Current Trends in Biotechnology and Pharmacy
Vol. 18(3) 1951-1961, July 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.3.40

The Role of trnH-pbsA Spacer Gene in *Eucalyptus* Species Identification and its Importance in Phylogenetics

G. Jayaraj

Centre For Research and Dept. of Botany and Biotechnology,
Andhra Loyola College (Autonomous), Vijayawada 520 008, Andhra Pradesh.
*Corresponding author: jgorantla3@gmail.com

Abstract

Our present study is made to use **t**he trnH-psbA spacer gene data to resolve some phylogenetic relationships among the *Eucalyptu*s speices. The species included are *E. moorei, E.affn. moorei, E. dives, E. mitchelliana, E. pauciflora* and E. *stellulata.* There were successful amplifications of DNA for the seven *Eucalyptus* species using gene trnH-psbA. The trnH-psbA spacer is the short with 476 bps. It is to be complimented with rbcL, MatK and trnH-psbA gene study. The trnH-psbA spacer is showing 21 SNPs out of 476 bps segment. The mean length is 733 bps long. Amplification success rate is almost 100%. For each species there are multiple good quality sequences available for the sequence alignment. The trnH-psbA spacer has discriminated 5 species clearly from the rest and the other 3 species are also separated from each other though the distance between them is not significantly noticeable.

Key words: trnH-psbA spacer, SNPs, variable gene, phylogenetics, complementary

Introduction

In the past few decades, the Genomic research has undergone tremendous changes. There have been several newer technical advances created to help the structural and functional aspects of genes, chromosomes and sometimes the entire genome. e.g. the sequencing of the entire Human genome, of *Arabidopsis, Rice, o*r *Popular* genomes. There are several other researches are taking place recently. It includes the *Eucalyptus* that has the sequencing of the entire genome. It has a wider application. Genetic information can help us with a good resolution of species boundaries. This will eventually may give insights into the patterns and rates of evolutionary diversification among species.

In DNA Barcoding we use a short selection DNA (portion of a gene) in order to identify a species. This is called DNA barcoding. The DNA barcoding is a newer system created to provide accurate and automatable species identifications by using short and variable standardized gene regions as species tags or species identity. This has initiated a new method/ technique and eventually led to the formation of Consortium for the Barcode of Life (5). Our objective of this project is to obtain DNA barcodes from all species of *Eucalyptus* for all over the world, from different geographical and climatic regions. Our challenge is also to find out a very suitable region which shows enough variation within it to discriminate among species yet conserved enough to be present.

A critique of barcoding

The advocates of DNA barcoding say that this technology would revitalize biological collections and speed up species identification and inventories. There are only 1.7 million specimens that have been identified by taxonomists

and about 10-20 million more which have not been named or explained.

The opponents of barcoding argue that this technology would destroy the traditional systematics and turn it into a service industry (6). These fears are allayed and in all cases DNA barcoding are applied only in conjunction with classical approach where species are simply unknown or no attempts have been made to delimit them. Therefore, barcoding as originally intended would be limited in its applicability. The DNA barcoding address by matching DNA sequences to 'known' species. As in the words of CBOL "barcoding is neither a substitution for alpha taxonomy nor about interfering phylogenies (20). Apart from being a diagnostic tool, barcode sequences per se and their ever-increasing taxonomic coverage could become an unprecedented resource for taxonomy and systematic studies. In future in plants multiple markers is likely to be a necessity and it is already being explored (5). It is possible some taxa can be established from the sequence variation alone and re-identified unequivocally while awaiting morphological analysis and formal description, i.e. the 'reverse taxonomy'. Using DNA barcoding the present research is carried out in order to find out the phylogenetic relationship of 6 closely related and 2 distantly related eucalypts.

It has been reported that the Chloroplasts are maternally inherited in most Angiosperms and for the eucalypts too this has been demonstrated by Byrne et al., (1993) (1). in *E. nitens* of eastern Australia. Schael et al., (1999) (19). showed a uniparental inheritance of the chloroplast genome and lack of recombination. The variations in the chloroplast genome are ordered accordingly. Since the Chloroplast genome is non- recombining, asexually inherited and evolves slowly, these characters are useful for the estimation of the extension of gene flow between species. Many studies in recent years have been focused on the chloroplast DNA variations in eucalypts because of their economic and ecological significance apart from other important genomic studies. Eucalypts have dominated different land scape of Australia over the thousands of centuries.

