KFRI RESEARCH REPORT NO: 510
(Final Report of Project KFRI RP: 617/2011)

Development of institutional capability for DNA barcoding of life forms

E.M. MURALIDHARAN
Forest Genetics & Biotechnology Division

T.B. SUMA
Forest Genetics & Biotechnology Division

V.B. SREEKUMAR
Forest Ecology & Biodiversity Conservation Division

KERALA FOREST RESEARCH INSTITUTE
(An Institution of Kerala State Council for Science, Technology and Environment)
Peechi 680653, Kerala, India.

November 2015

ABSTRACT OF THE PROJECT PROPOSAL

Project Number	:	KFRI RP 617/2011
Title	:	Development of institutional capability for barcoding of life forms
Objectives	:	1. Development of protocols for DNA barcoding of various life forms 2. To demonstrate the capability of DNA barcoding techniques through case studies
Project period	:	July 2011 - June 2014
Funding agency	:	KFRI Plan Grants
Investigators	:	Dr. E.M. Muralidharan Dr. T.B. Suma Dr. V.B. Sreekumar
Research Fellows	:	Anoja Kurian BS Anoop

ACKNOWLEDGEMENTS

The authors are thankful to Dr. K.V. Sankaran, Dr. P.S. Easa former Directors, Kerala Forest Research Institute (KFRI) for providing the facilities to perform the project work and for their keen interest. The authors' record their gratitude for the financial support provided by Kerala State Council for Science Technology and Environment (KSCSTE), Govt. of Kerala, to carry out the research work at Kerala Forest Research Institute (KFRI). We are indebted to Kerala Forest Department for giving us permission to collect the required samples from the natural distribution zones in the Kerala part of the Western Ghats. The meticulous laboratory and field works carried out by Ms. Anoja Kurian and Mr. BS. Anoop during the tenure of the project is highly appreciated. We would also like to extend our thanks to the members of editorial committee for their helpful suggestions and comments on the project report.

CONTENTS

1. INTRODUCTION

Carl Linnaeus formalized biological classification, with binomial system of nomenclature in 1700's which assigns each organism a genus and species name for identification purposes. Since then, of the 8.74 million species exists on earth, only 1.5 million species have only been described using conventional taxonomy. This means that more than 80 % of the living organisms are yet to be discovered. The current rate of extinction of 100-1000 species per million species points out that thousands of plants and animals are lost each year without being identified (Mora et al., 2011). Classical taxonomy which requires trained expertise for species identification falls short in the race to catalogue biological diversity before it disappears. To help discovering the hidden biodiversity and in order to provide a useful and standardized supplementary tool for species identification, the sequence based molecular tool of 'DNA barcoding' has been proposed in 2003. DNA barcoding has been developed for providing rapid, accurate and automatable species identification using standardized DNA sequences (Hebert *et al*., 2003). DNA barcodes allow even non-experts to precisely identify species, even from small bits of specimens, independent of developmental stages or environmental variations.

The initial goal of DNA barcoding was to find a universal barcode gene locus for the identification of living organisms. This was first experimented in metazoans by Hebert et al. (2003) and has drawn attention of taxonomists worldwide (Miller, 2007). In 2004, barcoding project was formally initiated by the establishment of the Consortium for the Barcode of Life (CBOL), which aims to develop a standard protocol for barcoding of living organisms. Consequently, CBOL has developed certain criteria for recommending a locus for DNA barcoding purposes that includes suitability of the locus for a wide range of taxa (universality), high variation between species and conserved nature within species, among others. Ideal barcodes also should be routinely retrievable with a single primer pair and be short enough to ease PCR amplification (Hollingsworth *et al*., 2009).

Barcoding is now a well established technique for species identification in animals using cytochrome C oxidase subunit (*COX 1*) in mitochondria which is characterised by the presence of multiple copies, absence of indels, introns and ease

in PCR amplification (Folmer *et al*., 1994; Hebert *et al*., 2003). Contrary to this, no universal single barcode locus for land plants is available. mtDNA has low substitution rates when compared to plastid DNA in plants (Wolfe *et al*., 1987). To facilitate and formalize the selection of a plant barcode, a plant working group was constituted by the Consortium of Barcode of Life (CBOL) in 2004 with representation from the different research groups/consortia. After evaluation of several candidate plastid gene loci, twelve chloroplast gene regions were suggested as plant DNA barcodes. Because of the high species discrimination and universality, maturase K *(matK*) and of the ease of PCR amplification, ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*), were recommended as core barcodes (Hollingsworth *et al*., 2009).

The intergenic spacer sequences such as *trnH-psbA, atpF- atpH, psbK-psbI* were suggested by CBOL as supplementary barcodes with highest species discrimination for *trnH-psbA* of size ~300 bp up to 1000 bp for DNA barcoding in plants (Kress *et al*., 2005). Combinations of this region along with *matK* were also widely studied (Chase *et al*., 2007, Newmaster *et al*., 2008; Erickson *et al*., 2008; Wiel *et al*., 2009). *atpF- atp H* encoding the ATP subunits CFO I -CFO III (Dragger and Hallick, 1993) and *psbK - psbI* genes which encodes for two low molecular mass polypeptides, K-I of the photosystem II (Meng *et al*., 1991) were also added to the list later on. However, due to the inconsistency in getting bidirectional unambiguous sequences, this was only considered as a supplementary barcode by CBOL (Lahaye *et al*., 2008). Eventhough supplementary chloroplast barcode regions, *rpoB, rpoC1, rpoC2, accD, ycF5* and *ndhJ* have high universality, they were eliminated as barcode because of their low species discrimination ability (Guisinger *et al*., 2008; Tsumara *et al*., 1996).

nrITS spacer sequences, a part of the rDNA cistron has been proposed as a potential candidate for barcoding in plants as well as in fungi. The rDNA cistron is a multigene family encoding the nucleic acid core of the ribosome. All the units of this multigene family evolve in a concerted manner and higher level of overall sequence homogeneity exists among copies of the rDNA within a species (Chase *et al*., 2007). The availability of universal primers, presence of multiple copies in cells, high universality and good species discriminatory power makes *nrITS* a potential

candidate for barcoding in plants as well as fungi (Kress *et al*., 2005; Chen *et al*., 2010; O'Brien H, *et al*., 2005). Molecular identification through DNA barcoding of fungi using *ITS* sequences has become an integrated and essential part of fungal ecology research and has provided new insights into the diversity and ecology of many different groups of fungi (Seifert *et al*. 2007; Chase and Fay, 2009). Molecular analysis allows to study abundance and species richness of fungi at a high rate and more reliably than conventional biotic surveys (O'Brien *et al*., 2005).

Many researchers have suggested that a multi-locus method will be required to obtain adequate species discrimination in plants (Hebert *et al*., 2004; Kress and Erickson, 2007; Erickson *et al*., 2008; Kane and Cronk, 2008; Lahaye *et al*., 2008; CBOL Plant Working Group, 2009; Chase and Fay, 2009). CBOL recommended the two-locus combination of *matK* + *rbcL* as the best plant barcode with a discriminatory efficiency of only 72% (CBOL Plant Working Group, 2009). Various combinations of plastid loci have been proposed including *rbcL* + *trnH-psbA* (Kress and Erickson, 2007), *rpoC1* + *rpoB* + *matK* or *rpoC1* + *matK* + *trnH-psbA* (Chase *et al*., 2007) and *matK* + *atpF-atpH* + *psbK-psbI* or *matK* + *atpFatpH* + *trnH-psbA* (Pennisi, 2007). These combined barcodes exhibit higher species discrimination than single-locus approaches. Researchers have recently proposed the use of the whole-plastid genome sequence in plant identification (Erickson *et al*., 2008; Nock *et al*., 2011; Yang *et al*., 2013). However since this requires high cost next generation sequencing (NGS), the approach is not widely accepted.

DNA Barcoding has extensive applications and has been used for ecological surveys (Dick and Kress, 2009) and biosystematics including cryptic taxon identification (Hebert *et al*., 2004; Lahaye *et al*., 2008; Xue and Li, 2011), to support ownership or intellectual property rights (Stewart, 2005), authentication of Ayurvedic raw drugs (Hou et al., 2013), to link biological samples to crime scenes in forensics (Yoon, 1993; Coyle *et al*., 2005; Mildenhall, 2006), timber forensics (Lowe and Cross, 2011; Dev *et al*., 2014), biosecurity, wild life forensics and food forensics (Galimberti *et al*., 2012; Huxley-Jones *et al*., 2012).

At the Kerala Forest Research Institute, research in biodiversity particularly in species discovery has been a priority since we are located in the hotspot of the Western Ghats. The technique of DNA barcoding is expected to support this endeavour in a substantial manner. Precise species identification is also required by researchers within the institute and by others in the region who depend on us for the species identification needs. The project was taken up to build the competence within the institute to undertake routine DNA barcoding work in all life forms.

To demonstrate the capability and the application of the DNA barcoding technology, two case studies were taken up. The difficulties encounter when conventional methods are used for identification of plant species and the potential of DNA barcoding to offer solutions was demonstrated.

2. OBJECTIVES

A precise molecular tool which can supplement the traditional species identification procedure, particularly in case of taxonomic complexities has long been envisaged by researchers. This particular project aims at standardizing a sequence based molecular tool, DNA barcoding for addressing biosystematic issues across all the life forms.

This study therefore has been undertaken with the following objectives:

- Standardisation of protocols for DNA barcoding of various life forms
- To demonstrate the capability of DNA barcoding techniques through case studies

3. STANDARDIZATION OF PROTOCOLS FOR DNA BARCODING OF VARIOUS LIFEFORMS

DNA barcoding is based on the principle that a short standardized sequence can distinguish individuals of a species, since between species genetic variation exceeds within species. Identification of an unknown specimen is performed through the identification of sequences from the recommended standard barcode regions. The species of a query sample (an unknown specimen) can be identified based on the similarity index of the developed barcode sequences to a reference sequence (from a voucher specimen) in the NCBI genbank public domain. DNA barcoding workflow begins with the collection of a tissue sample from the specimens to be barcoded, DNA extraction from the tissue, PCR amplification of the barcode region using universal primers, screening of the PCR products and an automated DNA sequencing to generate the barcode sequences. With the recent technological innovations and automation through high-throughput sequencing, the entire process can be completed at a reduced cost, within a short span of time (Hajibabaei *et al*., 2005).

PROTOCOL FOR DNA BARCODING IN PLANTS

Genomic DNA extractions can be performed from fresh, dried, degraded or fossil specimens for barcoding purposes. DNA extractions were performed from fresh as well as dried specimens adopting standard CTAB method (Doyle and Doyle, 1987). For each silica gel dried sample, genomic DNA was extracted from 50-100 mg tissue which was ground using a mortar and pestle with liquid nitrogen. Extractions were performed using the modified CTAB method as well as by commercial DNA isolation kit (Qiagen, USA) with required modifications. In the CTAB method, the powdered tissue sample was transferred to pre-warmed (65^{o}C) CTAB extraction buffer. RNase (10 mg/ 100 ml) was added to this slurry and the samples were incubated at 65^{o}C on a water bath for two hours with gentle inversion for every 5 minutes. After incubation, the lysate was allowed to cool to room temperature. An equal volume of chloroform: isoamylalcohol (24:1) was added to the lysate and mixed properly by inversion and centrifuged at

8,000 rpm for 5 minutes. The supernatant was pipetted out carefully and transferred to a 1.5 ml microcentrifuge tube. A half volume of 5M NaCl and double volume of ice cold ethanol were added to precipitate the DNA. The solution was centrifuged at 5000 rpm for 10 minutes. The DNA pellet was air dried to remove the residual ethanol. The pellet was dissolved in sterile water and stored in -20˚C deep freezer for further use. The isolated genomic DNA was subjected to electrophoresis to visualise the DNA. The samples were separated on 1.5 per cent agarose gel and then stained in ethidium bromide and visualised under UV transilluminator (Fig.1). DNA was quantified using a Nanodrop spectrophotometer (Nanodrop Fisher Thermo., USA).

