BIOLOGIA PLANTARUM 53 (1): 125-128, 2009

BRIEF COMMUNICATION

RAPD markers associated with quercetin accumulation in *Psidium guajava*

I.A. FERIA-ROMERO*[1], H. ASTUDILLO-DE LA VEGA**, M.A. CHAVEZ-SOTO**, E. RIVERA-ARCE**, M. LÓPEZ***, H. SERRANO*** and X. LOZOYA**

*Doctorado en Ciencias Biológicas, UAM, 84960, México, D.F., México**
*Centro Medico Nacional Siglo XXI, IMSS, 06720, México, D.F., México***
*Universidad Autonoma Metropolitana, 09340, México, D.F., México****

Abstract

We used a random amplified polymorphic DNA (RAPD) amplification method to identify molecular markers associated with high quercetin accumulation in the leaves of *Psidium guajava* L. trees, selected from four different Mexican agronomic regions. We identified six polymorphic RAPD fragments of 620, 590, 370, 690, 480 and 460 bp among individuals of *P. guajava*. Genetic linkage disequilibrium analysis revealed that three RAPD profiles considered as DNA markers (620/590 bp, 370 bp and 480/460 bp) had a positive, direct association with quercetin content. These informative molecular markers can be used for selective identification of plants with the highest accumulation of flavonoids.

Additional key words: guava leaves, DNA profile, linkage disequilibrium, flavonoid accumulation, QTL.

_____

Psidium guajava L. (*Mirtaceae*) originated in Mesoamerica is now widely distributed in many tropical and subtropical areas of the world. The fruit is extensively consumed either fresh or processed and the leaf extract is used in traditional herbal medicine for gastrointestinal disorder treatment (Lutterodt 1989) due to the presence of flavonols (quercetine glycosides) (Lozoya *et al.* 1994). In addition, quercetin glycosides and other flavonoids in guava leaves, take part in gene regulation and growth (Lozoya *et al.* 2002), and their accumulation is related to the geography, phenology of plant, and UV incidence (Havsteen 2002). Individuals with high capacity in flavonoid accumulation could be distinguished by their chemical analysis that is expensive and time consuming (Vargas-Alvarez *et al.* 2006). Beside the current knowledge of the genes related to flavonoid production, there are not conclusive data of a particular gene or group of genes related to quercetin glycosides accumulation (Hagsteen 2002). From this perspective, indirect, non-related indicators or markers are needed. Random amplified polymorphic DNA (RAPD) markers have been used for the assessment of genetic relationship among cultivars, to estimate chemical and genetic diversity between varieties, and to evaluate their correlation with active compounds in medicinal species (Mohapatra and Rout 2006, Padmesh *et al.* 2006, Narasimhan *et al.* 2006). RAPD markers have been used in guava characterization: Prakash *et al.* (2002) estimate the molecular diversity of 41 genotypes from different regions of India and Padilla-Ramírez *et al.* (2002) use them to evaluate the germplasm through fruit characteristics in a very narrow region of Mexico. To the best of our knowledge, there are no reports dealing with the study of genetic markers and flavonoid accumulation in *P. guajava* that could be used as an auxiliary or complementary method for quality control in plant drug production. The aim of this study was to find a group of DNA markers that could identify those *P. guajava* individuals with the highest quercetin content, using RAPD analysis in leaf samples from four different regions of Mexico with variable levels of guava fruit production.

_____

Received 6 February 2007, *accepted* 28 August 2007.
Abbreviations: QTL - quantitative trait loci; RAPD - random amplified polymorphic DNA.
Acknowledgements: The authors would like to thank Silvia Arzate-Trujillo and Akemi Sons-Matzumoto for technical assistance, as well as the people involved in the maintenance and processing of the *P. guajava L.* cultivars in the four sampled regions. I.A. Feria-Romero is under the Doctorado en Ciencias Biologicas, UAM. This work was supported by an IMSS-FOFOI FP-2003/039 grant.
[1] Corresponding author present address: M.Sc. Unidad de Investigación en Enfermedades Neurológicas, Centro Médico Nacional Siglo XXI, IMSS. Avenida Cuauhtémoc 330, Col. Doctores CP 06720, México, D.F., México; fax: (+525) 55780240, e-mail: irisferi@yahoo.com.mx.

