Development and Application of RAPD-SCAR Marker for Identification of *Phyllanthus emblica* Linn.

Warude Dnyaneshwar, Chavan Preeti, Joshi Kalpana,* and Patwardhan Bhushan

Bioprospecting Laboratory, Interdisciplinary School of Health Sciences, University of Pune; Pune–411 007, India.
Received July 4, 2006; accepted August 17, 2006

Correct genotype identification of medicinal plant material remains important for botanical drug industry. Limitations of chemical and morphological approaches for authentication have generated need for newer methods in quality control of botanicals. The present study was carried out to develop DNA based marker for identification of *Phyllanthus emblica* Linn. A putative marker (1.1 kb) specific for *P. emblica* was identified by Random Amplified Polymorphic DNA (RAPD) technique. Sequence Characterized Amplified Region (SCAR) marker was developed from the RAPD amplicon. The SCAR marker was found useful for identification of *P. emblica* in its commercial samples and Triphalachurna, a multi-component Ayurvedic formulation.

Key words *Phyllanthus emblica*; Euphorbiaceae; random amplified polymorphic DNA (RAPD); sequence characterized amplified region (SCAR)

Phyllanthus emblica Linn. Syn: *Emblica officinalis* Gaertn. (Indian Gooseberry) family-Euphorbiaceae fruit is one of the top selling botanicals having diverse applications in healthcare, food and cosmetic industry. It has been well studied for immunomodulatory,[1] anticancer,[2] antioxidant[3] and antiulcer[4] activities. An official drug of Ayurvedic Pharmacopoeia[5] and Indian Herbal Pharmacopoeia,[6] it forms a main ingredient of various multi-component formulations. Correct genotype identification of the plant material, therefore, remains important for protection of both the public health and industry. Chemoprofiling and morphological evaluation are routinely used for identification of the botanical. Chemical complexity and lack of therapeutic marker(s) are some of the limitations associated with chemical approach while subjective bias in morphological evaluation limits its use.

Molecular biology offers various techniques that can be applied for plant identification.[7] Genetic polymorphism in medicinal plants has been widely studied which helps in distinguishing plants at inter- and/or intra-species level.[8,9] PCR-based methods including Randomly Amplified Polymorphic DNA (RAPD)[10] can be effectively used for authentication of the medicinal plant material. Further, development of more specific, sensitive and reproducible markers like RAPD based Sequence Characterized Amplified Region (SCAR) can increase industrial application of the molecular techniques. These markers have been used for authentication of medicinal plant species of ginseng,[11] Artemisia[12] and other commercially important timbers like bamboo.[13]

In the present study, we have chosen six *Phyllanthus* species (other than *P. emblica*) that are commercially and therapeutically important; and are widely used in Ayurvedic system of medicine for treating various ailments.[14] SCAR marker for correct genotype identification of *P. emblica* from a species-specific RAPD amplicon is developed by comparative analysis. The marker is further used for authentication of commercial samples of *P. emblica* fruit powders (Amlachurna) as well as Triphalachurna, a multi-component formulation which contains fruit powders of *P. emblica*, *Terminalia chebula* Rotz. and *Terminalia belerica* Roxb.

MATERIALS AND METHODS

Plant Material Fresh leaf tissues of *Phyllanthus distichus* Linn., *Phyllanthus reticulatus* Poir., *Phyllanthus urinaria* Linn., *Phyllanthus simplex* Retz., *Phyllanthus niruri* Linn. and *Phyllanthus indofischeri* Bennet. were collected from different parts of India (three samples of each species) and authenticated in Agharkar research Institute, Pune, India. Leaf tissues of eleven cultivars of *P. emblica viz*.; NA-06, NA-07, NA-10, Kanchan, Chakaiya, Francis, Banarasi, Hathizool, Dongri, Bansired and Anand-01 were collected from standard cultivated plots of Mahatma Phule Agricultural University, Rahuri, India. The collected fresh tissues were powdered in liquid nitrogen and stored in deep freezer (−70 °C) till further processing.

Three commercial *P. emblica* fruit powder samples (Amlachurna) were purchased from different Ayurvedic crude drug suppliers and Triphalachurna marketed formulations were collected from local pharmacy and stored at cool, dry place.

