GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Genetic diversity of *Lagerstroemia* (Lythraceae) species assessed by simple sequence repeat markers

D. He[1], Y. Liu[1], M. Cai[1], H.T. Pan[1,2], Q.X. Zhang[1,2], X.Y. Wang[1] and X.J. Wang[1]

[1]Department of Ornamental Horticulture, College of Landscape Architecture, Beijing Forestry University, Beijing, China
[2]China National Engineering Research Center for Floriculture, Beijing Forestry University, Beijing, China

Corresponding authors: H.T. Pan / Q.X. Zhang
E-mail: htpan2000@yahoo.com.cn / zqx@bjfu.edu.cn

Genet. Mol. Res. 11 (3): 3522-3533 (2012)
Received December 20, 2011
Accepted March 15, 2012
Published September 26, 2012
DOI http://dx.doi.org/10.4238/2012.September.26.9

ABSTRACT. *Lagerstroemia* (crape myrtle) are famous ornamental plants with large pyramidal racemes, long flower duration, and diverse colors. However, little is known about the genetic structure and diversity of germplasm in *Lagerstroemia*. We genotyped 81 *L. indica* cultivars, five other species of *Lagerstroemia*, and 10 interspecific hybrids using 30 simple sequence repeat markers; 275 alleles were generated with a mean of nine alleles per locus. The mean polymorphism information content value, a measure of gene diversity, was 0.63, with a range from 0.25 to 0.86. The mean observed heterozygosity (0.51) tended to be lower than the mean expected heterozygosity (0.67). The mean F-statistics (F_{ST}, F_{IS}, and F_{IT}) were 0.05, 0.20, and 0.24, respectively, indicating a high level of genetic variation among cultivars. Clustering analysis based on genetic distance divided the 96 genotypes into three distinct groups, which corresponded with their genetic backgrounds and geographic regions. *L. indica* cultivars and the other five *L.* species were grouped into different sub-clusters. Chinese and North American cultivars were divided into different clusters. These data about the genetic relationship among cultivars demonstrated the potential value of *L. indica* cultivars and

other *Lagerstroemia* species for widening the genetic basis of breeding programs for this ornamental flower.

Key words: *Lagerstroemia*; SSR; Genetic diversity; Cluster analysis

INTRODUCTION

Lagerstroemia is a member of the family Lythraceae and comprises more than 50 species that are native to southeastern Asia and Australia (Brickell, 1996). *Lagerstroemia indica*, native to Southeast Asia, is a medium to large shrub with brown bark, glabrous trunk, and deciduous leaves that change from green to yellow, orange, or red in the fall (Wang et al., 2010). In China, *L. indica* has been widely cultivated in gardens for about 1800 years (Zhang, 1991).

Many new cultivars with colorful flowers, different growth habits, and disease resistance had been cultivated since the 1960s (Egolf, 1967, 1981, 1987, 1990; Pooler, 2006). *L. fauriei*, *L. speciosa*, *L. limii*, and *L. subcostata* have been introduced in crape myrtle breeding programs (Pooler, 2003; Pounders, 2007). *L. fauriei*, from central and southern Japan, is a very important species, partly for its special landscape performance and strong resistance to mildew diseases. In fact, a series of crape myrtle cultivars with colorful flowers, beautiful fall leaves and high mildew resistance were found to be hybrids of *L. fauriei* (Egolf and Andrick, 1978; Egolf, 1986). *L. speciosa*, native to South East Asia, India and the Philippines, has large flowers and is widely cultivated as an ornamental tree in tropical and subtropical areas. *L. limii* and *L. subcostata* are usually trees of about 4 to 14 m high and have smaller flowers. *L. caudata* has smaller, fragrant white flowers, blooms earlier than *L. indica*, and has been used to breed early-flowering and fragrant crape myrtle hybrid cultivars since 2008 (Cai et al., 2010).

Compared to traditional morphological traits, molecular markers are not influenced by the environment and reflect the real genetic diversity. They do not require previous pedigree information (Bohn et al., 1999). There are various types of molecular markers for genome analysis, among which simple sequence repeats (SSRs) are abundant, co-dominant, multi-allelic, highly polymorphic, and chromosome-specific. SSRs have been extensively used in genetic diversity studies in many plants, including *Citrus* (Barkley et al., 2006), *Theobroma cacao* (Sereno et al., 2006), *Malus sieversii* (Zhang et al., 2007), *Trifolium pretense* (Dias et al., 2008), and *Jatropha curcas* (Subramanyam et al., 2009).

The first investigation of the genetic diversity of *Lagerstroemia* was on 12 *L. fauriei* clones using AFLP and RAPD markers by Pooler (2003). Recently, SSRs have been used to estimate genetic diversity among *Lagerstroemia* species and cultivars (Rinehart and Pounders, 2010; Wang et al., 2011; Cai et al., 2011). However, there has been no report on genetic diversity of crape myrtle cultivars from different geographic regions. The objectives of our study were 1) to evaluate the genetic diversity among *Lagerstroemia* species and cultivars using SSR markers and 2) to compare genetic backgrounds and evaluate the possibility of collecting cultivars from different regions for crape myrtle breeding.

MATERIAL AND METHODS

Sampling of species and cultivars

A total of 96 genotypes, including five species (*L. speciosa*, *L. subcostata*, *L. limii*, *L.*

fauriei, and *L. caudata*), 81 *L. indica* cultivars, and 10 interspecific hybrids (Table 1), were used in our study. Among these accessions, 55 cultivars were introduced from North America and four cultivars from Europe. The other 22 cultivars were all from China, and the five *Lagerstroemia* species have been cultivated in China for many years. The 10 interspecific hybrids are from the crosses of *L. caudata* x *L. indica* and *L. indica* x *L. speciosa* and were created by our research group in 2008. The list of cultivars and species, with their plant size, flower color, genetic background, and their possible geographic regions, are listed in Table 1. All of these cultivars and species were cultivated as ornamental plant germplasm at the China National Engineering Research Center for Floriculture in Beijing.