***trnH-PsbA* gene and barcoding:**

Kress et al., (2005) (5). Shaw et al., (2005) (20). and other researchers showed that trnH-psbA region is one of the most valuable non-coding regions of plastid genome. It has been shown that this region with highest variable sites makes it a good candidate for the species discrimination. However, there are problems to align the sequences because of high rates of insertions/deletions. Kress et al., (2006) (12). found that *trnH-psbA* spacer is very short (less than 300 bp). The alignment of *trnH-psbA* spacer across bigger families of angiosperms remains highly ambiguous. There exists also great length deference. In contrast with the indels problem for phylogenetic construction, Kress et al., (2005) (5), think that indels would ultimately enhance the information needed for species identification. The highest divergence is provided by trnH-psbA spacer from the studies.

Materials and Methods

Finding the suitable regions of the genome

In the present study we are taking into consideration multi locus region *trnH-psbA* of the genome for the barcoding of *8 Eucalypt* species. The psbA-trnH intergenic spacer is one of the most variable non-coding regions of the plastid genome in Angiosperms. It has highest percentage of variable sites (20). In some group of plants, it is relatively short, having less than 300 bps (12). The The chloroplast gene rbcL encodes the large subunit of ribulose bisphosphate carboxylase in plants.

The taxa selection for barcoding

The taxa selected for the study includes *E. moorei, E.affn. moorei, E. dives, E. mitchelliana, E. pauciflora* and E. *stellulata* . These are considered to be a highly evolved group in eucalypts. Their mallee form is said to be of recent origin on the evolutionary scale of eucalypts and it is an adaptation to the poor soil and dry

climatic conditions (Ian Brooker personal communication). The other two species (in the present study) included as a close out- group are *E. globulus* and *E. nitens* and they evolved much before the mallee forms. Of the seven species in this study, *E. moorei*, *E. affn. moorei*, *E. dives*, and *E. pauciflora* came from one population each of Nerega region of New South Wales. The *E. mitchelliana* specimens were collected from one plantation of South Australia. *E. stellulata* collections came from one population of Black Mountain of Canberra, ACT. The *E. globulus* and *E. nitens* collections came from Central Victoria- one population each respectively. There were 30 specimens collected for each of the eight species for DNA barcoding from each population. After extracting the DNA from each specimen, the DNA is pooled for a each one.

DNA extraction

About 5-7g of leaves were cut, avoiding the thicker part of the midrib and petiole and any large dead regions. They were immersed in liquid nitrogen. Immediately after liquid N2 evaporates, grind 30-40 seconds in coffee grinder. The powder was resuspended in 40 ml of extraction buffer (recipe below) in small plastic beaker. Polytron (homogenize) for 20-30 seconds at 3/4 speed. This procedure is followed using a standard protocol.

Preparation of eucalypt extraction buffer

For 1 L buffer preparation 100 mM Tris, 100 mM Tris 12.1 g, 25 mM EDTA 50 ml, 0.35 M Sorbitol 64 g, 100 mM Boric acid 6.2 g 1 M

Preparation of eucalypt extraction buffer

For 1 L buffer preparation 100 mM Tris, 100 mM Tris 12.1 g, 25 mM EDTA 50 ml, 0.35 M Sorbitol 64 g, 100 mM Boric acid 6.2 g 1 M NaCl. 58.4 g were weighed and taken in a 2 L beaker. The pH to 8.0 was added before 2% PVP 40,000 20 g was added. Then the following chemicals 10% PEG 8000 100 g, 0.5% BSA 5 g, 0.1 % spermine 1 g, 0.1% spermidine 1 g were added. We prepared the wash buffer with the following chemicals for 1 L. 50 mM Tris50 ml (of 1M, pH8.0), 25 mM EDTA 50 ml (of 0.5 M EDTA, pH 8.0), 0.35 M Sorbitol 64 g. Sodium metabisulphite was added to extraction buffers and wash buffer to 10 mM just prior to use (about 0.2 g per 100 ml – or 10ml per liter of a 1M soln, = 95g/L). PVP 40,000 20 g was added. Then the following chemicals 10% PEG 8000 100 g, 0.5% BSA 5 g, 0.1 % spermine 1 g, 0.1% spermidine 1 g were added.

Primers The universal primers are selected for the genes matK, rbcL, and intergeneric spacers- ITS and trnH-psbA (11). After extracting the DNA individually from all the specimens of each species, the DNA for 30 specimens is pooled together for each is species. The DNA quality is checked using standard protocol before the PCR amplifications are done.

