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

Thirty-four *Musa* (Musaceae) expressed sequence tag-derived microsatellite markers transferred to *Musella lasiocarpa*

W.J. Li[1,2]*, H. Ma[1]*, Z.H. Li[1], Y.M. Wan[1], X.X. Liu[1] and C.L. Zhou[1]

[1]Biotechnology Department, Research Institute of Resource Insects, Chinese Academy of Forestry, Kunming, China
[2]Yunnan Reascend Tobacco Technology (Group) Company Limited, Kunming, China

*These authors contributed equally to this study.
Corresponding author: Z.H. Li
E-mail: lzh4949@163.com

Genet. Mol. Res. 11 (3): 2094-2098 (2012)
Received January 9, 2012
Accepted March 30, 2012
Published August 6, 2012
DOI http://dx.doi.org/10.4238/2012.August.6.13

ABSTRACT. We assembled 31,308 publicly available *Musa* EST sequences into 21,129 unigenes; 4944 of them contained 5416 SSR motifs. In all, 238 unigenes flanking SSRs were randomly selected for primer design and then tested for amplification in *Musella lasiocarpa*. Seventy-eight primer pairs were found to be transferable to this species, and 49 displayed polymorphism. A set of 34 polymorphic SSR markers was analyzed in 24 individuals from four wild *M. lasiocarpa* populations. The mean number of alleles per locus was 3.0, ranging from 2 to 7. The observed and expected heterozygosities per marker ranged from 0.087 to 0.875 (mean 0.503) and from 0.294 to 0.788

(mean 0.544), respectively. These markers will be of practical use for genetic diversity and quantitative trait loci analysis of *M. lasiocarpa.*

Key words: EST-SSRs; Molecular markers; *Musella*; Transferability; Polymorphism

INTRUDUCTION

Musella lasiocarpa, normally known as Chinese dwarf banana, native to southwestern China, represents the monotypic genus *Musella* in the family Musaceae (Wu and Kress, 2000). Its showy and lotus-like golden inflorescence with more than eight month anthesis has strongly attracted horticulturists; consequently, it has been introduced into various regions and botanical gardens around the world. Wild populations of *M. lasiocarpa* are found on cliffs within the watershed of the upper Yangtze River and its tributaries, in northern Yunnan and southern Sichuan Provinces in China. It is now difficult to find wild populations as appropriate habitats have been fragmented due to widespread expansion of agriculture (Ma et al., 2011). However, it is locally common in this region of southwestern China, maintained as a semi-cultivated plant. This taxon is extensively used by farmers as fodder for livestock, for erosion control, and as weaving material (Ma et al., 2011). Plants have been utilized as an edible vegetable, as medicine, in wine making, and even as a resource for honey production (Liu et al., 2003; Long et al., 2008). Considering the high ornamental value and endemism of *M. lasiocarpa*, we have been committed to investigate the diversity of its wild populations since 2004 and have documented nine wild populations (one of them contains less than 30 tufts) to date. We assessed the conservation status of its wild populations as Endangered, according to IUCN Red List criteria (IUCN, 2001), because the number of population locations and mature plants could fluctuate extremely in the near future.

Microsatellite (SSR) markers have proven to be a powerful tool to investigate genetic diversity, mapping and cultivar identification. However, there have been relatively few reports of SSR markers in *M. lasiocarpa* (Yang et al., 2009). Expressed sequence tag (EST)-derived SSR markers show a higher level of transferability to closely related species and fewer null alleles than genomic-SSRs and provide an efficient means for SSR screening (La Rota et al., 2005; Cristancho and Escobar, 2008; Li et al., 2010). We found no genetic maps or EST sequences for *M. lasiocarpa* in GenBank. However, it is reasonable to expect that useful SSR markers could be developed from *Musa* (Musaceae) EST sequences.

MATERIAL AND METHODS

Twenty-four individuals from four wild populations of *M. lasiocarpa* from Yunnan Province, China, were collected to screen for transferability and polymorphism of *Musa* EST-derived SSR markers. All of them had green petioles, pale green primary veins and yellow bracts. Plant genomic DNA from young leaves was extracted using the CTAB method (Doyle and Doyle, 1990) and stored at -20°C.

