SCIENTIFIC REPORTS

natureresearch

Received: 3 May 2019 Accepted: 16 September 2019 Published online: 01 October 2019

OPEN Genetic Diversity and DNA **Fingerprints of Three Important Aquatic Vegetables by EST-SSR** Markers

Xingwen Zheng^{1,2}, Teng Cheng¹, Liangbo Yang², Jinxing Xu², Jiping Tang², Kegiang Xie², Xinfang Huang³, Zhongzhou Bao⁴, Xingfei Zheng¹, Ying Diao⁵, Yongning You¹ & Zhongli Hu¹

Twenty-two sacred lotus (Nelumbo nucifera), 46 taros (Colocasia esculenta) and 10 arrowheads (Sagittaria trifolia) were used as materials and combined with EST-SSR (expressed sequence tag-simple sequence repeats) primers developed by our laboratory. Core primers were screened from a large number of primers that were able to distinguish all materials with a high frequency of polymorphisms. Six pairs, twenty pairs and three pairs of core primers were screened from sacred lotus, taro, and arrowhead, respectively. The SSR fingerprints of these three important aquatic vegetables, producing 17-, 87- and 14-bit binary molecular identity cards, respectively, were separately determined by using the core primers. Since there were few core primers of sacred lotus and arrowhead, 3 and 9 primer pairs with higher polymorphic information content (PIC), respectively, were selected as candidate primers. These core and candidate primers were used to identify the purities of No.36 space lotus, Shandong 8502 taro and Wuhan arrowhead, which were 93.3% (84/90), 98.9% (89/90) and 100.0% (90/90), respectively. The fingerprints, displayed as binary molecular identification cards of three important aquatic vegetables, were obtained, and their purity was successfully determined with EST-SSR labeling technology. Phylogenetic trees were also constructed to analyze the genetic diversity of 22 sacred lotus, 46 taros and 10 arrowheads. This study classifies and identifies germplasm resources and is an important reference to test the authenticity and variety purity of other aquatic vegetables in the future.

With their rich nutrients, fantastic flavor and taste, excellent medicinal value and health care functions, aquatic vegetables have been continuously well received by consumers as important economic crops. There are various kinds of aquatic vegetables, of which N. nucifera, C. esculenta and S. trifolia are important and widely planted in China, Southeast Asia, Australia, and Oceania; these aquatic vegetables have high economic value and a long history of cultivation^{1,2}.

SSR fingerprints are inherent in genomes. SSR markers are not affected by external and internal environments, including growth and development time. The advantages of SSR markers are that they are low-cost, fast (within hours), accurate, and reliable. On the one hand, the authenticity of a variety can be effectively identified according to the expected specificity of the DNA fragment. On the other hand, the quality and purity of seeds can be determined by the proportion in the population. Currently, many crop varieties have DNA fingerprints constructed by SSR, such as rice³, corn⁴, wheat^{5,6}, cotton⁷, and potato⁸. In addition, a set of crop varieties have DNA fingerprints constructed based on EST-SSR markers, such as Chinese cabbage9, Cherry¹⁰, Hemarthria sibirica¹¹, Blueberry¹², Eriobotrya japonica¹³, and Asparagus¹⁴.

The authenticity of the variety and quality and purity of seeds are extremely significant for crop and vegetable production¹⁵. However, currently, seed purity identification in China is still dominated by morphological

¹State Key Laboratory of Hybrid Rice, Lotus Engineering Research Center of Hubei Province, College of Life Sciences, Wuhan University, Wuhan, 430072, P.R. China. ²Guangchang White Lotus Research Institute, Guangchang, Jiangxi, 344900, China. ³Wuhan National Germplasm Repository for Aquatic Vegetables, Wuhan Vegetable Institute of Science, Wuhan, 430065, Hubei, P.R. China. ⁴Suzhou Institute of Vegetables, Suzhou, 215008, Jiangsu, China. ⁵School of Forestry and Life Sciences, Chongqing University of Arts and Sciences, Chongqing, 402160, P.R. China. Correspondence and requests for materials should be addressed to Y.Y. (email: shangruzhenren(@163.com) or Z.H. (email: huzhongli@whu.edu.cn)

Rank	Primers	Repeats	Fragment size range (bp)	Number of polymorphic fragments	The value of PIC
1	NL-61	(AC)6	212-227	2	0.69
2	NL-28	(AT)6	255-261	3	0.67
3	NL-P8	(TTC)6	153-187	3	0.53
4	NL-35	(GT)6	156-162	3	0.74
5	NL-1	(AGTG)6	242-251	3	0.53
6	NL-P9	(AT)7	180-185	3	0.52

Table 1. Basic information of the six lotus core primers.

Rank	Primers	Repeats	Fragment size range (bp)	Number of polymorphic fragments	The value of PIC
1	NL-60	(AG)9	188-210	2	0.19
2	NL-47	(CT)8	225-280	5	0.74
3	NL-P23	(AT)9	248-275	6	0.77

Table 2. Basic information of the three lotus candidate primers.

identification¹⁶. The method of morphological identification is not only time-consuming and expensive but also subject to environmental factors. Therefore, morphological identification cannot quickly and accurately determine purity or variety. However, identification by SSR markers is accurate, can be accomplished within a short time, is low-cost, and is immune to environmental and self factors. Additionally, EST-SSR markers facilitate easy access to base information and metastasis among related species, with a low cost. Thus, EST-SSR markers have vast potential for future development¹⁷.