Leaves collected from forty *P. guajava* individuals grown in four different Mexican regions (Calvillo, Aguascalientes; Coatepec H., Estado de Mexico; Jalpan, Queretaro; Zitacuaro, Michoacan) were identified according to Rivera-Arce *et al.* (2003), air-dried at 40 ºC in a forced-ventilation oven and powdered. One g of powder was extracted in a *Soxhlet* apparatus with 25 cm^3 of methanol for 3 h; then, the solvent was evaporated in vacuum and the residue was recovered. The extracted quercetin glycosides were hydrolyzed by 20 cm^3 of 0.5 M hydrochloric acid in methanol for 1 h; afterwards, methanol was added for a final volume of 25 cm^3 and a 1 cm^3 aliquot was diluted 1:10 with the same solvent. This solution was filtered through a 0.45 μm filter for HPLC analysis on a reverse-phase C_{18} column of 4.6 × 150 mm (*Varian*, Palo Alto, CA, USA) in a *Waters 2796* system (*Waters*, Milford, MA, USA), equipped with a *Waters 996* PDA detector maintained at 35 ºC and by using a 20 - 80 % increasing acetonitrile gradient in 10 mM phosphoric acid. The flow rate was kept constant at 0.6 cm^3 min^{-1} for a total running time of 35 min. All the operation conditions, data acquisition and analysis were controlled by *Millennium* software (*Waters*). Quercetin peak was compared to a commercial standard one (*Sigma-Aldrich*, St. Louis, MO, USA), based on retention time and quantified by interpolation on a calibration curve, as described by Vargas-Alvarez *et al.* (2006).

Young and disease-free leaves were washed in 70 % ethanol for 2 min and sterile, distilled water for 2 min, in sterile containers. The clean leaves were stored at -70 °C until use. A 150 mg leaf sample was cut into small pieces and ground to a fine powder under liquid nitrogen with a mortar and pestle. DNA was extracted from the sample using the *DNeasy Plant Mini Kit* (*Qiagen*, Hilden, Germany) according to the manufacturer's instructions. The extracted DNA was quantified and checked for purity and stored at -20 ºC until use. The DNA solutions were adjusted to a final concentration of 15 μg cm^{-3} prior to their use in the RAPD reactions.

Six 10-mer commercial primers (*Amersham Biosciences*, Piscataway, NJ, USA) were used for RAPD-polymerase chain reaction (PCR) amplification. Template DNA (15 ng) was amplified in a 0.025 cm^3 total PCR reaction including one RAPD analysis bead (*Amersham Biosciences*) to guarantee the reproducibility in the samples and 25 pmol of each primer. DNA amplification was performed using a *GeneAmp 9700* PCR thermocycler (*PE Biosystems*, Foster City, CA, USA) programmed for 1 step of 5 min at 94 °C followed by 45 cycles of 1 min at 94 °C, 1 min at 36 °C and 2 min at 72 °C and a final 10 min, 72 °C extension. To asses that RAPD amplification was reproducible, at least two individual experiments were performed with no significant differences (data not shown).

Amplified RAPD fragments were separated on a 2 % agarose gel stained with ethidium bromide, visualized under UV fluorescence, digitalized in a *BioImaging System* (*UVP*, Cambridge, UK) and evaluated with the *LabWorks v1.1.27* software (*UVP*).

One-way analysis of variance was used for statistical evaluation of total quercetin accumulation in the samples collected. Statistical evaluation of RAPD profiles and their correlation with quercetin accumulation was obtained by using the *SNPanalyzer ISTECH* web service facility, available at http://www.istech.info/istech/board/login_form.jsp. This software takes into account the frequency of a determined character or band and a specific chemical trait and returns a numerical value that is the statistical correlation for each pair of data. For this purpose, the banding patterns were classified according to polymorphic variations in amplified RAPD fragments, if they differed in at least two samples with the same primer. We consider that the RAPD profiles have an allelic behavior when they share some bands at the same position on the agarose gel. The homozygote showed a simple band, while heterozygote showed one, two or three bands for the same primers used with different samples. An estimation of major allele frequencies was obtained by χ^2 statistical and their correlation with quercetin accumulation trait, as a genetic linkage disequilibrium (LD) was obtained according to Lewontin (1964, 1995). Additionally, each banding pattern was associated with the quercetin concentration as a phenotypic trait. The association of RAPD profiles generated from each sample with different primers with quercetin accumulation was obtained through the expectation-maximization (E-M) algorithm of the *SNPanalyzer* software (Excoffier and Slatkin 1995).

