Molecular Ecology Notes (2005) **5**, 745–748 doi: 10.1111/j.1471-8286.2005.01050.x

PRIMER NOTE

Isolation and characterization of microsatellite loci from *Psidium guajava* L.

A. M. RISTERUCCI,* M. F. DUVAL,* W. ROHDE† and N. BILLOTTE*

**CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), UMR 1096 Polymorphismes d'Intérêt Agronomique, TA 40/03 Avenue Agropolis, 34398 Montpellier Cedex 5, France, †MPIZ (Max-Planck-Institut für Zuechtungsforschung), Carl-von-Linne-Weg 10 D 50829 Köln, Germany*

Abstract

A $(GA)_n$ and $(GT)_n$ microsatellite-enriched library was constructed and 23 nuclear simple sequence repeat (SSR) loci were characterized in the guava species (*Psidium guajava* L.). All SSR loci were found to be polymorphic after screening for diversity in different cultivars, and across-taxa amplification tests showed the potential transferability of most SSR markers in three other *Psidium* species. First to be published for *P. guajava*, this new SSR resource will be a powerful tool for genetic studies of guava, including cultivars identification and linkage mapping, as well as potentially for interspecific genetic studies within the genus *Psidium*.

Keywords: guava, microsatellite, *Psidium guajava*

Received 3 March 2005; revision accepted 20 April 2005

Guava (*Psidium guajava* L.; $2n = 22$) is native to northern South America and widely distributed in the tropical regions of America. Although considered minor in terms of commercial world trade, the economic importance of guava is related to the use of its fruit that contains high amounts of vitamins A and C, dietary fibre and calcium, to its multiple derived products such as like juice, cream, marmalade, desserts, etc. (Nakasone & Paull 1998) and to its pharmaceutical uses. Up to now, only dominant polymerase chain reaction (PCR)-based marker technologies such as random amplified polymorphic DNA (RAPD) (Prakash *et al*. 2002) were applied to study the guava molecular genetic diversity. The simple sequence repeat (SSR) or microsatellite codominant technique, which has proven its advantages and suitability in a large range of applications in genetics, was developed in order to improve the availability of best performing molecular tools for genetic studies and further marker-assisted breeding in guava and its close related species.

Total genomic DNA was extracted from freeze-dried leaf sample and then purified on anion-exchange microcolumns (Risterucci *et al*. 2000). DNA was restricted with *Rsa*I (Invitrogen) then used to construct a $(GA)_n$ and $(GT)_n$ microsatellite-enriched library following the protocol of Billotte *et al*. (1999). The enriched microsatellite fragments were cloned into pGEM-T (Promega) as indicated by the supplier and were used to transform Epicurian Coli XL1-Blue MRF' supercompetent cells (Stratagene). In all, 192 white transformant clones were then PCR-amplified for 35 cycles [94 °C (30 s), 52 °C (45 s), 72 °C (1 min 30 s); and a final elongation at 72 °C (8 min)] with the adaptator primer. PCR products were transferred onto Hybond-N+ nylon membranes (Amersham), which were hybridized at 56 °C with [γ^{32}P] dATP end-labelled $(GA)_{15}$ and $(GT)_{15}$. Among these clones 60% gave a satisfactory positive signal. Sequencing was performed on 48 clones using the universal T7 and SP6 sequencing primers and the BigDye Terminator version 3.1 Cycle Sequencing kit (Applied Biosystems) and run on an Applied Biosystems 3730 XL™ DNA analyser (Genome Express). Plasmidic and adaptor flanking sequences were eliminated and the presence of SSRs was detected using the MICROSAT software (CIRAD). A single redundant sequence was detected with SEQUENCHER version 4.0.5 (Gene Codes Corporation). Unique DNA sequences were analysed with BIOEDIT version 5.0.9 (Hall 1999). PCR primer pairs for microsatellite amplification were designed from a subset of 27 appropriate sequences using OLIGO 4.06 (National Biosciences), among which, 23 amplified the expected size fragment. PCR amplification was performed on 16 genomic DNA samples of *P. guajava* from various

