

RESEARCH ARTICLE

# Two New Potential Barcodes to Discriminate *Dalbergia* Species

Rasika M. Bhagwat<sup>1</sup>, Bhushan B. Dholakia<sup>1</sup>, Narendra Y. Kadoo<sup>1</sup>, M. Balasundaran<sup>2</sup>, Vidya S. Gupta<sup>1\*</sup>

**1** Plant Molecular Biology Group, Biochemical Sciences Division, CSIR-National Chemical Laboratory, Pune, Maharashtra, India, **2** Forest Genetics and Biotechnology Division, Kerala Forest Research Institute, Peechi, Thrissur, Kerala, India

\* [vs.gupta@ncl.res.in](mailto:vs.gupta@ncl.res.in)



**OPEN ACCESS**

**Citation:** Bhagwat RM, Dholakia BB, Kadoo NY, Balasundaran M, Gupta VS (2015) Two New Potential Barcodes to Discriminate *Dalbergia* Species. PLoS ONE 10(11): e0142965. doi:10.1371/journal.pone.0142965

**Editor:** Dirk Steinke, Biodiversity Institute of Ontario - University of Guelph, CANADA

**Received:** October 10, 2014

**Accepted:** October 29, 2015

**Published:** November 16, 2015

**Copyright:** © 2015 Bhagwat et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Nucleotide sequences of analyzed loci from all individuals were deposited in NCBI database ([S1 Dataset](#)); accession numbers - *matK*: KM276475-KM276412; *rbcl*: KM100059-KM099987; *trnH-psbA*: KM276322-KM276250 and *nriTS*: KM276165-KM276104).

**Funding:** The authors thank the Department of Biotechnology, New Delhi (project code GAP267426) and the CSIR (project code BSC0106) for the financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

DNA barcoding enables precise identification of species from analysis of unique DNA sequence of a target gene. The present study was undertaken to develop barcodes for different species of the genus *Dalbergia*, an economically important timber plant and is widely distributed in the tropics. Ten *Dalbergia* species selected from the Western Ghats of India were evaluated using three regions in the plastid genome (*matK*, *rbcl*, *trnH-psbA*), a nuclear transcribed spacer (*nriTS*) and their combinations, in order to discriminate them at species level. Five criteria: (i) inter and intraspecific distances, (ii) Neighbor Joining (NJ) trees, (iii) Best Match (BM) and Best Close Match (BCM), (iv) character based rank test and (v) Wilcoxon signed rank test were used for species discrimination. Among the evaluated loci, *rbcl* had the highest success rate for amplification and sequencing (97.6%), followed by *matK* (97.0%), *trnH-psbA* (94.7%) and *nriTS* (80.5%). The inter and intraspecific distances, along with Wilcoxon signed rank test, indicated a higher divergence for *nriTS*. The BM and BCM approaches revealed the highest rate of correct species identification (100%) with *matK*, *matK+rbcl* and *matK+trnH-psb* loci. These three loci, along with *nriTS*, were further supported by character based identification method. Considering the overall performance of these loci and their ranking with different approaches, we suggest *matK* and *matK+rbcl* as the most suitable barcodes to unambiguously differentiate *Dalbergia* species. These findings will potentially be helpful in delineating the various species of *Dalbergia* genus, as well as other related genera.

## Introduction

In DNA barcoding, the sequence of a short stretch of DNA is used for accurate species identification [1], supplementing the classical taxonomic methods [2]. Although DNA barcoding has been successfully used for discriminating animal species, applying this approach for discriminating plant species is more difficult due to many challenges [3]. Plant mitochondrial genomes exhibit low rates of nucleotide substitution and high rates of chromosomal rearrangements [4], while extensive gene duplication occurs in the nuclear genome [5]. Initial DNA barcoding

**Competing Interests:** The authors have declared that no competing interests exist.

studies in plants have proposed a few plastid coding as well as non-coding regions, such as *rbcl* and *trnH-psbA* [6], *matK*, *rpoB*, *rpoC1* and *trnH-psbA* [7] and *atpF/H*, *matK*, *psbK/I* and *trnH-psbA* [8] as promising candidates. However, the slow evolving coding regions of plastid genomes might not possess enough variation to discriminate closely related plant species and this could lower their potential as effective barcodes [9]. This can be overcome by analyzing the selected loci either individually or in combination [10, 11]. Recently evolved nuclear region, i.e. nuclear internal transcribed spacer from ribosomal gene (*nrITS*) has also been proposed as potential barcodes [12].

*Dalbergia* Linn. F. (Family: Fabaceae) is a genus of shrubs, lianas and trees. It is confined to the tropical regions of the world with Amazonia, Madagascar, Africa and Indonesia as the centers of diversity [13, 14]. About 200 species comprise the genus, of which nearly 35 are found in India with 10–15 species in the Western Ghats (WG) alone [14, 15]. The overall species diversity is high in WG Seven species are endemic to this region ([http://wgbis.ces.iisc.ernet.in/biodiversity/sahyadri\\_eneews/newsletter/issue38/article/index.htm](http://wgbis.ces.iisc.ernet.in/biodiversity/sahyadri_eneews/newsletter/issue38/article/index.htm)); hence, we choose to select WG as our study area. The *Dalbergia* genus is economically important for its quality timber. The wood of different *Dalbergia* species is used for specific purposes such as making furniture (*D. latifolia*, *D. sissoo*), boat building (*D. sissoo*) and manufacturing musical instruments (*D. melanoxyton*) [15]. Studies on tropical dry evergreen forests (TDEF) of India have indicated indiscriminate logging as one of the major factors responsible for the loss of commercial tree species, biodiversity. This is particularly the case for the species listed in Appendix II of the CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) document [16]. The Red list of IUCN (International Union for Conservation of Nature) has more than 30 *Dalbergia* species under endangered category (<http://www.iucnredlist.org>) including *D. cochinchinensis* and *D. latifolia* as vulnerable species. Similarly, APFORGEN (Asia Pacific Forest Genetic Resource Programme) has identified *D. latifolia* as a prime concern from a conservation point of view. Moreover, as the wood of *Dalbergia* species is illegally traded in some countries, it is difficult to prove their identity and take legal action in the absence of accurate tools and methods for species identification [16]. This has facilitated fraudulent marketing and sale of poor quality wood of other tree species in place of *Dalbergia*. In this context, DNA barcoding can help as a quick way of authenticating the wood of *Dalbergia* even for legal purpose if needed.

*Dalbergia* species are morphologically variable and possess a wide range of habitat preference. This makes it difficult to classify the New World and the Old World species into natural groups [17, 18]. Over the past several decades, many revisions based on morphological characters have made the taxonomic speciation in *Dalbergia* quite challenging [12, 17, 19–23]. Moreover, very limited information is available on the molecular taxonomy of *Dalbergia* genus. There is only one report [14] describing the phylogeny of *Dalbergia* species indicating its monophyletic nature of origin. The genus was included in the evolutionary study of Leguminosae [24] to analyze the relationship of *Machaerium* and *Aeschynomene* using *trnL* and nuclear ribosomal DNA sequences [25]. Very few studies have reported on the molecular analysis of Indian *Dalbergia* species [15, 26–29], making it imperative to conduct studies on the genus on various aspects including phylogeny, diversity and end-use quality using DNA markers and sequence based polymorphism in suitable genomic regions.

In the present study, the primary focus was to develop an accurate species identification method for *Dalbergia* genus and this was addressed by developing potential DNA barcodes for the genus. We have evaluated 37 primer pairs from plastid and nuclear genomes of which four loci (*rbcl*, *matK*, *trnH-psbA* and *nrITS*) were shortlisted and various statistical parameters were employed to demonstrate their potential as barcodes to unambiguously discriminate *Dalbergia* species.

## Materials and Methods

### Ethics statement

The locations involved in the study were not part of any protected area, reserve forests or national parks except for Chinar wildlife sanctuary and Parambikulam wildlife sanctuary. The samples from these areas were collected by Kerala Forest Research Institute (KFRI), Peechi, Kerala, which is a government organization having the requisite permissions. The exact GPS coordinates for the collection sites are not available. Further, none of these species are endangered or protected species.

### Sample collection

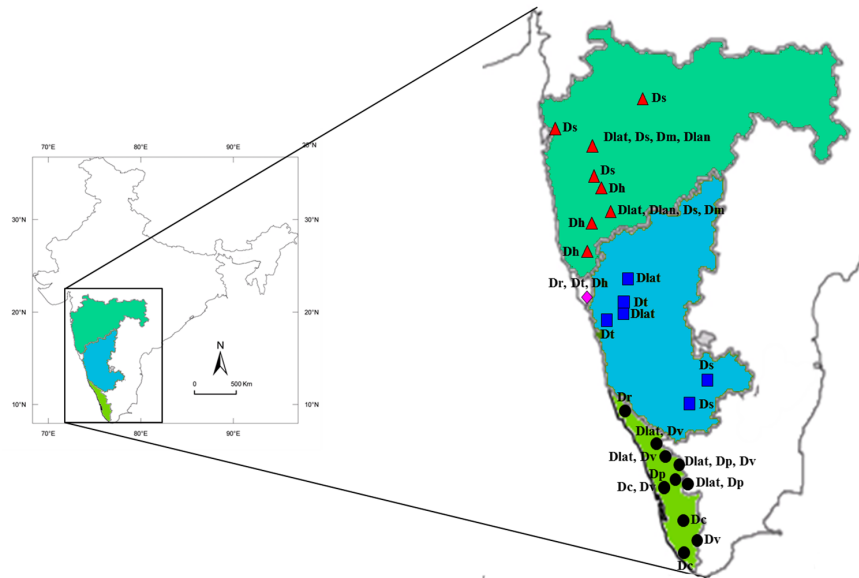
The study included 166 accessions from ten *Dalbergia* species representing three sections, section Sissoa (*Dalbergia latifolia*, *D. melanoxylon*, *D. sissoo*, *D. rubiginosa*, *D. horrida* and *D. tamarindifolia*), section Dalbergia (*D. volubilis*, *D. paniculata* and *D. lanceolaria*) [15] and section Selenobia (*D. candenatensis*) [20]. We focused on the locations in WG, which is one of the most important biodiversity hotspots in India (Fig 1 and S1 Dataset). Between 5 and 25 accessions of each species were collected from different locations to understand the effect of geographical isolation on intraspecific variation in barcoding. The samples were authenticated by KFRI and the Botanical Survey of India (BSI, Western Circle, Pune, India) and the voucher specimens from each species were deposited in their respective herbaria. *Pterocarpus marsupium*, which falls outside the *Dalbergia* clade and is native to WG, was used as an out-group in the present study [14].

### DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from fresh or dried leaf samples using the modified cetyltrimethylammonium bromide (CTAB) method [30]. At the time of initiating this study, since no specific region was recommended as universal plant barcode, based on available literature we selected the genomic loci corresponding to *matK* (7 primer pairs), *rpoC* (4 primer pairs), *rpoB* (5 primer pairs), *accD* (6 primer pairs), *ndhJ* (3 primer pairs), *ycf5* (4 primer pairs), *trnH-psbA* (5 primer pairs), *nrITS* (2 primer pairs) and *rbcl* (single primer pair) for developing the barcodes. As sequence information for most of these loci was not available for *Dalbergia* species, we attempted multiple sets of primers to amplify the respective loci from all the ten species. Thirty seven primer pairs were tested to identify the loci satisfying the set criteria for DNA barcoding. Four primer pairs (S2 Dataset) corresponding to *matK*, *rbcl*, *trnH-psbA* and *nrITS* produced highly specific amplifications (sharp bands on agarose gel) and gave good quality DNA sequences. Therefore, these were selected for further study. PCR amplifications were performed in a final volume of 20 or 25  $\mu$ L (S3 Dataset) and the amplicons were resolved on 1% agarose gel. Most of the PCR reactions yielded specific amplifications (i.e. sharp single bands on agarose gel) and these were directly used as templates for sequencing reactions. In the samples that generated multiple PCR products, bands corresponding to the expected size were eluted from the gel using PureLink<sup>®</sup> Quick Gel Extraction Kit (Invitrogen, USA) and used as templates in sequencing reactions. Sequencing was performed using Sanger chemistry in both ends of the DNA fragment using MegaBACE DYEnamic ET dye terminator kit with MegaBACE1000 DNA Analysis System (GE Healthcare, USA).

### Sequence analysis

For each sequence, the chromatograms were inspected and poor quality 5' and 3' DNA sequence ends were trimmed. Post trimming lengths were maintained at least 60% of the



**Fig 1. Map of India showing the locations of collection sites.** The map highlights three states of India across which the Western Ghats are spread. The expended view of the inset shows the location and geographical distribution of the actual sites.

doi:10.1371/journal.pone.0142965.g001

original read length, subject to the minimum average quality score of Q20. The sequences failing this criterion were rejected and re-sequenced. All the nucleotide variations were evaluated and confirmed by aligning the chromatograms from forward and reverse sequencing results. Sequences with 70% or more overlap were considered for creating consensus sequence for each amplicon [31]. Good quality sequences from all individuals were assembled and aligned using CLUSTALW 1.83 [32]. Conserved, variable and parsimony informative sites were determined using MEGA 5.0 [33]. Distance matrices and Neighbor-Joining (NJ) trees were established in MEGA using the best fit nucleotide substitution model (chosen with AICc) [34].

### Data analysis

Genetic distance was calculated using Kimura-2-Parameter (K2P) model [35]. The interspecific divergence between the species was studied using the following three parameters: (i) average inter specific distance; (ii) average theta prime ( $\theta'$ ), where  $\theta'$  is the mean pairwise distance within species, thus eliminating the biases associated with different individual count among species; and (iii) minimum inter specific distance. Three additional parameters were studied for the intraspecific divergence: (i) average intraspecific divergence, (ii) theta ( $\theta$ ) and (iii) average coalescent depth [36].

Wilcoxon signed rank tests were performed to check existence of significant divergence between the inter and intraspecific variability between the pairs of barcoding loci [11]. Consensus sequences were generated for all the ten *Dalbergia* species using TaxonDNA [37] with 1000 bootstraps. To analyze inter and intraspecific variation, sequence variants were generated with DnaSP 5.0 [38] using consensus sequences. Further, NJ trees were constructed in MEGA 5.0 with 1000 bootstraps. Based on the distance method using K2P parameter and a minimum sequence overlap of 300 bp, accurate species identification was performed by TaxonDNA or SpeciesIdentifier 1.7.7 [37] using two approaches: (i) Best match (BM) and (ii) Best close match (BCM). In these approaches, each sequence from the dataset was used as a query against the remaining sequences from the same dataset. With BM, a query sequence was identified by

searching the reference sequence for the best match with the smallest genetic distance to the query. The BCM approach required a threshold value, which was calculated for each locus from pairwise summary. The threshold was a value below which 95% of all intraspecific distances were observed, leading to an upper bound value on the similarity of a barcode match [37]. If both, the query and the subject sequences were from the same species, the identification was considered as successful. Whereas, if more than one query sequence from different species exhibited equally good match, then the samples were considered as ambiguous. Another character based analysis method, Barcoding with LOGic Formulae (BLOG), was also employed [39]. This method selected the unique nucleotide position of the sequence and derived a formula to differentiate among species. It also provided concise and meaningful classification rules [40].

## Results

### Amplification success

The success rate for PCR amplification and sequencing of bidirectional reads was the highest for *rbcL* (97.6%), followed by *matK* (97.0%) and *trnH-psbA* (94.7%), while *nrITS* exhibited the lowest rate (80.5%). Nucleotide sequences of analyzed loci from all individuals were deposited in NCBI database ([S1 Dataset](#); accession numbers—*matK*: KM276475-KM276412; *rbcL*: KM100059-KM099987; *trnH-psbA*: KM276322-KM276250 and *nrITS*: KM276165-KM276104). Using BLAST analysis, all the loci correctly identified 100% of the samples at genus level; while at species level, *nrITS* had the highest identification rate i.e. 60% followed by *rbcL* (50%), *matK* (20%) and *trnH-psbA* (10%). The low rate of species level identification might be due to the absence of species records in NCBI database and high percentage of in-dels especially in the case of *trnH-psbA* sequences.

### Nucleotide variation

The percentages of polymorphic informative (Pi) sites and variable sites were comparable for the respective loci. For *nrITS*, aligned length was 637 bp, with 29.83% sites variable and 28.89% polymorphic informative, which was the highest among all the loci (single locus as well as combination of loci). Based on the percentage of conserved sites, the most conserved loci were *rbcL* followed by *matK* and *matK+rbcL* ([Table 1](#)).

### Inter and intraspecific divergence

**Distance analysis and Wilcoxon signed rank test.** The *nrITS* locus showed greater interspecific divergence than the plastid loci (*matK*, *rbcL* and *trnH-psbA* and their combinations) using both average inter specific distance and  $\theta'$  parameters. However, in case of intraspecific divergence, *nrITS* and *rbcL* showed the highest and the lowest value, respectively. Thus, no single locus revealed the highest interspecific but the lowest intraspecific divergence ([Table 2](#) and [Fig 2](#)). When the Wilcoxon signed rank test was used to compare the loci, *nrITS* exhibited the highest interspecific divergence followed by *trnH-psbA*, whereas *rbcL* displayed the lowest intraspecific divergence ([Tables 3](#) and [4](#)).

**Barcode gap.** Barcode gap represents the absence of overlapping regions between inter and intraspecific distances. The barcode gap was absent for all the marker loci used in the present study, indicating overlaps between inter and intraspecific distances ([Fig 3](#)). However, the mean interspecific divergence was significantly higher than that of the corresponding intraspecific divergence for each of the loci. This was further confirmed by analysis carried out using TaxonDNA.

**Table 1. Summary statistics for potential barcode loci from ten *Dalbergia* species.**

Locus	<i>matK</i>	<i>rbcL</i>	<i>trnH-psbA</i>	<i>nrITS</i>	<i>matK+ trnH-psbA</i>	<i>matK+ rbcL</i>	<i>rbcL+ trnH-psbA</i>
No. of sequences analyzed	165	166	161	137	157	163	157
Total no. of sites	677	491	273	637	950	1168	764
Conserved sites	636 (93.94)	477 (97.15)	250 (91.58)	447 (70.17)	863 (90.84)	1113 (95.29)	724 (94.76)
Variable sites	41 (6.06)	14 (2.85)	23 (8.42)	190 (29.83)	87 (9.16)	55 (4.71)	40 (5.24)
Parsimony informative sites	40 (5.91)	14 (2.85)	15 (5.49)	184 (28.89)	55 (5.79)	54 (4.62)	29 (3.80)

**Note:** Values in parentheses are expressed in percentage.

doi:10.1371/journal.pone.0142965.t001

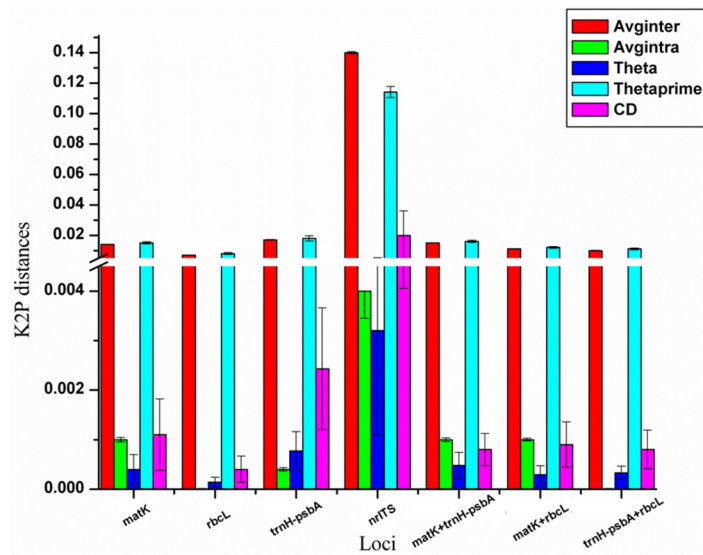
**Tree based analyses.** The sequence variants of each marker locus were determined using DnaSP 5.0 and MEGA 5.0 as mentioned previously. Among all loci, *nrITS* exhibited the maximum number of sequence variants (Table 5). By including all the sequence variants, seven NJ trees were constructed with *matK*, *rbcL*, *trnH-psbA* and *nrITS* either alone (Fig 4) or in combinations (Fig 5). All of them except *rbcL* revealed a separate cluster for each species and *rbcL* could not differentiate between *D. rubiginosa*, *D. candenatensis* and *D. tamarindifolia*. Interestingly, except *trnH-psbA* all other loci (*matK*, *rbcL*, *nrITS* and *matK+rbcL*) either alone or in combination were capable of grouping together all three species-clusters from the section *Dalbergia* (*D. volubilis*, *D. lanceolaria* and *D. paniculata*). This agrees with a previous report on genome size variation and evolution of *Dalbergia* species which found that *D. lanceolaria* and *D. paniculata* were closely related [15]. These observations indicated that *matK*, *nrITS*, *rbcL* and *matK+rbcL* could correctly identify the reported relationships among the *Dalbergia* species and hence, they could most likely be successful as barcodes for this genus.