Due to the limited research on molecular markers, most aquatic vegetables do not have established molecular marker fingerprints or DNA fingerprints constructed from EST-SSR markers. Furthermore, there is no way to identify the authenticity of varieties or the quality and purity of seeds by molecular markers. Therefore, the purpose of this study is to construct DNA fingerprints for aquatic vegetables (sacred lotus, taro and arrowhead) by using a large number of EST-SSR molecular markers developed in our laboratory and to identify the purity of these aquatic vegetables. We aim to provide theoretical knowledge for the construction of DNA fingerprints and purity identification of other aquatic vegetables.

Results and Analysis

Screening of SSR core and candidate primers for three important aquatic vegetables. A total of 11,178 SSR sites were detected by MISA software, and 6,568 pairs of primers were designed by Primer3.0 software. We used a Perl script to synthesize 38 pairs of primers according to electronic polymorphism prediction, and seventy-two pairs of primers were randomly synthesized. Three hundred twenty-five clear bands were obtained by amplifying DNA from 22 scared lotus cultivars with 80 pairs of primers, with an average of 4 bands per pair of primers. From these 80 pairs of primers, 6 pairs of core SSR-labeled primers with polymorphisms that were able to distinguish all 22 materials were selected (Table 1). Three pairs of EST-SSR primers with high PIC (polymorphism information content) were selected as candidate primers for variety purity identification (Table 2). The PIC of each EST-SSR primer pair was calculated using the following formula: PIC = $1 - \left(\sum_{i=1}^{n} q_i^2\right) - \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2q_j^2 q_j^2\right)$, where *n* is the number of alleles and *q* is the *i*th and *j*th allele frequency¹⁸.

In total, 5,278 SSR loci were detected by MISA software, and 2,858 pairs of primers were designed by Primer3.0 software. We randomly synthesized 100 pairs of primers. Two hundred thirteen clear bands were obtained by amplifying DNA from 46 taro cultivars with 72 pairs of primers, with an average of 3 bands per pair of primers. From these 72 pairs of primers, 20 pairs of core EST-SSR primers with a polymorphism frequency that could distinguish all 46 taro materials were selected (Table 3).

A total of 3,861 SSR loci were detected by MISA software, and 2,476 pairs of primers were designed by Primer3.0 software. We randomly synthesized 100 pairs of primers. One hundred ninety-three clear bands were obtained by amplifying DNA from 10 arrowhead cultivars with 78 pairs of primers, with an average of 2.5 bands per pair of primers. From these 78 pairs of primers, 3 pairs of core EST-SSR primers with a polymorphism frequency that could distinguish all 10 arrowhead materials were selected (Table 4). Nine pairs of EST-SSR primers with high PIC were selected as candidate primers for variety purity identification (Table 5). For every aquatic vegetable, at least one figure of polymorphism showed in supplementary file 2.

DNA Fingerprints of three important aquatic vegetables. According to the sequences of the primer alleles presented in Table 4, the scored results, marked as "0" and "1" for the SSR molecular marker bands of sacred lotus core primers, were counted into binary data as 17-bit binary molecular identity cards of 22 sacred lotus cultivars (Table 6).

Rank	Primers	Repeats	Fragment size range (bp)	Number of polymorphic fragments	The value of PIC
1	CE-24	(GCT)5	248-275	9	0.80
2	CE-67	(AT)7	152-188	13	0.82
3	CE-8	(GAG)5	273-309	8	0.81
4	CE-19	(TC)8	142-217	7	0.74
5	CE-34	(CT)6	114-142	4	0.67
6	CE-17	(AT)9	193-210	5	0.63
7	CE-98	(CCT)6	206-217	4	0.61
8	CE-84	(TC)6	228-242	4	0.61
9	CE-59	(CGA)5	238-263	3	0.53
10	CE-57	(GGT)5	238-253	4	0.47
11	CE-78	(CTT)6	147-170	2	0.41
12	CE-100	(AG)7	206-231	3	0.40
13	CE-28	(TGC)5	242-273	4	0.36
14	CE-63	(AT)7	105-121	2	0.35
15	CE-2	(CGGTGA)5	180-195	4	0.33
16	CE-37	(CTC)5	276-282	3	0.08
17	CE-45	(CAT)5	191	1	-
18	CE-64	(AT)6	273	1	-
19	CE-27	(GCT)5	105-116	3	0.54
20	CE-47	(GA)7	100-115	3	0.24

Table 3. Basic information of the twenty taro core primers.

Rank	Primers	Repeats	Fragment size range (bp)	Number of polymorphic fragments	The value of PIC
1	SS-50	(GGA)5	171-196	7	0.80
2	SS-86	(AGC)5	270-287	5	0.62
3	SS-84	(TA)7	213-220	2	0.33

Table 4. Basic information of the three arrowhead core primers.

Rank	Primers	Repeats	Fragment size range (bp)	Number of polymorphic fragments	The value of PIC
1	SS-82	(TC)9	135-160	6	0.84
2	SS-62	(GCTG)5	180-189	4	0.75
3	SS-68	(CGA)5	278-290	5	0.72
4	SS-95	(GGA)6	127-136	3	0.67
5	SS-43	(TC)8	250-263	3	0.65
6	SS-41	(GCC)5	192-201	4	0.63
7	SS-93	(CGC)6	222-235	4	0.63
8	SS-88	(TGC)5	186-198	4	0.58
9	SS-80	(CTC)6	142-153	3	0.57

....

Table 5. Basic information of the nine arrowhead candidate primers.

According to the sequences of the primer alleles presented in Table 5, the scored results, marked as "0" and "1" for the SSR molecular marker bands of taro core primers, were counted into binary data as 87-bit binary molecular identity cards of 46 taro cultivars (Table 7).

According to the sequences of the primer alleles presented in Table 6, the scored results, marked as "0" and "1" for the SSR molecular marker bands of arrowhead core primers, were counted into binary data as 14-bit binary molecular identity cards of 10 arrowhead cultivars (Table 8).

