International Journal of Academic Research and Development
ISSN: 2455-4197, Impact Factor: RJIF 5.22
www.academicsjournal.com
Volume 2; Issue 4; July 2017; Page No. 645-648

DNA barcoding as a means for identifying mangrove plants of Kerala

[*1] S Surya, [2] N Hari

[1] Research Scholar, CMS College, Kottayam, Kerala, India
[2] Assistant Professor, CMS College, Kottayam, Kerala, India

Abstract
DNA barcoding is currently gaining popularity due to its simplicity and high accuracy as compared to the complexity and subjective biases associated with morphology-based identification of taxa. The standard chloroplast DNA barcode for land plants recommended by the Consortium for the Barcode of Life (CBOL) plant working group needs to be evaluated for a wide range of plant species. The present study therefore tested the potential of the *rbcL* and *matK* marker for the identification of mangrove plants belonging to diverse families of coastal regions. Here, DNA barcoding is being used to identify mangrove plants found in Kerala and distinguished them from other similar species. Several challenges to the successful implementation of plant DNA barcoding are presented and discussed. Despite these challenges, DNA barcoding has the potential to uniquely identify mangrove plants and provide quality control and standardization of the plant material supplied to the pharmaceutical industry.

Keywords: DNA barcoding; *rbcL*; *matK;* mangrove plants; identification.

Introduction
The main aim of DNA barcoding is to establish a shared community resource of DNA sequences that can be used for organism identification and taxonomic clarification. In plants, establishing a standardized DNA barcoding system has been more challenging [1, 2, 3]. In DNA barcoding, the unique nucleotide sequence patterns of small DNA fragments (400–800 bp) are used as specific reference collections to identify specimens and to discover overlooked Species [4, 5]. Thus, the initial goal of the DNA barcoding process is to construct on-line libraries of barcode sequences for all known species that can serve as a standard to which DNA barcodes of any identified or unidentified specimens can be matched [6]. This can alleviate several inherent problems associated with traditional taxonomic identification, based on morphological characters, such as wrong identification of species due to phenotypic plasticity and genotypic variability of the characters, overlooking cryptic taxa, difficulty in finding reliable characters due to long maturity periods, etc.5. DNA barcoding, thus, can provide the taxonomists, conservationists and others who need the identification of species, a cost-effective and efficient tool, much as a barcode that identifies supermarket products [7, 8].
Several chloroplast gene regions are typically used as plant barcodes, with maturase K (*matK*) and ribulose 1,5-bisphosphatecarboxylase/oxygenase large subunit (*rbcL*) considered core barcodes [1]. Based on assessments of recoverability, sequence quality, and levels of species discrimination, the Consortium for the Barcode of Life (CBOL) plant working group has recommended a standard barcode comprising ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) and/or maturase K (*matK*) for the barcoding of all land plants [1]. However, the universality of barcode markers is hampered due to morphological/geographical variation and reticulate evolution in plant species [8]. The ongoing research on plant barcoding suggests that the development of universal DNA barcoding markers for land plants is challenging; even the choice of the correct loci has been debated [9, 10].

Materials and methods
Plants were randomly collected in the morning from 9.00 am until 12.00 noon. Whole plants and/or plant parts were cut using sharp scissors and placed in a sterile plastic bag. Morphological confirmation and species identification was done by Botanical Survey of India, Coimbatore (Table 1). DNA sequences were trimmed and manually aligned using the software Bio Edit [11]. Sequence data for 4 medicinal plants were downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/). The gene regions chosen were *rbcL*, *matK*, sequenced for each species.

Table 1: General information on fourteen species of medicinal plants from Pakistan collected for DNA barcoding in this study. Species in bold have not been sequenced for two or three of the primary barcoding regions.