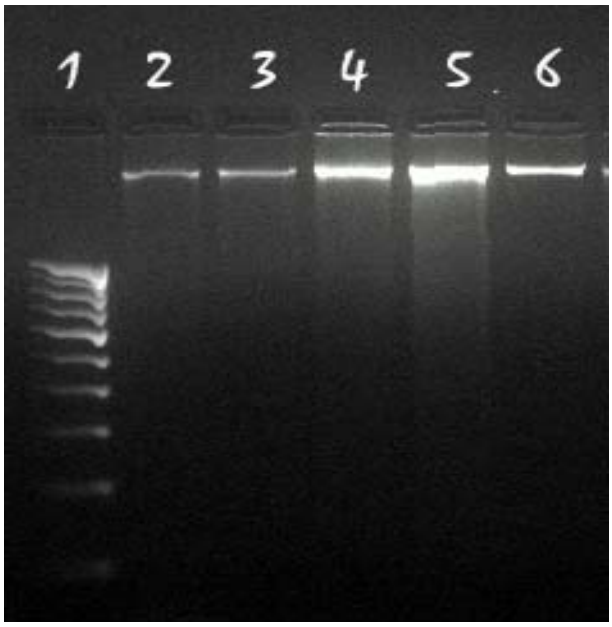

Figure 1. Total genomic DNA of plants

The PCR amplifications were carried out for the CBOL recommended core DNA barcodes *viz. rbcL, matK* as well as for the supplementary barcodes *trnH-psbA*, nuclear *ITS* (Internal Transcribed Spacer) regions *(ITS1 and ITS2*), *rpoB* and *rpoC* barcode regions (CBOL, 2009). Reported universal primer sequences for plant barcoding purposes were selected for the specific barcode loci (Table 1) (Vijayan and Tsou, 2010; China Plant BOL Group, 2011). PCR was performed to amplify the gene sequences in 20 µl reaction volume, containing 1 µl genomic DNA, 2 µl primer (20 picomoles), 200 µM dNTPs, 2 µL of 10x Taq buffer and 2U of

Taq DNA polymerase. DNA was amplified by a programmable thermal cycler PTC 200 (MJ Research Inc., USA). PCR reaction followed is given in Table 1. Annealing temperature varies from primer to primer as well as species to species.

PCR products were resolved on 2 % agarose gels stained with ethidium bromide and then purified using Nucleospin Elution Kit (Machery-Nagel, U.S.A.) (Fig. 2). PCR reaction has been scaled up to 50 µL volume for the purpose of elution. Elution of the PCR product was done by Nucleospin gel and PCR clean up kit as per the manufacturer's protocol (Machery-Nagel, U.S.A.). DNA sequencing was performed for the eluted PCR products in both forward and reverse directions employing Sanger's dideoxy chemistry.

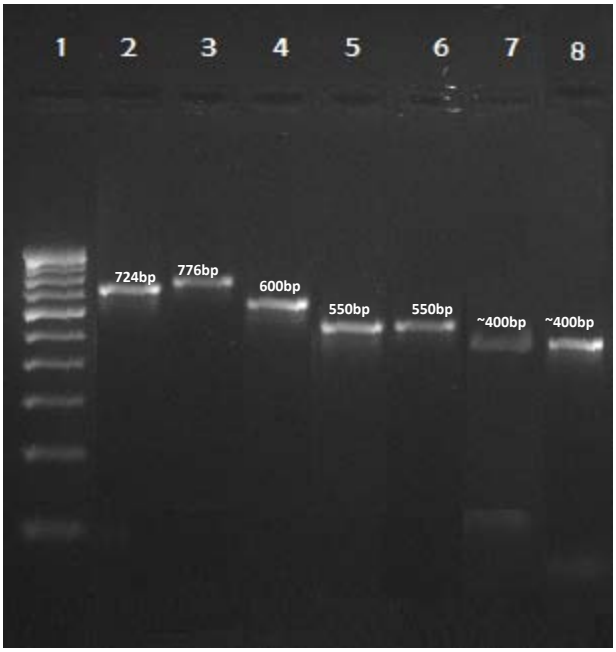

Figure 2. PCR products of plant barcode regions (lane 1-100bp ladder, lane 2-rbcL, lane 3-matK, lane 4-trnH-psbA, lane 5-rpoB, lane 6-rpoC, lane 7-ITS1, lane 8-ITS2)

Table 1. Primer sequences of seven candidate DNA barcodes and their reaction conditions

Barcode region	Primer	Primer sequence 5′-3′	Reaction conditions
rbcL	1F 724R	ATGTCACCACAAACAGAAAC TCGCATGTACCTGCAGTAGC	94°C 5 min. 94°C 1 min. 60°C 1min. 72°C 1min. 35 cycles 72 °C 10 min.
matk	472F 1248R	CCCRTYCATCTGGAAATCTTGGTT GCTRTRATAATGAGAAAGATTTCTGC	94°C 10 min. 94°C 1 min. 57°C 1min. 72°C 1.5 min. 35 cycles 72 °C 12 min.
psbA-trnH	*psbA* *trnH*	GTWATGCAYG AACGTAATGCTC CGCGCATGGTGGATTCACAATCC	94°C 5 min. 94°C 1 min. 60.5°C 1min. 72°C 1 min. 35 cycles 72 °C 10 min.
rpoC	*rpoC* *rpoC*	GGCAAAGAGGGAAGATTTCG CCATAAGCATATCTTGAGTTGG	94°C 5 min. 94°C 1 min. 57°C 1min. 72°C 1min. 35 cycles 72 °C 10 min.
rpoB	*rpoB F* *rpoBR*	AAGTGCATTGTTGGAACTGG GATCCCAGCATCACAATTCC	94°C 5 min. 94°C 1 min. 60°C 1min. 72°C 1min. 35 cycles 72 °C 10 min
ITS1	*ITSF* *ITSR*	TCCGTAGGTGAACCTGCGG GCTGCGTTCATCGATGC	94 °C-5 min. 94 °C-1 min. 63 °C -1 min. 72 °C-1 min. 35 cycles 72 °C-10 min.
ITS2	*ITS3F* *IT4R*	GCATCGATGAAGAACGCAGC TCCTCCGCTTATTGATATGC	94 °C-7 min. 94 °C-1 min. 59 °C -1 min. 72 °C-1 min. 35cycles 72 °C-10 min.

PROTOCOL FOR DNA BARCODING IN INSECTS

Insects comprise about 1 million species and are the most abundant life forms on earth. India has listed among the top ten mega diversity nations of the world with respect to insect diversity, comprising 7.10 % of the world insect fauna (Ghorpade *et al*., 2010). Diversification of the insect fauna is more extensive when compared to other animal phyla and hence species identification of insects is always a challenging task for the conventional taxonomists. Insect DNA barcoding refers to the identification of an insect species, by the use of 648bp conserved DNA sequences from the specific mitochondrial gene *CO1* (cytochrome oxidase 1), and provides a supplementary tool for insect species identification (Hebert *et al*., 2003).

Insect sample was collected and ground using mortar and pestle with 150 µl SDS Buffer [200 mM Tris- HCl (pH 8.0), 25 mM EDTA, 250 mM NaCl and 0.5 % sodium dodecylsulfate]. This was transferred into a 2 ml vial and 350 µl SDS buffer was added. RNase A solution (100 mg/ml) was added to the homogenate. Samples were incubated at 37 °C for 1 hour. After incubation, proteinase K (20mg/ ml) was added and mixed well. This was again kept for incubation at 50 °C for 1 hour. The homogenate was then extracted with 240 µl of chloroform:isoamyl alcohol (24:1) and was centrifuged for 10 minutes at 12,000 g. Supernatant was transferred to a new tube. To precipitate DNA, 500 µl chilled absolute ethanol was added and tube was centrifuged at 12,000 g for 15 min. The pellet was washed with 500 µl of 70% ethanol and centrifuged at 12,000 g for 3 minutes to remove residual salts. The pellet was dried in a vial at 37 °C for 30 minutes in an incubator or air dried at room temperature for overnight. The pellet was resuspended in 50 µl sterile double distilled water. The samples were separated on 1.5 % agarose gel and then stained in ethidium bromide and visualized under UV transilluminator (Fig.3). DNA was quantified using a Nanodrop spectrophotometer (Nanodrop, Fisher Thermo., USA).

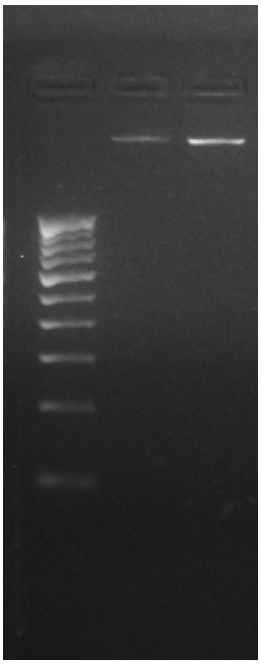

Figure 3. Total genomic DNA of insect

The *COI* barcode region was PCR amplified using the reported primer combinations C1-J-1718 F (5′-GGAGGATTTGGAAATTGATTAGTGCC- 3′) and C1-N-2191 (5′ - CCCGGTAAAATTAAAATATAAA -3′) (Folmer *et al.,* 1994). PCR was performed to amplify the *COI* sequence in 20 µl reaction volume, containing 1 µl genomic DNA, 2 µl primer (20 picomoles), 200 µM dNTPs, 2 µL of 10x Taq buffer and 2U of Taq DNA polymerase. DNA was amplified by a programmable thermal cycler PTC 200 (MJ Research Inc., USA). PCR reaction was performed with the following conditions of initial denaturation of 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 62 °C for 1 min, 72 °C for 1 min and a final extension of 72 °C for 10 min.

PCR products were resloved on 2 % agarose gels stained with ethidium bromide. PCR reaction has been scaled up to 50 µL volume for the purpose of elution and then purified using Nucleospin Elution kit (Machery-Nagel, U.S.A.) (Fig. 4). DNA sequencing was performed for the eluted PCR products in both forward and reverse directions employing Sanger's dideoxy chemistry.

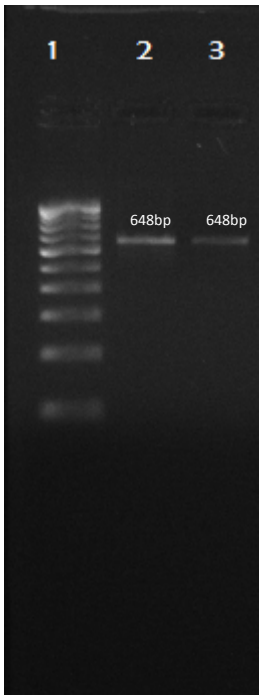

Figure 4. PCR products of insect *CO1* barcode region (lane1-100 bp ladder, lane 2 & 3- *COI*)

PROTOCOL FOR DNA BARCODING IN FUNGI

Internal transcribed spacer region (*ITS*) of nuclear DNA, a routinely used sequence data for analyzing fungal diversity in environmental samples, has recently been recommended as the standard region for fungal DNA barcoding. DNA barcoding for fungal species identification has immensely contributed to fungal ecology research and has generated information about the diversity and ecology of many different groups of fungi. This sequence based tool has the specific advantage of studying the ecology of fungal species even from the inconspicuous mycelial stage in the absence of fruiting bodies (Anderson and Cairney, 2004; Chase and Fay, 2009). Environmental barcoding employing sequence based barcodes has significantly advanced over the past decade as a reliable tool to study abundance and species richness of fungi (O'Brien *et al*., 2005). The internal transcribed spacer (*ITS*) of nuclear DNA (nrDNA) has also been proposed as the official primary barcoding marker for fungi (Deliberation of 37 mycologists from 12 countries at the Smithsonian's Conservation and Research Centre, 2007) (Bellemain *et al*., 2010).

