Evolution, systematics and taxonomy of *Elaeocarpus* (Elaeocarpaceae) in Australasia

Thesis submitted by

Yumiko Baba

BSc (Hons) Tokyo University of Agriculture

MSc The University of Edinburgh

in July 2013

for the degree of Doctor of Philosophy

Australian Tropical Herbarium

and the School of Marine and Tropical Biology

James Cook University

Cairns, Australia

STATEMENT ON THE CONTRIBUTION OF OTHERS

The chapters of this thesis are also manuscripts that have been published or are in preparation for submission. Several researchers have made contributions to these manuscripts as follow:

Chapter 1: P. Bannink produced a distribution map.

Chapter 3: M. Rossetto supplied primer sequences of microsatellite markers utilised in this chapter and technical guidance for genotyping and analyses. C.D. Kilgour provided assistance in collecting the samples. P. Bannink produced a distribution map. C. Micheneau assisted in the data analysis and provided theoretical background. Technical support for running the statistical program R was provided by A. Lechner.

Chapter 4: C.D. Kilgour provided assistance in collecting the samples.

Chapter 5: This chapter has been published as Baba, Y. and Crayn, D. (2012) *Elaeocarpus hylobroma* (Elaeocarpaceae): a new species endemic to mountain tops in north-east Queensland, Australia. Kew Bulletin, 67, 1–8. D. Crayn and P. Bostock provided technical and theoretical background. Assistance for production of line drawing was gained from W. Cooper. P. Bannink produced a distribution map. C.D. Kilgour supplied a picture for a figure. M. Coode contributed for the improvement of manuscript. Permission for reproduction of the article as a chapter in this thesis was obtained from Springer with a license number 3165350580769.

Herbarium Loans were obtained for all the chapters from Queensland Herbarium (BRI). Land access and permissions for the scientific collection were granted from Council of Northern Peninsula Area, a chairman of Lama Lama National Park (CYPAL), Kuku Yalanji People, other National Park rangers, Australian Wildlife Conservancy Picaninny Plain, the owners of Bromwell Station and Mackay Botanic Garden. Scientific collections were made under WISP05391710 and WITK04740210, issued by the Queensland Department of Science, Information Technology, Innovation and the Arts.

I was supported financially by an Australian Biological Resources Study (ABRS) scholarship, which funded me the second half of this study, came from the ABRS Grant No. 211-42 to D.M. Crayn, S.-N. Phoon and Y. Baba, titled Systematic studies in the Polystachyus-XIB group complex of *Elaeocarpus* L.. The Skyrail Rainforest Foundation provided two consecutive funding rounds to carry out Chapters 3, 4 and 5. The Australian Tropical Herbarium (ATH) made available significant in-kind and logistical support for this project.

STATEMENT ON THE PUBLICATION OF NEW TAXONOMIC NAMES

Under Article 30.7 (reproduced below) of the International Code of Botanical Nomenclature, *Melbourne Code* (McNeill et al. 2012), taxonomic names proposed in this thesis are not considered effectively published.

> *Article 30.7.* Publication on or after 1 January 1953 of an independent non-serial work stated to be a thesis submitted to a university or other institute of education for the purpose of obtaining a degree does not constitute effective publication unless the work includes an explicit statement (referring to the requirements of the *Code* for effective publication) or other internal evidence that it is regarded as an effective publication by its author or publisher.

ACKNOWLEDGMENTS

Special thanks are due to my principal supervisor, Professor Darren M. Crayn for the support, training and encouragement he gave me throughout my candidature. I also express my gratitude to co-supervisors, Professor Paul A. Gadek and Dr. Mark Harrington. I could not have wished for more enthusiastic, patient and inspiring supervisors.

I gratefully acknowledge the grant bodies, my parents and Australian Biological Resource Study for the generous financial support and for the academic tuition fees and living stipend; Skyrail Rainforest Foundation for research funding; Australian Systematic Botany Society and Australian Biological Resources Study for the support to attend conferences. My candidature would have been impossible without the support of these people and organizations.