We obtained 31,308 publicly available *Musa* EST sequences (11,155 from

Musa paradisiaca, 5289 from *Musa balbisiana*, 14,864 from *Musa acuminata*) from the NCBI database (http://www.ncbi.nlm.nih.gov/nucest) on September 10, 2010 and saved them as Fasta-formatted files. The 31,308 sequences were assembled into 21,129 non-redundant clusters (11,165 for *M. acuminata*, 5908 for *M. paradisiaca* and 4056 for *M. balbisiana*), using the CAP3 software (Huang and Madan, 1999). Unigenes containing SSRs were identified with SSRIT (http://www.gramene.org/db/markers/ssrtool). The minimum length of SSR was defined as di-, tri-, tetra-, penta-, and hexanucleotide (DNRs, TNRs, TeNRs, PNRs, and HNRs) for 6, 4, 3, 3, and 3 subunits, respectively. Across all three species, 5416 SSRs (234 were compound SSRs) were identified in 4944 unigenes. Primer pairs were designed with the help of PRIMER3 (Rozen and Skaletsky, 2000), using the following settings: optimal primer length 20 bp (maximum 25 bp, minimum 18 bp), optimal annealing temperature (Tm) 57°C (maximum 63°C, minimum 52°C, maximum temperature difference 5°C), product size 100-500 bp. The CG content varied from 35 to 65%.

The PCR was performed in a 20 µL volume containing 25 ng genomic DNA, 0.75 U *Taq* DNA polymerase, 1X PCR buffer (10 mM Tris-HCl, pH 8.0, 50 mM KCl), 1.8 mM $MgCl_2$, 0.13 mM dNTPs (Takara), and 0.25 µM of each forward and reserve primers. The PCR was carried out using the following thermal cycling program: initial denaturation at 95°C for 4 min, followed by 30 cycles of 95°C for 45 s, 45 s at primer-specific annealing temperature of 50° to 63°C, 72°C for 1 min, and a final extension step at 72°C for 10 min. PCR products were separated on 8% non-denatured polyacrylamide gel and then stained with ethidium bromide. The successful PCR products were subsequently sized on an ABI PRISM 3730xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The forward primers were labeled at the 5'-end with an FAM fluorescent dye.

The number of alleles (N_A), linkage disequilibrium, observed (H_O) and expected (H_E) heterozygosities were calculated by the PopGen version 1.31 software (http://www.ualberta.ca/~fyeh/). The fitness of genotypic frequency to the Hardy-Weinberg equilibrium (HWE) was determined using FSTAT version 2.9.3 (Goudet, 1995).

RESULTS AND DISCUSSION

In all, 78 of the 238 *Musa* EST-derived SSR primers successfully amplified PCR products at close to the size expected in *M. lasiocarpa*. The remaining primer pairs amplified no products, or gave weak or indistinct bands, or gave obviously larger sizes than expected. In the 34 polymorphic EST-SSR primer pairs that were identified, the N_A ranged from 2 to 7, with a mean of 3.0. Compared with the loci developed from the genome (Yang et al., 2009), the N_A per locus was relatively low. The H_O ranged from 0.087 to 0.875 (mean 0.503), and H_E ranged from 0.294 to 0.788 (mean 0.544) (Table 1). Among the 34 microsatellite loci, nine showed significant deviation from HWE ($P < 0.01$), indicating the possibility of null alleles or sampling problems. All loci were in linkage disequilibrium. The addition of 34 polymorphic microsatellite markers to the genomic-SSR markers (Yang et al., 2009) will prove valuable for future analysis of genetic diversity of wild populations and study of quantitative traits of *M. lasiocarpa*.

Table 1. Transferability of 34 *Musa* EST-SSR markers to *Musella lasiocarpa*.