For the development of DNA barcode, fungal mucelium was cultured on Potato Dextrose Agar (PDA) in a petridish. Fungal mycelia were scraped directly

from the surface of the fully grown PDA plate. The mycelial mass was ground using liquid nitrogen in a mortar and pestle to form a fine powder. After grinding, 2 ml of extraction buffer [100 mM Tris- HCl (pH 8.0), 20 mM EDTA, 0.5 M NaCl and 1 % Sodium dodecylsulfate] was added and the slurry was transferred to a 2 ml vial. 8 µl RNase was added and incubated at 65 °C for 20 minutes with a gentle inversion in every 5 minutes. After incubation, an equal volume of chloroform:isoamyl alcohol (24:1) was added and centrifuged at 16,000 g for 5 minutes. Aqueous phase was carefully transferred to a new micro centrifuge tube and the previous step was repeated. Final aqueous phase was transferred to a new tube and mixed with half volume of 5 M NaCl and double volume of ice cold ethanol. Samples were incubated at room temperature for 10 min and centrifuged for 15 min at 16,000 g to recover the pellet. After rinsing the pellet with 95 % ethanol, the pellet was allowed to air dry briefly and was subsequently resuspended in 50 µl sterile double distilled water. The isolated genomic DNA was subjected to electrophoresis to visualize the genomic DNA. The samples were separated on 1.5 % agarose gel and then stained in ethidium bromide and visualized under UV transilluminator (Fig. 5). DNA was quantified using a Nanodrop spectrophotometer (Nanodrop, Fisher Thermo., USA).

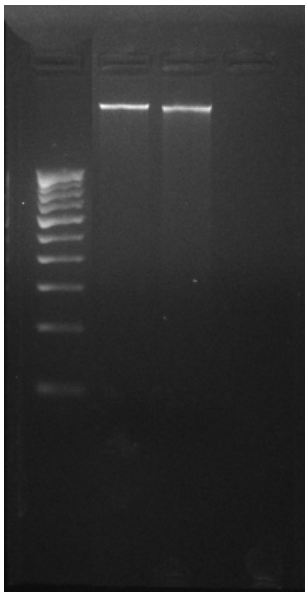

Figure 5. Total genomic DNA of fungi

ITS regions (*ITS1* and *ITS2*) considered as the universal barcode of fungi were amplified using the recommended primers, *ITS*1F TCCGTAGGTGAACCTGCGG, *ITS*2R GCTGCGTTCATCGATGC for *ITS*1 region and *ITS*3 GCATCGATGAAGAACGCAGC and *ITS*4 TCCTCCGCTTATTGATATGC for *ITS*2 region used primers (White *et. al.*, 1990). PCR was performed to amplify the gene sequences in 20 µl reaction volume, containing 1 µl genomic DNA, 2 µl primer (20 picomoles), 200 µM dNTPs, 2 µL of 10x Taq buffer and 2U of Taq DNA polymerase. DNA was amplified by a programmable thermal cycler PTC 200 (MJ Research Inc., USA). For *ITS*1 barcode region, the PCR parameters were initial denaturation of 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min, and a final extension of 10 min at 72 °C. *ITS2* barcode region was amplified using PCR parameters of initial denaturation of 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 59 °C for 1 min and 72 °C for 1 min, and a final extension of 10 min at 72 °C for 5 min.

PCR products were resolved on 2 % agarose gels stained with ethidium bromide and then purified using Nucleospin elution kit (Machery-Nagel, U.S.A.) (Fig. 6). PCR reaction has been scaled up to 50 µL volume for the purpose of elution. DNA sequencing was performed for the eluted PCR products in both forward and reverse directions employing Sanger's dideoxy chemistry.

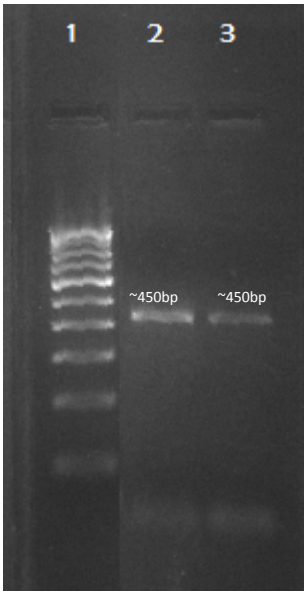

Figure 6. PCR products of fungal *ITS* regions (lane 1-100 bp ladder, lane 2-*ITS1*, lane 3-*ITS-2*)

SEQUENCE DATA ANALYSIS

The raw sequence chromatograms obtained after sequencing were edited using *BioEdit* Software v.7.0. The edited sequences were used for multiple sequence alignment in *CLUSTAL X* (Thompson *et al.,* 1997). The sequences after alignment were then subjected to BLAST sequence similarity search in NCBI GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The Kimura 2-parameter (K2P) nucleotide substitution model was used for calculating pairwise distances with *MEGA v.6.0* (Tamura *et al.,* 2013). Average pairwise intraspecific divergence (K2P distance), mean theta (θ), and average coalescent depth were used to characterize intraspecific divergence, as described in Meyer and Paulay (2005). Intraspecific as well as interspecific divergence parameters were compared for all the analyzed barcode regions along with their possible combinations. An ideal barcode should have high interspecific but low intraspecific divergence. DNA barcoding gaps were calculated by comparing intra and interspecific genetic distances (Kress and Erickson, 2008). Wilcoxon signed rank tests for testing the significance of inter-intraspecific divergence were also performed (Lahaye *et al.* 2008). Positive species identification by a candidate barcode was counted only if multiple individuals formed a monophyletic cluster in the Neighbor-Joining trees (*MEGA v.6.0*), according to the method of Starr *et al.* (2009). Clade support was estimated with 1,000 heuristic bootstrap replicates (100 random addition cycles per replicate, with tree bisection-reconnection and branch-swapping) to test the reliability of the inferred phylograms.

CASE STUDY 1

4. IDENTIFICATION OF MARKET ADULTERANTS IN EAST INDIAN SANDALWOOD USING DNA BARCODING

Sandalwood, the most valuable tropical hardwood, is mainly extracted from *Santalum album* L. (East Indian sandalwood) and *S. spicatum* R.Br (Australian sandalwood) (Hewson and George, 1984). In India, *S. album* is distributed naturally over 9400km^2 on the Deccan Plateau in the states of Karnataka, Kerala and Tamil Nadu (Srimathi *et al*., 1995). Indian sandalwood is the most appreciated in the world, for its fragrant heartwood and oil used for incenses, soaps, creams, perfumes, carvings, paintings and for religious rituals for over 4000 years. The global supply of natural Indian sandalwoods has reached critical levels due to illegal felling and over-harvest with the tree being recognised as a vulnerable species on the World Conservation Union's (IUCN) Threatened Species Red List (IUCN 2009). In India, official production of sandalwood has been declined significantly over the years due to the depletion of the existing natural resources (Ananthapadmanabha, 2012).

To supplement internal demands, which are 5,000 tonnes per annum, India has to adopt sandalwood import in substantial quantity and the import of substitutes especially the *'Tanzanian* or *African sandalwood'*, *Osyris lanceolata* Hochst. & Steud. was more than 3,000 tonnes in 2006 (Ananthapadmanabha, 2012). Consequently, the quality of superior sandalwood oil was affected and Government had to impose a ban on import of sandalwood in India. The East Indian sandalwood oil is very unique with a large number of essential molecules which are extremely difficult to replace with a synthetic substitute. The common sandalwood adulterants available in commercial market of India are the 'Nepal sandalwood', *Osyris wightiana* Wall. ex Wight (Santalaceae) and *'Indian bastard sandal', Erythroxylum monogynum* Roxb. (Erythroxylaceae) (Anupama *et al*., 2012). It is difficult to distinguish between *O. wightiana* and 'East Indian sandalwood' in grain or colour with the scent of the two woods differing only slightly. The heart wood is faintly fragrant and used for adulterating sandalwood (Shyaula, 2012). The heartwood of *E. monogynum*, is reddish-brown and with a pleasant odour (Oyen and Dung, 1999).

A number of lawsuits registered with regard to the adulteration of sandalwood and oil, get stalled due to the lack of technical tools for the correct identification of the source wood (Bhat *et al*., 2006). The ability to track or identify timber resources of economic value is therefore critically essential for the effective management and appropriate regulation of timber trade. DNA barcoding has been suggested as a valid technical tool even in the court of law (Lowe and Cross, 2011) and can replace other timber species verification means like wood anatomy (Bhat *et al.,* 2006), chemical composition (Deguilloux *et al*., 2002) and isotopic fingerprinting (Boner *et al*., 2007), which in majority cases are unable to resolve down to the species level. DNA barcoding, as facilitated by the Consortium for the Barcode of Life (CBOL) and the International Barcode of Life Project (iBOL), has developed a standardized set of gene regions and central database of reference samples (Barcode of Life Database – BOLD) to identify many species around the world (Hebert *et al*., 2003; Costion *et al*., 2011). Developments over the last 10 years in the extraction of DNA from dried and processed wood samples, leads to the development of reliable and efficient tracing methods particularly for forest trees affected by illegal logging and adulteration activities (Deguilloux *et al*., 2002; Liepelt *et al*., 2006; Rachmayanti *et al*., 2006; Lowe *et al*., 2010).

This study demonstrates the ability of DNA barcoding, using three standard plant barcodes, to distinguish East Indian sandalwood from its adulterant species *viz*. *O. wightiana* and *E. monogynum*.

MATERIAL AND METHODS

DNA extraction

Leaf samples of *S. album* as well as of the adulterants *O. wightiana* and *E. monogynum* were collected from the Marayur sandalwood reserve forest and Chinnar Wildlife Sanctuary located on the leeward side of the Western Ghats in the SouthWest region of India (10°15′N latitude and 77°11′E longitude). Voucher specimens for all the species were deposited at the KFRI herbaria maintained in the institute. Genomic DNA was extracted from silica gel dried leaves using DNeasy Plant Mini Kit

(Qiagen, U.S.A.) in accordance with the manufacturer's protocol with slight modifications.

PCR amplification and Sequencing

PCR amplifications were performed as per standard protocols in five samples each for the three standard barcode loci *viz. rbcL* (ribulose bis phosphate carboxylase) and *matK* (maturase K) as well as for the non-coding *trnH-psbA* intergenic spacer region of the plastid genome. Polymerase chain reaction (PCR) amplifications were carried out using CBOL (2009) recommended universal primers for *rbcL, matK* and *trnH-psbA* listed in Table 2.

Table 2. Barcode loci and primer information

Barcode loci	Primer information	Primer Annealing Temperature
rbcL	1F ATGTCACCACAAACAGAAAC 724R TCGCATGTACCTGCAGTAGC	60^0C
matK	472F CCRTCATCTGGAAATCTTGGTT 1248R GCTRTRATAATGAGAAAGATTCTGC	60^0C
TrnH-psbA	trnHF CGCGCATGGTGGATTCACAATCC psbA RGTWATGCAYGAACGTAATGCTC	60.8^0C

Sequence analysis, barcode design and phylogeny construction

Please refer to the methodology provided in page 19.

Sequences of *rbcL*, *matK* and *trnH-psbA* of *Buxus sempervirens* available in the NCBI nucleotide library (Accession nos. HE963366.1, HE966887.1, HE966523.1, respectively) were also included in the distance analysis. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were 1358 positions in the final dataset.

RESULTS

All selected barcode regions were amplified with the COBOL primers and the eluted products gave good read length in the forward direction. Molecular diagnosis of sandalwood adulterants were performed by a multi-locus DNA barcoding approach. The final aligned partial sequences had a length of 679 bp, 710 bp and 443 bp for *rbcL*, *matK* and *trnH-psbA* intergenic spacer respectively (NCBI Genbank accession

nos. KC503279–KC503287). Unique barcodes using *rbcL*, *matK* and *psbA-trnH* were generated for sandalwood and its adulterants through BOLD (http://www.boldsystems.org/) and the *rbcL* barcodes are provided in Figure 7.