Table 1. The detailed information of *Lagerstroemia* L. used in this study.

No.	Cultivars/line name	Plant size	Genetic background	Flower color (RHS)	Region
1	Tuscarora	Tree	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 71B	NA
2	Muskogee	Tree	*L. indica* 1/2, *L. fauriei* 1/2	Red Purple 63C	NA
3	Tuskegee	Tree	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 61C	NA
4	Choctaw	Tree	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 68C	NA
5	Dallas Red	Shrub	*L. indica*	Red 53B	NA
6	Carolina Red	Shrub	*L. indica*	Red 46B	NA
7	Sarah Favorite	Tree	*L. indica* 1/2, *L. fauriei* 1/2	White NN155C	NA
8	Miami	Tree	*L. indica* 5/8, *L. fauriei* 3/8	Red 55C	NA
9	Biloxi	Tree	*L. indica* 1/2, *L. fauriei* 1/2	Red Purple 63B	NA
10	Natchez	Shrub	*L. indica* 1/2, *L. fauriei* 1/2	Red Purple 65A	NA
11	Catawba	Shrub	*L. indica*	Red Purple 72A	NA
12	Osage	Tree	*L. indica* 1/2, *L. fauriei* 1/2	Red Purple 68B	NA
13	Sioux	Tree	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 68B	NA
14	Comanche	Tree	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple N66A	NA
15	Queens Lace	Dwarf	*L. indica*	Red 53B and White NN155C	NA
16	Yuma	Tree	*L. indica* 3/8, *L. fauriei* 3/8, *L. amabilis* 1/4	Purple 76A	NA
17	Lipan	Tree	*L. indica* 5/8, *L. fauriei* 3/8	Red Purple 72B	NA
18	Centennial Spirit	Shrub	*L. indica*	Red Purple 58A	NA
19	William Toovey	Tree	*L. indica*	Red Purple N66B	NA
20	Pecos	Shrub	*L. indica* 1/2, *L. fauriei* 1/2	Red Purple 72B	NA
21	Okmulgee	Dwarf	*L. indica*	Red 47A	NA
22	Acoma	Shrub	*L. indica* 3/4, *L. fauriei* 1/4	White NN155C	NA
23	Hopi	Shrub	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 68A	NA
24	Victor	Shrub	*L. indica*	Red Purple 63A	NA
25	Near East	Shrub	*L. indica*	Red Purple N66A	EA
26	Zuni	Shrub	*L. indica* 3/4, *L. fauriei* 1/4	Purple N78A	NA
27	Centennial	Shrub	*L. indica*	Purple N78A	NA
28	Prairie Lace	Shrub	*L. indica*	Red 52A and White NN155C	NA
29	Tonto	Shrub	*L. indica* 3/4, *L. fauriei* 1/4	Red Purple 61B	NA
30	Velma's Royal Delight	Shrub	*L. indica*	Red Purple 72A	NA
31	Caddo	Shrub	*L. indica*, *L. fauriei*	Red Purple 58C	NA
32	Powhatan NEW	Shrub	*L. indica*	Purple N78D	NA
33	New Orleans	Shrub	*L. indica*	Purple N78B	NA
34	Sacramento	Dwarf	*L. indica*	Red Purple 67B	NA
35	Delta Blush	Shrub	*L. indica*	Red Purple 68A	NA
36	Pink Blush	Shrub	*L. indica*	Red Purple 63D and 68B	NA
37	Lafayette	Shrub	*L. indica*	Purple 76B	NA
38	World's Fair	Dwarf	*L. indica*	Red Purple 68A	NA
39	Houston NEW	Dwarf	*L. indica*	Red Purple 67C	NA
40	Pixie White	Dwarf	*L. indica*	White NN155C	NA
41	Mardi Gras	Shrub	*L. indica*	Purple Violet N80B	NA
42	Baton Rouge	Dwarf	*L. indica*	Red Purple 63A	NA
43	Bourbon Street	Dwarf	*L. indica*	Red Purple 61D	NA
44	Bayou Marie	Dwarf	*L. indica*	Red Purple 68B	NA
45	Creole	Dwarf	*L. indica*	Red Purple 63B	NA
46	Cordon Bleu	Dwarf	*L. indica*	Purple 76A	NA
47	Bicolor	Shrub	*L. indica*	Red Purple N66B and White NN155C	NA

Continued on next page

Table 1. Continued.

