

Molecular Identification and Analysis of *Psidium guajava* L. from Indigenous Tribes of Taiwan

TSENG-WEI CHEN, CHANG-CHAI NG, CHUNG-YI WANG AND YUAN-TAY SHYU*

Department of Horticulture, National Taiwan University, 140, Sec. 4, Keelung Rd., Da-an District, Taipei City 106, Taiwan, R.O.C.

(Received: July 31, 2006; Accepted: November 29, 2006)

ABSTRACT

Psidium guajava L. is a perennial fruit tree in subtropical and tropical areas. In Taiwan, *P. guajava* has been used as anthro-pharmacological plants by aboriginal people to treat acute diarrhea, cough and intestinal spasmodic diseases. The classification and functional identification of *P. guajava* remains unsolved these days. In this study, molecular markers 18S rRNA, internal transcribe spacer (ITS) region of ribosomal DNA, *trnL* intron and *trnL-trnF* intergenic spacer (IGS) of chloroplast DNA (cpDNA), and random amplified polymorphic DNA (RAPD) were used for the molecular identification of 18 *P. guajava* samples from different indigenous tribes, 2 from non-indigenous tribe, and 12 commercial cultivars from markets in Taiwan. Molecular methods restricted fragment length polymorphism (RFLP) and denatured gradient gel electrophoresis (DGGE) are found time-consuming and less efficient as compared to RAPD, thus are not suitable for samples of high homology. In this study, 18S rDNA, ITS and cpDNA *trnL* intron and *trnL-trnK* intergenic spacer were also tested molecular marker; however, results analyzed by molecular algorithm UPGMA, Neighbor-Joining, Parsimony or Maximum likelihood showed no discriminations (data not shown). On the other side, ten 10-mer oligonucleotide primers were used in RAPD to amplify the specific genes from 32 guava samples. Four primers, OPB 17, OPG 6, OPY 15 and OPY 18, were able to direct the amplification and yielded a total of 82 polymorphic RAPD patterns. Thirty-two genotypes on the dendrogram were identified and were divided into two major groups, the uncultivated and commercial cultivars. Based on the cluster analysis, the red-flesh *Psidium* samples that were believed to have high medical function were grouped independently. The results suggest that RAPD is useful for the discrimination of uncultivated, cultivars and potential *Psidium* of high economy.

Key words: guava, *Psidium guajava*, RAPD, indigenous tribe

INTRODUCTION

Psidium guajava L., commonly named guava, is a perennial fruit tree in subtropical and tropical areas. It is native to South American countries and was introduced to India by the Portuguese during 17th century⁽¹⁾. *P. guajava* has high nutritional content and is especially rich in vitamin C. There are many varieties of uncultivated guava and imported guavas in Taiwan. Cultivated cultivars of *P. guajava* were introduced from India and America for quality improvement. There are many cultured and uncultured guava varieties including pearl guava, crystal guava, Thai guava, pear guava, and white, red and yellow flesh guava. Most of uncultured species are found in indigenous tribes of Taiwan. They were used as an effective remedy to treat and prevent diseases such as headache, cough⁽²⁻³⁾, spasm, inflammatory, pyrexia, acute diarrhea⁽⁴⁾, colic, flatulence, and gastric pain⁽⁵⁾.

Morphological traits are traditional phenotypic markers for the identification of plants. They may change with the cultivation and growth environment so that the identification is confusing. In order to identify red-flesh and white-flesh guava trees in indigenous tribes in a more systematic way, specific genetic markers for guavas are

developed. Recently, many molecular markers, such as restriction fragment length polymorphism (RFLP), amplified fragment-length polymorphism (AFLP), sequence-characterized amplified regions (SCAR), inter simple sequence repeat (ISSR), simple sequence repeat (SSR) and random amplified polymorphism DNA (RAPD), are used in horticultural crops research. The chloroplast DNA from tobacco often serves as the reference for plastid genomes⁽⁶⁾ and its complete nucleotide sequence and gene map were published in 1986⁽⁷⁾. Zhang⁽⁸⁾ established the phylogenetic relationships in *Carpha* by cladistic analyses based on chloroplast *trnL* intron and *trnL-trnF* intergenic spacer sequence data. RAPD markers have been used for cultivar identification and genetic diversity analysis among 25 *Feijoa sellowiana*⁽⁹⁾ cultivars and accessions in Italy and 41 genotypes of guava in India⁽¹⁰⁾. It makes the discrimination of cultivar easy, fast and inexpensive. In the present study, 18S rDNA, ITS and cpDNA *trnL* intron and *trnL-trnK* intergenic spacer and RAPD markers are used to identify 32 indigenous genotypes of guava in Taiwan. The study is aimed to understand the distribution of red- and white-flesh guava in indigenous tribes of Taiwan by molecular markers analyses. The preliminary results are useful in the discrimination of guava species. It is crucial to identify guavas, which may have potential to be developed into a medicinal plant.