My colleagues at the Australian Tropical Herbarium in Cairns are some of the most wonderful people to have shared the journey of my candidature with. Many thanks are due to Dr. Caroline Puente Lelievre, Kaylene Bransgrove, Lalita Simpson, Sook-Ngoh Phoon, Agustina Arobaya, Gary Wilson, the co-occupants of the post graduate room; Drs. Claire Micheneau, Katharina Schulte, Ashley Field for sharing their extensive knowledge of systematics, patient guidance and discussions; Andrea Lim, Frank Zich and the team of ATH for logistic and moral support. These people provided me with a great deal of personal and professional support.

These people are acknowledged for providing technical assistance in the various subjects, in field and sharing their invaluable knowledge of NE Queensland: Wendy and William Cooper, Rigel Jensen, Andrew Ford, Mark Coode, Stuart Worboys, Dr. Fanie Venter, Dr. David Rentz, Melissa Harrison, Dr. Soraya Gaze, Dr. Alex Lechner, Peter Bannink, Dr. Tapio Linderhaus, Charlie and Lewis Roberts, Dr. Craig Costion, Tony Roberts, Bruce Gray, Fay Adams and Irene Champion.

Many thanks to the following people for criticising chapter drafts: Dr. Claire Micheneau, Gary Wilson, Dr. Chris Quinn, Cameron Kilgour, Kaylene Bransgrove, Dr. David Rentz, Dr. Fanie Venter. I really appreciated the constructive comments and advice.

Lastly and most importantly, I would like to express my greatest gratitude to my family Cameron Kilgour, Kayoko and Eiji Baba, and friends I acquired over the years, especially Dr. Caroline Puente Lelievre, Kaylene Bransgrove and Barbara Forster. We went through some

long and hard times together but they were most significant and fun time in my life, and cannot thank you enough for sharing it with me from the beginning until, the very end.

GENERAL ABSTRACT

An important role of taxonomy is to document the biological world by discovering and determining the relationships of, and formally describing, organisms. This not only enhances our knowledge of biodiversity, but also provides fundamental information for other disciplines. The aim was to increase the understanding of the hierarchical relationships within the rainforest tree genus *Elaeocarpus* as the basis for a better understanding of the evolutionary processes that have given rise to the extant diversity. An hierarchical approach using different methods and datasets for different taxonomic questions was employed: phylogenetics, population genetics, morphometrics and traditional descriptive taxonomy.

The phylogenetic relationships of Elaeocarpaceae were investigated using nucleotide sequences of two plastid intergenic spacers, *trnL-trnF* and *trnV-ndhC*, and the nuclear encoded Internal Transcribed Spacer region. Maximum parsimony and Bayesian analyses of the combined plastid and nuclear data with enhanced taxon sampling produced a more detailed estimate of relationships within Elaeocarpaceae than previous studies. Monophyly of all the genera of Elaeocarpaceae except *Elaeocarpus*, *Aceratium* and *Sericolea* was confirmed. *Elaeocarpus*, *Aceratium* and *Sericolea* formed a strongly supported clade in the multigene tree in the Bayesian analysis but the determination of taxonomic rank for each group requires further investigation.

Some morphological groups such as Group V Subgroup A, Group VI Subgroup B and part of Group XI subgroup B proposed by Coode (1984) were each strongly supported as monophyletic based on the separate marker data sets as well as the combined data set. Additionally, samples of the *Elaeocarpus obovatus* species complex (Group V D + *E. coorangooloo*) formed a monophyletic group strongly supported in most analyses. Some clades showed correlation with geography. These areas are New Caledonia-Pacific and Asia.

The phylogenetic study provided an evolutionary framework within which to place the undescribed taxa in Australia. The undescribed taxa sampled for this study now have their positions in their respective groups confirmed: *E.* sp. Mt Misery was nested in Group VI B; *E.* Mt. Windsor Tableland was placed in Group XI Subgroup B. Further investigation of these entities was beyond the scope of this study, partly because insufficient material was available for a thorough analysis.

With the aim to solve long-standing problems of taxonomic delimitation within the *E. obovatus* species complex (Group V D + *E. coorangooloo*), genetic variation, diversity and relatedness were assessed using a population genetics approach with established microsatellite markers. While there needs a validation from other data source to confirm microsatellite profiles suggested that *E. arnhemicus* and *E. obovatus* may be tetraploids. Because the appropriate methodology for analysing polyploid and diploid species together in population genetics is yet to be standardised, a synthesis of three different approaches was utilised in this study: similarity based analysis (PCoA), model based analysis (STRUCTURE), hypothesis testing (Analysis of Molecular Variance (AMOVA) using Φ_{PT} and Multigroup Discriminant Function Analysis (MDFA)).