Variety purity identification of three important aquatic vegetables. Among 90 samples of the No.36 space lotus, the bands of 84 samples were regarded as the same characteristic band of the cultivar. The other 6 samples judged to be different cultivars were compared with the characteristic band. Three samples of differential alleles had eight bands, and the other 3 samples had one band. The purity of No.36 space lotus was 93.3% (84/90).

Rank	Name	Binary molecular identity card	Rank	Name	Binary molecular identity card
1	Cunsan lotus	01100101001101010	12	Wufei lotus	10001110010110011
2	Furong lotus	10100101001001100	13	Zhuoshang lotus	10001100010101010
3	Baixiang lotus	10100101010001100	14	Jiande red flower lotus	10001001010001010
4	Hongxiang lotus	10100101010011100	15	Yixian lotus	10001011010011010
5	Chuzhou white lotus	10010001001001110	16	Qiushui changtian lotus	10001101100101000
6	Tuxuan lotus	10010101001001100	17	Chongtai lotus	10001101010101011
7	White flower jian lotus	10010101001001110	18	Donggua lotus	10001001100001100
8	Red flower jianlotus	01010100001100010	19	No.36 space lotus	10001101010001100
9	Baiye lotus	01001101010001100	20	Riza lotus III-1	10001100010001101
10	Jingguanglotus II	10001101001001100	21	Riza lotus III-2	01001100010001101
11	Star peony	01010100010100010	22	Wuzhi lotus II	01001101001100010

 Table 6. Binary molecular identity of twenty-two lotus varieties.

Among 90 samples of Shangdong 8502 taro, the bands of 89 samples were regarded as the same characteristic band of the cultivar. The remaining sample judged to be a different cultivar with 4 differential alleles was compared with the characteristic band. The purity of Shangdong 8502 taro was 98.9% (89/90).

The bands of all 90 Wuhan arrowhead samples were regarded as the same characteristic band of the cultivar, so the 90 samples were identified as the same variety. The purity of Wuhan arrowhead was 100% (90/90).

Genetic diversity analysis of three important aquatic vegetables. A phylogenetic tree of three important aquatic vegetables was constructed (Fig. 1). The Jaccard similarity coefficient ranged from 0.50 to 0.99. Twenty-two individual plants were separated into three major groups, I, II, and III, at a Jaccard similarity coefficient level of 0.65 in sacred lotus (Fig. 1A). The results showed that there was genetic differentiation between lotus with flowers and lotus with roots; lotus with roots was concentrated in group I, while lotus with flowers was mainly concentrated in group III. Forty-six taros were separated into four major groups at a Jaccard similarity coefficient level of 0.74 (Fig. 1B). Two multi-headed taro Jiangan Baba taro and Dongxiang Goutou taro were clustered together, which indicated that there was obvious genetic differentiation between multi-headed and other taro cultivars. Some head taros clustered together in group III, and some clustered together with other types of materials; more than half of the multi-cormels taros were clustered in group II. Arrowhead was separated into two groups (Fig. 1C). A dendrogram with a scale from 0.70 to 0.98 based on Jaccard's similarity coefficient was constructed and clearly separated the 8 accessions from Jiangsu and Zhejiang provinces into 4 main clusters, while the remaining Guangxi sand arrowhead and Guangzhou arrowhead were classified into group II.

Discussion

DNA fingerprints are generally constructed by three methods: the characteristic locus method, the single primer method and the core primer combination method, the most efficient of which is the core primer combination method^{4,19-22}. Core primers can play important roles in preliminary research, generally leading to a good polymorphism frequency, stability, reproducibility and discriminability. Identifying the the core primers of sacred lotus, taro and arrowhead is a key step for SSR fingerprint construction and purity identification, and it is also a prerequisite for the commercial application of SSR fingerprints. In this study, polymorphic SSR primers were first screened as alternative primers, and a set of SSR primers were found to be the core primers for these germplasm resources, with the smallest number that could completely distinguish all germplasm resources. According to the expansion and complexity of samples and genomes, the number of core primers should be adjusted accordingly. For variety identification, it is better to identify more samples with fewer primers; therefore, the ideal method is to distinguish the most varieties with a single primer. However, this was difficult to achieve, even in arrowhead, the cultivar with the smallest number. As a result, the core primer combination method was adopted in this study to construct fingerprints. In this study, 6 core primers and 3 candidate primers were selected for sacred lotus, 20 core primers for taro, and 3 core primers and 9 candidate primers for arrowhead. Then, 17-, 87- and 14-bit binary molecular identification cards were successfully constructed for 22 sacred lotus cultivars, 46 taro cultivars and 10 arrowhead cultivars, respectively. The DNA fingerprint of sacred lotus was more useful than those of the others^{23,24}. The construction of SSR fingerprints for these three important aquatic vegetables is still in its infancy, and in particular, SSR fingerprints for taro and arrowhead have not been reported. SSR fingerprints were constructed by the core primers combination method, which provides useful methodological guidance for the construction of a standard DNA fingerprint database and performing aquatic vegetable mapping analysis in the future.

As a kind of molecular marker with the advantages of high allelic variation, codominance, simple and rapid detection, and good stability, SSRs have been maturely applied in many regions, such as genetic diversity analysis, fingerprint construction, trait marker and genetic linkage map construction. Many agronomists and geneticists have carried out extensive research and application of SSRs²⁵. For example, 30 pairs of core primers were used to construct a fingerprint and test the authenticity and purity of Zhongmian Institute 63 cotton. Compared with field trial identification, the results showed that 30 pairs of core primers achieved good identification²⁶. From 36 pairs of primers, 3 core primers, MCPI-5, MCPI-16 and MCPI-17, were selected. Identified by MCPI-5 and MCPI-16, the purity of 267 watermelon T-1 hybrids was 97.75%, consistent with the field morphological identification results. Hence, the core primers selected could be used for purity identification of watermelon T-1 hybrids²⁷.