Species	Family	Common Name
Avicennia marina	Avicenniaceae	Gray Mangrove
Avicennia officinalis	Avicenniaceae	Uppatti
Sonneratia alba	Sonneratiaceae	Mangrove apple
Sonneratia caseolaris	Sonneratiaceae	crabapple mangrove

Result and Discussion

Several potential challenges in sequencing of mangrove DNA because these plants contain biologically active secondary compounds, including tannins, alkaloids, and polysaccharides, all of which can inhibit DNA extraction and amplification by co-precipitating with or binding to DNA. If clean genomic DNA is obtained, appropriate primers are needed to amplify the targeted gene region. Certain gene regions, including *rbcL* and *matK* have universal primers that work for most plants.

In *Sonneratia* both the species show variation in gene sequences in *rbcL* gene, the base pair variations are in the position 406 and in *matK* 412,581,602 positions were changed. In *Avicennia* species the base pair variation show in the position 177, 230, 664, in *rbcL* and in *mat*K 333,334,336,345,509,529,607,614, 716, 816 base pair variation were noticed. (Plate-1, 2, 3 and 4).

Within a given genus, the *rbcL* and *matK* sequence might be identical for all species, which means that while *rbcL* and *matK* might verify the genus, it cannot verify the species. Variability within a given gene region may be quite different from one genus or family to the next, so the best way to increase the likelihood of a positive identification is to adequately sample related species for every target plant [12].

DNA sequences that are currently available in GenBank demonstrate the challenge of discrimination power. Table 2 lists barcoding sequences of 4 mangrove plants and indicates how well each sequence identifies the plant. The results vary by species and by gene region.

This leads to the question of what quality of match is required to use barcodes for identification. A match of 100% between a query sequence and a reference sequence is unambiguous at one level – each base pair is exactly matched. However, if the query sequence is 150 base pairs long, and the reference sequence is 2000 base pairs long, the 100% match might not be as meaningful. The match might be along a part of the gene region that is highly conserved, with little to no variation among many species. Although *rbcL* and *matK* are relatively long (approximately 1430 and 1550 bp respectively), not all portions evolve at the same rate, and submissions of reference sequences to GenBank do not always include the complete gene region [12].

The utility of a given barcoding region needs to be evaluated and confirmed for each different species whose identity is being verified. The data in Table 2 show that while *rbcL* and *matK* is a good barcode for *Avicennia* and *Sonneratia* with a 100% match to multiple vouchers.

Table 2: Sequence data from GenBank for 4 species of medicinal plants.

Species	Gene	% Match in GenBank
So Avicennia marina	rbcL, matK	99%, 100% similarity to different vouchers.
Avicennia officinalis	rbcL, matK	100%, 100% similarity to different vouchers.
Sonneratia alba	rbcL, matK	99%, 100% similarity different vouchers.
Sonneratia caseolaris	rbcL, matK	100%, 99%. Similarity to different vouchers.

Sequences were downloaded for *matK*, and *rbcL*. The % match indicates how closely the barcode sequences matched other accessions in Gen Bank, including other voucher sequences for the same species.

1 AAGTGTTGGA TTCAAAGCCG GTGTTAAAGA TTATAAACTG ACTTATTATA CTCCTGAATA
61 TCAAACCAAA GATACTGATA TCTTGGCAGC ATTCCGAGTA ACTCCTCAAC CTGGAGTTCC
121 GGCTGAGGAA GCAGGGGCTG CAGTAGCTGC TGAATCTTCT ACTGGTACCT GGACAACTGT
181 GTGGACCGAT GGGCTTACCA GCCTTGATCG TTATAAAGGA AGATGCTACA ACATCGAGCC
241 TGTTGCTGGA GAAGAAAATC AATATATATG TTATGTAGCT TACCCCTTAG ACCTTTTTGA
301 AGAAGGTTCT GTTACTAATA TGTTTACTTC CATTGTGGGT AATGTATTTG GGTTCAAAGC
361 CCTGCGTGCT CTACGTCTGG AGGATCTGAG AATCCCTATT GCCTATATTA AAACTTTCCA
421 AGGCCCGCCT CATGGTATCC AAGTTGAGAG AGATAAATTG AACAAGTATG GCCGTCCCCT
481 ATTGGGATGT ACTATTAAAC CTAAATTGGG ATTATCTGCT AAGAACTACG GTAGAGCGGT
541 TTATGAATGT CTTCGTGGTG GACTTGATTT TACGAAGGAT GATGAGAACG TGAACTCACA
601 ACCATTTATG CGTTGGAGAG ACCGTTTCTT ATTTTGTGCC GAAGCACTTT ATAAAGCACA
661 GAATGAAACT GGTGAAATCA AAGGGCATTA CTTGAAT 697