Correspondence: A. M. Risterucci, Fax: (+33) 46761 5605; E-mail: ange.risterucci@cirad.fr

origins (Cameroon, Colombia, Cuba, Florida, Hawaii and Martinique) as well as on a genomic DNA sample of each of the related species *Psidium acutangulum* D.C., *Psidium cattleianum* Sabine var. *lucidum* and *Psidium friedrichsthalianum* (O. Berg). The primers were end-labelled with γ-[^{33}P]-ATP and the PCR amplification was performed in a MJ Research PTC-100™ thermocycler in a 20-μL final volume of buffer [10 mM Tris-HCl (pH 8), 50 mM KCl, 0.01% w/v gelatin and 1.5 mM $MgCl_2$] containing 20 ng genomic DNA, 0.2 μM for each primer, 200 μM dNTP and 0.5 U *Taq* DNA polymerase (Qbiogen). The PCR programme was: initial denaturation at 94 °C for 4 min; 30 cycles at 94 °C for 45 s, 55 °C for 60 s and 72 °C for 60 s; and a final elongation step at 72 °C for 8 min. PCR products of each individual were subjected to electrophoresis according to Risterucci *et al*. (2000).

All SSR primer pairs amplified successfully in *P. guajava*, with an average total number of 4.5 alleles per locus and no more than two displayed bands (nuclear SSR loci) (Table 1). Analysis was performed using GDA software (Lewis & Zaykin 2001). Despite a limited sampling of individuals in this analysis, observed (H_O) and expected (H_E) heterozygosities according to Nei (1973) revealed rather high average polymorphism rates of 0.42 and 0.62, respectively, promoting the utility of the developed SSRs for population genetic studies and suggesting they will provide good discriminating power for cultivar identification in guava. Significant deviations from Hardy–Weinberg equilibrium (HWE) were recorded for eight loci. Linkage disequilibria (LD) were calculated using the composite LD option and not assuming HWE, with a Bonferroni correction. Only one significant LD was found between loci mPgCIR016 and mPgCIR019.

Cross-species amplification tests on the other *Psidium* species revealed clear SSR patterns (Table 2), except for four loci without any product, possibly because of a mutational polymorphism at an annealing site. Not less than 21, 16 and 12 primers pairs were PCR-functional, respectively, in

Table 1 Synopsis of characteristics of 23 nuclear SSR loci isolated from *Psidium guajava*

