horticulturae

MDPI

Article

Genome Wide Characterization, Comparative and Genetic Diversity Analysis of Simple Sequence Repeats in Cucurbita Species

Lei Zhu [1,2], Huayu Zhu [1], Yanman Li [1], Yong Wang [1], Xiangbin Wu [2], Jintao Li [2], Zhenli Zhang [2], Yanjiao Wang [2], Jianbin Hu [1,2], Sen Yang [1], Luming Yang [1,2,*] and Shouru Sun [1,2,*]

[1] Henan Key Laboratory of Fruit and Cucurbit Biology, Henan Agricultural University, Zhengzhou 450002, China; zhulei2018@henau.edu.cn (L.Z.); hyzhu@henau.edu.cn (H.Z.); liyanman2004@163.com (Y.L.); yongwang@henau.edu.cn (Y.W.); jianbinhu@henau.edu.cn (J.H.); senyang1989@henau.edu.cn (S.Y.)

[2] International Joint Laboratory of Horticultural Biology, College of Horticulture, Henan Agricultural University, 95 Wenhua Road, Zhengzhou 450002, China; wxb15993635378@163.com (X.W.); l18739921624@163.com (J.L.); Zhangzl15225961525@163.com (Z.Z.); jiaoblue123@163.com (Y.W.)

* Correspondence: lumingyang@henau.edu.cn (L.Y.); ssr365@henau.edu.cn (S.S.)

Abstract: Simple sequence repeats (SSRs) are widely used in mapping constructions and comparative and genetic diversity analyses. Here, 103,056 SSR loci were found in *Cucurbita* species by in silico PCR. In general, the frequency of these SSRs decreased with the increase in the motif length, and di-nucleotide motifs were the most common type. For the same repeat types, the SSR frequency decreased sharply with the increase in the repeat number. The majority of the SSR loci were suitable for marker development (84.75% in *Cucurbita moschata*, 94.53% in *Cucurbita maxima*, and 95.09% in *Cucurbita pepo*). Using these markers, the cross-species transferable SSR markers between *C. pepo* and other Cucurbitaceae species were developed, and the complicated mosaic relationships among them were analyzed. Especially, the main syntenic relationships between *C. pepo* and *C. moschata* or *C. maxima* indicated that the chromosomes in the *Cucurbita* genomes were highly conserved during evolution. Furthermore, 66 core SSR markers were selected to measure the genetic diversity in 61 *C. pepo* germplasms, and they were divided into two groups by structure and unweighted pair group method with arithmetic analysis. These results will promote the utilization of SSRs in basic and applied research of *Cucurbita* species.

Keywords: pumpkin; simple sequence repeat (SSR); syntenic relationship; cross-species markers; population structure

Citation: Zhu, L.; Zhu, H.; Li, Y.; Wang, Y.; Wu, X.; Li, J.; Zhang, Z.; Wang, Y.; Hu, J.; Yang, S.; et al. Genome Wide Characterization, Comparative and Genetic Diversity Analysis of Simple Sequence Repeats in Cucurbita Species. *Horticulturae* **2021**, *7*, 143. https://doi.org/10.3390/horticulturae7060143

Academic Editor: Yuyang Zhang

Received: 19 May 2021
Accepted: 5 June 2021
Published: 8 June 2021

1. Introduction

The *Cucurbita* genus (2n = 2x = 40), belonging to the Cucurbitaceae family, contains more than 13 species [1]. Most *Cucurbita* species are wild resources, and only three domesticated species, *Cucurbita maxima*, *Cucurbita moschata*, and *Cucurbita pepo*, are widely cultivated and have become important food crops globally [2]. At present, Asia has the largest pumpkin cultivation area, and China is the main producer of pumpkins. In 2012, the planting area of pumpkin was approximately 3.8×10^4 Hm^2 in China, and the total output reached 7.0×10^6 tons (http://www.fao.org/faostat/zh/#data/QC/visualize, 2020). Due to the fact of their long history of cultivation and domestication, *Cucurbita* species show a greater diversity in fruit shape, size, and color than other Cucurbitaceae species [3]. Furthermore, *Cucurbita* species have strong roots and exhibit good adaption to different biotic and abiotic stresses, such as cold, viruses, and salinity, and so they are widely used as rootstocks in grafting [4,5]. Although they are a common global crop, fundamental

genetic research on *Cucurbita* is lacking, and few studies have been conducted to improve the cultivation and breeding of this genus.

Simple sequence repeats (SSRs) are widely used in genetic mapping constructions, genetic diversity analyses, and genome-wide association studies due to the fact of their relative abundance, multi-allelism, co-dominance, and low cost [6,7]. In the Cucurbitaceae family, the whole-genome sequencing of *Cucumis sativus*, *Cucumis melon*, and *Citrullus lanatus* has been completed [8–10], and genome-wide SSR markers have been characterized and developed in these crops, which has greatly promoted their application in gene and quantitative trait locus (QTL) mapping as well as in comparative genomics [11–13]. A rough syntenic relationship between melon (2n = 2x = 24) and cucumber (2n = 2x = 14) chromosomes was revealed by comparative mapping using 199 SSR markers developed from cucumber [14]. Later, Yang et al. (2014) developed a higher density map of *Cucumis hystrix* containing 416 SSR markers, and 151 and 50 markers were derived from cucumber and melon, respectively. With these shared markers among the three *Cucumis* species, the chromosome-level syntenic relationships were well established, which was further confirmed by fluorescence in situ hybridization (FISH) [15]. Ninety-one syntenic blocks were divided between cucumber and melon, and 53 syntenic blocks were identified between cucumber and *Cucumis hystrix*. Furthermore, the genome-wide SSR markers developed from melon and watermelon have made it possible to more clearly define chromosomal syntenic relationships, and the complicated mosaic patterns of chromosome synteny between melon, watermelon, and cucumber have been well established based on cross-species SSR markers [12,15]. However, the syntenic relationships and chromosomal rearrangements between *Cucurbita* species and other Cucurbitaceae crops are still largely fragmented and incomplete.

Based on the conserved sequences among species or genera, some amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and SSR markers were developed in previous studies [16–20]. However, these restricted markers are insufficient for research on genetic diversity, genetic mapping, and comparative genomics. Esteras et al. (2012) constructed the first genetic map in pumpkin using 304 single nucleotide polymorphisms (SNPs) and 11 SSR markers and found that the linkage groups of pumpkin were partially homoeologous to cucumber chromosomes. The applications of these expressed sequence tag (EST)-SNP markers are still greatly limited due to the small numbers of markers, the high cost of enzymes, and the complicated operating procedure [21]. Due to the lack of genome-wide coverage and polymorphic markers, in-depth application and comparative analysis still cannot be conducted. With the development of high-throughput sequencing technology, there has been an increase in studies on *Cucurbita*, and the whole-genome sequences of three important cucurbit crops have become available. Based on these SNPs' data, the whole-genome synteny analysis indicated that both the *C. maxima* and *C. moschata* genomes underwent a whole-genome duplication (WGD) event and that pairs of *C. maxima* (or *C. moschata*) homoeologous regions are shared between chromosomes corresponding to the two sub-genomes [22]. Montero et al. also identified that the covered regions in most of the *C. pepo* genome had experienced a WGD event [23]. Furthermore, some transcriptomes of *Cucurbita* species have become available, and EST-SSRs were developed from them [24–28]. To date, the development of SSR markers in *Cucurbita* species is still limited.

The whole-genome sequences of *C. moschata*, *C. maxima*, and *C. pepo* have been completely assembled, which will greatly promote the large-scale development of SSR markers, allowing for the construction of high-resolution maps, gene mapping, and genome-wide association studies (GWAS). In this study, we conducted a genome-wide identification of SSR motifs in three *Cucurbita* species, analyzed the distribution and frequency of different repeat types, identified cross-species transferable SSR markers by in silico PCR analysis, and studied the chromosome synteny of *C. pepo* with other Cucurbitaceae crops. In addition, 66 core SSR markers were identified in *Cucurbita* genomes and used to evaluate the genetic diversity and population structure of 61 *C. pepo* germplasms. Our study will

be useful for research on the population structure, genetic diversity, molecular-assisted selection, and map-based cloning in *Cucurbita* species.