Rank	Name	Binary molecular identity card
1	Qionglai red-billed taro	0100011010000100010001000000000000001010
2	Xinjin white-billed taro	01001111000001000101001001000000111011101101111
3	Renshou Wuganqiang taro	01001111000001000101001001000000111011101111
4	Pengshan black stem taro	010011110000010001011010010000011000010000
5	Jiange red stem taro	0100111100000100010100100100000000011101101111
6	Dayi dryland taro	0100111010000100010001001000000010101010
7	Longmen dryland taro	01001111000001000100010110000001110111
8	Tongji dark stem taro	010011110000010001010101010100100111011101111
9	Chengbei Jintang taro	01001111000001000101010101000001110111
10	Baoning taro	0100111100101100100101010001000011011111
11	Baiqiao Jintang taro	0100111100000100010001011000000000000
12	Baoning red stem taro	01001111000001000101010101001001110111
13	Saijin taro	01001110101010001010010010000000010100111001000101
14	Xingwen dryland taro	01001111000001000101010101000001110111
15	Huilong taro	0100111100100100100101010000010010111111
16	Chengxi cyan stem taro	010011110000010001010101010100100111011101111
17	Yongan Wuganqiang taro	0100111100101100100101010001000011011111
18	Dazhu taro	01001111010001000101010101001001110111
19	Dongxiang red stem taro	010111110100010001010101010100100111011101111
20	Taihe red taro	01010111000001000100010110000001110111
21	Gongping black stem taro	0101111110000010001010101010100000111011101111
22	Lantian taro	0101111100000100010110100100000011000110001111
23	Qionglai black stem taro	0101111110000010001010101010100100111011101111
24	Jintang Wujiaoqing taro	01011111100000100010001011000000111011101111
25	Shuangliu wetland taro	0101111100101100100101010000000011111101111
26	Yumen Tangbao taro	0101111110001011001000010010000001110101
27	Xinchang betelnut taro	0101111010000100010100100100000001010101
28	Lipu taro	0101111110001011001000010010000001110101
29	Sichuan head taro I	0101111100000100101101001001001100001001111
30	Wanxing head taro	0101111110000010000011010110001000110101
31	Lujing big taro	0101111100100101000110101000100011010001111
32	Zhongpa head taro	010111111000001000100010010010011101010000
33	Yugan head taro	0101111100100101000110101100000001101001111
34	Qingfeng head taro	01011111001001010001101011000000011010001111
35	Jiangan Baba taro	010111110100010001110010010001000001111010
36	Dongxiang Goutou taro	0101111101000100011100100100000011110101
37	Ezhou taro	0101111100100100100101010000000110111111
38	WHU greenhouse taro	0100111100100100100101010001010011011111
39	WHU market taro	0100111001000100011100100011011100101111
40	Shandong 8502 taro	0100111101000100011100100011011100101111
41	Zhifang taro	01001111000001000101001000110100111011101101110000
42	WHU private plot taro	010000110010010010010101000000110111111
43	Paizhong taro	0100111100100100100101010000000110111111
44	Zhuji taro	0100111100100100100101010000000110111111
45	Zuonao Maliuyuan taro	01001111000001000101001000110000111011101111
46	Tangxun lake village taro	101011110010010010010000010010001101101

Table 7. Binary molecular identity of forty-six taro varieties.

Other crops, such as rice²⁸, rape²⁹, potato³⁰, maize³¹, eggplant^{32,33}, cabbage^{34,35}, and soybean³⁶, have also been able to be identified.

EST-SSR markers are new molecular markers based on the expression of sequence tag data combined with the characteristics of SSR. These markers combine the dual advantages of SSR and EST and expand the application prospects of SSR. In 2009, EST-SSR markers were applied to the identification of safflower purity, which became the first successful example of the use of EST-SSR markers for purity identification of a crop hybrid population³⁷. Subsequently, studies of the purity identification of rice, cabbage, melon (cucumber, melon, watermelon) and other crops have been carried out. The results showed that EST-SSR markers are superior and more sensitive than plant morphology identification in the field, and the purity of detection is generally 2–3% higher than the results from the field^{34,38–40}. However, the most commonly used method for the authenticity and purity detection

Rank	Name	Binary molecular identity card	Rank	Name	Binary molecular identity card
1	Baoying ziyuan arrowhead	00001100101001	6	Shengdang II arrowhead	00001000010011
2	Gaoyou great white arrowhead	00010010100011	7	Kunming arrowhead	10001000010011
3	Jiangxing produce shendang arrowhead	00001000110011	8	Guangxi sand arrowhead	00001001010001
4	Suzhou yellow arrowhead	00001000110001	9	Guangzhou arrowhead	00001010001001
5	Ziyuan II arrowhead	01000100110011	10	Wuhan arrowhead	00101000010101

Table 8. Binary molecular identity of ten arrowhead varieties.

.....

of aquatic vegetables is still morphological identification, which is accomplished by comparison with standard morphological characteristics or determination of morphological characteristic consistency in tested varieties. Due to the limited molecular genetics research, there are a few available SSR markers and no reports of the use of EST-SSR markers to detect the authenticity and purity of aquatic vegetables. Therefore, based on lotus, taro and arrowhead cultivars from all over China, the purity of No.36 space lotus, Shandong 8502 taro and Wuhan arrowhead was determined by EST-SSR and core primer PCR amplification. The purities of these three cultivars were 93.3%, 98.9% and 100%, respectively. It can be seen that these varieties all developed with rigorous breeding and screening. With less genetic variation and higher consistency, the three important aquatic vegetables were close to sacred lotus, taro and arrowhead pure varieties. Because of the genetic differentiation in this cultivar, No.36 space lotus with lower purity is a hybrid cultivar.