Ten mature trees from each area were randomly selected and evaluated for their quercetin content during different seasons throughout the year. Significant differences were observed in quercetin accumulation throughout seasons in some places. Whereas Queretaro and Michoacan did not have a significant season variation (5.57 ± 0.95 and 2.32 ± 0.43 mg g^{-1}, respectively), cultivars from Aguascalientes showed the highest quercetin content of 9.34 ± 1.14 mg g^{-1} followed by Estado de Mexico with 8.51 ± 1.94 mg g^{-1} in spring and the lowest in autumn (5.52 ± 2.02 and 5.53 ± 1.91 mg g^{-1}, respectively), showing a significant difference between seasons ($P < 0.001$). Interestingly, from all the collected trees, one Estado de Mexico individual analyzed in spring was the highest producer, with a quercetin content of 12.62 mg g^{-1} (data not shown). Our results are in agreement with those of Vargas-Alvarez *et al.* (2006), who stated that May and July 2000 were the peak times for the flavonoids (myricetin, quercetin, luteolin, and kaempferol) accumulation in mature guava leaves in the geographic regions included in this study.

RAPD analysis is a molecular biological tool that allows an easy evaluation of the genome constitution of different plant samples, and the reproducibility of the results depends on the standard method and the highest quality of reagents employed. The six decamer primers that were used generated a total of 91 RAPD fragments, resulting in different amplification fragment patterns

Table 1. The sequences of primers used for RAPD analysis and a summary of the frequency and linkage disequilibriums of the banding patterns obtained from *P. guajava* L. individuals using RAPD polymorphic fragments [bp]. The directionality of each nucleotide sequence is 5' to 3'. Polymorphic fragment analysis is indicated between parentheses. The major pattern band obtained in RAPD analysis correlating with linkage disequilibrium association is indicated by an asterisk.

Primer	Nucleotide sequence	Number of polymorphic fragments	Total number of fragments	RAPD profiles from polymorphic fragments	Frequency of major pattern band	Linkage disequilibrium association
1	GGTGCGGGAA	2 (590, 610)	22	1a (590)* 1b (610)	0.82	4c
2	GTTTCGCTCC	1 (370)	16	2a* 2b (370)	0.79	4c
3	GTAGACCCGT	1 (690)	15	3a* 3b (690)	0.98	4b
4	AAGAGCCCGT	2 (460, 480)	14	4a 4b (460) 4c (460/480)*	0.69	
5	AACGCGCAAC	0	12	5		
6	CCCGTCAGCA	0	12	6		