***Sonneratia alba rbcL* gene.**

1 AAGTGTTGGA TTCAAAGCCG GTGTTAAAGA TTATAAACTG ACTTATTATA CTCCTGAATA
61 TCAAACCAAA GATACTGATA TCTTGGCAGC ATTCCGAGTA ACTCCTCAAC CTGGAGTTCC
121 GGCTGAGGAA GCAGGGGCTG CAGTAGCTGC TGAATCTTCT ACTGGTACCT GGACAACTGT
181 GTGGACCGAT GGGCTTACCA GCCTTGATCG TTATAAAGGA AGATGCTACA ACATCGAGCC
241 TGTTGCTGGA GAAGAAAATC AATATATATG TTATGTAGCT TACCCCTTAG ACCTTTTTGA
301 AGAAGGTTCT GTTACTAATA TGTTTACTTC CATTGTGGGT AATGTATTTG GGTTCAAAGC
361 CCTGCGTGCT CTACGTCTGG AGGATCTGAG AATCCCTACT GCCTATATTA AAACTTTCCA
421 AGGCCCGCCT CATGGTATCC AAGTTGAGAG AGATAAATTG AACAAGTATG GCCGTCCCCT
481 ATTGGGATGT ACTATTAAAC CTAAATTGGG ATTATCTGCT AAGAACTACG GTAGAGCGGT
541 TTATGAATGT CTTCGTGGTG GACTTGATTT TACGAAGGAT GATGAGAACG TGAACTCACA
601 ACCATTTATG CGTTGGAGAG ACCGTTTCTT ATTTTGTGCC GAAGCACTTT ATAAAGCACA
661 GAATGAAACT GGTGAAATCA AAGGGCATTA CTTGAAT 697

Plate 1: *Sonneratia caseolaris rbcL* gene.

1 AGAGGACCTA TTTTCACATT TAGATTATGC GTCAGATGTA CTAATACCCT ATCCCACCCA
61 TTTTGAAATC TTGGTTCAAA CCCTTCGCTA CTGGGTGAAG GATGCCTCTT CTTTGCATTT
121 TTTACGTTTC TTTTTCTACG AGTATTGTAA TTGGAATAGT CTTATTACTC CCCAAAAACA
181 TATTTCCATT TTTTTAAAGG GTAATCCAAG ATTTTTCTTG TTCCTATATA ATTCTTTTGC
241 ATGTGAATAC GAATCTATTT TCCTTTTTCT CCGTAATCAA TCTTCTCATT TCCGGTCAAC
301 ATCTTCTGGA GTCTTTTTTG AGCGAATCTA TTTCTATGTA AAAATAGAAA ATCTTGTTGA
361 AGTTTTTTTT GATAATGATT TTCGGGACAT TCGATCCTTC TTCAAGGATT CTTTCATGCA
421 TTATGTTAGA TATCAAGGAA AATCAATTCT GGCTTCAAAA GATACACCCT TTCTGATGAA
481 TAAATGGAAA TACTACCTTG TCAATTTATG GCAATATCAT TTTTACGCGT GGTCCCAACC
541 AGGAAGGATC AATCTAAACC GATTAGGAAA GTATTCTCTT GACTTTTTGG GCTATTTTTC
601 AAACGTGCAA CTCACTTTTT CAGTGGTACG AAGTAAAATG CTCGAAAATT CATTTCTAGT
661 AAATACTGCT ATGAAGAAGC TCGAAACAAT AGTTCCCGTT ATTCCTTTGA TTGGATCGTT
721 GTCTAAAGCG AAATTTTGTA ACGTATTAGG ATATCCCGTT AGTAAGTCGA CCCGGACCGA
781 TTCATCAGAT TCTGATATTA TCGACCGATT TGTGCGTATA TGCAGAAATC TTTCTCATTA
841 TCACAGTGGA TCCTCAAAAA AAA 863