2. Methods

2.1. Plant Materials

All of the pumpkin accessions used in this study were introduced from the National Crop Germplasm Resource Platform of China (platform of vegetable germplasm resources) in 2018. Four of the accessions came from Russia, one from America, and 56 accessions were from 17 provinces in China. The number and sources are shown in Table S1.

2.2. Genome SSR Identification and Development in Cucurbita Genomes

The genome information of watermelon, melon, cucumber, and pumpkin was downloaded from http://cucurbitgenomics.org/ (2020). To develop a set of higher polymorphic SSR primers for the future study, the criteria used for microsatellite identification in this study was from 2 to 8 bp, and mononucleotides were not considered due to the difficulty in distinguishing bona fide microsatellites from sequencing or assembly error. The microsatellite identification tool (MISA) was used to identify and analyze SSR markers including perfect and compound microsatellites. The specific screening details were as follows: repeats with a minimum length of 18 bp (for di- and tetra-nucleotides), 20 bp (for penta-nucleotides), 24 bp (for hexa-nucleotides), 21 bp (for hepta-nucleotides), and 24 bp (for octa-nucleotides). The oligonucleotide primers for these SSRs were designed according to the flanking genomic sequence using Primer3 software (v.1.1.4). Primers were designed to generate amplicons of 100–300 bp in length with the following minimum, optimum, and maximum values for Primer3 parameters: primer length (bp): 18–20–24; Tm (°C): 50–55–60. Other parameters used the default program values.

2.3. In Silico PCR and Synteny Analysis of Cross-Species SSR Markers

Using the SSR markers from pumpkin (*C. pepo* MU-CU-16) genome as a reference, we comparatively analyzed the genome SSR information of cucumber (Gy14), melon (DH92), watermelon (97103), *C. moschata* cv. Rifu, and *C. maxima* cv. Rimu. This was performed with a custom Perl script that used the NCBI BLASTN program as a search engine with an expected value of 10 and filtering. We allowed up to five nucleotide mismatches at the 5′-end of the primer, no mismatches at the 3′-end, and a minimum of 90% overall match homology. To establish the syntenic relationships of chromosomes between *C. pepo* with *C. sativus*, *C. lanatus*, *C. melo*, *C. maxima*, and *C. moschata*, we discarded these SSR markers with multiply physical locations in the same genome, only retaining the SSR markers in the genomes which had a single in silico PCR product. In addition, these shared SSR markers located on the unanchored scaffolds of the chromosome were further filtered. The SSR marker-based syntenic relationships were finally visualized with visualization blocks in Circos software v.0.55 [29].

2.4. Genomic DNA Extraction, PCR Amplification, and Electrophoresis Detection

Genomic DNA of all the materials was extracted using 1 g of young leaf sample with the cetyl trimethyl ammonium bromide (CTAB) method [30]. The extracted DNA was dissolved in 1× Tris-EDTA (TE) buffer (Solarbio, Cat: T1121). The concentration and purity were detected by the Nanodrop-2000 nucleic acid analyzer. The extracted DNA was diluted to 30 ng/μL as working solution and kept at 4 °C.

Each PCR reaction contained 1 μL of template DNA, 0.5 μM each of forward and reverse primers, 5 μL mastermix (GenStar, Cat: A012-105), and 3 μL ddH_2O. The amplification was carried out as follows: An initial denaturing step at 95 °C for 5 min, 94 °C for 30 s, followed by 6 cycles of 68–58 °C for 45 s. Each cycle was reduced by 2 °C, each annealing time was 1 min, and 72 °C for 1 min; 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min. In the last cycle, primer extension was performed at 72 °C for 10 min.

The PCR products were analyzed by 9% polyacrylamide gel electrophoresis, and a 100 bp DNA ladder was used as the reference marker. After electrophoresis, silver staining was performed to display the PCR products, and photos were taken for preservation.

2.5. Calculation of Clustering

The heterozygosity (He), observer gene number (Na), effective alleles (Ne), observed heterozygosity (Ho), and the Shannon–Weaver index (I) were calculated using Pop-gen software v.1.32 (Canada, University of Alberta). Polymorphic information content (PIC) of SSR markers was computed using EXCEL (China, WPS of JINSHAN). When the PIC of an SSR marker was below 0.25, it was considered as a low polymorphic marker, and a marker was considered highly polymorphic if its PIC was above 0.5.

These amplification bands of each SSR primer were separated using polyacrylamide gel-electrophoresis. The band patterns were visualized with silver staining, and gel images were taken with a digital camera. In the same location, the presence of a band was marked as "1", the absence of a band was marked as "0", and a missing band was marked as "−1". In this study we used Genalex-6 software [31] to conduct the matrix calculation of SSR marker data which had been assigned a value, then transformed it into a triangle matrix, saved it as a mega-file, finally, imported the mega-file into the Mega-6.0 software (USA, Tamura, K team), and selected the unweighted pair group method with arithmetic (UPGMA) algorithm in the "phylogeny" dropdown menu to draw the cluster diagram [32].

The software Structure v.2.3 (USA, UChicago; Britain, Oxon) was used to analyze the population structure [33,34]. An admixture model and correlated allele frequencies were used to estimate the number of the populations. For each of the K-values (ranging from 1 to 5), ten independent runs were performed with a burn-in period of 100,000 followed by 500,000 Markov chain Monte Carlo runs. The optimal K-values depends on the peak of K = mean (|Ln"P(D)|)/(sdLnP(D)). Based on the structure results, the most probable K-value was analyzed using Structure Harvester (http://taylor0.biology.ucla.edu/struct_harvest/, 2020).

3. Result

3.1. The Frequency and Distribution of Different SSR Types in Cucurbita Genomes

A total of 103,056 microsatellite sequences were identified in the *Cucurbita* genome, including 34,375 SSR loci in the 269.9 Mb draft genome sequence of *C. moschata* cv. Rifu, 30,577 SSR loci in the 271.4 Mb draft genome sequence of *C. maxima* cv. Rimu, and 38,104 SSR loci in the 263 Mb draft genome sequence of *C. pepo* MU-CU-16 (Table S2). *Cucurbita pepo* had the largest number of markers with the smallest reference genome size, indicating the highest average density of markers (145 SSR/Mb). To obtain more information, we used *C. pepo* with a higher marker density as the control for the following comparative genomic analysis.

Here, we analyzed repeat types ranging from di-nucleotide to octa-nucleotide. Among all of these nucleotide motifs, di-nucleotide motifs (41.0%) were the most common type, accounting for 41.78%, 39.90%, and 41.01% of the total SSR loci discovered in *C. moschata*, *C. maxima*, and *C. pepo*, respectively, followed by tri-nucleotide motifs (16.97%, 19.19%, and 17.88%, respectively), whereas octa-nucleotide motifs (3.78%, 3.76%, and 3.38%, respectively) were the least represented repeat type in the three *Cucurbita* genomes (Table S2). In general, the frequency of the total SSR loci decreased with the increase in motif length, except for hepta-nucleotide SSRs.

We further examined the distribution of SSR motifs with regard to their repeat numbers (Figure 1). For all the repeat types, with an increase in the repeat number, the SSR frequency decreased sharply, and this change was more obvious in the longer SSR motifs (Figure 1). Consequently, the mean repeat numbers in the di-nucleotides were the highest of all of the repeat types. The analysis of individual SSR types revealed that some specific motifs were more prevalent than others in each class (Figure S1). For example, the AT motif was the most frequent di-nucleotide type in all three genomes, accounting for 31.61%

(in *C. moschata*), 28.81% (in *C. maxima*), and 30.45% (in *C. pepo*) of the total di-nucleotide loci. Similarly, the AAT, AAAT, AAAAT, AAAAAT, AAAAAAT, and AAAAAAAT motifs (AATAATAT motif in *C. maxima*) were the most frequent types in each class. These results indicated that AnT-rich motifs were the most abundant in all SSR motifs in the *C. moschata*, *C. maxima*, and *C. pepo* genomes.