DNA Isolation DNA was isolated from fresh as well as commercial samples using modified cetyl trimethyl ammonium bromide (CTAB) extraction method.[15] In brief, fresh leaf tissue (1—1.5 g) or dried fruit powders (0.5 g) or Triphalachurna (1.0 g) was ground in liquid nitrogen with polyvinyl pyrrolidone (PVP). Freshly prepared extraction buffer (containing 0.3% [v/v] β-mercaptoethanol) was added to adjust the pH to 7.5—8. The suspension was incubated at 65 °C for 30 min with intermittent pH monitoring. After cooling at room temperature for 5 min an equal volume of chloroform–isoamyl alcohol (IAA) (24 : 1) was added and centrifuged at 665 ***g*** for 10 min at 25 °C. DNA from aqueous layer was precipitated by adding 1/10th volumes of 3 M sodium acetate and 2 volumes of chilled ethanol at 4 °C. The mixture was centrifuged at 665 ***g*** at 4 °C for 5 min. DNA pellet was washed with 70% (v/v) ethanol and dissolved in Tris-Cl-EDTA (pH 8.0) (TE) buffer after drying in SpeedVac (Savants, U.S.A.). DNA concentration was determined by taking absorbance at 260 nm, according to Sambrook and Russell (2001).[16]

RAPD-PCR PCR for amplifying the DNA preparations was carried out in a 25-μl vol of reaction mixture. A reaction tube contained 25 ng of DNA, 0.6 U of *Taq* DNA polymerase

* To whom correspondence should be addressed. e-mail: kalpana@unipune.ernet.in

Table 1. *P. emblica* Specific SCAR Primer Devised from Sequenced RAPD Amplicon PE01

RAPD Primer	SCAR primer	Number of base pairs (bp)	Sequence (5′–3′)	G+C content (%)	Annealing temperature
OPA-16	D1	23	CAG ATC TCG TGT AAA AAG CGT TG	44	55 °C
	D2	22	TGC AGT GAA TTC CAA GTG TTT C	41	

enzyme, 100 μM of each dNTP, 1X *Taq* DNA polymerase buffer with 1.5 mM $MgCl_2$, and 10 pmol decanucleotide primer. Amplifications were carried out in a DNA thermal cycler (Eppendorf, mastercycler gradient) using following parameters: 94 °C for 5 min; 45 cycles at 94 °C for 1 min, 36 °C for 1 min, and 72 °C for 2 min; and a final extension at 72 °C for 5 min. PCR products were subjected to agarose gel (1.5% [w/v]) electrophoresis in 0.5X TBE buffer, along with 1 kb and 100 bp DNA ladders (Banglore Genie, India) as size markers. DNA was stained with ethidium bromide and photographed under UV light.

Screening Strategy and Identification of Specific RAPD Amplicon Eighty Random primers (OPA1-20, OPB1-20, OPC1-20, OPH1-20) (Qiagen Operon Technologies Inc., Alameda, CA, U.S.A.) were screened by RAPD for identification of specific marker. Amplification of DNA from the seven *Phyllanthus* species including eleven *P. emblica* cultivars was carried out using a single primer. The amplicon, which was monomorphic to all the *P. emblica* cultivars but absent in other six *Phyllanthus* species was identified.

Cloning of the RAPD Amplicon The putative marker amplified by the random primer OPA-16 was excised form 1.5% agarose gel with sterile gel slicer and purified using Clean Genei Gel Extraction kit (Banglore Genei, India). A-tailing to the amplicon was done by taking 5 μl of the amplified fragment, 0.9 U of *Taq* DNA polymerase, 2.5 μl 1X reaction buffer with 1.5 mM $MgCl_2$, 0.2 mM dATP in a 10 μl reaction volume and incubating at 70 °C for 30 min. Two microliters of the A-tailed DNA was ligated into a pGEM®-T easy vector (Promega Co., U.S.A.) following the supplier's instructions. The ligated vector was introduced into competent *Escherichia coli* strain XL1blue following protocol for transformation by calcium chloride as described by Sambrook *et al.* (2001). Ten distinct white colonies were picked up from the LB-ampicillin plate and recombinant DNA was isolated from each overnight grown colony. Confirmation of the clones was done by amplifying the DNA using SP6 and T7 promoter primers (Banglore Genei, India).