**Similarity based approach.** To evaluate the accuracy of these potential barcodes in species assignments, the BM and BCM parameters from TaxonDNA analysis were used (Table 6). Finding a standard threshold for BCM approach is difficult as there is a large variation in inter and intraspecific divergence across all loci in different plant systems [9]. Moreover, our approach to use multiple accessions of each species, as suggested by Pettengill and Neel [9] has ensured that the basic requirement was fulfilled and therefore, we chose to use calculated thresholds. The calculated threshold value per locus varied from 0.12% in *rbcL+trnH-psbA* to 1.2% in *nrITS*. With the BM and BCM approaches, the success rate of correct identification was unambiguously 100% for *matK*, *matK+trnH-psbA* and *matK+rbcL* and 0% incorrect identification (Table 6).

**Table 2. Inter and intraspecific divergence values for potential barcode loci.**

Distance parameters	<i>matK</i>	<i>rbcL</i>	<i>trnH-psbA</i>	<i>nrITS</i>	<i>matK+trnH-psbA</i>	<i>matK+rbcL</i>	<i>trnH-psbA+rbcL</i>
Average interspecific distance	0.014±5.74E-05	0.007±3.61E-05	0.017±1.11E-4	0.140±6.45E-4	0.015±6.45E-05	0.011±4.55E-05	0.010±5.03E-05
Theta prime (θ')	0.015±7.43E-4	0.008±5.6E-4	0.018±1.677E-3	0.114±3.656E-3	0.016±8.2E-4	0.012±5.98E-4	0.011±7.07E-4
Smallest interspecific distance	0.014±7.12E-4	0.008±5.43E-3	0.017±1.627E-3	0.156±7.922E-3	0.015±8.17E-4	0.011±5.66E-4	0.011±6.83E-4
Average intraspecific distance	0.001±4.50E-05	0.000±1.52E-05	0.000±3.52E-05	0.004±5.48E-4	0.001±3.77E-05	0.001±2.78E-05	0.000±1.50E-05
Theta (θ)	0.000±2.96E-4	0.000±1.03E-4	0.001±3.85E-4	0.003±2.112E-3	0.000±2.61E-4	0.000±1.77E-4	0.000±1.29E-4

doi:10.1371/journal.pone.0142965.t002



**Fig 2. Distribution of inter and intraspecific divergence.** The plot depicts inter and intraspecific divergence parameters for various loci. Avginter: Average inter specific distance, Avgintra: Average intraspecific distance, Theta, Theta prime, CD: coalescence depth.

doi:10.1371/journal.pone.0142965.g002

**Character based approach.** The data analysis resulted into logic formulae as well as revealed information regarding correctly classified, wrongly classified and not classified species. Only the analysis done using *matK*, *nrITS*, *matK+rbcL* and *matK+trnH-psbA* loci could assign the characteristic nucleotide positions for all the species with 100% correct classification (Table 7).

### Overall performance of the loci

The different parameters used for screening potential barcode loci were ranked based on their performance on a scale of 1–10. In case of NJ trees, the ranking was done based on clustering of the species. Those loci which separated all the species irrespective of intraspecific variation were given ten marks, while for the remaining loci, the scale was determined based on the number of species clubbed together. For inter- and intraspecific distances, the difference between the maximum and minimum distance was calculated to determine the scale for each locus. For BM and BCM methods, the percent values corresponding to correct, ambiguous and incorrect classification were used to rank the loci. A similar methodology was also applied for BLOG. Finally, for Wilcoxon signed rank test, the locus which performed the best in a pair in both, inter and intraspecific distance determinations, was ranked the highest (Table 8).

### Discussion

Paul Hebert’s research in 2003 on species identification using short stretches of DNA from a well characterized region of the genome, gave birth to the concept of DNA barcoding [41]. Initial efforts proved the reliability of mitochondrial cytochrome c oxidase 1 (*cox1*) gene as an impressive barcode in animals [42]. However, initial research on plant DNA barcoding suggested that species discrimination in plants with a single universal locus is difficult. This is primarily due to various phenomena such as polyploidy, hybridization, heteroplasmy etc., which result in the formation of continuous range of variable characters and making delineation a difficult task. Alternatively, sufficient time is often required to accumulate mutations in organisms

**Table 3. Wilcoxon signed-rank tests results for interspecific divergence of the indicated loci.**

W+	W-	Inter relative ranks	Results
<i>matK</i>	<i>rbcl</i>	W+ = 85063, W- = 15, n = 412, p = 0	<i>matK</i> > <i>rbcl</i>
<i>matK</i>	<i>matK + rbcl</i>	W+ = 73114, W- = 39, n = 412, p = 0	<i>matK</i> > <i>matK+rbcl</i>
<i>matK</i>	<i>rbcl+trnH-psbA</i>	W+ = 68612, W- = 6078, n = 412, p = 0	<i>matK</i> > <i>rbcl+trnH-psbA</i>
<i>matK</i>	<i>trnH-psbA</i>	W+ = 29484, W- = 46370, n = 412, p = 0	<i>trnH-psbA</i> > <i>matK</i>
<i>matK</i>	<i>nrITS</i>	W+ = 1, W- = 85077, n = 412, p = 0	<i>nrITS</i> > <i>matK</i>
<i>matK</i>	<i>matK+trnH-psbA</i>	W+ = 27245, W- = 41019, n = 412, p = 0.001	<i>matK+trnH-psbA</i> > <i>matK</i>
<i>rbcl</i>	<i>nrITS</i>	W+ = 0, W- = 85078, n = 412, p = 0	<i>nrITS</i> > <i>rbcl</i>
<i>rbcl</i>	<i>trnH-psbA</i>	W+ = 5996, W- = 76625, n = 412, p = 0	<i>trnH-psbA</i> > <i>rbcl</i>
<i>rbcl</i>	<i>matK + rbcl</i>	W+ = 0, W- = 84255, n = 412, p = 0	<i>matK+rbcl</i> > <i>rbcl</i>
<i>rbcl</i>	<i>matK+trnH-psbA</i>	W+ = 0, W- = 84666, n = 412, p = 0	<i>matK+trnH-psbA</i> > <i>rbcl</i>
<i>rbcl</i>	<i>rbcl+trnH-psbA</i>	W+ = 1924, W- = 60204, n = 412, p = 0	<i>rbcl+trnH-psbA</i> > <i>rbcl</i>
<i>trnH-psbA</i>	<i>matK+rbcl</i>	W+ = 63125, W- = 17476.5, n = 412, p = 0	<i>trnH-psbA</i> > <i>matK+rbcl</i>
<i>trnH-psbA</i>	<i>matK+trnH-psbA</i>	W+ = 47083.5, W- = 28771.50, n = 412, p = 0	<i>trnH-psbA</i> > <i>matK+trnH-psbA</i>
<i>trnH-psbA</i>	<i>rbcl+trnH-psbA</i>	W+ = 73380.5, W- = 6020.5, n = 412, p = 0	<i>trnH-psbA</i> > <i>rbcl+trnH-psbA</i>
<i>trnH-psbA</i>	<i>nrITS</i>	W+ = 0, W- = 85078, n = 412, p = 0	<i>nrITS</i> > <i>trnH-psbA</i>
<i>nrITS</i>	<i>matK+rbcl</i>	W+ = 85077, W- = 1, n = 412, p = 0	<i>nrITS</i> > <i>matK+rbcl</i>
<i>nrITS</i>	<i>matK+trnH-psbA</i>	W+ = 85077, W- = 1, n = 412, p = 0	<i>nrITS</i> > <i>matK+trnH-psbA</i>
<i>nrITS</i>	<i>rbcl+trnH-psbA</i>	W+ = 85078, W- = 0, n = 412, p = 0	<i>nrITS</i> > <i>rbcl+trnH-psbA</i>
<i>matK + rbcl</i>	<i>matK+trnH-psbA</i>	W+ = 37, W- = 73116, n = 412, p = 0	<i>matK+trnH-psbA</i> > <i>matK+rbcl</i>
<i>matK + rbcl</i>	<i>rbcl+trnH-psbA</i>	W+ = 49627, W- = 22004, n = 412, p = 0	<i>matK+rbcl</i> > <i>rbcl+trnH-psbA</i>
<i>matK+trnH-psbA</i>	<i>rbcl+trnH-psbA</i>	W+ = 79759, W- = 41, n = 412, p = 0	<i>matK+trnH-psbA</i> > <i>rbcl+trnH-psbA</i>

doi:10.1371/journal.pone.0142965.t003

which are responsible for separation of closely related species. However, the lack of such sufficient genetic variation hampers species level discrimination of plants by DNA barcoding [8]. This problem is exaggerated in woody plants because of longer generation time and lower mutation rate. It is also difficult to differentiate species in taxonomically complex groups where species are narrowly defined. Additionally, large ancestral population sizes and low levels of within species gene flow for plastid markers create difficulty in barcode based identification [3, 8]. In order to resolve these problems, several attempts have been made to establish DNA barcodes using multiple genes from different plant genomes for specific families such as Myristicaceae [43], Lemnaceae [44], Zingiberaceae [45], Podocarpaceae [46] or genera such as *Paeonia* [47], *Acacia* [48], *Paphiopedilum* [49], *Parnassia* [50] and *Gossypium* [51]. However, from different studies, it appears that finding a universal barcode or even a barcode at family level is difficult and it may be possible to establish a discriminating barcode only at genus level [52].