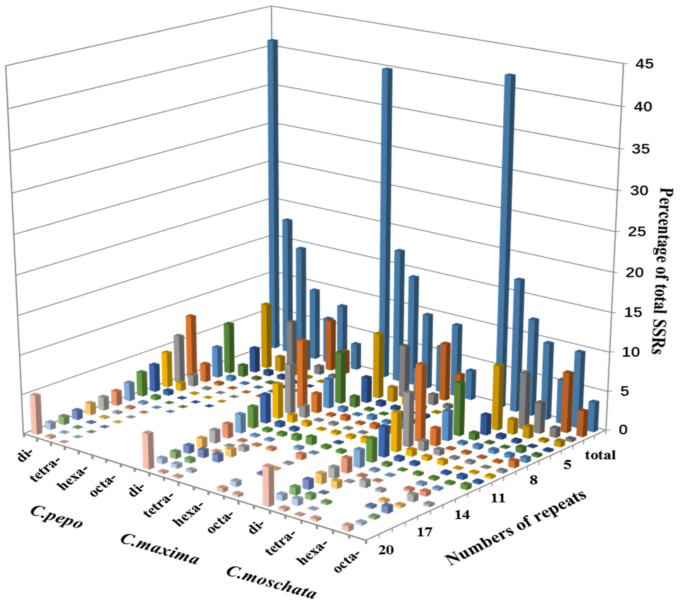

Figure 1. Distribution of SSR motif repeat numbers and relative frequency in *Cucurbita* genome. The vertical axis shows the abundance of microsatellites that have different motif repeat numbers (from 3 to >20) with different colors.

We also investigated the SSR density in each chromosome of the three *Cucurbita* species and found that the density of microsatellite loci was not correlated with the chromosome size (Table S3). For example, in the *C. moschata* genome, the SSR density of the longest chromosome (Chr04) had a medium density of SSRs, while Chr02, which is much shorter than Chr04, had the highest SSR density. A similar trend was also observed in the other two genomes, indicating that the distribution of SSRs was uneven in the *Cucurbita* chromosomes. To better understand the distributions of different SSR motifs, we further checked their frequencies on each chromosome (Figure 2). Our results showed that the distribution of different SSR types on the chromosomes corresponded with their frequencies and SSR density in the *Cucurbita* whole genomes.

The genomic sequences containing these microsatellites were screened for PCR primer design, and 94,272 SSR microsatellite loci were found to contain suitable flanking sites for SSR primer design. While *C. moschata* had the lowest proportion of SSRs suitable for primers design (84.75%), the percentages in *C. maxima* and *C. pepo* reached 94.53% and 95.09%, respectively (Table S2). Though the di-nucleotide repeat types were the most frequent in all three genomes, they did not exhibit good performance in primer design. Interestingly, the hexata-nucleotide repeat types had the highest ratio of SSRs suitable for primer design in all three genomes, followed by penta-nucleotide repeat types, indicating that the longer motifs were more suitable for primer design in *Cucurbita* species. Finally, a total of 91,248 SSR primers (28,194 in *C. moschata*, 28,061 in *C. maxima*, and 34,993 in *C. pepo*) were designed, with some primers including more than one SSR locus as the compound SSR (Tables S4–S6).

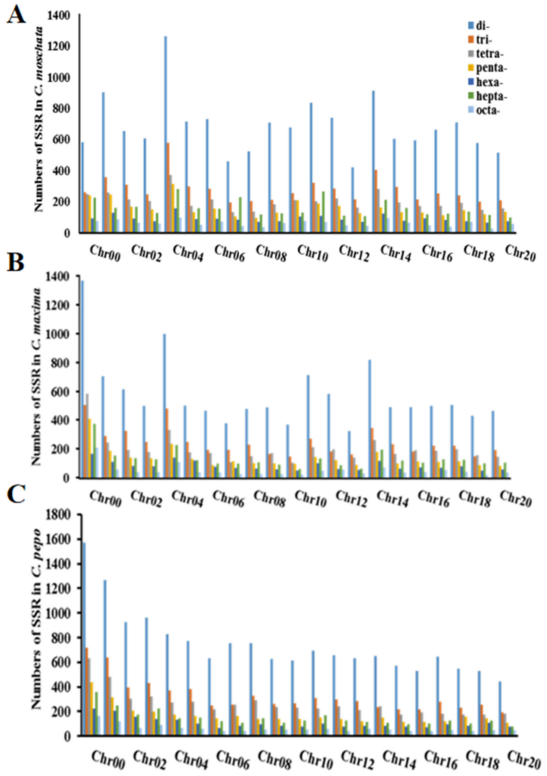

Figure 2. (**A**) The distribution of SSR repeat types on each chromosome in *C. moschata*. (**B**) The distribution of SSR repeat types on each chromosome in *C. maxima*. (**C**) The distribution of SSR repeat types on each chromosome in *C. pepo*. The vertical axis shows the number of microsatellites from di-nucleotide to octo-nucleotide which are discriminated by different colors. The horizontal axis shows different chromosomes of *C. ssp*, and LG00 means all the chromosome unanchored scaffolds.

3.2. Chromosome Synteny Relationships of C. pepo with Other Cucurbitaceae Species

In order to understand the universality and correlation of SSR markers among Cucurbitaceae crops, we compared and analyzed the cross-species SSR markers between *C. pepo* and other Cucurbitaceae species by in silico PCR. We identified 391 cross-species SSR markers between *C. pepo* and *C. sativus*, 425 cross-species SSR markers between *C. pepo* and *C. melo*, 717 cross-species SSR markers between *C. pepo* and *C. lanatus*, 11,732 cross-species SSR markers between *C. pepo* and *C. maxima*, and 15,274 cross-species SSR markers between *C. pepo* and *C. moschata* (Tables S7–S11). The collinear blocks to inversion blocks ratio was 26:26 between the *C. pepo* and *C. sativus* genomes, 25:36 between the *C. pepo* and *C. melo* genomes, 51:38 between the *C. pepo* and *C. lanatus* genomes, 154:158 between the *C. pepo* and *C. maxima* genomes, and 153:152 between the *C. pepo* and *C. moschata* genomes. Interestingly, the ratio of collinear blocks to inversion blocks was nearly 1:1 among the three *Cucurbita* species. Each *C. pepo* chromosome shared 3–36 SSR markers with *C. sativus*, *C. lanatus*, or *C. melo*. However, most of the *C. pepo* chromosome shared a larger number of SSR markers (3-1,436) with *C. maxima* or *C. moschata*. The *C. pepo* syntenic block, CpeCma7, had the largest number of shared SSR markers (i.e., 296) between *C. pepo* chromosome Cpe1 and *C. maxima* chromosome Cma4.

The physical positions of those common shared markers were compared. The main syntenic relationships between *C. pepo* and other Cucurbitaceae species are listed in Table 1, and the syntenic relationships visualized for *C. pepo* with *C. lanatus*, *C. melo*, and *C. sativus* are shown in Figure 3. The main syntenic relationships among the chromosomes revealed complex mosaic patterns. In Figure 3, each *C. pepo* chromosome was syntenic to more than two chromosomes in other Cucurbitaceae species. The *C. pepo* chromosomes Cpe9 and Cpe16 had the simplest syntenic pattern with watermelon, and each of them was mainly syntenic to one watermelon chromosome (Table 1). Cpe9 was syntenic to watermelon

chromosome W5, and 14 commonly shared SSR markers were found between Cpe9 and W5. From the markers CpeSSR15544 to CpeSSR16107, there were three blocks belonging to watermelon chromosome W5, and each block contained at least four SSR markers. According to the continuous physical positions of these markers on both of the reference genomes, the syntenic blocks CpeWM37 and CpeWM38 showed an inversion pattern, and the syntenic block CpeWM39 showed a collinear pattern between *C. pepo* and *C. lanatus*. Similar comparisons were carried out between *C. pepo* and *C. sativus* or *C. pepo* and *C. melo* using the cross-species SSR markers. The *C. pepo* chromosomes Cpe7, Cpe8, Cpe11, and Cpe20 had the simplest syntenic pattern with *C. sativus*, and each of them was only syntenic to one cucumber chromosome. Meanwhile, the simplest syntenic patterns between *C. pepo* and *C. melo* were mainly found on chromosomes Cpe15, Cpe18, Cpe19, and Cpe20. The most complicated syntenic pattern was found on *C. pepo* chromosome Cpe1, which corresponded to five chromosomes of *C. moschata*, four chromosomes of *C. maxima*, seven chromosomes of *C. lanatus*, three chromosomes of *C. sativus*, and five chromosomes of *C. melo*.