* Author for correspondence. Tel: +886-2-33664850; Fax: +886-2-23661441; E-mail: tedshyu@ntu.edu.tw

MATERIALS AND METHODS

I. Plant Materials

The guava tree leaves collected from different indigenous tribes of Taiwan are listed in Table 1. Guava leaves were collected during from April to July, 2005. Fifteen to 20 leaves were collected from a guava tree in each indigenous tribe. Only the leaf in the sun shined position was collected.

II. Total Genomic DNA Extraction

Plant leaves were rinsed with distilled water, dried and stored at -80°C . Genomic DNA was extracted from the leaves by the cetyltrimethylammonium bromide (CTAB) method⁽¹¹⁾ with some modifications. Leaf samples were ground into powder with mortar and pestle in liquid nitrogen. Approximately 0.1 g of dried leaf powder was mixed with 1 mL of extraction buffer (2% CTAB, 20 mM ethylenediaminetetraacetic acid (EDTA), 100 mM tris/HCl, pH 8.0, 1.4 M NaCl). Mixture was centrifuged at $13,200 \times g$. The guava DNA precipitate was washed with 500 μL of 70% alcohol twice and a clean precipitate was obtained. The DNA was dissolved in 20 μL of sterile water after vacuum drying for 20 min and stored at -20°C for further use.

III. DNA Quantification

The DNA quantification was carried out using a spectrophotometer (Beckman, CoulterTM DU[®]640, USA). One microliter of each guava DNA extracts was diluted with 99 μL of deionized water and the absorbances at 260 and 280 nm were measured. The concentrations absorbances were calculated and expressed in ng/mL. The final concentrations of guava DNA stock solutions were adjusted to 100 ng/mL.

IV. Polymerase Chain Reaction

Specific DNA fragments were amplified was carried out by polymerase chain reaction (PCR). The total volume of reaction mixture was 25 μL that contained 0.5 mM primers, 1X buffer, 0.5 unit DNA polymerase (DyNAzyme TM II, FINNZYMES Inc., Riihitontuntie, Finland), 200 mM dNTP, and 200 ng genomic DNA. The cycles No. of reaction depends on the amplified regions of genomic DNA. The specific primers of 18S rDNA, ITS and cpDNA *trnL* intron and *trnL-trnF* intergenic spacer regions are listed in Table 2⁽¹²⁻¹³⁾. For 18S rDNA amplification, the PCR was programmed for 35 cycles of: 94°C for 30 sec, 55°C for 30 sec, and 72°C for 3 min with an initial denaturation step at 94°C for 3 min and an additional 7-min extension step at 72°C . For ITS gene amplification, the PCR was programmed for 35 cycles of: 95°C for 30 sec, 58°C for 30 sec, and 72°C for 1.5 min with an

initial denaturation at 95°C for 3 min and a final extension step at 72°C for 7 min. The PCR of *trnL* intron and *trnL-trnF* intergenic spacer gene amplification was carried out in 30 cycles of: 96°C for 1 sec, 54°C for 5 sec, and 72°C for 1 min with an initial denaturation step 3 min at 96°C and a final extension at 72°C for 10 min. For RAPD method, four primers (Table 3)⁽¹⁰⁾ directed the amplification of highly reproductive and the highest numbers of

Table 1. Commercial cultivars and plant materials from different indigenous tribes of Taiwan

