khanazarov, A.A.; Chernova, G.M.; Venglovskiy, B.I.; Ashimov, K.S.; Kenjabaev, S.K. Man-made green monuments of Central Asia: Some examples of Uzbekistan. *Schweizerische Zeitschrift fur Forstwesen* **2003**, *154*, 226–228. [CrossRef]
75. Aubaile, F. Pathways of diffusion of some plants and animals between Asia and the Mediterranean region. *Rev. d'Ethnoécologie* **2012**, *1*, 1–26. [CrossRef]
76. Dehghan, B.; Vahdati, K.; Rezaee, R.; Hassani, D. Persian walnut (*Juglans regia* L.) grafting as influenced by different bench grafting methods and scion cultivars. *J. Appl. Hortic.* **2009**, *11*, 56–58. [CrossRef]
77. Pollegioni, P.; Woeste, K.E.; Chiocchini, F.; Del Lungo, S.; Olimpieri, I.; Tortolano, V.; Clark, J.; Hemery, G.E.; Mapelli, S.; Malvolti, M.E. Ancient Humans Influenced the Current Spatial Genetic Structure of Common Walnut Populations in Asia. *PLoS ONE* **2015**, *10*, e0135980. [CrossRef]
78. Miller, A.J.; Gross, B.L. From forest to field: Perennial fruit crop domestication. *Am. J. Bot.* **2011**, *98*, 1389–1414. [CrossRef]
79. Harris, D.R. Domestication of plants in the old world: The origin and spread of cultivated plants in west Asia, Europe and the Nile valley. *Agr. Hist. Rev.* **2001**, *49*, 226–227.

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.