American Journal of Botany

American Journal of Botany: e17–e19. 2012.

AJB Primer Notes & Protocols in the Plant Sciences

New microsatellite markers for garlic, *Allium sativum* (Alliaceae)[1]

Camila P. Cunha[2], Eulália S. S. Hoogerheide[3], Maria I. Zucchi[4], Mariza Monteiro[4], and José B. Pinheiro[2,5]

[2]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz," Universidade de São Paulo, Av. Pádua Dias 11, Caixa Postal 83, Piracicaba, São Paulo 13400-970, Brazil; [3]Embrapa Agrossilvipastoril, Av. dos Jacarandás 2639, Sinop, Mato Grosso 78550-003, Brazil; and [4]Agência Paulista de Tecnologia dos Agronegócios, Rodovia SP 127, Vila Fátima, Piracicaba, São Paulo 13400-970, Brazil

- *Premise of the study:* A new set of microsatellite or simple sequence repeat (SSR) markers for garlic, an important medicinal spice, was developed to aid studies of genetic diversity and to define efficient strategies for germplasm conservation.
- *Methods and Results:* Using a $(CT)_8$- and $(GT)_8$-enriched library, a total of 16 SSR loci were developed and optimized in garlic. Ten loci were found to be polymorphic after screening 75 accessions. The parameters used to characterize the loci were observed and expected heterozygosity, number of alleles, Shannon Index, and polymorphism information content (PIC). A total of 44 alleles were identified, with an average of 4.4 alleles per loci. The vast majority of loci were moderate to highly informative according to PIC and the Shannon Index.
- *Conclusion:* The new SSR markers have the potential to be informative tools for genetic diversity, allele mining, mapping and associative studies, and in the management and conservation of garlic collections.

Key words: Alliaceae; *Allium sativum*; garlic; genetic diversity; germplasm; simple sequence repeat markers.

Garlic (*Allium sativum* L., Alliaceae) has been used since as early as 5000 BC as a medicine and food condiment. Beneficial compounds to our health such as oligosaccharides, steroidal glycosides, essential oil, flavonoids, anthocyanins, lectins, prostaglandins, fructan, pectin, adenosine, and vitamins are found in garlic bulbs and leaves, but the organosulfur compound allicin is responsible for its medicinal properties (Arzanlou and Bohlooli, 2010). The regular consumption of garlic prevents cardiovascular diseases, diabetes, asthma, and cancer (Rana et al., 2011). Furthermore, its antioxidant and antimicrobial activities protect the immunologic system. Because of the medical efficacy and nutraceutical properties of garlic, there is increased interest in the development of new varieties.

The center of origin of garlic is Central Asia, where the greatest number of varieties of the genus are found. Commercial cultivars are diploid ($2n = 16$), sterile, and obligate clonally propagated. Although male fertility in wild garlic has been observed in nature and genetic linkage maps from S_1 (family of plants produced by self-pollination) are already available (Ipek et al., 2005; Zewdie et al., 2005), the selection of spontaneous or induced mutations still plays an important role in garlic breeding. Breeding programs depend on genetic variability and identification of traits of interest, which require a rigorous characterization of accessions within germplasm collections. Significant levels of genetic variability were detected in garlic germplasm by morphological traits, isozymes, and molecular markers, e.g., random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and, recently, simple sequence repeat (SSR) markers (Ma et al., 2009). Today, SSRs are the marker of choice for a broad number of genetic studies because of their high polymorphism, codominance, genomic abundance, and facility in laboratory usage. The number of available polymorphic SSR markers for garlic in the literature is still reduced compared to other minor crops. This fact reinforces the need to develop new SSR markers. In the current study, 16 new garlic SSR loci are described, enhancing the number of tools available for genetic studies of this important crop.

METHODS AND RESULTS

Total genomic DNA was extracted from lyophilized leaf samples from a single individual using a modified cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1990). A microsatellite-enriched genomic library was constructed following the adapted protocol from Billotte et al. (1999). The genomic DNA of garlic was digested by the *Rsa*I enzyme (Invitrogen, Carlsbad, California, USA) and microsatellite fragments were selected by $(CT)_8$ and $(GT)_8$ motif probes. The fragments in the enriched library were cloned into pGEM-T (Promega, Madison, Wisconsin, USA) and used to transform TOP10 *Escherichia coli* competent cells (Invitrogen). The positive clones were selected using the beta-galactosidase gene and grown overnight in liquid medium containing ampicillin. A total of 384 clones were sequenced in an ABI 377 automated sequencer (Applied Biosystems, Foster City, California, USA) using the BigDye terminator cycle sequencing kit (Applied Biosystems). Microsatellite sequences were identified using WebSat (Martins et al., 2009).