Table 1. The main syntenic relationships of *C. pepo* with other Cucurbitaceae species.

C. pepo	*C. moschata*	*C. maxima*	*C. lanatus*	*C. sativus*	*C. melo*
Cpe1	Cmo3(4), Cmo4(1,436), Cmo9(5), Cmo10(3), Cmo17(24)	Cma3(5), Cma4(1,103), Cma9(9), Cma17(19)	Cla1(5), Cla5(21), Cla6(4), Cla7(13), Cla8(5), Cla10(3), Cla11(14)	Csa3(6), Csa5(28), Csa6(3)	Cme3(4), Cme6(4), Cme7(3), Cme9(5), Cme10(16)
Cpe2	Cmo1(913), Cmo10(3), Cmo18(12)	Cma1(674), Cma18(8)	Cla5(8), Cla7(8), Cla10(13), Cla11(7)	Csa3(7), Csa4(6), Csa7(3)	Cme1(3), Cme4(10), Cme7(9)
Cpe3	Cmo4(3), Cmo14(1,080)	Cma4(4), Cma14(859)	Cla5(24), Cla7(8), Cla10(36)	Csa3(34), Csa4(4)	Cme4(25), Cme6(17), Cme7(6)
Cpe4	Cmo10(3), Cmo11(822)	Cma11(640)	Cla2(6), Cla3(5), Cla6(12), Cla10(3)	Csa1(15), Csa3(3)	Cme2(11)
Cpe5	Cmo2(904), Cmo10(8)	Cma2(692), Cma10(8)	Cla1(8), Cla2(20), Cla9(5)	Csa3(5), Csa5(8), Csa6(8)	Cme4(3), Cme5(4), Cme9(6), Cme11(10)
Cpe6	Cmo9(551)	Cma9(396)	Cla5(12), Cla8(4), Cla9(5), Cla11(9)	Csa3(6), Csa4(3)	Cme4(6), Cme7(9)
Cpe7	Cmo5(3), Cmo12(587), Cmo14(20)	Cma5(4), Cma12(452)	Cla2(3), Cla8(6)	Csa2(5)	Cme3(5), Cme5(3)
Cpe8	Cmo6(785)	Cma6(431)	Cla5(13), Cla10(16)	Csa3(6)	Cme4(10), Cme6(6)
Cpe9	Cmo18(544), Cmo19(6)	Cma2(3), Cma18(440)	Cla5(14)	Csa1(4), Csa3(5), Csa5(5)	Cme6(4), Cme10(3), Cme12(4)
Cpe10	Cmo3(659), Cmo18(5)	Cma3(547), Cma18(6)	Cla1(23), Cla4(13)	Csa4(3), Csa6(19)	Cme8(20)
Cpe11	Cmo5(707), Cmo10(3)	Cma5(574)	Cla2(4), Cla8(17), Cla11(14)	Csa2(14)	Cme3(11), Cme5(9)
Cpe12	Cmo17(665)	Cma17(516)	Cla6(12), Cla9(15)	Csa6(9), Csa7(14)	Cme1(17), Cme11(6)
Cpe13	Cmo8(9), Cmo15(649)	Cma4(3), Cma8(7), Cma15(468)	Cla1(21), Cla8(11), Cla11(8)	Csa2(3), Csa5(13), Csa6(6)	Cme3(8), Cme9(12)
Cpe14	Cmo16(565)	Cma16(409)	Cla5(4), Cla7(17), Cla10(14)	Csa3(15), Csa4(8)	Cme6(13), Cme7(5)
Cpe15	Cmo19(493)	Cma19(356)	Cla2(16), Cla7(4), Cla9(11)	Csa3(4), Csa7(11)	Cme1(12)
Cpe16	Cmo20(526)	Cma20(394)	Cla2(23)	Csa2(4), Csa6(4)	Cme5(5), Cme11(10)
Cpe17	Cmo4(3), Cmo8(634), Cmo9(7), Cmo14(3), Cmo17(3)	Cma8(472), Cma14(3), Cma17(4)	Cla6(11), Cla9(17)	Csa6(3), Csa7(3)	Cme1(8), Cme11(4)
Cpe18	Cmo10(500), Cmo14(3)	Cma10(354), Cma14(3)	Cla3(8), Cla6(22)	Csa1(16)	Cme2(17)
Cpe19	Cmo7(658)	Cma7(462)	Cla1(20), Cla4(7)	Csa4(4), Csa6(16)	Cme8(23)
Cpe20	Cmo13(468)	Cma13(330)	Cla1(4), Cla3(12), Cla4(4)	Csa1(11)	Cme12(8)

The syntenic relationships among different *Cucurbita* species were simple and clear. For instance, each of the 20 chromosomes in *C. pepo* was mainly syntenic with one chromosome in *C. moschata* or *C. maxima* (Figure 4), implying that the chromosomes in the *Cucurbita* genomes were highly conserved during evolution. Our results also showed that there were three main relationship patterns among the *C. pepo*, *C. maxima*, or *C. moschata* genomes, including (1) the eleven linear relationship chromosomes between *C. pepo* and *C. maxima* or *C. moschata* such as Cpe2–Cmo1–Cma1. Most of the cross-markers in the corresponding chromosomes showed collinear patterns. (2) There were eight inverted

relationship chromosomes between *C. pepo* and *C. maxima* or *C. moschata*. For example, the chromosome Cpe1 of *C. pepo* was inverted to the chromosome Cmo4 of *C. moschata* and Cma4 of *C. maxima*. (3) There was a mosaic pattern between *C. pepo* and *C. maxima* or *C. moschata*, for example, Cpe4–Cmo11–Cma11.

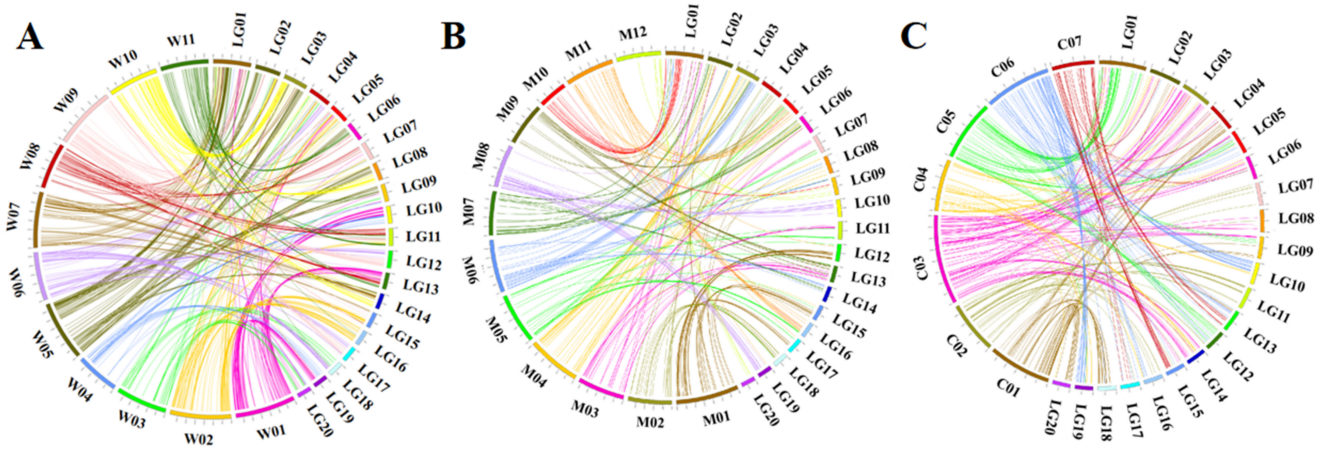

Figure 3. Syntenic relationships of *C. pepo* with (**A**) *C. lanatus*, (**B**) *C. melo*, and (**C**) *C. sativus*. Chromosome synteny between *C. pepo* and *C. sativus* was based on 391 cross-species markers; synteny between *C. pepo* and *C. melo* was based on 425 cross-species markers; synteny between *C. pepo* and *C. lanatus* was based on 717 cross-species markers. W1–W11 represent *C. lanatus*' eleven chromosomes, M01–M12 represent *C. melo*'s twelve chromosomes, C01–C07 represent *C. sativus*'s seven chromosomes, and LG01–LG20 represent *C. pepo*'s twenty chromosomes. Syntenic blocks are connected by the same color lines from *C. pepo* chromosomes.