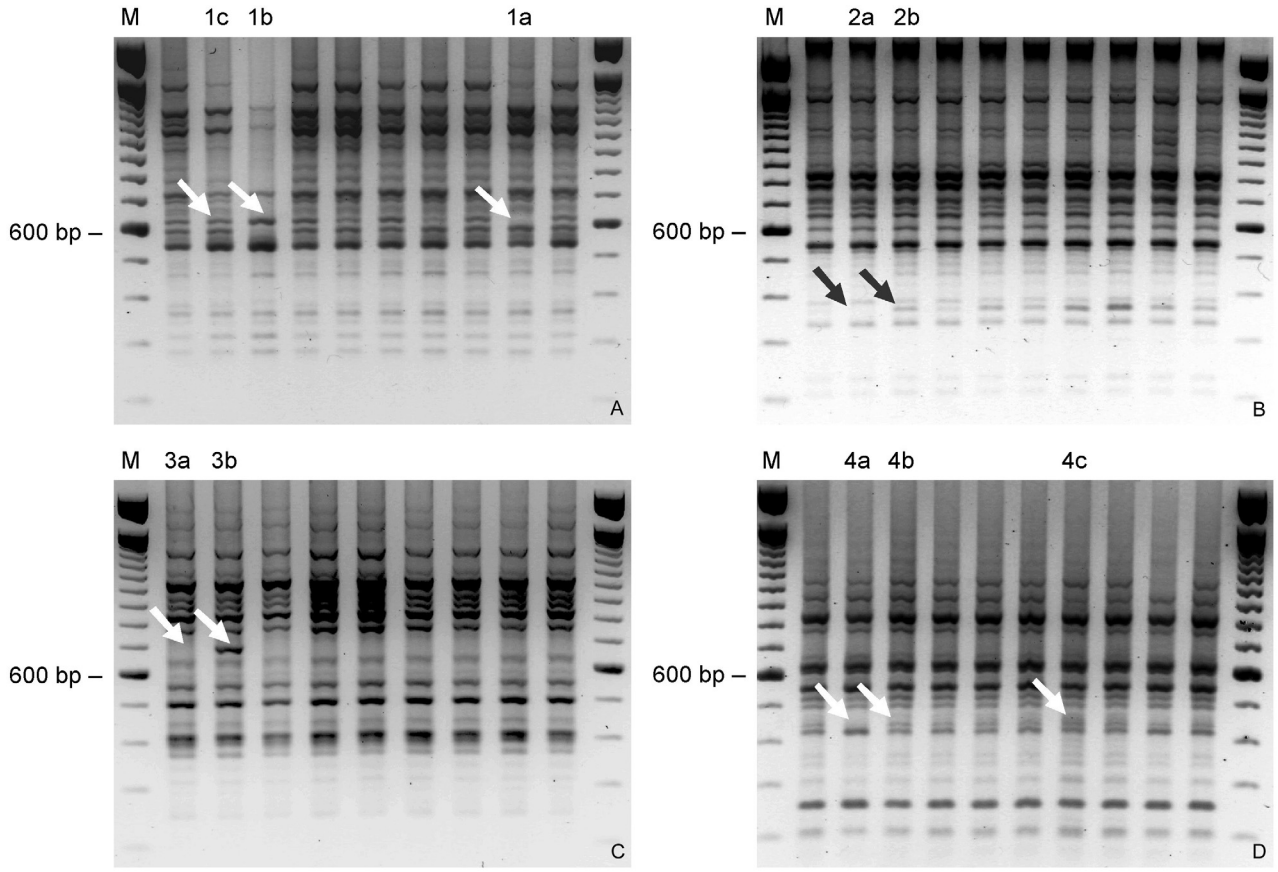

Fig. 1. RAPD profiles of *P. guajava* from different Mexican states. *A*: 1a, 1b and 1c represent polymorphic RAPD profiles using primer 1 for Michoacan individuals. *B*: 2a and 2b represent polymorphic RAPD profiles using primer 2 for Aguascalientes individuals. *C*: 3a and 3b represent polymorphic RAPD profiles using primer 3 for Queretaro individuals. *D*: 4a, 4b and 4c represent polymorphic RAPD profiles using primer 4 for Aguascalientes individuals. *Lane M*: 100 bp ladder. *Arrows* indicate informative fragments with allelic behavior.

(Table 1); sharing a 74 to 100 % by all the analyzed individuals, depending on the primer used. The validation of the RAPD technique was implemented by considering fragments from 200 to 1000 bp (Welsh and McClelland 1990, Williams *et al.* 1990), and band pattern reproducibility was determined in at least two individual

experiments, as indicated above. Also, the similarity obtained among all collected samples allowing the identification of polymorphic differences, was used as an indicator of reproducibility. Among the 91 generated fragments, only 6 were polymorphic in at least one individual, corresponding to primers 1, 2, 3 and 4 (*arrows* in Fig. 1). An allelic behavior upon RAPD amplification in these primers, along with the correlation between RAPD profile and a specific Mexican state, was observed. Pattern 1b was present only in Michoacan, patterns 1c and 2b in Aguascalientes and Michoacan, pattern 3b in Queretaro and Michoacan and pattern 4a in Aguascalientes, Estado de Mexico and Michoacan. The amplified fragments for primers 5 and 6 were shared by all individuals without polymorphic fragments.

In a preliminary correlation between band patterns and quercetin accumulation, we found two individuals containing $\geq$ 10.5 mg(quercetin) g^{-1}(leaf d.m.) collected in spring from Aguascalientes and Estado de Mexico, sharing the same banding patterns. A statistical analysis in all RAPD patterns with *SNPanalyzer* software and E-M algorithm allowed us to determine that 1a, 2a, 3a and 4c are the most informative combinations (MIC), due to the high frequencies of the 590 bp fragment (allele 1a) and 480 and 460 bp fragments (compound allele 4c); alleles 2a and 3a are the lacking of the 360 or 690 bp bands, respectively. This MIC refers to the correlation between DNA fragments and the quercetin accumulation. In order to explore the possible interaction among all these allele fragments, we evaluate their linkage disequilibrium (LD) association. The major LD for allele 4c was with 1a and 2a fragments, whereas for allele 4b, significant LD values were with the 3a allele, the lack of the 690 bp fragment (Table 1).