No.	Cultivars/line name	Plant size	Genetic background	Flower color (RHS)	Region
48	Purple Velvet	Shrub	*L. indica*	Purple N78A	NA
49	Chisam Fire	Dwarf	*L. indica*	Red Purple 63A	NA
50	Chickasaw	Dwarf	*L. indica* 5/8, *L. fauriei* 3/8	Red Purple 64B	NA
51	Pocomoke	Shrub	*L. indica* 5/8, *L. fauriei* 3/8	Red Purple 67B	NA
52	David Hanging Basket	Shrub	*L. indica*	Red Purple 62C and 68R	NA
53	Arapahoe	Shrub	*L. indica* 11/16, *L. fauriei* 1/16, *L. limii* 1/4	Purple N78D	NA
54	Apalachee	Shrub	*L. indica* 1/2, *L. fauriei* 1/2	Purple 75A	NA
55	Dynamite	Tree	*L. indica*	Red 53B	NA
56	Berlingot Menthe	Shrub	*L. indica*	Red 53B and White NN155C	NA
57	Niver	Dwarf	*L. indica*	Violet 85D	EU
58	Petite Pink	Dwarf	*L. indica*	Red Purple N57C	EU
59	Rosea Nova	Dwarf	*L. indica*	Red Purple 68A	EU
60	Violet	Shrub	*L. indica*	Purple N78B	EU
61	Hong Die Fei Wu	Shrub	*L. indica*	Red Purple 67A	EA
62	Duo Hua Zi	Shrub	*L. indica*	Purple 77B	EA
63	Ceng Yun Ji Xue	Shrub	*L. indica*	White NN155C	EA
64	Duo Hua Fen	Shrub	*L. indica*	Red Purple 67C	EA
65	Huan Fen Liu Yun	Tree	*L. indica*	Red Purple 62D and White NN155C	EA
66	Bing Qing Yu Die	Tree	*L. indica*	White NN155C	EA
67	Fu Jian Za Zhong Bai	Shrub	*L. indica*, *L. limii*	White NN155C	EA
68	Zi Sha	Tree	*L. indica*	Purple 76B	EA
69	Fen Jing	Shrub	*L. indica*	Red Purple 63B	EA
70	Zi Jin	Tree	*L. indica*	Purple 77C	EA
71	Hong Die Fei Wu II	Shrub	*L. indica*	Red Purple 61 B	EA
72	Hong Ri Ying Xue	Shrub	*L. indica*	White NN155C	EA
73	Bai Yun Ying Xia	Tree	*L. indica*	White NN155C	EA
74	Liu Yue Fei Xue	Shrub	*L. indica*	Red 53B and White NN155C	EA
75	Xiao Hua Zi	Tree	*L. indica*, *L. subcostata*	Red Purple N66D	EA
76	Zhi Zhi Fen	Shrub	*L. indica*	Red Purple 67C	EA
77	BZ	Shrub	unknown	Unknown	EA
78	Zhi Zhi Zi	Shrub	*L. indica*	Violet N81C	EA
79	Qiao Jia Ren	Shrub	*L. indica*	Red Purple N66C and White NN155C	EA
80	XTSC3	Shrub	*L. indica*	Red Purple N74B	EA
81	BH	Tree	*L. indica*	Red Purple N57D	EA
82	FD	Tree	*L. indica* 1/2, *L. speciosa* 1/2	Purple Violet N80C	EA
83	DD	Tree	*L. indica* 1/2, *L. speciosa* 1/2	Unknown	EA
84	*L. speciosa*	Tree	*L. speciosa*	Purple 78B	EA
85	*L. subcostata*	Tree	*L. subcostata*	White NN155C	EA
86	*L. limii*	Tree	*L. limii*	Unknown	EA
87	*L. fauriei*	Tree	*L. fauriei*	Red Purple 58B	EA
88	*L. caudata*	Tree	*L. caudata*	White NN155C	EA
89	WH1	Tree	*L. indica*, *L. caudata*	White NN155C	EA
90	WH2	Tree	*L. indica* 1/2, *L. caudata* 1/2	Red Purple N74B	EA
91	WH3	Tree	*L. indica* 1/2, *L. caudata* 1/2	White NN155C	EA
92	WH4	Tree	*L. indica* 1/2, *L. caudata* 1/2	White NN155C	EA
93	WH5	Tree	*L. indica* 1/2, *L. caudata* 1/2	Purple 75A	EA
94	WH6	Tree	*L. indica* 1/2, *L. caudata* 1/2	Red Purple N74C	EA
95	WH7	Tree	*L. indica* 1/2, *L. caudata* 1/2	Red Purple 72C	EA
96	WH8	Tree	*L. indica* 1/2, *L. caudata* 1/2	White NN155C	EA

RHS = Royal Horticulture Society; NA = North American; EU = Europe; EA = East Asia. Some pedigree information refers to Wang et al. (2011).

DNA extraction

Total genomic DNA was isolated from young leaf tissues, using the FastDNA kit (TianGen) and following the manufacturer protocol. DNA concentrations were estimated with Smart Ladder (TianGen) on a 1% (w/v) agarose gel containing Gel Red at 0.1 µg/mL 1X TAE buffer, with bands visualized under UV light (Zhu et al., 2011).

SSR screening and analysis

A total of 28 SSR primers were synthesized based on published information (Wang et al., 2010, 2011; Cai et al., 2011). Primers 29 and 30 were designed by Primer Premier 5.0 (Palo Alto, Canada) from the sequence data provided by GenBank. These primers were prescreened on eight randomly selected genotypes to select optimal primers for further analysis. The details of the nucleotide sequences of SSR primers are shown in Table 2.

Table 2. Nucleotide sequences of 30 simple sequence repeat (SSR) primers used in this study.