Sequencing, SCAR Primer Designing and Amplification of the Genomic Region The recombinant plasmid was purified by phenol, chloroform and isoamyl alcohol treatment following protocol of Sambrook and Russell (2001).[16] Both ends of each DNA insert were sequenced on an ABI 373 automated sequencer (Applied Biosystems, Inc.) using a DyeDeoxy™ Terminator Cycle Sequencing kit (Applied Biosystems) as recommended by the manufacturer. The vector specific promoter primers SP6 and T7 were used for sequencing. Based on the sequenced RAPD amplicon a pair of SCAR oligonucleotide primer (D1 and D2), which could amplify approximately 343 bp of the genomic *P. emblica* DNA was designed (Table 1). D1 was designed as forward primer and D2 as reverse primer. The sequences were custom synthesized from Integrated DNA Technologies Inc., U.S.A.

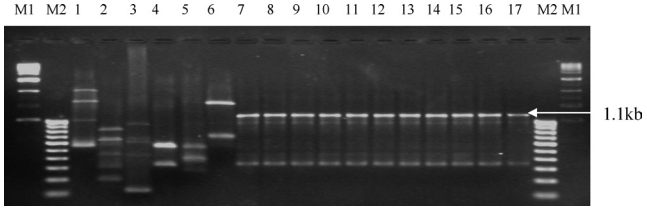

Fig. 1. RAPD Profiles of *Phyllanthus* Species Amplified with OPA-16 on 1.5% Agarose Gel

M1: mole wt marker (1 kb), M2: mole wt marker (100 bp), lane 1: *P. destichus*, lane 2: *P. urinaria*, lane 3: *P. reticulatus*, lane 4: *P. niruri*, lane 5: *P. simplex*, lane 6: *P. indofischeri.* lanes 7—17: *P. emblica* cultivars *viz*; 7: NA-06, 8: NA-07, 9: NA-10, 10: Kanchan, 11: Chakaiya, 12: Francis, 13: Banarasi, 14: Hathizool, 15: Dongri, 16: Bansired and 17: Anand-01.

The SCAR primers pair was used for PCR amplifications of genomic DNA from the seven *Phyllanthus* species (including 11 *P. emblica* cultivars). Thermal cycling conditions for amplification using SCAR primers were optimized as: 94 °C for 5 min; 45 cycles at 94 °C for 30 s, 55 °C for 30 s., and 72 °C for 2 min; and a final extension at 72 °C for 5 min.

PCR amplification for authentication of commercial samples was done by using the SCAR primer pair with the above-mentioned thermal cycling conditions.

Analysis of Sequence Data The DNA sequence was submitted to GeneBank (Accession numbers DX575975). Homology searches were performed within GenBank's non-redundant database using the BLAST 2.2.8 (Basic Local Alignment Search Tool) algorithm at http://www.ncbi.nlm.nih.gov/BLAST/ of the National Center for Biotechnology Information (NCBI), with the program BLASTX.

RESULTS

Identification of RAPD Marker for *P. emblica* High molecular weight genomic DNA was isolated from all the fresh tissues. The procedure yielded 400—600 ng of DNA per 100 mg of tissue. Fragmentation was observed in the DNA isolated from commercial samples. An absorbance (A_{260}/A_{280}) ratio of 1.6—1.8 indicates insignificant levels of contaminating proteins and polysaccharides. Of the 80 primers screened, 44 primers produced distinct, reproducible amplification profile for all the screened DNAs. Among these, 21 primers revealed polymorphism for the screened *Phyllanthus* species and were found monomorphic for *P. emblica* cultivars. Primers OPA-16 consistently amplified a single, intense band of approx. 1.1 kb for all the *P. emblica* cultivars, which was absent in rest of the 6 *Phyllanthus* species (Fig. 1). This band named as PE01 was selected as putative *P. emblica* specific marker.