With the purpose of identifying a genetic factor that was involved in the variation of quercetin accumulation, we suggested a DNA molecular marker analysis to identify genomic characteristics related to quercetin variation. Many authors have treated the RAPD technique as equivalent to single genetic alleles in such a way that when population genetics statistics are used, there are relationships to the observed phenotype (Becerra and Paredes 2000, Padilla-Ramírez *et al*. 2002). According to our analysis, the frequency of fragments resulting in MIC, correlates to quercetin content, a dominant trait with a Lod score of 9.498 for a quantitative trait loci (QTL) relationship between the 1a and 2a with 4c banding patterns ($P < 0.05$), a significant high value that correlates with our phenotype trait.

In conclusion, we found clear and defined amplification fragments linked to quercetin glycoside accumulation in *P. guajava* individuals and demonstrated that, even when different individuals of the same region were analyzed, the presence of these amplified fragments correlated with the quercetin accumulation capability of a specific individual.

References

Becerra, V., Paredes, M.: [Use of biochemical and molecular markers in genetic diversity studies.] - Agr. tec. **60**: 270-281, 2000. [In Spanish]

Excoffier, L., Slatkin, M.: Maximum likelihood estimation of molecular haplotype frequencies in a diploid population. - Mol. Biol. Evol. **12**: 921-927, 1995.

Havsteen, B.H.: The biochemistry and medical significance of the flavonoids. - Pharmacol. therap. **96**: 67-202, 2002.

Lewontin, R.: The detection of linkage disequilibrium in molecular sequence data. - Genetics **140**: 377-388, 1995.

Lewontin, R.: The interaction of selection and linkage II. Optimum models. - Genetics **30**: 757-782, 1964.

Lozoya, X., Meckes, M., Abou-Zaid, M., Tortoriello, J., Nozzolillo, C., Thor, J.: Quercetin glycosides in *Psidium guajava* L. leaves and determination of a spasmolytic principle. - Arch. med. Res. **25**: 11-15, 1994.

Lozoya, X., Reyes-Morales, H., Chavez-Soto, M.A., Martinez-Garcia, M.C., Soto-Gonzalez, Y., Doubova, S.V.: Intestinal anti-spasmodic effect of a phytodrug of *Psidium guajava* folia in the treatment of acute diarrheic disease. - J. Ethnopharmacol. **83**: 19-24, 2002.

Lutterodt, G.D.: Inhibition of gastrointestinal release of acetylcholine by quercetin as a possible mode of action of *Psidium guajava* leaf extracts in the treatment of acute diarrhoeal disease. - J. Ethnopharmacol. **25**: 235-247, 1989.

Mohapatra, A., Rout, G.R.: Optimization of primer screening for evaluation of genetic relationship in rose cultivars. - Biol. Plant. **50**: 295-299, 2006.

Narasimhan, S., Padmesh, P., Nair, G.M.: Assessment of genetic diversity in *Coscinium fenestratum*. - Biol. Plant. **50**: 111-113, 2006.

Padilla-Ramírez J.S., González-Gaona E., Esquivel-Villagrana F., Mercado-Silva E., Hernández-Delgado S., Mayek-Pérez N.: [Characterization of outstanding guava germplasm from the Calvillo-Cañones region, México.] - Rev. Fitotec. Mex. **25**: 393-299, 2002. [In Spanish]

Padmesh, P., Reji, J.V., Jinish Dhar, M.: Estimation of genetic diversity in varieties of *Mucuna pruriens* using RAPD. - Biol Plant. **50**: 367-372, 2006.

Prakash, D.P., Narayanaswamy, P., Sondur, S.N.: Analysis of molecular diversity in guava using RAPD markers. - J. hort. Sci. Biotechnol. **77**: 287-293, 2002.

Rivera-Arce, E., Gattuso, M., Lozoya, X.: Anatomical identity parameters of the crude drug *Psidii guajavae-folium*. - Pharmaceut. Biol. **41**: 516-521, 2003.

Vargas-Alvarez, D., Soto-Hernández, M., González-Hernández, V.A., Engleman, E.M., Martínez-Garza, A.: [Kinetics of accumulation and distribution of flavonoids in guava (*Psidium guajava* L.).] - Agrociencia **40**: 109-115, 2006. [In Spanish]

Welsh, J., McClelland, M.: Fingerprinting genomes using PCR with arbitrary primers. - Nucleic Acids Res. **18**: 7213-7218, 1990.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V.: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. - Nucleic Acids Res. **18**: 6531-6535, 1990.