

Molecular identification of Mango, *Mangifera indica L.var. totupura*

Sankar Jagarlamudi¹, Rosaiah G², Ravi Kumar Kurapati¹ & Rajasekhar Pinnamaneni^{3*}

¹Institute of Biological Sciences, Vijayawada, Andhra Pradesh, India; ²Department of Botany, Acharya Nagarjuna University, Nagarjuna Nagar, Andhra Pradesh, India; ³Department of Biotechnology, Sreenidhi Institute of Science and Technology, Yammampet, Ghatkesar, R.R.District, Andhra Pradesh, India; Rajasekhar Pinnamaneni - Email: pinnamaneniraj@gmail.com; *Corresponding author

Received September 20, 2010; Accepted January 10, 2011; Published February 15, 2011

Abstract:

Mango (*Mangifera indica*) belonging to Anacardiaceae family is a fruit that grows in tropical regions. It is considered as the King of fruits. The present work was taken up to identify a tool in identifying the mango species at the molecular level. The chloroplast *trnL-F* region was amplified from extracted total genomic DNA using the polymerase chain reaction (PCR) and sequenced. Sequence of the dominant DGGE band revealed that *Mangifera indica* in tested leaves was *Mangifera indica* (100% similarity to the ITS sequences of *Mangifera indica*). This sequence was deposited in NCBI with the accession no. GQ927757.

Key words: Mango, *Mangifera indica*, *trnL-F*.

Abbreviations: AFLP-Amplified fragment length polymorphism, cpDNA-Chloroplast DNA, DDGE-Denaturing gradient gel electrophoresis, DNA-Deoxyribo nucleic acid, EDTA-Ethylenediamine tetraacetic acid, HCl-Hydrochloric acid, ISSR-Inter simple sequence repeats, ITS- Internal transcribed spacer, MATAB-Methyl Ammonium Bromide, Na₂SO₃-Sodium sulphite, NaCl-Sodium chloride, NCBI- National Centre for Biotechnology Information PCR-Polymerase chain reaction, PEG-Polyethylene glycol, RAPD-Randomly amplified polymorphic DNA, *trnL-F*-Transfer RNA genes start codon-termination codon

Background:

Mango (*Mangifera indica*) is a fruit that grows in tropical regions belonging to Anacardiaceae Lindl., the cashew family which includes more than 700 species in 82 genera that are primarily distributed pantropically. *Mangifera indica* L. originated in a region including the north-eastern part of India (Assam), the western part of Myanmar, and Bangladesh. Mango was domesticated in this region and has been cultivated in India for 4000 years. Cultivation spread first to Malaysia and South-East Asia, supposedly expanded by Buddhist monks. Purseglove suggested that Phoenicians and Arabs spread the crop from India to East-Africa where it has been cultivated since the 10th century. The global spread of mango outside its original centres of domestication probably did not occur until the beginning of the European voyages of the 15th and 16th centuries, when the Portuguese took the mango to West-Africa and from there to Brazil at the beginning of the 18th century, and when the Spanish introduced polyembryonic mango types from The Philippines to Central America through the Pacific trading ports of Mexico and Panama.

A regional inventory of fruit genetic resources has recently been undertaken in the French West Indies (FWI). The objective is to conserve the FWI fruit variability *in-situ* and *ex-situ*, as well as to try local varieties under standard conditions. The project includes a genetic diversity study of collected material in comparison with the diversity observed in the germplasm bank maintained locally in Guadeloupe [1]

Different markers have been used to study mango diversity: RAPD markers in Florida [2], India [3], [4] Venezuela [5] and Brazil [6], AFLP [7], and ISSR [8], [9] Nevertheless, these molecular markers are dominant

and they do not permit the differentiation of heterozygous from homozygous accessions.

The first co-dominant markers developed were isozymes [10], [11], [12]. More recently three research groups chose to develop SSR markers [13], [14] including our own [15]. These molecular markers are co-dominant, specific and highly variable. These qualities make them highly suitable to study diversity in supposedly related populations or cultivars. The present work was taken up to identify a tool in identifying the mango species at the molecular level.

Materials and Methodology:

Plant materials and DNA isolation:

Fresh leaves of Mango, *Mangifera indica* L.var.*totupura* used in this study for DNA extraction were collected from Nuziveedu, Krishna District, Andhra Pradesh, India. Plant tissue was ground in one of the ways: by hand with a mortar and pestle or in tubes with sterile glass beads and sand placed in a tissue disruptor.