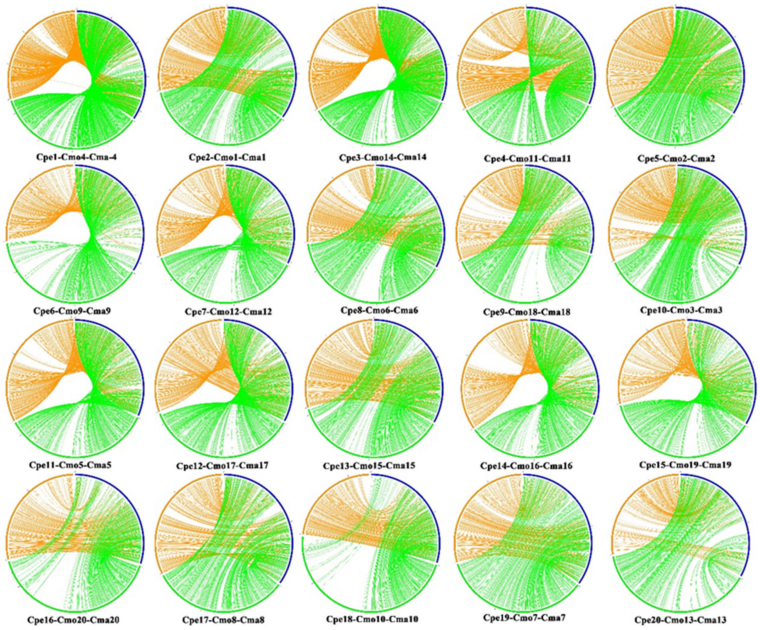

Figure 4. Chromosome synteny of *C. pepo* (blue) with *C. moschata* (green) and *C. maxima* (yellow). The physical positions of chromosomes of each crop in the figure are arranged clockwise. Chromosome synteny between *C. pepo* and *C. moschata* was based on 14,276 cross-species markers; synteny between *C. pepo* and *C. maxima* was based on 10,655 cross-species markers. Cpe1–Cpe20 represent *C. pepo*'s twenty chromosomes, Cmo1–Cmo20 represent *C. moschata*'s chromosomes, and Cma1–Cma20 represent *C. maxima* chromosomes. The syntenic relationship between *C. pepo* and *C. moschata* are connected with the green color lines, and the syntenic relationship between *C. pepo* and *C. maxima* are connected with the yellow color lines.

3.3. The Genetic Diversity and Population Structure Analysis of the C. pepo Germplasm

In our preliminary study, approximately 400 SSR markers were screened using 61 accessions of *C. pepo* germplasm. Finally, a total of 66 core SSR markers were selected based on the allelic number, the genomic coverage, and the efficiency of PCR amplification (Table S12). These markers exhibited clear band spectrums and were evenly distributed on the chromosomes. In this study, 276 alleles were detected by the 66 SSR markers in the 61 *C. pepo* accessions with an average of 4.18 loci per SSR marker. The number of Na ranged from two to nine. The highest number of Na was nine, which was detected by SSR010246, SSR026560, SSR026918, SSR027656, and SSR026980, followed by SSR011546, SSR003315, and SSR026797 with eight alleles. The number of Ne varied from 1.03 to 6.07 with an average of 2.31. The SI ranged from 0.083 to 1.96 with an average of 0.83. The PIC value ranged from 0.03 to 0.83 with an average of 0.43 (Table S13).

We further used a model-based approach for population structure analysis of the 61 *C. pepo* accessions. According to the results of the structural operation, when K = 2, ΔK showed a significant peak value, indicating that the 61 accessions used in this study could be obviously divided into two groups (Figure S2), named group I and group II. The five *C. pepo* subsp. *ovifer* accessions (2, 29, 30, 31, and 45) were clustered into group I (8.20%), and all of them were wild materials. Most of the *C. pepo* subsp. *pepo* accessions were clustered into group II (91.80%), which were all cultivated materials (Figure 5A). This indicated that the SSR markers we used could clearly distinguish the cultivated materials from the wild materials. The backgrounds of the cultivated accessions were narrow, except for accession 45 in group I, which should have a complex genetic background, similar to accession 14 and 16 in group II. The UPGMA analysis revealed that the 61 *C. pepo* accessions were divided into two clusters (Figure 5B), which was consistent with their population structure. The five *C. pepo* subsp. *ovifer* accessions were clustered together at the base of the phylogenetic tree, which further supported our population structure analysis.

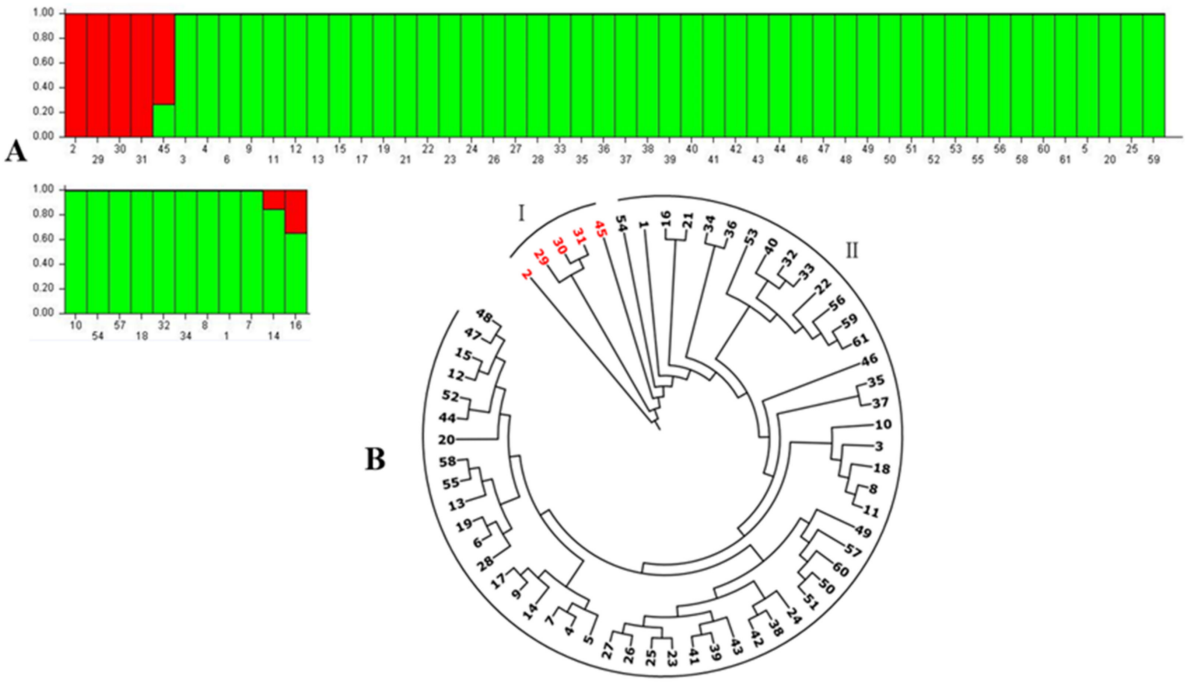

Figure 5. The genetic diversity of the 61 accessions based on SSR markers. (**A**) Population structure of 61 accessions in *C. pepo* by the model-based analysis. The scale of the *y*-axis represents the percentage of genetic components, and the *x*-axis represents the different materials. (**B**) Phylogenetic tree of 61 accessions in the *C. pepo* by UPGMA analysis; Group I (red makers) and Group II correspond to the structure analysis.