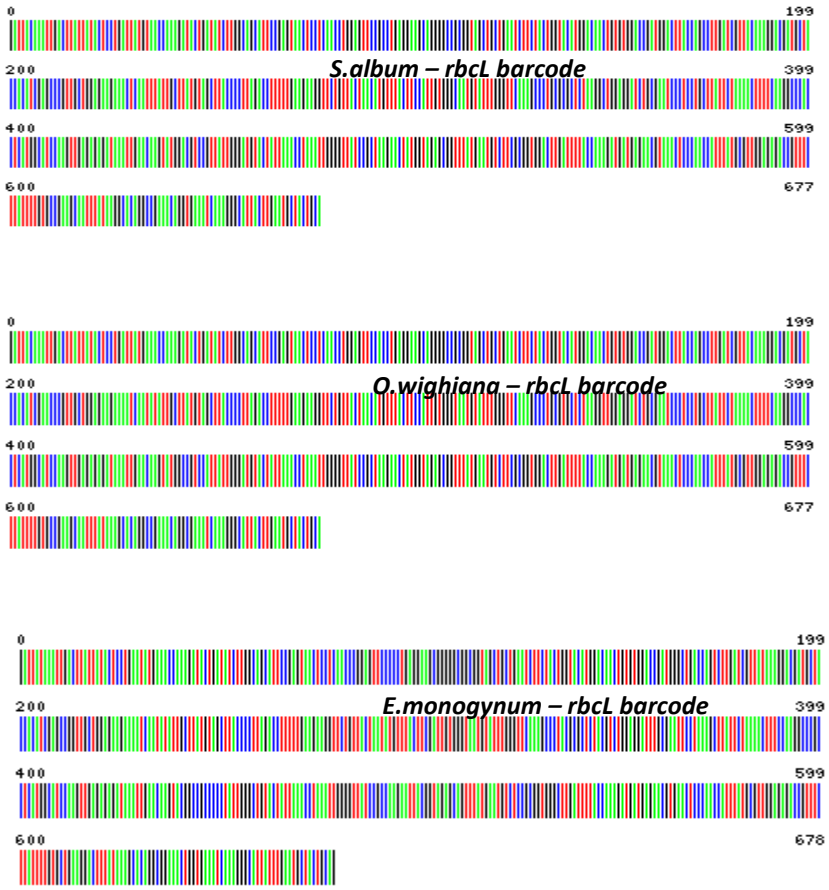

Figure 7. *rbcL* DNA barcodes for *S. album*, *O. wightiana* and *E. monogynum*

Of the total number of aligned nucleotides, conserved sites, variable but uninformative and parsimony informative sites showed variation among the three standard barcodes (Table 3). Generally, a very low percentage of parsimony informative characters of 1% or 2% were observed with all the three analysed barcodes. Transitions were more common than transversions except for *trnH-psbA* and hence transition/trasversion ratio was >1 in *rbcL* and *matK*. We manually examined the sequence alignments for SNPs and found unique SNPs in sandalwood and its adulterants (Fig. 8). Unique SNPs were observed with *rbcL* and *trnH-psbA* of *E. monogynum* and also in the *matK* barcode of *O. wightiana*.

Table 3. SNPs and Tajima's test (1993) statistics for the three barcode loci

Loci	Conserved	Variable	Parsimony informative	Overall transition/trasversion ratio (R)
rbcL	617	61	9 (1%)	1.6
matK	560	110	15 (2%)	1.4
trnH-psbA	242	153	7 (2%)	0.7

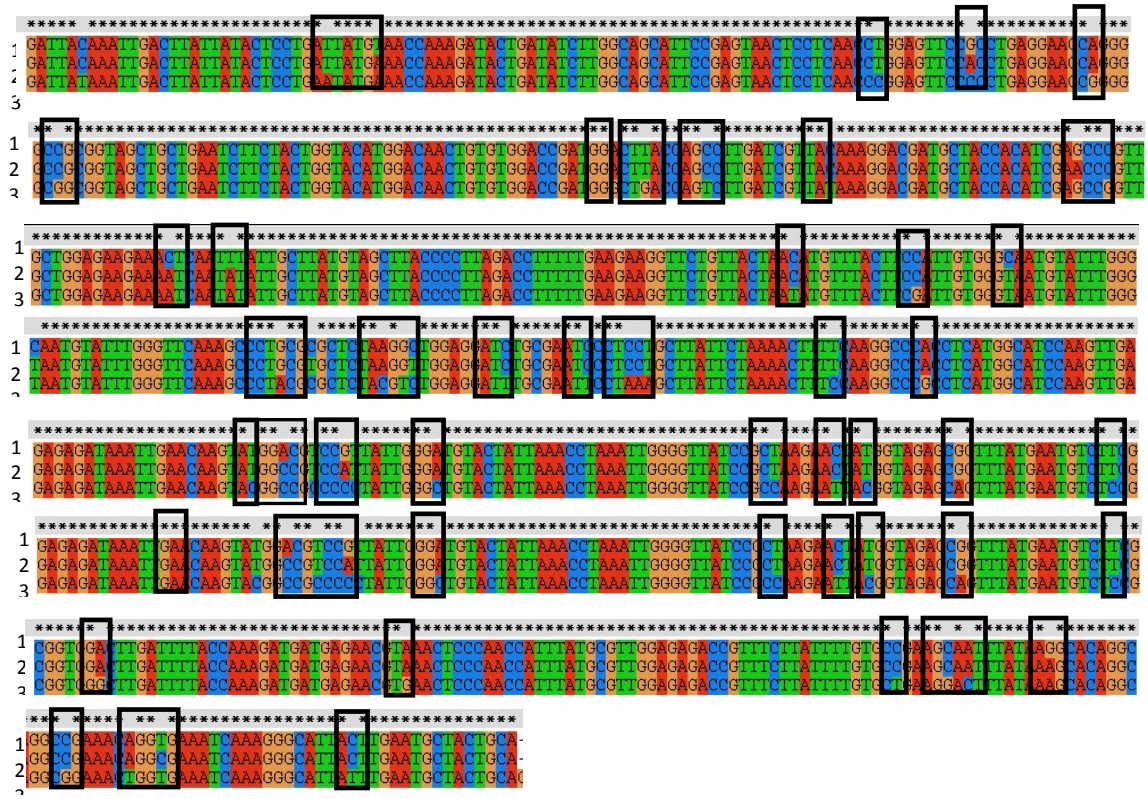

Figure 8. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) (highlighted in boxes) in *rbcL* sequences of *S. album* (1), *O. wightiana* (2) and *E. monogynum* (3)

Estimates of genetic distances among sandalwood and its adulterants calculated using the Kimura 2-parameter method in MEGA 4 (Kimura, 1980) with a total of 1358 positions in the combined dataset. The genetic distance between *S. album/ O. wightiana* was 0.037 and that of *S. album/ E. monogynum* was 0.080. Aligned molecular sequence data from all the three barcodes were used to generate a dendrogram which depicted the genetic relationship among sandalwood (Fig. 9) and its common adulterants. *S. album* and *O. wightiana* formed the most genetically similar cluster.

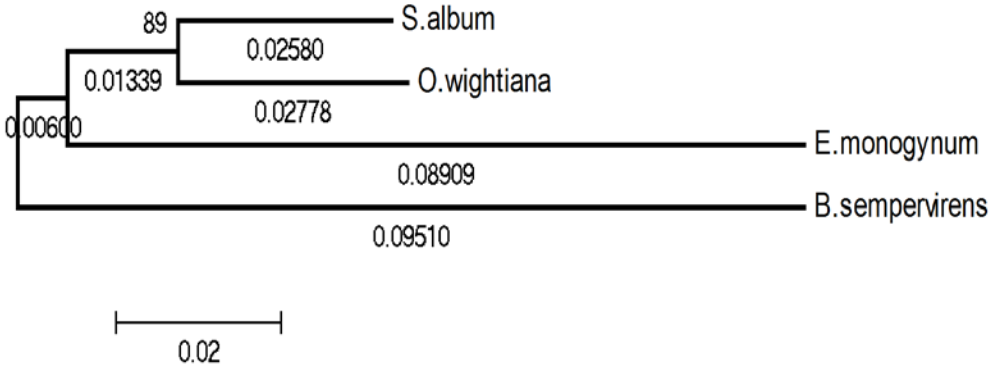

Figure 9. Cladogram depicting genetic relationships of sandalwood and its adulterants (values indicate genetic distances)

DISCUSSION

DNA barcoding offers numerous potential applications in the field of forest forensics including the identification of timber species in the timber trade industry as well as to monitor the illegal trade of wildlife, such as protected or endangered species or to identify the species of origin in the case of commercially processed food (Shivji *et al*., 2002). DNA barcoding has been employed to identify an endangered tropical timber species, *Gonystylus bancanus* (Asif and Cannon, 2005) as well as to identify wood of *Shorea* species and its origin by the Forestry and Forest Products Research Institute (FFPRI, Tsukuba), the Forest Research Centre, Sabah (FRC) as well as by the Forest Research Institute Malaysia (FRIM) (Fuji, 2007). In the Royal Botanic Gardens, Kew, DNA barcoding has been employed to detect illegal logging and origin of endangered timber species of the genera *Aquilaria* and *Cedrela* (Chase *et al*., 2005).

The present study demonstrated the usefulness of DNA barcodes to distinguish common wood adulterants from the original sandalwood. Out of the three barcodes tested, unique SNPs were identified with two of the standard barcodes *viz. rbcL*, *trnH-psbA* in *E. monogynum* and with *matK* barcode in *O. wightiana* which can be efficiently used for the genetic identification of sandalwood adulterants. SNPs have been used previously also for the timber tracking of wood adulterants (Degen and Fladung, 2008). Among the two adulterants, *O. wightiana* was found to be more genetically similar to *S. album* as revealed in the distance based dendrogram. Earlier studies on genetic grouping based on rDNA sequences revealed close genetic affinities of *S. album* and *O. wightiana* (Anupama *et al.,* 2012). In the barcodes, transitions, which mostly do not result in amino acid changes, were more frequent than transversions. Since barcode genes are more conserved, transitions which do not make any change in the amino acid sequence occur more frequently and DNA sequences in general are more likely to undergo transition than transversion with the ti/tv ratio above unity (Holmquist, 1983).

The Indian species of sandalwood (*S. album*) commands the highest price of upto A$105,000 per tonne at auction due to its superior oil yields and long-term market acceptance. An unfortunate consequence due to these rising prices is that some dealers are even using polished and scented sticks of most inferior quality woods which are difficult to identify by traditional means (Page *et al*., 2012). The ability of the standard DNA barcodes to differentiate sandalwood adulterant species demonstrated in the present study can thus be effectively utilized as a diagnostic adulterant monitoring tool by enforcement agencies (customs/border inspection) and traders of sandalwood and its products. The new technological advancements in the timber forensics have minimised the need for basic skills as well as dramatically reduced the cost, so that it can complement an integrated approach even in the court of law.

CASE STUDY 2

5. SPECIES DISCRIMINATION THROUGH DNA BARCODING FOR THE GENUS *SALACIA* OF THE WESTERN GHATS IN INDIA

The tropical genus *Salacia* L. (Celastraceae) comprises of 150 species, 21 of which are distributed in India and 8 of them in the Kerala region of the Western Ghats (Sasidharan, 2004; Ramamurthy and Venu; 2005, Udayan *et al*., 2012). The genus includes scandent or erect shrubs, small trees and lianas with opposite/ subopposite leaves, fascicled/ panicled flowers with intrastaminal conical disc and drupaceous/ subglobose fruits. Plants of this genus are used in traditional medicine as acrid, bitter, thermogenic, urinary, astringent, anodyne, anti-inflammatory, depurative, vulnerary, liver tonic and stomachic. The roots have also been used for treating itch, rheumatism, diabetes and venereal diseases such as gonorrhoea (Paarakh *et al*., 2008). Many *Salacia* species serve as a source of pharmaceutical chemicals used in the Ayurvedic system of medicine in India.

Species identity within the genus is conventionally established based on morphological features. However, identification of species within the genus *Salacia* is often difficult due to the high level of morphological similarities. To ensure quality and assured therapeutic effects in the raw drugs extracted from the genus, proper identification at the species level is critical. *Salacia reticulata* is the most widely recommended species for treating diabetes in Ayurveda, but availability of the species has diminished considerably in India (Udayan and Pradeep, 2012). Several other species of *Salacia* are heavily extracted from the natural areas of distribution and the roots sold in the raw drug market as that of *S. reticulata*, 'ponkarandi or ekanayakam' (in local language) and go undetected due to difficulties in the correct identification of species (Udayan and Pradeep, 2012).

