Biotechnology

E-ISSN No : 2581-6195 | Volume : 7 | Issue : 2 | Feb 2024

A NOTE ON ITS AND DNA BARCODING -ITS IMPORTANCE IN SPECIES PHYLOGENETICS

G. Jayaraj

ABSTRACT

ITS sequence data in the present research has provided some useful insights resolving phylogenetic relationship among Eucalyptus species. It is to be complimented with rbcL, MatK and trnH-psbA gene study. There are 77 SNPs in an average in ITS fragment of 746 bps. Therefore, it contains highest percentage of variable characters in the gene. This study has shown a good distance between E. mitchelliana and the rest of species. The position of E. globulus and E. nitens is not clear. There is an overlap about E.dives and E.pauciflora. They are mixed up in E. stellulata, E. moorei and affin moorei are not separated by this gene. ITS, though more variable, it is found to be not so helpful for species identification in the Eucalyptus species.

KEYWORDS: ITS, SNPs, variable gene, phylogenetics, complementary

Centre For Research and Dept. of Botany and Biotechnology, Andhra Loyola College (Autonomous), Vijayawada, Andhra Pradesh

HOW TO CITE THIS ARTICLE:
G. Jayaraj (2024).A Note On Its And Dna Barcoding -Its Importance In Species Phylogenetics, International Educational Journal of Science and Engineering (IEJSE), Vol: 7, Issue: 2, 09-20

INTRODUCTION

In the last few decades, the Genomic research has undergone tremendous changes and there have been several newer technical advances created to better 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, or Popular genomes. There are several other researches which taking place. It includes the *Eucalyptus* that has the sequencing of the entire genome. It has a wider application. Yasi information form genetic helps us with a Yasi good resolution of species boundaries, eventually it may give insights into the patterns and rates of evolutionary diversification among species.

We have seen great explorations on the genetic diversity of different plant species at the phenotypic level for the past century. We can observe a variation of genetic material within the species but what is so astonishing is the genetic variations at the level of conserved regions among distantantly related species. 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 (CBOL, http: barcoding.si.edu). Our objective of this project is to obtain DNA barcodes from all species all over the world, from different geographical and climatic regions. Till date we have only two million species are formally described and an estimated 10-100 million more are yet to be explored and to be identified. Though there are great efforts mad by professional taxonomists to collect and preserve, still there a great deal of information about the biological diversity surrounding us is inaccessible. We may lose it eternally even before we come to know them. Our objectives objectives of DNA barcoding include facilitating species discoveries based on cluster analysis, Species identification and promoting development of handheld DNA sequencing technology. We have no standard protocol still in for DNA barcoding in land plants. We have low levels of variation in the mitochondrial DNA of land plants as compared to animals where a portion of the mitochondrial gene *Cox1 (Co1)* is used successfully. Ther e is a standard barcode for animals. Therefore, we look into the plastid DNA for as a suitable region in barcoding. We see there is very slow rate of evolution in plastid genome in plants. Since it shows very low levels of variation in the plastid DNA, therefore, there is need for looking at more than one region for variations. Our challenge is also to find a very suitable region which shows enough variation within it to discriminate among species yet conserved enough to be present.

ITS gene and barcoding:

ITS is taken from the internal transcribed spacer portion of ribosomal DNA or *nrITS*. We use it for a long time for plant barcoding. (Balldwin, 1992; Chase et al., 2005; Kress et al., 2005). We find from here the results which are very positive

Research Paper

while using the ITS spacer. The ITS is often said to have 3-4 times more variable sites that evolve up more rapidly. The *nrITS* are also subject to gene conversions/concerted evolution (Wendel et al., 1995; Chase et al., 2003 in most organisms. In a process brings out a single copy type that is maintained amidst other variant copies. The occurrence of multiple and divergent copies makes ITS region suitable to be accepted one or a standard barcoding region. However, in some land plants multiple copies are maintained and also in some angiosperms several divergent and still functional copies are routinely detected. (Rapini et al., (2007).

ITS and Eucalyptus

We have seen a wide use of *ITS* spacer for molecular data analysis in order to understand the phylogenetics of Eucalyptus species. Steane et al., (1999) reported that the ITS sequence data are useful for resolving phylogenetic relationship among the Eucalyptus but its resolving power at and below the series level is low. The research findings of Steane et al., (2002), the *ITS* sequence data provided some useful insights. This has reported by the same team about the phylogenetic relationship between sections and subgenera of Eucalyptus and related genera. It suggests that *ITS* sequence data to be complimented with *rbcL, matK* and *trnH-psbA*. When we combine of molecular data and morphological data, it would help to understand the levels of taxonomic relationship.

CRITIQUING OF BARCODING:

Those who promote DNA barcoding say that this technology would revitalize biological collections and speed up species identification and inventories. Till now there are only 1.7 million specimens that have been identified by taxonomists and about 10-20 million more are not named or explained.

Those who oppose DNA barcoding argue that this technology would destroy the traditional systematics and turn it into a service industry (Ebach and Holdrege 2005). Their fears are removed when all cases DNA barcoding is applied only in conjunction with classical approach where species are simply unknown or no attempts have been made to delimit them. Therefore, barcoding intended originally it would be limited in its applicability. Its barcoding addresses 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 (Schindler and Miller 2005). Apart from being a diagnostic tool, barcode sequences per se and their ever-increasing taxonomic coverage, it could lead 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 (Chase et al., 2005 Kress et al 2005). 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' (Tautz et al., (2005). By 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.