SSR locus	EMBL Accession no.	Repeat motif	Clone size (bp)	Primer sequences (5′–3′)		Characteristics in *P. guajava*†			
				Forward	Reverse	No. of alleles	Allele size range	H_O‡	H_E§
mPgCIR01	AJ639775	$(GA)_{17}$	237	TAGTGCTTTGGTTGCTT	GCAGGTGGATATAAGGTC	3	236–250	0.167	0.542*
mPgCIR02	AJ639753	$(GA)_{20}$	224	AGTGAACGACTGAAGACC	ATTACACATTCAGCCACTT	3	202–230	0.417	0.448
mPgCIR03	AJ639754	$(GA)_{40}$	158	TTGTGGCTTGATTTCC	TCGTTTAGAGGACATTTCT	6	118–164	0.273	0.773*
mPgCIR04	AJ639755	$(GA)_{25}$	148	TTCAGGGTCTATGGCTAC	CAACAAGATACAGCGAACT	4	126–150	0.333	0.552
mPgCIR05	AJ639756	$(GA)_{31}$	252	GCCTTTGAACCACATC	TCAATACGAGAGGCAATA	5	224–280	0.727	0.756
mPgCIR07	AJ639757	$(CA)_{13}AA(GAA)_3$	149	ATGGAGGTAGGTTGATG	CGTAGTAATCGAAGAAATG	5	148–160	0.500	0.639
mPgCIR08	AJ639758	$(GA)_{12}$	214	ACTTTCGGTCTCAACAAG	AGGCTTCCTACAAAAGTG	4	210–224	0.500	0.681
mPgCIR09	AJ639759	$(GA)_{19}$	173	GCGTGTCGTATTGTTTC	ATTTTCTTCTGCCTTGTC	5	156–176	0.417	0.691*
mPgCIR10	AJ639760	$(CT)_{12}$	261	GTTGGCTCTTATTTTGGT	GCCCCATATCTAGGAAG	5	262–320	0.417	0.646
mPgCIR11	AJ639761	$(CT)_{17}$	298	TGAAAGACAACAAACGAG	TTACACCCACCTAAATAAGA	5	298–314	0.250	0.552
mPgCIR13	AJ639762	$AC)_{12}(AT)_4G(GA)_2$	245	CCTTTTTCCCGACCATTACA	TCGCACTGAGATTTTGTGCT	5	240–260	0.750	0.642
mPgCIR14	AJ639763	$(GA)_{11}$	185	TAAACACAACAAGGGTCA	CAGTTTTCATATCGTCCTC	2	184–186	0.417	0.330
mPgCIR15	AJ639764	$(GA)_8GG(GA)_9$	147	TCTAATCCCCTGAGTTTC	CCGATCATCTCTTTCTTT	5	144–172	0.333	0.531*
mPgCIR16	AJ639765	$(TC)_{25}$	292	AATACCAGCAACACCAA	CATCCGTCTCTAAACCTC	7	268–296	0.667	0.806
mPgCIR17	AJ639766	$(CT)_{23}$	231	CCTTTCGTCATATTCACTT	CATTGGATGGTTGACAT	4	230–240	0.417	0.691*
mPgCIR18	AJ639767	$(GA)_{23}$	195	TAAGCTGCATGTGTGC	ATGGCTTTGGATGAAA	3	192–204	0.500	0.538
mPgCIR19	AJ639768	$(CT)_{16}$	274	AAAATCCTGAAGACGAAC	TATCAGAGGCTTGCATTA	5	258–280	0.417	0.712
mPgCIR20	AJ639769	$(CT)_{14}(CA)_{17}$	266	TATACCACACGCTGAAAC	TTCCCCATAAACATCTCT	4	270–298	0.333	0.521
mPgCIR21	AJ639770	$(AG)_{15}GG(AG)_7$	154	TGCCCTTCTAAGTATAACAG	AGCTACAAACCTTCCTAAA	4	150–164	0.500	0.722*
mPgCIR22	AJ639771	$(GT)_9(GA)_{14}$	235	CATAAGGACATTTGAGGAA	AATAAGAAAGCGAGCAGA	5	236–252	0.333	0.649*
mPgCIR23	AJ639772	$(TA)_4(GT)_7$	185	GTCTATACCTAATGCTCTGG	CCCAGGAAAATCTATCAC	4	184–198	0.417	0.521
mPgCIR25	AJ639773	$(GA)_{24}$	124	GACAATCCAATCTCACTTT	TGTGTCAAGCATACCTTC	6	104–130	0.500	0.764
mPgCIR26	AJ639774	$(GT)_2(GA)_{17}$	185	CTACCAAGGAGATAGCAAG	GAAATGGAGACTTTGGAG	7	180–198	0.167	0.667*

Guava microsatellite loci were coded according to CIRAD's standard: mPgCIRxxx; m corresponds to microsatellite, Pg to *P. guajava* and CIR to CIRAD.

†Geographic origin of *P. guajava* individuals analysed: Cameroon (1), Martinique (1), Hawaii (3); Florida (6); Colombia (1); Cuba (4). ‡H_O, observed heterozygosity = percentage of individuals which are heterozygous on the locus; §H_E, expected heterozygosity according to Nei (1973).

*Indicates significant deviation from Hardy–Weinberg equilibrium; $P < 0.05$.