In this study, three plastid loci *viz.* ribulose-bisphosphate carboxylase (*rbcL),* maturase K *(matK)* and the non-coding chloroplast intergenic spacer *(trnH-psbA)* and a nuclear locus *viz*. the internal transcribed sequence (*ITS2),* were employed to develop species specific barcodes for identifying eight species and one variety of the genus *Salacia*. Our objective was to arrive at suitable barcodes that can accurately

and efficiently identify the botanical origins of the raw drugs out of the genus *Salacia.*

MATERIALS AND METHODS

Taxon sampling

Eight species and one variety of the genus *Salacia* (*S. beddomei*, *S. chinensis*, *S. fruticosa*, *S. macrosperma*, *S. malabarica*, *S. oblonga*, *S. vellaniana*, *S. agasthiamalana* and one variety *S. oblonga* var. *kakkayamana* were included in the present analysis. Six species and one variety of the genus are endemic to the Western Ghats region of peninsular India, either with restricted distribution only to the type locality (*S. agasthiamalana, S. vellaniana* and *S. oblonga* var. *kakkayamana*) or to a few sparsely distributed populations (*S. beddomei, S. macrosperma* and *S. malabarica*). *S. fruticosa* is widely distributed throughout the Western Ghats with a few individuals in each distribution zones. Non-endemic species, *S. oblonga* and *S. chinensis* were collected from four and six localities in the Western Ghats respectively. A map indicating the taxon sampling locations of the eight *Salacia* species and one variety is provided as Figure 10. Voucher specimens for all the species were deposited in the Kerala Forest Research Institute (KFRI) herbarium (www.kfriherbarium.org).

DNA extraction, amplification and sequencing

Genomic DNA was extracted from silica gel dried leaves using DNeasy Plant Mini Kit (Qiagen, U.S.A.) in accordance with the manufacturer's protocol with slight modifications. PCR amplifications as per standard protocols were performed for the four recommended standard barcode regions *rbcL*, *matK*, *trnH-psbA* of the plastid genome as well as for *ITS2* (CBOL Plant Working Group, 2009) of the nuclear genome, using previously reported primers (Table 4).

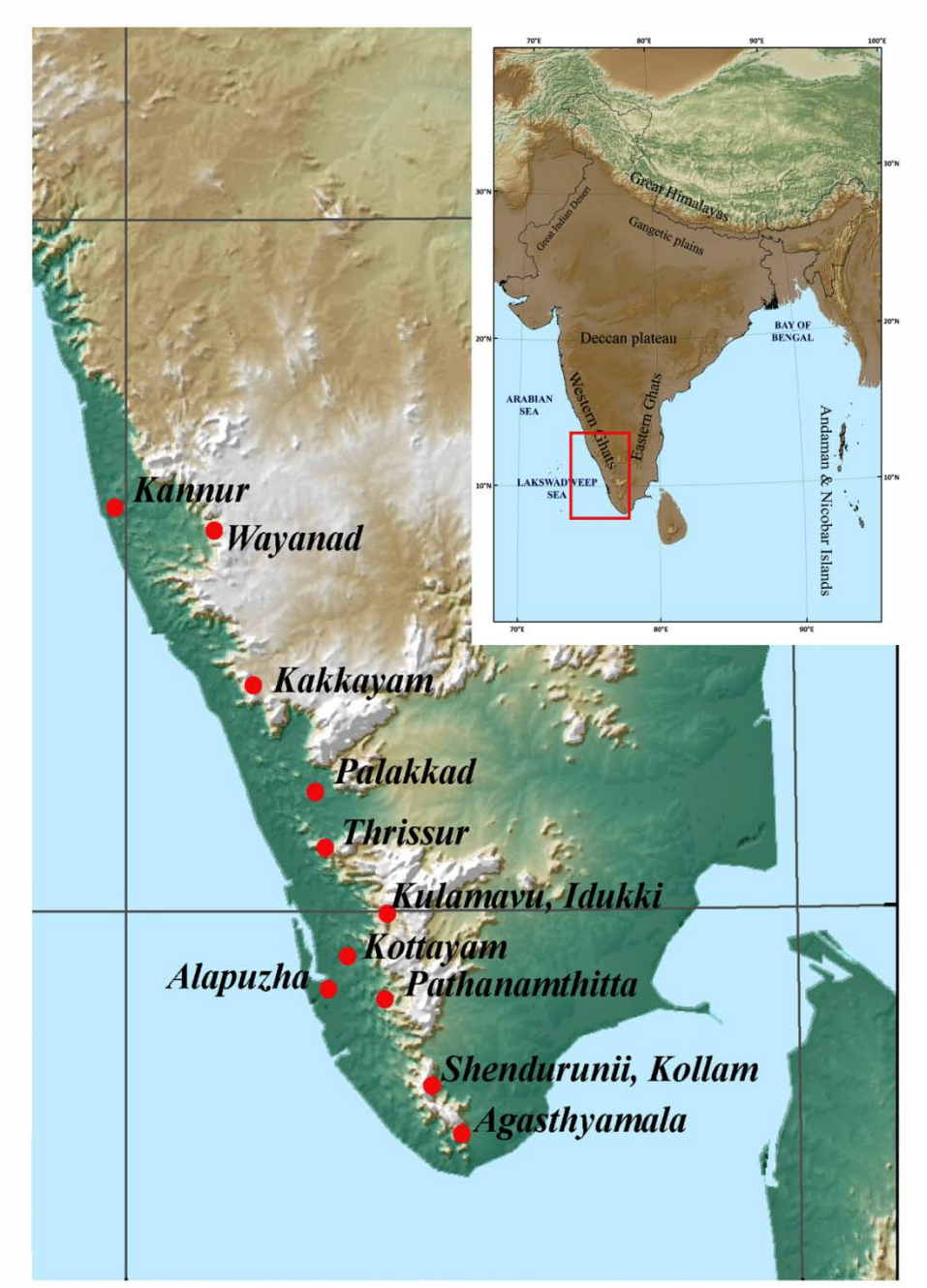

Figure 10. The localities of *Salacia* species in peninsular India used for this study

Table 4. Barcode loci and primer information

Barcode loci	Primer information (5′-3′)	Primer Annealing Temperature
rbcL (Fay et al. 1997)	1F ATGTCACCACAAACAGAAAC 724R TCGCATGTACCTGCAGTAGC	60^0C
matK (Yu et al. 2011)	472F CCRTCATCTGGAAATCTTGGTT 1248R GCTRTRATAATGAGAAAGATTTCTGC	60^0C
trnH-psbA (Kress et al. 2005)	*trnH* CGCGCATGGTGGATTCACAATCC *psbA* RGTWATGCAYGAACGTAATGCTC	60.8^0C
ITS2 (White et al. 1990)	ITS3-GCATCGATGAAGAACGCAGC ITS4-TCCTCCGCTTATTGATATGC	59^0C

Sequence alignment and data analysis

Please refer to the methodology provided in page 19.

RESULTS

All the four analyzed barcode regions (*rbcL, matK, trnH-psbA* and *ITS2*) were amplified using recommended specific primers and PCR efficiency was 100%. All the sequences generated out of the study were submitted to GenBank (http://www.ncbi.nlm.nih.gov/genbank/) as well as to BOLD (http://www.barcodinglife.org) (Table 5). Sequence length and basic sequence statistics like conserved sites, variable sites, singletons and transition/ transversion ratio, based on *CLUSTAL X* alignment as well as with alignment explorer in *MEGA 4.0,* are provided in Table 6. The number of variable sites was greater for *ITS2* as compared to other barcode regions.

Table 5. NCBI Genbank accession numbers and BOLD process ID of all the 60 samples

Species name	Sample ID	Genbank accession numbers				BOLD Process ID
		rbcL	*matK*	*trnH-psbA*	*ITS2*	*ITS2*
S. fruticosa	S.fru1	KR091312	KR091372	KR091432	KR091492	SALAC001-15
S. fruticosa	S.fru02	KR091313	KR091373	KR091433	KR091493	SALAC002-15
S. fruticosa	S.fru03	KR091314	KR091374	KR091434	KR091494	SALAC003-15
S. fruticosa	S.fru04	KR091315	KR091375	KR091435	KR091495	SALAC004-15
S. fruticosa	S.fru05	KR091316	KR091376	KR091436	KR091496	SALAC005-15
S. fruticosa	S.fru06	KR091317	KR091377	KR091437	KR091497	SALAC006-15
S. fruticosa	S.fru07	KR091318	KR091378	KR091438	KR091498	SALAC007-15
S. fruticosa	S.fru08	KR091319	KR091379	KR091439	KR091499	SALAC008-15
S. fruticosa	S.fru09	KR091320	KR091380	KR091440	KR091500	SALAC009-15
S. fruticosa	S.fru10	KR091321	KR091381	KR091441	KR091501	SALAC010-15
S. fruticosa	S.fru11	KR091322	KR091382	KR091442	KR091502	SALAC011-15
S. fruticosa	S.fru12	KR091323	KR091383	KR091443	KR091503	SALAC012-15
S. fruticosa	S.fru13	KR091324	KR091384	KR091444	KR091504	SALAC013-15
S. fruticosa	S.fru14	KR091325	KR091385	KR091445	KR091505	SALAC014-15
S. oblonga	S.ob01	KR091326	KR091386	KR091446	KR091506	SALAC015-15
S. oblonga	S.ob02	KR091327	KR091387	KR091447	KR091507	SALAC016-15
S. oblonga	S.ob03	KR091328	KR091388	KR091448	KR091508	SALAC017-15
S. oblonga	S.ob04	KR091329	KR091389	KR091449	KR091509	SALAC018-15
S. beddomei	S.bed01	KR091330	KR091390	KR091450	KR091510	SALAC019-15
S. beddomei	S.bed02	KR091331	KR091391	KR091451	KR091511	SALAC020-15
S. beddomei	S.bed03	KR091332	KR091392	KR091452	KR091512	SALAC021-15
S. beddomei	S.bed04	KR091333	KR091393	KR091453	KR091513	SALAC022-15
S. beddomei	S.bed05	KR091334	KR091394	KR091454	KR091514	SALAC023-15
S. beddomei	S.bed06	KR091335	KR091395	KR091455	KR091515	SALAC024-15
S. beddomei	S.bed07	KR091336	KR091396	KR091456	KR091516	SALAC025-15
S. beddomei	S.bed08	KR091337	KR091397	KR091457	KR091517	SALAC026-15
S. chinensis	S.chi01	KR091338	KR091398	KR091458	KR091518	SALAC027-15
S. chinensis	S.chi02	KR091339	KR091399	KR091459	KR091519	SALAC028-15
S. chinensis	S.chi03	KR091340	KR091400	KR091460	KR091520	SALAC029-15