Table 1. Primer sequences used in this study (listed 5'- to -3' end)

psbA- trnH	GTTATGCATGAACGTAATGCTC (1) CGCGCATGGTGGATTCACAATCC (2)

PCR amplification

The non-coding as well as coding regions of matK*, trnH-psbA, ITS and rbcL were* amplified and sequenced by following the protocol usnign the universal. primer pairs (Table 3.1) with TaqF2 (Fisher Biotech, Australia) polymerase.

All PCR amplifications were performed in 20 µl reaction with specific primers annealing temperature. The PCR reactions were consisted of 2 µl of 10xbuffer, 1 µl of 10mM dNTP, 1.6 µl of 25 mM $MgCl_2$, 0.5 µl of each primer (20µM), 1µl pooled genomic DNA (~45ng/µl), 12.9 µl distilled water and 0.5 µl Taq F2 DNA polymerase (5 units/µl; Fisher Biotech, Australia). The amplifications were performed on an ABI thermal cycler (GeneAmp^R PCR System 2700) with initial denaturing at 94°C for 1 min, 35 cycles of 94°C 30 s, primers specific annealing temperature for 30 s and 2-3 min at 72°C followed by a 10 min extension at 72°C.

EgrNAM1 genomic sequences were amplified from randomly selected sixteen trees and both parents DNA using PfuTurbo DNA polymerase (Strata gene, USA). The 100 µl PCR reactions were performed in 10 µl of 10xPCR buffer, 1 µl of 25mM each dNTPs, 2.5 µl of 20 µM each primer mix, 2 µl of 100 ng/µl DNA template, 2 µl of 2.5 U/µl PfuTurbo DNA polymerase (Strata gene, USA) and 80 µl of distilled water. DNA was amplified using the same thermal cycler following same amplification program as used in above mentioned genes with initial denaturing at 95˚C instead of 94˚C. All PCR products were confirmed by gel electrophoresis using a 1% agarose gel and purified by QIAGEN gel extraction kit (Hilden, Germany). Purified PCR product was then quantified by gel electrophoresis comparing with a 100bp gene ladder (Fermentas, Australia).

Cloning of PCR products

All purified DNA amplicons were ligated into the pGEM-T Easy vector using the pGEM-T Easy vector kits (Promega, USA). Ten microliters ligation mixtures contained 5 µl of 2xRapid Ligation Buffer, 1µl of pGEM-Teasy vector (50ng), 3 µl PCR products (25-30ng/µl), 1µl of T4 DNA Ligase (3U/µl) and 1 µl distilled water. A positive control was included for checking transformation and ligation efficiency, and used control insert DNA instead of PCR product as template. A no template negative control was also included. Two microliters of each ligation reaction were transformed into bacterial cells (JM109 and DH5α) by heat-shocking for 50 second at 42˚C water bath, and plated out in LB/ampicillin/IPTG/X-Gal medium, and then incubated at 37˚C for overnight. Twenty-four white colonies were picked and cultured in 5 ml LB/ampicillin medium for overnight at 37˚C and plasmid DNA isolated using the QIAprep Spin Miniprep kit (Hilden, Germany). In the case of *EgrNAM1*, DNA from 18 trees were ligated, transformed and plated out separately following the procedure described above. A single positive colony was picked from each individual and cultured for plasmid DNA preparation. The sizes of all inserts were verified by digestion with *Not*I followed by gel electrophoresis. Each 10 µl digestion reaction was consisted of 1 µl of buffer D, 0.1 µl of BSA, 1 µl of plasmid DNA, 0.20 µl of *Not*I (Invitrogen, USA) and 7.70 µl of distilled water and was incubated at 37˚C in a water bath for 2-3 hours.

Sequencing

Twenty-four different amplicons of each fragment of *EgrHB1*and EgrPAAPA*,* and 18 amplicons of each fragment of *EgrNAM1* were sequenced in both directions using pGEM-Teasy vector Forward (5’ GTAAAACGACGGCCAGT 3’) and Reverse (5’ CAGGAAACAGCTATGAC 3’) primers. Further sequences of large gene fragments were obtained using internal primers. Sequencing was carried out using Big Dye Terminator version 3.1 reagents and an ABI PRISM sequence analyzer using 1/8 reaction volume. Plasmid DNA (0.8 µl) was added to 14.2 µl of distilled water, 1µl of Big Dye version 3.1 mix, 3.5 µl of 5x sequencing buffer and 0.5 µl of each primer (10µM). Cycle sequencing used an initial step at 94˚C for 5 min, then 30 cycles of denaturation at 96˚C for 10 s, annealing at 50˚C for 5 s, and extension at 60˚C for 4 min. The products were precipitated using ethanol, dried down under vacuum and sent to the John Curtin School of Medical Research (JCMSR) for gel separation.