Elaeocarpus arnhemicus was supported as an entity distinct from the other two groups by the cluster analyses, AMOVA (Φ_{PT}= 0.43) and MDFA. All of the *E. obovatus* populations are weakly supported as a single entity by the majority of the clustering methods, and this group is strongly supported as distinct from *E.* sp. Mt Bellenden Ker based on AMOVA (Φ_{PT}= 0.31) and MDFA. Although some differentiation was found between *E. obovatus* North and South populations in the STRUCTURE analysis, it was decided that as a working hypothesis *E. obovatus* should be regarded as a single genetic entity because there exists a large sampling gap between them. STRUCTURE analysis detected some genetic admixture between *E. arnhemicus* and *E. obovatus*, *E.* sp. Mt Bellenden Ker and *E. coorangooloo*, and between *E. obovatus* and *E.* sp. Mt. Bellenden Ker. AMOVA indicated only up to 23 % of the variation was shared between each pair, with the likely explanation being retained ancestral polymorphism in both cases. Taken together the results suggest that there are three distinct genetic groups corresponding to *E. arnhemicus, E. obovatus*, and *E.* sp. Mt. Bellenden Ker.

Morphological variation within the *E. obovatus* complex was evaluated against the working hypothesis (the existence of three entities) that resulted from the population genetic study. The results of PCA, PCoA, Cluster analysis, Multigroup Discriminant Function Analysis and Classification Tree analysis revealed that *E. obovatus*, *E. arnhemicus*, and *E.* sp. Mt. Bellenden Ker are morphologically discrete on the basis of fruit and vegetative characters. Taken together, the results of both the genetic and morphological analyses indicate that recognition of *E.* sp. Bellenden Ker at species rank is justified. Hence the two named species, *E. obovatus* and *E. arnhemicus*, are maintained and *E.* sp. Mt. Bellenden Ker is newly described as *E. biracemosus* Y.Baba & Crayn. A dichotomous key to all entities and fully revised accounts of *E. arnhemicus*, *E. coorangooloo* and *E. obovatus* are provided.

A detailed investigation of the long-standing putatively recognised taxon *E.* sp. Mossman Bluff (D.G.Fell 1666) was undertaken and resulted in the description of a new species from the Australian Wet Tropics. The taxon was formally named *E. hylobroma* Y.Baba & Crayn and a full description and a line drawing of the species was produced. The position of this species as sister to the morphologically distinct Group V was strongly supported by Bayesian analysis of the combined sequence data plus indel data. Since the broader relationships of this clade are unclear, the species was tentatively assigned to Group V with the subgroup assignment suspended until more evidence becomes available.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1 General introduction

An important role of taxonomy is to document the biological world by discovering and determining the relationships of, and formally describing, organisms. This not only enhances our knowledge of biodiversity, but also provides fundamental information for other disciplines. Taxonomic information (names and taxon concepts) is a basis of communication in all of biology.

The primary purpose of this project is to assess the hierarchical relationships within the rainforest tree genus *Elaeocarpus* L. as the basis for 1) a taxonomy which better reflects the evolutionary relationships, and 2) a better understanding of the evolutionary processes that have given rise to the extant diversity. This study draws on morphological, molecular, biogeographical and field survey data. The specific aims are to:

- Investigate phylogenetic relationships within *Elaeocarpus* as a basis for evaluation of the existing classification and assessment of putatively new Australian taxa;
- Address species-level problems (species complexes and putative new species) in Australian *Elaeocarpus* and describe new taxa where appropriate.

1.1 Overview of the chapters

Chapter 1 provides a brief introduction to the genus *Elaeocarpus* L. in the context of the Australian flora.

Chapters 2 to 4 present the results of molecular and morphological studies that aimed to test taxonomic hypotheses at generic, infrageneric, specific and infraspecific ranks. In Chapter 2, molecular phylogenetic analyses are presented with the aim of understanding the evolutionary relationships within the family and the genus. The results are used to evaluate the existing infrageneric classification and resolve the broad relationships of putatively new Australian taxa as a basis for in depth assessment of their limits and status. In chapters 3 and 4, morphological and molecular datasets were analysed to test species boundaries within the *E. obovatus* species complex.