There are few reports on DNA barcoding of tropical tree species [16, 31, 53] which include Amazonian as well as Indian forest trees. These studies have used *nrITS*, *matK*, *rbcl* and *trnH*-

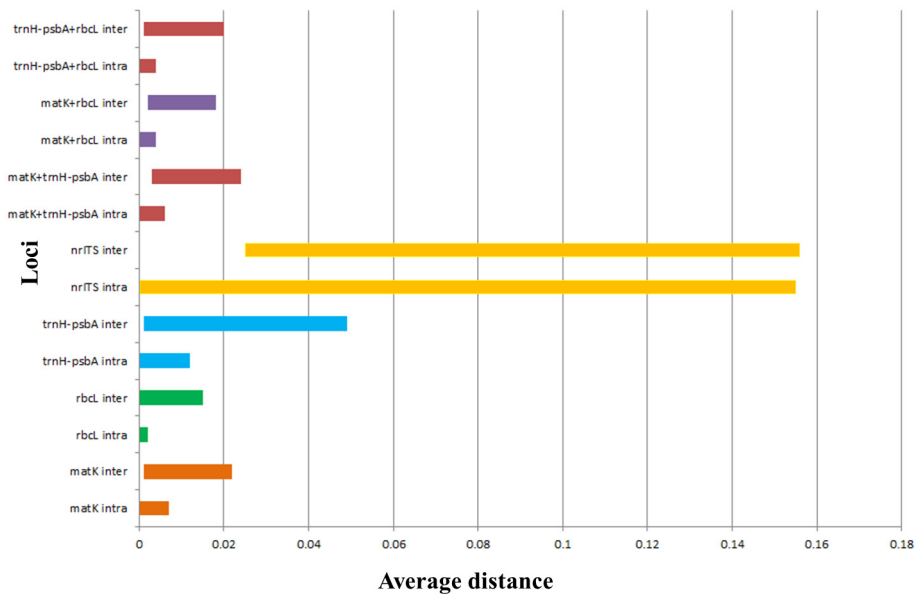


**Table 4. Wilcoxon signed-rank test results for intraspecific divergence of the indicated loci.**

W+	W-	Inter relative ranks	Results
<i>matK</i>	<i>rbcL</i>	W+ = 141, W- = 12, n = 53, p = 0.002	<i>matK</i> > <i>rbcL</i>
<i>matK</i>	<i>rbcL+trnH-psbA</i>	W+ = 181.50, W- = 49.50, n = 53, p = 0.020	<i>matK</i> > <i>rbcL+trnH-psbA</i>
<i>matK</i>	<i>matK + rbcL</i>	W+ = 18.50, W- = 2.50, n = 53, p = 0.084	<i>matK</i> = <i>matK+rbcL</i>
<i>matK</i>	<i>matK+trnH-psbA</i>	W+ = 66, W- = 39, n = 53, p = 0.369	<i>matK</i> = <i>matK+trnH-psbA</i>
<i>matK</i>	<i>trnH-psbA</i>	W+ = 135.5, W- = 74.5, n = 53, p = 0.250	<i>matK</i> = <i>trnH-psbA</i>
<i>matK</i>	<i>nrlITS</i>	W+ = 102.50, W- = 932.50, n = 53, p = 0	<i>nrlITS</i> > <i>matK</i>
<i>rbcL</i>	<i>nrlITS</i>	W+ = 10, W- = 1071, n = 53, p = 0	<i>nrlITS</i> > <i>rbcL</i>
<i>rbcL</i>	<i>trnH-psbA</i>	W+ = 28, W- = 63, n = 53, p = 0.212	<i>rbcL</i> = <i>trnH-psbA</i>
<i>rbcL</i>	<i>rbcL+trnH-psbA</i>	W+ = 42, W- = 49, n = 53, p = 0.793	<i>rbcL+trnH-psbA</i> = <i>rbcL</i>
<i>rbcL</i>	<i>matK + rbcL</i>	W+ = 6.50, W- = 146.50, n = 53, p = 0.001	<i>matK+rbcL</i> > <i>rbcL</i>
<i>rbcL</i>	<i>matK+trnH-psbA</i>	W+ = 8.5, W- = 111.5, n = 53, p = 0.003	<i>matK+trnH-psbA</i> > <i>rbcL</i>
<i>trnH-psbA</i>	<i>nrlITS</i>	W+ = 151.5, W- = 1024.5, n = 53, p = 0	<i>nrlITS</i> > <i>trnH-psbA</i>
<i>trnH-psbA</i>	<i>matK+rbcL</i>	W+ = 80.5, W- = 129.5, n = 18, p = 0.356	<i>trnH-psbA</i> = <i>matK+rbcL</i>
<i>trnH-psbA</i>	<i>matK+trnH-psbA</i>	W+ = 74.5, W- = 135.5, n = 53, p = 0.250	<i>trnH-psbA</i> = <i>matK+trnH-psbA</i>
<i>trnH-psbA</i>	<i>rbcL+trnH-psbA</i>	W+ = 63, W- = 28, n = 53, p = 0.212	<i>trnH-psbA</i> = <i>rbcL+trnH-psbA</i>
<i>nrlITS</i>	<i>matK+rbcL</i>	W+ = 1034, W- = 47, n = 53, p = 0	<i>nrlITS</i> > <i>matK+rbcL</i>
<i>nrlITS</i>	<i>matK+trnH-psbA</i>	W+ = 1122.5, W- = 102.5, n = 53, p = 0	<i>nrlITS</i> > <i>matK+trnH-psbA</i>
<i>nrlITS</i>	<i>rbcL+trnH-psbA</i>	W+ = 1209, W- = 16, n = 53, p = 0	<i>nrlITS</i> > <i>rbcL+trnH-psbA</i>
<i>matK+rbcL</i>	<i>matK+trnH-psbA</i>	W+ = 42, W- = 63, n = 53, p = 0.485	<i>matK+trnH-psbA</i> = <i>matK+rbcL</i>
<i>matK+rbcL</i>	<i>rbcL+trnH-psbA</i>	W+ = 163, W- = 47, n = 53, p = 0.028	<i>matK+rbcL</i> > <i>rbcL+trnH-psbA</i>
<i>matK+trnH-psbA</i>	<i>rbcL+trnH-psbA</i>	W+ = 167.50, W- = 22.50, n = 53, p = 0.002	<i>matK+trnH-psbA</i> > <i>rbcL+trnH-psbA</i>

doi:10.1371/journal.pone.0142965.t004

*psbA* loci. However, there are scanty reports on DNA barcoding of trees exclusively from WG of India. A study on 143 tree species from tropical dry evergreen forests in India covering 114 genera and 42 families revealed that combination of *matK* and *rbcL* loci gave the highest success in accurate identification [16]. Similarly, DNA barcoding of medicinal plants from the family Fabaceae revealed 80% and 96% success at species and genus level, respectively using *matK* locus, while the *ITS2* locus gave more than 80% success at species level and 100% success at genus level [54]. However, none of the above mentioned studies included *Dalbergia*. A recent study on tropical tree species from India (149 species from 82 genera and 38 families) included three *Dalbergia* species and suggested that *ITS* and *trnH-psbA* might not be highly successful [31]. Efforts to resolve the sister species complex of *Acacia* from Fabaceae using *rbcL*, *trnH-psbA* (same primer sequence as we have used in our study) and *matK* recommended all the



**Fig 3. The barcoding gap.** Graph of smallest interspecific and largest intraspecific distances highlighting the overlapping divergence.