There are available evidences show us that many species in eucalypts evolved during the last few hundred thousand years. This gave rise to the evolution of large group of very closely related taxa. It is reported that hybridization is relatively common phenomenon within different subgenera of the genus (Griffin, Burgess and Wolf, 1988). It is much debated about to what extent the hybridization is taking place. (Ladiges 1997, Potts and Wiltshire 1997). We place the basis of species divergence among the eucalypts as a model of simple speciation divergence through vicariance and allopatric speciation. (in the absence of reticulate evolution –genetic exchange between branches in phylogeny). There are breeding barriers to the formation of F1 generation or further generations of hybrids coming from spatial or ecological background (Griffin et al.1988). The barriers are also created by different flowering times (Potts and Reid 1985a; Davidson e t al., 1987). The other barriers are coming from pollen-pistil interactions (Ellis et al., 1991). It is also due to poor F1 hybrid vigor as reported by Lopez et al., (1988). From their extensive work on hybridization in eucalypts (Griffin et al. 1988) reported that among the 528 species, there is a 55% of individuals that were found to be involved in at least one natural hybrid combination. The evidences from recent hybridization and putative introgressed populations are commonly reported in eucalypts (Shaw et al. 1984; Potts and Reid 1985b; Sampson et al. 1995).

It is reported that the Chloroplasts are maternally inherited in most Angiosperms and for the eucalypts too this has been demonstrated by Byrne et al., (1993) in *E. nitens* of eastern Australia. Schael et al., (1999) showed a uniparental inheritance of the chloroplast genome and lack of recombination. There are 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.

From the 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 scapes of Australia over the thousands of centuries. (Byrne and Moran., 1994; Steane et al. 1998; Jackson et al. 1999; McKinnan et al. 1999; Byrne and Macdonald 2000; Freeman et al. 2002; Byrne and Hinds 2003).

The RFLP finding on subgenus *Symphyomyrtus* are suggestive of a sharing of cpDNA haplotypes among species within series (McKinnan et al. 1999; Steane et al. 1998 and Jackson et al. 1999). To a considerable extent of intraspecific polymorphism is also reported. The explanations are a lesson on the consistent pattern of cpDNA haplotypes that can be correlated with geographical circumscribed regions rather than with morphological species. Steane et al (1998) demonstrated that the existence of wide spread cpDNA lineage between *E. dalrympleana, E. urnigera and E. gunnii* of central Tasmania and *E. cypellocarpa* of mainland Australia. Jackson et al (1999) discovered at least three cpDNA lineages in *E. globulus*. The lineages were shared extensively with other species. The creation of different haplotypes is brought out through

gene flow between taxa by an incomplete lineage sharing of ancestral polymorphism (sorting). Doyle (1998) has reported that it is often difficult to distinguish the lineage sharing since the pattern of incongruence is rather similar. The geographical structuring of cpDNA haplotypes across Tasmania have been identified (Steane et al. 1998; Jackson et al. 1999; McKinnan et al. 1999; McKinnan et al. 2001; and Freeman et al. 2001). These findings from these studies have been used as evidence for explanation of wide hybridization and introgression which brings about evolution.

Byrne and Hines (2003) did a phylogeographic analysis in *E. loxophleba* which is endemic to south-western Australia. From the similar phytogeographical patterns' evidence in the study is an explanation of a shared response to the influences of historical processes. Our conclusions from the above findings lead to a hypothetical formulation of 'species-specific' hybridizations and introgressions and they were major factors in distribution of chloroplast variations in *E. loxophleba*. Our studies have brought forth an understanding on the complexity of the issue of evolution history and the phylogenetic relationships of the genus *Eucalyptus*.

Ladiges and Bayly (2006; Bayly et al.2008; Ocheing et al.2007) discovered ITS pseudogenes in the ribosomal DNA of eucalypts. The paralogues that are reported from the this study are found to be highly differentiated when compared to the functional gene sequences in closely related taxa. Since earlier studies (Steane et al. 1999; 2002) used extensively ITS sequences data for phylogenetic analysis in eucalypts, it has been cautioned by the same researchers about the anomalous or controversial ITS phylogenetic constructions in eucalypts when compared with the data coming from morphology or ESTs of rDNA or the microsatellites. However, it is suggested that rDNA paralogues' use might provide us useful information for variations in phylogenetic studies and for other investigations.

MATERIALS AND METHODS :

The finding of suitable regions of the genome.

The DNA barcoding mechanism is a diagnostic tool for species identification- using a short, standard DNA region. Our challenge is to find a suitable genomic region for a wide range of taxa in plants. The sequence variation must be high enough between species so that they can be discriminated from one another and it must be low enough within species that a clear threshold between intra- and inter specific genetic variation can be defined. The Mitochondrial cytochrome C oxidase gene (Co1) segment is used in many animal groups for barcoding. It is said to have been showing necessary universality and variability and more than 95% unambiguous identification in most cases of major animal clades where the studies are done. However, many of the mitochondrial genes are not proved suitable enough for barcoding in plants because of their low mutation rates. This is also the factor of rapidly changing structure of mitochondrial genome in plants. The DNA barcoding has great scientific applications in ecology and evolution. Because of this, an universal barcode is being explored by many scientists and it is yet to be agreed upon. Kress et al. (2005) proposed originally a trnH-psbA spacer of plastid region as suitable universal barcode for land plants. Chase et al. (2005) has suggested rpocl, rpoB and matK or rpocl, matK and trnH-psbA as a good barcoding region. Taberlet (2006) put forward the trnL intron as suitable plant barcode. Kress and Erickson (2007) proposed a combination of trnH-psbA with rbcL for plant barcoding (combining from the proposal of Newmaster et al. 2006). Lahaye et al (2007) identified a portion of matK gene as a universal DNA barcode for flowering plants. Considering the said reasons from the above proposal the present study uses a multi locus region – ***matK, trnH-psbA, ITS*** and ***rbcL*** of the genome for the barcoding of 8 eucalypt species.