Table 2 Number of alleles observed in some *Psidium* genera using primers developed for *Psidium guajava*

	P. acutangulum (1 accession)		*P. cattleianum* (2 accessions)		*P. friedrischsthalianum* (1 accession)	
SSR locus	Allele number	Allele size range	Allele number	Allele size range	Allele number	Allele size range
mPgCIR01	1 NS	236	1 NS	236	1 NS	238
mPgCIR02	1 NS	202	1 NS	202	2/1 NS	204–210
mPgCIR03	—		—		—	
mPgCIR04	1 NS	150	1	164	—	
mPgCIR05	—		—		—	
mPgCIR07	2 NS	152–160	—		1 NS	148
mPgCIR08	1 NS	214	—		—	
mPgCIR09	1	168	—		—	
mPgCIR10	2 NS	262–320	2	242–250	1	258
mPgCIR11	2 NS	302–312	5/2 NS	300–340	2/1 NS	306–312
mPgCIR13	1 NS	250	7/1 NS	206–240	2	204–210
mPgCIR14	2 NS	184–186	—		—	
mPgCIR15	2	150–162	1	114	1	122
mPgCIR16	2 NS	292	1 NS	280	—	
mPgCIR17	1 NS	238	3	208–254	2	206–242
mPgCIR18	1 NS	200	4	174–196	2	206–216
mPgCIR19	1 NS	280	1	274	—	
mPgCIR20	1 NS	288	2	260–294	—	
mPgCIR21	2 NS	150–162	7/1 NS	136–178	1	126
mPgCIR22	1 NS	244	1	220	—	
mPgCIR23	1 NS	198	3/2 NS	176–202	1 NS	184
mPgCIR25	2 NS	122–130	—		—	
mPgCIR26	1 NS	184	2 NS	182–190	2/1 NS	176–188

Number of alleles per accession may exceed two for *P. cattleianum* ($2n = 6x$ or $2n = 8x$). Number of accessions screened is under the species name. NS indicates nonspecific amplification product (allele) and a dash indicates no amplification.

P. acutangulum, *P. cattleianum* var. *lucidum* and *P. friedrichsthalianum*. One or two alleles per locus were produced in *P. acutangulum* ($2n = 44$) and in *P. friedrichsthalianum* ($2n = 44$). The *P. cattleianum* Sabine var. *lucidum* specimen displayed higher allele numbers per locus, up to six, in accordance with the high polyploid status of this species, which has been estimated to be hexaploid or octaploid (Atchison 1947; Hirano & Nakasone 1969).

Our $(GA)_n$ and $(GT)_n$ microsatellite-enriched library and derived SSR markers are the first to be published in *P. guajava* and are available for cultivars identification, pedigree analysis, germplasm diversity as well as genetic mapping studies. SSRs are putative transferable markers across *Psidium* taxa and so a potentially useful molecular resource for any researcher studying the genetics of the genus *Psidium*. SSR markers will be better investigated in the future based on a larger sample of the whole genetic diversity of guava.

Acknowledgements

We gratefully acknowledge Daniel Ducelier (Cirad Martinique), Patrick Fournier and Fabrice Le Bellec (Cirad Guadeloupe) and Frédéric Normand (Cirad Réunion) for providing foliar samples for this study. We also thank Dieter Becker (MPIZ Germany) for his help in extracting the DNA sample used to build the SSR-enriched library.

References

Atchison E (1947) Chromosome numbers in the Myrtaceae. *American Journal of Botany*, **34**, 159–164.

Billotte N, Lagoda PJL, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, **54** (4), 277–288.

Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hirano RT, Nakasone HY (1969) Chromosome numbers of ten species and clones in the genus *Psidium*. *Journal of the American Society for Horticultural Science*, **94**, 287–289.

Lewis PO, Zaykin D (2001) *GENETIC DATA ANALYSIS: Computer Program for the Analysis of Allelic Data, Version 1.0 (d16c)*. Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

Nakasone HY, Paull RE (1998) *Tropical Fruits*. CAB International, Wallingford, UK.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA*, **70**, 3321–3323.

Prakash DP, Narayanaswamy P, Suresh N, Sondur SN (2002) Analysis of molecular diversity in guava using RAPD markers. *Journal of Horticultural Science and Biotechnology*, **77** (3), 287–293.

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics*, **101**, 948–955.