doi:10.1371/journal.pone.0142965.g003

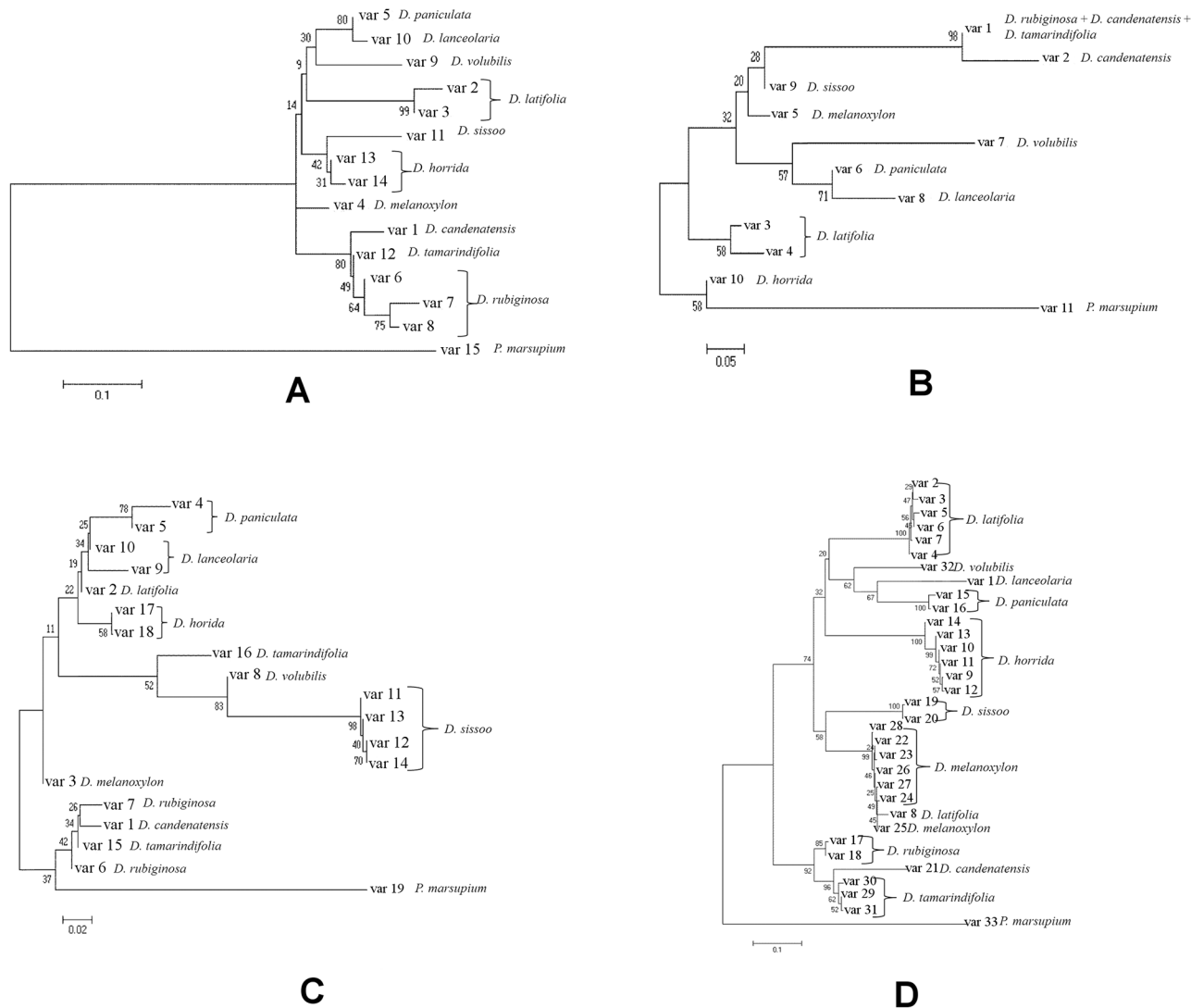
three regions for barcoding [48]. On the contrary, studies on *Aspalathus* using *ITS* (different primers than the ones used in our study), *psbA-trnH* and *trnT-trnL* concluded that all the three loci were unable to resolve the species [55]. It was observed that the output from *matK* analysis was variable based on the plant systems as well as on the combination of primers used for analysis. However, the Consortium for the Barcode of Life (CBOL) proposed 90% success with *matK* for plants. Our study also identified *matK* as one of the potential loci for DNA barcoding. Thus, *matK*, *nrITS* and *rbcl* individually or in their combinations could be explored as the potential DNA barcodes in various plant genera [53].

**Table 5. Distribution of sequence variants among the ten *Dalbergia* species across all loci.**

Species	Number of sequence variants						
	<i>matK</i>	<i>rbcl</i>	<i>trnH-psbA</i>	<i>nrITS</i>	<i>matK+rbcl</i>	<i>matK+trnH-psbA</i>	<i>rbcl+trnH-psbA</i>
Dc	1	2	1	1	2	1	2
Dlat	2	2	1	7	2	2	2
Dm	1	1	1	7	1	1	1
Dp	1	1	2	2	1	1	1
Dr	3	1	2	2	3	4	2
Dv	1	1	1	1	1	1	1
Dlan	1	1	2	1	1	2	2
Ds	1	1	4	2	1	4	4
Dt	1	1	2	3	1	2	2
Dh	2	1	2	6	2	2	2

**Species codes:**—Dc: *D. candenatensis*, Dlat: *D. latifolia*, Dm: *D. melanoxylon*, Dp: *D. paniculata*, Dr: *D. rubiginosa*, Dv: *D. volubilis*, Dlan, *D. lanceolaria*, Ds: *D. sissoo*, Dt: *D. tamarindifolia*, Dh: *D. horrida*.

doi:10.1371/journal.pone.0142965.t005

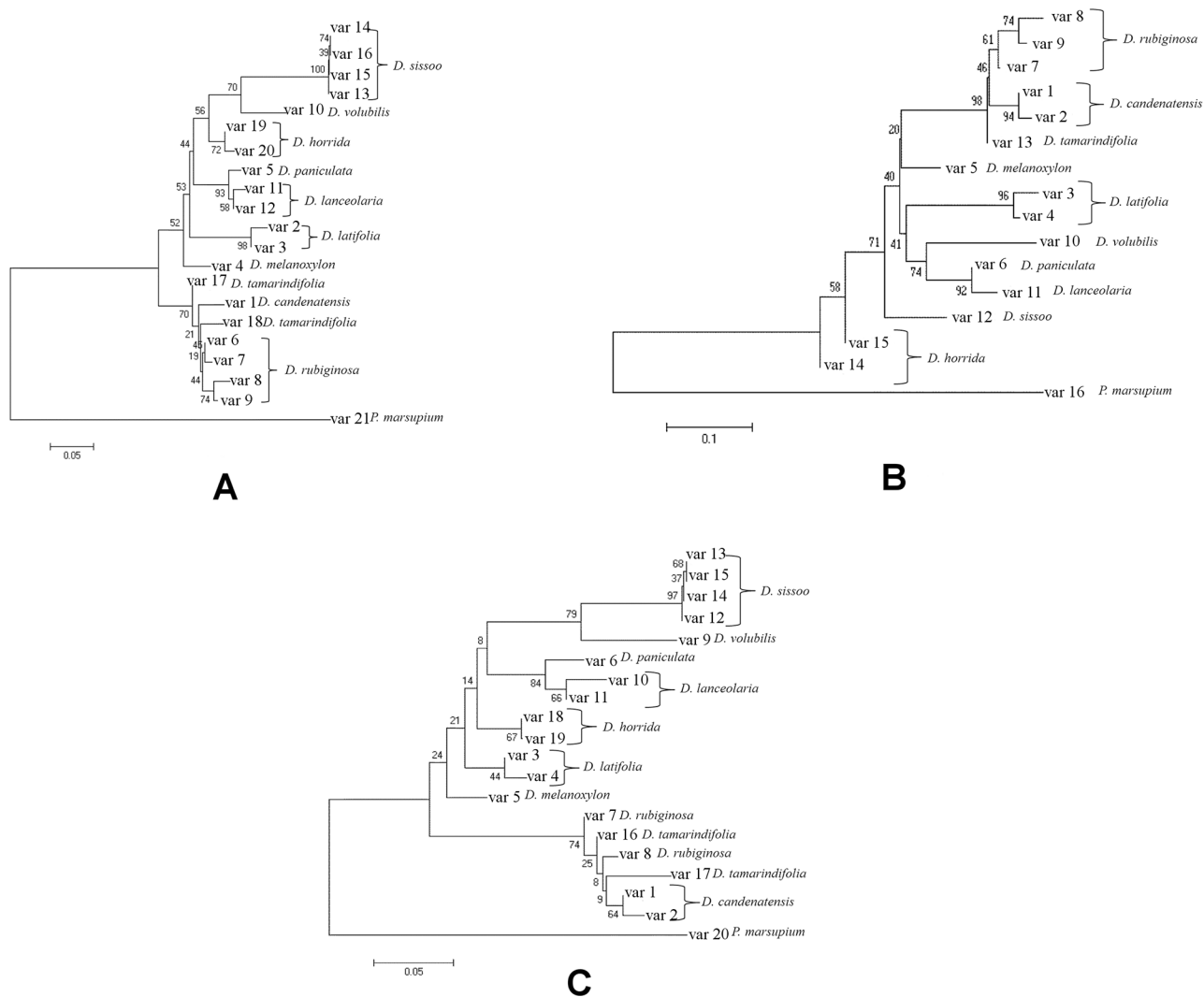


**Fig 4. Single locus NJ trees.** NJ trees were constructed using MEGA 5.0 based on K2P distance model—**A**, *matK*; **B**, *rbcL*; **C**, *trnH-psbA*, **D**, *nrITS*.

doi:10.1371/journal.pone.0142965.g004

### Assessment of the four candidate barcodes in *Dalbergia* genus

In the present study, the amplification and sequencing success rate in *Dalbergia* ranged from 80.5% (for *nrITS*) to 97.6% (for *rbcL*). While the *rbcL* locus was reported to be easy to amplify and sequence across a broad range of plant taxa, but offers low species resolution, the rapidly evolving *matK* locus, is known for its high discriminatory power with low universality [56]. Hence, the *matK* is popular for species discrimination in case of angiosperms [3]. However, mixed results ranging from high success rate [56, 57] to poor discrimination [3, 11] have been reported for *matK*. Even in the present study, *matK* showed good resolving power and although *trnH-psbA* showed good universality and higher discrimination, it also has variable length, presence of homopolymers, inversions and insertion of *rps19* gene [58–60]. Similarly, while the *nrITS* locus is a commonly used nuclear marker for phylogenetic studies [5], it was, however, not preferred for barcoding studies initially because of fungal contamination, paralogous gene copies and problems in recovery [8]. In our study, similarity search using BLAST did



**Fig 5. NJ trees with combined loci.** NJ trees constructed using MEGA 5.0 based on K2P distance model—**A**, *matK+rbcL*; **B**, *matK+trnH-psbA*; **C**, *rbcL + trnH-psbA*.

doi:10.1371/journal.pone.0142965.g005

not reveal any problem of fungal contamination in *nrITS* sequences; however, the sequencing success was low (80%), which might be due to the presence of divergent gene copies as reported earlier [5]. In case of *trnH-psbA* which gave 94.7% sequencing success, our data revealed the presence of T and A repeats, without any insertion of *rps19* gene when checked by BLAST.

