Plant Omics Journal POJ

POJ 7(5):345-352 (2014) ISSN:1836-3644

Development of 19 transferable *Cucurbita pepo* EST-SSR markers for the study of population structure and genetic diversity in pumpkin (*Cucurbita Moschata*)

Weihua Mao[1+], Yingjun Xu[2+], Shengchun Xu[3], Lingwei Ye[3], Hui Wang[3], Xu Gao[3], Xiaorong Mo[4]* and Yaming Gong[3]*

[1]Center of Analysis and Measurement, Zhejiang University, Hangzhou 310058, China
[2]Cixi Agricultural Technology and Popularization Center, Cixi 315300, China
[3]Institute of Vegetables, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China
[4]State Key Laboratory of Plant Physiology and Biochemistry, College of Life Science, Zhejiang University, Hangzhou 310058, China

[+]These authors are equally contributed to this work

***Corresponding author: xiaorong@zju.edu.cn; gongym07@126.com**

Abstract

Cucurbita moschata is an economically important species worldwide, but little is known about its genic and genomic information. Few molecular markers have been developed for *C. moschata* so far, which is not sufficient for breeding programs. The aim of this study was to examine the feasibility of development EST-SSR markers from *Cucurbita pepo* to *C. moschata* and to evaluate their potential for genetic analysis in *C. moschata* cultivars. Eighty-two *C. pepo* EST-SSRs were tested against 29 *C. moschata* cultivars from different geographical origin in China. As a result, a high proportion (84.14%) of markers was transferable to *C. moschata*. From them, nineteen markers exhibited polymorphism. A total of 53 alleles were identified with an average of 2.789 alleles per locus in all the tested samples. The polymorphic information content (*PIC*), the observed heterozygosity (H_O), the expected heterozygosity (H_E), and the Shannon's information index (*I*) were estimated to average 0.393, 0.125, 0.478, and 0.760, respectively. Population structure, principal coordinates and phylogenetic analysis all revealed that these 29 *C. mosch*ata cultivars tended to group into two subpopulations, which was related to fruit shape rather than geographic origin. Moreover, every subpopulation possessed its own population-specific alleles. This is the first report describing the development of transferable EST-SSR markers in *C. moschata* and their application to genetic analysis, which will offer an approach for future marker development and marker-assisted breeding in *C. moschata*.

Keywords: EST-SSR, *C. moschata*, *C. pepo*, Transferability, Genetic diversity.
Abbreviations: EST-SSRs_expressed sequence tag-derived simple sequence repeats; Na_number of alleles; Ne_number of effective alleles; Np_number of specific alleles; PIC_polymorphism information content; H_O_observed heterozygosity; H_E_expected heterozygosity; I_Shannon's Information index; PCoA_Principal coordinates analysis.

Introduction

Cucurbita moschata (2n = 2x = 40) is a Cucurbitaceae family member originating in Mexico and northern South America. Today, it is an economically important crop worldwide because of its nutritional quality, medicinal value, and wide ranging environmental adaptation to heat, humidity, disease and insects. *C. moschata* is also one of the most morphologically variable species in the plant kingdom in terms of fruit shape, size and color (Wu et al., 2011). Despite its agricultural and biological importance, much less genic and genomic information is available for *C. moschata* than for other Cucurbitaceae species, and the number of available molecular markers is still limited (Gwanama et al., 2000; Ferriol et al., 2004a; Wu et al., 2011). The development of more reliable and efficient molecular markers is urgently needed in order to facilitate *facilitate C. moschata breeding*.

*EST-SSR molecular markers are powerful tools for mapping and*marker-assisted breeding, possessing several intrinsic advantages (Varshney et al., 2005). Since they are derived from transcripts, EST-SSRs serve as a kind of functional marker, directly related to genes that are known to be functional. In addition, they are inexpensive to develop compared to other types of molecular markers. More importantly, an increasing amount of evidences demonstrate that EST-SSRs are highly transferable between related species (Scott et al., 2000; Decroocq et al., 2003; Saha et al., 2004; Varshney et al., 2005; Ellis and Burke, 2007). For example, Sim et al. (2009) tested the transferability of 165 cereal EST-SSRs to ryegrass, and found high levels of transferability (57%) between cereals and ryegrass. Moreover, 67% of the transferable markers proved to be polymorphic within the tested ryegrass cultivars. Xu et al. (2012) used EST-SSRs developed from pea to investigate transferability in relative species *Vicia faba.* A total of 22 primers pairs (53.7%) were transferable to *V. faba* and all of them revealed polymorphism. Recently, EST-SSRs derived from sugarcane ESTs were also used to the assessment their cross-species/genera transferability among four Saccharum species, two closely related genera and four cereals. This