Sequences analysis

Sequences were verified manually and contigs were assembled using the computer software program MEGA version 3.1 (Kumar et al., 2004). Multiple sequence alignments were made using the same program and adjusted manually. All chromatograms and SNPs were visually checked using Sequencer 4.6 (Gene Codes, Corporation, Ann Arbor. Michigan, USA) to exclude any sequencing errors. Philip analysis is done using the Philip version. Reference.

Results and Discussion

There were successful amplifications of DNA for the eight eucalypt species using 4 DNA barcode, trnH-psbA.. The trnH-psbA spacer is the short with 476 bps. The mean length is 733 bps long. Amplification success rate is almost 100%. For each species in each gene there are multiple good quality sequences available for the sequence alignment. They varied 4- 10 good quality sequences for each species in each gene. Looking at the Single Nucleotide Polymorphic (SNPs) sites, the trnH-psbA spacer showing 21 SNPs out of 476 bps segment. The trnH-psbA spacer has 4.41% variability and the rbcL showing only 1.57% variability.

gor1700505glo		A	A	C	C	G	T	T	
'gor170_0515mit'		T	A	G	T	G	C	T	
'gor170_0526nit'		A	A	C	T	A	T	T	
'gor170_0534ste'		A	A	G	T	G	C	T	
'gor170_0695Amo'		A	A	G	T	G	C	T	
'gor170_0706moo'		A	A	G	T	G	C	T	
'gor170_0719pau'		A	A	G	T	G	C	C	
'gor170_0731div'		A	C	G	T	G	C	T	

Figure 2: Variable sites

trnH-psbA spacer shows higher divergence (up to 0.29) The intraspecific divergence values are found to be very negligible (up to 0.008). In the present the study two of the trnH-psbA is able to meet the barcode criteria well by showing greater interspecific variability.

Discussion

Until now there are a few successful findings to give us most suitable DNA barcodes for land plants. These studies helped to focus on the generic level discrimination using barcodes or above this level. One of the prime objectives of this investigation was to compare the cpDNA region of *E. moorei* complex (*E. moorei* and *E. affn. moorei*). The study also included very closely related taxa such as *E. dives, E. stellulata, E. pauciflora* and *E. mitchelliana*. There is a question of the true (disputed) identity of *E. affn. moorei* coming from a disjunct population with diagnostically different morphological features from that of *E. moorei*. The following genes have provided very useful information for the phylogenetic study of eight taxa of eucalypts.

trnH-psbA

It is relatively shorter and the average length is 476 bp. It is the second most variable segment in the present study. The amplification and sequencing are easily done. There are alignment problems because of large number of indels. For example, E. *globulus* sequences contain 28 indels in the short fragment length of 476.

Table 2: Barcode genes' sequence analysis and SNPs comparisons

Gene	No. Bases	No. SNPs	No. of Taxa useful SNPs	% SNPs
trnH-psbA	476	21	7	4.41

gor1700505glo	A	A	C	C	G	T	T	-	-	A	A	A	C	C	G	G	T	T	T	T	T
'gor170_0505glo'	A	A	C	C	G	T	T	-	-	A	A	A	C	C	G	G	T	T	T	T	T
'gor170_0507glo'	A	A	C	C	G	T	T	-	-	A	A	A	A	T	G	G	T	T	T	T	T
'gor170_0509glo'	A	A	C	C	G	T	T	-	-	A	A	A	A	T	G	G	T	T	T	T	T
'gor170_0515mit'	T	A	G	T	G	C	T	T	A	A	A	A	A	T	A	A	T	T	T	T	A
'gor170_0517mit'	T	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
'gor170_0518mit'	T	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
'gor170_0521mit'	T	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
'gor170_0526nit'	A	A	C	T	A	T	T	T	A	A	A	A	A	T	A	G	C	T	T	T	T
'gor170_0528nit'	A	A	C	T	A	T	T	T	A	G	A	A	A	T	G	G	T	T	T	T	T
'gor170_0529nit'	A	A	C	T	A	T	T	T	A	A	A	A	A	T	A	G	T	T	T	T	T