In chapter 5, a traditional taxonomic approach (intuitive assessment of morphological characters) was used to describe a new Australian species of *Elaeocarpus*, and based on the results of the phylogenetic analysis provisionally assign it to an infrageneric grouping.
Finally, a summary and conclusions are presented in Chapter 6.

1.2 Ecological context: climatic history of the rainforests of the eastern coast of Australia

The current distribution of Australian tropical and subtropical rainforests has been shaped by climatic and geological history (Moritz et al., 1997). Rainforest was once widespread and intact across the continent, but the development of drier, cooler and more seasonal climates subsequent to the separation of Australia from Antarctica *ca*. 35 million years ago (MYA) caused the contraction of rainforest into a 'mesotherm archipelago' (Moritz et al., 1997; Nix, 1991) of fragmented patches distributed along the east coast (Nix, 1991). Major units within this archipelago have been defined including the 'Wet Tropics (15-19 degrees S)', 'Central Queensland (20-22 degrees S)' and 'Southern Queensland (27-32 degrees S, including Northern New South Wales)' (Webb & Tracy, 1981).

The most recent climatic oscillation that significantly influenced the contemporary distributions of rainforests in Australia occurred in the late Quaternary (*sensu* VanDerWal et al., 2009). The impact of this oscillation on rainforests was different in each region. In the Wet Tropics Bioregion (Thackway & Cresswell, 1995), the rainforest retreated to upland refugia, and the total area of survival was substantial, whereas in the south, the total area of refugia was hypothesised to be much more restricted (Moritz et al., 1997). In the Wet Tropics Bioregion, the locality of many refugia has been identified (Graham et al., 2010; Graham et al., 2006; VanDerWal et al., 2009), and recognitions of vicariance events are well advanced (Joseph & Moritz, 1993; 1994).

Rossetto and others (2009) pointed out the importance of consideration of macroecology in population genetics, to understand spatial and temporal distributional patterns. For example, in one of their studies (2007) it was highlighted that the patterns of genetic diversity (*He*) of *Elaeocarpus grandis* F.Muell., distributed from northern New South Wales (NSW) to the tip of Cape York Peninsula along the east coast, was higher in populations in Wet Tropics than in populations in Northern NSW, which suggests that the level of divergence was partially reinforced by the rainforest contractions during the glacial period. The origin of *Elaeocarpus* was hypothesised to be at least 30 million years ago (MYA), which implies that the diversification of the genus occurred during the unstable phase of rainforest development in

Australia. There may be some more examples of genetic consequences derived from the evolutionary process expected in Australian *Elaeocarpus*.

1.3 Elaeocarpaceae Juss.

The Elaeocarpaceae is primarily a tropical to sub-tropical family with some genera extending into the temperate zone. The family is known from all continents except North America and Africa. It is one of six families in the order Oxalidales Heintze, the others being Brunelliaceae Engler, Cephalotaceae Dumortier, Connaraceae R. Brown, Cunoniaceae R. Brown, Oxalidaceae R. Brown and Huaceae A. Chevalier (APG III, 2009) The Elaeocarpaceae comprises 12 genera and c. 550 species of trees and shrubs (Coode, 2004). Australia harbours the greatest genus diversity with nine genera represented. Its absence from North America, Africa and Europe has led to the suggestion that the family is of southern origin (Raven & Axelrod, 1974).

Recent molecular genetic studies have shed the light into some aspects of the evolutionary history of the Elaeocarpaceae that could not be discovered by traditional morphological studies alone. For example, the Elaeocarpaceae was formally placed in Malvales Berchtold & J. Presl (Cronquist 1981; Takhtajan 1997) and has been recognised as a part of Tiliaceae Juss. (Bentham & Hooker, 1862) due to superficial similarities in floral structure. Molecular studies, however, strongly support the placement of the Elaeocarpaceae in Oxalidales, distant from the Tiliaceae (Malvales) (Bradford and Barnes 2001; Crayn *et al.* 2006) and inclusion of the Tremandraceae R.Br. ex DC. in Elaeocarpaceae. The monophyly of all traditionally recognised genera was supported with moderate to high bootstrap values, except for *Elaeocarpus*, which appeared to be paraphyletic. The recently described species, *E. sedentarius*, was not nested within a clade with the other species of *Elaeocarpus* included in the analysis (Maynard, 2004; Crayn *et al.* 2006). This apparent paraphyly, however, was based on a limited sample set and was only weakly supported in some analyses, therefore, further molecular work is required before a full grasp of the phylogenetic relationships of *Elaeocarpus* can be achieved.