work revealed the cross transferability rate ranged from 87.0% to 93.4% in Saccharum complex, 80.0% to 87.0% in allied genera, and 76.0% to 80.0% in cereals (Singh et al., 2013). Therefore, due to their effectively transfer polymorphic across taxa, EST-SSRs could be used not only for comparative mapping and evolutionary studies, but also for analyses of related species lacking genic and genomic information (Varshney et al., 2005; Ellis and Burke, 2007). Currently, most studies in the Cucurbitaceae family have focused only on a minority of the crops, such as cucumber (Guo et al., 2010), melon (Fernandez-Silva et al., 2008), and *Cucurbita pepo* (Blanca et al., 2011). Knowledge about the *C. moschata* genome is so sparse that there are only a few sequences deposited in the public database, which greatly limits the development of *C. moschata* EST-SSRs. It had been suggested that *C. moschata* and *C. pepo* have the closest genetic relationship in the Cucurbita genus (Sanjur et al., 2002; Gong et al., 2008), and a recent report showed 25 *C. pepo* markers (83.33%) could be transferred to one *C. moschata* accession, suggesting the feasibility of transferring EST-SSRs from *C. pepo* to *C. moschata* (Blanca et al., 2011). In the present study, we will develop a subset of *C. pepo* EST-SSRs for their transferability and polymorphism to *C. moschata* by cross-species amplification, and then evaluate their efficiency in population structure assessment and diversity analysis in the different *C. moschata* individuals for the first time.

Results

The transferability and polymorphism of C. pepo EST-SSR markers in C. moschata

To assess the transferability of *C. pepo* EST-SSRs to *C. moschata*, 82 *C. pepo* EST-SSRs primer pairs were selected from the cucurbit genomics database (http://www.icugi.org/cgi-bin/ICuGI/EST/SSR.cgi?organism =cucurbita_pepo) and tested to verify the potential of cross-species amplification using twenty-nine widely cultivated *C. moschata* cultivars from different regions of China (Fig.1; Table 1). A total of 69 primer pairs (84.14%) showed positive amplification products with the expected allele sizes. These transferable primer pairs were subsequently utilized to screen for polymorphic loci among all tested cultivars. As a result, 19 loci (27.53%) revealed polymorphisms (Fig. 2), and the number of alleles per locus ranged from 2 to 7 with a mean value of 2.789 (Table 2). The polymorphism information content (*PIC*) value ranged from 0.674 (PU00396) to 0.162 (PU001551), with an average of 0.393. Two loci (PU00396 and PU01491) were highly polymorphic (*PIC*>0.5), sixteen loci were moderately polymorphic (0.25<*PIC*<0.5), and only one locus (PU001551) exhibited low polymorphism (*PIC*<0.25). The observed heterozygosity (H_O) values varied from 0.00 to 0.448, with a mean of 0.125, while the expected heterozygosity *(HE)* values ranged from 0.171 to 0.705, with an average of 0.478. The average Shannon's Information index (*I*) was 0.760, and ranged from 0.355 (PU01551) to 1.529 (PU00396). In addition, the result of BLASTX show that ten of these novel EST-SSR associated unigenes were similar to seven known genes and three uncharacterized protein (Table 2).

Genetic analysis in C. moschata

To evaluate the usefulness of these EST-SSRs, we used them to analyze the genetic relationships among the 29 Chinese *C. moschata* cultivars in this study. The population structure of the cultivated *C. moschata* was examined using the STRUCTURE program to estimate the optimum number of genetically distinct populations (K) according the method of Evanno et al. (2005). There was a clear peak in the value of ΔK at K= 2, which indicated the presence of two main subpopulations in *C. moschata* (Fig.3A). Interestingly, the classification of populations appeared to be in consonance with fruit shape rather than geographic origin. As shown in Fig.3B, POP1 consisted of 20 flattened-fruit cultivars, collected from Gansu (1), Hebei (3), Heilongjiang (3), Hunan (3), Jiangsu (1), Jilin (2), Liaoning (1), Shandong (2), Shanghai (1), Zhejiang (2), Guangdong (1), while the other population consisted of seven pyriform-fruit cultivars and two curved-fruit cultivars, gathered from Hebei (1), Hunan (1), Shanxi (1), Zhejiang (2), Beijin (3), Sichuan (1). Principal coordinates analysis (PCoA) and phylogenetic analysis also showed that all of the tested cultivars could be divided into two major clusters, which in general supported the classification inferred by the population structure analysis (Fig.4；Fig.5).

Assessment of the genetic diversity of the subpopulations

In order to reveal the genetic diversity of the subpopulations, the genetic diversity analysis was carried out for the two inferred subpopulations. The result showed that a total of 25 population-specific alleles were found. As showed in Fig.2, the allele with 110 bp was only detected in POP1, while the allele with 124 bp only existed in POP2. However, the average number of specific alleles in POP1 was 0.895, which was more than that in POP2, suggesting that there should be more population-specific alleles in POP1 (Table 3). In addition, POP1 showed higher gene diversity, with the expected heterozygosity (H_E) of 0.271 and Shannon's Information index (*I*) of 0.477. For each subpopulation, the expected heterozygosity (H_E) was higher than the observed heterozygosity (H_O) (Table 3).

