# DNA barcoding of Andaman Padauk (Pterocarpus dalbergioides Roxb.,) an endemic tree species of Andaman and Nicobar Islands, India

**Jaisankar Iyyappan**
ICAR- Central Island Agricultural Research Institute
**Arun Kumar De**
ICAR- Central Island Agricultural Research Institute
**Prabhu Pari**
pariprabhu594@gmail.com
ICAR- Central Island Agricultural Research Institute
**Manasseh Moses Ezekiel**
ICAR-Central Island Agricultural Research Institute

**Research Article**

**Keywords:** Pterocarpus dalbergioides, DNA barcoding, ITS2, matK, rbcL, endemic species, conservation, Andaman and Nicobar Islands

**Posted Date:** April 26th, 2024

**DOI:** https://doi.org/10.21203/rs.3.rs-4287497/v1



**Additional Declarations:** No competing interests reported.

# Abstract

The research investigates the genetic distinctiveness of *Pterocarpus dalbergioides* Roxb., commonly known as Andaman padauk, an endemic tree species of the Andaman and Nicobar Islands. The study employs DNA barcoding techniques, focusing on three barcode loci (ITS2, *matK*, and *rbcL*), to discern the species from closely related counterparts within the *Pterocarpus* genus. Sampling from 30 distinct locations across the Andaman and Nicobar Islands, genomic DNA isolation, PCR amplification, and sequencing were done. Polymorphism analysis revealed varying degrees of genetic diversity across the three barcode loci, with ITS2 demonstrating the highest discriminatory power. Phylogenetic analyses based on ITS2, *matK*, and *rbcL* sequences elucidated distinct species-specific clusters, reaffirming the endemic nature of *P. dalbergioides* to the Andaman and Nicobar Islands. Notably, ITS2 proved superior in species resolution compared to plastid barcodes *(matK* and *rbcL*). The study highlighted the utility of DNA barcoding in accurately identifying species, particularly in distinguishing closely related taxa within the *Pterocarpus* genus. The findings highlight the ecological and economic significance of *P. dalbergioides* as a valuable timber species and emphasize the importance of DNA barcoding in combating illegal trade and ensuring the sustainable management of endemic tree species. Overall, the research contributes to our understanding of the genetic diversity and conservation of *P. dalbergioides*, offering insights into its evolutionary relationships and aiding in the development of conservation strategies.


# Introduction

*Pterocarpus dalbergioides Roxb.*, also known as Andaman Padauk, is an endemic tree species of the Andaman and Nicobar Islands and belongs to the family Fabaceae (Arunkumar and Joshi 2014). The islands of Andaman and Nicobar have a tropical rainforest shade with highly unique flora because of the tropical humid environment and insular nature of the territory (Rao 1996). The Andaman and Nicobar Islands made of a blended verdure in with components from Indian, Myanmar, Malaysian and endemic flower strains (Mondal and Surekhalandge 2019). Andaman padauk is a dominant species in both semi-evergreen and moist deciduous Andaman forests, contributing high stem density and basal area in comparison to other species (Arunkumar and Joshi 2014; Prasad et al. 2008). This species holds significant cultural importance, serving as an important source of valuable timber and holding the esteemed position of being the state tree of the Andaman and Nicobar Islands (Arunkumar and Joshi 2014). Its versatility extends to ornamental and decorative applications, making it the primary timber tree in the Andaman and Nicobar Islands. The heartwood of this tree is prized for crafting high-quality furniture, contributing to a noteworthy economic transformation on the Island (Rao 2000; Mondal and Surekhalandge 2019). Beyond its utilitarian aspects, the Andaman pad auk wood is esteemed for its diverse array of colors, finding applications in a wide range of uses (Parkinson 1923). This species is recognized by various local names, including Andaman red wood, East Indian mahogany, and narra (Mondal and Surekhalandge 2019). Over the last thirty years, the population of this species has experienced a significant decline of 30% due to extensive harvesting practices (Mondal and Surekhalandge 2019). The detrimental effects of logging activities and a diminished capacity for

regeneration have collectively contributed to a negative trajectory in the population of the Andaman padauk tree. Due to the threat of extinction and low natural seed germination, the tree has been designated as a reserved tree. The FAO has emphasized the need for conservation initiatives in response to the species restricted distribution pattern and economic significance (Prasad et al. 2008). Illicit logging and the illegal timber trade pose substantial challenges on both domestic and international scales, exerting detrimental impacts not only on individual species but also on entire forest ecosystems. Consequently, this issue affects both consumers and producers alike. Global efforts to tackle this problem encompass the implementation of legislation that prohibits or restricts the trade of timber derived unlawfully. (Jiao et al. 2018). The species have been listed vulnerable in the IUCN red list (https://www.iucnredlist.org ). Strict restrictions are imposed globally for trade through the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (Dormontt et al. 2015). Species of *Pterocarpus* are spread globally across countries specifically in tropical Asia, Africa, and South America (Ng 1992; Jansen et al. 1995; Chen et al. 2004). In India variety of *Pterocarpus* species namely *Pterocarpus santalinus*, *P. marsupium*, *P. indicus* and *P. dalbergioides* are present widely across India. In which most of the species has medicinal uses, ornamental and furniture making (Sukhadiya et al. 2019; Arunkumar and Joshi 2014; Senthilkumar et al. 2020). The application of DNA barcoding serves as a pivotal tool in addressing significant challenges, including the identification of illicit logging, illegal timber trade, and forensic wood analysis, which are prevalent issues of considerable concern both at local and global level (Jiao et al. 2018; Hassold et al. 2016; Jiao et al. 2020). The need for species identification in forensic applications is propelled by escalating global apprehensions and the imperative for safeguarding biodiversity. This urgency is further emphasised by notable declines in biological diversity across various geographical, temporal, and biological dimensions. (Tittensor et al. 2014). Various DNA barcodes, including *rpoC1*, *matK*, *rbcL, trnH-psbA*, *trnH*-ITS, *trnH-psbA* and ITS, have previously been used to assess genetic variability in plants. These indicators have been determined to be a strong possibility for identifying families, genus, and species (Malik et al. 2018; De Mattia et al. 2011; De Mattia et al. 2012; Fatima et al. 2019). The Plant Working Group of the Consortium for the Barcode of Life (CBOL) recommended *matK* and *rbcL* sections as universal barcodes for plants (Ford et al. 2009). In the current study, ITS2, *matK*, and *rbcL* barcoding loci were specifically chosen for amplification to elucidate the interspecific regions of a timber tree. The primary objective of this investigation focus on the demarcation of endemic Andaman padauk tree from closely related counterparts by DNA barcoding and support its endemic nature.

## Materials and Methods

# Study Area

The study was undertaken at Andaman and Nicobar Islands, a conglomeration of 572 oceanic islands extending from 6° to 14° North latitudes and from 92° to 94° East longitudes. Situated inside the junction of Bay of Bengal and Indian Ocean on one facet and South China Sea and the Pacific Ocean on the opposite facet, these islands are a part of an archipelago unfold alongside 1120 km North to South. The

predominant part of this archipelago constitutes five most important islands, four in Andaman (North, South, Middle and Little Andaman) and one in Nicobar (Great Nicobar). The weather of Andaman and Nicobar Islands is hot and humid throughout the year. The season can be divided into dry and rainy seasons with a mean annual temperature of 26.4°C. The average annual precipitation is around 3,100 mm, with mean relative humidity ranging from 65 to 89 percent.

# Plant Materials and Collection Sites

In the present study, candidate plus trees of Andaman padauk were identified from 30 distinct locations encompassing South Andaman, North Andaman, Middle Andaman, Little Andaman, and Katchal (Nicobar group of islands) of the Andaman and Nicobar Islands (Table 1). Geo reference map was prepared for collection sites of Andaman padauk, the Andaman and Nicobar Island boundary shape file (.shp) was downloaded from http://www.diva-gis.org/gdata. The downloaded shape files (Country boundaries) are imposed in ArcGIS 10.5v (source from SAC-ISRO).