S. chinensis	S.chi04	KR091341	KR091401	KR091461	KR091521	SALAC030-15
S. chinensis	S.chi05	KR091342	KR091402	KR091462	KR091522	SALAC031-15
S. chinensis	S.chi06	KR091343	KR091403	KR091463	KR091523	SALAC032-15
S. macrosperma	S.ma01	KR091344	KR091404	KR091464	KR091524	SALAC033-15
S. macrosperma	S.ma02	KR091345	KR091405	KR091465	KR091525	SALAC034-15
S. macrosperma	S.ma03	KR091346	KR091406	KR091466	KR091526	SALAC035-15
S. macrosperma	S.ma04	KR091347	KR091407	KR091467	KR091527	SALAC036-15
S. macrosperma	S.ma05	KR091348	KR091408	KR091468	KR091528	SALAC037-15
S. macrosperma	S.ma06	KR091349	KR091409	KR091469	KR091529	SALAC038-15
S. vellaniana	S.vel01	KR091350	KR091410	KR091470	KR091530	SALAC039-15
S. vellaniana	S.vel02	KR091351	KR091411	KR091471	KR091531	SALAC040-15
S. vellaniana	S.vel03	KR091352	KR091412	KR091472	KR091532	SALAC041-15
S. vellaniana	S.vel04	KR091353	KR091413	KR091473	KR091533	SALAC042-15
S. vellaniana	S.vel05	KR091354	KR091414	KR091474	KR091534	SALAC043-15
S. vellaniana	S.vel06	KR091355	KR091415	KR091475	KR091535	SALAC044-15
S. agasthiamalana	S.aga01	KR091356	KR091416	KR091476	KR091536	SALAC04515
S. agasthiamalana	S.aga02	KR091357	KR091417	KR091477	KR091537	S47AC046-15
S. agasthiamalana	S.aga05	KR091358	KR091418	KR091478	KR091538	SALAC047-15
S. agasthiamalana	S.aga04	KR091359	KR091419	KR091479	KR091539	SALAC048-15
S. agasthiamalana	S.aga05	KR091360	KR091420	KR091480	KR091540	SALAC049-15
S. oblonga var. *kakkayamana*	S.ka01	KR091361	KR091421	KR091481	KR091541	SALAC050-15
S. oblonga var. *kakkayamana*	S.ka02	KR091362	KR091422	KR091482	KR091542	SALAC051-15
S. oblonga var. *kakkayamana*	S.ka03	KR091363	KR091423	KR091483	KR091543	SALAC052-15
S. oblonga var. *kakkayamana*	S.ka04	KR091364	KR091424	KR091484	KR091544	SALAC053-15
S. oblonga var. *kakkayamana*	S.ka05	KR091365	KR091425	KR091485	KR091545	SALAC054-15
S. oblonga var. *kakkayamana*	S.ka06	KR091366	KR091426	KR091486	KR091546	SALAC055-15
S. oblonga var. *kakkayamana*	S.ka07	KR091367	KR091427	KR091487	KR091547	SALAC056-15
S. oblonga var. *kakkayamana*	S.ka08	KR091368	KR091428	KR091488	KR091548	SALAC057-15
S.malabarica	S.mal01	KR091369	KR091429	KR091489	KR091549	SALAC058-15
S.malabarica	S.mal02	KR091370	KR091430	KR091490	KR091550	SALAC059-15
S.malabarica	S.mal03	KR091371	KR091431	KR091491	KR091551	SALAC060-15

Table 6. Summary statistics of the four analysed barcodes

	rbcL	*matK*	*trnH-psbA*	*ITS2*
Length range (bp)	701	743	643	410
Conserved sites	685	720	559	353
Variable sites	10	18	39	44
Singletons	3	12	23	11
Transition/ transversion ratio	2.7	1.2	1.8	2.2
Nearest *DISTANCE* method (Identification Efficiency (1%)	0	0	40	100

Genetic divergence was estimated using six parameters (Table 7). *ITS2* barcode exhibited the highest interspecific divergence (i.e., average interspecific distance; theta prime) followed by *trnH-psbA, matK* and *rbcL* and the minimum interspecific genetic distance was zero in all these four barcode loci. At the intraspecific level, maximal intraspecific genetic distance (coalescent depth) was highest for *trnH-psbA* followed by *ITS2* and lowest for *matK*. Wilcoxon's signed rank test demonstrated that the interspecific divergence measures displayed by *ITS2* barcode were significantly higher than that of other barcode regions.

DNA barcoding gap was evident for *ITS2* whereas it was less conspicuous for *trnH-psbA* as well as for *matK*. For *rbcL*, the intraspecific distance was higher than that of interspecific genetic distance. Nearest distance method was used to test the suitability of different barcodes for species identification. *ITS2* barcode possessed 100% species identification efficiency in all the tested samples whereas the identification efficiency was 40% for *trnH-psbA* and the rest of the barcode loci (*rbcL*, *matK*) failed to discriminate species.

Table 7. Analysis of interspecific divergence among congeneric species and intraspecific variation within species

Marker	*rbcL*	*matK*	*trnH-psbA*	*ITS2*
Average interspecific distance	0.0012±0.003	0.005±0.005	0.026±0.019	1.362±1.008
Average Theta prime	0.004±0.005	0.004±0.004	0.028±0.017	0.725±0.962
Minimum interspecific distance	0	0	0	0
Average intraspecific distance	0.001±0.004	0	0.011±0.015	0.052±0.1484
Average theta	0.003±0.002	0.003±0.003	0.026±0.016	0.698±0.628
Average coalescent depth	0.0031±0.000	0.001±0.001	0.027±0.036	0.0045±0.021

The topology of the four NJ derived phylograms revealed two or three major distinct clusters. Most species of *Salacia,* except *S. marcrosperma*/ *S. beddomei,* formed well supported monophyletic clusters in the generated phylograms. *S. macrosperma* and *S. beddomei* were clustered together in the *ITS2* (Fig. 11) and *ITS2*/ *trnH-psbA* derived phylograms while they displayed monophyletic clusters in the rest of the phylograms with many shared nucleotides. *S. oblonga* and its variety '*kakkayamana*' also displayed similar nucleotide changes and both clustered together in the generated phylograms. The widely distributed *S. fruticosa* showed many unique base pair changes in their haplotypes (Fig. 12). *Salacia chinensis* can be considered as a standalone species and generated strongly supported (100 %) monophyletic clusters in all the phylograms. Reciprocal monophyly was shown by *S. fruticosa*, *S. chinensis*, *S. agasthyamalana* and *S. macrosperma* in the phylogram generated out of the combined dataset of the four barcodes.

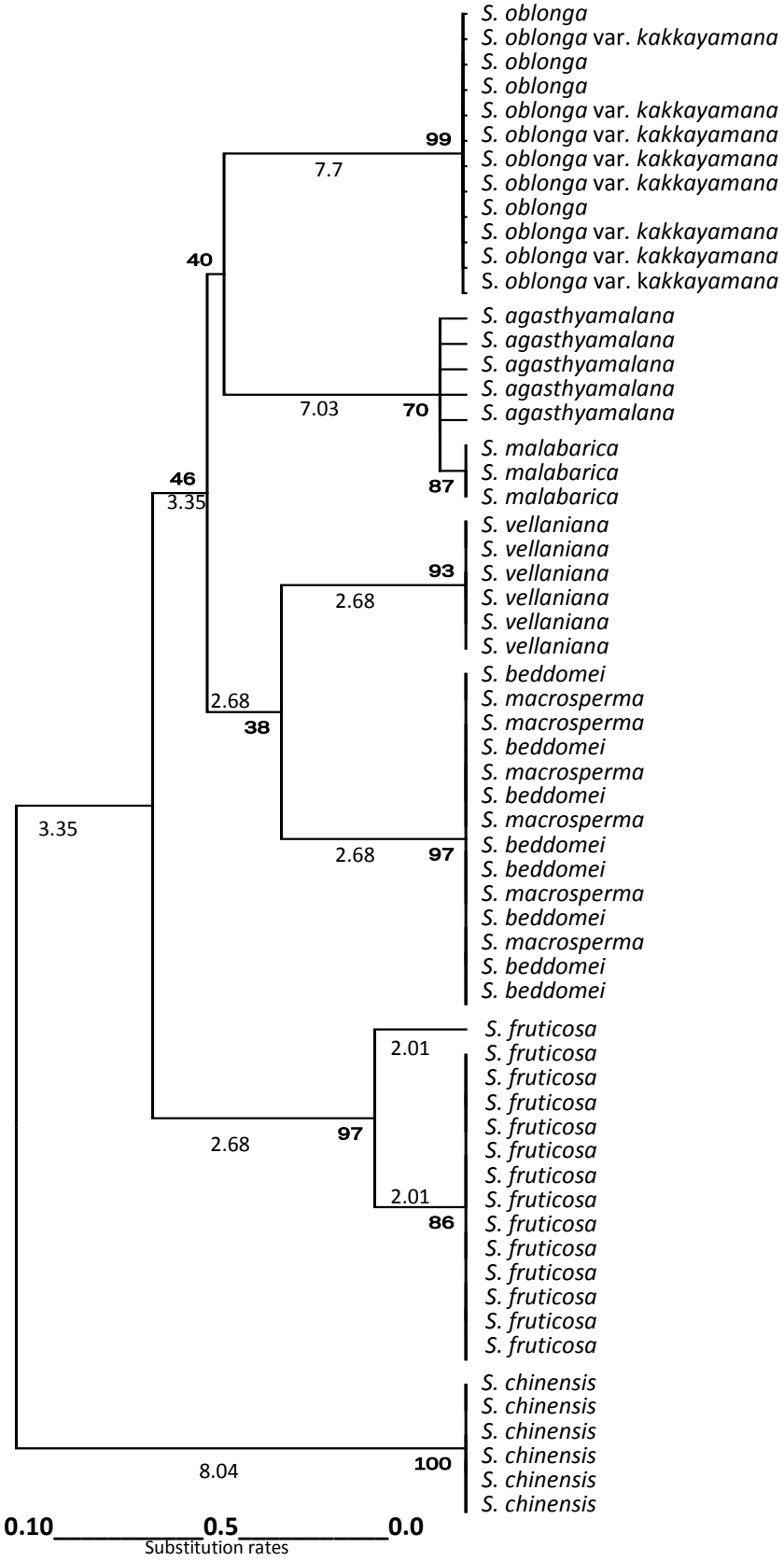

Figure 11. Phylogenetic tree of *ITS2* sequences. The lengths of the nodes represent the substitution rate, which is defined as the percentage of substitution sites per alignment length. Boot strap values≥35 are shown on the branch points. Phylogenetic trees were drawn as NJ trees adopting Kimura 2-parameter in the *MEGA 4.0.*

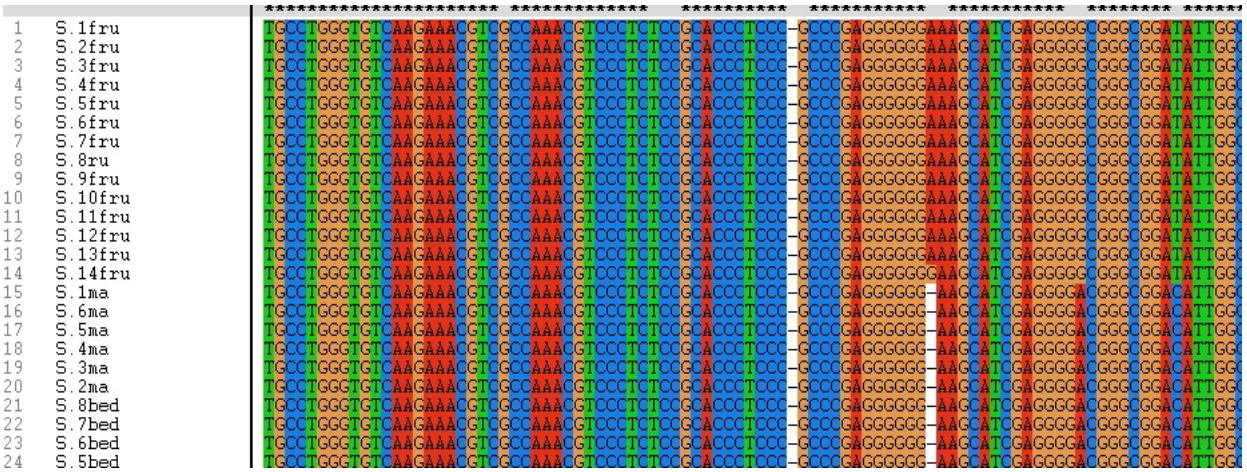

Figure 12. Unique base pair changes in *ITS2* sequences of *S. fruticosa* (S.fr)

DISCUSSION

DNA barcoding technique has been utilized as a powerful tool to confirm species identity in cases of taxonomic complexities (Pang *et al*., 2011; Xin *et al*., 2012; Hou *et al*., 2013). An ideal DNA barcode region should be evolutionarily conserved so as to have lower variability within a species and greater variability among congeneric species of a genus (Taberlet *et al*., 2007). Angiosperm DNA barcoding was proposed originally by the CBOL Plant Working Group with two standard barcode regions (*rbcL* and *matK*) and *trnH-psbA* and the *ITS2* region were recommended later on (Kress and Erickson, 2007; Chiou *et al*., 2007). The *ITS2* region has been selected as a barcode candidate because of the potential in phylogenetic analysis both at the genus and species levels (Chen *et al*., 2010; Xin *et al*., 2013).