Number	Name of tribes and economical cultivars	Source of selections (township / county)
7	Songhe	Heping / Taichung
11	Shilin	Taian / Miaoli
13	Chunghsing	Taian / Miaoli
15	Gangkau	Fengbin / Hualien
17	Iwan	Chengkung / Taitung
20	—	Sanchih / Taipei ^{a,d}
24	Hongye	Yanping / Taitung
28	Donghe	Yanping / Taitung
29	Jiafeng	Donghe / Taitung ^a
30	Balin	Fuxing / Taoyuan ^a
31	Kagil	Fuxing / Taoyuan
32	Tawan	Fuxing / Taoyuan
33	Fushan	Wulai / Tapei
34	Gangkau	Manchou / Pingtung
36	Manchou	Manchou / Pingtung ^a
37	Laiyi	Laiyi / Pingtung ^a
38	Saijia	Laiyi / Pingtung
39	Maer	Laiyi / Pingtung
41	Qinhe	Taoyuan / Kaohsiung
42	—	Renmei / Kaohsiung ^{a,d}
E1	Pearl guava	FTHESb
E2	Crystal guava	FTHES
E3	Red guava	FTHES
E4	White guava	FTHES
E5	20th century guava	FTHES
E6	<i>Psidium</i> 'Odorata'	FTHES
E7	Thai guava	FTHES
E8	G3-48	FTHES
E9	Seedless guava	FTHES
E10	Pear guava	FTHES
E11	Sao Tome guava	São Tomé and Príncipe ^c
E12	Chungshan moon guava	FTHES

^aNo. 20, 29, 30, 36, 37 and 42 belong to red-flesh guava tree.

^bFTHES: Fengshan Tropical Horticultural Experiment Station.

^cSão Tomé and Príncipe: The capital of The Democratic Republic of São Tomé and Príncipe.

^dNo. 20 and 42 are non-indigenous tribe guava tree.

Table 2. Primer sequences used for PCR amplifications and sequencing

	Region	Primer	Nucleotide Sequence (5' to 3')	Source
rDNA	18S	NS1 (F ^a)	GTA GTC ATA TGC TTG TCT C	White <i>et al.</i> ^b
		NS4 (R ^a)	CTT CCG TCA ATT CCT TTA AG	White <i>et al.</i> ^b
	ITS	ITS1 (F)	TCC GTA GGT GAA CCT GCG G	White <i>et al.</i> ^b
		ITS4 (R)	TCC TCC GCT TAT TGA TAT GC	White <i>et al.</i> ^b
cpDNA	<i>trnL</i> intron and	B49317 (F)	CGA AAT CGG TAG ACG CTA CG	Pierre ^c
	<i>trnL-trnF</i> IGS	A50272 (R)	ATT TGA ACT GGT GAC ACG AG	Pierre ^c

^aF: forward; R: reverse.

^bWhite *et al.*, 1990.

^cPierre, 1991.

Table 3. Ten 10-mer random primers used for molecular polymorphism analysis of *Psidium guajava* L

Primer	Nucleotide sequence (5' to 3')	No. of fragments amplified	Source
OPB 11	GTA GAC CCG T	—	Prakash <i>et al.</i> ^a
OPB 17 ^b	AGG GAA CGA G	24	Prakash <i>et al.</i> ^a
OPG 6 ^b	GTG CCT AAC C	22	Prakash <i>et al.</i> ^a
OPG 19	GTC AGG GCA A	—	Prakash <i>et al.</i> ^a
OPJ 1	CCC GGC ATA A	—	Prakash <i>et al.</i> ^a
OPY 14	GGT CGA TCT G	—	Prakash <i>et al.</i> ^a
OPY 15 ^b	AGT CGC CCT T	17	Prakash <i>et al.</i> ^a
OPY 18 ^b	GTG GAG TCA G	19	Prakash <i>et al.</i> ^a
OPY 19	TGA GGG TCC C	—	Prakash <i>et al.</i> ^a
OPY 20	AGC CGT GGA A	—	Prakash <i>et al.</i> ^a

^aPrakash *et al.*, 2002.

^bThe four best primer, OPB 17, OPG 6, OPY 15, OPY 18, which producing significant and producible polymorphic RAPD patents, were selected and used for final analysis.

diverse fragments. The PCR consisted of 35 cycles with initial denaturation at temperature 95°C for 3 min and final extension at temperature 72°C for 7 min. Each cycle included denaturation at 95°C for 30 sec, annealing at 30°C for 1 min and extension at 72°C for 2 min. Amplicons were resolved on a 2% agarose gels by electrophoresis at 100 V for 60 min.