Cloning and Sequencing of RAPD Marker PE01 was cloned and sequenced (Fig. 5). Restriction digestion analysis using restriction enzymes *Bam*HI and *Eco*RI revealed a band of 1.1 kb on 1.5% w/v agarose gel confirmed presence of in-

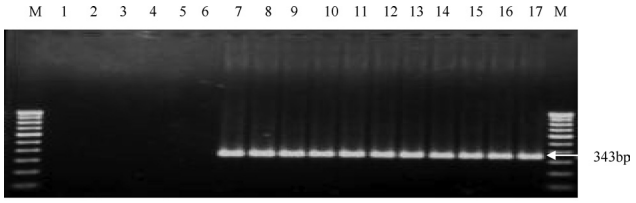

Fig. 2. PCR Amplification of *Phyllanthus* Species Using D1 and D2 on 1.5% Agarose Gel

M: mole wt marker (100 bp) lane 1: *P. destichus*, lane 2: *P. urinaria*, lane 3: *P. reticulatus*, lane 4: *P. niruri*, lane 5: *P. simplex*, lane 6: *P. indofischeri.* lanes 7—17: *P. emblica* cultivars *viz*; 7: NA-06, 8: NA-07, 9: NA-10, 10: Kanchan, 11: Chakaiya, 12: Francis, 13: Banarasi, 14: Hathizool, 15: Dongri, 16: Bansired and 17: Anand-01.

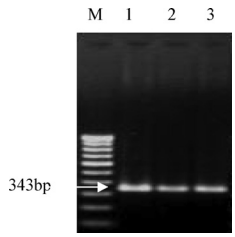

Fig. 3. PCR Amplification of Genomic DNA from Three Commercial Fruit Powder of *P. emblica* (Amlachurna) Using D1 and D2 on 1.5% Agarose Gel

M: mole wt marker (100 bp).

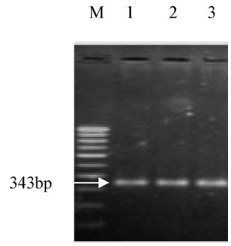

Fig. 4. PCR Amplification of Genomic DNA from Triphalachurna Using D1 and D2 on 1.5% Agarose Gel

M: mole wt marker (100 bp).

```
   1 AGCCAGCGAA CAGAGAGTGC TATGAATTTA ACCTTAAAAG TAACTTTCAA AGTACAATTA
       OPA-16
  61 ACCTGAGAAA  AACTAGTACT CTTAATCAGC CTACCAATCA GCAAGCCAAG CAATGTTTAA
 121 TGGATGGTGT ATTAAAGATC TAAACGAAGA TTGAAAACGT ATTACCGATA TAACAGATGG
 181 TCGACATATC AGGGAAAAGC AAGTTTTACC ATTACACGGA CAATGCCAAA TTGACAATAA
 241 TGTTACACAA AATTTCTTTA TAATGTTACA CAAAATTTCT TTCTTTGTAT CTCAATAATT
 301 CTATGACAAG TGCTACTCTG CAGCTATCTA ATGTTACATG AAAGGTTTGC TTGTGCTTTC
 361 AGATCTCGTG TAAAAAGCGT TGCATTTAAA TCAATAAATA GCAGGCTTTT TCTTCCAATT
       D1
 421 GGGAAGGATA AAGCCTGGCC GTAAATTCTT TGGTGACCCA TTTACATGTC ATTTGCCTTT
 481 TACTACGCTT GTAACTGGTC TCGCCGTAGG GCTTCTTGAA GCTGAATAAC TTGCTGTTGT
 541 TGGTCTGGTG TTAATGAACT CAACTGCTCA GGTGTAAGGC TGAGCACTTG CTGCAGCAAG
 601 GCAGAGTCAA CATCTGATGG AAGCTGCAAC TGGGAAAAAT CAGAACCGTA AAATTTGAGC
 661 TTCCACAAAA GAAAACATAA GAAACACTTG GAATTCACTG CAGAATATGT TGAATAATGC
                             D2
 721 AAACCAAAGC TTCATTTTCT TGATCAGCCT TGACATATAT ACCAAACAAT TATAAGAGGC
 781 TACAATCAAA GTGATTTATT ACAACTTATG CATAATCATT TCATTACCGA CAATGATTGT
 841 GAATGAAATT TTGAAGCCTC CATGAAGATA ATGGGCTGAT CCATGTAAAC TGAAAACTAG
 901 ATTAGTGCAA TACACTTGTA CCTATAAAGA TGCTATATTA CCTGAGAATT TTGTTTCTCC
 961 ACATATGGTG ACACTTCCGG CTGAGGAAGG AGATTTGCGG GCATGGAATC GACTGTAAAT
1021 GCTTTGGAAG GTCCCGAACC CAAAGCATTA GATGCACTGG AACCACCAGG AAGAAATGAG
1081 CTCCTTCCAT CATCCAGTTT GAGTAACTTC GCTGGCT
                          OPA-16
```