Discussion

Due to their derivation from transcripts, EST-SSR markers possess higher transferability to related species compared with genomic SSR markers (Varshney et al., 2005; Ellis and Burke, 2007). Therefore, transferability is one of the most useful features of EST-SSR markers. It was reported that the transferability of *M. truncatula* EST-SSRs to faba bean, chickpea and pea was two fold higher than that of genomic SSRs (Gutierrez et al., 2005). Now, the ability to effectively transfer EST-SSRs across taxa had been demonstrated in a number of cases (Sim et al., 2009; Xu et al., 2012; Singh et al., 2013), which indicated that developing transferable EST-SSRs shoud be a feasible method for the plant species with little information on SSRs or ESTs. However, the rate of transferability also depended on the phylogenetic distance and level of sequence conservation between the studied species (Peakall et al., 1998; García-Moreno et al., 2010). It could vary from 10 to 100% with the different tested species (Ellis and Burke, 2007). Unfortunately, although cross transferability of EST–SSRs have been demonstrated in many crops, studies on the transferability of EST–SSRs for *Cucurbita* have been scarce. Fernandez-Silva et al. (2008) investigated the transferability of 132 *C. melo* EST-SSR primers to *Cucurbita* species, and found very low levels of transferability in *C. pepo (10.6%), C. maxima and C. moschata (12.1%)*,concluding that transfer of melon EST-

Table 1. Names and origins of the 29 widely adapted *C. moschata* cultivars from China used in the study.

Cultivar No.	Origin	Fruit shape	Cultivar No.	Origin	Fruit shape
1	Beijin（39°53'59"N,116°17'59"E）	Pyriform	16	Hunan (28°7'12"N,112°35'24"E)	Flattened
2	Beijin（39°53'59"N,116°17'59"E）	Curved	17	Jiangsu (32°2'23"N,118°46'48"E)	Flattened
3	Beijin（39°53'59"N,116°17'59"E）	Pyriform	18	Jilin (43°32'23"N,125°11'23"E)	Flattened
4	Gansu（36°1'48"N,103°43'48"E）	Flattened	19	Jilin (43°32'23"N,125°11'23"E)	Flattened
5	Guangdong (23°9'36"N,113°13'48"E)	Flattened	20	Liaoning (41°28'48"N,123°15'0"E)	Flattened
6	Hebei (38°1'48"N,114°28'48"E)	Pyriform	21	Shandong (36°38'59"N,117°0'0"E)	Flattened
7	Hebei (38°1'48"N,114°28'48"E)	Flattened	22	Shandong (36°38'59"N,117°0'0"E)	Flattened
8	Hebei (38°1'48"N,114°28'48"E)	Flattened	23	Shanghai (34°30'0"N,121°25'59"E)	Flattened
9	Hebei (38°1'48"N,114°28'48"E)	Flattened	24	Shanxi (37°52'11"N,112°31'48"E)	Pyriform
10	Heilongjiang (45°44'25"N,126°39'23"E)	Flattened	25	Sichuan (30°40'12"N,104°3'36"E)	Curved
11	Heilongjiang (45°44'25"N,126°39'23"E)	Flattened	26	Zhejiang (30°14'35"N,120°9'36"E)	Flattened
12	Heilongjiang (45°44'25"N,126°39'23"E)	Flattened	27	Zhejiang (30°14'35"N,120°9'36"E)	Flattened
13	Hunan (28°7'12"N,112°35'24"E)	Flattened	28	Zhejiang (30°14'35"N,120°9'36"E)	Pyriform
14	Hunan (28°7'12"N,112°35'24"E)	Pyriform	29	Zhejiang (30°14'35"N,120°9'36"E)	Pyriform
15	Hunan (28°7'12"N,112°35'24"E)	Flattened			

Table 2. Characterization of 19 EST-SSR primers in *C. moschata*.

