Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Development of 107 SSR markers from whole genome shotgun sequences of Chinese bayberry (*Myrica rubra*) and their application in seedling identification*#

Hui-min JIA[1], Yu-tong SHEN[1], Yun JIAO[1], Guo-yun WANG[2], Xiao DONG[1], Hui-juan JIA[1], Fang DU[1], Sen-miao LIANG[3], Chao-chao ZHOU[4], Wei-hua MAO[5], Zhong-shan GAO†‡[1]

([1]*Department of Horticulture, Zhejiang University, Hangzhou 310058, China*)
([2]*Fruit Research Institute, Yuyao 315400, China*)
([3]*Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China*)
([4]*Yuyao Agricultural and Forestry Bureau, Yuyao 315400, China*)
([5]*Bio-Macromolecules Analysis Lab, Analysis Center of Agrobiology and Environmental Sciences, Zhejiang University, Hangzhou 310058, China*)
†E-mail: gaozhongshan@zju.edu.cn

Received Feb. 21, 2014; Revision accepted July 15, 2014; Crosschecked Oct. 15, 2014

Abstract: Chinese bayberry (*Myrica rubra* Sieb. et Zucc.) is one of the important subtropical fruit crops native to the South of China and Asian countries. In this study, 107 novel simple sequence repeat (SSR) molecular markers, a powerful tool for genetic diversity studies, cultivar identification, and linkage map construction, were developed and characterized from whole genome shotgun sequences. M13 tailing for forward primers was applied as a simple method in different situations. In total, 828 alleles across 45 accessions were detected, with an average of 8 alleles per locus. The number of effective alleles ranged from 1.22 to 10.41 with an average of 4.08. The polymorphic information content (PIC) varied from 0.13 to 0.89, with an average of 0.63. Moreover, these markers could also be amplified in their related species *Myrica cerifera* (syn. *Morella cerifera*) and *Myrica adenophora*. Seventy-eight SSR markers can be used to produce a genetic map of a cross between 'Biqi' and 'Dongkui'. A neighbor-joining (NJ) tree was constructed to assess the genetic relationships among accessions, and the elite accessions 'Y2010-70', 'Y2012-140', and 'Y2012-145', were characterized as potential new genotypes for cultivation.

Key words: Chinese bayberry, Simple sequence repeat (SSR), Genetic diversity
doi:10.1631/jzus.B1400051 **Document code:** A **CLC number:** S667.6

1 Introduction

Chinese bayberry (*Myrica rubra* Sieb. et Zucc.), a member of the Myricaceae, is native to China and neighboring Asian countries and has been cultivated in southern China for more than two thousand years (Chen *et al.*, 2004). Of the six *Myrica* species in China, only *M. rubra* is commercially cultivated for fresh fruit and processed products, especially in Zhejiang Province, which has the best known cultivars 'Biqi' and 'Dongkui'. In 2013, the cultivated area was over 340000 ha, and the annual production reached 1.2 million tons. Studies have shown that the fruit is rich in anthocyanins, which have a wide range of pharmacological properties (Zhang *et al.*, 2011) and can also be used in the food industry to replace

‡ Corresponding author
* Project supported by the Ningbo Science and Technology Bureau (No. 2012C10012), the Science and Technology Project of Zhejiang Province (No. 2012C12904), and the Yuyao Financial Bureau for Chinese Bayberry Breeding Program, China
# Electronic supplementary materials: The online version of this article (http://dx.doi.org/10.1631/jzus.B1400051) contains supplementary materials, which are available to authorized users
© Zhejiang University and Springer-Verlag Berlin Heidelberg 2014

synthetic dyes (Bao *et al.*, 2005). The bayberry fruit is usually consumed fresh, but can also be canned, and is used in products such as juice, jam, and wine. For the fresh market, Chinese bayberry is bred for good shipping qualities, color, appearance, size, and flavor, while the processing market always needs cultivars to meet product specifications such as low drip loss, exceptional flavor, and high soluble solid content. Chinese bayberry is regarded as a highly nutrient fruit, and the adverse reaction to this fruit is rare (Wang *et al.*, 2012).

There are abundant germplasm resources of Chinese bayberry in China, and 305 accessions have been recorded with 268 named as cultivars (Zhang *et al.*, 2009b). However, less than 50 cultivars have been planted as a pure line on a large scale, and only two cultivars, 'Biqi' and 'Dongkui', have been widely exported from Zhejiang to other provinces. There are many lines within a cultivar group in some local regions. New cultivar selections of Chinese bayberry largely depend on seedling identification and sporadic sport mutations of current cultivars (Xie *et al.*, 2011). Fruit color is a straightforward trait to select: bright red is usually more attractive than black, and nowadays, light yellow and pink fruit also attract consumers' attention. Recent initiatives to select elite accessions from the natural individuals within a cultivar group, with variable fruit maturity date, high and stable yield, larger fruit size and different flavor, cater to the needs of fresh consumption or processing. Using a molecular marker to identify the new breeding line is needed before regional testing of cultivars.