In this study, the genetic relationship of a lotus germplasm resources is not related to its geographical origin, as reported in previous reports⁴¹. At present, lotus can be divided into three types according to its morphological characteristics and its main parts, namely, lotus with seeds, lotus with flowers and lotus with roots⁴¹⁻⁴³. The differentiation of lotus with roots and lotus with flowers in this study was consistent with previous research⁴⁴. There is a certain correlation between the genetic relationship and phenotypic traits of lotus. This type of diversity evaluation will provide important reference materials for breeding research of these cultivars in the future. Forty-six taros were separated into four major groups. The main branches were a mixture of materials from different regions and types, and this result indicated that there is no obvious correlation between morphological characteristics or geographic distribution and kinship, possibly due to the cultivar classification of taro based on the variation of its morphological characteristics, which was probably caused by mutations of a few major genes. There is a certain correlation between the regional differentiation of arrowhead cultivars and their genetic differentiation. In Jiangsu and Zhejiang provinces, which belong to the Yangtze River basin, there is a certain differentiation of arrowhead cultivars, and the results of this study suggest that arrowhead is limited by region, especially water limitation. Lower gene exchange led to abundant genetic diversity, which was beneficial to arrowhead breeding. This study can server as a significant reference for the authenticity and purity detection of aquatic vegetable cultivars. Moreover, EST-SSR markers have higher species transferability⁴⁵ and can also be used to study other related species.

Materials and Methods

Plant materials. Twenty-two sacred lotus cultivars were used to construct SSR fingerprints, including 14 sacred lotus with seeds cultivars, 7 sacred lotus with flowers cultivars and 1 sacred lotus with root cultivar (Table 9). Forty-six taro cultivated varieties (Table 10) and 10 arrowhead cultivars (Table 11) were used to construct SSR fingerprints.

Extraction of genomic DNA from three important aquatic vegetables. From the above young leaves of these materials, genomic DNA⁴⁶ was extracted by the modified CTAB method. Among the materials listed above, genomic DNA from the varieties used for purity identification was collected and extracted from the young leaves of 90 plant samples. DNA was extracted by 1.0% agarose gel electrophoresis and stored at -20 °C for later use.

Primer amplification. SSR marker detection was performed using a Perl program known as MicroSatellite (MISA, http://pgrc.ipk-gatersleben.de/misa) from the S. sagittifolia transcriptome, C. esculenta transcriptome and N. nucifera transcriptome. Using effective EST-SSR primers developed in our laboratory, genomic DNA was amplified by PCR. The primer sequences for sacred lotus and arrowhead are shown in supplementary file 1. Based on the obtained bands, EST-SSR markers were selected as the core primers with a high polymorphism frequency that could distinguish all cultivars. These core primers were employed to amplify each of the 90 samples and 5 control samples of genomic DNA of No.36 space sacred lotus, Shandong 8502 taro, and Wuhan arrowhead. Because of the lack of core primers for sacred lotus and arrowhead, 3 and 9 pairs of candidate primers were added to the samples. The reaction mixture (15μ L) contained 1.5μ L of buffer ($10 \times$ dilution), 0.8μ L of MgCl₂ (20 mM), 0.6 µL of dNTP (10 mM), 0.5 µL of TaqDNA polymerase, 1 µL of forward primer (10 µM), 1 µL of reverse primer $(10 \mu M)$, $2 \mu L$ of template DNA and 7.6 μL of ddH₂O. The amplification conditions included an initial denaturation step at 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, and annealing extension at 54-65 °C, elongation at 72 °C for 30 s and a final step at 72 °C for 5 min. Ten microliters of PCR bromophenol blue buffer was added to the PCR products and placed in the gene amplification instrument (BIOER co.,LTD., Hangzhou, China) at 94 °C to degenerate for 5 min, then quickly cooled on ice and tested with a polyacrylamide gel. The silver staining method was used to display the color, and the electrophoresis results were observed on a slide lamp.



Figure 1. Dendrogram showing the genetic relationships of three important aquatic vegetables. (**A**) Genetic relationships among 22 individual sacred lotus plants based on EST-SSR markers. (**B**) Genetic relationships among 46 individual taros based on EST-SSR markers. (**C**) Genetic relationships among 10 arrowheads based on EST-SSR markers.

Data statistics and analysis. According to the polyacrylamide electrophoresis results, a band of the same fragment size was recorded as a marker allele and scored as 1, 0 and 2 for a band, no band and a deletion, respectively. The EST-SSR locus allelic variation was counted for sacred lotus, taro and arrowhead, and the statistical data were input into a computer. Data analyses were performed by using the NTSYSpc package version 2.1⁴⁷. The SSR markers were identified as core primers with indicated polymorphisms that could distinguish all cultivars. A set of binary data were obtained from the "0" and "1" data produced by these markers as the binary molecular identity cards of these cultivars.