‘gor170_0530nit’	A	A	C	T	A	T	T	T	A	A	A	A	A	T	G	G	T	T	T	T	T
‘gor170_0532nit’	A	A	C	T	A	T	T	T	A	A	A	A	A	T	A	G	T	T	T	A	T
‘gor170_0534ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0535ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	A	G	T	T	T	T	A
‘gor170_0536ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0538ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0541ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	A	G	T	T	T	T	A
‘gor170_0542ste’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	A	G	T	T	T	T	G
‘gor170_0695Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0696Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0698Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	C	T	T	G
‘gor170_0700Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0702Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0703Amo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0706moo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0708moo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0710moo’	A	A	G	T	G	C	T	T	A	A	G	A	A	T	G	G	T	T	T	T	G
‘gor170_0712moo’	A	A	G	T	G	C	T	T	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0714moo’	A	A	G	T	G	C	T	T	G	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0716moo’	A	A	G	T	G	C	T	T	A	A	G	A	A	T	G	G	T	T	T	T	G
‘gor170_0719pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0720pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	A	G	T	T	T	T	A
‘gor170_0722pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	A	G	T	T	T	T	A
‘gor170_0724pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	G	G	T	T	G	T	G
‘gor170_0725pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0727pau’	A	A	G	T	G	C	C	G	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0731div’	A	C	G	T	G	C	T	T	A	A	A	C	A	T	G	G	T	T	T	T	G
‘gor170_0732div’	A	C	G	T	G	C	T	T	A	A	A	C	A	T	G	G	T	T	T	T	G
‘gor170_0733div’	A	C	G	T	G	C	T	T	A	A	A	C	A	T	G	G	T	T	T	T	G
‘gor170_0734div’	A	C	G	T	G	C	T	T	A	A	A	C	A	T	G	G	T	T	T	T	G
‘gor170_0737div’	A	C	G	T	G	C	T	G	A	A	A	A	A	T	G	G	T	T	T	T	G
‘gor170_0739div’	A	C	G	T	G	C	T	T	A	A	A	C	A	T	G	G	T	T	T	T	G

Figure 1: trnH-PsbA spacer sequences

The others have 7-9 indels. The taxonomically useful SNPs have clearly discriminated 5 species, namely, *E. dives, E. globulus, E. mitchelliana, E. nitens* and *E. pauciflora.* The *E. stellulata* has shown 5 single nucleotide substitutions and indicating that 50% chances of recognizing a separate species based on the molecular data. There are also 3 single nucleotide substitutions in the case of *E. moorei* that are different from the *E. affn. moorei*. The 62.5% success rate of interspecific variations using this barcode makes it a good candidate as ear-

lier reported (11) that can be recommended as one of the most useful non-coding regions for cp DNA barcode studies.

Table. 4 Sequence length and percent intra-specific sequence divergence for four plastid regions of 8 eucalypts

Species	Gene
	trnH-psbA length(bp) %divergence
E. affin. *moorei*	447/1 0.002
E. moorei	447/2 0.004
E. dives	447/2 0.004
E. globulus	447/2 0.004
E. mitchelliana	447/0 0.000
E. nitens	447/2 0.004
E. pauciflora	447//0 0.000
E. stellulata	447/2 0.004

We can explain the lack of clear discrimination between *E. moorei and E. affn. moorei* and *E. stellulata* because of strong chloroplast sharing among the closely related eucalypts. As it is reported earlier that hybridization or introgressions are very common among the species of the same subgenus in eucalypts which come from the same geographical regions.

Implications of the present study and future prospects

The present research findings highlight the importance of matK (coding region) and trnH-psbA to a great. Two of them are better suited for low-level taxonomic investigations than other coding and non-coding barcodes so far reported. The matK which has been recommended as an important barcoding gene recently (15). Kress, 2007 (13) proved again its great resolving power at interspecific discrimination. The trnH-psbA is also found to be easily amplifiable with the universal primers and has shown good number of variable sites in the sequence analysis. In spite of large number of indels, this gene has yielded relatively significant PIC values. The trnH-psbA spacers discriminated 5 species clearly from the rest and the other 3 species are also separated from each other though the distance between them is not significantly noticeable.

This present work has focused on a small group consisting of five blue ashes, one Peppermint and two symohyomyrts (eight closely related) Eucalyptus taxa. These are said to be one of the most rapidly evolving group of individuals in the genus *Eucalyptus.*

There are many reports of interspecific hybridizations and introgressions. It is not always easy to distinguish the hybrids because the hybrids share the maternal plastid DNA. Therefore, the choice of this study actually enters into a problem-group area in a sense this might be a good start with a problem group and apply the DNA barcodes for the 8 closely related taxa is challenging. Sometimes we also know the fact that discrimination of some taxa might be lost with greater taxonomic and geographic sampling. This gives us a grasp of things on a minor scale before launching a large-scale study.