Locus	SSR motif	Forward primer sequence (5'–3')	Reverse primer sequence (5'–3')	Putative function	Size	*Na*	*Ne*	*PIC*	*I*	*Ho*	H_E
PU000007	(CCG)5	GGAGACCCCATTTTCCTGTT	AAGCACTCACTTCCCCTTCA	Spermidine synthase-like	105	3	1.976	0.437	0.849	0.148	0.494
PU000104	(TGC)8	TCGGATCCACTCCCATAGAC	AAACCATGCCTCTGACAACC	Homoserine kinase-like	233	2	1.979	0.372	0.688	0.069	0.495
PU000308	(TC)8	GGGGGAAGCCATTTATTCTC	CCATGGCCTTACAAGAAGGA		194	3	1.736	0.347	0.670	0.207	0.424
PU000396	(TCT)5	CCTGCGATTGTAAGATGCAA	CAGACCTCGAAAGCCAAGTC		221	7	3.391	0.674	1.529	0.379	0.705
PU000705	(CGA)7	AAAGAAAAGCTCGCTTGGTG	TCCATGCCCACATCAACTAA		122	4	1.967	0.450	0.930	0.000	0.492
PU000716	(TATGA)6	TCCGGTAGAGAATCATTCCG	ATCATCCATCATCGACCCAT		189	2	1.942	0.367	0.678	0.069	0.485
PU000750	(CCTTT)5	AAAGTTGTCCAAAACCGACG	GGCTCGAGAATCTGAAAACG	Uncharacterized protein	228	3	2.314	0.477	0.919	0.207	0.568
PU001491	(CTT)5	CACAGACGCAGAAACAGAGG	AAGGACAAGATCCATGGGG		158	3	2.951	0.587	1.090	0.448	0.661
PU001527	(CTC)5	CAACTGCTACACCAAGCCAA	GGAGGTGGGGGTAGAGAGAC		160	2	1.824	0.350	0.644	0.069	0.452
PU001551	(GAAT)5	AAGGAGGAGGAGGTAGCAGC	CCTCGCACTGAGACAACAAA	Tau class glutathione transferase	161	3	1.206	0.162	0.355	0.111	0.171
PU001748	(ATT)6	TGCAATCCATAATTCATAATT TAACC	CACACGACGCTGTGAAAGAT	DNA polymerase epsilon subunit 3-like	188	2	1.787	0.344	0.632	0.034	0.441
PU001821	(AT)7	CGGCAGTGATGACTGAGAAA	CGCTACCATGTGGACAGAGA	Drm3-like protein isoform 1	204	2	1.748	0.336	0.619	0.000	0.428
PU001869	(GATAT)5	TCCGGTAGAGAATCATTCCG	ATCATCCATCATCGACCCAT		177	3	1.967	0.388	0.743	0.069	0.492
PU002054	(CTC)5	CAACTGCTACACCAAGCCAA	GGAGGTGGGGGTAGAGAGAC	Vinorine synthase-like	164	2	1.859	0.355	0.655	0.034	0.462
PU002121	(ATT)6	ATCGATAAGAAAGACGGCGA	CCTTCAGGTCCGAGTTTCAA		199	2	1.906	0.362	0.668	0.037	0.475
PU002147	(GAAGAG)5	GCTCTGTGAAGGAGGACGAC	CCGCTGCATAGGGAATGTAT	Uncharacterized protein	239	4	2.159	0.439	0.881	0.069	0.537
PU002158	(TC)6	GTCGGGGTAGGGAGAAACTC	AGTCGGAACAACGAATCCAC		251	2	1.800	0.346	0.637	0.074	0.444
PU002184	(TGA)5	GGGGATGCAAAAAGTGAAAA	CCCCAACAAAATCCCTTCTT	Uncharacterized protein	124	2	1.748	0.336	0.619	0.000	0.428
PU002200	(TTC)7	GGGCTTCGAGGACATATTGA	GCCGAAAAGCTTTCTCACAC	Nuclear complex protein 4 homolog B-like	219	2	1.774	0.341	0.628	0.357	0.436

Data provided for each primer include locus name, repeat type, primer sequence, size of expected product, number of alleles detected (N_A), number of effective alleles (*Ne*), polymorphism information content (*PIC*), Shannon's information index (*I*), observed heterozygosity (H_O), expected heterozygosity (H_E).

Fig 1. Geographic distribution of the tested *C. moschata* cultivars in this study. Dot indicated sampling site identified in Table 1.

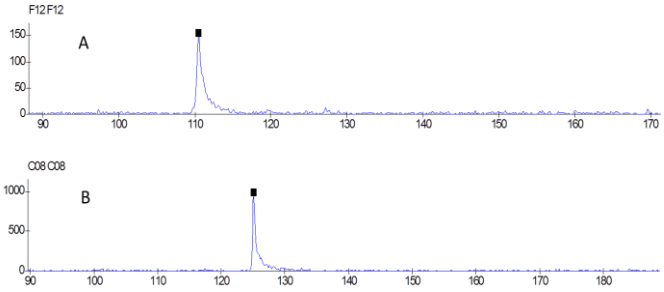

Fig 2. Representative electropherogram of marker PU002184 in two C. moschata cultivars A and B. A represented cultivars 27 and showed a allele with 110 bp, B represented cultivars 28 and showed a allele with 124 bp. X axis indicated allele size in base pairs (bp), Y axis indicated relative fluorescence unit (RFU).