Locus	GenBank accession No.	Primer sequence (5'-3')	Repeat	Tm (°C)	Allele size (bp)
SSR1	GQ402050	F: <FAM>AGAGAAAGAGAGGAGCGGGAGT R: ACCTTCTTCCCCAATTCAATCC	$(GT)_6GC(GT)_6$	55	213
SSR2	GQ402053	F: <FAM>ACGTTTAGCACACCGGTACTGT R: GGAAGCACATCACTATGGCAAG	$(GT)_7$	55	181
SSR3	GQ402058	F: <FAM>GGTGGAGATGCTAACAAGCAAG R: GGATTTTTGCTGTAGGGTGATT	$(TG)_{16}$	55	161
SSR4	GQ402073	F: <FAM>CCCTCATACCTTCTTTATCAAGTCA R: ATCCCCACAAAATCTCTCCTTC	$(AC)_{13}$	55	215
SSR5	GQ402085	F: <FAM>TGGGATCGATGCTATTAATGTTG R: TACACCAATTCACACCTCCACTC	$(GT)_8$	55	207
SSR6	GQ402087	F: <FAM>ACATGGCTCCCATCACACAG R: CAGGGGATTCTTGTTTTGCTTT	$(AC)_6$	55	160
SSR7	GQ424407	F: <ROX>CTCTCAAATGACCCTCTT R: TTGAGTAATAACAAGTCCC	$(AAAG)_5$	48	255
SSR8	GQ424409	F: <HEX>GAGTTCATGCAGTTAGGT R: ATATCGGATTTATCTTCC	$(AAG)_6$	48	130
SSR9	GQ424414	F: <ROX>GGAAGAGGGATTGGAACC R: TCTCACTGAAAGAAACTA	$(GA)_5G(GGA)_4$	48	134
SSR10	GQ424415	F: <HEX>ACGGTAGATAAGGTGAGC R: GGTTTCGTATCGTCGTAG	$(CT)_9T(CTGT)GT(CTGT)_4$	52	166
SSR11	GQ424416	F: <HEX>TACTGGGTATCCGTTTCT R: ACAGGTGCATTTACTTCC	$(CT)_6CC(CT)_{12}$	50	317
SSR12	GQ438235	F: <FAM>TTCTGACCCAGCAGTAAA R: CGTATCTCATCTGTAGCGTA	$(AGGT)_4$	50	138
SSR13	GQ438241	F: <HEX>GGGAATTTGGGATATGGA R: TAAAGAAACGACCGAGCC	$(AAAG)_5$	52	179
SSR14	GQ248217	F: <HEX>GTCACAGGTTACCGAATC R: ATGTAAATGGTGAGGAGG	$(AATC)_5$	50	253
SSR15	GQ248218	F: <ROX>TTCTTGTCTTGGGTATCGC R: GAGCCAGTATTGTCTTCACG	$(CCTT)_7$	50	228
SSR16	GQ248219	F: <HEX>TTCTTCCACTTCCTCCTT R: CAGCCCACATTAACTTTT	$(AG)_{12}$	50	204
SSR17	GQ248220	F: <FAM>AAAGACGCAGAAGGATGG R: CGATTAGTTTCAGCTCGT	$(AG)_{20}$	50	420
SSR18	GQ248228	F: <FAM>GGACCAGATTGTAAATGC R: CTGCTCCTAATATCAGTGTC	$(CT)_{15}$	50	289
SSR19	GQ248233	F: <HEX>TAGTCCATTCATGTCAAG R: GGATTCACCAAACTACTT	$(AG)_{14}$	52	246
SSR20	GQ424383	F: <FAM>TTTGGTGGTAGTGGGAGT R: GTGTCTGCATGGCTGTAA	$(CTGT)_6$	54	305
SSR21	GQ424384	F: <ROX>CCTAACAAGAAAGGAACAG R: TTTCAGGACATCAGCACC	$(AG)_{11}$	60	144
SSR22	GQ424386	F: <HEX>CCTCCTCCTGCCACTCCTCT R: CCCGTCGTCTCCTCAGTTCTC	$(AAG)_4ATG(AAG)_2$	54	194
SSR23	GQ424397	F: <FAM>CAACAGTAAAATTGGAGC R: AGTAGTGATTCGGGTGGA	$(CTTT)_7$	55	144
SSR24	HQ677290	F: <HEX>ATCAATGTCCACAACAACTTGCC R: GGTTTGGTCGATTTGGTTCAGTTA	$(AAG)_7$	55	158

Continued on next page

Table 2. Continued.

Locus	GenBank accession No.	Primer sequence (5'-3')	Repeat	Tm (°C)	Allele size (bp)
SSR25	HQ677242	F: <FAM>TACACTCCCTCCCATTCAGATTGT R: GCTGCCTGAATCAGTGAAGAGAGT	$(AGA)_7$	55	97
SSR26	HQ677282	F: <HEX>ACGTATCAACCGAATGACCACTTT R: GAATTCAAAGCTCAAGTGGGGAC	$(TTC)_{10}$	55	137
SSR27	HQ677293	F: <HEX>GTCTCACTCTCTCAACTCAAGGGC R: TGAGAAAGAATTTTTCCTGAACCG	$(TCT)_6$	55	137
SSR28	HQ677322	F: <HEX>ATGTACACCCGAAACCCTTTAGGT, R: TCCATGTCTTGTCACAGCCTCTAC	$(TCT)_7$	55	131
SSR29	GQ402090	F: <FAM> TGTCACTTCTGCAAATAT R: AACTACTGCCATCATACT	$(TC)_7$	50	229
SSR30	GQ402051	F: <FAM>GTGTTGGGAGTCAGATGG R: ACAGCCGTTCGACATTAA	$(AC)_6$	48	191

Tm = melting temperature.

The forward primers were synthesized with an additional 18 nucleotides from the M13 universal primer appended to the 5'-end (Schuelke, 2000). The SSR reactions were performed in a 10-μL reaction volume that contained 20 ng genomic DNA, 5 μL 2X Taq PCR Master Mix (Biomiga), and 50 ng each forward and reverse primer. The M13 universal primer was labeled either with a blue (FAM), green (HEX), or red (ROX) fluorescent tag (Sangon, Shanghai, China).

PCR began with denaturation at 94°C for 3 min, followed by 30 cycles of PCR amplification, each consisting of 30 s at 94°C for denaturation, 30 s at 48°-60°C (primer specific) for annealing, and 30 s at 72°C for extension, and a final extension of 72°C for 5 min. The PCR products (0.5 μL) with different fragment sizes and different fluorescent labels were pooled and combined with 9 μL Hi-Di formamide and 0.5 μL LIZ-500 (Applied Biosystems, USA) size standard and analyzed on an ABI3730xl DNA Analyzer (Applied Biosystems). The GeneMapper v4.0 software (Applied Biosystems) was used to determine the polymorphic information content (PIC).

Data analysis

The summary statistics reflected the genetic diversity level, including the observed number of alleles per locus (N_A), PIC, Shannon's information index (I), observed heterozygosity (H_O), expected heterozygosity (H_E), Nei's genetic distance, F-statistics (F_{ST}, F_{IS}, and F_{IT}) (Wright, 1969) and gene flow (N_m), which were calculated using POPGENE version 1.31 (Yeh et al., 1999). We carried out analysis of molecular variance (AMOVA) to partition the genetic variance between regions, and between individuals within regions, using the Arlequin version 2000 software (Schneider et al., 2000). Genetic distances between individuals were calculated based on shared allele distance to create a matrix. The cluster results of 96 genotypes were conducted by PowerMarker version 3.25 (Liu and Muse, 2005), and dendrograms were viewed with MEGA 4 (Tamura et al., 2007).