Through the long cultivation history of Chinese bayberry, there has been no report on cross-breeding for cultivar improvement. We have previously published a report on pollen from a mutated branch of cultivars (usually female plants), which allow crossing between cultivars (Jiao *et al.*, 2013). As a result of crossing 'Biqi'×'Dongkui' in 2011, we obtained offspring plantlets in 2013. Because fruit trees are large and take many years to produce fruit, the identification of molecular markers linked to the gender would confer tremendous advantages in the breeding and female genotype selection. The construction of a genetic linkage map of Chinese bayberry is a first step towards this long-term goal.

Simple sequence repeat (SSR) markers, which are co-dominant markers, are ubiquitous in genomes of plant species, highly polymorphic and reproducible, and follow the Mendelian inheritance. Nowadays, it is much easier to develop SSR markers from genome or transcriptome databases (Jiao *et al.*, 2012; Yue *et al.*, 2014). They have been routinely used in the construction of molecular linkage maps, and in genetic diversity analysis, marker-assisted selection, fingerprinting and cultivar identification in fruit trees such as apple (Celton *et al.*, 2009) and peach (Testolin *et al.*, 2000). *M. rubra* is dioecious, resulting in individual plants being highly heterozygous. Consequently, the SSR marker system would be useful in the construction of genetic maps. Thirteen polymorphic microsatellite loci have been developed from a genomic library in *M. rubra* (Terakawa *et al.*, 2006), and eleven from a library of expressed sequence tags (ESTs) (Zhang *et al.*, 2009a). With the development of next-generation sequencing technology, a fast and cost-effective approach to develop SSR markers from a whole genome survey has been reported (Jiao *et al.*, 2012). One hundred and nine EST-SSRs developed from the transcriptome have been published (Zhang *et al.*, 2012). However, the number of markers was insufficient to construct a high density genetic linkage map for breeding purposes. In this paper, we identify additional polymorphic SSR markers from whole genome shotgun sequences in *M. rubra* and use these markers to analyze the genetic diversity of Chinese bayberry and to identify new cultivars. Our objective is to develop more SSRs from the genomic sequences to construct an SSR-based linkage map.

2 Materials and methods

2.1 Plant materials and DNA extraction

A set of 45 accessions was used to test polymorphism in newly developed SSR markers. This set comprised 43 accessions of the cultivated species (*M. rubra*) and two related species (*M. adenophora* and *M. cerifera*) collected from different provinces in China (Table 1).

Young leaves were collected and frozen in liquid nitrogen and total genomic DNA was extracted from 1 g frozen leaf tissue with a modified version of the cetyltrimethylammonium bromide (CTAB) procedure described by Zhang *et al.* (2009b). The DNA concentration was diluted to 20 ng/μl based on the quality and quantity assessed on 0.01 g/ml agarose gels using standard DNA markers (TaKaRa, Dalian, China).

Table 1 Forty-five bayberry accessions included in this study

No.	Accession	Sex	Fruit/flower color[a]	Region
1	Biqi	♀	Black	Cixi, Ningbo, Zhejiang, China
2	Biqi12	♀	Black	Yuyao, Ningbo, Zhejiang, China
3	C2010-4	♀♂	Red	Cixi, Ningbo, Zhejiang, China
4	C2010-55	♂	Red	Cixi, Ningbo, Zhejiang, China
5	Chise	♀	Red	Hangzhou, Zhejiang, China
6	Dafudayexidi	♀	Black	Suzhou, Jiangsu, China
7	Dahongpao	♀	Red	Huzhou, Zhejiang, China
8	Daji	♀	Red	Suzhou, Jiangsu, China
9	Dayeguang	♀	Black	Huzhou, Zhejiang, China
10	Dingao	♀	Black	Wenzhou, Zhejiang, China
11	Dongkui	♀	Red	Taizhou, Zhejiang, China
12	Fenhong	♀	Pink	Yuyao, Ningbo, Zhejiang, China
13	Guangdongdamei	♀	Red	Guangzhou, Guangdong, China
14	Guangdongheimei	♀	Black	Guangzhou, Guangdong, China
15	GX2011-22	♂	Orange	Guilin, Guangxi, China
16	Heiruilin	♀	Black	Taibei, Taiwan, China
17	Jinqiantan	♀	Black	Hangzhou, Zhejiang, China
18	Liuyemei	♀	Red	Jinhua, Zhejiang, China
19	*Myrcia adenophora*	♀	Red	Hechi, Guangxi, China
20	*Myrica cerifera*	♂	Yellow	Cixi, Ningbo, Zhejiang, China
21	Niuyemei	♀	Black	Jingzhou, Hunan, China
22	Ruiguangmei	♀	Black	Japan
23	Shangchongmei	♀	Black	Jingzhou, Hunan, China
24	Shuijing	♀	Light yellow	Yuyao, Ningbo, Zhejiang, China
25	Songjiang	♀	Black	Cixi, Ningbo, Zhejiang, China
26	Songmaoli	♀	Black	Hangzhou, Zhejiang, China
27	Tumei	♀	Black	Wenzhou, Zhejiang, China
28	W2011-1	♂	Yellow	Wenzhou, Zhejiang, China
29	Wenfangmei	♀	Red	Guixi, Jiangxi, China
30	Wenlinbenmei	♀	Black	Taizhou, Zhejiang, China
31	Xianghong	♀	Red	Hangzhou, Zhejiang, China
32	Y2010-70	♀	Red	Yuyao, Ningbo, Zhejiang, China
33	Y2012-1	♂	Red	Yuyao, Ningbo, Zhejiang, China
34	Y2012-137	♀	Black	Yuyao, Ningbo, Zhejiang, China
35	Y2012-139	♀	Black	Yuyao, Ningbo, Zhejiang, China
36	Y2012-140	♀	Pink	Yuyao, Ningbo, Zhejiang, China
37	Y2012-142	♀	Red	Yuyao, Ningbo, Zhejiang, China
38	Y2012-145	♀	Light yellow	Yuyao, Ningbo, Zhejiang, China
39	Y2012-2	♂	Yellow-red	Yuyao, Ningbo, Zhejiang, China
40	Y2012-46	♀	Black	Yuyao, Ningbo, Zhejiang, China
41	Yangpinmei	♀	Black	Taizhou, Zhejiang, China
42	Yongxuan56	♀	Red	Ningbo, Zhejiang, China
43	Yuyaobai12	♀	White	Yuyao, Ningbo, Zhejiang, China
44	Zaoqimimei	♀	Red	Yuyao, Ningbo, Zhejiang, China
45	Zaoshenganmei	♀	Black	Anhai, Fujian, China