Locus	Accession No.	Repeat motif	Primer sequences(5'-3')	Ta (°C)	*M. lasiocarpa*				
					Size range	N_A	H_O	H_E	P_{HWE}
Smu2	FL641808	$(CTC)_5$	F: TTCCAACCTGCCACCCTCCT R: GCCTTTGATTCCGGCGAAGC	56	255-273	2	0.708	0.467	0.991
Smu10	FL646854	$(CTT)_5$	F: TGCGGAAGTGCAGTGTGACC R: TGCACTTGCGAGCAGACGTT	54	199-220	5	0.542	0.788	0.001*
Smu19	FL648830	$(GAGTCT)_6$	F: TCATGCGCTCAAGCTCCGTT R: AAGGGTCCGATTCTTCGGCG	54	226-250	4	0.167	0.694	0.000*
Smu24	FL650445	$(GAA)_7$	F: GGGATAGAGCAGCAGACCGC R: CTGGCAGCGAACCGATCCAT	61	164-179	4	0.583	0.715	0.048
Smu29	FL648383	$(ATCG)_4$	F: GCTTCATGGAGCTGCGGGAG R: ACGATGAAGGAGGCCGAGGA	51	288-312	2	0.458	0.361	0.771
Smu148	FL649365	$(GAT)_6$	F: GCAGAAACAGGCACCGGGTA R: TGTGTTCGGGTGCTCCATGT	57	219-228	2	0.750	0.507	0.980
Smu153	FL646992	$(GA)_{13}$	F: GAGTCAGTGAGCGAGGAGCG R: AGGGCGAGGGTTCTTCTCGT	53	211-215	3	0.208	0.543	0.000*
Smu155	FL650274	$(AT)_9$	F: GAGATGGCCTGTGGCGTCAA R: GAAAGCTTCTCGAGGGCCGA	55	175-179	2	0.348	0.510	0.024
Smu157	FL649749	$(GTG)_4$	F: GAGACCCGACACTTCGGTGG R: TCCGTGTGCCGAAAGGTCTG	53	202-214	3	0.087	0.552	0.000*
Smu158	FL649571	$(GTG)_4$	F: GACGGAGACCCGACACTTCG R: AAGCCCGAGATCTTGCCTGC	52	293-296	2	0.833	0.497	0.999
Smu162	FL651169	$(AG)_{20}$	F: GAAGCGGCGACTTTGATCGC R: AGTCGGGCTTCTTGTGAGCG	54	195-215	7	0.417	0.616	0.005*
Smu180	FL646390	$(TCG)_5$	F: CGCTCCACTCCACGCTTCTT R: ACGATATGCGGTGCCCATCC	52	211-236	4	0.583	0.481	0.871
Smu181	FL647640	$(TCG)_5$	F: CGCTCCACTCCACGCTTCTT R: GGTGCCCATCCTTCTCCGTG	63	212-230	3	0.636	0.661	0.384
Smu182	FL646375	$(GCC)_8$	F: CGCCTCCGAAACCCTTCCTC R: CCGCGTTCCCACAGTATCCG	54	189-192	2	0.391	0.476	0.303
Smu184	FL651442	$(CCT)_8$	F: CGAGTCGACGCCTGTGAGAG R: ACTGCTTGATGGCGTCCTGG	51	147-171	4	0.667	0.687	0.324
Smu193	FL647754	$(CCT)_9$	F: CCGTCTCTTCCGCTGCTCTC R: CTGAGCCCGTCGAATCTCGG	55	178-187	2	0.522	0.394	0.863
Smu198	FL650994	$(TAAACC)_8$	F: CATCCTCGTGGTGCAGGCAA R: CCACCAACAACGAGGTAGCCA	56	229-241	2	0.542	0.503	0.690
Smu221	FL650890	$(TCGA)_4$	F: CAGCGAGTCAGTGAGCGAGG R: CTCTGCGATCTGGTCGTCGG	54	194-202	3	0.261	0.564	0.001*
Smu237	FL646654	$(CTG)_7$	F: ACCACTGGCTGGGTAGGCAT R: TGCATTCTCAACAAGAAATGAAGGT	50	201-225	5	0.583	0.752	0.003*
Smu239	FL646915	$(AGA)_{12}$	F: ACAAGATGAGGCCACTGCGG R: ACAAACGCGAACCCTCTGCT	51	295-307	2	0.667	0.454	0.980
Mms10	FL646660	$(TCA)_8(GT)_8$	F: GCTTTCGGCCTCGGAGAAGT R: GAGAGAGAGAGAGAACACAGTTCA	53	135-147	4	0.652	0.737	0.124
Musf8	DN238224	$(TC)_{14}$	F: CACTGCACCTCGGTGGCAAA R: TCGATCCACCTGAGCAACACGA	51	214-218	3	0.565	0.670	0.102
Musf21	FL665190	$(TTC)_{11}$	F: TCCACCAATAGCAGGCAACACC R: CCATGCCCAAGCTTCCGCTAA	52	233-242	3	0.174	0.618	0.000*
Musf22	FL665824	$(TG)_{19}$	F: CCGTCACCCATCCCTTGGGTA R: ACCGCGTTCATCACCCGTTT	54	187-191	2	0.667	0.507	0.905
Musf24	FL663908	$(AGA)_{11}$	F: GCAGCAGAGTGATCAAAGGAGGC R: CGCATGGAACCACCGGACAA	59	233-239	2	0.261	0.294	0.217
Musf25	FL663530	$(AGATGA)_6$	F: TGGAGGGACTTGCAGCCCAA R: GCCACGAACGCACAACCTCA	55	157-175	3	0.783	0.618	0.949
Musf29	ES433855	$(CTC)_{10}$	F: TGTAAGCGCAACCCTGCCAC R: TCAGCTGATTTGACGCGGGC	54	214-226	3	0.875	0.547	1.000
Musf30	FL660562	$(AG)_{12}$	F: CGGTGTCAGTAGCAGCAGTGG R: CGCTTACATCGGCGAGCGAA	60	169-173	2	0.583	0.479	0.776
Musf31	FL660501	$(GAA)_8$	F: ATCAGCGAGTCGTCGGAGCA R: GCAGCAGCAGCAGCTAACGA	58	186-192	2	0.250	0.422	0.009*
Musf34	ES433050	$(GAA)_8$	F: TGGCGATGCCCAGGAACTCA R: TCGAGACGTCCGTCGACAAGT	52	149-152	2	0.208	0.311	0.022