The characteristic band of a variety is the common characteristic band of most sample individuals. The band of each sample was compared with the characteristic band, and different allele numbers between cultivars

Rank	Name	Category	Source	Rank	Name	Category	Source
1	Cunsan lotus**	Lotus with seeds	Xiangtan, Hunan	12	Wufei lotus**	Lotus with flowers	Beijing
2	Furong lotus	Lotus with seeds	Xiangtan, Hunan	13	Zhuoshang lotus	Lotus with flowers	Beijing
3	Baixiang lotus	Lotus with seeds	Xiangtan, Hunan	14	Jiande red flower lotus	Lotus with flowers	Jiande, Zhejiang
4	Hongxiang lotus	Lotus with seeds	Xiangtan, Hunan	15	Yixian lotus	Lotus with flowers	Nanjing, Jiangsu
5	Chuzhou white lotus	Lotus with seeds	Jinhua, Zhejiang	16	Qiushui changtian lotus	Lotus with flowers	Guangchang, Jiangxi
6	Tuxuan lotus	Lotus with seeds	Jinhua, Zhejiang	17	Chongtai lotus	Lotus with flowers	Wuhan, Hubei
7	White flower jian lotus**	Lotus with seeds	Jianning, Fujian	18	Donggua lotus**	Lotus with flowers	Changsha, Hunan
8	Red flower jianlotus	Lotus with seeds	Jianning, Fujian	19	No.36 space lotus*	Lotus with seeds	Guangchang, Jiangxi
9	Baiye lotus	Lotus with seeds	Guangchang, Jiangxi	20	Riza lotus III-1	Lotus with seeds	Wuhan, Hubei
10	Jingguanglotus II	Lotus with seeds	Guangchang, Jiangxi	21	Riza lotus III-2	Lotus with seeds	Wuhan, Hubei
11	Star peony	Lotus with seeds	Guangchang, Jiangxi	22	Wuzhi lotus II**	lotus with root	Wuhan, Hubei

Table 9. Twenty-two lotus varieties used in this study. *90 materials for purity identification, **Control sample for purity identification.

Rank	Name	Category	Area	Rank	Name	Category	Area
1	Qionglai red-billed taro	Taro with multi-cormels	Yangan, Qionglai, Sichuan	24	Jintang Wujiaoqing taro	Taro with multi-cormels	Qingjiang, Jintang, Sichuan
2	Xinjin white-billed taro	Taro with multi-cormels	Wanxing, Xinjin, Sichuan	25	Shuangliu wetland taro	Taro with multi-cormels	Jitian, Shuangliu, Sichuan
3	Renshou Wuganqiang taro	Taro with multi-cormels	Shigao, Renshou, Sichuan	26	Yumen Tangbao taro**	Taro with multi-cormels	Yumen, Yanbian, Sichuan
4	Pengshan black stem taro	Taro with multi-cormels	Pengshan, Sichuan	27	Xinchang betelnut taro	Head taro	Xinchang, Dayi, Sichuan
5	Jiange red stem taro	Taro with multi-cormels	Jiange, Guangyuan, Sichuan	28	Lipu taro	Head taro	Jiangyang, Luzhou, Sichuan
6	Dayi dryland taro **	Taro with multi-cormels	Anren, Dayi, Sichuan	29	Sichuan head taro I	Head taro	Chengdu, Sichuan
7	Longmen dryland taro	Taro with multi-cormels	Longmen, Mianyang, Sichuan	30	Wanxing head taro	Head taro	Wanxing, Xinjin, Sichuan
8	Tongji dark stem taro	Taro with multi-cormels	Tongji, Zhongjiang, Sichuan	31	Lujing big taro	Head taro	Lujing, Yingshan, Sichuan
9	Chengbei Jintang taro	Taro with multi- cormels	Chengbei, Jiange, Sichuan	32	Zhongpa head taro	Head taro	Zhongpa, Yuechi, Sichuan
10	Baoning taro	Taro with multi-cormels	Baoning, Langzhong, Sichuan	33	Yugan head taro**	Head taro	Yugan, Yanbian, Sichuan
11	Baiqiao Jintang taro	Taro with multi-cormels	Baiqiao, Cangxi, Sichuan	34	Qingfeng head taro	Head taro	Qingfeng, Jianyang, Sichuan
12	Baoning red stem taro	Taro with multi-cormels	Baoning, Langzhong, Sichuan	35	Jiangan Baba taro	Multi-headed taro	Jiangan, Yibin, Sichuan
13	Saijin taro	Taro with multi-cormels	Saijin, Yilong, Sichuan	36	Dongxiang Goutou taro**	Multi-headed taro	Dongxiang, Yihan, Sichuan
14	Xingwen dryland taro	Taro with multi-cormels	Xingwen, Bazhou, Sichuan	37	Ezhou taro**	Taro	Ezhou, Hubei
15	Huilong taro	Taro with multi-cormels	Huilong, Yingshan, Sichuan	38	WHU greenhouse taro	Taro	Wuhan University, Hubei
16	Chengxi cyan stem taro	Taro with multi-cormels	Chengxi, Dazhu, Sichuan	39	WHU market taro	Taro	Wuhan University, Hubei
17	Yongan Wuganqiang taro	Taro with multi-cormels	Yongan, Gaoping, Sichuan	40	Shandong 8502 taro *	Taro	Shandong
18	Dazhu taro	Taro with multi-cormels	Dazhu, Daxian, Sichuan	41	Zhifang taro	Taro	Zhifang, Jiangxia, Hubei
19	Dongxiang red stem taro	Taro with multi-cormels	Dongxiang, Xuanhan, Sichuan	42	WHU private plot taro	Taro	Wuhan University, Hubei
20	Taihe red taro	Taro with multi-cormels	Taihe, Shehong, Sichuan	43	Paizhong taro	Taro	Paizhong, Xiantao, Hubei
21	Gongping black stem taro	Taro with multi-cormels	Gongping, Zizhong, Sichuan	44	Zhuji taro	Taro	Zhuji, Xiantao, Hubei
22	Lantian taro	Taro with multi-cormels	Jiangyang, Luzhou, Sichuan	45	Zuonao Maliuyuan taro	Taro	Louhe, Xiantao, Hubei
23	Qionglai black stem taro	Taro with multi-cormels	Qionglai, Sichuan	46	Tangxun lake village taro	Taro	Hongshan, Wuhan, Hubei

Table 10. Forty-six taro varieties used in this study. *90 materials for purity identification, ** Control sample for purity identification.