Conclusion

In the Phylip tree construction based on Nucleic Acid sequence Maximum likelihood method has given us comparative confidence limits of interspecific divergence of eight closely related Eucalyptus species. Among the eight taxa, there are five species with significant confidence levels. They are as following. *E. dives*

Figure 3: trnH-PsbA - Phylogenetic tree (UPGMA)

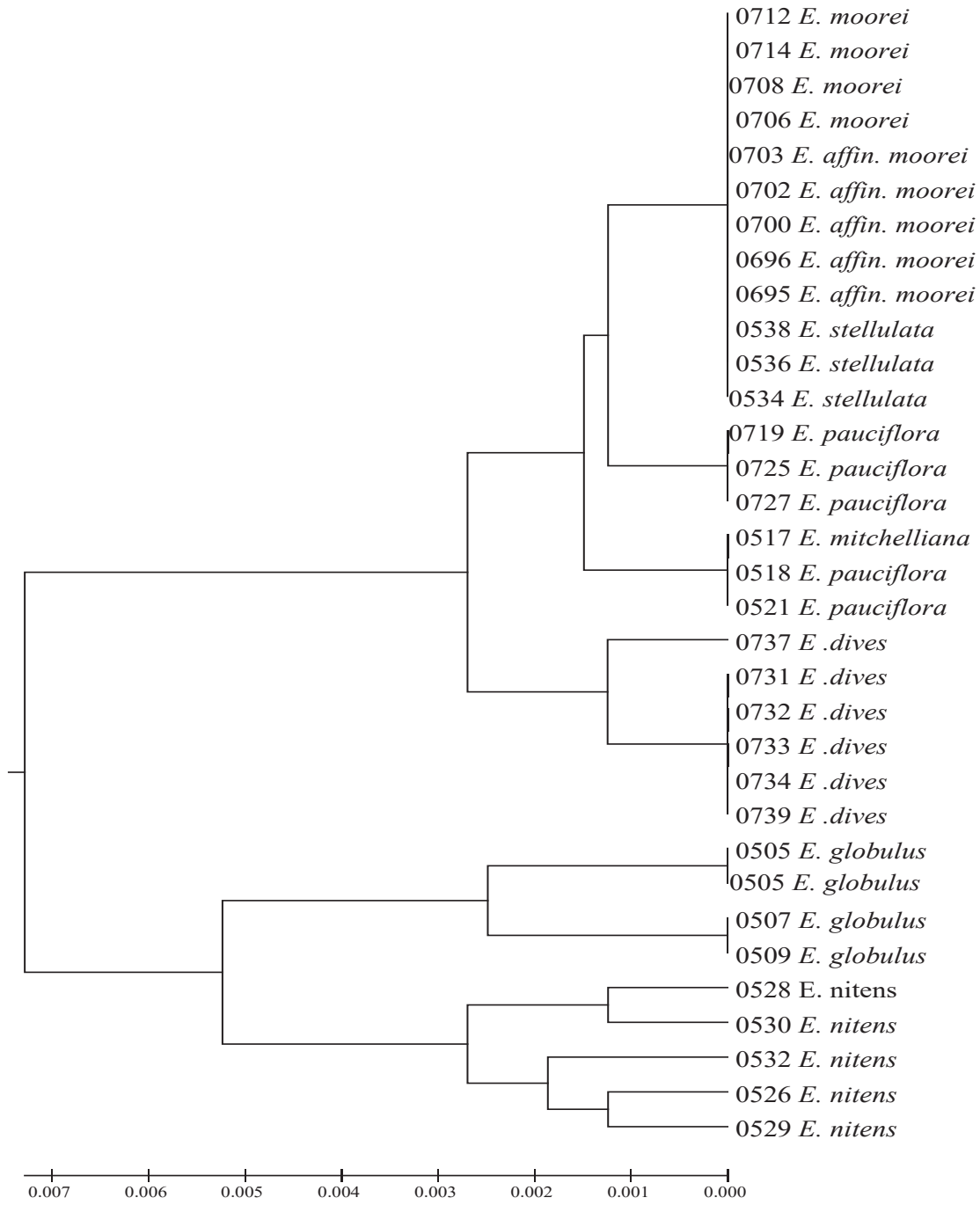

Figure 3: trnH-PsbA - Phylogenetic tree (UPGMA)

0.17 confidence, *E. pauciflora* 0.18 confidence, *E. mitchelliana* 0.37 confidence, *E. globulus* 010 confidence, and *E. nitens* 0.24 confidence. These levels of confidence are positively significant. In other three cases of *E. stellulata*, *E. moorei* and *E. affn. moorei* the confidence limits are insignificant.