X. DNA Sequencing

The PCR products of guava specimens were sequenced by Mission Biotech Co., Taiwan on a ABI PRISM 377-96 DNA- Sequencer, Perkin-Elmer, CA, USA.

XI. Cladistic Analysis

Each DNA amplification was repeated three times and the result of the bands on agarose gels were marked as present (1) or absent (0). The RAPD polymorphism was analyzed and expressed as a genetic dissimilarity matrix using NTSYS-pc (Numerical Taxonomy System, version 2.0, Exeter Software, NY, USA software). Dice similarity index $SD = 2N_{ab}/(N_a + N_b)^{(14)}$ was used to calculate

the pairwise similarity matrix, where N_{ab} indicates the number of shared bands between a pair of genotypes a and b, N_a means the number scored bands in genotype a, and N_b means the number of scored bands in genotype b. The similarity of genotypes was analyzed by unweighted pair-group method analysis (UPGMA), and the result of clusters analysis was demonstrated as a dendrogram. After sequencing the PCR products, Neighbor-joining method (NJ), Parsimony method (PA) and Maximum-likelihood algorithm (ML) were applied to the cluster analysis. The results were also expressed as dendrogram.

RESULTS AND DISCUSSION

The sequences of 18s rDNA, ITS, and cpDNA *trnL* and *trnL-trnF* intergenic spacer were determined and then processed by PA, NJ, and ML algorithms for the construction of dendrograms. The dendrogram of 18S rDNA produced by PA method indicated high similarity of these guava trees (Figure 1). Nine uncultivated guavas and 3 commercial cultivars were chosen for cluster analysis. Three cultivars were grouped into a cluster and they were

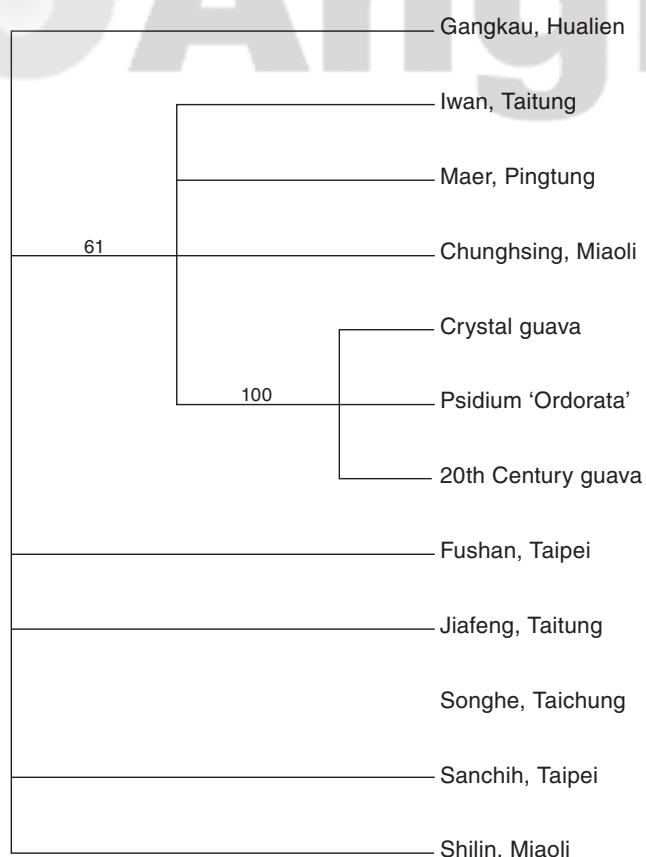


Figure 1. Phylogenetic relationship of 13 guava specimens based on applying Parsimony method (PA) for cluster similarity analysis of rDNA 18S region. The probability of commercial cultivars and sub-cluster are 100% and 61% respectively (Bootstrap = 1000).

also grouped with several uncultivated ones into another cluster with respective reoccurrence rates of 100% and 61%. The sequences of rDNA ITS and cpDNA *trnL* and *trnL-trnF* intergenic spacer were also determined and analyzed using PA, NJ, and ML. The dendrograms of these two regions were constructed and displayed low degree of discrimination (data not shown). It suggested that 18S rDNA and its ITS regions of guava were not suitable for intra-species phylogenetic analysis due to their highly conserved DNA sequences. Besides, cpDNA *trnL* and *trnL-trnF* intergenic spacer, that belongs to the non-coding region and are inherited through the maternal linkage, thus producing low degree of discrimination.