A total of 16 different nucleotide sequences, containing SSR motifs of interest, were used to design primers in the SSR flanking regions using Primer3 software

[1]Manuscript received 8 June 2011; revision accepted 16 August 2011.

The authors thank Talita Val and Jaqueline Campos for technical assistance and Francisco Resende for providing the garlic accessions. This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (grants 2008/02936-5 and 2009/03577-1).

[5]Author for correspondence: baldin@esalq.usp.br

doi:10.3732/ajb.1100278

(http://frodo.wi.mit.edu/primer3/; Rozen and Skaletsky, 2000). These sequences are available at the National Center for Biotechnology Information website (NCBI; http://www.ncbi.nlm.nih.gov) and can be found using the GenBank accession numbers provided in Table 1. The new set of primers was tested in a total of 75 garlic accessions from "Luiz de Queiroz" College of Agriculture (ESALQ), Agronomic Institute from Campinas (IAC), and Embrapa Vegetable Crops (Embrapa Hortaliças) (Appendix 1). The accessions were maintained in a garden environment at ESALQ. PCR was performed in 20 μL final volume, consisting of 1× buffer (Invitrogen), 20 ng of genomic DNA, 1.5 mM of $MgCl_2$, 0.2 μM of each primer, 0.2 mM of dNTPs, and 1 U *Taq* DNA polymerase (Invitrogen).

The amplification program used is described in Ma et al. (2009), consisting of an initial denaturing step at 94°C for 3 min; followed by 30 cycles at 94°C for 30 s, at the specific annealing temperature of each pair of primers (Table 1) for 45 s, and at 72°C for 1 min; 10 cycles at 94°C for 30 s, at 2°C below the specific annealing temperature of each pair of primers for 45 s, and at 72°C for 1 min; and a final elongation step at 72°C for 10 min. Amplification products were confirmed by electrophoresis in 7% denaturing polyacrylamide gels stained in a sodium hydroxide protocol (Sanguinetti et al., 1994). Allele scoring was carried out using 10 bp DNA Ladder (Invitrogen) as a size standard. The characterization of polymorphic SSR loci was performed using observed (H_o) and expected heterozygosity (H_e), number of alleles (N_a), polymorphism information content (*PIC*), and Shannon Index (I_S) obtained using the software MS-Toolkit (http://animalgenomics.ucd.ie/sdepark/ms-toolkit; Park, 2001) and MSA (Dieringer and Schlötterer, 2003). Deviations from Hardy–Weinberg equilibrium (HWE) were calculated by exact tests using TFPGA version 1.3 (Miller, 1997), and the *P* value was obtained by the conventional Monte Carlo method using 10 batches and 1000 permutations per batch. The significance threshold was adjusted using Bonferroni correction.

The description of 16 SSR loci is shown in Table 1. From the screening of 75 garlic accessions, six primers were monomorphic (Asa04, Asa06, Asa20, Asa23, Asa27, and Asa59) and did not reveal differences among accessions. Ten loci were polymorphic and are characterized in further detail in Table 2. A total of 44 alleles were identified using the polymorphic SSR loci, ranging from two to eight alleles per loci. Observed and expected heterozygosity ranged from 0 to 1 and 0.212 to 0.809, respectively (Table 2). Eight out of 10 loci were considered moderate to highly informative (*PIC* > 0.3 and I_S > 0.6) and, therefore, adequate for genetic studies. The average *PIC* and I_S were 0.518 and 1.073, respectively. In addition, monomorphic and low informative loci, Asa16 and Asa18, might be useful in a different set of accessions. Eight loci deviated significantly from expected HWE proportions ($P > 0.005$), with the exception of Asa07 and Asa18. The deviation from HWE was expected because the accessions used in this study are part of germplasm collections, in which genotypes are maintained by clonal propagation, using bulbils, and no panmixia is found.

CONCLUSIONS

In combination with previously available SSR markers, the polymorphic markers developed in this study will be useful for further studies of genetic diversity, allele mining, mapping and associative studies, in the management and conservation of garlic collections, and potentially for interspecific genetic studies within the genus *Allium*. The genetic variability within germplasm collections, assessed by those SSR markers, is a source of information for breeding programs to engineer new varieties with higher nutritional value.