[a] Fruit color for cultivar and flower color for male. ♀: female plant; ♂: male plant

2.2 SSR mining and primer design

The MISA program (MIcroSAtellite identification tool; http://pgrc.ipk-gatersleben.de/misa/misa.html) was employed for microsatellite mining. In this study, the SSRs were considered to contain motifs with two nucleotides and a minimum of eight contiguous repeat units. Based on MISA results, sequences already present in the public database were removed (Jiao *et al.*, 2012). Primer3 v2.23 (http://primer3.sourceforge.net) was used to design the primer pairs. The expected polymerase chain reaction (PCR) product size of the markers was set from 120 to 300 bp and the primer size from 18 to 22 bp, with the annealing temperature set at 60 °C. The M13 tail (sequence: TGTAAAACGACGGCCAGT) was added to the 5' end of each forward primer pair, with an annealing temperature of 53 °C. The tail was alternatively labeled with the following four dyes: NED (yellow), PET (red), FAM (blue), and HEX (green). Primers were synthesized by Invitrogen Trading Co., Ltd., Shanghai, China. Two cultivars ('Biqi' and 'Dongkui') and *M. cerifera* were selected to check the polymorphism of the 400 SSR markers, and those heterozygous markers either in 'Biqi' and/or 'Dongkui' were analyzed for polymorphism, using the 45 accessions as described above.

2.3 PCR reaction

PCR was carried out as described previously by Jiao *et al*. (2012). The 20 μl reaction mixtures contained 10× PCR buffer (plus Mg^{2+}), 0.2 mmol/L of each deoxyribonucleoside triphosphate (dNTP), 5 pmol of each reverse, 4 pmol of tail primer, 1 pmol of the forward primer, 0.5 U of recombinant Taq (rTaq) polymerase (TaKaRa, Dalian, China), and a 20-ng genomic DNA template. DNA amplification was in an Eppendorf Mastercycler (Germany) programmed at 94 °C for 5 min for initial denaturation, 32 cycles at 94 °C (30 s)/58 °C (30 s)/72 °C (30 s), and then 8 cycles of 94 °C (30 s)/53 °C (30 s)/72 °C (30 s), followed by a final extension at 72 °C for 10 min. Each PCR product was run on 0.01 g/ml agarose gel to check the quality. PCR products with four different colors were diluted, mixed with an internal size standard LIZ500 (Applied Biosystems), loaded on an ABI 3130 genetic analyzer and analyzed using GeneMapper Version 4.0 software (Applied Biosystems, Foster City, CA, USA).

2.4 Data analysis

The number of alleles (N_a), number of effective alleles (N_e), observed heterozygosity (H_o), and expected heterozygosity (H_e) were calculated using GenAlEx 6.4 (Peakall and Smouse, 2006). The polymorphic information content (PIC) and Chi-square test for Hardy-Weinberg equilibrium were calculated using PowerMarker Version 3.25 (Liu and Muse, 2005). Cluster analysis was conducted using the neighbor-joining algorithm as implemented in TREECON ver.1.3 b (van de Peer and de Wachter, 1994). Bootstraps were performed for 1000 iterations to test the reliability of the branches (Felsenstein, 1985).

2.5 Fruit characteristic analysis

The width and length of each fruit were measured using a digital vernier caliper. Fruit and the fruit core were weighed and the edible ratio estimated. A texture analyzer (TA-XT2i Plus, Stable Micro System Ltd., Surrey, UK) fitted with a 5.0 mm-diameter head was used for fruit firmness analysis. Total soluble solids (TSSs) and titratable acids (TAs) were measured as described by Zhang *et al.* (2005). TSSs of thirty fruits (two measurements per fruit) were determined with a digital refractometer (PR-101, Atago, Japan). For TA analysis, 1 g of fruit mesocarp tissue derived from a segment of the flesh was ground with 5 ml of distilled water. After filtration and centrifugation for 5 min at 12000×*g*, the supernatant was brought to 10 ml with distilled water. After heating for 5 min at 100 °C to eliminate CO_2, the water was titrated with fresh 10 mmol/L NaOH to pH 8.2. The results were expressed in mmol H^+/g fresh weight (FW). The samples for TA analysis were in triplicate.