The *mat*K gene was formerly called *orf*K. It is 1500 bps in length is a maturase coding gene. Its position is can be found within the intron of the chloroplast gene *trn*K. These genes are found at the upstream of the *psb*A gene 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 (Shaw et al., 2007). In some group of plants it is relatively short, having less than 300 bps (Kress et al., 2006). The ITS is an internal transcribed spacer region of nuclear ribosomal DNA. It is highly variable in Angiosperms. The ITS region of the nuclear ribosomal Cistron (18S-5.8S-26S) is most commonly sequenced locus for plant molecular systematics. Two smaller fragments namely ITS1 and ITS2 adjoining the 5.8S locus can be amplified. The *rbcL* is a ribulose-1,5-bisphosphate carboxylase/oxygenase gene.The chloroplast gene rbcL encodes the large subunit of ribulose bisphosphate carboxylase in plants

The Taxa Selection for Barcoding.

The group containing *E. moorei, E.affn. moorei, E. dives, E. mitchelliana, E. pauciflora* and *E. stellulata* is considered 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 eight 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 each species.

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 resuspend in 40 ml of extraction buffer (recipe below) in small plastic beaker. Polytron (homogenise) for 20-30 seconds at 3/4 speed. I(t was then filtered through two layers of muslin

(only one layer if it is the very fine muslin) into tube in ice. The tubes were shaken vigorously and then spun at 2000 rpm (750 x g) for 10 minutes. The supernatant was poured off. The tubes were inverted on paper towel to drain for more than one minute. About 5 ml of wash buffer was added and the tubes were resuspend by vortexing (or with paint brush for species with gummy pellets). About 3 ml of 20% SDS, 20% Triton X-100 were added. They were mixed gently and kept at room temperature for at least five minutes, with occasional gentle mixing by inversion. About 4.0 ml of NaCl/CTAB mix. (34.4g CTAB + 162g NaCl per liter) were added. The tubes were incubated at 65°C for 20-30 minutes, with occasional mixing. chloroform (24:1) to 45 ml was added to the extract the tubes were inverted for at least 2 minutes. (Or placed on shaker for 20-60 minutes. When the aqueous phase retained the green color, it was extracted for longer). It was spun at 5000 rpm (5000 x g) for 10 minutes. The upper aqueous layer we transferred to a new tube. (If interphase layer is solid enough, the aqueous layer can be poured off – otherwise use a plastic disposable pipette.) About 2/3 volume 100% isopropanol was added. The was hooked (spool) out or spun down if unhook able. It was wash in 1.5 ml of 50% isopropanol/0.3M NH4OAc for several hours or overnight. They were air dried to precipitate for 15-20 minutes. They were resuspend in a minimal volume of TE (pH 8.0), (usually 200-500 µl).

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 20g 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 sol, = 95g/L).

PVP 40,000 20g 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 son, = 95g/L).

The Primers:

The primers are selected are universal for the genes matK, rbcL, and intergeneric spacers- ITS and trnH-psbA. (CBOL universal primers. Kress et al., 2005). (Table 1). 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.

ITS	ATGTCACCACAAACAGAAAC (1) TCGCATGTACCTGCAGTAGC (2)
ITS	ATGTCACCACAAACAGAAAC (1) TCGCATGTACCTGCAGTAGC (2)

Table 1

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 using 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 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 (Stratagene, 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 (Stratagene, USA) and 80 µl of distilled water. The 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). The 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). The 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. There is a positive control was included for checking transformation and ligation efficiency, and used control insert DNA instead of PCR product as template. There is no template negative control was also included. Two microliters of each ligation reaction was 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. We have included 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. We also included a single positive colony was picked from each individual and

cultured for plasmid DNA preparation. These 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 NotI (Invitrogen, USA) and 7.70 µl of distilled water and was incubated at 37°C in a water bath for 2-3 hours.

Sequencing

There are 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. Moreover sequences of large gene fragments were obtained using internal primers. Sequencing was carried out using BigDye 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 BigDye 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. These 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

These sequences were verified manually and contigs were assembled using the computer software program MEGA version 3.1 (Kumar et al., 2004). We made a multiple sequence alignments were made using the same program and adjusted manually. All our chromatograms and SNPs were visually checked using Sequencher 4.6 (Gene Codes, Corporation, Ann Arbor. Michigan, USA) to exclude any sequencing errors. Phylip analysis is done using the Phylip version..reference..