RESULTS

Allelic diversity at SSR loci

One third of 30 SSR loci were developed from *L. caudata*, and the rest were devel-

oped from *L. indica*. Of these 30 SSR loci, dinucleotide repeats and tetranucleotide repeats both represented 40%, and the remaining six SSR loci (20%) were trinucleotide repeats. The 30 SSR markers that generated 275 alleles were used to estimate the genetic diversity among 96 accessions. The N_A revealed by each marker ranged from 4 to 15, with an average of 9 per marker (Table 3). The PIC value ranged from 0.25 (SSR22) to 0.86 (SSR3), with an average of 0.63.

Table 3. Statistical analyses of 30 simple sequence repeat (SSR) loci.

Locus	N_A	PIC	I	H_O	H_E	Nei'	F_{IS}	F_{IT}	F_{ST}	N_m
SSR1	10	0.76	1.78	0.65	0.79	0.78	-0.03	0.03	0.06	3.83
SSR2	12	0.73	1.81	0.42	0.75	0.75	0.47	0.51	0.06	3.66
SSR3	15	0.86	2.29	0.73	0.87	0.87	0.01	0.06	0.05	5.15
SSR4	11	0.58	1.34	0.28	0.63	0.63	0.48	0.53	0.08	2.83
SSR5	15	0.81	2.04	0.75	0.83	0.83	0.13	0.17	0.05	4.55
SSR6	11	0.85	2.11	0.78	0.87	0.86	0.02	0.07	0.05	5.12
SSR7	8	0.45	1.06	0.32	0.47	0.46	0.33	0.37	0.07	3.31
SSR8	8	0.56	1.22	0.25	0.62	0.62	0.61	0.63	0.05	5.21
SSR9	12	0.78	1.99	0.64	0.81	0.80	0.23	0.28	0.07	3.59
SSR10	10	0.75	1.75	0.81	0.78	0.77	-0.06	0.01	0.07	3.52
SSR11	9	0.65	1.47	0.38	0.71	0.70	0.39	0.41	0.03	8.23
SSR12	4	0.36	0.76	0.45	0.39	0.38	-0.19	-0.16	0.03	8.60
SSR13	5	0.59	1.18	0.64	0.66	0.66	-0.13	-0.01	0.11	1.99
SSR14	5	0.51	1.05	0.49	0.57	0.57	0.05	0.11	0.06	4.00
SSR15	8	0.63	1.44	0.71	0.66	0.66	0.03	0.05	0.03	8.58
SSR16	11	0.63	1.54	0.19	0.66	0.66	0.63	0.64	0.02	10.29
SSR17	15	0.84	2.14	0.81	0.86	0.86	-0.13	-0.04	0.08	2.78
SSR18	8	0.72	1.60	0.45	0.76	0.75	0.36	0.37	0.02	10.37
SSR19	14	0.79	2.05	0.32	0.81	0.81	0.59	0.61	0.03	7.60
SSR20	4	0.46	0.91	0.18	0.55	0.55	0.76	0.76	0.03	8.55
SSR21	7	0.26	0.64	0.15	0.27	0.27	0.35	0.43	0.12	1.76
SSR22	6	0.25	0.59	0.26	0.26	0.26	-0.05	0.02	0.07	3.53
SSR23	6	0.43	0.95	0.32	0.46	0.46	0.23	0.33	0.13	1.61
SSR24	7	0.61	1.34	0.69	0.64	0.64	0.01	0.05	0.04	5.99
SSR25	10	0.71	1.66	0.73	0.74	0.73	-0.01	0.02	0.03	9.04
SSR26	10	0.81	1.97	0.65	0.83	0.83	0.13	0.15	0.02	10.98
SSR27	6	0.60	1.28	0.41	0.64	0.64	0.46	0.50	0.07	3.27
SSR28	15	0.73	1.75	0.39	0.76	0.76	0.42	0.45	0.05	5.18
SSR29	8	0.74	1.71	0.67	0.78	0.77	0.18	0.19	0.02	14.98
SSR30	5	0.59	1.20	0.75	0.64	0.64	-0.22	-0.13	0.07	3.41
Mean	9	0.63	1.49	0.51	0.67	0.67	0.20	0.24	0.05	4.50

N_A = observed number of alleles; PIC = polymorphic information content; I = Shannon's information index; H_O = observed heterozygosity; H_E = expected heterozygosity; N_m = gene flow.

The I computed in our experiment was 1.49 and ranged from 0.59 (SSR22) to 2.29 (SSR3). Nei's genetic distance for the SSR loci ranged from 0.26 (SSR22) to 0.87 (SSR3), with an average of 0.67. The mean H_O (0.51) was lower than the H_E (0.67). Of the 30 SSR loci, the mean F_{ST} was 0.05, while F_{IS} and F_{IT} were 0.20 and 0.24, respectively. N_m varied from 14.98 (SSR29) to 1.61 (SSR23), with an average of 4.50.

Cluster analysis and genetic diversity

The genetic relationships between the accessions were determined based on shared allele distance for the 275 fragments. Overall, the 96 accessions were divided into three major clusters: A, B, and C (Figure 1).