The overall interspecific distances were high compared to intraspecific distances and no significant barcode gap was observed in the present study. Usually in the closely related plant species, plastid regions such as *rbcL* and *matK* do not generate a barcode gap [57]. Several studies have also revealed the absence of barcode gap in different plant systems such as *Agalinis* [9], *Parnassia* [50], *Gossypium* [51] medicinal plants [12] and *Dioscorea* [61]. Furthermore in the NJ tree based analysis, *nrITS*, *matK* and *trnH-psbA* and their combinations formed separate clusters for each species. However, *rbcL* could not differentiate *D. rubiginosa*, *D. candenatensis* and *D. tamarindifolia*, which could be because of the conserved nature of the gene [62]. Similar

**Table 6. Results from similarity based analysis using TaxonDNA.**

Regions	Best Match			Best Close Match				
	Correct	Ambiguous	Incorrect	Correct	Ambiguous	Incorrect	Sequence without any match closer than threshold	Threshold (%)
<i>matK</i>	165 (100.00)	0 (0.00)	0 (0.00)	165 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.74
<i>rbcL</i>	126 (75.90)	40 (24.09)	0 (0.00)	126 (75.90)	40 (24.09)	0 (0.00)	0 (0.00)	0.20
<i>trnH-psbA</i>	109 (67.70)	51 (31.67)	1 (0.62)	108 (67.08)	51 (31.67)	0 (0.00)	2 (1.24)	0.37
<i>nrlTS</i>	134 (99.30)	0 (0.00)	1 (0.73)	135 (99.26)	0 (0.00)	1 (0.73)	0 (0.00)	1.20
<i>matK+ trnH-psbA</i>	157 (100.00)	0 (0.00)	0 (0.00)	157 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.52
<i>matK+ rbcL</i>	163 (100.00)	0 (0.00)	0 (0.00)	163 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.25
<i>rbcL+trnH-psbA</i>	135 (85.98)	21 (13.37)	1 (0.63)	135 (85.98)	21 (13.37)	0 (0.00)	1 (0.63)	0.12

**Note:** Values in parentheses are expressed in %.

doi:10.1371/journal.pone.0142965.t006

**Table 7. Character based approach for species identification in *Dalbergia*.**

Locus	cc	wc	nc	Formula									
				Dc	Dlat	Dm	Dp	Dr	Dv	Dlan	Ds	Dt	Dh
<i>matK</i>	100	0	0	362 = A	84 = T	206 = A	28 = G, 166 = T, 368 = G	440 = G	7 = C	166 = A	51 = G	292 = T, 362 = T, 440 = T	368 = G, 422 = A
<i>rbcL</i>	76.56	0	23.44	339 = A	191 = T	19 = T	35 = T, 186 = G	-	485 = T	35 = C	19 = A, 179 = A, 86 = A, 191 = C, 458 = G, 485 = C	-	458 = T
<i>trnH-psbA</i>	69.35	0	30.65	24 = T	-	12 = C	118 = G	139 = A	52 = C, 114 = A, 228 = A	26 = T	114 = G, 228 = A	52 = A, 228 = A	114 = G, 228 = G
<i>nrlTS</i>	100	0	0	107 = C	122 = C, 456 = T	83 = C	621 = A	132 = T, 231 = G	43 = C	132 = T, 456 = C	128 = A	107 = G	539 = C, 637 = C
<i>matK+rbcL</i>	100	0	0	1052 = T	711 = T	697 = A	95 = A, 657 = T, 711 = G, 1011 = A	931 = G OR 783 = A	1011 = G	657 = A	1018 = G	783 = T, 801 = A, 931 = T, 1052 = C	922 = T
<i>matK+trnH-psbA</i>	100	0	0	362 = A	84 = T	206 = A	795 = G	440 = G OR 292 = A	422 = C, 905 = A	166 = A	422 = A, 905 = A	292 = T, 362 = T, 440 = T	422 = A, 905 = G
<i>rbcL+trnH-psbA</i>	86.99	0	13.01	515 = T	191 = T	503 = C	609 = G	630 = A	495 = T, 734 = C	95 = A, 191 = C, 734 = A	495 = G	-	191 = C, 503 = A, 686 = A, 734 = A

**Note:** cc: correctly classified, wc: wrongly classified, nc: not classified

**Species codes:-** Dc: *D. candenatensis*, Dlat: *D. latifolia*, Dm: *D. melanoxyton*, Dp: *D. paniculata*, Dr: *D. rubiginosa*, Dv: *D. volubilis*, Dlan, *D. lanceolaria*, Ds: *D. sissou*, Dt: *D. tamarindifolia*, Dh: *D. horrida*.

doi:10.1371/journal.pone.0142965.t007

**Table 8. Comparative ranking of loci used in DNA barcoding of *Dalbergia*.**

Parameters	<i>matK</i>	<i>rbcL</i>	<i>trnH-psbA</i>	<i>nrITS</i>	<i>matK+rbcL</i>	<i>matK+ trnH-psbA</i>	<i>rbcL+ trnH-psbA</i>
Barcode	10	7	8	8	10	8	8
Inter and intra specific distances	10	7	8	8	10	8	8
Best match and best close match	10	7.2	4.9	6.6	10	10	3.3
BLOG	10	7	7	7.5	10	8.6	6.4
Wilcoxon Signed Rank test	4	5	9	5.5	3.5	5	6
<b>Total</b>	<b>44</b>	<b>33.2</b>	<b>36.9</b>	<b>35.6</b>	<b>43.5</b>	<b>39.6</b>	<b>31.7</b>

**Note:** Larger values indicate better performance.

doi:10.1371/journal.pone.0142965.t008

behavior of *rbcL* was also reported in *Carex* [58]. Together this suggested that individually *rbcL* might not serve as a good barcode but can be utilized in combination with other loci.

A recent report on DNA barcoding of eight *Dalbergia* species from Vietnam recommended *ITS* locus as a potential barcode based on UPGMA analysis and nucleotide diversity [63]. It has been reported that being a multigene family, 18s-26s rDNA is subjected to concerted evolution. In certain cases, *ITS1* [64, 65] and *ITS2* [12, 60, 65, 66] have been used as separate loci for DNA barcoding. However, point mutations displayed by *ITS1* and *ITS2* also contribute to high intra-specific variations [67]. We used the complete ITS region (ITS1-5.8S-ITS2) as a single barcoding locus. In our study, *nrITS* showed high intraspecific variation with high species discrimination, leading to incorrect identification with BM and BCM. However, DNA barcoding of eight *Dalbergia* species from Vietnam [63], did not use the species from the current study. A reanalysis of the data from NCBI for the species used in the Vietnam study along with dataset from our study revealed a high number of sequence variants for most of the species (S1 Fig). Moreover, from the available sequence data in NCBI for the Vietnam study [63], we could find only one *nrITS* sequence each for *D. dialoides*, *D. entadoides* and *D. hancei* making it difficult to assay the intraspecific variation. It was therefore, not possible to comment on either the intraspecific diversity of these species, which is an important factor for DNA barcoding or the suitability of *nrITS* as the potential barcode for *Dalbergia* species. It is essential to sample enough number of accessions for each of these species, ideally from different geographical locations, to sample the intraspecific variation from the entire distributional range [53].

## Conclusions

In the present study 7–26 accessions of ten *Dalbergia* species each collected from different geographic locations in WG region of India were screened using 37 primer pairs from nuclear and plastid genes. Four loci (*rbcL*, *matK*, *trnH-psbA* and *nrITS*) and their combinations were further evaluated with five different analyses and ranked based on their performance. These studies have revealed *matK* and *matK+rbcL* loci as the most suitable barcodes to discriminate *Dalbergia* species.

## Supporting Information

**S1 Fig. NJ tree.** Combined analysis of *nrITS* sequences submitted by Phong et al. [63] with those generated in this study, revealing high intraspecific variation and several sequence variants for most species.  
(TIF)

**S2 Fig. NJ tree.** Representative tree for *matK+rbcL* using all the individuals without any division. Dc: *D. candenatensis*, Dlat: *D. latifolia*, Dm: *D. melanoxydon*, Dp: *D. paniculata*, Dr: *D.*

*rubiginosa*, Dv: *D. volubilis*, Dlan: *D. lanceolaria*, Ds: *D. sissoo*, Dt: *D. tamarindifolia*, Dh: *D. horrida*.

(TIF)

**S1 Dataset. Sample details.** List of all samples with collection details and GenBank accession numbers.

(DOCX)

**S2 Dataset. Primer details.** Primers used in DNA barcoding of *Dalbergia* species.

(DOCX)

**S3 Dataset. PCR reaction details.** PCR conditions for *matK*, *rbcl*, *trnH-psbA* and *nrITS*

(DOCX)

## Acknowledgments

RMB is thankful to Council of Scientific and Industrial Research (CSIR) for senior research fellowship; Dr. Sachin Punekar (Biospheres, Pune, India) and Dr. P. Tetali (Temple Rose Construction, Private Ltd, Pune), Mr. Amol Kasodekar and Mr. Amol Jadhav (CSIR-NCL, Pune) for their help during sample collections; Dr. Neelesh Dahanukar (IISER, Pune) and Dr. Shobha Rao (Research & Training Society for Initiatives in Nutrition and Development, Pune) for the help in data analysis; and Dr. Anargha Wakhare (Department of Geography, Nowrosjee Wadia College, Pune) for her help in preparing the map. Dr. Dhanasekaran Shanmugam, (CSIR-NCL, Pune) is gratefully acknowledged for thorough reading of the manuscript. Financial support in the form of Department of Biotechnology (DBT) grant (GAP267426) and CSIR grant (Project code: BSC0106) to CSIR-NCL is gratefully acknowledged.

## Author Contributions

Conceived and designed the experiments: VSG NYK. Performed the experiments: RMB BBD MB NYK. Analyzed the data: RMB BBD NYK. Contributed reagents/materials/analysis tools: MB VSG NYK. Wrote the paper: RMB BBD NYK VSG.