RESULTS:

We found out that there were successful amplifications of DNA for the eight eucalypt species using 4 DNA barcode genes namely, matK, ITS, trnH-psbA and rbcL. These sequence length obtained is quite long with good number of variable sites. From among the genes matK is the longest with 855bp (base pairs) followed by ITS spacer which is 746 bps long. The rbcL (gene amplified) has 744 bps. The trnH-psbA spacer is the shortest with 476 bps. The mean length is 733 bps long. Amplification success rate is almost 100%. In the each species in each gene there are multiple good quality sequences available for the sequence alignment. It varied 4- 10 good quality sequences for each species in each gene. While looking at the Single Nucleotide Polymorphic (SNPs) sites, the ITS gene showed the highest number of variable sites. The number SNPs recorded for ITS are 77 from 746 bp sequence. The next gene which showed most variable site is matK with 37 SNPs among the 713 bps. After this is followed by the trnH-psbA spacer showing 21 SNPs out of 476 bps segment. The rbcL is showing least number of variable sites with only 11 SNPs recorded in a sequence of 700 bps.

Among the comparisons of Taxonomically useful SNPs (SNPs that might potentially distinguish taxa) for phylogenetic studies, the ITS gene is found to be having largest number variable sites in the present study. It shows up 17 taxonomically useful SNPs. It is the ITS is found to be most variable with 12.89% variations. It is followed by the matK with 5.1% variations. The trnH-psbA spacer has 4.41% variability and the rbcL showing only 1.57% variability.By comparison ITS has much higher interspecific divergence value (up to 0.4) than the others..

DISCUSSION:

So far there are a few successful findings to give us most suitable DNA barcodes for land plants. Our study 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 coda region of *E. moorei complex (E. moorei and E. affn. moorei)*. This 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*. We find out that the following genes have provided very useful information for the phylogenetic study of eight taxa of eucalypts.

ITS: This is found to be with highest number of variable sites. We can see there are 77 SNPs in an average fragment of 746bps. It the fore it contains highest percentage of variable characters in the gene. None the less when it comes to the interspecific discrimination the interspecific resolution power is only 50%. In spite of this gene is successfully used earlier for phylogenetic studies at higher levels, like Series or Subgenera level , its usefulness for lower level has not been demonstrated (Steane 1998). The above sequence data failed to separate species because of the large number of informative characters in taxa resulted in large number of character-state changes leading to high homoplasy. The above informative characters present in the data have very low levels of confidence to separate the species. None the less it has been cautioned about the ITS paralogues (pseudo genes) by the previous studies. We have always the question of ITS's subjection to conversion/ concerted evolution (Wendel et al. 1995; Chase et al. 2003). We see there is a presence of multiple and divergent copies of ITS. This phylogenetic tree in the present case has shown also a big difference from the other tree genes. This has put a good distance between *E. mitchelliana* and the rest which is not in consonance with the taxonomic construction. It also created confusion about the position of *E. globulus* and *E. nitens*. The other individuals like *E. dives, E. pauciflora* are mixed up and there is an overlap created between these two. In the rest, *E. stellulata, E. moorei* and *E. affn. moorei* are also not separated by this gene. From this we can say ITS though most variable, it is not helpful for species identification in the present study. This confirms the some earlier reports of its' unsuitability for barcoding especially at the interspecific divergence.

Implications of the present study and future prospects:

In the present research findings highlights the importance of ITS to some extent and for DNA barcoding in closely related eucalypts. We find that two of them are better suited for low-

level taxonomic investigations than other coding and non-coding barcodes so far reported. of indels, this gene has yielded relatively significant PIC values. Among this, 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. The next useful one is ITS.

Our work has focused on a small group consisting of five Blue ashes, one Peppermint and two symohyomyrts (eight closely related) *Eucalyptus taxa*. These species are said to be one of the most rapidly evolving group of individuals in the genus *Eucalyptus*. We have many reports of interspecific hybridizations and introgressions. This is not always easy to distinguish the hybrids because the hybrids share the maternal plastid DNA. Our choice of study therefore 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. It gives us a grasp of things on a minor scale before launching a large scale study. . It could be further tested by including many individuals from the same subgenus to test the utility of this barcode. Moreover a study like this one on a small group (8 taxa) can be useful to explain their taxonomic relationships in detail . This may be a good tool for predicting the overall levels of variations that are likely to be found in a large scale.

CONCLUSIONS

In the present 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. From among the eight taxa, there are five species with significant confidence levels. They are as following. *E. dives* 0.17 confidence, *E. pauciflora* 0.18 confidence, *E. mitchelliana* 0.37 confidence, *E. globulus* 010 confidence, and *E. nitens* 0.24 confidence. From what we have got, 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. All these form a one taxonomic complex. We could 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. affn. moorei*. In 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 has been 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 .It 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. Some times we can also use some low-copy nuclear genes to resolve the identity and position of *E. aft . moorei.*

We have seen 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. There is no consensus as yet if an individual or a multilocus barcodes that would work very well in plants belonging to different families coming from various geographical locations of the planet. We see that some taxonomist's view (Chase et al 2003) that DNA barcodes based on uniparentally inherited markers can never reflect the complexity that exists in nature.

There are ambiguities created at the species level concerned by barcodes in some cases. There some taxonomists have a suspicion or skepticism of barcodes. They are critical of this word. It is for most taxonomists important to have a reasonable barcode based on multiple low copy nuclear DNA loci, a multi-locus barcode system (MBC). It would mean looking 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.