Results of the present study recommend *ITS2* as the most promising DNA barcode region for identifying *Salacia* species in terms of universality, specific genetic divergence and discriminality among congeneric species. Among the four analysed barcode regions, *ITS2* was the shortest in length and was relatively easy to amplify using the recommended universal primers. Even though *matK, rbcL* and *trnH-psbA* were successfully amplified (100 %), the three markers were less powerful than *ITS2*

in identifying *Salacia* species. Genetic divergence parameters derived out of *ITS2* sequences displayed a significant difference between average inter and intraspecific genetic distances and DNA barcoding gap was also highest. According to the nearest distance method, identification accuracies of *ITS2* barcode were 100% at the species level. Even though in previous studies on other genera, *trnH-psbA* has been accepted as a potential barcode (Kress *et al*., 2005; Song, 2009), it has been reported to be less efficient in distinguishing closely related plants (Chen *et al*., 2010). The present study also showed a lower efficiency of this non-coding chloroplast intergenic spacer region (*trnH-psbA*) (40% identification efficiency) as compared to the *ITS2* barcode. *ITS2* has been previously reported as a candidate marker for taxonomic classification because of its correct identification rate and amplification efficiency (Coleman, 2003, 2007; Li *et al*., 2011; Hollingsworth, 2011) and has been recommended for the barcoding of seed plants (China Plant BOL group *et al*., 2011). The potential of *ITS2* region to identify closely related species has been previously reported for some medicinal plants (Chen et al. 2010). Even though multiple copies of the nuclear *ITS2* gene are present, they generally display a high amount of sequence similarity and can often be treated as a single locus (Alvarez and Wendel, 2003; Coleman, 2007). *ITS2* as a DNA barcode had been used as a confirmatory tool in resolving taxonomic complexities in the genus *Astragalus* and *A. membranaceus* was recommended to be a variety of *A. mongholicus* (Dong *et al*., 2003). Similarly, 100% homology of *ITS2* sequences and striking morphological similarities were also reported in *Caragana rosea* and *C. sinica* (Hou *et al*., 2008). *ITS2* has been used as a potential barcode in identifying medicinal plants of the family Fabaceae (Gao *et al*., 2010), as well as in the family Polygonaceae (Yao *et al*., 2009), for confirming the species identity in the genus *Nyssa* (Wang *et al*., 2012) as well as to distinguish medicinal plant *Boerhaavia diffusa* from its adulterant species (Selvaraj *et al*., 2012). *ITS2* together with *trnH-psbA* is used for discriminating the Chinese medicinal plant, *Lonicera japonica* from its morphologically similar adulterant species (Hou *et al*., 2013).

Species identification in the genus *Salacia* is difficult when based solely on the morphological characteristics and an accurate method is indispensable to ensure proper and safe use and assured efficacy of extracted raw drugs. The *ITS2* barcode

with significant species discrimination power generated through the present study can serve as a supplementary tool for confirming the species identity of the raw drugs derived from the genus. The monophyletic clustering displayed by most species demonstrated the efficiency of the barcode to identify *Salacia* species. Five of the studied species (*S. fruticosa, S. vellanina, S. chinensis, S. malabarica, S. agastyamalana*) displayed well supported distinct monophyletic clustering in the generated phylograms. The reciprocal monophyletic clustering displayed by *S. macrosperma/ S. beddomei* in most of the cases suggest that they are closely related sister species. All the vegetative characters of these species are substantially the same and the only difference between these two species is the numerous flowers in dense fascicles in *S. macrosperma* and some variation in fruit characteristics (Udayan and Pradeep, 2012). Reciprocal monophyly and sharing of haplotypes between sister species is an expected transitional stage for diverging lineages (Tajima, 1983; Maddison, 1997). In all the derived phylograms, *S. oblonga* and its variety (*S. oblonga* var. *kakkayamana*) showed 100 % homology and are clustered together. Morphological differences among them exist only in the floral characteristics such as globose nature of flowers, colour and size of flowers (Udayan *et al*., 2012, 2013). All other species of *Salacia* even though morphologically similar displayed well supported monophyletic clustering in the generated phylograms with unique haplotypes for each species.

6. SUMMARY AND CONCLUSIONS

The present project could standardise protocols for DNA barcoding and barcode sequence data analysis for deriving interpretations on the identity of the unknown specimens in the case of plants, insects and fungi. The case studies carried out could demonstrate the capability of DNA barcoding technique in solving taxonomic complexities as well as in discriminating the original species from the substitute/ adulterant species.

East Indian Sandalwood (*Santalum album*) in commercial markets is highly prone to adulteration. A number of cases were registered with regard to the adulteration of East Indian sandalwood, but the lack of technical tools for the precise species identification of the source wood stalled most of the cases. The ability to track or identify timber resources of economic value is essential for the effective management and appropriate regulation of timber trade. The standard DNA barcode regions *rbcL*, *matK* and *trnH-psbA* of choloroplast genomic sequences recommended by the Consortium of Barcode of Life (CBOL) were analysed for their potential to identify wood adulterants of sandalwood. Standard polymerase chain reactions with CBOL recommended primers were performed for all the three barcode loci. The PCR products after gel elution were sequenced and alignments were performed using *CLUSTALX*.

The project could develop species specific DNA barcodes for the original East Indian sandalwood and their market adulterants. Just like the Universal Product Code (UPC) that is used for tracking trade items widely, the developed species specific DNA barcodes would help to distinguish the market adulterants of sandalwood from the original wood. Single nucleotide polymorphisms observed in the adulterants of original sandalwood clearly indicated the utility of DNA barcoding to keep track the wood adulterants in Indian sandalwood industry. Recent technological innovations and high level of automation have reduced the expenses and the technical skills required for the execution of DNA barcoding technique. Practical checks at the credible voluntary wood certification agencies as well as in the court of law could be made tamper proof by the implementation of DNA barcoding along with the existing traditional techniques for timber monitoring.

The second case study was carried out to prove the efficiency of DNA barcode sequences to discriminate morphologically similar species' of the genus *Salacia*. The genus *Salacia* is a source of many important pharmaceutical chemicals used in the Ayurvedic system of medicine in India. Owing to the morphological similarities, taxonomical complexities exist within the genus. To ensure quality and assured therapeutic effects in the raw drugs from the genus, proper identification at the species level is critical. The main objective of this study was to arrive at suitable DNA barcodes that can accurately and efficiently identify the potential medicinal species of the genus. Among the barcode loci analyzed, *ITS2* DNA barcode exhibited highest interspecific divergence followed by *trnH-psbA, matK* and *rbcL*. A clear barcoding gap was evident for the *ITS2* barcode region where as it was less conspicuous for *trnH-psbA* as well as for *matK*. *ITS2* barcode could discriminate all the analyzed *Salacia* species with 100 % accuracy. We therefore propose DNA barcoding with *ITS2* to confirm the taxonomic identity of the raw drugs in the market and to ensure safe supply of traditional medicines derived out of different species of the genus *Salacia*. Further, DNA barcoding based on *ITS2* region can be employed for the biosystematics of the genus *Salacia.*

Altogether, the studies on the genus *Salacia* support the claim that *ITS2* is a valuable barcode locus for discriminating congeneric species and that DNA barcoding can be used as a valid tool for biosystematics. We propose that applying DNA barcoding technology with *ITS2* barcode can resolve classification problems in the genus *Salacia* at the species level. Further, the application of DNA barcoding for identification of *Salacia* species can establish taxonomic identity of the raw drugs in the market and ensure safe supply of traditional medicines derived from the genus. These case studies established the capability of DNA barcoding technique for the identification of the botanical origins at species level of commercially important plant products.

7. REFERENCES

Alvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogen Evol 29: 417–434

Ananthapadmanabha HS (2012) Expert sandalwood marketing report, TFS sandalwood project report, pp 69-71

Anderson IC, Cairney JWG (2004) Diversity and ecology of soil fungal communities: increased understanding through the application of molecular techniques. Environmental Microbiology 6(8):769-779

Anupama C, Balasundaran M, Sujanapal P (2012) Phylogenetic relationships of *Santalum album* and its adulterants as inferred from nuclear DNA sequences. Int J Agriculture and Forestry 2: 150-156

Asif MJ, Cannon CH (2005) DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). Plant Molecular Biology Reporter 23: 185-192

Bellemain E, Carlsen T, Brochmann **C,** Coissac **E,** Taberlet **P,** Kauserud H (2010) ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. BMC Microbiology 10:189

Bhat KV, Balasundaran M, Balagopalan M (2006) Identification of *Santalum album* and *Osyris lanceolata* through morphological and biochemical characteristics and molecular markers to check adulteration. KFRI Research Report 307. 22p.

Boner M, Sommer T, Erven C, Forstel H (2007) Stable isotopes as a tool to trace back the origin of wood. In: Proceedings of the international workshop, Fingerprinting methods for the identification of timber origins, October 8-9, Bonn/Germany, pp 47-57

CBOL– Plant Working Group (2009) A DNA barcode for land plants. Proceedings of the National Academy of Sciences, USA 106: 12794–12797

Chase MW, Cowan RS, Hollingsworth PM, Berg CVD, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M (2007) A proposal for a standardized protocol to barcode all land plants. TAXON 56: 295–299

Chase MW, Fay MF (2009) Ecology. Barcoding of plants and fungi. Science 325: 682–683

Chase MW, Salamin N, Wilkinson M (2005) Land plants and DNA barcodes: short-term and long-term goals. Philos Trans R Soc Lond B Biol Sci 360: 1889–1895

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin, Y, Leon C (2010) Validation of the *ITS2* region as a model DNA barcode for identifying medicinal plants species. PLoS ONE 5: e8613

China Plant BOL Group, Li DZ, Gao LM, Li HT, Wang H *et al* (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (*ITS*) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci USA 108: 19641–19646

Chiou SJ, Yen JH, Fang CL, Chen HL, Lin, TY (2007) Authentication of medicinal herbs using PCR-amplified *ITS2* with specific primers. Planta Med 73: 1421–1426

Coleman AW (2003) *ITS2* is a double-edged tool for eukaryote evolutionary comparisons. Trends in Genetics 19: 370–375

Coleman AW (2007) Pan-eukaryote *ITS2* homologies revealed by RNA secondary structure. – Nucleic Acids Res 35: 3322–3329

Costion C, Ford A, Cross H (2011) Plant DNA barcodes can accurately estimate species richness in poorly known floras. PLoS ONE 6: E26841. www.plosone.org

Coyle HM, Lee C, Lin W, Lee HC, Palmbach TM (2005) Forensic botany: using plant evidence to aid in forensic death investigation. Croatian Medical Journal 46: 606–612

Degen B, Fladung, M (2008), Use of DNA-markers for tracing illegal logging. In: Degen B (ed) Proceedings of the international workshop "Fingerprinting methods for the identification of timber origins" October 8–9, Landbauforschung, VTI Agriculture and Forestry Research, Sonderheft 321, Germany, pp 6–14

Deguilloux MF, Premonge MH, Petit RJ (2002) Novel perspectives in wood certification and forensics: dry wood as a source of DNA. Proc R Soc Lond 269: 1039-1046

Dev SA, Muralidharan EM, Sujanapal P, Balasundaran M (2014) Identification of market adulterants in East Indian sandalwood using DNA barcoding. Annals of Forest Science 71: 517–522

Dick CW, Kress WJ (2009) Dissecting Tropical Plant Diversity with Forest Plots and a Molecular Toolkit. Bioscience 59, 9: 745–755

Dong TTX, Ma XQ, Clarke C, Song ZH, Ji ZN, Lo CK, Karl W, Tsim KWK (2003) Phylogeny of *Astragalus* in China: molecular evidence from the DNA sequences of 5S rRNA spacer, ITS, and 18S rRNA. Journal of Agricultural and Food Chemistry 51: 6709–6714

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 19:11-15

Drager RG, Hallick RB (1993) A novel *Euglena gracilis* chloroplast operon encoding 4 ATP synthase subunits and 2 ribosomal-proteins contains 17 Introns. Current Genetics 23: 271– 280

Erickson DL, Spouge J, Resch A, Weigt LA, Kress J W (2008) DNA barcoding in land plants: developing standards to quantify and maximize success. Taxon 57: 1304–1316

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology & Biotechnology, 3: 294–299

Fay MF, Swensen SM, Chase MW (1997) Taxonomic affinities of Medusagyne oppositifolia (Medusagynaceae). Kew Bulletin 52: 111-120

Fuji T (2007) Outline of the research project »Methods to identify wood species and origin of timber of Southeast Asia. Proceedings of the international symposium on development of improved methods to identify *Shorea* species wood and its origin, 19: 1699- 1704

Galimberti A, De Mattia F, Losa A, Bruni I, Federici S, Casiraghi M, Martellos S, Labra M (2012) DNA barcoding as a new tool for food traceability. Food Research International 50: 55–63

Gao T, Yao H, Song J, Zhu Y, Liu C. Chen S (2010) Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evolutionary Biology 10: article 324.