## References

1. Hebert PDN, Gregory TR. The promise of DNA barcoding for taxonomy.—*Syst Biol.* 2005; 54(5):852–9. doi: [10.1080/10635150500354886](https://doi.org/10.1080/10635150500354886) ISI:000232883700014. PMID: [16243770](https://pubmed.ncbi.nlm.nih.gov/16243770/)
2. Ren BQ, Xiang XG, Chen ZD. Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. *Mol Ecol Resour.* 2010; 10(4):594–605. doi: [10.1111/j.1755-0998.2009.02815.x](https://doi.org/10.1111/j.1755-0998.2009.02815.x) ISI:000278676300002. PMID: [21565064](https://pubmed.ncbi.nlm.nih.gov/21565064/)
3. Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, Graham SW, Barrett SCH, et al. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Mol Ecol Resour.* 2009; 9:130–9. doi: [10.1111/j.1755-0998.2009.02652.x](https://doi.org/10.1111/j.1755-0998.2009.02652.x) ISI:000265227700013. PMID: [21564972](https://pubmed.ncbi.nlm.nih.gov/21564972/)
4. Palmer JD. Evolution of chloroplast and mitochondrial DNA in plants and algae. In: MacIntyre RJ, editor. MacIntyre ILI (ed) Monographs in evolutionary biology: Molecular evolutionary genetics. Plenum, New York 1985. p. 131–240.
5. Alvarez I, Wendel JF. Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol.* 2003; 29(3):417–34. doi: [10.1016/S1055-7903\(03\)00208-2](https://doi.org/10.1016/S1055-7903(03)00208-2) ISI:000186738000005. PMID: [14615184](https://pubmed.ncbi.nlm.nih.gov/14615184/)
6. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci.* 2005; 102(23):8369–74. doi: [10.1073/pnas.0503123102](https://doi.org/10.1073/pnas.0503123102) ISI:000229650500053. PMID: [15928076](https://pubmed.ncbi.nlm.nih.gov/15928076/)
7. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, et al. A proposal for a standardised protocol to barcode all land plants. *Taxon.* 2007; 56(2):295–9. ISI:000247420000004.

8. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. Plos One. 2011; 6(5). doi: [10.1371/journal.pone.0019254](https://doi.org/10.1371/journal.pone.0019254) ISI:000291052200009.
9. Pettengill JB, Neel MC. An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus *Agalinis* (Orobanchaceae). Am J Bot. 2010; 97(8):1391–406. doi: [10.3732/Ajb.0900176](https://doi.org/10.3732/Ajb.0900176) ISI:000280481800015. PMID: [21616891](https://pubmed.ncbi.nlm.nih.gov/21616891/)
10. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, et al. Land plants and DNA barcodes: short-term and long-term goals.—Phil Trans R Soc. B-Biol Sci. 2005; 360(1462):1889–95. doi: [10.1098/rstb.2005.1720](https://doi.org/10.1098/rstb.2005.1720) ISI:000232719300009.
11. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: The coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. Plos One. 2007; 2(6). doi: [10.1371/journal.pone.0000508](https://doi.org/10.1371/journal.pone.0000508) ISI:000207451500017.
12. Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. Plos One. 2010; 5(1). doi: [10.1371/Journal.Pone.0008613](https://doi.org/10.1371/Journal.Pone.0008613) ISI:000273414100007.
13. Ribeiro RA, Ramos ACS, Filho JPD, Lovato MB. Genetic variation in remnant populations of *Dalbergia nigra* (Papilionoideae), an endangered tree from the Brazilian Atlantic forest. Ann Bot-London. 2005; 95(7):1171–7. doi: [10.1093/Aob/Mci128](https://doi.org/10.1093/Aob/Mci128) ISI:000229583500010.
14. Vatanparast M, Klitgard BB, Adema FACB, Pennington RT, Yahara T, Kajita T. First molecular phylogeny of the pantropical genus *Dalbergia*: implications for infrageneric circumscription and biogeography. S Afr J Bot. 2013; 89:143–9. doi: [10.1016/j.sajb.2013.07.001](https://doi.org/10.1016/j.sajb.2013.07.001) ISI:000328808400014.
15. Hiremath SC, Nagasampige MH. Genome size variation and evolution in some species of *Dalbergia* Linn.f. (Fabaceae). Caryologia. 2004; 57(4):367–72. ISI:000228079000007.
16. Nithaniyal S, Newmaster SG, Ragupathy S, Krishnamoorthy D, Vassou SL, Parani M. DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. Plos One. 2014; 9(9):e107669. Epub 2014/09/27. doi: [10.1371/journal.pone.0107669](https://doi.org/10.1371/journal.pone.0107669) PMID: [25259794](https://pubmed.ncbi.nlm.nih.gov/25259794/); PubMed Central PMCID: [PMC4178033](https://pubmed.ncbi.nlm.nih.gov/PMC4178033/).
17. Bentham G. Synopsis of Dalbergieae, a tribe of the Leguminosae. 1860; J Proc Linn Soc., Bot. IV (Supplement):1–134.
18. Carvalho Ad. Systematic studies in the genus *Dalbergia* L. f. in Brazil: University of Reading; 1989.
19. Prain D. The species of *Dalbergia* of South-eastern Asia. Ann Roy Bot Gard. (Calcutta). 1904;(10):1–114.
20. Thothathri K. Taxonomic revision of the tribe Dalbergieae in the Indian subcontinent: Botanical Survey of India (Calcutta); 1987. 244 p.
21. Carvalho A. A synopsis of the genus *Dalbergia* (Fabaceae: Dalbergieae) in Brazil. Brittonia. 1997; 49(1):87–109. doi: [10.2307/2807701](https://doi.org/10.2307/2807701)
22. Sunarno B, Ohashi H. *Dalbergia* (Leguminosae) of Borneo. J Japan Bot. 1997; 72(4):198–220.
23. Niyomdham C. An account of *Dalbergia* (Leguminosae-Papilionoideae) in Thailand. Thailand Forest Bulletin (BOT). 2002; 30:124–66.
24. Lavin M, Pennington RT, Klitgaard BB, Sprent JI, de Lima HC, Gasson PE. The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. Am J Bot. 2001; 88(3):503–33. doi: [10.2307/2657116](https://doi.org/10.2307/2657116) ISI:000167595000017. PMID: [11250829](https://pubmed.ncbi.nlm.nih.gov/11250829/)
25. Ribeiro RAMatt L; Lemos-Filho José Pires; Filho Carlos Victor Mendonça; Santos Fabrício Rodrigues dos; Lovato Maria Bernadete. The genus *Machaerium* (Leguminosae) is more closely related to *Aeschynomene* sect. *Ochopodium* than to *Dalbergia*: Inferences from combined sequence data. Phytochemistry. 2007; 32(4):762–71(10).
26. Mohana GS, Shaanker RU, Ganeshiaiah KN, Dayanandan S. Genetic relatedness among developing seeds and intra fruit seed abortion in *Dalbergia sissoo* (Fabaceae). Am J Bot. 2001; 88(7):1181–8. doi: [10.2307/3558328](https://doi.org/10.2307/3558328) ISI:000170012400004. PMID: [11454617](https://pubmed.ncbi.nlm.nih.gov/11454617/)
27. Rout GR, Bhattacharya D, Nanda RM, Nayak S, Das P. Evaluation of genetic relationships in *Dalbergia* species using RAPD markers. Biodivers Conserv. 2003; 12(2):197–206. doi: [10.1023/A:1021996020947](https://doi.org/10.1023/A:1021996020947) ISI:000180344500002.
28. Arif M, Zaidi NW, Singh YP, Haq QMR, Singh US. A comparative analysis of ISSR and RAPD markers for study of genetic diversity in Shisham (*Dalbergia sissoo*). Plant Mol Biol Report. 2009; 27(4):488–95. doi: [10.1007/s11105-009-0097-0](https://doi.org/10.1007/s11105-009-0097-0) ISI:000270780900009.
29. Bakshi M, Sharma A. Assessment of genetic diversity in *Dalbergia sissoo* clones through RAPD profiling. J Fores Res. 2011; 22(3):393–7. doi: [10.1007/s11676-011-0183-7](https://doi.org/10.1007/s11676-011-0183-7)
30. Richards E, Reichardt M, Rogers S. Preparation of genomic DNA from plant tissue. Curr Protoc Mol Biol. 1994; 1:2.3.1–2.3.7.