To enhance the polymorphism of genotypes of guava, ten 10-mer primers were used for DNA amplification. Only four primers (OPB17, OPG6, OPY15, and OPY18)⁽¹⁰⁾ were selected for fingerprinting. Since the amplification directed by these four primers produced significant polymorphism and yielded a total of 82 polymorphic RAPD patterns as shown in Figure 2. Genetic similarity matrix (Table 4) was generated using NTSYS-pc based on the marked scores of the polymorphic RAPD patterns where score 1 or 0 was assigned to the present or absent band. It is shown in genetic similarity matrix that the high-

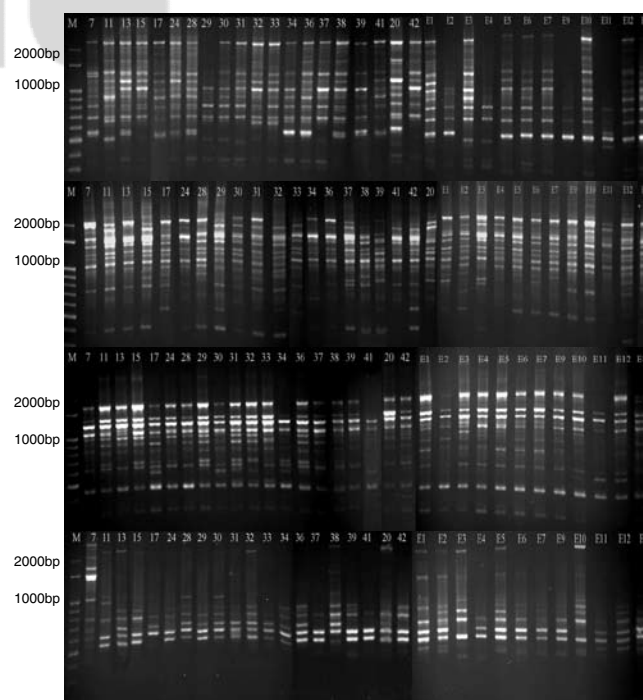


Figure 2. RAPD markers of 32 guava produced by (A) OPB 17, (B) OPG 6, (C) OPY 15, (D) OPY 18.

est genetic similarity is 87% between Chungshin guava and Jiafeng guava, and the least one is 33% among Sam Tome guava, Red guava and the guava collected from Gangkau, Hualien. The genetic similarity among uncultivated guavas or commercial cultivars is higher than that between uncultivated and commercial cultivars.

Thirty two genotypes on the dendrogram were distinguished and divided into two major groups as shown in Figure 3, in which their origins based on geographical locations of different genotypes are indicated. Uncultivated guavas collected from indigenous tribes were grouped under G1 cluster except for the guava collected from the tribe of Qinhe, Kaohsiung. Most commercial cultivars were grouped under G2 cluster. Poor reproducibility of RAPD markers may explain exclusion of Qinhe guava from cluster G1. According to Dai *et al.*⁽¹⁵⁾, the discrimination of *Lilium formosanum* and *L. longiflorum* could be achieved using four 10-mer primers, namely OPB17, OPG6, OPY15 and OPY18, with high reproducibility. Their result demonstrated that four arbitrary oligonucleotide 10-mers could direct the production of 86 reproducible bands and over 80% polymorphism was observed. In this study, two commercial cultivars grown in high altitude from Sam Tome & Principe, and Southern Africa were used as reference groups. Their grouping status, same as the guavas from the Qinhe tribe sampled from high altitude over 700 m in the mountain of Kaohsiung, was excluded from G1 and G2. This might indicate that RAPD is useful to differentiate samples from different geographical areas or various climates.