Table 4: DNA sequence homology of trnH-PsbA gene (%)

Gene	Species	*E.dives*	*E.globulus*	*E.mitchelliana*	*E.Affin.moorei*	*E.moorei*	*E.nitens*	*E.pauciflora*	*E.stellulata*
	E. dives	**1**							
	E. globulus	**0.99**	**1**						
	E. mitchelliana	**0.99**	**0.98**	**1**					
trnH-psbA	***E. affin. moorei***	**0.99**	**0.98**	**0.99**	**1**				
	E. moorei	**0.99**	**0.98**	**0.99**	**0.99**	**1**			
	E. nitens	**0.98**	**0.98**	**0.98**	**0.98**	**0.98**	**1**		
	E. pauciflora	**0.98**	**0.99**	**0.98**	**0.99**	**0.99**	**0.97**	**1**	
	E. stellulata	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**		**1**

These form a one taxonomic complex. We can also call this an aggregated species based on molecular data. But the taxonomic identity of *E. steullulata* is well defined based on morphological characters. There is an ambiguity about the identity of the disputed E. *aft. moorei*. Our present study using 4 different cp DNA and nrDNA-ITS barcodes have not resolved the issue fully although it has indicated us some haplotypes of *E. moorei*/ *E. affn*. *moorei*. It is recommended (Ian Brooker personal communication, 2008) to make a collection from a population of Blackheath, New South Wales, for this species complex in order to do a comparative study based on morphological characters as well as molecular data. This might either clarify or brings out useful information for identifying *E. affn. moorei* as a different species or consider it still part *E. moorei* species complex. We can also use some low-copy nuclear genes to resolve the identity and position of *E. affn . moorei*.

There are many success stories so far reported using either a single cp DNA barcode or in combination two or three in different individual groups of land plants. But there is no consensus as yet if an individual or a multilogues barcodes that would work very well in plants belonging to different families coming from various geographical locations of the planet. Some taxonomist's view (Chase et al 2003) (9). that DNA barcodes based on uniparentally inherited markers can never reflect the complexity that exists in nature.

There are ambiguities created as far as species limits are concerned by barcodes in some cases. Therefore, some taxonomists have a suspicion or skepticism of barcodes. They are critical of this work. For most taxonomists it is important to have a reasonable barcode based on multiple low copy nuclear DNA loci, a multi-locus barcode system (MBC). This would mean look-

ing for identification of conserved flanking regions containing variable sites. These sites may be introns of appropriate size. These conserved sites would serve as universal PCR priming locations. The reasons for MBC are because of detection of hybridization/ introgression cannot be reliably done by examining a single cp DNA region.

The starting presupposition was to identify all the 8 species of this group using 4 cp DNA barcodes coding and non-coding and expected them to be taxonomically discriminated based on molecular data. The findings are not far from the objectives. Though there is no 100% perfection of the results, these have clarified and illumined the understanding that matK gene and trnH-psbA spacer are better suited for low-level molecular phylogenetic studies in eucalypts. Therefore, it is not a question of mathematical precision of usefulness of data but the value and the significance of the information it has provided us in order to enhance the ongoing search for the most suited barcode regions for plant identification in a wider concern for recording and preserving the biodiversity on our planet .The regions included for this kind of study are trnH-psbA spacer, there are many other non-coding regions of cpDNA investigated but they are not recorded or not explored yet. Because of these reasons we have little information about relative rate of evolution among different non-coding regions. Each research group designed its own experiments to test different barcodes on different group of plants. Some of the works are on a large scale. However, there is no consensus as yet. Apart from various recommendations, some of the latest such as Lahaye et al (2007) (15). which correctly classified 90% of the species by using matK and trnH-psbA (either alone or in combination)? The final agreements seem to be in the direction of using multiple regions than one. The latest CBOL' conference in Taipei proposal is for using matK, trnH-psbA and atpF-H. The present findings are in the direction of above research work.

References

1. Byrne, M., G. F. Moran, et al. (1993). "Restriction map and maternal inheritance of chloroplast DNA in *Eucalyptus nitens*." Journal of Heredity 84(3): 218-220.