3 Results

3.1 Character of SSR markers

The effectiveness of 400 primer pairs was analyzed in two cultivars ('Biqi' and 'Dongkui') and *M. cerifera*, and 354 (88.5%) SSRs were amplified in 'Biqi' and 'Dongkui' with expected sizes of PCR products, and 296 (74%) in *M. cerifera*. One hundred and seven markers identified as heterozygous either in 'Biqi' or 'Dongkui' were used for genotyping the set of 45 accessions. These sequence data have been deposited in the GenBank Data Library under the

accession numbers from KF914760 to KF914866, which are accessible. The 107 SSR forward and reverse primers and characteristics in 45 bayberry accessions are listed in Table S1. The PCR product size ranged from 115 to 330 bp, and the detected N_a from 3 to 17, with an average of 8 alleles. N_e ranged from 1.22 to 10.41, with an average of 4.08. The PIC at each locus ranged from 0.13 to 0.89 with an average of 0.63. H_o was between 0.04 (ZJU190) and 0.96 (ZJU258), while H_e was between 0.18 (ZJU215) and 0.90 (ZJU181). There was a significant deviation from the Hardy-Weinberg equilibrium (P<0.05) of 92% of the SSR primers (98 primer pairs). Partial correlation analysis showed highly positive correlations between the repeat unit length and PIC (P<0.01, r=0.3301). This study showed that these 107 SSR markers had high rates of transferability (96.2% and 88.8%, respectively) for two related species, *M. adenophora* and *M. cerifera*, suggesting that these markers could be used for genetic diversity analyses in other *Myrica* species (Table S1). For constructing the bayberry molecular marker linkage maps, we selected a total of 78 primer pairs showing heterozygous loci either in 'Biqi' or 'Dongkui', two parents of the F_1 population.

3.2 Genetic relationship analysis

A total of 107 SSR markers were used to evaluate genetic diversity and relatedness among the 45 accessions. The SSR markers clearly discriminated between *M. rubra*, *M. adenophora*, and *M. cerifera*, separating them into distinct clusters (Fig. 1). According to neighbor-joining cluster analysis, the *M. rubra* was divided into five groups. These were labeled as 'A-1', 'A-2', 'A-3', 'A-4', and 'A-5'.

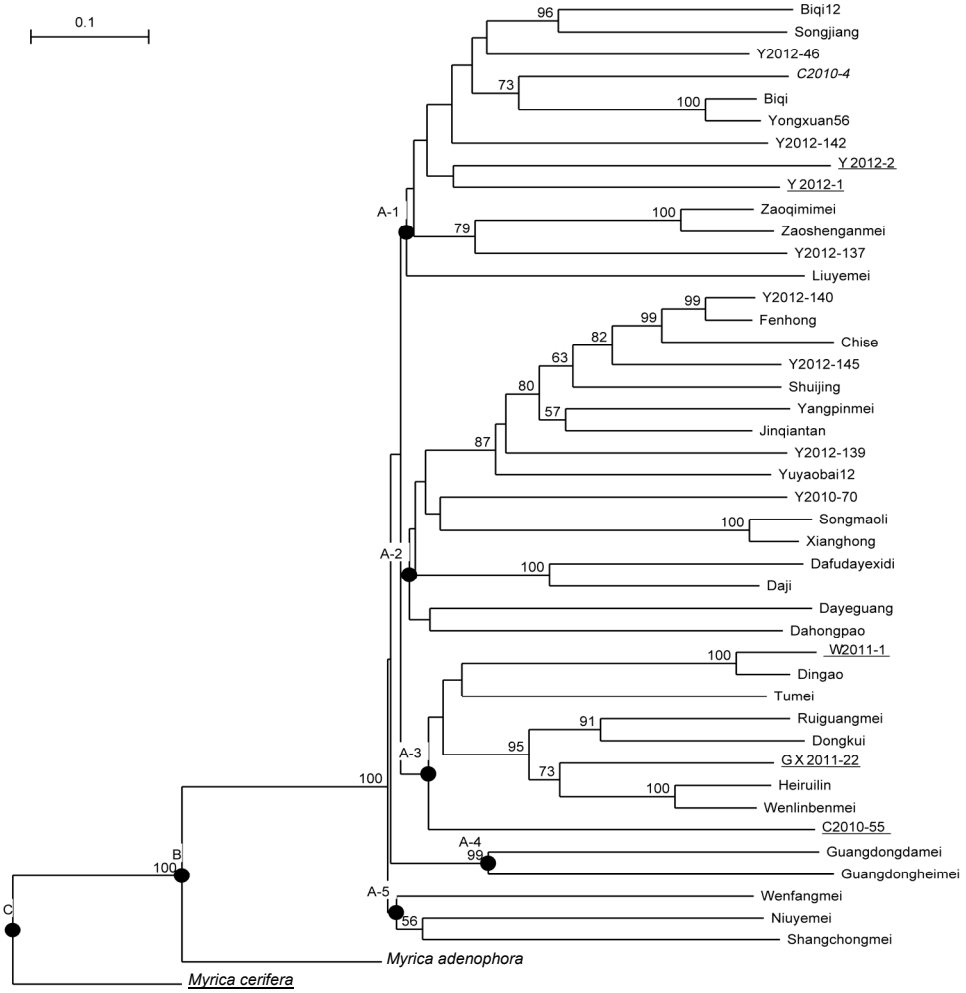

Fig. 1 Dendrogram for 45 Chinese bayberry accessions derived from neighbor-joining cluster analysis based on 107 highly polymorphic SSR markers