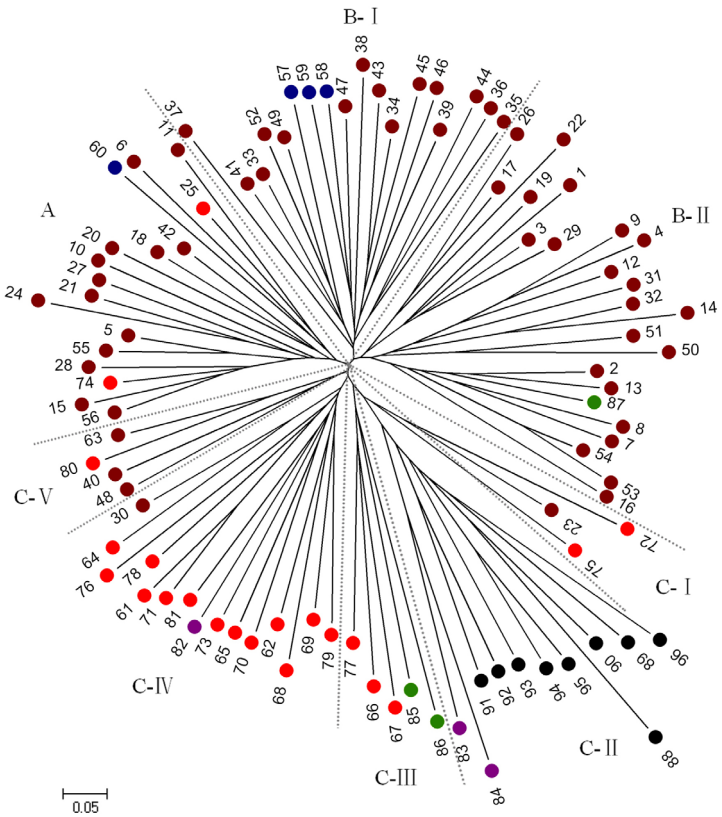

Figure 1. Cluster analysis of 96 individuals. Brown represents the cultivars collected from North American, blue represents the cultivars collected from Europe, red represents the cultivars from China, green represents the species (*Lagerstroemia subcostata*, *L. limii*, *L. fauriei*), purple represents *L. speciosa* and its interspecific hybrids, black represents *L. caudata* and its interspecific hybrids.

Cluster A comprised 17 cultivars: 14 cultivars collected from North America, No. 60 from Europe, and Nos. 25 and 74 from East Asia. Among the 17 cultivars, Nos. 10 and 20 had a common parent *L. fauriei* (No. 87). The rest of the cultivars originated from *L. indica* and the flower colors were red or purple.

Cluster B was composed of *L. fauriei* with 40 cultivars, of which three (Nos. 57, 58 and 59) were from Europe; the other cultivars were collected from North America. Two sub-clusters of cluster B were named B-I and B-II. Cluster B-I comprised 18 genotypes. This sub-cluster corresponded to growth habit; most were dwarf and weeping cultivars. Cluster B-II contained 20 of the 22 cultivars, which were interspecific hybrids between *L. fauriei* (No. 87) and *L. indica*.

Cultivars in Cluster C were cultivated in China except for Nos. 23, 30, 40, and 48, and originated from *L. indica*. All the interspecific hybrids shared the genetic background of *L. indica*, and were grouped in this cluster. Cluster C was further divided into five sub-clusters. Cluster C-IV comprised 15 cultivars: 14 Chinese cultivars, which were long-term selections from cultivated population of *L. indica* and No. 82. Clusters C-I and V contained Nos. 72, 75

and 80 from China and Nos. 23, 30, 40, and 48 from North America, respectively. Clusters II and III consisted of four species and their hybrids. *L. speciosa* (No. 84) and No. 83 are closely related, because *L. speciosa* was one of the direct parents of No. 83. Accession Nos. 89 to 96 shared a common parent, *L. caudata* (No. 88) in cluster C-II. In cluster C- III, Nos. 77, 66, and 67 were all cultivars from China and clustered with *L. subcostata* (No. 85) and *L. limii* (No. 86).

DISCUSSION

SSR primers developed from both *L. caudata* and *L. indica* produced clean amplification of PCR products with the expected allele sizes from 96 genotypes, indicating that a high proportion of SSRs from related species can be used in amplifying crape myrtles (Cai et al., 2011). The mean number of alleles per SSR locus (9) detected in this study was higher than that reported by Wang et al. (2011) (6.6) from 57 *L. indica* cultivars, five *L. fauriei* cultivars and 37 interspecific hybrids and higher than that detected by Cai et al. (2011) (5.8) from 50 Chinese crape myrtle cultivars. The higher number of alleles per locus in our study could reflect the differences between genotypes and the co-dominant SSR markers used in genetic diversity analysis. Compared to other studies on the genetic diversity of crape myrtle using AFLP and RAPD (Pooler, 2003; Yang et al., 2004; Gu et al., 2010), SSR markers showed a higher efficiency of polymorphism evaluation. The SSR loci that produced a higher number (12-15) of alleles, such as SSR2, SSR3, SSR5, SSR9, SSR17, SSR19, and SSR28, revealed a high gene diversity (PIC value), which ranged from 0.73 to 0.86 per locus. The average PIC among dinucleotide repeats was 0.7, while the PIC for tri- and tetranucleotide repeats was 0.67 and 0.55. In general, the SSRs with dinucleotide repeats showed higher allele diversity than did tri- and tetranucleotide repeats, and a direct relationship exists between marker information content and the number of repeat units (Innan et al., 1997; Ali et al., 2008; Cai et al., 2011). The PIC values in our study indicated that these markers are useful for molecular research, for differentiating genotypes and clustering them for genetic diversity analysis (Narasimhamoorthy, 2008).