Criteria for selection included trees exceeding 20 meters in height, possessing a clear bole with a robust crown, and exhibiting a girth at breast height (GBH) ranging from 200 cm to 450 cm, along with a diameter at breast height (DBH) exceeding 100 cm. The sampling duration for seed collection extended from April to June 2019. Seed samples were gathered from the identified plants meeting the above specified criteria. Subsequently, these collected seed samples underwent an overnight soaking in water before being sowed in the mother bed in order to facilitate germination. To isolate DNA extraction, young leaf samples were collected from the seedlings raised in the nursery.

Table 1
Locations of collection sites of Andaman padauk

| Acc. No | Place | Latitude | Longitude | Altitude (M) | Tree Height (M) | GBH (CM) | District |
|---|---|---|---|---|---|---|---|
| 1 | Jirkatang | 11˚50’20”N | 92˚39’14”E | 15 | 24 | 320 | SA |
| 2 | Sippighat I | 11˚36’41”N | 92˚40’51”E | 24 | 22 | 210 | SA |
| 3 | Sippighat II | 11˚36’42”N | 92˚40’57”E | 28 | 21 | 220 | SA |
| 4 | Sippighat III | 11˚36’43”N | 92˚40’44”E | 14 | 23 | 190 | SA |
| 5 | Rutland | 11˚26’29”N | 92˚38’25”E | 157 | 23 | 210 | SA |
| 6 | Burmanallah | 11˚32’56”N | 92˚43’08”E | 38 | 24 | 396 | SA |
| 7 | Wrightmyo | 11˚47’11”N | 92˚42’34”E | 13 | 21 | 321 | SA |
| 8 | Mannarghat | 11˚45’40”N | 92˚42’22”E | 15 | 19 | 417 | SA |
| 9 | Hutbay | 10˚38’37”N | 92˚30’42”E | 78 | 22 | 195 | SA |
| 10 | Nimbudera | 12˚43’24”N | 92˚53’08”E | 24 | 24 | 420 | NA |
| 11 | Gandhi jetty | 12˚18’28”N | 92˚47’16”E | 7 | 26 | 235 | NA |
| 12 | Nimbudera II | 12˚43’10”N | 92˚53’02”E | 22 | 22 | 178 | NA |
| 13 | Kadamtala | 12˚20’00”N | 92˚47’04”E | 49 | 26 | 390 | NA |
| 14 | Adajig | 12˚15’30”N | 92˚48’05”E | 36 | 34 | 198 | NA |
| 15 | Bakultala | 12˚30’18”N | 92˚51’45”E | 7 | 22 | 148 | NA |
| 16 | Rangath | 12˚30’11”N | 92˚53’45”E | 87 | 21 | 198 | NA |
| 17 | Bakultala II | 12˚29’58”N | 92˚52’24”E | 27 | 27 | 150 | NA |
| 18 | Billiground | 12˚40’06”N | 92˚52’53”E | 12 | 24 | 190 | NA |
| 19 | Kadamtala II | 12˚30’39”N | 92˚46’41”E | 30 | 22 | 135 | NA |
| 20 | Lamiya Bay | 13˚12’19”N | 93˚01’22”E | 32 | 26 | 480 | NA |
| 21 | Buddha Nallah | 12˚33’21”N | 92˚52’17”E | 17 | 23 | 185 | NA |
| 22 | Lampagram | 13˚14’29”N | 93˚00’47”E | 39 | 21 | 210 | NA |

| Acc. No | Place | Latitude | Longitude | Altitude (M) | Tree Height (M) | GBH (CM) | District |
|---|---|---|---|---|---|---|---|
| 23 | Kalpong | 13˚13’59”N | 92˚57’48”E | 17 | 24 | 420 | NA |
| 24 | Pembroke Bay | 13˚10’56”N | 92˚49’43”E | 24 | 24 | 375 | NA |
| 25 | Kalapahad | 13˚09’55”N | 92˚49’33”E | 22 | 21 | 195 | NA |
| 26 | Prolobjig | 12˚23’44”N | 92˚53’19”E | 16 | 24 | 480 | NA |
| 27 | Long Island | 12˚23’34”N | 92˚56’06”E | 73 | 23 | 390 | NA |
| 28 | SastriNallah | 12˚14’55”N | 92˚49’35”E | 78 | 20 | 167 | NA |
| 29 | Kalighat | 13˚07’35”N | 92˚57’57”E | 78 | 22 | 325 | NA |
| 30 | Katchal | 7˚57’18”N | 93˚21’53”E | 108 | 23 | 275 | Nicobar |

• NA – North Andaman

• SA – South Andaman

# Genomic DNA Isolation

DNA isolated from the leaf of one-year-old seedlings. Around 2 gm of leaf samples from the respective accessions were used for genomic DNA isolation. Genomic DNA was extracted by CTAB extraction method using a commercially available DNA extraction kit (HiPurA$^{TM}$Plant DNA Isolation Kit (CTAB Method), HiMedia Laboratories Pvt. Ltd., Mumbai, India) as per the protocol mentioned by the manufacturer. The isolated DNA samples were dissolved in Tris-EDTA (TE) buffer. The quality and concentration of isolated DNA samples were checked in a Bio Spectrometer (Eppendorf, Hamburg, Germany) at 260 and 280 nm wavelengths. DNA samples with 260/280 ratio between 1.7 to 1.9 were considered for further processing and stored at -80°C.

# PCR Amplification and Sequencing of Candidate DNA Barcodes

For DNA barcoding of Andaman padauk, three established DNA barcoding loci namely Ribulose-1,5-bisphosphate carboxylase large subunit (*rbcL*), Maturase K (*matK*) and Internal Transcribed Spacer 2 (ITS2) were chosen. DNA amplification was carried out in a 25 µl reaction mixture containing 2.5 µl 10X *Taq* reaction buffer with 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 1 µM each forward and reverse primer, 1 IU *Taq* DNA polymerase (GCC Biotech Pvt. Ltd., West Bengal, India) and approximately 50 ng genomic DNA. The sequences of the primers used are presented in Table 2. PCR reactions were carried out in a thermal cycler (Prima-96™, HiMedia Laboratories Pvt. Ltd., Mumbai, India) with the following cycling conditions; initial denaturation at 95 ˚C for 5 minutes, 35 cycles of denaturation at 95 ˚C for 30 seconds, annealing at

50 ˚C for 45 seconds, extension at 72 ˚C for 1 minute and a final extension at 72 ˚C for 7 minutes. Confirmation of the amplicons was done in 1.5% agarose gels following which the products were purified and sent to a commercial company (Eurofins Scientific India Pvt. Ltd., Bengaluru, India) for Sanger sequencing in both directions. The generated sequences were edited using Sequencher v 5.4.6 (Gene Codes Corporation, USA).