Ghorpade RP, Chopra A, Nikam TD (2010) In vitro zygotic embryo germination and propagation of an endangered *Boswellia serrata* Roxb., a source of boswellic acid. *Physiology* and *Molecular Biology* of *Plants* 16:159–165

Guisinger MM, Kuehl JV, Boore JL, Jansen RK (2008) Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. Proceedings of the National Academy of Sciences, USA 105: 18424–18429

Hajibabaei M, deWaard JR, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie PM, Hebert PDN (2005) Philosophical Transactions of the Royal Society of *London*. Series *B*: Biological Sciences 360: 1959–1967

Hall TA (1999) A user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series 41: 95-98

Hebert PDN, Cywinska A, Ball S L, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences 270: 313–321

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astrapes fulgerator. Proceedings of the National Academy of Sciences of the United States of America 101: 14812–14817

Hewson HJ, George AS (1984) Santalaceae In: George AS (ed) Flora of Australia, Australian Government Publishing Service, Canberra, pp 29-67

Hollingsworth ML, Clark AA, Forrest LL, Richardson J, Pennington RT, Long DG, Cowan R, Chase MW, Gaudeul M, Hollingsworth PM (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. Molecular Ecology Resources 9: 439–457

Hollingsworth PM (2011) Refining the DNA barcode for land plants. Proc Natl Acad Sci USA 108: 19451–19452

Holmquist R (1983) Transitions and transversions in evolutionary descent: an approach to understanding. J Mol Evol 19: 134-144

Hou DY, Song JY, Shi LC, Ma XC, Xin TY, Han JP, Xiao W, Sun ZY, Cheng RY, Yao H (2013) Stability and accuracy assessment of identification of traditional Chinese materia medica using DNA barcoding: a case study on Flos Lonicerae Japonicae. Biomed research international 549073

Hou X, Liu JE, Zhao YZ (2008) Molecular phylogeny of *Caragana* (Fabaceae) in China. Journal of Systematic Evolution 46: 600–607

Huxley-Jones E, Shaw JL, Fletcher C, Parnell J, Watts PC (2012) Use of DNA barcoding to reveal species composition of convenience seafood. Conservation Biology 26: 367–371

IUCN (2009) IUCN Red List of Threatened Species. Version 2009.1, www.iucnredlist.org

Kane NC, Cronk Q (2008) Botany without borders: barcoding in focus. Molecular Ecology 17: 5175-5176

Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120

Kress JW, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *trnH-psb*A spacer region. PLoS One 2: e508

Kress JW, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences, USA 102: 8369–8374

Kress WJ, Erickson DL (2008) DNA Barcoding--a Windfall for Tropical Biology? Biotropica 40: 405-408

Lahaye R, van der Bank M, Bogarin D (2008) DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci USA 105: 2923–2928

Larkin MA, Blackshields G, Brown NP (2007) Clustal W and ClustaL X version 2. Bioinformatics 23: 2947-2948

Li DZ, Gao LM, Li HT, Wang H, Ge XJ *et al* (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (*ITS*) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci 108: 19641-19646

Liepelt S, Sperisen C, Deguilloux MF (2006) Authenticated DNA from ancient wood remains. Ann Bot 98: 1107–1111.

Linnaeus C (1753) Species Plantarum. Stockholm: Laurentii Salvii

Lowe AJ, Cross HB (2011) The application of DNA to timber tracking and origin verification. Journal of the International Association of Wood Anatomists 32: 251-262

Lowe AJ, Wong KN, Tiong TS (2010) A DNA method to verify the integrity of timber supply chains, confirming the legal sourcing of Merbau timber from logging concession to sawmill. Silvae Genetica 59: 263-268

Maddison WP (1997) Gene trees in species trees. Syst Biol 46: 523–536

Meng BY, Wakasugi T, Sugiura M (1991) Two promoters within the *psbK–psbI–trnG* gene-cluster in tobacco chloroplast DNA. Current Genetics 20: 259–264

Meyer C P, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 3: 2229–2238

Mildenhall D (2006) Hypericum pollen determines the presence of burglars at the scene of a crime: an example of forensic palynology. Forensic Science International 163: 231–235

Miller SE (2007) DNA barcoding and the renaissance of taxonomy. Proceedings of the National Academy of Science USA 104: 4775–4776

Mora C, Tittensor DP, Adl S, Simpson AGB and Worm B (2011) How Many Species Are There on Earth and in the Ocean? PLoS Biology 9(8)

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. Molecular Ecology Notes 8: 480–490

Nock CJ, Waters DLE, Edwards MA, Bowen SG, Rice N (2011) Chloroplast genome sequences from total DNA for plant identification. Plant Biotechnology Journal 9: 328–333

O'Brien HE, Parrent JL, Jackson JA, Moncalvo J, Vilgalys R (2005) Fungal community analysis by large-scale sequencing of environmental samples. Applied and Environmental Microbiology 71: 5544–5550

Oyen LPA, Dung NX (1999) Plant resources of south-east Asia No. 19, Essential-oil plants, Backhuys, Leiden, the Netherlands, 1999.

Paarakh PM, Patil LJ, Thanga SA (2008) Genus *Salacia*: a comprehensive review. Journal of Natural Remedies 8: 116–131

Page T, Tate H, Bunt C, Potrawiak A, Berry A (2012) Opportunities for the smallholder sandalwood industry in Vanuatu. ACIAR technical reports No.79, Australian Centre for International Agricultural Research, Canberra. 67 p

Pang X, Song J, Zhu Y (2011) Applying plant DNA barcodes for Rosaceae species identification. Cladistics 27: 165–170

Pennisi E (2007) Taxonomy. Wanted: A barcode for plants. Science 318:190–191

Rachmayanti Y, Leinemann L, Gailing O (2006) Extraction, amplification and characterization of wood DNA from Dipterocarpaceae. Plant Mol Biol Rep 24: 45–55

Ramamurthy K, Venu P (2005) In: Daniel P (ed), *Flora of Kerala* Vol. 1, Calcutta, Botanical Survey of India.

Ross HA, Sumathi M, Li WLS (2008) Testing the reliability of genetic methods of species identification via simulation. Syst Biol 57: 216–230

Sasidharan N (2004) Biodiversity documentation for Kerala. *Part 6. Flowering Plants,* Kerala Forest Research Institute, Peechi, Kerala.

Seifert KA, Samson RA, Dewaard JR, Houbraken J, Levesque CA, Moncalvo JM, Louis-Seize G, Hebert PDN (2007) Prospects for fungus identification using C01 DNA barcodes, with *Penicillium* as a test case. Proceedings of the National Academy of Sciences of the United States of America 104: 3901-3906

Selvaraj D, Shanmughanandan D, Sarma RV, Joseph JC, Srinivasan RV, Ramalingam S (2012) DNA barcode *ITS* effectively distinguishes the medicinal plant *Boerhaavia diffusa* from its adulterants. –Genomics Proteomics Bioinformatics 10: 364–367

Shivji M, Clarke C, Pank M (2002) Genetic identification of pelagic shark body parts for conservation and trade monitoring. Conservation Biology 16: 1036-1047

Shyaula SL (2012) A review on genus *Osyris*: phytochemical constituents and traditional uses. J. Natural Pharmaceuticals 3: 61-70

Song JY (2009) Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J Ethnopharm 124: 434–439

Srimathi RA, Kulkarni HD, Ventkatesan K (1995) Recent advances in research and management of sandal (*Santalum album* L.) in India, Associated publishing Co., New Delhi, 1995.

Starr JR, Naczi RF, Chouinard BN (2009) Plant DNA barcodes and species resolution in sedges (Carex, Cyperaceae). Molecular Ecology Resources 9, Suppl 1: 151-163

Stewart CN Jr. (2005) Monitoring the presence and expression of transgenes in living plants. Trends in Plant Science 10: 390–396

Taberlet P, Coissac E, Pombanon E *et al* (2007) Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. Nucleic Acids Res 35: e14

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437–460

Tamura K, Dudley J, Nei M (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA 6: Molecular Evoluionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30: 2725-2729

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The *CLUSTAL_X* windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Resources 25: 4876–4882

Tsumura Y, Kawahara T, Wickneswari R, Yoshimura K (1996) Molecular phylogeny of Dipterocarpaceae in Southeast Asia using RFLP of PCR-amplified chloroplast genes. Theoretical and Applied Genetics 93: 22–29

Udayan PS, Pradeep AK (2012) Notes on the occurrence of *Salacia reticulata* (Hippocrataceae) in India. Acta Botanica Hungarica 54: 433-441

Udayan PS, Yohannan R, Devipriya MS, Devipriya V, Pradeep AK (2012) A new species of *Salacia* (Hippocrataceae) from South India. Edinburgh Journal of Botany 69: 1-4

Udayan PS, Yohannan R, Devipriya MS, Devipriya V, Pradeep AK (2013) *Salacia vellaniana* Udayan, Yohannan & Pradeep (Celastraceae), a new species from India. Candollea 68: 147-149

van de Wiel CCM, van der Schoot J, van Valkenburg JLCH, Duistermaat H, Smulders MJM (2009) DNA barcoding discriminates the noxious invasive plant species, floating pennywort (Hydrocotyle ranunculoides L.f.), from non-invasive relatives. Molecular Ecology Resources 9: 1086–1091

Vijayan K, Tsou C (2010) DNA barcoding in plants: taxonomy in a new perspective. Current Science (Bangalore) 99: 1530–1541

Wang N, Jacques MB, Milne MB, Zhang CQ, Yang JB (2012) DNA barcoding of Nyssaceae (Cornales) and taxonomic issues. Botanical Studies 53: 265-274

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: a guide to methods and applications. (Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds). Academic Press, New York, USA: 315–322

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. Proceedings of the National Academy of Sciences, USA 84: 9054–9058

Xin TY, Yao H, Gao SH *(* 2013) Super food *Lycium barbarum* (Solanaceae) traceability via an internal transcribed spacer 2 barcode. Food Research International 54(2): pp. 1699–1704

Xin TY, Yao H, Luo K (2012) Stability and accuracy of the identification of *Notopterygii rhizoma* et Radix using the *ITS/ITS2* barcodes. Acta Pharmaceutica Sinica 47: 1098–1105

Xue CY, Li DZ (2011) Use of DNA barcode sensu lato to identify traditional Tibetan medicinal plant Gentianopsis paludosa (Gentianaceae). Journal of Systematic Evolution 49: 267-270

Yang JB, Tang M, Li HT, Zhang ZR, Li DZ (2013) Complete chloroplast genome of the genus Cymbidium: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evolutionary Biology 13: 84

Yao H, Song JY, Ma XY, Liu C, Li Y, Xu XH, Han JP, Duan LS, ChenSL (2009) Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *psbA-trnH* intergenic region. Planta Med 75: 667–669

Yoon CK (1993) Forensic science. Botanical witness for the prosecution. Science 260: 894–895

Yu J, Xue JH and Zhou SL (2011) New universal *matK* primers for DNA barcoding angiosperms. J Syst Evol 49: 176-181.