Our 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. Our findings are not far from the objectives. Even though there is no 100% perfection of the results, 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

From 1995 there has been a great deal of increase in using the noncoding cp DNA and the coding matK data for molecular systematic investigations. The regions TS and rbcL. There are many other non-coding regions of cpDNA investigated but they are not recorded or not explored yet. For 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 our works are on a large scale. However there is no consensus as yet. From among various recommendations, some of the latest such as Lahaye et al (2007), which correctly classified 90% of the species by using matK and trnH-psbA (either alone or in combination)? In the end the final agreements seem to be in the direction of using multiple regions than one.. The present findings are in the direction of above research work.

REFERENCE

1. Baldwin, B. G. (1993). "MOLECULAR PHYLOGENETICS OF CALYCADENIA (COMPOSITAE) BASED ON ITS SEQUENCES OF NUCLEAR RIBOSOMAL DNA - CHROMOSOMAL AND MORPHOLOGICAL EVOLUTION REEXAMINED." American Journal of Botany 80(2): 222-238.
2. Baldwin, B. G. (1999). "New combinations and new genera in the North American tarweeds (Compositae-Madiinae)." Novon 9(4): 462-471.
3. Bayly, M. J., F. Udovicic, et al. (2008). "Ribosomal DNA

pseudogenes are widespread in the eucalypt group (Myrtaceae): implications for phylogenetic analysis." Cladistics 24(2): 131-146.
4. Bayly, M. J., F. Udovicic, et al. (2008). "Ribosomal DNA pseudogenes are widespread in the eucalypt group (Myrtaceae): implications for phylogenetic analysis." Cladistics 24(2): 131-146.
5. Byrne, M. and G. F. Moran (1994). "Population divergence in the chloroplast genome of Eucalyptus Nitens." Heredity 73: 18-28.
6. Byrne, M., G. F. Moran, et al. (1994). "Detection and inheritance of RFLPs in Eucalyptus nitens." Theoretical and Applied Genetics 89(4): 397-402.
7. 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.
8. Byrne, M., G. F. Moran, et al. (1993). "Restriction Map and Maternal Inheritance of Chloroplast DNA in Eucalyptus nitens." J Hered 84(3): 218-220.
9. Chase, M. W. (2005). "Classification of Orchidaceae in the age of DNA data." Curtis's Botanical Magazine 22(Part 1): 2-7.
10. Chase, M. W. (2005). "Relationships between the families of flowering plants." Plant diversity and evolution: genotypic and phenotypic variation in higher plants: 7-23.
11. 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.
12. Doyle, J. J., M. Morgante, et al. (1998). "Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (Glycine subgenus Glycine)." Molecular Biology and Evolution 15(2): 215-218.
13. Ebach, M. C. and C. Holdrege (2005). "More Taxonomy, Not DNA Barcoding." BioScience 55(10): 823-824.
14. Ellis, M. F., M. Sedgley, et al. (1991). "Interspecific Pollen-Pistil Interaction in Eucalyptus L'Her. (Myrtaceae): The Effect of Taxonomic Distance." Ann Bot 68(3): 185-194.
15. Freeman, J. S., H. D. Jackson, et al. (2001). "Chloroplast DNA phylogeography of Eucalyptus globulus." Australian Journal of Botany 49(5): 585-596.
16. Goldman, D. H., J. V. Freudenstein, et al. (2001). "Phylogenetics of Arethuseae (Orchidaceae) Based on Plastid matK and rbcL Sequences." Systematic Botany 26(3): 670-695.
17. Griffin, A. R., I. P. Burgess, et al. (1988). "PATTERNS OF NATURAL AND MANIPULATED HYBRIDIZATION IN THE GENUS EUCALYPTUS LHERIT - A REVIEW." Australian Journal of Botany 36(1): 41-66.
18. Griffin, A. R., I. P. Burgess, et al. (1988). "PATTERNS OF NATURAL AND MANIPULATED HYBRIDIZATION IN THE GENUS EUCALYPTUS LHERIT - A REVIEW." Australian Journal of Botany 36(1): 41-66.
19. Hebert, I. D. H. a. P. D. N. (2004). "Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes." Canadian Journal of Zoology 82(5): 749-754.
20. Hines, B. and M. Byrne (2001). "Genetic differentiation between mallee and tree forms in the Eucalyptus loxophleba complex." Heredity 87: 566-572.
21. Hilu, K. W., T. Borsch, et al. (2003). "Angiosperm phylogeny based on matK sequence information." Am. J. Bot. 90(12): 1758-1776.
22. Hilu, K. W., T. Borsch, et al. (2003). "Angiosperm phylogeny based on matK sequence information." Am. J. Bot. 90(12): 1758-1776.
23. Kumar, S., K. Tamura, et al. (2004). "MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment." Briefings in Bioinformatics 5: 150-163.
24. Ladiges, P. and F. Udovicic (2005). "Evolution of southern hemisphere Myrtaceae: molecular and morphological evidence." Plant Systematics and Evolution 251(1): 1-1.
25. Ladiges, P. Y., S. R. Ariati, et al. (2006). "Biogeography of the Acacia victoriae, pyrifolia and murrayana species groups in arid Australia." Journal of Arid Environments 66(3): 462-476.
26. 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.
27. 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.
28. Jackson, H. D., D. A. Steane, et al. (1999). "Chloroplast DNA evidence for reticulate evolution in Eucalyptus (Myrtaceae)." Molecular Ecology 8(5): 739-751.
29. Jackson, H. D., D. A. Steane, et al. (1999). "Chloroplast DNA evidence for reticulate evolution in Eucalyptus (Myrtaceae)." Molecular Ecology 8(5): 739-751.
30. Kress, W. J., V. A. Funk, et al. (2005). "Mapping biological diversity." Plant Conservation: A Natural History Approach: 209-235.
31. 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.
32. 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.
33. 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.
34. 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.
35. McKinnon, G. E., R. E. Vaillancourt, et al. (2001). "Chloroplast sharing in the Tasmanian eucalypts." Evolution Int J Org Evolution 55(4): 703-11.
36. McKinnon, G. E., R. E. Vaillancourt, et al. (2001). "Chloroplast sharing in the Tasmanian eucalypts." Evolution 55(4): 703-711.
37. McKinnon, G. E., R. E. Vaillancourt, et al. (2001). "Chloroplast sharing in the Tasmanian eucalypts." Evolution 55(4): 703-711..
38. Miller, A. J. and B. A. Schaal (2006). "Domestication and the distribution of genetic variation
39. Newmaster, S. G., A. J. Fazekas, et al. (2006). "DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach." Canadian Journal of Botany-Revue Canadienne De Botanique 84(3): 335-341.
40. 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.
41. Newmaster SG, F. A., Steeves RAD, Janovec J (2008). "Testing candidate plant barcode regions in the Myristicaceae." Molecular Ecology Resources 8: 480-490.
42. Ochieng, J. W., D. A. Steane, et al. (2007). "Microsatellites retain phylogenetic signals across genera in eucalypts (Myrtaceae)." Genetics and Molecular Biology 30: 1125-1134.
43. Potts, B. M. and J. B. Reid (1985). "ANALYSIS OF A HYBRID SWARM BETWEEN EUCALYPTUS-RISDONII HOOK F AND EUCALYPTUS-AMYGDALINA LABILL." Australian