Rank	Name	Category	Area
1	Baoying ziyuan arrowhead	Ziyuan arrowhead	Baoying, Jiangsu
2	Gaoyou great white arrowhead**	Great white arrowhead	Gaoyou, Jiangsu
3	Jiangxing produce shendang arrowhead	Shendang arrowhead	Jiangxing, Zhejiang
4	Suzhou yellow arrowhead	Suzhou yellow arrowhead	Suzhou vegetable research institute, Jiangsu
5	Ziyuan II arrowhead**	Ziyuan arrowhead	Suzhou vegetable research institute, Jiangsu
6	Shengdang II arrowhead	Shengdang arrowhead	Suzhou vegetable research institute, Jiangsu
7	Kunming arrowhead**	Kunming arrowhead	Kunming, Yunnan
8	Guangxi sand arrowhead**	Guangxi sand arrowhead	Baipeng, Liujiang, Guangxi
9	Guangzhou arrowhead**	Guangzhou arrowhead	Guangzhou, Guangdong
10	Wuhan arrowhead*	Wuhan arrowhead	Wuhan University, Hubei

Table 11. Ten arrowhead varieties used in this study. *90 materials for purity identification, **Control sample for purity identification.

(>1) were determined to indicate different varieties, while (=0) was determined to indicate the same variety. $P = X/Y \times 100\%$, where X is the number of individuals of the same variety, Y is the total number of individuals identified and P is the value of purity.