Table 2
Details of the primers used for barcoding

| Barcode loci | Primer sequences (5'-3') |
|---|---|
| **rbcL** | F: ATGCGATACTTGGTGTGA |
| | R: TAGCCCCGCCTGACCTGA |
| **matK** | F: CCRTCATCTGGAAATGTTGGTT |
| | R: GCTRTRATAATGAGAAAGATTTCTGC |
| **ITS2** | F: ATGTCACCACAAACAGAAAC |
| | R: TCGCATGTACCTGCAGTAGC |

# Bioinformatics analysis

Ribulose-1,5-bisphosphate carboxylase large subunit (*rbcL*), Maturase K (*matK*) and Internal Transcribed Spacer 2 (ITS2) sequences of 18 different *Pterocarpus* species from different parts of the world were retrieved from GenBank (www.ncbi.nlm.nih.gov) along with that 30 Andaman padauk barcoding (*rbcL*, *matK* and ITS2) sequences generated in our study and *Eucalyptus globulus* subsp. *maidenii* (ITS2), *Eucalyptus haemastoma* (*matK*), *Eucalyptus globulus* (*rbcL*) were used as outgroup. All the sequences aligned using ClustalW programme (Thompson et al. 1997). in MEGAX (Kumar et al. 2018). Polymorphism and population diversity parameters like number of polymorphic sites, haplotype number and diversity, nucleotide diversity and DNA divergence between populations were computed by DnaSp v 6 (Rozas et al. 2017). Maximum Composite Likelihood model (Tamura et al. 2004). implemented in MEGAX was used to compute pair-wise genetic distance among different sequences. Phylogenetic relationship among the sequences was inferred using the Neighbor-Joining method (Saitou and Nei 1987). with the Tamura-Nei model (Tamura et al. 1993). as implemented in MEGAX following 1,000 bootstrap replications. Bayesian phylogenetic analysis of the sequences was established using MCMC model in BEAST v1.10.4 (Suchard et al. 2018). To determine the evolutionary relationship among different sequences, median-joining networks were constructed in PopART ver. 1.7 (Leigh and Bryant 2015). with default settings. For phylogenetic tree construction, we trimmed extra nucleotides from our sequences and GenBank retrieved sequences to make a homogeneous length.

# Results

Sequence information of the three barcoding loci (ITS2 complete sequence, *rbcL* partial sequence, and *matK* partial sequence) were generated for thirty Andaman padauk isolates from different geographical regions from all three districts of Andaman and Nicobar Islands (Fig. 1). (South Andaman 9, North and Middle Andaman 20, and Nicobar 1). Two chloroplast DNA barcoding loci (*matK* and *rbcL*) recommended by the Consortium for the Barcode of Life (CBOL) and one nuclear barcoding locus (ITS2) (Chen et al. 2004). were used for molecular DNA barcoding of Andaman padauk. The PCR amplified products of rbcL, matK, and ITS2 are presented in Fig. 2.

# 2.1 Polymorphism Analysis

Polymorphism Analysis of ITS2 sequence of the 30 isolates detected in which 2 variable /polymorphic sites including one singleton variable site and one parsimony informative site were found. A total of 30 haplotypes were observed with haplotype diversity of 0.421 ± 0.087 (HD ± SD). The nucleotide diversity (Pi ± SD) of the sequences were 0.00181 ± 0.00041. The ITS2 based polygenetic relationship is presented in Fig. 3 (a).

Three well defied clusters were observed, in which two clusters were found in north and middle Andaman and the sequences under the third cluster were distributed in all the three districts of Andaman and Nicobar. The *matK* polygenetic analysis indicate seven polymorphic sites with two singleton variable sites and five parsimony informative sites. Total of five haplotypes are detected with haplotype diversity (HD ± SD) of 0.662 ± 0.073. The nucleotide diversity (Pi ± SD) of the sequences was 0.00233 ± 0.00036. The geographical distribution of the sequences under five haplotypes are presented in Fig. 3 (b). Five well defined clusters of each haplotype were detected, one from south Andaman specific cluster and another three clusters were from north and middle Andaman, the rest three clusters were not of any region specific. The *rbcL* based polygenetic analysis indicate only two polymorphic sites, in which both were parsimony informative sites. Total of three haplotypes were detected with haplotype diversity of (HD ± SD) 0.480 ± 0.093. Nucleotide diversity of the sequence was (Pi ± SD) 0.00097 ± 0.00020. The geographical distribution of the haplotypes is presented in Fig. 3 (c). No region specific clusters were found. The concatenated sequences of the three markers were detected in fifteen haplotypes, within haplotype diversity of 0.915 ± 0.031. The nucleotide diversity was 0.0016 ± 0.00019. The polygenetic relationship and geographical distribution of the haplotype is presented in Fig. 3 (d). Total of nine haplotypes were found in north and middle Andaman specific, three were south Andaman specific and rest three were not region specific.

# ITS2

The neighbor-joining phylogenetic tree and Bayesian phylogenetic tree, both based on ITS2 data, demonstrate distinct species specific clusters, revealing separate groupings for different species of genus *Petrocorpus.* shown in Fig. 4. Phylogenetic analysis revealed the close relationship of Andaman padauk (*P.dalbergioides*) with *P. marsupium* from India and *P. santalinus* from China and India. Moreover, the ITS-based network tree demonstrated distinct clustering, with all 30 Andaman isolates forming separate and identifiable clusters. *P. santalinus* from India and China; *P. marsupium* from Coimbatore; *P.*

*tinctorius* from Republic of Congo and Tanzania; *P. acapulcensis* of Venezuela and Mexico; *P.soyauxii* from Cameroon and Gabon; *P. indicus* from India, Indonesia and Malaysia; *P.angolensis* from Angola and Zimbabwe: *P. rohri* from Columbia, Gautemala, Mexico, Brazil and Bolivia; *P. mildbraedii* of Nigeria and Tanzania; *P.erinaceus* of Nigeria and Ghana formed separate groupings with respective species identity with very few exception. i.e. *P. ternatus* and *P. zehntneri* from Brazil grouped with *P. rohri.*

# matK

The evolutionary relationship between Andaman padauk (*P.dalbergioides*) and various *Pterocarpus* species worldwide was analyzed using *matK*-based data, depicted in Fig. 5. The analysis revealed two major clades: The Clade 1 includes *P. dalbergioides* and *P. santalinus* from India and China, *P. macrocarpus* from Thailand, *P. marsupium* from India, *P. angolensis* from Angola and South Africa, and other related species from diverse regions such as Myanmar, Cambodia, Malaysia, Vietnam, Indonesia, Papua New Guinea, and Puerto Rico. The Clade 2 encompasses *Pterocarpus* species like *P. acapulcensis*, *P. tinctorius*, *P. rohri*, *P. soyauxii*, *P. ternatus*, *P. erinaceus*, *P. officinalis*, *P. orbiculatus*, *P. mildbraedii*, and others, spread across regions like Venezuela, Congo, Tanzania, Brazil, Colombia, Mexico, Guatemala, Bolivia, Zimbabwe, Mozambique, Nigeria, Ghana, Puerto Rico, and more. The specific genetic analysis showed certain Andaman padauk isolates (*P. dalbergioides* Andaman S1, S3, S21, S30, S6, S28, and S29) to be genetically close to *P. santalinus* from India and China. Another group of Andaman padauk isolates (*P. dalbergioides* Andaman S7, S20, S23, S24, S25, S26, and S27) appeared genetically proximate to *P. santalinus* of India and China, along with distinct genetic clusters observed in other isolates of *P. dalbergioides* from the Andaman region. In the network trees depicting the relationship among *Pterocarpus* species, Andaman padauk (*P. dalbergioides*) clustered with *P. santalinus*, *P. marsupium*, *P. indicus*, *P. macrocarpus*, and *P. angolensis*. Overall, the matK-based evolutionary analysis highlighted distinct genetic groupings and relationships among various *Pterocarpus* species, elucidating the genetic closeness of Andaman padauk to specific counterparts within the *Pterocarpus* genus across different geographic locations.