2. Byrne, M. and B. Hines (2004). "Phylogeographical analysis of cpDNA variation in *Eucalyptus loxophleba* (*Myrtaceae*)." Australian Journal of Botany 52(4): 459-470.

3. Byrne, M., B. MacDonald, et al. (2002). "Phylogeographical patterns in chloroplast DNA variation within the *Acacia acuminata* (*Leguminosae : Mimosoideae*) complex in Western Australia." Journal of Evolutionary Biology 15(4): 576-587.

4. Chase, M. W., S. Knapp, et al. (2003). "Molecular systematics, GISH and the origin of hybrid taxa in Nicotiana (Solanaceae)." Annals of Botany 92(1): 107-127.

5. Chase, M. W., N. Salamin, et al. (2005). "Land plants and DNA barcodes: short-term and long-term goals." Philosophical Transactions of the Royal Society B-Biological Sciences 360(1462): 1889-1895.

6. Ebach, M. C. and C. Holdrege (2005). "More Taxonomy, Not DNA Barcoding." BioScience 55(10): 823-824.

7. Freeman, J. S., H. D. Jackson, et al. (2001). "Chloroplast DNA phylogeography of Eucalyptus globulus." Australian Journal of Botany 49(5): 585-596.

8. Freeman, J. S., H. D. Jackson, et al. (2001). "Chloroplast DNA phylogeography of Eucalyptus globulus." Australian Journal of Botany 49(5): 585-596.

9. Kress, W. J. and T. Htun (2003). "A second species of Smithatris (*Zingiberaceae*) from Myanmar." Novon 13(1): 68-71.

10. Kress, W. J., V. A. Funk, et al. (2005). "Mapping biological diversity." Plant Conservation: A Natural History Approach: 209-235.

11. Kress, W. J., A. Z. Liu, et al. (2005). "The molecular phylogeny of Alpinia (*Zingiberaceae*): A complex and polyphyletic genus of gingers." American Journal of Botany 92(1): 167-178.

12. Kress, W. J. and G. A. Krupnick (2006). "The future of Floras: new frameworks, new technologies, new uses." Taxon 55(3): 579-580.

13. Kress, W. J., M. F. Newman, et al. (2007). "An analysis of generic circumscriptions in tribe Alpinieae (*Alpiniodeae : Zingiberaceae*)." Gardens' Bulletin (Singapore) 59(Part 1-2): 113-127.

14. Kumar, S., K. Tamura, et al. (2004). "MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment." Briefings in Bioinformatics 5: 150-163.

15. Jackson, H. D., D. A. Steane, et al. (1999). "Chloroplast DNA evidence for reticulate evolution in Eucalyptus (Myrtaceae)." Molecular Ecology 8(5): 739-751.

16. Lahaye, R., J. Klackenberg, et al. (2007). "Phylogenetic relationships between derived Apocynaceae s.l. and within Secamonoideae based on chloroplast sequences." Annals of the Missouri Botanical Garden 94(2): 376-391.

17. McKinnon, G. E., D. A. Steane, et al. (1999). "Incongruence between chloroplast and species phylogenies in Eucalyptus subgenus Monocalyptus (Myrtaceae)." American Journal of Botany 86(7): 1038-1046.

18. McKinnon, G. E., D. A. Steane, et al. (1999). "Incongruence between chloroplast and species phylogenies in Eucalyptus subgenus Monocalyptus (Myrtaceae)." American Journal of Botany 86(7): 1038-1046.

19. Newmaster SG, F. A., Ragupathy S (2006). "DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach." Canadian Journal of Botany-Revue Canadienne De Botanique 84: 335-341

20. Schaal, B. A., D. A. Hayworth, et al. (1998). "Phylogeographic studies in plants: problems and prospects." Molecular Ecology 7(4): 465-474.

21. Shaw, J., E. B. Lickey, et al. (2005). "The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis." Am. J. Bot. 92(1): 142-166.

22. Schindel, D. E. and S. E. Miller (2005). "DNA barcoding a useful tool for taxonomists." Nature 435(7038): 17-17.

23. Steane, D. A. and D. J. Mabberley (1998). "Rotheca (Lamiaceae) revived." Novon 8(2): 204-206.

24. Taberlet, P., E. Coissac, et al. (2007). "Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding." Nucleic Acids Research 35(3).

25. Tautz, M. M. a. D. (2005). "Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences." Philos Trans R Soc Lond B Biol Sc 360: 1462.