DNA extraction:

DNA was extracted from fresh young leaves using a MATAB (Methyl ammonium Bromide) protocol described by Risterucci *et al.*, (2000) [16]. Fresh leaves were grounded to a fine powder in liquid nitrogen, and incubated for 30 min at 74°C with extraction buffer (1.4 M NaCl, 100 mM Tris HCl pH 8.0, 20 mM EDTA, 10 mM Na₂SO₃, 1% PEG 6000, 2% MATAB). After being cooled to 20°C, an equal volume of chloroform-isoamyl alcohol (24:1 v/v) was added, 20-μl reaction mixtures followed by emulsification. The tube was then centrifuged at 10000 g for 20 min and

DNA was precipitated from the supernatant with cold isopropanol. DNA concentrations were estimated on a 0.8% agarose gel stained with ethidium bromide.

DNA amplification and sequencing:

The chloroplast *trnL-F* regions were amplified from extracted total genomic DNA using the polymerase chain reaction (PCR) method. The universal primers 1, 2, 3 and 4 (Helini Biomolecules, Chennai) of Taberlet *et al.*, 1991[17] (**Table 1** see Supplementary material) (**Figure 1**) were used to amplify *trnL-F*. Thermal cycling parameters were an initial denaturation of 2 minutes at 97°C; 30 cycles of 94°C for one minute, annealing at 48°C for two minutes, and elongation at 72°C for two minutes; followed by an elongation step of 72°C for 16 minutes.

Eluting DNA from agarose gel fragments:

Ethidium bromide stained agarose gel was visualized under a transilluminator. The fragment of interest was excised with a clean razor blade. After removing the excess liquid, the agarose fragment was placed in the spin column. The tube was centrifuged at 5500 rpm for not more than 45 seconds for the elution of DNA. The eluent was checked by running on an agarose gel and observed on a transilluminator for the presence of ethidium bromide stained DNA. The eluted DNA was used directly in manipulation reactions. This DNA fraction was subjected for sequencing (Helini Biomolecules, Chennai).

Sequencing and Phylogenetic placement:

The eluted PCR product was directly sequenced using *trnL-F* specific primers at Ohmlina Centre for Molecular Research, Chennai. Sequencing reactions were carried out with ABI PRISM Dye Terminator Cycle Sequence Ready Reaction Kit (Applied Biosystems Inc., USA). The obtained sequence was compared to the sequences in GenBank using the BLAST algorithm [18] to search for close evolutionary relatives.

GenBank accession numbers:

The representative sequence of the *Mangifera indica* species was deposited in GenBank of National Centre for Biotechnology Information (NCBI).

Results and Discussion:

Collection of Plant material:

Fresh leaves of Mango, *Mangifera indica* L.var.*totupura* were collected from Nuziveedu, Krishna District, Andhra Pradesh, India were used in this study for DNA extraction.

Amplification of DNA coding for *trnL-F*:

The genomic DNA of the *Mangifera indica* was subjected for the isolation of the DNA coding for *trnL-F* (**Figure 2**) by using Polymerase chain reaction. The bands were cut and eluted and the DNA so obtained was subjected for sequencing.