SSRs was not an effective method for marker development in *Cucurbita*. By contrast, recent research revealed that EST-SSRs developed from *C. pepo* had a high level of transferability within the *Cucurbita* genus, and 83.33% of the *C. pepo* markers could transfer to *C. moschata* (Blanca et al., 2011). However, the polymorphism of these markers had not been tested. In the present study, we found that among 82 *C. pepo* primer pairs, 69 produced positive cross-amplification in *C. moschata*, which corresponded to an 84.14 % transfer rate. This high transfer rate is similar to the previous report (Blanca et al., 2011)**.** The higher level of transferability between *C. moschata* and *C. pepo* might due to their close genetic relationship in the genus *Cucurbita* (Sanjur et al., 2002; Gong et al., 2008). Moreover, among these transferable markers, we found that 19 loci (27.53%) were polymorphic, which was less than the 30.1% polymorphic rate noted by Gong et al. (2008) based on 170 *C. pepo* genomic microsatellites analyzed in three *C. moschata* accessions. This observation was in agreement with previous reports that EST-SSR markers had lower polymorphism than genomic SSRs in crop plants because of greater sequence conservation in transcribed regions. Given that genomic sequence information is still very limited in *C. moschata*, the high levels of transferability and substantial polymorphism found in this study would contribute to future marker development in *C. moschata*.

Generally, the polymorphism information content (*PIC*) value is the most useful index to evaluate the discriminatory power of a marker, based on allelic diversity and frequency (Shete et al., 2000; Nagl et al., 2011; Saeed et al., 2011). Locus polymorphism can be classified as low (*PIC* < 0.25), medium (0.5 > *PIC* > 0.25) or high (*PIC* > 0.5) (Xie et al., 2010; Babaei et al., 2012). In the present study, the average polymorphism information content (*PIC*) value for all loci was 0.393, and only one locus had low polymorphism, suggesting the usefulness of these markers for germplasm evaluation and genetic variation analysis. In addition, the Shannon's information index (*I*) is also acommon diversity index to represent the effectiveness of microsatellite loci to reveal the variation (Babaei et al., 2012). Interestingly, the average value of *I* and *He* in the present study was 0.760 which was higher than that detected by Wu et al. using AFLP in *C. moschata* (2011). This result not only indicated abundant genetic variation for the germplasm studied herein, but also demonstrated the effectiveness of these EST-SSRs to reveal the variation (Babaei et al., 2012). Based on these variability parameters, the new EST-SSRs described in this study would be useful tools for genetic and genetic diversity among the different *C. moschata* genotypes by EST-SSR markers has not previously been carried out. In order to reveal the potential of EST-SSR markers, we used our novel EST-SSR markers to analyze the genetic diversity and genetic structure of *C. moschata*. The results of the population structure analysis, the neighbour joining (NJ) tree, and principal coordinates analysis all supported division of the tested cultivars into two major clusters (Fig. 3-5). The flattened-fruit cultivars formed one cluster, while the pyriform- and curved-fruit cultivars were clustered in the other subgroup. However, pyriform- and curved-fruit accessions could not be further distinguished in this study. This suggested that there shoud be relatively high genetic similarity between these two fruit shape types. Therefore, development of more EST-SSR markers for *C. moschata* will help to distinguish the different fruit shape types in more detail. Our results also found that the genetic relationship among *C. moschata* cultivars was related to their fruit shape rather than geographic origin. However, in recently reported AFLP analysis, the *C. moschata* accessions could not be grouped clearly according to fruit shape (Wu et al., 2011). The different results obtained with EST-SSRs and AFLPs might due to the different information provided by the two different marker systems（Ferriol et al., 2003; Ferriol et al., 2004a; Ferriol et al., 2004b). EST-SSRs come from coding regions, which are involved in morphological and agronomic traits. Furthermore, the different markers may cover different genomic regions. Interestingly, our results were in good correspondence with the results of Xu et al. (2011), who examined genetic diversity of the bottle gourd based on EST-SSRs, and found that the chinese bottle gourd germplasm could be divided according to fruit shape. Further genetic diversity analyses of the subpopulations showed that every population possessed its own population-specific alleles,

Table 3. Genetic diversity parameters for the two subpopulations.

Population	*Na*	*Ne*	*Np*	*Ho*	H_E	*I*
Pop1	2.368	1.537	0.895	0.136	0.271	0.477
Pop2	1.895	1.341	0.421	0.102	0.158	0.291

Na, the average number of alleles; *Ne*, the average number of effective alleles; *Np*, the average number of specific alleles; *I*, Shannon's information index; *Ho*, observed heterozygosity; H_E, expected heterozygosity.