Fig. 5. Nucleotide Sequence of the RARD Amplicon PE01

sert in the vector. The recombinant was sequenced using primers SP6 and T7. The first ten nucleotides of the sequence obtained matched completely with the corresponding RAPD primer used.

Sequence Data Analysis The length of the PE01 marker sequence obtained was 1117 bp with 38% G+C content (A=371; C=211; G=212; T=323). BLAST results revealed that the sequence has partial homology with known plant nucleotide sequences at different sequence-similarity levels. Considerable similarity was found with RNA/nucleic acid binding protein (45.4 bits, and E-value=0.004) of *Arabidopsis thaliana* and proline-rich protein (41.2 bits, E-value=0.008) of *Oryza sativa.*

Amplification Using SCAR Primers The designed SCAR primer pair was used to amplify genomic DNA from the 7 *Phyllanthus* species (including 11 of *P. emblica* cultivars). A single, distinct and brightly resolved band of 343 bp was obtained in DNA isolated from all the 11 cultivars and no non-specific amplification was observed in the other six *Phyllanthus* species (Fig. 2). Reduction of the annealing temperatures did not generate any fragment other than the SCAR, confirming the specificity of the SCAR primer for all the *P. emblica* cultivars.

The SCAR primers were further used to amplify DNA from commercial samples of *P. emblica.* A sharp and reproducible band (343 bp) was observed for the three fruit powder samples (Amlachurna) (Fig. 3) and for the DNA isolated from the three Triphalachurna formulations (Fig. 4).

DISCUSSION

In this study, we have developed RAPD-SCAR marker for identification of *P. emblica.* RAPD analysis can reveal high degree of polymorphism; does not require prior DNA sequence information of the species and is easy to perform. Therefore, various researchers have explored its application for authentication of traditional Chinese medicines like ginseng,[17] *Echinacea*,[18] *Atractylodes*.[19] In our RAPD analysis, significant genetic polymorphism was observed among the *Phyllanthus* species. Polymorphism was also detected among the *P. emblica* cultivars, which was expected due to cross-pollination.[20] Since fruits of all *P. emblica* cultivars have equal commercial importance we selected the monomorphic band PE01 for SCAR marker development. In SCAR, pairs of 20—25 bp oligonucleotide primers specific to the sequence of polymorphic bands can be used to amplify the characterized regions from genomic DNA under stringent conditions, which makes these markers more specific and dependable as compared to RAPD markers.[21] Based on the sequence of PE01, SCAR primer pair (D1, D2) was designed. Genomic DNA from most of the commercial *P. emblica* containing samples was found to be fragmented. Therefore, D1 and D2 were designed to amplify a small region (343 bp) of PE01 to widen their application. As shown in Fig. 2, D1 and D2 generated 343 bp band in all *P. emblica* cultivars, while no amplification was observed in other *Phyllanthus* species.

The genetic polymorphism observed among the cultivars is interesting and can be used to develop markers for cultivar identification. Studies comparing the therapeutic efficacy of various cultivars are needed and development of such culti-

var specific markers would then be relevant. We are working on metabolite profiling and genetic characterization of *P. emblica* cultivars.

Utility of D1 and D2 for identification of *P. emblica* in commercial samples of dried fruit powder and Triphalachurna was tested. SCAR primers D1 and D2 amplified the expected 343 bp DNA fragment in all tested samples confirming the presence of *P. emblica.* These results substantiate the applicability of the designed primers as a qualitative diagnostic tool for identification of *P. emblica.* However, for quantitative analysis of *P. emblica* content in commercial samples advanced technique such as real time PCR could be tried. Further, there is pool of material that can be used as adulterant for crude and processed *P. emblica* fruits. The adulterant may be phyllogenetically close or distinct (*e.g.* dried fruit pieces of pumpkin) from *P. emblica* and we are in process of developing primers for identification of such frequently used adulterants.