TABLE 1. Characteristics of 16 microsatellite primers developed in *Allium sativum.* Shown for each primer pair are the forward and reverse sequence, repeat type, annealing temperature when run individually, size of the original fragment, and GenBank accession number. All values are based on the screening of 75 accessions of garlic.

Loci	Primer sequence (5′–3′)	Repeat motif	T_a (°C)	Allele size (bp)	GenBank Accession No.
Asa04	F: AGACTTTTGGAGGCTAGGGC	$(TCC)_5(TCC)_4(TCC)_5$	54	264	JN084085
	R: CCCTGGTCTCTTTCAACCAA				
Asa06	F: GGGGTGTTACATTCTCCCCT	$(TG)_5$	57	192	JN084086
	R: ACCGCCTGATTTTGCATTAG				
Asa07	F: CTCGGAACCAACCAGCATA	$(TG)_7$	58	229–235	JN084087
	R: CCCAAACAAGGTAGGTCAGC				
Asa08	F: TGATTGAAACGAATCCCACA	$(GT)_8$	56	209–257	JN084088
	R: GGGGGTTACCTGAACCTGTTA				
Asa10	F: TTGTTGTTCTGCCATTTT	$(AC)_7$	48	225–239	JN084089
	R: GATCTAAGCCGAGAGAAA				
Asa14	F: TCTATCTCGCTTCTCAGGGG	$(GT)_7$	48	220–234	JN084090
	R: GCTGACAGAAGTAGTCTTTCC				
Asa16	F: CACGACTTTTCCTCCCATTT	$(TG)_5C(GT)_6$	48	148–154	JN084091
	R: GCTAATGTTCATGTCCCCAGT				
Asa17	F: TCCACGACACACACACACAC	$(CA)_{12}(CT)_{28}$	56	126–196	JN084092
	R: ATGCAGAGAATTTGGCATCC				
Asa18	F: TCAAGCTCCTCCAAGTGTCC	$(TG)_8$	45	254–264	JN084093
	R: TCGGGATATGACAGCATTTG				
Asa20	F: GAAGCAGCAAAGATCCAAGC	$(G)_{12}$	48	260	JN084094
	R: CGTGCAGAACTTAACCTT				
Asa23	F: TGGAGGGGGAAAAAGGATAG	$(GA)_5$	55	271	JN084095
	R: TGTGAAGCAAGTGGGATCAA				
Asa24	F: TTGTTGTGCCGAGTTCCATA	$(GT)_4(GT)_3(GT)_5$	48	149–161	JN084096
	R: CAGCAATTTACCAAAGCCAAG				
Asa25	F: GCACTTCACTTTCCCCATTC	$(CT)_3(CT)_{27}$	51	117–127	JN084097
	R: GGCGACGGTGAAGAGAGAG				
Asa27	F: GGGAGAGAATGGCTTGATTG	$(TC)_{17}(TC)_5$	55	127	JN084098
	R: GGACAGCATCATCACCAC				
Asa31	F: CAGAGACTAGGGCGAATGG	$(CTT)_7$	50	237–243	JN084099
	R: ATGATGATGACGACGACGAG				
Asa59	F: CGCTTACTATGGGTGTGTGTC	$(ATCA)_3$	50	290	JN084100
	R: CAAGTGGGAGACTGTTGGAG				

Note: T_a = annealing temperature.

TABLE 2. Results using polymorphic SSR primers to screen 75 *Allium sativum* accessions.

Loci	N_a	H_o	H_e	I_S	*PIC*	*P* HWE[a]
Asa07	3	0.467	0.511	0.765	0.398	0.2806[n.s.]
Asa08	8	1.000	0.809	1.760	0.775	0.0000*
Asa10	5	0.160	0.661	1.248	0.608	0.0000*
Asa14	4	0.819	0.720	1.323	0.667	0.0000*
Asa16	2	0.000	0.322	0.500	0.269	0.0000*
Asa17	5	0.000	0.576	1.024	0.502	0.0000*
Asa18	2	0.239	0.212	0.366	0.189	0.5801[n.s.]
Asa24	7	1.000	0.715	1.562	0.710	0.0000*
Asa25	6	1.000	0.737	1.497	0.689	0.0000*
Asa31	2	0.000	0.501	0.690	0.374	0.0000*
Average	4.4	0.468	0.576	1.073	0.518	—

Note: H_e = expected heterozygosity; H_o = observed heterozygosity; HWE = Hardy–Weinberg equilibrium; I_s = Shannon Index; N_a = number of alleles; n.s. = not significant; *PIC* = polymorphism information content.