Journal of Botany 33(5): 543-562.
44. Rapini, A., C. van den Berg, et al. (2007). "Diversification of Asclepiadoideae (Apocynaceae) in the New World." Annals of the Missouri Botanical Garden 94(2): 407-422.
45. Salazar, G. A., M. W. Chase, et al. (2003). "Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): evidence from plastid and nuclear DNA sequences." Am. J. Bot. 90(5): 777-795.
46. Schaal, B. A., D. A. Hayworth, et al. (1998). "Phylogeographic studies in plants: problems and prospects." Molecular Ecology 7(4): 465-474.
47. Schindler, T., R. Bergfeld, et al. (1995). "Arabinogalactan Proteins in Maize Coleoptiles - Developmental Relationship to Cell-Death During Xylem Differentiation but Not to Extension Growth." Plant Journal 7(1): 25-36.
48. Shaw, M. J., B. M. Potts, et al. (1984). "VARIATION WITHIN AND BETWEEN EUCALYPTUS-NITIDA HOOK F AND EUCALYPTUS-COCCIFERA HOOK F." Australian Journal of Botany 32(6): 641-654.
49. Shaw, W. H., A. Muirhead, et al. (2001). "A glutathione-dependent formaldehyde dehydrogenase gene of symbiotic cyanobacterium Anabaena azollae." Symbiosis 30(1): 41-55.
50. Steane, D. A., M. Byrne, et al. (1998). "Chloroplast DNA polymorphism signals complex interspecific interactions in Eucalyptus (Myrtaceae)." Australian Systematic Botany 11(1): 25-40.
51. Steane, D. A., M. Byrne, et al. (1998). "Chloroplast DNA polymorphism signals complex interspecific interactions in Eucalyptus (Myrtaceae)." Australian Systematic Botany 11(1): 25-40.
52. Steane, D. A., D. Nicolle, et al. (2002). "Higher-level relationships among the eucalypts are resolved by ITS-sequence data." Australian Systematic Botany 15(1): 49-62.
53. Steane, D. A., D. Nicolle, et al. (2002). "Higher-level relationships among the eucalypts are resolved by ITS-sequence data." Australian Systematic Botany 15(1): 49-62.
54. Steane, D. A., D. Nicolle, et al. (2002). "Higher-level relationships among the eucalypts are resolved by ITS-sequence data." Australian Systematic Botany 15(1): 49-62.
55. Steane, D. A., D. Nicolle, et al. (2002). "Higher-level relationships among the eucalypts are resolved by ITS-sequence data." Australian Systematic Botany 15(1): 49-62.
56. Steane, D. A., D. Nicolle, et al. (2002). "Higher-level relationships among the eucalypts are resolved by ITS-sequence data." Australian Systematic Botany 15(1): 49-62.
57. 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
58. Wendel, J. F., A. Schnabel, et al. (1995). "AN UNUSUAL RIBOSOMAL DNA-SEQUENCE FROM GOSSYPIUM GOSSYPIOIDES REVEALS ANCIENT, CRYPTIC, INTERGENOMIC INTROGRESSION." Molecular Phylogenetics and Evolution 4(3): 298-313.
59. Wiltshire, R. J. E., B. M. Potts, et al. (1998). "Genetic control of reproductive and vegetative phase change in the Eucalyptus risdonii E-tenuiramis complex." Australian Journal of Botany 46(1): 45-63.
60. Wolf, P. G., R. A. Murray, et al. (1997). "Species-independent, geographical structuring of chloroplast DNA haplotypes in a montane herb Ipomopsis (Polemoniaceae)." Molecular Ecology 6(3): 283-291.
61. Wolf, P. G., E. Sheffield, et al. (1991). "ESTIMATES OF GENE FLOW, GENETIC SUBSTRUCTURE AND POPULATION HETEROGENEITY IN BRACKEN (PTERIDIUM-AQUILINUM)." Biological Journal of the Linnean Society 42(4): 407-423.