References

- Kreike, C. M., Van Eck, H. J. & Lebot, V. Genetic diversity of taro, Colocasia esculenta (L.) Schott, in Southeast Asia and the Pacific. *Theoretical and Applied Genetics* 109, 761–768, https://doi.org/10.1007/s00122-004-1691-z (2004).
- 2. Li, O.-J. et al. in Advances in Applied Biotechnology 637–642 (Springer, 2015).
- 3. Chakravarthi, B. K. & Naravaneni, R. SSR marker based DNA fingerprinting and diversity study in rice (Oryza sativa. L). African Journal of Biotechnology 5 (2006).
- 4. Wang, F.-G. *et al.* Development and characterization of a core set of SSR markers for fingerprinting analysis of Chinese maize varieties. *Maydica* **56** (2011).
- Hayden, M. J., Khatkar, S. & Sharp, P. Targetting microsatellites (SSRs) in genetic linkage maps of bread wheat. Australian Journal of Agricultural Research 52, 1143–1152 (2001).
- 6. Wendehake, K. *et al.* Construction and analysis of a microsatellite-based database of European wheat varieties. *Theoretical and Applied Genetics* **106**, 67–73 (2002).
- Meng, K., Weihua, Y. & Hongxia, X. Construction of DNA fingerprinting and analysis of genetic diversity with SSR markers for cotton major cultivars in China. *China Agriculture Science* (2011).
- Duan, Y. et al. Construction of fingerprinting and analysis of genetic diversity with SSR markers for eighty-eight approved potato cultivars (Solanum tuberosum L.) in China. Acta agronomica sinica 35, 1451–1457 (2009).
- Li, L. et al. Construction Chinese cabbage (Brassica rapa L.) core collection and its EST-SSR fingerprint database by EST-SSR molecular markers. Genomics and Applied Biology 28, 76–88 (2009).
- Cantini, C., Iezzoni, A. F., Lamboy, W. F., Boritzki, M. & Struss, D. DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. *Journal of the American Society for Horticultural Science* 126, 205–209 (2001).
- Huang, L. et al. Constructing DNA fingerprinting of Hemarthria cultivars using EST-SSR and SCoT markers. Genetic resources and crop evolution 61, 1047–1055 (2014).
- Rowland, L. J. et al. Construction of a genetic linkage map of an interspecific diploid blueberry population and identification of QTL for chilling requirement and cold hardiness. *Molecular Breeding* 34, 2033–2048, https://doi.org/10.1007/s11032-014-0161-9 (2014).
- 13. He, Q. et al. Genetic diversity and identity of Chinese loquat cultivars/accessions (Eriobotrya japonica) using apple SSR markers. *Plant Molecular Biology Reporter* 29, 197–208 (2011).
- Caruso, M., Federici, C. T. & Roose, M. L. EST–SSR markers for asparagus genetic diversity evaluation and cultivar identification. *Molecular Breeding* 21, 195–204 (2008).
- 15. George, R. A. Vegetable seed production. (CABI, 2009).
- 16. Zhao, Z. et al. Genetic purity testing of loose-curd cauliflower hybrids using SSR markers and grow out test. Seed Science and Technology 40, 209–214 (2012).
- 17. Chapman, M. A. et al. Development, polymorphism, and cross-taxon utility of EST-SSR markers from safflower (Carthamus tinctorius L.). Theoretical and Applied Genetics 120, 85–91 (2009).
- Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics* 32, 314 (1980).
- 19. Fang, D. & Roose, M. Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theoretical and Applied Genetics* **95**, 408–417 (1997).
- 20. Archak, S. *et al.* DNA fingerprinting of Indian cashew (Anacardium occidentale L.) varieties using RAPD and ISSR techniques. *Euphytica* **130**, 397–404 (2003).
- 21. Hao, Q. et al. Studies on Paeonia cultivars and hybrids identification based on SRAP analysis. Hereditas 145, 38-47 (2008).
- Bushman, B. S. et al. Orchardgrass (Dactylis glomerata L.) EST and SSR marker development, annotation, and transferability. Theoretical and applied genetics 123, 119–129 (2011).
- 23. Kong, D.-z., LIU, Y.-P. & Yang, Q.-S. Identification of eight small lotus cultivars by DNA fingerprints [J]. Journal of Nanjing Agricultural University 2 (2009).
- Yang, L. et al. Application of SSR Fluorescent Labeling in the Construction of Fingerprint for 80 Cultivars in Lotus (Nelumbo spp.). Molecular Plant Breeding 3, 986r–993 (2017).
- 25. Sundaram, R. *et al.* Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica* **163**, 215–224 (2008).
- 26. Fu, X. & Yang, F. Identification Genuineness and purity of CCRI 63 Using SSR. Markers. China Cotton 43, 17-20 (2016).
- Jiang, L., Zhang, G. & Zhang, J. Purity identification of watermelon hybrid combination T-1 seeds using SSR molecular marker. Journal of Northwest A&F University 44, 93–98 (2016).
- Ye-Yun, X., Zhan, Z., Yi-Ping, X. & Long-Ping, Y. Identification and purity test of super hybrid rice with SSR molecular markers. *Rice. Science* 12, 7–12 (2005).
- Tommasini, L. et al. The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity
 and stability testing of rape (Brassica napus L.) varieties. Theoretical and Applied Genetics 106, 1091–1101 (2003).
- Ghislain, M. et al. Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theoretical and applied Genetics 108, 881–890 (2004).
- Wu, M., Jia, X., Tian, L. & Lv, B. Rapid and reliable purity identification of F 1 hybrids of Maize (Zea may L.) using SSR markers. Maize Genomics and Genetics 1 (2010).
- Kumar, A., Dadlani, M., Kumar, R. & Jacob, S. R. Identification and validation of informative SSR markers suitable for ensuring the genetic purity of brinjal (Solanum melongena L.) hybrid seeds. *Scientia Horticulturae* 171, 95–100 (2014).
- Jha, N., Jacob, S., Nepolean, T., Jain, S. & Kumar, M. SSR markers based DNA fingerprinting and it's utility in testing purity of eggplant hybrid seeds. Quality Assurance and Safety of Crops & Foods 8, 333–338 (2016).
- Liu, G. et al. Seed genetic purity testing of F1 hybrid cabbage (Brassica oleracea var. capitata) with molecular marker analysis. Seed science and technology 35, 477–486 (2007).
- 35. Ye, S. *et al.* Genetic purity testing of F1 hybrid seed with molecular markers in cabbage (Brassica oleracea var. capitata). *Scientia Horticulturae* **155**, 92–96 (2013).
- 36. Sun, H. et al. Application of SSR Markers for Purity Testing of Commercial Hybrid Soybean (Glycine max L.) (2018).
- Naresh, V., Yamini, K., Rajendrakumar, P. & Kumar, V. D. EST-SSR marker-based assay for the genetic purity assessment of safflower hybrids. *Euphytica* 170, 347–353 (2009).
- 38. Liu, L. *et al*. Evaluation of genetic purity of F1 hybrid seeds in cabbage with RAPD, ISSR, SRAP, and SSR markers. *Hortscience* **42**, 724–727 (2007).
- Moorthy, K. K. et al. Identification of informative EST-SSR markers capable of distinguishing popular Indian rice varieties and their utilization in seed genetic purity assessments. Seed Science and Technology 39, 282–292 (2011).
- Kwon, Y. Use of EST-SSR markers for genetic characterisation of commercial watermelon varieties and hybrid seed purity testing. Seed Science and Technology 41, 245–256 (2013).

- Guo, H., Li, S., Peng, J. & Ke, W. Genetic diversity of Nelumbo accessions revealed by RAPD. Genetic Resources and Crop Evolution 54, 741–748 (2007).
- 42. Nguyen, Q. Lotus for export to Asia: an agronomic and physiological study. A report for the Rural Industries Research and Development Corporation. NSW Agriculture. *Horticultural Research & Advisory Station. Gosford (NSW)*, 4–5 (2001).
- Guo, H. Cultivation of lotus (Nelumbo nucifera Gaertn. ssp. nucifera) and its utilization in China. Genetic Resources and Crop Evolution 56, 323–330 (2009).
- Han, Y.-C. et al. Analyses of genetic relationships in Nelumbo nucifera using nuclear ribosomal ITS sequence data, ISSR and RAPD markers. Aquatic Botany 87, 141–146 (2007).
- 45. Lindqvist, C. et al. An expressed sequence tag (EST) library from developing fruits of an Hawaiian endemic mint (Stenogyne rugosa, Lamiaceae): characterization and microsatellite markers. BMC Plant Biology 6, 16 (2006).
- 46. Doyle, J. J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11-15 (1987).
- 47. Rohlf, F. J. NTSYS-pc: numerical taxonomy and multivariate analysis system. (Exeter Publishing, 1988).

Acknowledgements

This research was supported by National Key Technologies R&D Program (No. 2012BAD27B01).

Author Contributions

X. Zheng designed and performed the experiments, analyzed the results, drew the figures and wrote. the manuscript. T. Cheng worked on some experiments. L. Yang, J. Xu, J. Tang, K. Xie and Z. Bao provided some materials. X. Zheng, X. Huang and Y. Diao provided some scientific advice and corrections. Y. You and Z. Hu analyzed the results and approved the final manuscript. All authors have read and approved the manuscript in its final form.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-50569-3.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019