[a] Significance threshold adjusted using Bonferroni correction: *P* (5%) ≤ 0.005.

* Statistically significant at 5% level.

LITERATURE CITED

ARZANLOU, M., AND S. BOHLOOLI. 2010. Introducing of [*sic*] green garlic plant as a new source of allicin. *Food Chemistry* 120: 179–183.

BILLOTE, N., P. J. L. LAGODA, A. RISTERUCCI, AND F. BAURENS. 1999. Microsatellite-enriched libraries: Applied methodology for the development of SSR markers in tropical crops. *Fruits* 54: 277–288.

DIERINGER, D., AND C. SCHLÖTTERER. 2003. Microsatellite analyser (MSA): A platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3: 167–169.

DOYLE, J. J., AND J. L. DOYLE. 1990. Isolation of plant DNA fresh tissue. *Focus (San Francisco, Calif.)* 12: 13–15.

IPEK, M., E. A. IPEK, E. S. G. ALQUIST, AND E. P. W. SIMON. 2005. Demonstration of linkage and development of the first low-density genetic map of garlic, based on AFLP markers. *Theoretical and Applied Genetics* 110: 228–236.

MA, K.-H., J.-G. KWAG, W. ZHAO, A. DIXIT, G.-A. LEE, H.-H. KIM, I.-M. CHUNG, ET AL. 2009. Isolation and characteristics of eight novel polymorphic microsatellite loci from the genome of garlic (*Allium sativum* L.). *Scientia Horticulturae* 122: 355–361.

MARTINS, W. S., D. C. S. LUCAS, K. F. S. NEVES, AND D. J. BERTIOLLI. 2009. WebSat: A web software for microsatellite marker development. *Bioinformation* 3: 282–283 Website http://wsmartins.net/websat [accessed 12 December 2011].

MILLER, M. P. 1997. TFPGA: Tools for population genetic analyses; version 1.3. Northern Arizona University, Flagstaff, Arizona, USA.

PARK, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. Ph.D. thesis, University of Dublin, Dublin, Ireland.

RANA, S. V., R. PAL, K. VAIPHEI, S. K. SHARMA, AND R. P. OLA. 2011. Garlic in health and disease. *Nutrition Research Reviews* 24: 60–71.

ROZEN, S., AND H. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Bioinformatics: Methods and protocols. Humana Press Inc., Totowa, New Jersey, USA.

SANGUINETTI, C. J., E. DIAS NERO, AND A. J. G. SIMPSON. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *BioTechniques* 17: 915–919.

ZEWDIE, Y., M. J. HAVEY, J. P. PRINCE, AND M. M. JENDEREK. 2005. The first genetic linkages among expressed regions of the garlic genome. *Journal of the American Society for Horticultural Science* 130: 569–574.

APPENDIX 1. Identification of 75 *Allium sativum* accessions from "Luiz de Queiroz" College of Agriculture (ESALQ), Agronomic Institute from Campinas (IAC), and Embrapa Vegetable Crops (Embrapa Hortaliças) germplasm.

ESALQ: Crespo and São José

IAC: BGH-0525, BGH-4823, BGH-5947, BGH-5963, Canela de Ema, Chinês I-4653, Sacaia Goiânia, Roxinho I-5063, Formosa I-4713, and Peruano Bisão ESALQ

Embrapa Hortaliças: Araguari, Barbado, Branco Dourado, Branco Mineiro, Catiguá, Chileno (PR), Caturra, Centralina A, Chinês Real, Cuiabá, Dourado, Gigante de Inconfidentes, Gigante de Lavínia, Gigante Roxão, Gravatá, Hozan, Inconfidentes II, Inhumas A, Inhumas Casca Roxa, Jacobina, Jundiaí, Juiz de Fora, Juréia, Mexicano, Mexicano A, Mexicano B, Morano Arequipeno, Mucugê, Novo Cruzeiro, Paraíba III, Peruano, Pinheiral, Piracicabano Amaralino, Roxo Dourado, Santa Izabel, Tempero de Bode, Ugarte, DDR6024, DDR6804, DDR6807, DDR6811, DDR6822, PI38383, PI540318, PI540351, RAI27, RAI41, RAI75, RAI159, RAI751, RE PSK, RE6820, UO73, UO74, UO79-3, UO94-5, UO94-11, WE8407, WE10735, WE12832, Chonan, Quitéria, and Ito