Accessions: underlined, male (androphyte) plant; in italic, *C2010-4*, monoecious plant in normal, common female plants. The numbers are bootstrap values based on 1000 iterations. Only bootstrap values greater than 50 are indicated

The first cluster, 'A-1', consisted of 12 accessions from Zhejiang Province and one from Fujian Province. The two androphytes 'Y2012-1' and 'Y2012-2' clustered into this group were close to cultivars of the same geographic origin. As reported previously by Jiao *et al.* (2012), the rare monoecious individual 'C2010-4' was closely related to a main local cultivar 'Biqi'. 'Zaoshenganmei' from Fujian Province was closely related to the Zhejiang accession 'Zaoqimimei'. Sixteen accessions were in the subgroup 'A-2', 'Daji' and 'Dafudayexidi', two from Jiangsu Province, and 14 from Zhejiang Province. The pink fruit accession 'Y2012-140' was close to 'Fenhong', while 'Shuijing' and 'Y2012-145' (both light yellow or white fruit type) were clearly separated in the cluster, with a short genetic distance. The elite accession 'Y2010-70', with red fruit, was also clustered in this subgroup. The accessions 'Songmaoli' and 'Xianghong' (both from Hangzhou) were close together. The two accessions 'Dafudayexidi' and 'Daji', from Jiangsu Province, were clustered together, demonstrating a relationship between genetic background and geographic origin. Subgroup 'A-3' included nine accessions, from 'W2011-1' to 'C2010-55'. Androphyte 'W2011-1' was close to a cultivar from the same geographic origin. It is worth noting that accessions from Taiwan of China ('Heiruilin') and Japan ('Ruiguangmei') were clustered with those cultivars from Zhejiang Province (China), as has been described previously (Zhang *et al.*, 2009b; Xie *et al.*, 2011). Two accessions originating from Guangzhou were included in subgroup 'A-4', implying these accessions may have independent genetic backgrounds. Subgroup 'A-5' included three accessions, two from Hunan Province and one from Jiangxi Province, indicating that these accessions may share a common ancestor.

3.3 Cultivar identification

SSR markers are a powerful tool for identification of Chinese bayberry cultivars. The black fruit cultivar 'Biqi', originating from Zhejiang Province, has been introduced to other provinces 20 years ago, although its yield and commercial quality are constantly declined. In addition, the light yellow (white) and pink fruit types, such as 'Shuijing' and 'Fenhong', have demanded the highest price of Chinese bayberry in recent years, due to their appealing color and delicious taste. The elite accessions 'Y2010-70', 'Y2012-140', and 'Y2012-145' with black-red, pink, and light yellow fruit, respectively, have been studied for more than five years, and introduced to other cities of Zhejiang Province due to their good fruit quality (Table 2, Fig. 2).

Based on our data, we can demonstrate that 'Y2010-70', 'Y2012-140', and 'Y2012-145' are genetically different lines of Chinese bayberry. The red fruit of line 'Y2010-70' has high acidity and the strong typical Chinese bayberry aroma, making it suitable for processing.

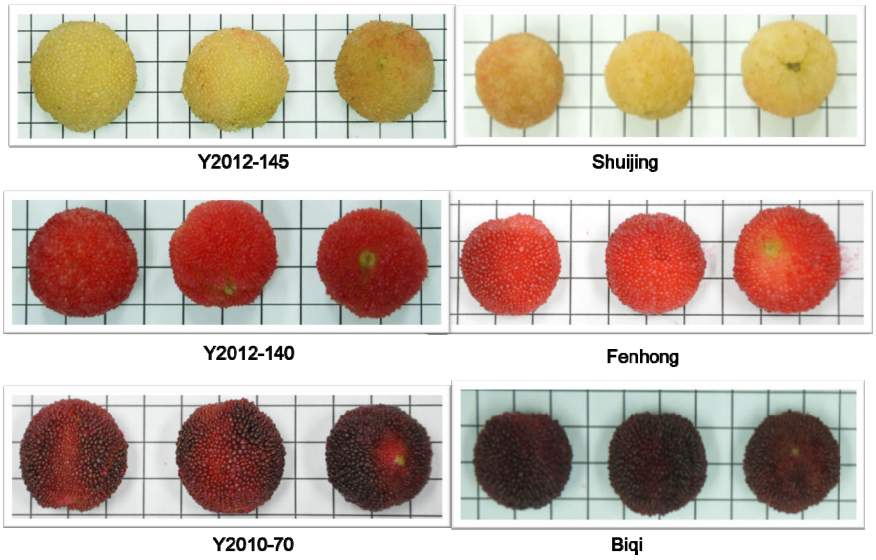

Fig. 2 Fruit size and color of three selections and three reference cultivars
Each grid cell represents one square centimeter

Table 2 Main fruit characters of the identified three new accessions and three local cultivars