# rbcL

A neighbor-joining tree was generated using *rbcL* gene sequences, the analysis unveiled distinct clustering patterns unlike ITS2 and *matK* there is no clear species delineation was observed Fig. 6. Six different clusters were formed without any species-specific clusters. cluster 1 included Andaman padauk (*P.dalbergioides*) isolates S1, S2, S9, S11, S12, S22, S23, S24, and S28, showing genetic proximity to *P.soyauxii* of Cameroon, *P.officinalis* of Puerto Rico, and *P. macrocarpus* of Myanmar. Cluster 2 Isolates S29, S18, S7, S14, and S21 clustered with *P. indicus* of Indonesia, *P. marsupium* of India, *P. santalinus* of India, and *P.tinctorius* of Congo. Group 3 comprised isolates S15 and S25, closely associated with *P.indicus* of Malaysia, *P. marsupium* of India, *P. officinalis* of Guatemala, and *P.rohrii* of Brazil. Cluster 4, isolates S4, S10, and S19 (Group 4) displayed genetic similarities with *P. angolensis* of Angola and Mozambique, *P. santalinus* of China, *P.macrocarpus*of Myanmar, and *P.indicus* of China. Group 5, represented by S5, S13, S20, and S30, showed close relations to *P.acapulcensis* of Venezuela, *P.santalinus* of India and China, *P. marsupium* of India, and *P.indicus*of China. Finally, Group 6

encompassing isolates S3, S8, S17, S27, S6, S16, and S26 exhibited proximity to various species including *P.indicus* of Papua New Guinea, *P.mildbraedii*of Nigeria, *P.orbiculatus* of Mexico, *P.rohrii*found across Mexico, Guatemala, Bolivia, Brazil, Colombia, and French Guiana, *P. ternatus* of Brazil, and *P. santalinus* of India.

# Discussion

The primary objective of this investigation is to elucidate the distinctiveness of the endemic Andaman padauk species (*Pterocarpus dalbergioides*) in comparison to other closely related *Pterocarpus* species, with help of DNA barcoding. DNA barcoding proves to be an ideal tool for the discriminate and identification of species. In this study, three DNA barcoding regions were employed, namely the nuclear ribosomal DNA region Internal Transcribed Spacer 2 (ITS2), and two plastid DNA barcodes, ribulose-1,5-bisphosphate carboxylase large subunit (*rbcL*) and maturase K (*matK*). Notably, ITS2 demonstrated the highest discriminatory power among the three barcodes examined. The superior identification capabilities of the nuclear DNA region ITS2 over the plastid barcodes *rbcL* and *matK* are in line with findings from previous studies. This reaffirms the effectiveness of ITS2 in distinguishing between closely related species within the *Pterocarpus* genus (Chen et al. 2010; Shaw et al. 2007; Ly et al. 2015; Yao et al. 2010; Pang et al. 2013; Han et al. 2016). While the chloroplast DNA regions *rbcL* and *matK* were initially advocated as core barcodes for seed plants by the Consortium for the Barcode of Life Plant Working Group, our study revealed a lower species resolution for *Pterocarpus* when using these two regions, particularly the *rbcL* region. This observation aligns with the findings of a study conducted by (Jiao et al. 2018). Despite their widespread use in phylogenetic analyses, with over 130,000 sequences available in Genbank, both *rbcL* and *matK* exhibited limitations in accurately resolving species within the *Pterocarpus* genus. Kress et al. (2005) previously demonstrated that the *rbcL* sequence evolves slowly, and this barcode is recognized for having the lowest divergence among studied plastid genes in flowering plants. Consequently, on average, it may not be sufficiently informative for species-level identification. Additionally, *matK* has demonstrated varying success rates in discriminating species across different taxonomic groups. For instance, it effectively discriminates more than 90% of species in the Orchidaceae, but its discriminatory power falls below 49% for species in the nutmeg family. In contrast, our study highlights the superior performance of the nuclear DNA region ITS2 in providing enhanced species resolution for *Pterocarpus*. These observations emphasize the need for a careful selection of DNA barcodes based on their efficacy within specific taxonomic groups. In this context, the ITS2 loci emerge as the most suitable single barcode region for distinguishing *Pterocarpus* (Chase et al. 2007; Kang et al. 2017) DNA extracted from wood of *Pterocarpus* also shows better amplification and high discrimination with ITS2 (But et al. 2023). In this investigation, a comprehensive analysis was undertaken by comparing the Internal Transcribed Spacer 2 (ITS2) sequences of 30 Andaman padauk accessions with those of 16 closely related *Pterocarpus* species obtained from GenBank, originating from diverse global locations. The resulting Neighbor-Joining (NJ) tree, constructed based on the ITS2 sequence data, revealed distinct species-specific clusters. Notably, all 30 Andaman padauk accessions formed a unique cluster clearly different from other related species and similar reports were not found

globally, it providing molecular evidence that supports the endemism of *Pterocarpus dalbergioides* to the Andaman and Nicobar Islands. This finding is consistent with previous research conducted by (Jaisankar et al. 2020; Prasad et al. 2008). Similarly, DNA barcoding techniques were employed to confirm the distinctiveness of a narrowly distributed endemic tree within the Rubiaceae family on Palau Island in Micronesia. The results indicated that this tree is likely a distinct species within the *Timonius* genus, as revealed in the study conducted by (Costion et al. 2016). Beyond its taxonomic significance, the Andaman padauk (*P. dalbergioides*) holds ecological and economic importance as an endemic tree species with high timber value. Recognizing the critical role of DNA barcoding in species identification, this study emphasizes its potential application in curbing illegal trading and ensuring accurate species identification in the market. Illegal trading poses a substantial threat to the Andaman padauk due to its valuable timber. By leveraging DNA barcoding, a robust and reliable method for species identification, we can establish a stringent mechanism to verify the authenticity of Andaman padauk products in the market. This approach acts as a powerful deterrent against the misrepresentation of other tree species as Andaman padauk, mitigating the risk of unauthorized trade and ensuring the sustainability of the species.

# Conclusion

In conclusion, the study successfully utilized DNA barcoding to distinguish Andaman padauk (*Pterocarpus dalbergioides*) from closely related species, emphasizing the efficacy of ITS2 in species resolution. Molecular evidence supports the endemic nature of Andaman padauk, emphasising its ecological and economic significance. DNA barcoding offers a robust method for accurate species identification, crucial for conservation and trade regulation.

# Declarations

### Ethical approval:

This article contains no studies on human or animal objects

### Competing Interests:

The authors declare that they have no competing interests.

### Acknowledgements

The authors express their gratitude to Department of Environment, Forests and Climate change for approval to collect the seeds for this study. The authors are also thankful to the Director of ICAR-CIARI, Port Blair for providing Institute facilities and constant encouragement during the study period.

### Funding

This study was supported by the Department of Biotechnology (DBT), Ministry of Science & Technology, Government of India, India (grant number BT/PR29179/FCB/125/7/2018).

**Authors' contributions:**

IJ conceived and supervised the overall execution of the project, PP prepared the manuscript, AKD & EMM assisted in analysis & sample collection.