31. Tripathi AM, Tyagi A, Kumar A, Singh A, Singh S, Chaudhary LB, et al. The internal transcribed spacer (ITS) region and *trnH-psbA* are suitable candidate loci for DNA barcoding of tropical tree species of India. *Plos One*. 2013; 8(2). doi: [10.1371/journal.pone.0057934](https://doi.org/10.1371/journal.pone.0057934) ISI:000315519000170.
32. Thompson JD, Higgins DG, Gibson TJ. Clustal-W—Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22(22):4673–80. doi: [10.1093/nar/22.22.4673](https://doi.org/10.1093/nar/22.22.4673) ISI: A1994PU19900018. PMID: [7984417](https://pubmed.ncbi.nlm.nih.gov/7984417/)
33. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28(10):2731–9. doi: [10.1093/molbev/msr121](https://doi.org/10.1093/molbev/msr121) ISI:000295184200003. PMID: [21546353](https://pubmed.ncbi.nlm.nih.gov/21546353/)
34. Padhye A, Pandit R, Patil R, Gaikwad S, Dahanukar N, Shouche Y. Range extension of Ferguson's Toad *Duttaphrynus scaber* (Schneider) (Amphibia: Anura: Bufonidae) up to the northern most limit of Western Ghats, with its advertisement call analysis. *J Threat Taxa*. 2013.
35. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J Mol Evol*. 1980; 16(2):111–20. doi: [10.1007/Bf01731581](https://doi.org/10.1007/Bf01731581) ISI: A1980KW57300003. PMID: [7463489](https://pubmed.ncbi.nlm.nih.gov/7463489/)
36. Chen R, Jiang LY, Liu L, Liu QH, Wen J, Zhang RL, et al. The *gnd* gene of *Buchnera* as a new, effective DNA barcode for aphid identification. *Sys Entomol*. 2013; 38(3):615–25. doi: [10.1111/Syen.12018](https://doi.org/10.1111/Syen.12018) ISI:000320560100011.
37. Meier R, Shiyang K, Vaidya G, Ng PKL. DNA barcoding and taxonomy in diptera: A tale of high intra-specific variability and low identification success. *Syst Biol*. 2006; 55(5):715–28. doi: [10.1080/10635150600969864](https://doi.org/10.1080/10635150600969864) ISI:000246721800001. PMID: [17060194](https://pubmed.ncbi.nlm.nih.gov/17060194/)
38. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. 2003; 19(18):2496–7. doi: [10.1093/bioinformatics/btg359](https://doi.org/10.1093/bioinformatics/btg359) ISI:000187217700029. PMID: [14668244](https://pubmed.ncbi.nlm.nih.gov/14668244/)
39. Weitschek E, Van Velzen R, Felici G, Bertolazzi P. BLOG 2.0: a software system for character-based species classification with DNA barcode sequences. What it does, how to use it. *Mol Ecol Resour*. 2013; 13(6):1043–6. doi: [10.1111/1755-0998.12073](https://doi.org/10.1111/1755-0998.12073) ISI:000325627700008. PMID: [23350601](https://pubmed.ncbi.nlm.nih.gov/23350601/)
40. Bertolazzi P, Felici G, Weitschek E. Learning to classify species with barcodes. *BMC Bioinformatics*. 2009; 10. doi: [10.1186/1471-2105-10-S14-S7](https://doi.org/10.1186/1471-2105-10-S14-S7) ISI:000271765800007.
41. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc Biol Sci*. 2003; 270(1512):313–21. Epub 2003/03/05. doi: [10.1098/rspb.2002.2218](https://doi.org/10.1098/rspb.2002.2218) PMID: [12614582](https://pubmed.ncbi.nlm.nih.gov/12614582/).
42. Hebert PD, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci*. 2003; 270 Suppl 1:S96–9. Epub 2003/09/04. doi: [10.1098/rsbl.2003.0025](https://doi.org/10.1098/rsbl.2003.0025) PMID: [12952648](https://pubmed.ncbi.nlm.nih.gov/12952648/).
43. Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J. Testing candidate plant barcode regions in the Myristicaceae. *Mol Ecol Notes*. 2008; 8(3):480–90. doi: [10.1111/j.1471-8286.2007.02002.x](https://doi.org/10.1111/j.1471-8286.2007.02002.x)
44. Wang WQ, Wu YR, Yan YH, Ermakova M, Kerstetter R, Messing J. DNA barcoding of the Lemnaceae, a family of aquatic monocots. *BMC Plant Biol*. 2010; 10. doi: [10.1186/1471-2229-10-205](https://doi.org/10.1186/1471-2229-10-205) ISI:000283249100002.
45. Shi LC, Zhang J, Han JP, Song JY, Yao H, Zhu YJ, et al. Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. *J Syst Evol*. 2011; 49(3):261–6. doi: [10.1111/j.1759-6831.2011.00133.x](https://doi.org/10.1111/j.1759-6831.2011.00133.x) ISI:000291236500012.
46. Little DP, Knopf P, Schulz C. DNA barcode identification of Podocarpaceae—the second largest Conifer family. *Plos One*. 2013; 8(11). doi: [10.1371/journal.pone.0081008](https://doi.org/10.1371/journal.pone.0081008) ISI:000327652100057.
47. Zhang JM, Wang JX, Xia T, Zhou SL. DNA barcoding: species delimitation in tree peonies. *Science in China Series C-Life Sciences*. 2009; 52(6):568–78. doi: [10.1007/s11427-009-0069-5](https://doi.org/10.1007/s11427-009-0069-5) ISI:000267396600010.
48. Newmaster SG, Ragupathy S. Testing plant barcoding in a sister species complex of pantropical *Acacia* (Mimosoideae, Fabaceae). *Mol Ecol Resour*. 2009; 9:172–80. doi: [10.1111/j.1755-0998.2009.02642.x](https://doi.org/10.1111/j.1755-0998.2009.02642.x) ISI:000265227700017. PMID: [21564976](https://pubmed.ncbi.nlm.nih.gov/21564976/)
49. Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB. DNA barcoding of endangered Indian *Paphiopedilum* species. *Mol Ecol Resour*. 2012; 12(1):82–90. doi: [10.1111/j.1755-0998.2011.03071.x](https://doi.org/10.1111/j.1755-0998.2011.03071.x) PMID: [21951639](https://pubmed.ncbi.nlm.nih.gov/21951639/)
50. Yang JB, Wang YP, Moller M, Gao LM, Wu D. Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). *Mol Ecol Resour*. 2012; 12(2):267–75. doi: [10.1111/j.1755-0998.2011.03095.x](https://doi.org/10.1111/j.1755-0998.2011.03095.x) ISI:000299930300009. PMID: [22136257](https://pubmed.ncbi.nlm.nih.gov/22136257/)

51. Ashfaq M, Asif M, Anjum ZI, Zafar Y. Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (*Gossypium*: Malvaceae). *Mol Ecol Resour.* 2013; 13(4):573–82. doi: [10.1111/1755-0998.12089](https://doi.org/10.1111/1755-0998.12089) ISI:000320396300003. PMID: [23480447](https://pubmed.ncbi.nlm.nih.gov/23480447/)
52. Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Res.* 2011:1–11. doi: [10.1093/nar/gkr732](https://doi.org/10.1093/nar/gkr732)
53. Gonzalez MA, Baraloto C, Engel J, Mori SA, Petronelli P, Riera B, et al. Identification of Amazonian trees with DNA barcodes. *Plos One.* 2009; 4(10):e7483. Epub 2009/10/17. doi: [10.1371/journal.pone.0007483](https://doi.org/10.1371/journal.pone.0007483) PMID: [19834612](https://pubmed.ncbi.nlm.nih.gov/19834612/); PubMed Central PMCID: PMC2759516.
54. Gao T, Yao H, Song JY, Liu C, Zhu YJ, Ma XY, et al. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. *J Ethnopharmacol.* 2010; 130(1):116–21. doi: [10.1016/j.jep.2010.04.026](https://doi.org/10.1016/j.jep.2010.04.026) ISI:000279886900017. PMID: [20435122](https://pubmed.ncbi.nlm.nih.gov/20435122/)
55. Edwards D, Horn A, Taylor D, Savolainen V, Hawkins J. DNA barcoding of a large genus, *Aspalathus* L. (Fabaceae). *Taxon.* 2008; 57(4):1317–27.
56. Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. *Proc Natl Acad Sci.* 2009; 106(31):12794–7. doi: [10.1073/pnas.0905845106](https://doi.org/10.1073/pnas.0905845106) ISI:000268667600043. PMID: [19666622](https://pubmed.ncbi.nlm.nih.gov/19666622/)
57. Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci.* 2008; 105(8):2923–8. doi: [10.1073/pnas.0709936105](https://doi.org/10.1073/pnas.0709936105) ISI:000253567900033. PMID: [18258745](https://pubmed.ncbi.nlm.nih.gov/18258745/)
58. Starr JR, Naczi RFC, Chouinard BN. Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). *Mol Ecol Resour.* 2009; 9:151–63. doi: [10.1111/j.1755-0998.2009.02640.x](https://doi.org/10.1111/j.1755-0998.2009.02640.x) ISI:000265227700015. PMID: [21564974](https://pubmed.ncbi.nlm.nih.gov/21564974/)
59. Whitlock BA, Hale AM, Groff PA. Intraspecific inversions pose a challenge for the *trnH-psbA* plant DNA barcode. *Plos One.* 2010; 5(7). doi: [10.1371/journal.pone.0011533](https://doi.org/10.1371/journal.pone.0011533) ISI:000279822300007.
60. Pang XH, Liu C, Shi LC, Liu R, Liang D, Li H, et al. Utility of the *trnH-psbA* intergenic spacer region and its combinations as plant DNA barcodes: A meta-analysis. *Plos One.* 2012; 7(11). doi: [10.1371/journal.pone.0048833](https://doi.org/10.1371/journal.pone.0048833) ISI:000311151900046.
61. Sun XQ, Zhu YJ, Guo JL, Peng B, Bai MM, Hang YY. DNA barcoding the *Dioscorea* in China, a vital group in the evolution of monocotyledon: Use of *matK* gene for species discrimination. *Plos One.* 2012; 7(2). doi: [10.1371/journal.pone.0032057](https://doi.org/10.1371/journal.pone.0032057) ISI:000302871500108.
62. Albert VA, Backlund A, Bremer K, Chase MW, Manhart JR, Mishler BD, et al. Functional constraints and *rbcL* evidence for land plant phylogeny. *Ann Mol Bot Gard.* 1994; 81(3):534–67. doi: [10.2307/2399902](https://doi.org/10.2307/2399902) ISI:A1994PA50800006.
63. Phong DT, Tang DV, Hien VTT, Ton ND, Van HN. Nucleotide diversity of a nuclear and four chloroplast DNA regions in rare tropical wood species of *Dalbergia* in Vietnam: a DNA barcode identifying utility. *Asian J Appl Sci.* 2014; 02(02):116–25.
64. Campbell CS, Wright WA, Cox M, Vining TF, Major CS, Arsenault MP. Nuclear ribosomal DNA internal transcribed spacer 1 (ITS1) in *Picea* (Pinaceae): sequence divergence and structure. *Mol Phylogenet Evol.* 2005; 35(1):165–85. doi: [10.1016/j.ympev.2004.11.010](https://doi.org/10.1016/j.ympev.2004.11.010) ISI:000227602600012. PMID: [15737589](https://pubmed.ncbi.nlm.nih.gov/15737589/)
65. Blaaliid R, Kumar S, Nilsson RH, Abarenkov K, Kirk PM, Kauserud H. ITS1 versus ITS2 as DNA meta-barcodes for fungi. *Mol Ecol Resour.* 2013; 13(2):218–24. doi: [10.1111/1755-0998.12065](https://doi.org/10.1111/1755-0998.12065) ISI:000315032600007. PMID: [23350562](https://pubmed.ncbi.nlm.nih.gov/23350562/)
66. Han J, Shi L, Chen X, LIN Y. Comparison of four DNA barcodes in identifying certain medicinal plants of Lamiaceae. *J Syst Evol.* 2012; 50(3):227–34. doi: [10.1111/j.1759-6831.2012.00184.x](https://doi.org/10.1111/j.1759-6831.2012.00184.x)
67. Baldwin BG. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Mol Phylogenet Evol.* 1992; 1(1):3–16. doi: [10.1016/1055-7903\(92\)90030-K](https://doi.org/10.1016/1055-7903(92)90030-K) ISI:000207480900002. PMID: [1342921](https://pubmed.ncbi.nlm.nih.gov/1342921/)