Continued on next page

Table 1. Continued.

Locus	Accession No.	Repeat motif	Primer sequences (5'-3')	Ta (°C)	*M. lasiocarpa*				
					Size range	N_A	H_O	H_E	P_{HWE}
Musf37	ES436430	$(CAG)_8$	F: CAGCTTCAGCCACAGTCGCA R: GCATGTTGGTTGGCCCTGCT	62	238-247	2	0.609	0.510	0.890
Musf39	ES432796	$(AG)_{14}$	F: CGGTGCCAATGGTGCAAACG R: AGCATGGCTGAGGTAGTCGCA	60	182-192	4	0.591	0.606	0.364
Musf41	ES432772	$(AG)_{16}$	F: CCGATCTCCCATGGCAGCAAA R: TGCATCCGCGACGATGTCTG	56	248-272	4	0.667	0.646	0.491
Musf50	DN239657	$(GCC)_{10}$	F: GCTCGCCAACGGTAATGGCA R: AAGGTGAGCCAGAGGCAGCA	56	177-186	2	0.273	0.304	0.363

*Significant deviation from Hardy-Weinberg equilibrium ($P < 0.01$). Ta = annealing temperature; N_A = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity.

ACKNOWLEDGMENTS

Research supported by the Special Fund for Forestry Research in the Public Interest Granted by the State Forestry Administration of China (#200904053) and the Yunnan Natural Science Foundation (#2009ZC188M).

REFERENCES

Cristancho M and Escobar C (2008). Transferability of SSR markers from related Uredinales species to the coffee rust *Hemileia vastatrix*. *Genet. Mol. Res.* 7: 1186-1192.

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Goudet J (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *J. Hered.* 86: 485-486.

Huang X and Madan A (1999). CAP3: A DNA sequence assembly program. *Genome Res.* 9: 868-877.

IUCN (2001). IUCN Red List Categories and Criteria, Version3.1. Prepared by the IUCN Species Survival Commission. IUCN, Gland, Switzerland, and Cambridge.

La Rota M, Kantety RV, Yu JK and Sorrells ME (2005). Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley. *BMC Genomics* 6: 23.

Li XY, Shangguan LF, Song CN, Wang C, et al. (2010). Analysis of expressed sequence tags from *Prunus mume* flower and development of simple sequence repeat markers. *BMC Genet.* 11: 66.

Liu AZ, Kress WJ and Long CL (2003). Customary use and conservational attention to *Musella lasiocarpa* (Musaceae): a monotypic genus endemic to China. *Econ. Bot.* 57: 279-281.

Long CL, Selena A, Wang XR, Liu YT, et al. (2008). Why *Musella lasiocarpa* (Musaceae) is used in Southwest China to feed pigs. *Econ. Bot.* 62: 182-186.

Ma H, Pan QJ, Wang L, Li ZH, et al. (2011). *Musella lasiocarpa* var. *rubribracteata* (Musaceae), a new variety from Sichuan, China. *Novon* 21: 349-353.

Rozen S and Skaletsky HJ (2000). Primer 3: Bioinformatics Methods and Protocols. In: Methods in Molecular Biology (Krawetz S and Misener S, eds.). Humana Press, New Jersey, 365-386. Available at [http://frodo.wi.mit.edu/primer3/]. Accessed November 1, 2010.

Wu DL and Kress WJ (2000). Flora of China: Musaceae. In: Flagellariaceae through Marantaceae (Wu CY and Raven PH, eds.). Vol. 24. Science Press, Beijing, 314-318.

Yang CY, Huang Y and Long CL (2009). Isolation and characterization of 17 polymorphic microsatellite loci for *Musella lasiocarpa* (Musaceae). *HortScience* 44: 2041-2042.