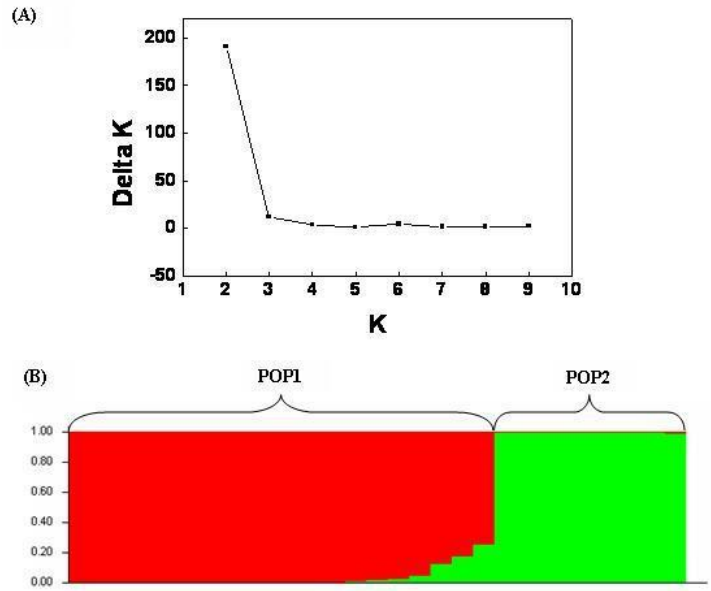

Fig 3. The population structure analysis of the tested cultivars in *C. moschata*. (A) Delta K values for different numbers of populations assumed (K) in the STRUCURE analysis. (B) Two subpopulations were assigned when K = 2. The distribution of the cultivars to the different subpopulations was shown with different colors (POP1: red, POP2: green).

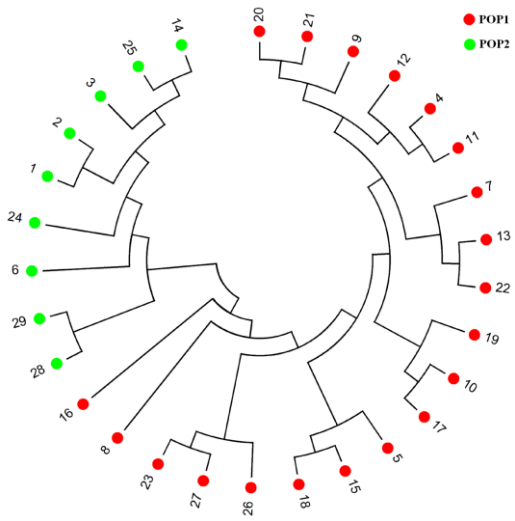

Fig 4. Unrooted neighbor-joining tree (NJ) of the 29 tested cultivars in *C. moschata*. Different colors indicated the different subpopulations inferred from the population structure analysis. The red dots represented the cultivars which belonged to POP1, while the green dots represented the cultivars which belonged to POP2.

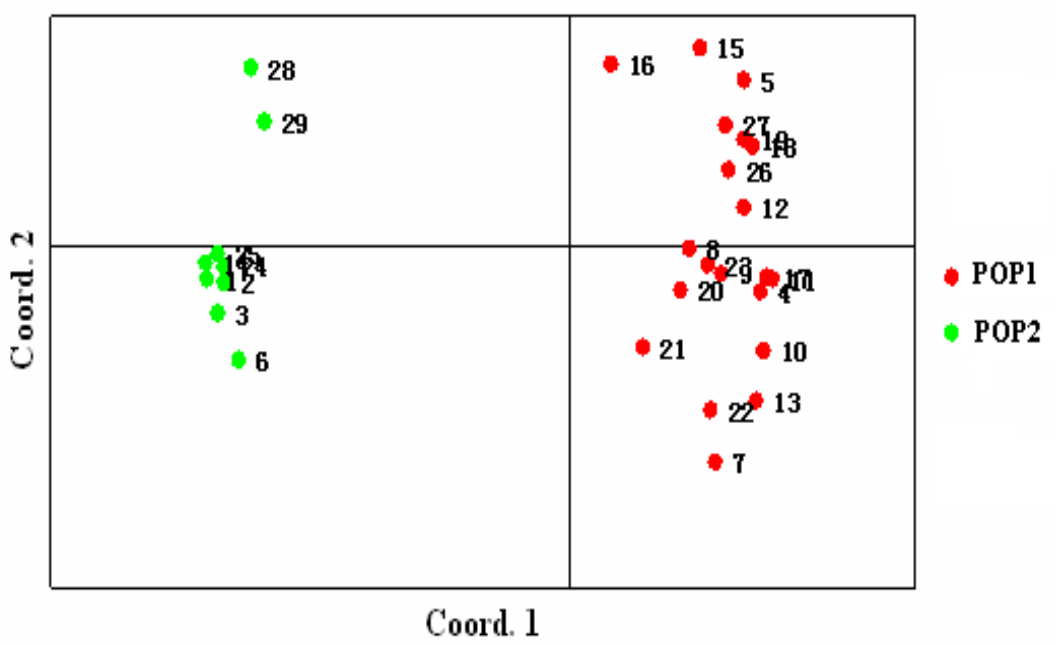

Fig 5. Principal coordinates analysis (PCoA) of the 29 tested cultivars in *C. moschata*. Different colors indicated the different subpopulations inferred from the population structure analysis. The red dots represented the cultivars which belonged to POP1, while the green dots represent the cultivars which belonged to POP2.

which provided us useful information for core collection and breeding in *C. moschata*. For example, in order to broaden the genetic diversity, we should collect and preserve genetic resources originating from different populations. Meanwhile, crosses between sub-populations would broaden the genetic base and benefit breeding and cultivar improvement.