Variety	Diameter (mm)	Weight (g)	Firmness (N)	TSS (°Brix)	TA (mmol H^+/g FW)
Y2012-145	31.63±0.47A	16.97±0.82A	2.73±0.05AB	11.04±0.09AB	0.24±0.00C
Shuijing	26.33±0.70CD	9.33±0.50D	2.34±0.22B	11.39±0.01A	0.23±0.01CD
Y2012-140	29.48±0.23AB	13.52±0.14B	2.36±0.12B	10.12±0.04B	0.21±0.01CD
Fenhong	25.82±0.15D	8.95±0.09D	3.19±0.06A	8.63±0.05C	0.28±0.01B
Y2010-70	28.59±0.23BC	12.61±0.06BC	2.56±0.07AB	11.03±0.07AB	0.36±0.01A
Biqi	26.84±0.81BC	10.72±0.08CD	2.34±0.09B	9.16±0.13C	0.19±0.01D

TSS: total soluble solids; TA: titratable acid; FW: fresh weight. Data are expressed as mean±SD (n=3). Values followed by different letters (A–D) in the same column are significantly different at P<0.01

4 Discussion

In our study, we aimed to develop additional SSR markers for *M. rubra* and used 45 accessions to evaluate the character of the markers. These markers can be generally used in genetic diversity analysis of cultivars, androphyte and related species. These markers also can be used in construction of SSR-based linkage map in *Myrica.*

4.1 SSR polymorphism

SSRs from genome and transcriptome sequences have been widely used for genetic diversity analysis and map construction (Emanuelli *et al.*, 2013; Frascaroli *et al.*, 2013; Würschum *et al.*, 2013). In our study, the mean value of PIC (0.63) was less than that (0.67) previously reported for genomic-SSRs (Jiao *et al.*, 2012), but significantly higher than the figure (0.44) for EST-SSRs (Zhang *et al.*, 2012). Our results are in agreement with that reported in lettuce (Rauscher and Simko, 2013). Compared with EST-SSRs, genomic-SSRs have relatively high polymorphism and good genome coverage, so they can be used in constructing a linkage map. Among 107 polymorphic SSRs, 98 were AG/CT repeat type, the same main dinucleotide repeat motifs of Chinese bayberry in the genome and transcriptome (Zhang *et al.*, 2012). The SSRs developed from *M. rubra* could also be used in related species *M. cerifera* and *M. adenophora*, showing a high transferability. *M. cerifera* is from the USA, *M. adenophora* is grown as wild trees and the fruit is used for medical purposes. The bayberry genome has a high level of heterozygosity (Jiao *et al.*, 2012), also reflected by over 70% of the SSR loci being heterozygous for most accessions. The most homozygous (58/107 SSRs) accession was found to be 'Y2012-145'. This selection may be used as material for the whole genome sequencing project of Chinese bayberry.

4.2 Genetic relationship

The lowest similarity coefficient was observed between the species *M. cerifera* and *M. rubra*, as reported previously using amplified fragment length polymorphisms (AFLPs) (Zhang *et al.*, 2009b) and SSRs (Xie *et al.*, 2011; Jiao *et al.*, 2012). Accessions from the same geographical region were clustered more closely than those from different regions. Moreover, some accessions, originating from different provinces in China, clustered within the same subgroups, suggesting the gene flow among those provinces (Zhang *et al.*, 2009b; Xie *et al.*, 2011; Jiao *et al.*, 2012).

The Taiwan cultivar 'Heiruilin' clustered into subgroup A-3 and was most closely related to 'Wenlinbenmei' from Zhejiang Province as reported previously (Zhang *et al.*, 2009b; Xie *et al.*, 2011). This indicates that these cultivars may share a common ancestor, although this remains to be further clarified. 'Ruiguangmei' from Japan has a close relationship with the Zhejiang Province cultivar 'Dongkui'. Our data, along with previous analysis (Zhang *et al.*, 2009b; Xie *et al.*, 2011), support the hypothesis that the Japanese cultivar 'Ruiguangmei' (Zuiko) was introduced from China (Handa and Kajiura, 1991).

The two famous cultivars 'Biqi' and 'Dongkui', widely cultivated in most regions, separated into different subgroups ('A-1' and 'A-3'), consistent with the previous studies using SSR markers (Xie *et al.*, 2011; Jiao *et al.*, 2012), while they have been clustered together in previous studies using AFLP (Zhang *et al.*, 2009b). This inconsistency might be due to the different molecular markers used.

Androphyte accessions of *M. rubra* were distributed in two different subgroups (A-1 and A-3), which were closely related to the gynophyte accessions from the same geographic origin. This result suggests that the androphyte and gynophyte share a common ancestry.

4.3 Cultivar identification

Molecular technology is a reliable method to identify cultivars of Chinese bayberry. 'Yongxuan56', with bigger fruit and better fruit quality, has been shown to be a new line of 'Biqi' based on AFLP analysis (Zheng *et al.*, 2006). In the new Chinese bayberry lines and cultivars, the fruit color is mainly black and red, such as 'Zaoqimimei' (Qi *et al*., 2003) and 'Zijing' (Huang *et al.*, 2013). However, no new cultivar with light yellow (white) or pink fruit color has been reported. 'Y2010-70', with high quality and an early-mature date, and 'Y2012-140', with pink fruit and good flavor, are two elite selections within the red-colored bayberry group, while 'Y2012-145', with larger fruit than 'Shuijing', is derived from the light yellow (or white) color group. Based on our SSR data and the characteristics of the fruit, the elite accessions 'Y2010-70', 'Y2012-140', and 'Y2012-145' can be demonstrated as new varieties of Chinese bayberry.