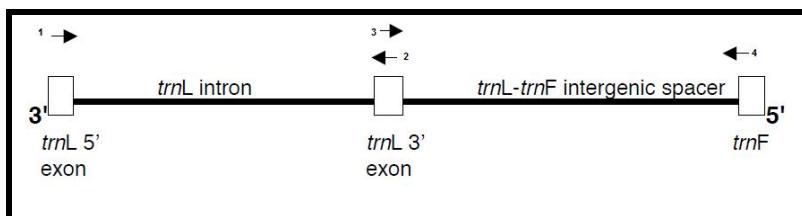


Figure 1: Approximate location of *trnL-F* primers used in this study

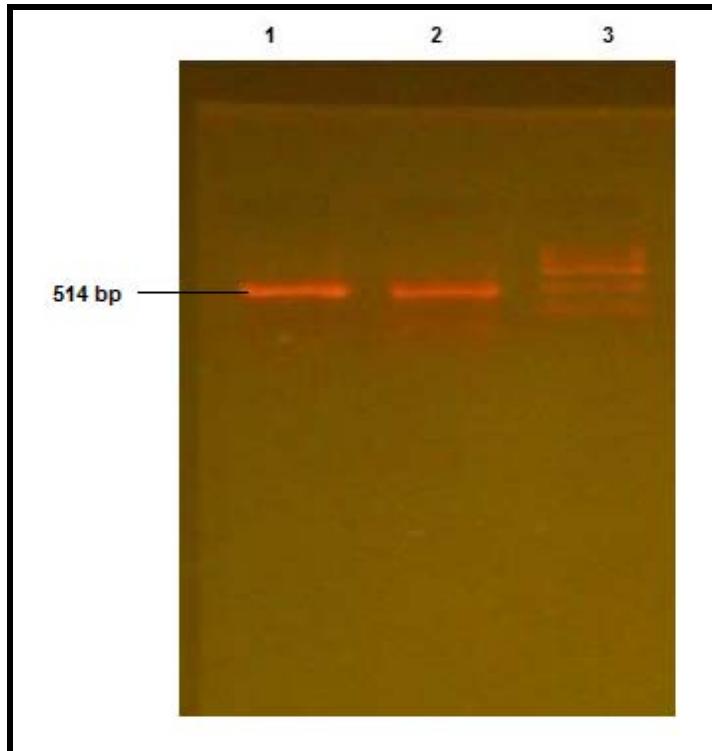


Figure 2: Agarose gel showing amplified *trnL-F* DNA of *Mangifera indica* L.var.*totupura*. Lane 1 and 2: Amplified DNA; Lane 3: 1 kb DNA Ladder

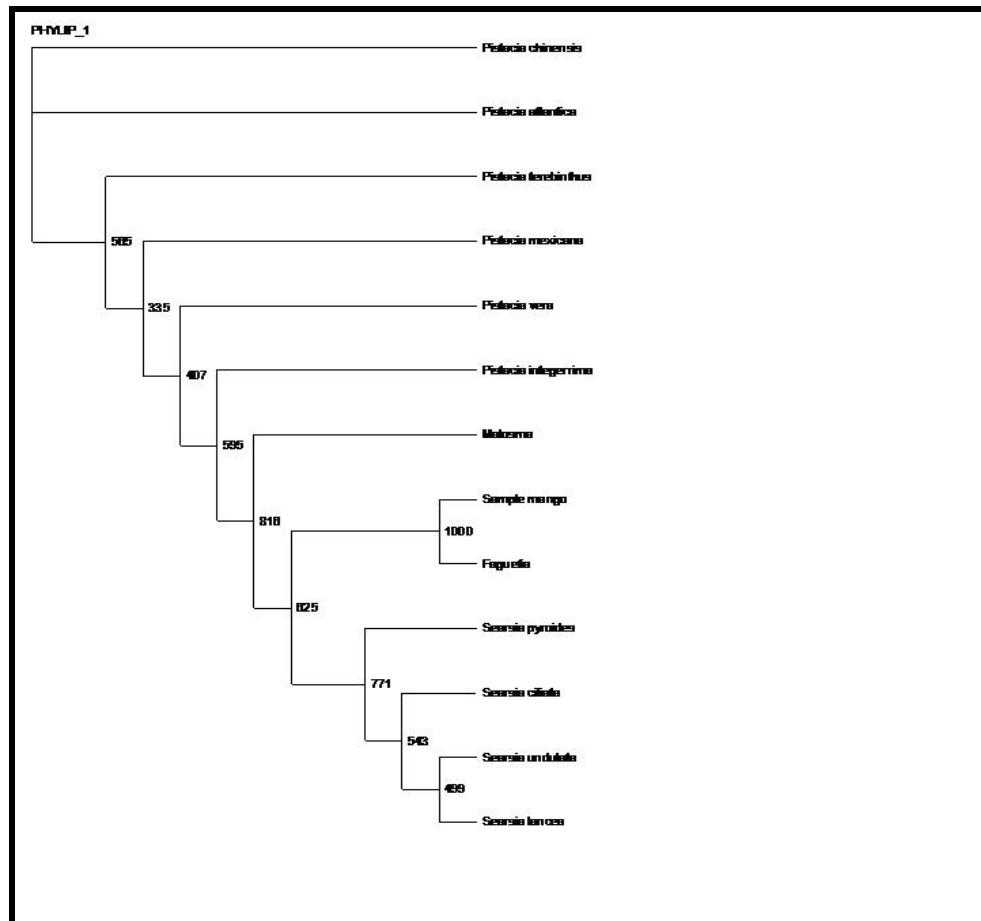


Figure 3: Phylogenetic affiliation of *Mangifera indica* based on *trnL-F*

Sequence analysis:

The sequence analysis demonstrated that all the corresponding bands on agarose gel belonged to *Mangifera indica*. Upon sequencing of the amplified DNA, the data obtained corresponds to 514 bases for *Mangifera indica*. The sequence so obtained was as follows.

```

CCAAATTGATTGAGCTTGGTATGGAAACCTACTAAGTGATAACT
TTCAAATTTCAGAGAAAACCCTGGAATC
AAAAATGGCAATCCTGAGCCAATCCTATTTACGAGAACAA
AAACAACAAGGGGTCAAGACGGAGA
AAAAAAGGATAGTGCAGAGACTCAATGGAAGCTGTTCTAAC
AATGGAGTTGATTGCTTTGGTA
AAGAAAGGAATTCTCTATCGAATATCGAAAGGCCATAACGGA
TGAAGGATAAGCCTATATACACTATGT
ATAAGTAATGAAAAAATACACTATGTATACGTAATGAAAAAGG
ATCTCAAAATGACGCCGAATCCTT
TTTGTCTTCTTGAAGAACTAATTAATCAATCGGACGAGAA
AAAGATAGAGTCCCATTCTACATGC
CAATATCAATACTGCAACAATGAAATTATAGTAAGAGGAAA
ATCCGTCGACTTTAGAAATCGTGAGGG
TTCAAGTCCTCCTATCCCCCAAAA

```

Phylogenetic affiliation of the *Mangifera indica* based on the ITS sequences:

Sequence of the dominant DGGE bands revealed that *Mangifera indica* in tested leaves was *Mangifera indica* (100% similarity to the ITS sequences of *Mangifera indica*) with the accession no. GQ927757. This sequence can be found in NCBI Genome Databank. The phylogenetic tree was shown in Figure 3.

The dataset consisted of 514 characters (100%) were parsimony informative. The matrix was easily manually aligned. Coding gaps as binary characters, missing data had no affect on the topology and very little affect on branch support. The 100% bootstrap consensus tree is shown in Figure 2. The *trnL-F* gene of mango was amplified and it was sequenced. The sequence was blasted against nr database and the top hits were taken and aligned by multiple sequence alignment. The picture shows the phylogenetically related sequences. As our sequence is the first entry in genbank, there was no homologous sequence for mango. In future if mango sample was searched in database, our sequence will be the first hit. The topology is consistent with the maximum parsimony tree, which is more resolved within the Anacardiaceae clades.

The chloroplast DNA (cpDNA) *trnT-F* region in land plants consists of the transfer RNA genes *trnT_{ugu}*, *trnL_{Uua}*, and *trnF_{gaa}* arranged in tandem and separated by noncoding spacer regions. The region is positioned in the large single copy region, approximately 8 kb downstream of *rbcL*. The *trnL* gene of cyanobacteria and a number of chloroplast genomes, including that of all land plants, contains a group I intron positioned between the U and the A of the UAA anticodon loop. This intron is inferred from phylogenetic analysis to have been present in the cyanobacterial ancestor of the plastid lineages of Rhodophyta, Chlorophyta, and Glaucochlorophyta and to have been subsequently vertically transmitted [19].

The succession of conserved *trn* genes and the apparent absence of gene rearrangements in the *trnT-F* region made the design of plant universal primers possible [17]. As a consequence, the *trnL-F* region, comprising the *trnL* intron and *trnL-F* spacer, has become one of the most widely used

chloroplast markers for phylogenetic analyses in plants [20]. The accumulation of an increasingly large number of sequences of the *trn(T-L)-F* region from a wide range of plants has allowed further study of structures, functions, and evolution in different orders of flowering plants [21], in basal angiosperms [22], in land plants [21], in bryophytes [23], and in Gnetales [24].

Sequences from the *trnL-F* region (excluding the *trnT-L* region and *trnL* 5' exon) have recently been used, in combination with those from further chloroplast markers *rbcL* and *matK*, as a source of characters for phylogenetic reconstruction in the tropical flowering plant family Annonaceae Juss. These phylogenies have been used to answer questions about morphological character evolution [25], classification (Mols *et al.*, 2004), biogeography [26] [27], and molecular dating [28]. These markers appeared to contain complementary phylogenetic signals, as is expected from different sequences sampled from the plastid genome [29], and were thus applied in combined analyses. The combined analyses yielded better resolved phylogenies, subject to higher levels of support, than those derived from individual markers.