Materials and Methods

Plant materials and DNA extraction

Twenty-nine widely cultivated cultivars from different geographical origins in China were collected to evaluate the transferability and polymorphism of the markers (Fig.1; Table 1). Genomic DNA from each individual was extracted from young leaf tissue using a modified CTAB method (Steward and Via, 1993).

Transferability analysis and EST-SSR marker development

A total of eighty-two *C. pepo* primer pairs were chosen and downloaded from the cucurbit genomics database (http://www.icugi.org/cgi-bin/ICuGI/EST/SSR.cgi?organism=cucurbita_pepo). In order to accurately assess the transferability of marker, these chosen SSRs are exclusive of the markers which had been tested by Blanca et al. (2011). The forward primer of each pair was labeled with HEX or 6-FAM fluorescent dye. PCR was performed in 20 μL reaction mixtures that contained 20 ng template DNA, 1 unit Taq DNA polymerase (TaKaRa, Dalian, Liaoning, China), 0.2 mM dNTPs, 0.2 μM each primer, and 1 × PCR buffer (containing Mg^{2+}). The PCR was carried out on a PTC-225 Peltier Thermal Cycler (MJ Research, Waltham, MA) as follows: denaturation at 95°C for 3 min, followed by 30 cycles of 30 s at 95°C, 30 s at appropriate annealing temperature, 1 min at 72°C and final extension at 72°C for 10 min. The PCR-amplified products were detected using the MegaBACE 1000 DNA analysis system (Amersham Biosciences, Piscataway, NJ). The ET550-R size standard (GE Healthcare, Piscataway, NJ) and Genetic Profiler version 2.2 (GE Healthcare, Piscataway, NJ) were used to analyse the size of the amplified fragments as previously described (Gong et al., 2010). Only the primer pairs which could product distinct amplification product with the expected allele sizes were considered transferable. The transferable EST-SSRs were subsequently used to screen for polymorphic loci among all tested cultivars. To determine the putative function of these novel EST-SSRs, all the EST-SSR associated unigenes were blasted against GenBank using the BLASTX (http://www.ncbi.nlm.nih.gov) with the expected value < 10^{-10}.

Statistical analysis

Genetic diversity

The number of alleles per locus including the average number of alleles (*Na*), the average number of effective alleles (*Ne*), the observed heterozygosity (*Ho*), the expected heterozygosity (H_E), Shannon's information index (*I*), major allele frequency and the number of specific alleles were calculated using the GenAlEx 6.5 software (Peakall and Smouse, 2012). Polymorphism information content (*PIC*) values were calculated using the formula developed by Anderson et al.,(1993).

Population structure

The population structure was analyzed with the STRUCTURE software (version 2.3.3), using an admixture model with k values ranging from 1 to 10. For each value of k, the analysis was run with five independent replicates of 100,000 Markov chain iterations (Evanno et al., 2005). The optimal number of clusters (K) was identified by STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest/).

Principal coordinates analysis (PCoA) and phylogenetic analysis

Principal coordinates analysis (PCoA) were also performed using the GenAlEx 6.5 software (Peakall and Smouse, 2012). A phylogenetic tree was constructed based on the neighbour-joining method using MEGA 4.0 (Tamura et al., 2007).

Conclusion

In conclusion, we have developed 19 *C. moschata* EST-SSR markers by cross-species amplification with *C. pepo* EST-SSRs, and systematically assessed their efficiency in analysis of population structure and genetic diversity in *C. moschata.* Being highly transferable and polymorphic, these novel markers will not only be widely useful for genetic diversity, resource conservation, and marker-assisted breeding in *C. moschata*, but also promote comparative genomics in other Cucurbitaceae species, which will open the door for future marker development and breeding studies in *C. moschata*.

Acknowledgments

The research was supported by the National Natural Science Foundation of China (30971701, 31372072), Zhejiang Provincial Important Science & Technology Specific Projects (2012C12903), the Public Welfare Project of Zhejiang Province (No. 2014C32032), and the Shaoxing Public Welfare Project (No. 2013B70051).

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome. 36:181-186.

Babaei N, Abdullah NAP, Saleh G, Abdullah TL (2012) Isolation and characterization of microsatellite markers and analysis of genetic variability in *Curculigo latifolia* Dryand. Mol Biol Rep. 39:9869-9877.