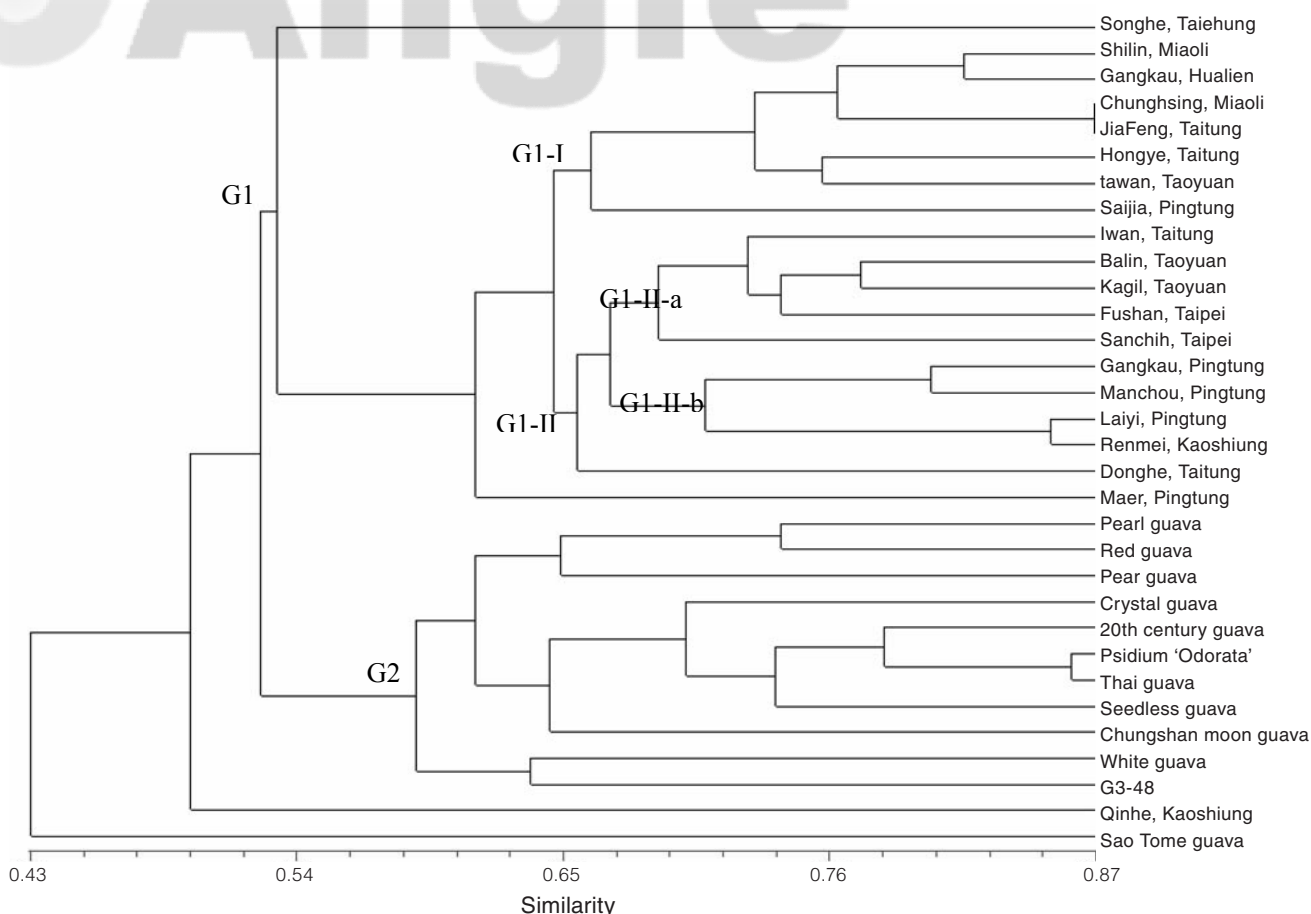


Figure 3. Dendrogram of 32 guava specimens based on cluster similarity analysis of RAPD markers.

ity of these two tribes. Most of guavas in G1-II-a grow at a higher altitude compared with guavas in G1-II-b. As to the subgroup G1-I guava were collected from tribes of altitude over 250 m, such as Hongye & Jiafeng, Taitung, Chunghsing & Shilin, Miaoli and Tawan, Taoyuan or from other area such as Gangkao, Hualien and Saijia, Pingtung. They all belong to white-flesh guava and share similar genotypes.