# References

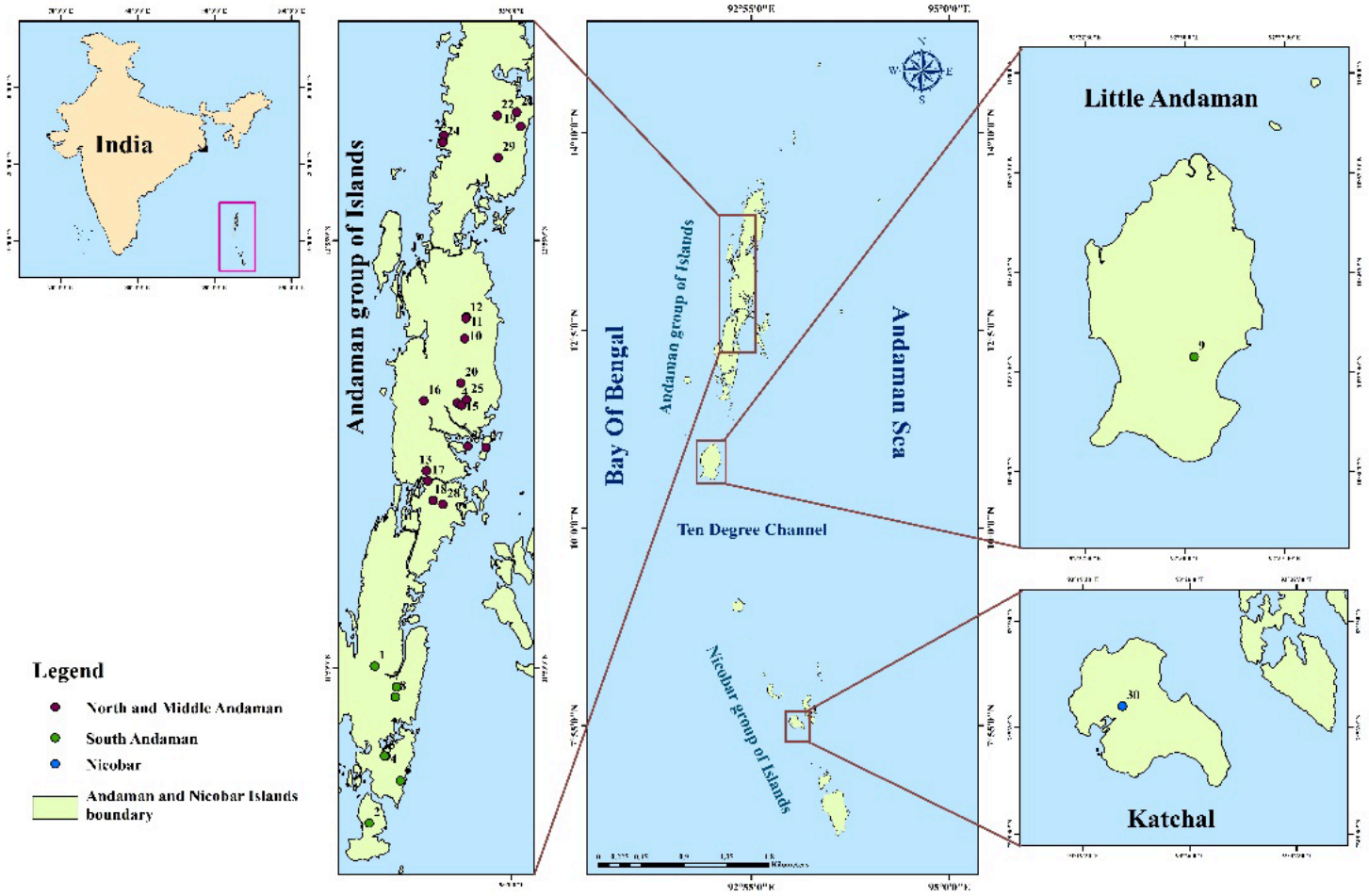
1. Arunkumar AN, Joshi G (2014) *Pterocarpus santalinus* (Red Sanders) an Endemic, Endangered Tree of India: Current Status, Improvement and the Future. *J. Trop. For. Sci* 4(2) 1-10 https://doi.org/10.31357/jtfe.v4i2.2063
2. Barstow M (2018) *Pterocarpus dalbergioides*. The IUCN Red List of Threatened Species:http://dx.doi.org/10.2305/IUCN.UK.2018-1.RLTS.T33261A67802958.en (assessed on 10 Oct 2017).
3. But GW, Wu HY, Siu TY et al (2023) Comparison of DNA extraction methods on CITES-listed timber species and application in species authentication of commercial products using DNA barcoding. *Sci. Rep.* 13(1):151. https://doi.org/10.1038/s41598-022-27195-7
4. Chase MW, Cowan RS, Hollingsworth PM et al (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon*.56(2):295-9. https://doi.org/10.1002/tax.562004
5. Chen QD, Li XM, Zeng J, L et al (2004) Introduction status and perspectives of *Pterocarpus* in China. *Sci. Technol.* 2, 38-41.
6. Chen S, Yao H, Han J et al (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PloS One* 5(1): e8613. https://doi.org/10.1371/journal.pone.0008613
7. Costion CM, Kress WJ, Crayn DM (2016) DNA barcodes confirm the taxonomic and conservation status of a species of tree on the brink of extinction in the Pacific. *PLoS One* 11(6): e0155118. https://doi.org/10.1371/journal.pone.0155118
8. De Mattia F, Bruni I, Galimberti A et al (2011) A comparative study of different DNA barcoding markers for the identification of some members of Lamiacaea. *Food Res. Int.* 44(3), 693-702 https://doi.org/10.1016/j.foodres.2010.12.032 .
9. De Mattia F, Gentili R, Bruni I et al (2012) A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. *Bot. J. Linn.* 169(3): 518-29. https://doi.org/10.1111/j.1095-8339.2012.01251.x
10. Dormontt, EE, Boner M, Braun B et al (2015) Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biol. Conserv.* 191: 790-8. https://doi.org/10.1016/j.biocon.2015.06.038

11. Fatima S, Zaidi NW, Verma N et al (2019) Comparative analysis of different DNA barcoding loci to discriminate among three Brassica species. *Plant Syst. Evol*. 305(9): 741-750.
12. Ford CS, Ayres KL, Toomey N et al (2009). Selection of candidate coding DNA barcoding regions for use on land plants. *Bot. J. Linn.* 159(1): 1-1. https://doi.org/10.1111/j.1095-8339.2008.00938.x
13. Han J, Pang X, Liao B et al (2016). An authenticity survey of herbal medicines from markets in China using DNA barcoding. *Sci. Rep.* 6(1): 18723. https://doi.org/10.1038/srep18723
14. Hassold S, Lowry PP, Bauert MR (2016) DNA barcoding of Malagasy rosewoods: towards a molecular identification of CITES-listed *Dalbergia* species. *PLoS One*. 11(6): e0157881. https://doi.org/10.1371/journal.pone.0157881
15. Jaisankar I, Sundaram S, Jerard BA et al (2020) Seed microstructures of the Andaman padauk (*Pterocarpus dalbergoides*). *Curr. Sci.* 119(3): 562-6. https://www.jstor.org/stable/27138894
16. Jansen PCM, Westphal E, Wulijarni-Soetjipto N et al (1995) Plant resources of South-East Asia 5(2). Timber trees: Minor commercial timbers. Leiden: Backhuys Publishers, 655 p. https://ui.adsabs.harvard.edu/link_gateway/1996KewBu..51..825G/doi:10.2307/4119744
17. Jiao L, Lu Y, He T et al (2020). DNA barcoding for wood identification: Global review of the last decade and future perspective. *IAWA J.* 41(4): 620-43. https://doi.org/10.1163/22941932-bja10041
18. Jiao L, Yu M, Wiedenhoeft AC et al (2018). DNA barcode authentication and library development for the wood of six commercial *Pterocarpus* species: the critical role of *Xylarium* specimens. *Sci. Rep.* 8(1): 1945. https://doi.org/10.1038/s41598-018-20381-6
19. Kang Y, Deng Z, Zang R et al (2017). DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. *Sci. Rep.*7(1): 12564. https://doi.org/10.1038/s41598-017-13057-0
20. Kress WJ, Wurdack KJ, Zimmer EA et al (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences* 102(23): 8369-74. https://doi.org/10.1073/pnas.0503123102
21. Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6): 1547. https://doi.org/10.1093%2Fmolbev%2Fmsy096
22. Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. *Mol Biol Evol.* 6(9): 1110-6. doi: 10.1111/2041-210X.12410
23. Lv T, Teng R, Shao Q et al (2015). DNA barcodes for the identification of *Anoectochilus roxburghii* and its adulterants. *Planta*. 242: 1167-74. https://doi.org/10.1007/s00425-015-2353-x
24. Malik S, Kumar M, Kaur S, et al (2018) Barcoding of Indian Berberis species based on *rbcL* and *matK* gene fragments. *3 Biotech* 8(3): 160.
25. Mondal S, Surekhalandge V (2019) Ecological Status of *Pterocarpus dalbergioides*, the Endemic Timber Tree of Andaman and Nicobar Islands: Current Status and Management Issues. *Forests*, 10(8): 712.
26. Ng FSP (1992) *Pterocarpus indicus*-the majestic N-fixing tree, Nitrogen Fixing Tree Association (NFTA) *Pterocarpus indicus*-the majestic N-fixing tree *NFT Highlights* (92-02).