Blanca J, Canizares J, Roig C, Ziarsolo P, Nuez F, Pico B (2011) Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). BMC Genomiccs **.**12:104.

Decroocq V, Fave MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet.106: 912-922.

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity. 99:125–132.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14:2611–2620.

Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Picó B, Nuez F, Arús P, Garcia-Mas J, Monforte AJ (2008) Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.). Theor Appl Genet. 118:139-150.

Ferriol M, Picó B, de Cordova PF, Nuez F (2004a) Molecular diversity of a germplasm collection of squash *(Cucurbita moschata*) determined by SRAP and AFLP markers. Crop Sci. 44:653-664.

Ferriol M, Picó B, Nuez F (2003) Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. Theor Appl Genet. 107:271-282.

Ferriol M, Picó B, Nuez F (2004b) Morphological and molecular diversity of a collection of *Cucurbita maxima* landraces. J Am Soc Hortic Sci. 129:60-69.

García-Moreno MJ, Velasco L, Pérez-Vich B (2010) Transferability of non-genic microsatellite and gene-based sunflower markers to safflower. Euphytica. 175:145-150.

Gong L, Stift G, Kofler R, Pachner M, Lelley T (2008) Microsatellites for the genus Cucurbita and an SSR-based genetic linkage map of *Cucurbita pepo* L. Theor Appl Genet. 117:37-48.

Gong YM, Xu SC, Mao WH, Hu QZ, Zhang GW, Ding J, Li ZY (2010) Generation and characterization of 11 novel EST derived microsatellites from *Vicia faba* (Fabaceae). Am J Bot. 97:E69-E71.

Guo SG, Zheng Y, Joung JG, Liu SQ, Zhang ZH, Crasta OR, Sobral BW, Xu Y, Huang SW, Fei, ZJ (2010) Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. BMC Genomics. 11:384.

Gutierrez MV, Patto MCV, Huguet T, Cubero JI, Moreno MT, Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor Appl Genet.110: 1210-1217.

Gwanama C, Labuschagne MT, Botha AM. (2000) Analysis of genetic variation i*n Cucurbita moschata* by random amplified polymorphic DNA (RAPD) markers. Euphytica. 113:19–24.

Nagl N, Taski-Ajdukovic K, Popovic A, Curcic A, Danojevic D, Kovacev L (2011) Estimation of genetic variation among related sugar beet genotypes by using RAPD. Genetika-belgrade. 43:575–582.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol Biol Evol. 15:1275–1287.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 28:2537-2539.

Saeed A, Hovsepyan H, Darvishzadeh R, Imtiaz M (2011) Panguluri SK. Nazaryan, R. Genetic diversity of Iranian accessions, improved lines of chickpea (*Cicer arietinum* L.) and their wild relatives by using simple sequence repeats. Plant Mol Biol Rep. 29: 848–858.

Saha MC, Mian MAR, Eujayl I, Zwonitzer JC,Wang LJ, May GD (2004). Tall fescue EST-SSR markers with transferability across several grass species. Theor Appl Genet. 109: 783-791.

Sanjur OI, Piperno DR, Andres TC, Wessel-Beaver L (2002) Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from mitochondrial gene: implications for crop plant evolution and areas of origin. Proc Natl Acad Sci USA. 99:535–540.

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet. 100: 723-726.

Shete S, Tiwari H, Elston RC (2000) On estimating the heterozygosity and polymorphism information content value. Theor Popul Biol. 57:265-271.

Sim SC, Yu JK, Jo YK, Sorrells ME, Jung G (2009) Transferability of cereal EST-SSR markers. Genome. 52: 431-437.

Singh RK, Jena SN, Khan S, Yadav S, Banarjee N, Raghuvanshi S, Bhardwaj V, Dattamajumder SK, Kapur R, Solomon S (2013) Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane Gene. 524: 309-329.

Steward CN, Via LE (1993) A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. Biotechniques. 14:748-750.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23:48-55.

Wu JX, Chang ZJ, Wu QS, Zhan HX, Xie SL (2011) Molecular diversity of Chinese *Cucurbita moschata* germplasm collections detected by AFLP markers. Sci Hortic. 128:7-13.

Xie WG, Zhang XQ, Cai HW, Liu W, Peng Y (2010) Genetic diversity analysis and transferability of cereal EST-SSR markers to orchardgrass (*Dactylis glomerata* L). Biochem Syst Ecol. 38:740-749

Xu P, Wu XH, Luo J, Wang BG, Liu YH, Ehlers JD, Wang S, Lu ZF, Li GJ (2011) Partial sequencing of the bottle gourd genome reveals markers useful for phylogenetic analysis and breeding. BMC Genomics. 12:467.

Xu SC, Gong YM, Mao WH, Hu QZ, Zhang GW, Fu W, Xian, QQ (2012) Development and characterization of 41 novel EST-SSR markers for *Pisum sativum* (Leguminosae). Am J Bot. 99: E149-E153.