4.4 Cross-breeding perspective in *Myrica*

Cross-breeding is a traditional breeding method, and use of a hybrid (F_1) to construct a genetic linkage map is essential for marker-assisted breeding and selection (MAS). Identification of viable pollens from authorized cultivars makes cross-breeding possible. Studies have shown that pollen characteristics of different bayberry plants grown in various regions are quite dissimilar. As the 'Dongkui' pollen has a high fruit setting rate, with an average of 21.47% (Chai *et al*., 2012; Jiao *et al.*, 2013), we carried out a cross-breeding program between 'Biqi'×'Dongkui' cultivars using occasional pollens from the 'Dongkui' cultivar (usually only female followers). Polymorphic SSRs that are heterozygosis in either of the parents can be used in linkage mapping of F_1 populations. In previous studies, the numbers of SSRs from whole genome shotgun sequences and EST heterozygosis in 'Biqi' and 'Dongkui' were given as 135 and 80, respectively (Jiao *et al.*, 2012; Zhang *et al.*, 2012). In this study, we developed 78 SSR markers, which can be applied in linkage mapping of the F_1 populations resulting from the cross-breeding between 'Biqi' and 'Dongkui'. Therefore, around 300 SSR markers are available for the linkage mapping work till now.

5 Conclusions

Chinese bayberry genome sequences are a good resource for developing SSR markers. In our study, 107 SSR markers were successfully used to evaluate the genetic diversity and identify cultivars, providing a valuable genetic and genomic tool for further genetic research and varietal development in the Chinese bayberry, such as diversity studies, construction of a genetic map, and molecular breeding.

Acknowledgements

We acknowledge Dr. Qi-kang GAO from the Bio-Macromolecules Analysis Lab, Analysis Center of Agrobiology and Environmental Sciences of Zhejiang University, China for his assistance in the lab work.

Compliance with ethics guidelines

Guo-yun WANG and Zhong-shan GAO received research grants from the Ningbo Science and Technology Bureau (No. 2012C10012) and the Yuyao Financial Bureau for Chinese Bayberry Breeding Program; Sen-miao LIANG received research grant from the Science and Technology Project of Zhejiang Province (No. 2012C12904). Hui-min JIA, Yu-tong SHEN, Yun JIAO, Xiao DONG, Hui-juan JIA, Fang DU, Chao-chao ZHOU, and Wei-hua MAO declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Bao, J.S., Cai, Y.Z., Sun, M., *et al*., 2005. Anthocyanins, flavonols, and free radical scavenging activity of Chinese bayberry (*Myrica rubra*) extracts and their color properties and stability. *J. Agric. Food Chem*., **53**(6): 2327-2332. [doi:10.1021/jf048312z]

Celton, J.M., Tustin, D., Chagné, D., *et al*., 2009. Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. *Tree Genet. Genomes*, **5**(1):93-107. [doi:10.1007/s11295-008-0171-z]

Chai, C.Y., Xu, S.Q., Zhou, H.F., *et al*., 2012. Studies on pollination and metaxenia of *Myrica rubra*. *J. Fujian For. Sci. Tech*., **39**(4):30-33 (in Chinese). [doi:10.3969/j.issn.1002-7351.2012.04.09]

Chen, K.S., Xu, C.J., Zhang, B., *et al.*, 2004. Red bayberry: botany and horticulture. *In*: Janick, J. (Ed.), Horticultural Reviews, Volume 30. John Wiley & Sons, Inc., p.83-114. [doi:10.1002/9780470650837]

Emanuelli, F., Lorenzi, S., Grzeskowiak, L., *et al*., 2013. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biol*., **13**:39. [doi:10.1186/1471-2229-13-39]

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**(4):783-791.

Frascaroli, E., Schrag, T.A., Melchinger, A.E., 2013. Genetic diversity analysis of elite european maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. *Theor. Appl. Genet*., **126**(1):133-141. [doi:10.1007/s00122-012-1968-6]

Handa, T., Kajiura, I., 1991. Isozyme analysis of yamamomo (*Myrica rubra* Sieb. et Zucc.) cultivars. *Jpn. J. Breed*., **41**(2):203-209. [doi:10.1270/jsbbs1951.41.203]

Huang, Y.H., Yu, W.S., Guo, Z.H., *et al.*, 2013. A new Chinese bayberry cultivar 'Zijing'. *Acta Hort. Sin.*, **4**(4): 791-792 (in Chinese).

Jiao, Y., Jia, H.M., Li, X.W., *et al.*, 2012. Development of simple sequence repeat (SSR) markers from a genome survey of Chinese bayberry (*Myrica rubra*). *BMC Genomics*, **13**:201. [doi:10.1186/1471-2164-13-201]

Jiao, Y., Wang, G.Y., Chai, C.Y., *et al.*, 2013. Morphology of pollen grains from the plant with different type of sex by scanning electron microscope (SEM) and the viablity of pollens in red bayberry. *South China Fruits*, **42**(1):12-15 (in Chinese).