27. Pang X, Shi L, Song J et al (2013). Use of the potential DNA barcode ITS2 to identify herbal materials. *J.Nat. Med.* 67: 571-5. https://doi.org/10.1007/s11418-012-0715-2
28. Parkinson CE (1923) The forest flora of the Andaman Islands: An account of the trees, shrubs and principal climbers of the Islands. Superintendent, Government Central Press.
29. Patil UH, Dattatraya KG (2011) *Pterocarpus marsupium*: A valuable medicinal plant in diabetes management. *Int J Appl Biol Pharm* 2 (3): 6-13.
30. Prasad R, Maikhuri RK, Singh K et al (2008) Conservation of *Pterocarpus dalbergioides*: A State Tree of Andaman and Nicobar Islands, India. *Int J Biodivers Sci Ecosyst Serv Manag*, 4(3): 125-131.
31. Rao A (1996). V. S. S. N. Flora of Andaman and Nicobar Islands. In Botanical Survey of India.
32. Rao RS (2000) Resin tapping and oil yield of *Pterocarpus dalbergioides* Roxb. in relation to season and tree diameter. *Indian For.* 126(9): 962-969.
33. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC et al (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol.* 34(12) : 3299-302. https://doi.org/10.1093/molbev/msx248
34. Saitou N, Nei M et al (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4(4): 406-25. https://doi.org/10.1093/oxfordjournals.molbev.a040454
35. Senthilkumar N, Baby Shalini T, Lenora et al (2020) *Pterocarpus indicus* Willd: A Lesser Known Tree Species of Medicinal Importance. *Asian J. Res. Bot.*3(4): 20-32.
36. Shaw J, Lickey EB, Schilling EE et al (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. of Bot.* 94(3): 275-88. https://doi.org/10.3732/ajb.94.3.275
37. Suchard MA, Lemey P, Baele G et al (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 4(1): vey016. doi: 10.1093/ve/vey016
38. Sukhadiya M, Dholariya C, Behera LK et al (2019). Indian kino tree (*Pterocarpusmarsupium* roxb.): biography of excellent timber tree species. *MFP NEWS* 29(1): 4-8.
39. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol.* 10(3): 51cc2-26. https://doi.org/10.1093/oxfordjournals.molbev.a040023
40. Tamura K, Nei M, Kumar S (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl Acad Sci. U.S.A.* 101(30):11030-5. https://doi.org/10.1073/pnas.0404206101
41. Thompson JD, Gibson TJ, Plewniak F et al (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25(24): 4876-82. https://doi.org/10.1093/nar/25.24.4876
42. Tittensor DP, Walpole M, Hill SL et al (2014) A mid-term analysis of progress toward international biodiversity targets. *Science*. 346 (6206): 241-4. https://doi.org/10.1126/science.1257484

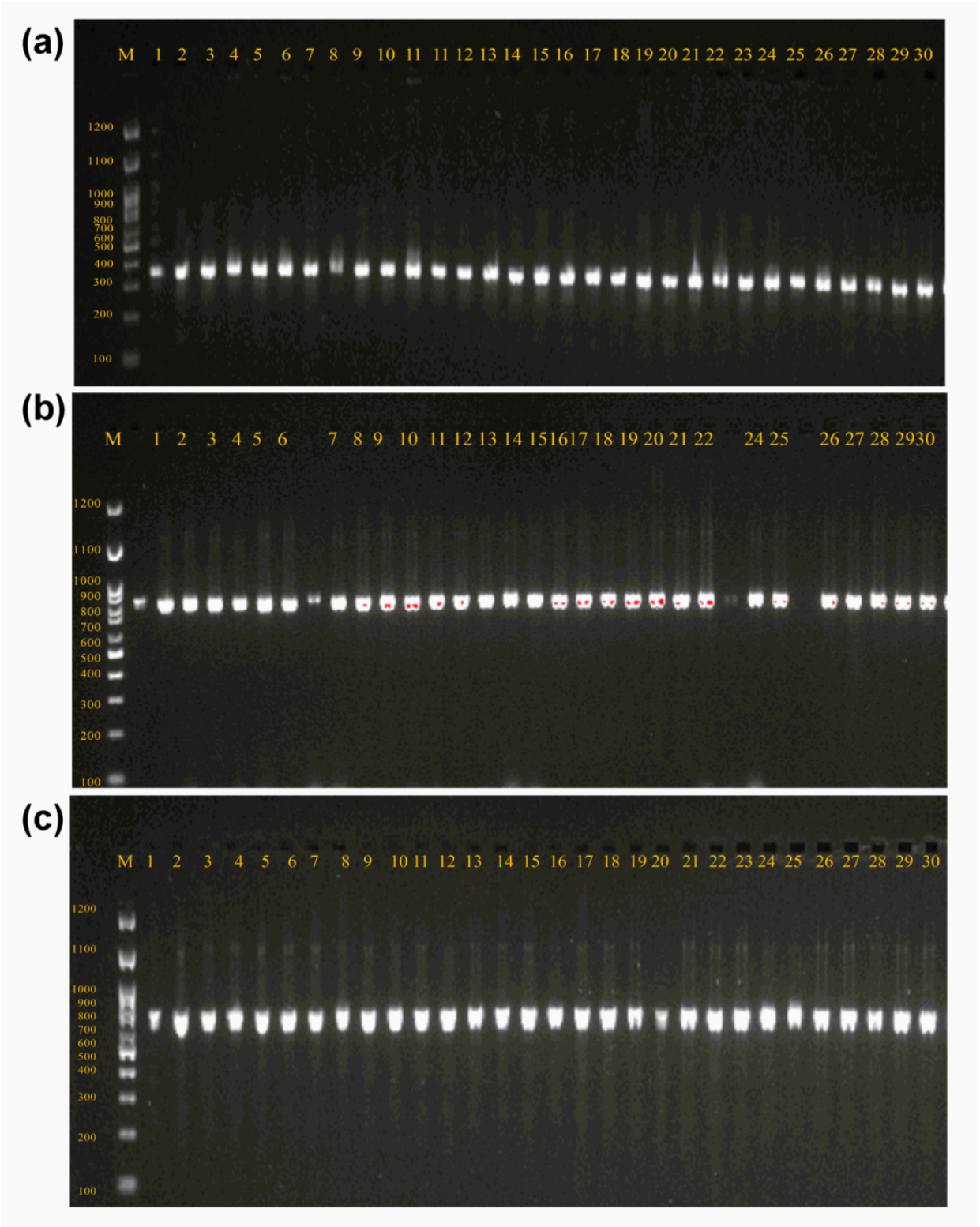
43. Yao H, Song J, Liu C et al (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. *PloS One*. 5(10): e13102. https://doi.org/10.1371/journal.pone.0013102