***Sonneratia alba matK* gene**

1 AGAGGACCTA TTTTCACATT TAGATTATGC GTCAGATGTA CTAATACCCT ATCCCACCCA
61 TTTTGAAATC TTGGTTCAAA CCCTTCGCTA CTGGGTGAAG GATGCCTCTT CTTTGCATTT
121 TTTACGTTTC TTTTTCTACG AGTATTGTAA TTGGAATAGT CTTATTACTC CCCAAAAACA
181 TATTTCCATT TTTTTAAAGG GTAATCCAAG ATTTTTCTTG TTCCTATATA ATTCTTTTGC
241 ATGTGAATAC GAATCTATTT TCCTTTTTCT CCGTAATCAA TCTTCTCATT TCCGGTCAAC
301 ATCTTCTGGA GTCTTTTTTG AGCGAATCTA TTTCTATGTA AAAATAGAAA ATCTTGTTGA
361 AGTTTTTTTT GATAATGATT TTCGGGACAT TCGATCCTTC TTCAAGGATT CGTTCATGCA
421 TTATGTTAGA TATCAAGGAA AATCAATTCT GGCTTCAAAA GATACACCCT TTCTGATGAA
481 TAAATGGAAA TACTACCTTG TCAATTTATG GCAATATCAT TTTTACGCGT GGTCCCAACC
541 AGGAAGGATC AATCTAAACC GATTAGGAAA GTATTCTCTT TACTTTTTGG GCTATTTTTC
601 AGACGTGCAA CTCACTTTTT CAGTGGTACG AAGTAAAATG CTCGAAAATT CATTTCTAGT
661 AAATACTGCT ATGAAGAAGC TCGAAACAAT AGTTCCCGTT ATTCCTTTGA TTGGATCGTT
721 GTCTAAAGCG AAATTTTGTA ACGTATTAGG ATATCCCGTT AGTAAGTCGA CCCGGACCGA
781 TTCATCAGAT TCTGATATTA TCGACCGATT TGTGCGTATA TGCAGAAATC TTTCTCATTA
841 TCACAGTGGA TCCTCAAAAA AAA 863

Plate 2: *Sonneratia caseolaris matK* gene

1 AAGTGTTGGA TTCAAAGCGG GTGTTAAAGA GTACAAATTG ACTTATTATA CTCCTGAATA
61 CGAAACCAAA GATACTGATA TCTTGGCAGC ATTCCGAGTA ACTCCTCAAC CTGGAGTTCC
121 GCCTGAAGAA GCAGGGGCCG CGGTAGCTGC CGAATCTTCT ACTGGTACATGGACAAGCGT
181 GTGGACCGAT GGACTTACCA GCCTTGATCG TTACAAAGGG CGATGCTACA ACATCGAGCC
241 CGTTCCTGGC GAAACAGATC AATATATCTG TTATGTAGCT TACCCTTTAG ACCTTTTTGA
301 AGAAGGTTCT GTTACTAACA TGTTTACTTC CATTGTAGGA AATGTATTTG GATTCAAAGC
361 CCTGCGTGCT CTACGTCTGG AAGATCTGCG AATCCCTCCT GCTTATATTA AAACTTTCCA
421 AGGCCCACCT CATGGGATCC AAGTTGAGAG AGATAAATTG AACAAGTATG GTCGTCCCCT
481 GCTGGGATGT ACTATTAAAC CTAAATTGGG GTTATCTGCT AAAAACTATG GTAGAGCATG
541 TTATGAATGT CTTCGCGGGG GACTTGATTT TACCAAAGAT GATGAGAACG TGAACTCCCA
601 ACCATTTATG CGTTGGAGAG ATCGTTTCTT ATTTTGTGCC GAAGCAATTT ATAAAGCACA
661 GGCGGAAACA GGTGAAATCA AAGGGCATTA CTTGAAT 697