ITS																					
E.dives	G	A	A	A	A	G	A	G	C	C	C	G	G	G	C	C	T	C	A		
E.globulus	G	G	G	G	G	G	G	G	T	G	A	A	A	C	C	C	C	C	C		
E.mitchelliana	A	A	A	A	A	A	A	C	C	C	C	G	G	G	T	T	T	T	A		
E.affin.moorei	G	A	A	A	A	G	A	G	C	C	C	G	G	G	C	C	T	C	A		
E.moorei	G	A	A	A	A	G	A	G	C	C	C	G	G	G	C	C	T	C	A		
E.nitens	G	G	G	G	A	G	G	G	T	G	A	A	A	C	C	C	C	C	C		
E.pauciflora	G	A	A	A	A	G	A	G	C	C	C	G	G	G	C	C	T	T	A		
E.stellulata	G	A	A	A	A	G	A	G	C	C	C	G	G	G	C	C	T	C	A		

Figure 1: ITS Variable sites

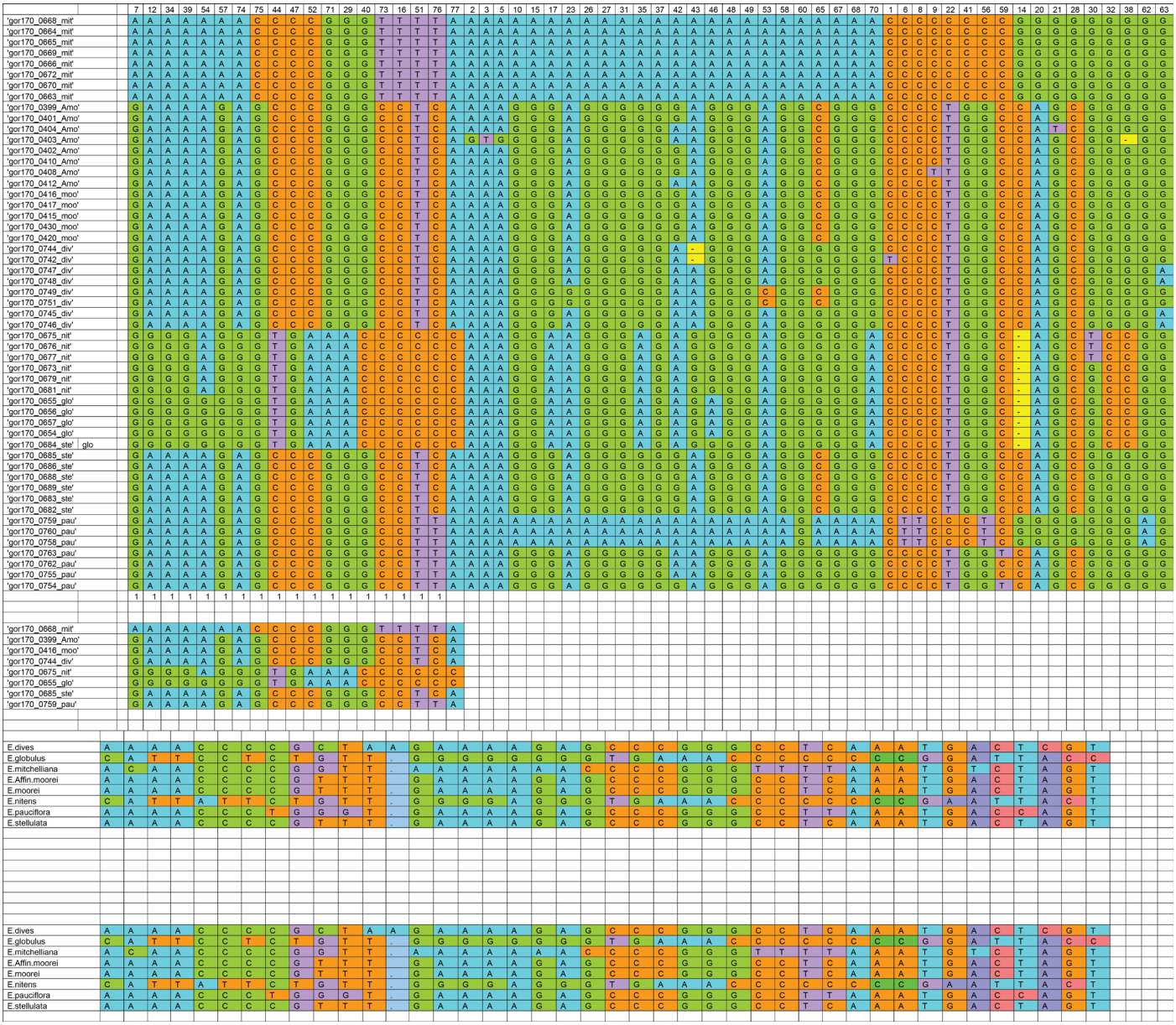

Figure 2: DNA Sequences showing variable nucleotides

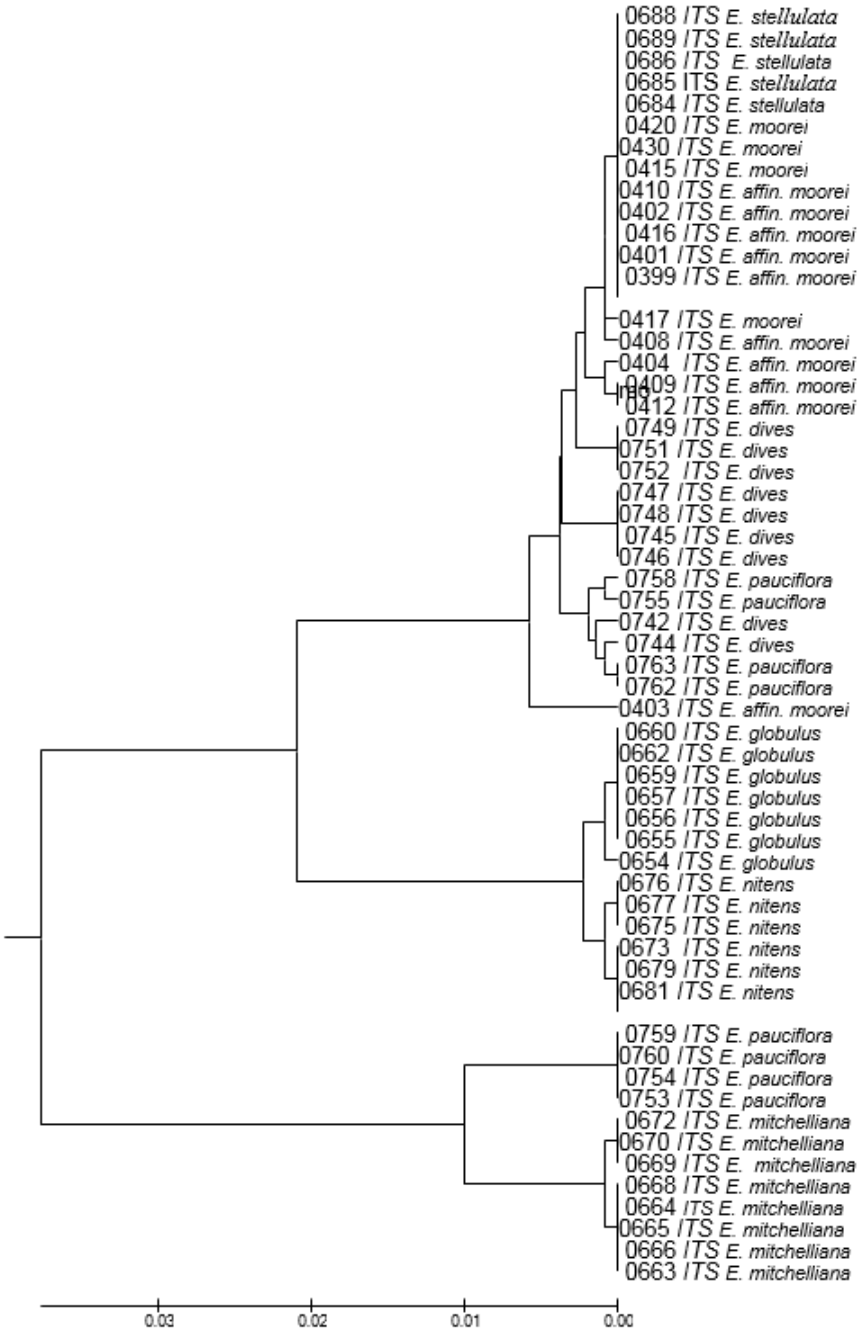

Figure 4 The Phylogenetic tree

ITS	ATGTCACCACAAACAGAAAC (1) TCGCATGTACCTGCAGTAGC (2)
ITS	ATGTCACCACAAACAGAAAC (1) TCGCATGTACCTGCAGTAGC (2)

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

GENE length 597bp	SPE-CIES	E.dives	E.globu-lus	E.mitch-elliana	E.affn. moorei	E.moorei	E.nitens	E.pauci-flora	E.stellu-lata
ITS	E. dives		0.064	0.088	0.051	0.026	0.056	0.085	0.055