Liu, K., Muse, S.V., 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, **21**(9):2128-2129. [doi:10.1093/bioinformatics/bti282]

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol. Ecol. Notes*, **6**(1):288-295. [doi:10.1111/j.1471-8286.2005.01155.x]

Qi, X.J., Liang, S.M., Zheng, X.L., *et al.*, 2003. A new Chinese bayberry cultivar 'Zaoqimimei'. *Acta Hort. Sin.*, **30**(6): 759 (in Chinese).

Rauscher, G., Simko, I., 2013. Development of genomic SSR markers for fingerprinting lettuce (*Lactuca sativa* L.) cultivars and mapping genes. *BMC Plant Biol.*, **13**(1):11. [doi:10.1186/1471-2229-13-11]

Terakawa, M., Kikuchi, S., Kanetani, S., *et al.*, 2006. Characterization of 13 polymorphic microsatellite loci for an evergreen tree, *Myrica rubra*. *Mol. Ecol. Notes*, **6**(3): 709-711. [doi:10.1111/j.1471-8286.2006.01318.x]

Testolin, R., Marrazzo, T., Cipriani, G., *et al.*, 2000. Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome*, **43**(3):512-520. [doi:10.1139/g00-010]

van de Peer, Y., de Wachter, R., 1994. TREECON for windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.*, **10**(5):569-570. [doi:10.1093/bioinformatics/10.5.569]

Wang, H.Y., Gao, Z.S., Yang, Z.W., *et al.*, 2012. Anaphylaxis and generalized urticaria in a woman eating Chinese bayberry fruit. *J. Zhejiang Univ.-Sci. B (Biomed. & Biotechnol.)*, **13**(10):851-854. [doi:10.1631/jzus.B1200150]

Würschum, T., Langer, S.M., Longin, C.F.H., *et al.*, 2013. Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.*, **126**(6):1477-1486. [doi:10.1007/s00122-013-2065-1]

Xie, R.J., Zhou, J., Wang, G.Y., *et al.*, 2011. Cultivar identification and genetic diversity of Chinese bayberry (*Myrica rubra*) accessions based on fluorescent SSR markers. *Plant Mol. Biol. Rep.*, **29**(3):554-562. [doi:10.1007/s11105-010-0261-6]

Yue, X.Y., Liu, G.Q., Zong, Y., *et al.*, 2014. Development of genic SSR markers from transcriptome sequencing of pear buds. *J. Zhejiang Univ.-Sci. B (Biomed. & Biotechnol.)*, **15**(4):303-312. [doi:10.1631/jzus.B1300240]

Zhang, B., Kang, M.X., Xie, Q.P., *et al.*, 2011. Anthocyanins from Chinese bayberry extract protect β cells from oxidative stress-mediated injury via HO-1 upregulation. *J. Agric. Food Chem.*, **59**(2):537-545. [doi:10.1021/jf1035405]

Zhang, S.M., Xu, C.J., Gao, Z.S., *et al.*, 2009a. Development and characterization of microsatellite markers for Chinese bayberry (*Myrica rubra* Sieb. & Zucc.). *Conserv. Genet.*, **10**(5):1605-1607. [doi:10.1007/s10592-008-9804-x]

Zhang, S.M., Gao, Z.S., Xu, C.J., *et al.*, 2009b. Genetic diversity of Chinese bayberry (*Myrica rubra* Sieb. et Zucc.) accessions revealed by amplified fragment length polymorphism. *Hortscience*, **44**(2):487-491.

Zhang, S.Y., Li, X., Feng, C., *et al.*, 2012. Development and characterization of 109 polymorphic EST-SSRs derived from the Chinese bayberry (*Myrica rubra*, Myricaceae) transcriptome. *Am. J. Bot.*, **99**(12):e501-e507. [doi:10.3732/ajb.1200156]

Zhang, W.S., Chen, K.S., Zhang, B., *et al.*, 2005. Postharvest responses of Chinese bayberry fruit. *Postharvest Biol. Technol.*, **37**(3):241-251. [doi:10.1016/j.postharvbio.2005.05.005]

Zheng, J.T., Zhang, W.S., Bao, L., *et al.*, 2006. AFLP identification and characterization of Yongxuan 56, a new line of Biqi Chinese bayberry. *J. Fruit Sci.*, **23**(3):397-400 (in Chinese).

List of electronic supplementary materials

Table S1 Characteristics of the 107 polymorphic SSR markers validated in 45 bayberry accessions

中文概要：

本文题目：基于杨梅基因组鸟枪法测序开发的 107 个 SSR 标记及其在品系鉴定中应用

Development of 107 SSR markers from whole genome shotgun sequences of Chinese bayberry (*Myrica rubra*) and their application in seedling identification

研究目的：开发多态性高的基因组简单重复序列（SSR）标记，为杨梅遗传多样性、遗传图谱构建及品种鉴定相关研究奠定基础。

创新要点：从杨梅全基因组鸟枪法测序数据中新开发了 107 个 SSR 标记，并应用于品种间遗传多样性分析及品系鉴定。

研究方法：应用生物信息学软件从基因组测序数据中搜索 SSR 位点并设计相应引物，添加通用的 M13 尾巴和荧光标识筛选多态性高的 SSR 标记。

重要结论：开发的 107 个基因组 SSR 标记可以广泛应用于杨梅种质资源和遗传育种研究，鉴定出三个杨梅新品系。

关键词组：杨梅；简单序列重复（SSR）标记；遗传多样性