***Avicennia marina, rbcL* gene**

1 AAGTGTTGGA TTCAAAGCGG GTGTTAAAGA GTACAAATTG ACTTATTATA CTCCTGAATA
61 CGAAACCAAA GATACTGATA TCTTGGCAGC ATTCCGAGTA ACTCCTCAAC CTGGAGTTCC
121 GCCTGAAGAA GCAGGGGCCG CGGTAGCTGC CGAATCTTCT ACTGGTACATGGACAACCGT
181 GTGGACCGAT GGACTTACCA GCCTTGATCG TTACAAAGGG CGATGCTACG ACATCGAGCC
241 CGTTCCTGGC GAAACAGATC AATATATCTG TTATGTAGCT TACCCTTTAG ACCTTTTTGA
301 AGAAGGTTCT GTTACTAACA TGTTTACTTC CATTGTAGGA AATGTATTTG GATTCAAAGC
361 CCTGCGTGCT CTACGTCTGG AAGATCTGCG AATCCCTCCT GCTTATATTA AAACTTTCCA
421 AGGCCCACCT CATGGGATCC AAGTTGAGAG AGATAAATTG AACAAGTATG GTCGTCCCCT
481 GCTGGGATGT ACTATTAAAC CTAAATTGGG GTTATCTGCT AAAAACTATG GTAGAGCATG
541 TTATGAATGT CTTCGTGGGG GACTTGATTT TACCAAAGAT GATGAGAACG TGAACTCCCA
601 ACCATTTATG CGTTGGAGAG ATCGTTTCTT ATTTTGTGCC GAAGCAATTT ATAAAGCACA
661 GGCTGAAACA GGTGAAATCA AAGGGCATTA CTTGAAT 697

Plate 3: *Avicennia officinalis, rbcL* gene.

1 AGAGGACAAT TTTTCACATT TAAATTTTGT GTTAGATGTA CTAATACCCC ACCCTGTCCA
61 TGTAGAAATC TTGGTTCAAA CTCTTCGCTA TTGGTTAAAA GATGCCTCTT CTTTGCATTT
121 ATTACGATTC TTTCTCAACG AGTATTGTAA TTGGAATAGT TTTATTTTGC CAAAGAAAGA
181 CGGTTCCTCT TTTTCAAAAA GAAATCAAAG ATTATTCTTA TTCTTATATA ATTCTCATGT
241 ATGGGAATAT GAATCCATTT TCGTCTTTCT ACGTAATCAA TCTGCTCATT TACGATCAAC
301 ATCTTCTGGA GTTCTTCTTG AACGAATCTA TTTCTATGGA AAAATGGAAC GTCTTGTGAA
361 CGTCTTTGTT AAGGTTAAGG ATTTTCGGTC GAACCCCAGG TTGATCAAGG AACCTTGCAT

421 GCATTATATT AGGTATCAAA GAAAATCCAT TCTGGCTTCA AAAGGGATGT CGCTTTTCAT
481 GAATAAATGG AAATGTTACC TTGTCACTCT TTGGCAATGG CATTTTTCGC TGTGGTTTCA
541 GCCAAGAAGG ATTTATATAA ACCAATTAGC CAATCATTCC TTTGAATTCT TGGGCTATCT
601 TTCAAGTGTG CGGATGAACC CTTCAGTGAT ACGGAGTCAA ATTCTCGAAA ATGCATTTCT
661 AATCAATAAT GCTATTAAGA AGTTTGATAC TCTTGTTCCA ATTATTCCTC TGATTATGTC
721 ATTGGCTAAA GCAAAATTTT GTAACGTATT AGGCCATCCT ATTAGTAAGC CGGTTTGGGC
781 TGATTTATCA GATTCTAATA TTATTGACCG ATTTCGACGT ATATGCAGAA ATTTTTCTCA
841 TTATCATAGC GGATCTTCCA CAAAAA 866