	E. glob-ulus	0.064		0.098	0.078	0.055	0.016	0.107	0.053
	E. mitch-elliana	0.088	.0.098		0.105	0.078	0.095	0.087	0.107
	E. affin. moorei	0.051	0.078	0.105		0.40	0.073	0.100	0.154
	E. moorei	0.026	0.055	0.078	0.040		0.050	0.080	0.045
	E. nitens	0.056	0.016	0.095	0.073	0.050		0.102	0.050
	E. pauci-flora	0.085	0.107	0.087	0.100	0.080	0.102		0.092
	E. stellu-lata	0.055	0.053	.0.107	0.154	0.045	.0.050	0.092	

Table. 2 Sequence length and percent interspecific sequence divergence for four plastid regions of 2 Subgenera of 8 eucalypts.

GENE	No. Bases	No. SNPs	No.Tax.useful SNPs	% SNPs
ITS	597	77	19	12.89

Table 3. Barcode genes' sequence analysis and SNPs comparisions

SPECIES	GENE
	ITS length (bp) % divergence
E. affin. moorei	597/15 0.0251
E. moorei	597/5 0.008
E. dives	597/7 0.011
E. globulus	597/6 0.010
E. mitchelliana	597/7 0.011
E. nitens	597/5 0.008
E .pauciflora	597/7 0.011
E. stellulata	597/6 0.010

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