Conclusion:

The above study can be utilized in identifying the Mango species directly by using any part of the tissue in live form or even from the fossil specimens. The gene sequences of *trnL-F* will be the same in any part of the plant because of the presence of chloroplast and its genome is same throughout. The sequence as reported in this present work is of *Mangifera indica L.var. totupura*. There was no sequence available in the database for mango till now and hence it can be exploited for the identification of mango species up to the level of varieties as the sequences will be more or less conserved with minor variations.

References:

- [1] http://www.hort.purdue.edu/newcrop/morton/mango_ars.html
- [2] RJ Schnell *et al.* *Theoretical Applied Genetics* (1995) **90**: 269
- [3] NVH Kumar & P Narayanaswamy, *Journal of Horticultural Science and Biotechnology* (2001) **76**(5): 529
- [4] JL Karihaloo *et al.* *Journal of Horticultural Science and Biotechnology* (2003) **78**(3): 285
- [5] JA Lopez-Valenzeula *et al.* *Hortscience* (1997) **32**(6): 1105
- [6] VA Barbosa de Souza & P Sarmanho da Costa Lima, *Acta Horticulturae* (2004) **645**: 303
- [7] W Eiadthong *et al.* *J Am Soc Hortic Sci.* (2000) **125**(2): 160
- [8] W Eiadthong *et al.* *Scientia Horticulturae* (1999) **82**(1/2): 57
- [9] A Gonzalez *et al.* *Acta Horticultura*. (2002) **575**(1): 139
- [10] C Degani *et al.* *HortScience* (1992) **27**: 252
- [11] GS Pascua *et al.* *Philippine Journal of Crop Science*, (1996) **21**(Supplement 1): 10
- [12] W Eiadthong *et al.* *Thai Journal of Agricultural Science* (1998) **31**(4): 555
- [13] C Honsho & K Nishiyama, *Molecular Ecology Notes* (2005) **5**: 152
- [14] RJ Schnell *et al.* *Molecular Ecology Notes* (2005) **5**: 625
- [15] MF Duval *et al.* *Molecular Ecology Notes* (2005) **5**(4): 824
- [16] AM Risterucci *et al.* *Theor Appl Genet* (2000) **101**: 948
- [17] P Taberlet *et al.* *Plant Molecular Biology* (1991) **17**: 1105 [PMID: 1932684]
- [18] SF Altschul *et al.* *Nucleic Acids Research* (1997) **25**: 3389 [PMID: 9254694]
- [19] A Besendahl *et al.* *Current Genetics* (2000) **37**: 12 [PMID: 10672439]
- [20] D Quandt *et al.* *Monographs in Systematic Botany* (2004) **98**: 13
- [21] FT Bakker *et al.* *Molecular Biology and Evolution* (2000) **17**: 1146 [PMID: 10991703]
- [22] T Borsch *et al.* *Journal of Evolutionary Biology* (2003) **16**: 558 [PMID: 14632220]
- [23] D Quandt & M Stech, *Plant Biology* (2004) **6**: 545 [PMID: 15375725]
- [24] H Won & SS Renner, *Journal of Molecular Evolution* (2005) **61**: 425 [PMID: 16155750]
- [25] H Sauquet *et al.* *Botanical Journal of the Linnean Society* (2003) **142**: 125
- [26] JE Richardson *et al.* *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* (2004) **359**: 1495 [PMID: 15519968]
- [27] MD Pirie *et al.* *Journal of Biogeography* (2006) **33**: 31
- [28] MD Pirie *et al.* *Regnum Vegetabile* (2005) **143**: 149-174
- [29] MW Chase & AV Cox, *Australian Systematic Botany* (1998) **11**: 215

Edited by B Lakshmi

Citation: Sankar *et al.* Bioinformation 5(10): 405-409 (2011)
License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

Supplementary material:

Table 1: Sequences of the *trnL-F* primers used in this study

Primer	5'-3' Sequence
1	CGAAATCGGTAGACGCTACG
2	GGGGATAGAGGGACTTGAAC
3	GGTTCAAGTCCCTCTATCCC
4	ATTGAACTGGTGACACGAG