# Figures



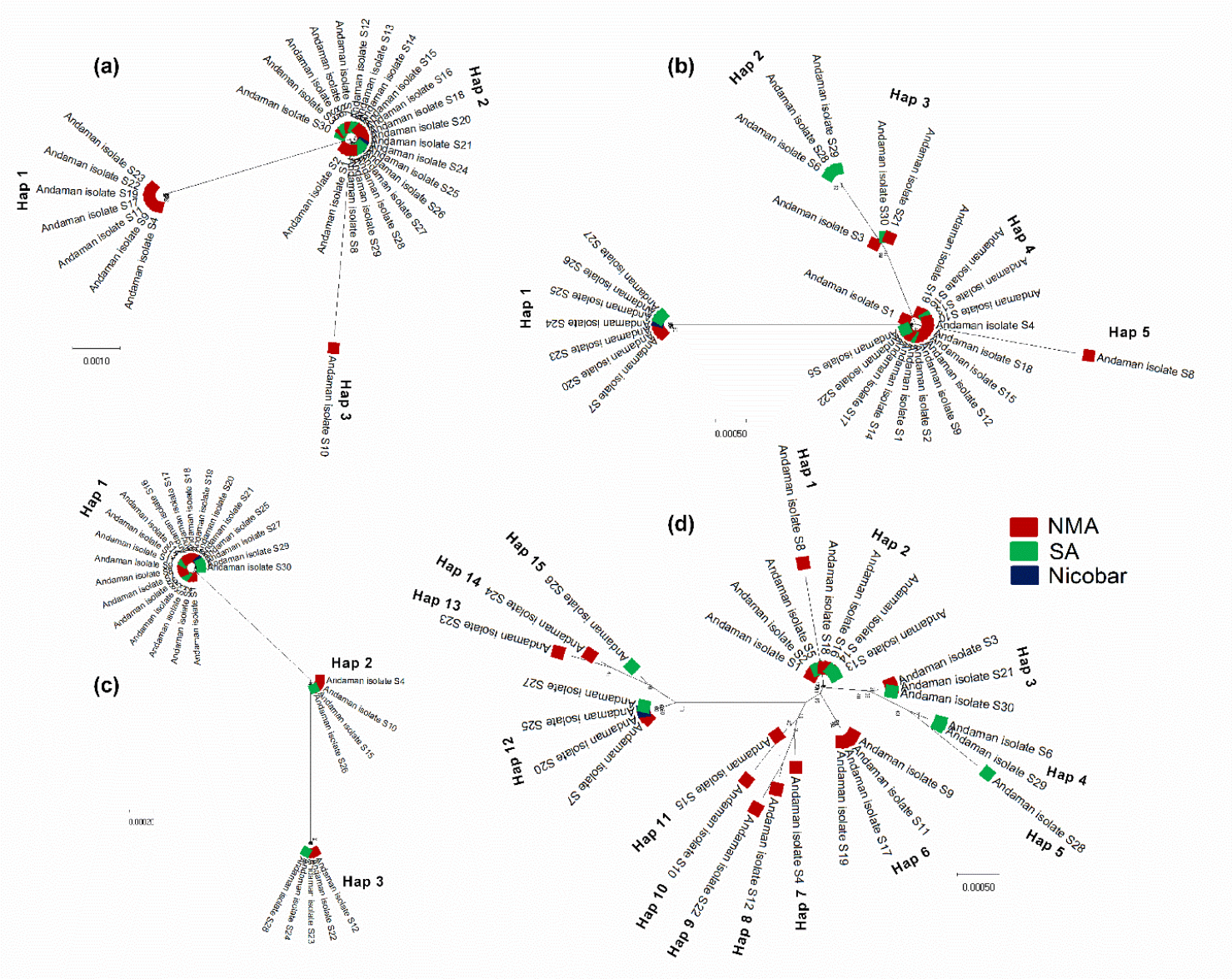

## Figure 1

Geo reference map of collection sites of Andaman padauk (*P. dalbergioides*).