Acknowledgements Part of this work has been carried out under University Grants Commission, New Delhi Major Research project. D.W. and P.C. thank Council for Scientific and Industrial Research, Technology and Business Development Division, New Delhi for financial assistance and Mahatma Phule Agricultural University, Rahuri for providing the plant cultivars. Generous support of National AIDS Research Institute, Pune for cloning and sequencing is gratefully acknowledged. Critical evaluation of the work by Dr. Abhay Jere and Dr. Abhay Harsulkar is acknowledged.

REFERENCES

1) Ganju L., Karan D., Chanda S., Srivastava K. K., Sawhney R. C., Selvamurthy W., *Biomed. Pharmacother.*, **57**, 296—300 (2003).
2) Jose J. K., Kuttan G., Kuttan R., *J. Ethnopharmacol.*, **75**, 65—69 (2001).
3) Bhattacharya A., Ghosal S., Bhattacharya S. K., *Indian J. Exp. Biol.*, **38**, 877—880 (2000).
4) Al-Rehaily A. J., Al-Howiriny T. A., Al-Sohaibani M. O., Rafatullah S., *Phytomedicine*, **9**, 515—522 (2002).
5) Government of India, "The Ayurvedic Pharmacopoeia of India," Controller of Publications, Delhi, 2001.
6) "Indian Herbal Pharmacopoeia," Indian Drug Manufacturer's Association, Mumbai, 2002.
7) Techen N., Crockett S. L., Khan I. A., Scheffler B. E., *Curr. Med. Chem.*, **11**, 1391—1401 (2004).
8) Joshi K., Chavan P., Warude D., Patwardhan B., *Curr. Sci.*, **87**, 157—165 (2004).
9) Weder J. K., *J. Agric. Food Chem.*, **50**, 4456—4463 (2002).
10) Williams J. G. K., Kubelik A. R., Livak K. J., Rafalski J. A., Tingey S. V., *Nucleic Acids Res.*, **18**, 6531—6535 (1990).
11) Wang J., Wai-Yan H., Ngan F., But P. P., Shaw P., *Planta Med.*, **67**, 781—783 (2001).
12) Lee M. Y., Doh E. J., Chae H. P., Young H. K., Eung S. K., Ko B. S., Oh S. E., *Biol. Pharm. Bull.*, **29**, 629—663 (2006).
13) Das M., Bhattacharya S., Pal A., *Annals Bot.*, **95**, 835—841 (2005).
14) Kirtikar K. R., Basu B. D., "Indian Medicinal Plants," Vol. III, International Book Distributor, Dehradun, India, 1996, pp. 2117—2127.
15) Warude D., Chavan P., Joshi K., Patwardhan B., *Plant Mol. Biol. Reporter*, **21**, 1—6 (2003).
16) Sambrook J., Russell D. W., "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, New York, 2001.
17) Um J. Y., Chung H. S., Kim M. S., Na H. J., Kwon H. J., Kim J. J., Lee K. M., Lee S. J., Lim J. P., Do K. R., Hwang W. J., Lyu Y. S., An N. H., Kim H. M., *Biol. Pharm. Bull.*, **24**, 872—875 (2001).
18) Wolf H. T., Zundorf I., Winckler T., Bauer R., Dingermann T., *Planta Med.*, **65**, 773—774 (1999).
19) Chen K. T., Su Y. C., Lin J. G., Hsin L. H., Su Y. P., Su C. H., *Acta Pharmacol. Sin.*, **22**, 493—497 (2001).
20) Pareek O. P., "Germplasm Survey, Collection and Evaluation," In Annual Report, AICRP on Arid Zone Fruits, NRC for Arid Horticulture, Bikaner, India, 1998, pp. 13—21.
21) Paran I., Michelmore R. W., *Theor. Appl. Genet.*, **85**, 985—993 (1993).