4. Discussion

4.1. Frequency, Distribution, and Characterization of Microsatellites in Three Cucurbita Genomes

With the development of sequencing technology, the discovery and mining of genomic SSR loci has successfully been applied in many plant species, such as cotton [35,36], foxtail millet [37], cucumber [11], watermelon [13], tobacco [38], and melon [12]. *Cucurbita moschata*, *C. maxima*, and *C. pepo* are important species that are cultivated worldwide, and their graft genomes were released several years ago. However, there remains little information on the development of genome-wide SSR markers in *Cucurbita* species, which has strongly limited their genetic research. In the present study, genome-wide microsatellites were identified and characterized in the three *Cucurbita* species. A total of 34,375, 30,577, and 38,104 SSR loci were detected in the *C. moschata*, *C. maxima*, and *C. pepo* genomes, respectively. The smallest genome size and maximum number of microsatellites were detected in *C. pepo*, indicating that there was no direct correlation between genome size and the number of microsatellites. The density of the SSR markers in the three *Cucurbita* species was approximately 113–145 SSR/Mb, which is lower than that in cucumber (552 SSR/Mb) but comparable to that in melon (109 SSR/Mb) and watermelon (111 SSR/Mb) [11–13]. In addition to the natural differences among different genomes, many other factors could affect the deviations in SSR density such as the software and parameters used for microsatellite detection. We suspect that the main reason for the difference in SSR density between *Cucurbita* species and cucumber was the different selection criteria for the SSR loci, e.g., the repeat types (di- to octa-nucleotides versus mono- to penta-nucleotides) and the minimum lengths (18 bp versus 12 bp).

We further analyzed the distribution and frequency of microsatellites in the three *Cucurbita* species (Figures 1 and 2). In most cases, a negative correlation was observed between the microsatellite frequency and the number of repeat units. Consistent with previous studies in watermelon and melon, the di-nucleotide repeats were the most abundant SSRs, followed by tri-, tetra-, penta-, hepta-, hexa-, and octo-nucleotide repeats [12,13]. This is something that varies in different species. For example, the density of tetra-nucleotide repeats was highest in *C. sativus* (164.2 SSR/Mb), *Populus trichocarpa* (144.9 SSR/Mb), *Medicago truncatula* (102.8 SSR/Mb), and *Vitis vinifera* (171.3 SSR/Mb), whereas the density of tri-nucleotide repeats was the highest in Arabidopsis thaliana (146.6 SSR/Mb), Glycine max (103.1 SSR/Mb), and Oryza sativa (220.1 SSR/Mb) [11]. Some studies have revealed that the di-nucleotide motifs with high repeat numbers are more abundant and polymorphic compared to those with short repeat units [39]. The reason is that di-nucleotide repeats are much less frequent in coding regions than in non-coding regions [40,41]. It is also reported that the exon region contains more triplet SSRs than other repeats, and triplet SSR motifs may be related to high frequencies of certain amino acids [42,43]. These SSRs in the coding sequence may have the potential to affect all aspects of genetic functions including gene regulation, development, and evolution. However, the function of genes that contain SSRs and the role of these SSR motifs in plant genes are less studied and poorly understood [44]. It is interesting to note that many bacterial SSRs in the intergenic regions have regulatory functions [45], and whether these SSR motifs in the intergenic regions of *Cucurbita* species play a role in specialization or gene regulation should be further studied.

The low number of repeat motifs was predominant, and the AT-rich motifs in particular contributed a large proportion of all types of di-nucleotide repeats in the *Cucurbita* species (Figure S1). The AT or AAT type is more common in dicots [13], which is consistent with our results. Recently, the characterization of SSR markers in bitter gourd showed that the tri-nucleotide repeat units were the main type, with an overrepresentation of A/T, AT/AT, AAT/ATT, and AAAT/ATTT motifs in all kinds of repeat types [46]. This has also been found in other genomes [11,47,48]. On the contrary, the frequency of the GC or CCG type was much lower at the genomic level [49,50], and the GC, TC, or GA types have relatively stable structures. Most of the AT types are distributed in non-genic regions, while the TC/GA types are primarily distributed in coding sequences [38].

4.2. Chromosome Synteny Analysis between C. pepo and Other Cucurbitaceae Species

Chromosome synteny analysis has been conducted in many species, such as cucumber, watermelon, and melon, but few studies have been conducted on the chromosome synteny among different *Cucurbita* species or between *Cucurbita* species and other Cucurbitaceae crops. In this study, the genome-wide SSR development from the three *Cucurbita* genomes provided the possibility to identify their syntenic relationships at a high-resolution level via in silico PCR analysis. Though the sizes of the pumpkin genomes are similar to that of other sequenced Cucurbitaceae species, the number of cross-species SSR markers in the *Cucurbita* genus is much higher. Compared to hundreds of shared markers in previous studies [14], we identified many more cross-species transferable SSR markers in the *Cucurbita* genus that were used for chromosome synteny analysis. The WGD event in *Cucurbita*, which has not been observed in other sequenced Cucurbitaceae species, such as cucumber [8], melon [10], and watermelon [9], may be a possible reason leading to the high abundance of SSR markers.

According to the cross-species transferable SSR markers, 52, 61, and 89 syntenic blocks distributed on all chromosomes were identified between *C. pepo* with cucumber, melon, and watermelon, respectively (Figure 3). Similar homoeologous blocks were detected by whole-genome comparison [22], suggesting that the cross-species transferable SSR markers are useful and reliable in genome comparisons and chromosome synteny analyses. In most cases, there were multiple synteny blocks detected between *C. pepo* and other Cucurbitaceae species due to the fact of chromosome fission. The most complicated syntenic pattern existed on chromosome Cpe1 of *C. pepo*, which was syntenic to seven watermelon chromosomes, indicating that complicated structural changes occurred after their divergence from a common ancestor. The ratio of collinear blocks to inversion blocks was nearly 1:1 in *Cucurbita*, and the reason for this may be that genome duplication and inter-chromosomal exchanges occurred randomly during chromosome evolution.

Based on the cross-species transferable SSR markers, we identified more highly conserved syntenic blocks among *Cucurbita* species than melon, cucumber, or watermelon. We found that each block among three *Cucurbita* species of the same genus contained many more shared common SSR markers, and these homoeologous chromosomes were much conserved, which further confirmed their close evolutionary relationships in the Cucurbitaceae family. For example, the *C. pepo* syntenic block contained more markers than that in melon [12]. Due to the WGD during chromosome evolution and speciation, the number of the chromosomes and cross-markers increased. However, those blocks were highly conserved during chromosome evolution among different Cucurbitaceae species. The chromosomal pair analysis by cross-species SSR markers showed that there were eight large-scale inversions on different chromosomes between *C. pepo* and *C. moschata* or between *C. pepo* and *C. maxima*, indicating that *C. pepo* experienced more complex evolutionary processes (Figure 4). Interestingly, Chr4 contained a mosaic region among *Cucurbita* species. The reason might be due to the fact of genome duplication, large-scale inter-chromosomal exchanges, or long-term evolutionary forces. Whether the partial inversion of chromosome 4 in *C. pepo* will affect the mapping, cloning, and study of some traits is worth exploring in the future.

4.3. The Genetic Diversity and Population Structure of C. pepo Germplasm

Previously, because of the scarcity of genomic sequences, there were limited molecular markers available to study the genetic diversity and population structure of *Cucurbita* species. Though the genetic diversity of *Cucurbita* species has been evaluated using sequence-related amplified polymorphism (SRAP), AFLP, SSR, RAPD, and inter-simple sequence repeats (ISSRs), most of the markers used have high randomness, lack precise location information, and have low genomic coverage and poor polymorphism, which greatly limit their application [18,51,52]. With the draft genome available for three cucurbit crops, we developed 91,248 SSR markers with precise physical locations on chromosomes and evaluated the genetic diversity of 61 pumpkin accessions using 66 core SSR markers.

The population structure of 61 accessions revealed that the background of some materials was mixed between group I and group II, suggesting that these accessions may have undergone gene exchange between two subspecies. The materials were collected from different provinces in China, and they were obviously classified into two subspecies, subsp. *ovifer* (or subsp. *texana*) and subsp. *pepo*, which is consistent with previous studies [21,51]. However, the three subspecies of *C. pepo* classified by Decker are *C. pepo* subsp. *fraterna* (Bailey) Andres, *C. pepo* subsp. *texana* (Scheele) Filov, and *C. pepo* subsp. *pepo* [53]. The putative ancestor for *C. pepo*, namely, subsp. *fraterna* from northeastern Mexico, has been considered a wild gourd [54]. The population structure and UPGMA results indicated that these accessions of *C. pepo* in China come from the common ancestor. Thus, there have great prospects for germplasm improvement.