***Avicennia marina, matK* gene**

1 AGAGGACAAT TTTTCACATT TAAATTTTGT GTTAGATGTA CTAATACCCC ACCCTGTCCA
61 TGTAGAAATC TTGGTTCAAA CTCTTCGCTA TTGGTTAAAA GATGCCTCTT CTTTGCATTT
121 ATTACGATTT TTTCTCAACG AGTATTGTAA TTGGAATAGT TTTATTTTAC CAAAGAAAGA
181 TGGTTTCTCT TTTTCAAAAA GAAATCAAAG ATTATTCTTA TTCTTATATA ATTCTCATGT
241 ATGGGAATAT GAATCCATTT TCGTCTTTCT ACGTAATCAA TCTGCTCATT TACGATCAAC
301 ATCTTCTGGA GTTCTTCTTG AACGAATCTA TTGTTCTGGA AAAACGGAAC GTCTTGTGAA
361 CGTCTTTGTT AAGGTTAAGG ATTTTCGGTC GAACCCCAGG TTGATCAAGG AACCTTGCAT
421 GCATTATATT AGGTATCAAA GAAAATCCAT TCTGGCTTCA AAAGGGATGT CGCTTTTCAT
481 GAATAAATGG AAATGTTACC TTGTCACTTT TTGGCAATGG CATTTTTCTC TGTGGTTTCA
541 GCCAAGAAGG ATTTATATAA ACCAATTAGC CAATCATTCC TTTGAATTCT TGGGCTATCT
601 TTCAAGCGTG CGGATGAACC CTTTAGTGAT ACGGAGTCAA ATTCTCGAAA ATGCATTTCT
661 AATCAATAAT GCTATTAAGA AGTTTGATAC TCTTGTTCCA ATTATTCCTC TGATTGTGTC
721 ATTGGCTAAA GCAAAATTTT GTAACGTATT AGGCCATCCT ATTAGTAAGC CGGTTTGGGC
781 TGATTTATCA GATTCTAATA TTATTGACCG ATTTCAACGT ATATGCAGAA ATTTTTCTCA
841 TTATCATAGC GGATCTTCCA CAAAAA

Plate 4: *Avicennia officinalis, matK* gene

Conclusion

In conclusion, this study provides preliminary assessment data that will be useful for wider application of DNA barcoding in mangrove plants. With the help of primers, we found that *rbcL* and *matK* will be very useful for the barcoding of some mangrove plant species in Kerala. However, further protocol development to intensify clean DNA extraction, PCR amplification strategies, including the development of new primers, and local authenticated databases could play important roles in efficient utilization of plant barcoding.

Acknowledgment

The authors thank to CMS College Kottayam, Kerala for providing required facilities to carry out this research work.

References

1. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. PLoS ONE. 2011; 6(5):19254.
2. Hollingsworth PM, Forrest LL, Spouge M, Hajibabaei S, Ratnasingham M, van der Bank MW, *et al.* A DNA barcode for land plants. Proc. Nat. Acad. Sci. 2009; 106(31):12794-12797.
3. Kress JW, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. USA. 2005; 102:8369-8374.
4. CBOL–Plant Working Group, A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA. 2009; 106:12794-12797.
5. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Mol. Ecol. Notes. 2007; 7:355-364.
6. Blaxter M. Counting angels with DNA. Nature. 2003; 42:122-124. 2005; 435, 177.
7. Schindel DE, Miller SE. DNA barcoding a useful tool for taxonomists. *Nature*, CBOL Plant Working Group A DNA barcode for land plants. *Proc. Natl.* Acad. Sci. U. S. A. 2009; 06:12794-12797.
8. Roy S, Tyagi A, Shukla V, Kumar A, *et al.* Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. 2010; 1(5):13674.
9. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, *et al.* Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. 2008; 1(3):2802.
10. De Groot GA, During HJ, Maas JW, Schneider H, *et al.* Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. 2011; 1(6):16371.
11. Hall TA, Bio Edit a user 3 friendly biological sequence alignment editor and analysis program for Windows 95/98/NT", Nucleic Acids Symposium series. 1999, 41:95398.
12. Schori M, Showalter AM. DNA barcoding as a means for medicinal plants of Pakisthan. Pak. J. (Medicinal Plants: Conservation & Sustainable use), 2011, 43:1-4.