The H_E (0.67) was higher than the H_O (0.51), which suggests that heterozygosity is deficient for these loci, possibly because of the presence of null alleles at these loci. In such cases, heterozygous plants carrying one null allele may be scored as homozygous for the readable allele (Dias et al., 2008). The hierarchical AMOVA indicated different levels of genetic variance among populations and among individuals within populations. Only 5% (F_{ST}) of the overall variation was the result of differences between the three geographic regions. The higher gene flow observed in our study might have been related to the sampling of populations over a larger geographic area (Mehes, 2009) and because the cultivars came from different regions. In our study, the results indicated a higher genetic diversity compared to other studies (Nan et al., 2003; Sereno et al., 2006; Ali et al., 2008), which probably resulted from the allogamy of the Lythraceae. Thus, the results in this study demonstrated that SSR markers are effective tools for improving the genus taxonomy of *Lagerstroemia*, especially in solving the confusion surrounding the exact number of species in this genus (Rinehart and Pounders, 2010).

In the cluster analysis, nearly all the crape myrtle cultivars were clustered together, while the *Lagerstroemia* species and interspecific hybrids were grouped together, which was similar to the results based on SSR markers of Rinehart and Pounders (2010), Wang et al. (2011), and Cai et al. (2011). There were several exceptions, i.e., the *L. indica* cultivars Nos.

19 and 32 clustered with *L. fauriei* (No. 87) while Nos. 66 and 77 clustered with *L. subcostata* (No. 85) or *L. limii* (No. 86). Nos. 19 and 32 were purchased from a commercial source. When they were propagated, transported, and sold, mislabeling could have occurred because of their similar flower color and growth habit (Wang et al., 2011). On the other hand, Nos. 19, 32, 66, and 77 lack a full and certain description and may be hybrids, although they were labeled as *L. indica*.

The cluster analysis showed that cultivars from the same geographic regions tended to group together (Zhu et al., 2011). Cultivars from North America mainly grouped in Clusters A and B, while Chinese cultivars grouped in Cluster C. There was an interesting exception in that Nos. 15, 56, and 74 are from different geographic regions but grouped together. Interestingly, according to the morphological database, they have many similar traits, such as the same flower color (Red 53B and White NN155C) and plant size (shrub). Based on this result, we concluded that the three cultivars may have a similar genetic background.

Within clusters, there are some sub-clusters corresponding to growth habit. Nos. 34 and 39, which originated from a bud mutation of No. 38, are dwarf cultivars. Likewise, Nos. 33, 35, 36, 37, 43, 46, and 49, which originated from seedling selection by Chopin from a checklist (http://www.usna.usda.gov/Research/Herbarium/Lagerstroemia/), are dwarf and weeping cultivars. The hybrids, namely Nos. 1, 2, 3, 4, 7, 8, 9, 12, 13, 14, 16, and 17, are trees or large shrubs, probably because they have the genetic background of *L. fauriei*. Clustering resulting from growth habit is likely to reflect shared pedigrees (Dirr et al., 2005; Wang et al., 2011) or the same breeding method.

Most cultivars in Cluster A originated from *L. indica*, and the flower colors were red or purple. Zhang et al. (2008) reported that the red or purple flower color is probably the result of specific anthocyanin accumulation that was inherited from *L. indica*. White flower color is associated with *L. fauriei*. Generally, the higher percentage of *L. fauriei* in a cultivar's genetic background, the lighter the flower color will be (Wang et al., 2011). Our results showed a similar trend as in the report of Wang. Nos. 2, 7, 8, 13, 16, and 54 clustered closely with *L. fauriei* (No. 87), contained half or a quarter of the genetic material of *L. fauriei* in their background, and produced light lavender or pink flowers. The other cultivars in sub-cluster B-II that contained less genetic background from *L. fauriei* had dark lavender or pink flowers.

Among the five species, *L. fauriei* was clustered with North American cultivars because it was a parent of these cultivars. The remaining four species were clustered with the Chinese cultivars. Eight interspecific hybrids (Nos. 89 to 96) were hybrids of *L. caudata* (No. 86) and Chinese cultivars, and they had different levels of fragrance because *L. caudata* is an aromatic species. *L. speciosa* and *L. caudata* were clustered together, while *L. subcostata* and *L. limii* were clustered closely. Species and their interspecific hybrids formed new cultivar groups, indicating that these accessions represent new resources for breeding programs.

Several genotypes that shared a common parentage were clustered into different groups instead of being clustered together. For example, Nos. 10 and 20 (Cluster A) and No. 23 (Cluster C-I) share *L. fauriei* in their parentage but were not grouped in Cluster B-II. Similarly, No. 82 (Cluster C-IV) has the parentage of *L. speciosa* (Cluster C-II), but they were not clustered together, which demonstrated that genetic relationships based on molecular markers do not always agree with the results estimated by pedigree information, because of unrealistic assumptions for estimating the co-ancestry coefficient (Ali et al., 2008).

In conclusion, based on the results of our study, the use of microsatellite markers was

highly informative for *Lagerstroemia*. Long time geographic isolation seems to have led to significant genetic differentiation among *Lagerstroemia* germplasm resources. The phylogenetic relationship and genetic variation of *Lagerstroemia* germplasm will provide a rich source of materials for the selection of appropriate parents for breeding programs.

ACKNOWLEDGMENTS

Research supported by the Fundamental Research Funds for the Central Universities of China (#JD2010-1) and the 12th Five Years Key Programs for Science and Technology Development of China (#2012BAD01B07).

REFERENCES

Ali ML, Rajewski JF, Baenziger PS and Gill KS (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Mol. Breed.* 21: 497-509.

Barkley NA, Roose ML, Krueger RR and Federici CT (2006). Assessing genetic diversity and population structure in a *Citrus* germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor. Appl. Genet.* 112: 1519-1531.

Bohn M, Utz HF and Melchinger AE (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs and SSRs and their use for predicting progeny variance. *Crop Sci.* 39: 228-237.

Brickell C (1996). Encyclopedia of Garden Plants. Macmillan Press, New York, 250-252.

Cai M, Meng R, Pan HT and Gao YK (2010). Isolation and characterization of microsatellite markers from *Lagerstroemia caudata* (Lythraceae) and cross-amplification in other related species. *Conservat. Genet. Resour.* 2: 89-91.