The *Cucurbita* genus contains several economically important crops, but its breeding has lagged behind the other Cucurbitaceous crops. Limited high-quality cultivars cannot meet the production requirements. Thus, different breeding programs can be facilitated using marker assisted selection. The whole-genome SSR markers detected in this study will promote the development and utilization in basic and applied research of *Cucurbita* species.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/horticulturae7060143/s1, Figure S1: The top five types of each SSR repeat motif and their frequencies in *C. moschata*, *C. maxima*, and *C. pepo*, Figure S2: The optimal K-values analysis by using Structure Harvester, Table S1: The list of the *C. pepo* introduction accessions, Table S2: The distribution of different nucleotide repeats in the genome of three *Cucurbita* species, Table S3: The distribution of SSR loci on different chromosomes in *C. moschata*, *C. maxima*, and *C. pepo*, Table S4: The identified SSR markers in *C. moschata*, Table S5: The identified SSR markers in *C. maxima*, Table S6: The identified SSR markers in *C. pepo*, Table S7: List of cross-species SSR markers between *C. pepo* and *C. sativus* identified by in silico PCR, Table S8: List of cross-species SSR markers between *C. pepo* and melon identified by in silico PCR, Table S9: List of cross-species SSR markers between *C. pepo* and watermelon identified by in silico PCR, Table S10: List of cross-species SSR markers between *C. pepo* and *C. maxima* identified by in silico PCR, Table S11: List of cross-species SSR markers between *C. pepo* and *C. moschata* identified by in silico PCR, Table S12: The total SSR markers in *C. pepo* genetic diversity and population structure analysis, Table S13: Polymorphism and allelic diversity of SSR markers in *C. pepo* materials.

Author Contributions: L.Z. and H.Z. performed the data analysis and wrote the manuscript. Y.L., Y.W. (Yong Wang), and X.W. conducted validation of SSR polymorphism, genetic diversity, and population structure analysis. J.L., Z.Z., and Y.W. (Yanjiao Wang). participated in the SSR genetic diversity analysis. J.H., S.Y., and Y.L. conducted the germplasm collection. S.S. and L.Y. designed the experiments and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Major Public Welfare Project of Henan Province (202102110202) and the Science and Technology Project of China National Tobacco Corporation (Henan Tobacco Company) (2018410000270095).

Data Availability Statement: All generated or analyzed data during this study are reflected in the present article.

Acknowledgments: The authors are highly grateful to Yufeng Wu of Nanjing Agricultural University with the Circos analysis. The *C. pepo* accessions used in this study were introduced from the National Crop Germplasm Resource Platform of China. We are also indebted to the editor and reviewers for critically evaluating the manuscript and providing constructive comments for improvement.

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

References

1. George, E.B.; Ronald, J.T. Toxic plants of North America. *Choice Curr. Rev. Acad. Libraries* **2013**, *12*, 2202–2203.
2. Loy, J.B. Morpho-physiological aspects of productivity and quality in squash and pumpkins (*Culcurbita* spp.). *Crit. Rev. Plant Sci.* **2004**, *23*, 337–363. [CrossRef]

3. Savage, J.A.; Haines, D.F.; Holbrook, N.M. The making of giant pumpkins: How selective breeding changed the phloem of *Cucurbita maxima* from source to sink. *Plant Cell Environ.* **2015**, *38*, 1543–1554. [CrossRef] [PubMed]
4. Davis, A.R.; Perkins-Veazie, P.; Sakata, Y.; Lopez-Galarza, S.; Maroto, J.V.; Lee, S.G.; Huh, Y.C.; Sun, Z.Y.; Miguel, A.; King, S.R.; et al. Cucurbit grafting. *Crit. Rev. Plant Sci.* **2008**, *27*, 50–74. [CrossRef]
5. Lee, J.M.; Kubota, C.; Tsao, S.J.; Bie, Z.; Echevarria, P.H.; Morra, L.; Oda, M. Current status of vegetable grafting: Diffusion, grafting techniques, automation. *Sci. Hortic.* **2010**, *127*, 93–105. [CrossRef]
6. Lv, J.; Qi, J.J.; Shi, Q.X.; Shen, D.; Zhang, S.P.; Shao, G.J.; Li, H.; Sun, Z.Y.; Weng, Y.Q.; Shang, Y.; et al. Genetic Diversity and Population Structure of Cucumber (*Cucumis sativus* L.). *PLoS ONE* **2012**, *7*, e46919. [CrossRef] [PubMed]
7. Li, Y.; Wen, C.; Weng, Y. Fine mapping of the pleiotropic locus B for black spine and orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB transcription factor. *Theor. Appl. Genet.* **2013**, *126*, 2187–2196. [CrossRef] [PubMed]
8. Huang, S.W.; Li, R.Q.; Zhang, Z.H.; Li, L.; Gu, X.F.; Fan, W.; Lucas, W.J.; Wang, X.W.; Xie, B.Y.; Ni, P.X.; et al. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* **2009**, *41*, 1275–1281. [CrossRef] [PubMed]
9. Guo, S.G.; Zhang, J.G.; Sun, H.H.; Salse, J.; Lucas, W.J.; Zhang, H.Y.; Zheng, Y.; Mao, L.Y.; Ren, Y.; Wang, Z.W.; et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* **2013**, *45*, 51–58. [CrossRef]
10. Garcia-Mas, J.; Benjak, A.; Sanseverino, W.; Bourgeois, M.; Mir, G.; González, V.M.; Hénaff, E.; Câmara, F.; Cozzuto, L.; Lowy, E.; et al. The genome of melon (*Cucumis melo* L.). *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11872–11877. [CrossRef]
11. 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 Genom.* **2010**, *11*, 569. [CrossRef]
12. Zhu, H.; Guo, L.; Song, P.; Luan, F.; Hu, J.; Sun, X.; Yang, L. Development of genome-wide SSR markers in melon with their cross-species transferability analysis and utilization in genetic diversity study. *Mol. Breed.* **2016**, *36*, 153. [CrossRef]
13. 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]
14. Li, D.W.; Cuevas, H.E.; Yang, L.M.; Li, Y.H.; Garcia-Mas, J.; Zalapa, J.; Staub, J.E.; Luan, F.S.; Reddy, U.; He, X.M.; et al. Syntenic relationships between cucumber (*Cucumis sativus* L.) and melon (*C. melo* L.) chromosomes as revealed by comparative genetic mapping. *BMC Genom.* **2011**, *12*, 396. [CrossRef] [PubMed]
15. Yang, L.M.; Koo, D.H.; Li, D.W.; Zhang, T.; Jiang, J.M.; Luan, F.S.; Renner, S.S.; Henaff, E.; Sanseverino, W.; Garcia-Mas, J.; et al. Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in Cucumis. *Plant J.* **2014**, *77*, 16–30. [CrossRef] [PubMed]
16. Hi, L.Y.; Jeong, J.H.; Hong, K.H.; Dong, K.B. Use of Random Amplified Polymorphic DNAs for Linkage Group Analysis in Interspecific Hybrid F_2 Generation of *Cucurbita*. *Hortic. Environ. Biotechnol.* **1995**, *36*, 323–330.
17. Brown, R.N.; Myers, J.R. A genetic map of squash (*Cucurbita* sp.) with randomly amplified polymorphic DNA markers and morphological markers. *J. Am. Soc. Hortic.* **2002**, *127*, 568–575. [CrossRef]
18. Paris, H.S.; Yonash, N.; Portnoy, V.; Mozes-Daube, N.; Tzuri, G.; Katzir, N. Assessment of genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers. *Theor. Appl. Genet.* **2003**, *106*, 971–978. [CrossRef]
19. Zraidi, A.; Stift, G.; Pachner, M.; Shojaeiyan, A.; Gong, L.; Lelley, T. A consensus map for *Cucurbita pepo*. *Mol. Breed.* **2007**, *20*, 375–388. [CrossRef]
20. Gong, L.; Stift, G.; Kofler, R.; Pachner, M.; Lelley, T. Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theor. Appl. Genet.* **2008**, *117*, 37–48. [CrossRef]
21. Esteras, C.; Gomez, P.; Monforte, A.J.; Blanca, J.; Vicente-Dolera, N.; Roig, C.; Nuez, F.; Pico, B. High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. *BMC Genom.* **2012**, *13*, 80. [CrossRef] [PubMed]
22. Sun, H.H.; Wu, S.; Zhang, G.Y.; Jiao, C.; Guo, S.G.; Ren, Y.; Zhang, J.; Zhang, H.Y.; Gong, G.Y.; Jia, Z.C.; et al. Karyotype Stability and Unbiased Fractionation in the Paleo-Allotetraploid *Cucurbita* Genomes. *Mol. Plant* **2017**, *10*, 1293–1306. [CrossRef] [PubMed]
23. Montero-Pau, J.; Blanca, J.; Bombarely, A.; Ziarsolo, P.; Esteras, C.; Marti-Gomez, C.; Ferriol, M.; Gomez, P.; Jamilena, M.; Mueller, L.; et al. De novo assembly of the zucchini genome reveals a whole-genome duplication associated with the origin of the *Cucurbita* genus. *Plant Biotechnol. J.* **2018**, *16*, 1161–1171. [CrossRef]
24. Blanca, J.; Canizares, J.; Roig, C.; Ziarsolo, P.; Nuez, F.; Pico, B. Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genom.* **2011**, *12*, 104. [CrossRef]
25. Wyatt, L.E.; Strickler, S.R.; Mueller, L.A.; Mazourek, M. An acorn squash (*Cucurbita pepo* ssp. ovifera) fruit and seed transcriptome as a resource for the study of fruit traits in Cucurbita. *Hortic. Res.* **2015**, *2*, 14070. [CrossRef]
26. Xanthopoulou, A.; Psomopoulos, F.; Ganopoulos, I.; Manioudaki, M.; Tsaftaris, A.; Nianiou-Obeidat, I.; Madesis, P. De novo transcriptome assembly of two contrasting pumpkin cultivars. *Genom. Data* **2016**, *7*, 200–201. [CrossRef] [PubMed]
27. Aliki, X.; Ganopoulos, I.; Psomopoulos, F.; Manioudaki, M.; Moysiadis, T.; Kapazoglou, A.; Osathanunkul, M.; Michailidou, S.; Kalivas, A.; Tsaftaris, A.; et al. De novo comparative transcriptome analysis of genes involved in fruit morphology of pumpkin cultivars with extreme size difference and development of EST-SSR markers. *Gene* **2017**, *622*, 50–66. [CrossRef]
28. Vitiello, A.; Scarano, D.; D'Agostino, N.; Digilio, M.C.; Pennacchio, F.; Corrado, G.; Rao, R. Unraveling zucchini transcriptome response to aphids. *Peer J. PrePrints* **2016**, *4*. [CrossRef]

29. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [CrossRef] [PubMed]
30. Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4325. [CrossRef]
31. 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]
32. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]
33. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
34. 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]
35. Lu, C.R.; Zou, C.S.; Zhang, Y.P.; Yu, D.Q.; Cheng, H.L.; Jiang, P.F.; Yang, W.C.; Wang, Q.L.; Feng, X.X.; Prosper, M.A.; et al. Development of chromosome-specific markers with high polymorphism for allotetraploid cotton based on genome-wide characterization of simple sequence repeats in diploid cottons (*Gossypium arboreum* L. and *Gossypium raimondii* Ulbrich). *BMC Genom.* **2015**, *16*, 55. [CrossRef] [PubMed]
36. Wang, Q.; Fang, L.; Chen, J.D.; Hu, Y.; Si, Z.F.; Wang, S.; Chang, L.J.; Guo, W.Z.; Zhang, T.Z. Genome-Wide Mining, Characterization, and Development of Microsatellite Markers in *Gossypium* Species. *Sci. Rep.* **2015**, *5*, 10638. [CrossRef] [PubMed]
37. Zhang, S.; Tang, C.J.; Zhao, Q.; Li, J.; Yang, L.F.; Qie, L.F.; Fan, X.K.; Li, L.; Zhang, N.; Zhao, M.C.; et al. Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in Foxtail millet [*Setaria italica* (L.) P. Beauv.]. *BMC Genom.* **2014**, *15*, 78. [CrossRef] [PubMed]
38. Wang, X.W.; Yang, S.; Chen, Y.D.; Zhang, S.M.; Zhao, Q.S.; Li, M.; Gao, Y.L.; Yang, L.; Bennetzen, J.L. Comparative genome-wide characterization leading to simple sequence repeat marker development for Nicotiana. *BMC Genom.* **2018**, *19*, 500. [CrossRef]
39. Weber, J.L. Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. *Genomics* **1990**, *7*, 524–530. [CrossRef]
40. Li, Y.C.; Korol, A.B.; Fahima, T.; Beiles, A.; Nevo, E. Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Mol. Ecol.* **2002**, *11*, 2453–2465. [CrossRef] [PubMed]
41. Wang, Z.; Weber, J.L.; Zhong, G.; Tanksley, S.D. Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* **1994**, *88*, 1–6. [CrossRef] [PubMed]
42. Morgante, M.; Hanafey, M.; Powell, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* **2002**, *30*, 194–200. [CrossRef] [PubMed]
43. Li, Y.C.; Korol, A.B.; Fahima, T.; Nevo, E. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **2004**, *21*, 991–1007. [CrossRef] [PubMed]
44. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* **2005**, *23*, 48–55. [CrossRef] [PubMed]
45. Zhao, Z.X.; Guo, C.; Sutharzan, S.; Li, P.; Echt, C.S.; Zhang, J.; Liang, C. Genome-Wide Analysis of Tandem Repeats in Plants and Green Algae. *G3 Genes Genomes Genet.* **2014**, *4*, 67–78. [CrossRef]
46. Cui, J.; Cheng, J.; Nong, D.; Peng, J.; Hu, Y.; He, W.; Zhou, Q.; Dhillon, N.P.S.; Hu, K. Genome-Wide Analysis of Simple Sequence Repeats in Bitter Gourd (*Momordica charantia*). *Front. Plant Sci.* **2017**, *8*, 1103. [CrossRef] [PubMed]
47. Kim, T.S.; Booth, J.G.; Gauch, H.G., Jr.; Sun, Q.; Park, J.; Lee, Y.H.; Lee, K. Simple sequence repeats in Neurospora crassa: Distribution, polymorphism and evolutionary inference. *BMC Genom.* **2008**, *9*, 31. [CrossRef]
48. Cheng, J.; Zhao, Z.; Li, B.; Qin, C.; Wu, Z.; Trejo-Saavedra, D.L.; Luo, X.; Cui, J.; Rivera-Bustamante, R.F.; Li, S.; et al. A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in Capsicum. *Sci. Rep.* **2016**, *6*, 18919. [CrossRef] [PubMed]
49. Tangphatsornruang, S.; Somta, P.; Uthaipaisanwong, P.; Chanprasert, J.; Sangsrakru, D.; Seehalak, W.; Sommanas, W.; Tragoonrung, S.; Srinives, P. Characterization of microsatellites and gene contents from genome shotgun sequences of mungbean (*Vigna radiata* (L.) Wilczek). *BMC Plant Biol.* **2009**, *9*, 137. [CrossRef] [PubMed]
50. 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]
51. Ferriol, M.; Pico, B.; Nuez, F. Genetic diversity of a germplasm collection of Cucurbita pepo using SRAP and AFLP markers. *Theor. Appl. Genet.* **2003**, *107*, 271–282. [CrossRef] [PubMed]
52. Ntuli, N.R.; Tongoona, P.B.; Zobolo, A.M. Genetic diversity in *Cucurbita pepo* landraces revealed by RAPD and SSR markers. *Sci. Hortic.* **2015**, *189*, 192–200. [CrossRef]
53. Decker, D.S. Origin (s), Evolution, and Systematics of *Cucurbita pepo* (Cucurbitaceae). *Econ. Bot.* **1988**, *42*, 4–15. [CrossRef]
54. Paris, H.S. Germplasm enhancement of *Cucurbita pepo* (pumpkin, squash, gourd: Cucurbitaceae): Progress and challenges. *Euphytica* **2015**, *208*, 415–438. [CrossRef]