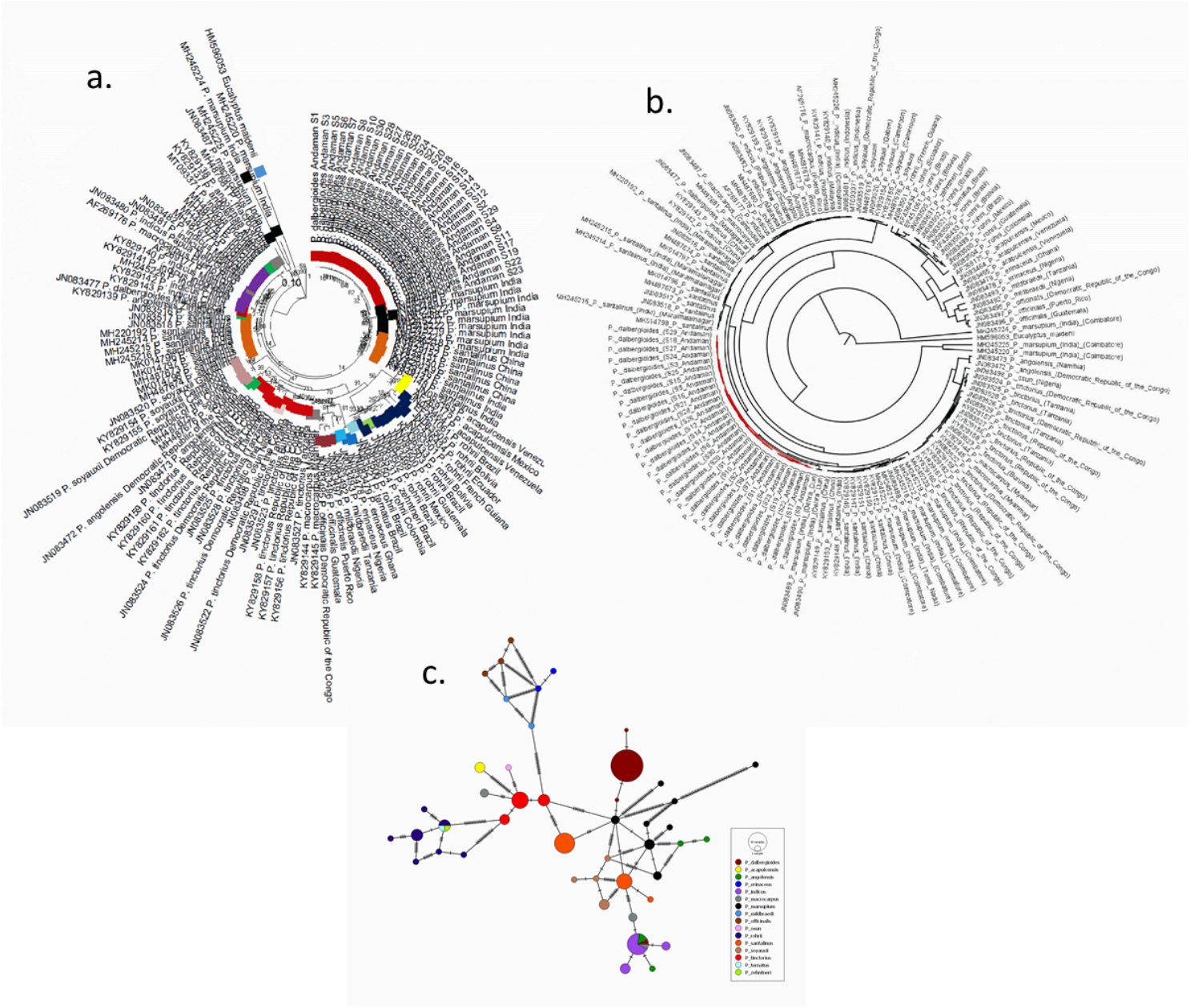

## Figure 2

Gel image of Barcoding primers a) ITS2, (b) *matK*, (c) *rbcL.*



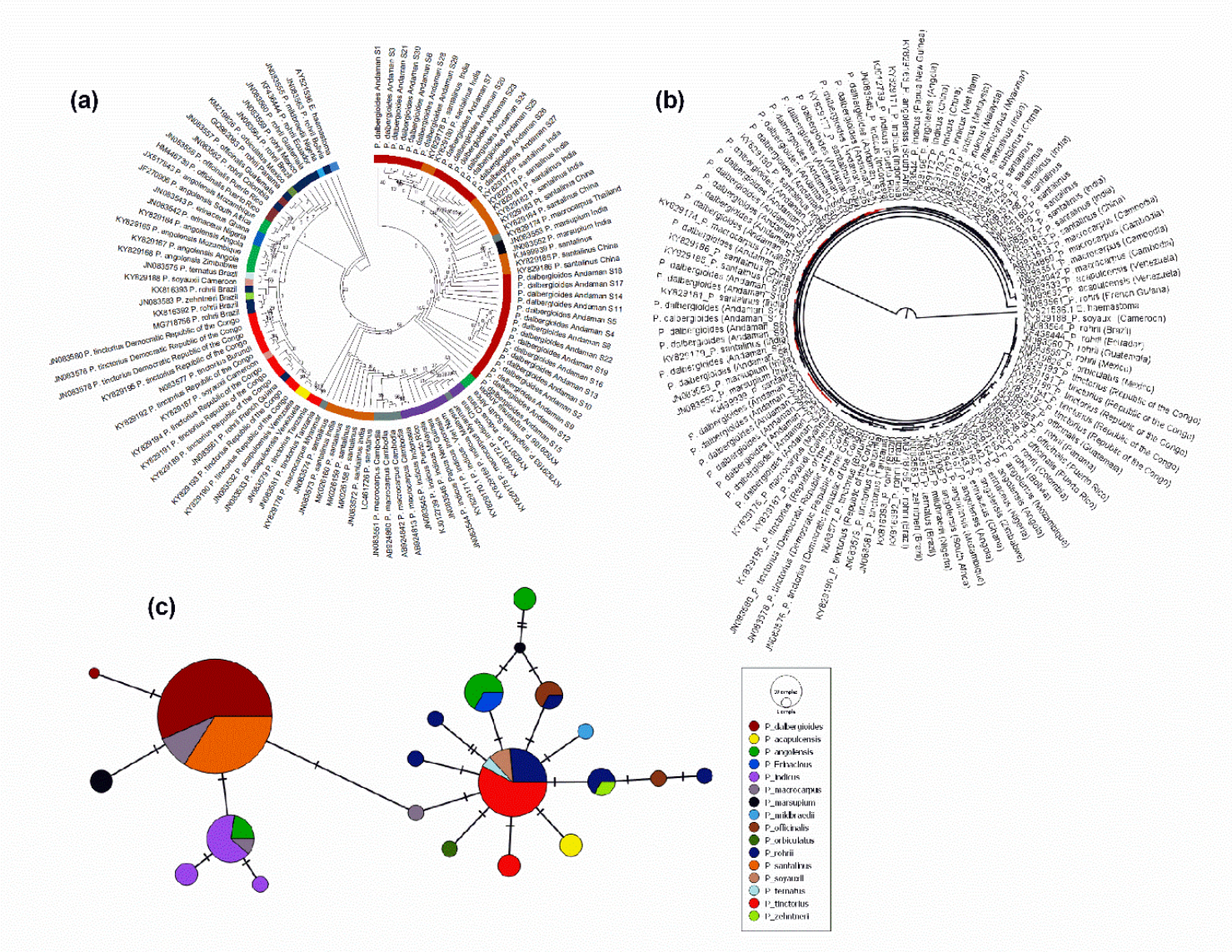

## Figure 3

Polymorphism Analysis of (a) ITS2, (b) *matK,* (c)*rbcL*, (d) Concatenated sequences



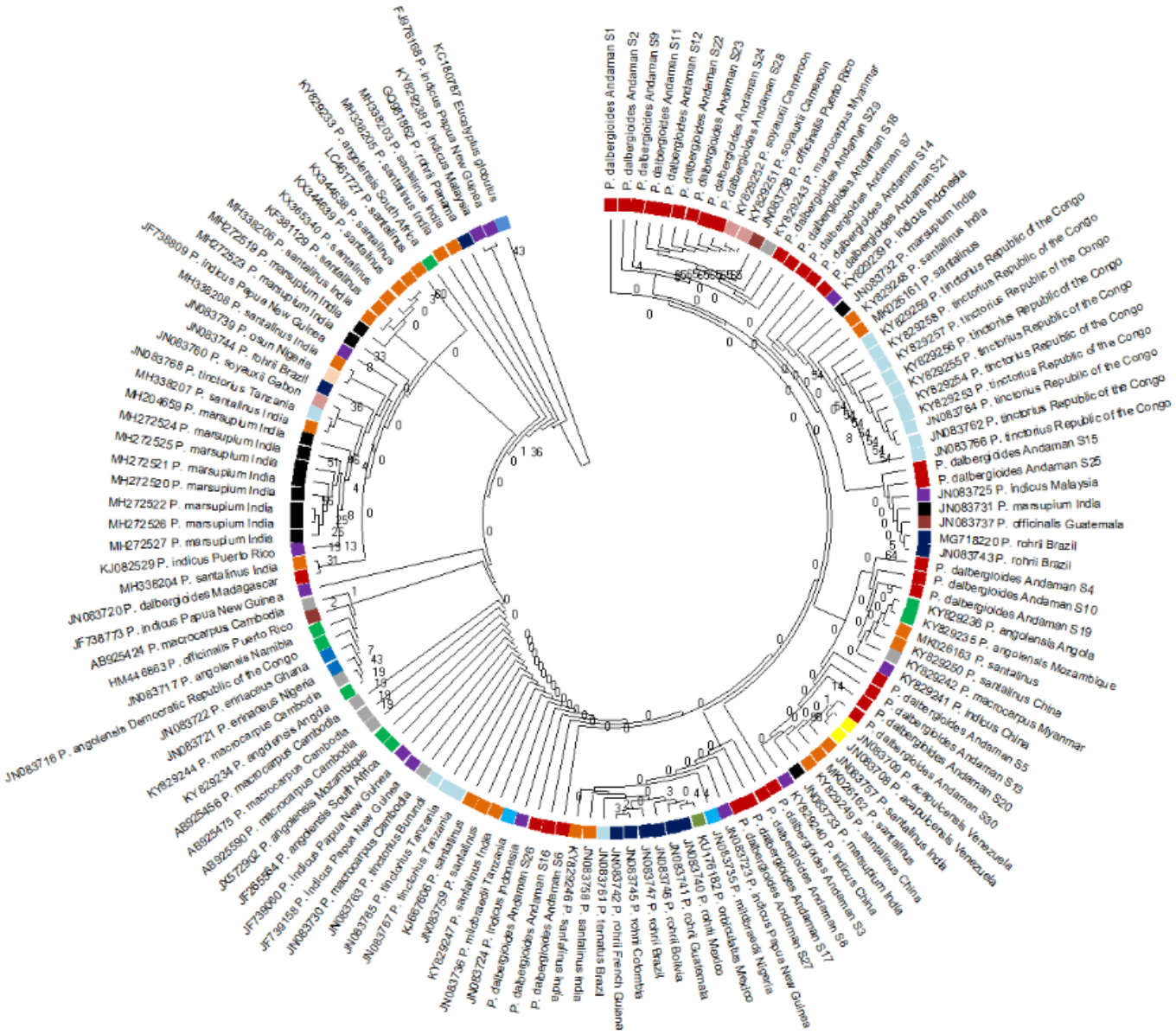

## Figure 4

ITS based phylogenetic relationship of Andaman padauk with different *Pterocarpus* species all over the world. (a) Neighbor Joining (NJ) phylogenetic tree, (b) Bayesian phylogenetic tree and (c) Network tree




## Figure 5

Matk based phylogenetic relationship of Andaman padauk with different *Pterocarpus* species all over the world. (a) Neighbor Joining (NJ) phylogenetic tree, (b) Bayesian phylogenetic tree and (c) Network tree




## Figure 6

*rbcL* based phylogenetic relationship of Andaman padauk with different *Pterocarpus* species all over the world.