Cai M, Pan HT, Wang XF and He D (2011). Development of novel microsatellites in *Lagerstroemia indica* and DNA fingerprinting in Chinese *Lagerstroemia* cultivars. *Sci. Hortic.* 131: 88-94.

Dias PMB, Julier B, Sampoux JP and Barre P (2008). Genetic diversity in red clover (*Trifolium pratense* L.) revealed by morphological and microsatellite (SSR) markers. *Euphytica* 160: 189-205.

Dirr MA, Waters V and Kardos J (2005). New Protected Woody Plant Introductions from the University of Georgia. Department of Horticulture, University of Georgia, Athens. Available at [http://www.canr.org/Dirr%20 WoodyPlts-05.pdf]. Accessed November 27, 2010.

Egolf DR (1967). Four new *Lagerstroemia indica* L. cultivars. -'Catawba', 'Conestoga', 'Potomac' and 'Powhatan'. *Baileya* 15: 7-13.

Egolf DR (1981). 'Muskogee' and 'Natchez' *Lagerstroemia*. *HortScience* 16: 576-577.

Egolf DR (1986). 'Acoma', 'Hopi', 'Pecos' and 'Zuni' *Lagerstroemia*. *HortScience* 21: 1250-1252.

Egolf DR (1987). 'Apalachee', 'Comanche', 'Lipan', 'Osage', 'Sioux' and 'Yuma' *Lagerstroemia*. *HortScience* 22: 674.

Egolf DR (1990). 'Choctaw' *Lagerstroemia*. *HortScience* 25: 992-993.

Egolf DR and Andrick AO (1978). The Lagerstroemia Handbook/Checklist: A Guide to Crape Myrtle Cultivars. American Association of Botanical Gardens and Arboreta, Las Cruces.

Gu CH, Bao ZY, Wang SX and Zhang QX (2010). AFLP analysis on the genetic relationship of *Lagerstroemia subcostata*, *Lagerstroemia limii* and 37 cultivated varieties. *Mol. Plant Breed.* 8: 730-735.

Innan H, Terauchi R and Miyashita NT (1997). Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. *Genetics* 146: 1441-1452.

Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129.

Mehes M, Nkongolo KK and Michael P (2009). Assessing genetic diversity and structure of fragmented populations of eastern white pine (*Pinus strobus*) and western white pine (*P. monticola*) for conservation management. *J. Plant Ecol.* 2: 143-151.

Nan P, Shi SH, Peng SL and Tian CJ (2003). Genetic diversity in *Primula obconica* (Primulaceae) from Central and Southwest China as revealed by ISSR markers. *Ann. Bot.* 91: 329-333.

Narasimhamoorthy B, Saha M, Swaller CT and Bouton JH (2008). Genetic diversity in switchgrass collections assessed by EST-SSR markers. *Bioenerg. Res.* 1: 136-146.

Pooler MR (2003). Molecular genetic diversity among 12 clones of *Lagerstroemia fauriei* revealed by AFLP and RAPD markers. *HortScience* 38: 256-259.

Pooler MR (2006). 'Arapaho' and 'Cheyenne' *Lagerstroemia*. *HortScience* 41: 855-856.

Pounders C (2007). Evaluation of interspecific hybrids between *Lagerstroemia indica* and *L. speciosa*. *HortScience* 42: 1317-1322.

Rinehart TA and Pounders CT (2010). Estimating diversity among *Lagerstroemia* species and hybrids using SSR markers. *Acta Hort.* 885: 285-290.

Schneider S, Roessli D and Excoffier L (2000). Arlequin ver. 2.000: A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory. University of Geneva, Geneva.

Schuelke M (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18: 233-234.

Sereno ML, Albuquerque PSB, Vencovsky R and Figueira A (2006). Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conservat. Genet.* 7: 13-24.

Subramanyam K, Muralidhararao D and Devanna N (2009). Genetic diversity assessment of wild and cultivated varieties of *Jatropha curcas* (L.) in India by RAPD analysis. *Afr. J. Biotechnol.* 8: 1900-1910.

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

Wang XW, Dean D, Wadl P and Hadziabdic D (2010). Development of microsatellite markers from crape myrtle (*Lagerstroemia* L.). *HortScience* 45: 842-844.

Wang XW, Wadl PA, Pounders C and Trigiano RN (2011). Evaluation of genetic diversity and pedigree within crape myrtle cultivars using simple sequence repeat markers. *J. Am. Soc. Hort. Sci.* 136: 116-128.

Wright S (1969). Evolution and the Genetics of Populations. Vol. 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago.

Yang YL, Zhang YD, Zhang XY and Lei XH (2004). Study on genetic diversity of wild *Lagerstroemia* from Baokang. *J. Huazhong Agr. Univ.* 23: 667-670.

Yeh FC, Yang RC and Boyle T (1999). POPGENE (Microsoft Windows-Based Freeware for Population Genetic Analysis) Version 1.31. Centre for International Forestry Research, University of Alberta, Edmonton.

Zhang C, Chen X, He T, Liu X, et al. (2007). Genetic structure of *Malus sieversii* population from Xinjiang, China, revealed by SSR markers. *J. Genet. Genomics* 34: 947-955.

Zhang J, Wang LS, Gao JM, Shu QY, et al. (2008). Determination of anthocyanins and exploration of relationship between their composition and petal coloration in crape myrtle (*Lagerstroemia hybrid*). *J. Integr. Plant Biol.* 50: 581-588.

Zhang QX (1991). Studies on cultivars of crape myrtle (*Lagerstroemia indica*) and their uses in urban greening. *J. Beijing For. Univ.* 13: 57-66.

Zhu RR, Gao YK, Xu LJ and Zhang QX (2011). Genetic diversity of *Aquilegia* (Ranunculaceae) species and cultivars assessed by AFLPs. *Genet. Mol. Res.* 10: 817-827.