Molecular markers have been used as a tool to investigate the plant germplasm diversity recently. Banding patterns can be converted into informative data for pedigree analyses. The shortcoming of RAPD method is the reproducibility in amplification. In this study, the PCR reactions were performed in optimal conditions and informative RAPD fragments were obtained with high reproducibility. RAPD analysis is efficient and accurate for the investigation of distribution of commercial, red-flesh, white-flesh guava or uncultivated guavas. The RAPD analysis is useful in the fingerprinting of each guava sample. The geographical locations, growth altitude, and climates may contribute the polymorphic RAPD of guava trees in Taiwan. It is believed this result is beneficial for further research on the guava functionality as traditional remedies.

ACKNOWLEDGEMENTS

We would like to thank Department of Health, Taiwan, ROC. for funding and Council of Agriculture, Executive, Taiwan, ROC. for plant samples. We also greatly appreciate Hong-Ye Hsieh from Fengshan Tropical Horticultural Experiment Station, for his helps in the collection of commercial cultivars.

REFERENCES

1. Menzel, C. M. 1985. Guava: an exotic fruit with potential in Queensland. *Queensl. Agric. J.* 111: 93-98.
2. Khan, M. I. H. and Ahmad, J. 1985. A pharmacognostic study of *Psidium guajava* L. *Int. J. Crude Drug Res.* 23: 95-103.
3. Jaiarj, P., Khoohaswan, P., Wongkrajang, Y., Peungvicha, P., Suriyawong, P., Saraya, M. L. S. and Ruangsomboon, O. 1999. Anticough and antimicrobial activities of *Psidium guajava* Linn. Leaf extract. *J. Ethnopharm.* 67: 203-212.
4. Lin, J., Puckree, T. and Mvelase, T. P. 2002. Anti-diarrhoeal evaluation of some medicinal plants used by

- Zulu traditional healers. J. Ethnopharm. 79: 53-56.
5. Lozoya, X., Reyes-Morales, H., Chávez-Soto, M. A., Martínez-García, M. C., Soto-González, Y. and Doubova, S. V. 2002. Intestinal anti-spasmodic effect of a phytodrug of *Psidium guajava folia* in the treatment of acute diarrheic disease. J. Ethnopharm. 83: 19-24.
 6. Wakasugi, T., Sugita, M., Tsudzuki, T. and Sugiura, M. 1998. Updated gene map of tobacco chloroplast DNA. Plant Mol. Biol. Rep. 16: 231-241.
 7. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Sinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N. and Shimada, H. S. 1986. The complete nucleotide sequence of the tobacco chloroplast genome. Plant Mol. Biol. Rep. 4: 110-147.
 8. Zhang, X., Marchant, A., Wilson, K. L. and Bruhl, J. J. 2004. Phylogenetic relationships of *Cyperaceae* and its relatives (Schoeneaceae, Cyperaceae) inferred from chloroplast *trnL* intron and *trnL-trnF* intergenic spacer sequences. Mol. Phylo. Evol. 31: 647-657.
 9. Dettori, M. T. and Palombi, M. A. 2000. Identification of *Feijoa sellowiana* Berg accessions by RAPD markers. Sci. Hortic. 86: 279-290.
 10. Prakash, D. P., Narayanasway P. and Sonder S. N. 2002. Analysis of molecular diversity in guava using RAPD markers. J. Hortic. Sci. Biotech. 77: 287-293.
 11. Porebski, S. L., Bailey, G. and Baum, R. B. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharides and polyphenol compounds. Plant Mol. Biol. Rep. 5: 8-15.
 12. White, T. J., T. Bruns, S. Lee, and J. W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In "PCR Protocols: A Guide to Methods and Applications". pp. 315-322. Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J. eds. Academic Press. New York, U. S. A.
 13. Pierre, T., Ludovic, G., Guy, P. and Jean, B. 1991. Universal primers for amplification of three non-coding regions of chloroplast. Plant Mol. Bio. 17: 1105-1109.
 14. Wilde, J., Waugh, R. and Powell, W. 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor. Appl. Genet. 36: 871-877.
 15. Dai, T. E., Lee, W. L. and Shii, C. T. 2006. Studies on genetic variability of *Lilium formosanum* and *L. longiflorum* mediated morphological and molecular markers. J. Taiwan Soc. Hort. Sci. 52: 37-52.