1.4 *Elaeocarpus* L.

Elaeocarpus is the largest (approx. 350 spp.) of the 12 genera recognised in the family. The others are *Aceratium* DC. (approx. 20 spp.), *Aristotelia* L'Hér (approx. 5 spp.), *Crinodendron* Molina (approx. 5 spp.), *Dubouzetia* Pancher ex Brongn. & Griseb (12 spp.), *Peripentadenia* L.S.Sm. (2 spp.), *Platytheca* Steetz (2 spp.), *Sericolea* Schltr. (approx. 16 spp.), *Sloanea* L. (approx. 150 spp.), *Tetratheca* Sm. (approx. 40 spp.), *Tremandra* R.Br. ex DC. (2 spp.), and *Vallea* Mutis ex L. (2 spp.) (Coode, 2004). *Elaeocarpus* is predominantly distributed in the Old

World tropics and subtropics from Madagascar eastward to Hawaii, and from Japan southward to New Zealand (Figure 1.1), with its highest diversity in Borneo and New Guinea (Zmarzty 2001). In Australia, *Elaeocarpus* is an important element in mesothermal and megathermal rainforests. For instance, the genus is a recognisable feature of c. 30% of rainforest in the Wet Tropics Bioregion in Regional Ecosystems in Queensland, where the genus is most diverse (Queensland Herbarium, 2013).

The ecology of *Elaeocarpus*, including dispersal mechanisms and pollination vectors, is largely unknown. Limited studies have shown a diversity of pollination vectors: nocturnal insects (Weber, 1994), honey bees (Roubik, 2005), small parrots (Brown & Hopkins, 1995), honeyeater birds (Castro & Robertson, 1997) and insects (moths, beetles and flies) (Devy & Davidar, 2003). In only a few cases has the dispersal vectors for species of *Elaeocarpus* been determined: frugivorous birds (Clout & Hay, 1986), bats (Corlett, 1990; Nakamoto et al., 2009), cassowaries (Stocker & Irvine, 1983), and bush rats (Maynard et al., 2008).

The cytology of *Elaeocarpus* is relatively understudied. Of the studies that have been undertaken, most counted chromosomes of somatic cells, reporting 2n = 28 in *E. photiniaefolius* Hook.et Arn (Ono, 1975), 2n = 30 in *E. hookerianus* (Allan, 1961; Rattenbury, 1947; de Lange et al., 2004), *E. "lancaefolius"* (probably *E. lanceifolius* Roxb.) (Mehara & Sareen, 1973) and *E. speciosus* Brongn. & Gris (Carr & McPherson, 1986), and 2n = 32 in *E. angustifolius* Blume (Gamage & Schmidt, 2009). A few studies have used Metaphase II chromosomes in pollen grains, suggesting base chromosome numbers for *E. speciosus* and *E. "lancaefolius"* (probably *E. lanceifolius* Roxb.) of x = 15 (Mehara & Sareen, 1973; Carr & McPherson, 1986). The only known case of polyploidy in Elaeocarpaceae is *Dubouzetia elegans*, with n = c. 90 (probably dodecaploid on x = 15; Carr & McPherson, 1986).

Much remains to be investigated regarding the taxonomy and systematics of the species within the genus. Although revisions at the local level have been published (Backer and Bakhuizen van den Brink, 1963; Coode, 1978; Coode, 1984; Ridley, 1922; Smith, 1981; Tang and Phengklai, 2007; Zmarzty, 2001), the lack of a recent monograph of the genus means that there is considerable uncertainty in regard to the number and circumscription of taxa. For a large genus such as *Elaeocarpus* with rich morphological diversity, a reliable infrageneric classification is required to provide the framework for meaningful comparisons. Furthermore because morphological characters are often prone to excessive homoplasy, morphological character sets should be tested against independent datasets such as DNA sequences.

The only revision of the Australian and New Zealand species of *Elaeocarpus* was published over 30 years ago (Coode, 1984), and since then several new species have been described, bringing the number of recognised Australian taxa (species and sub-species) to 27 (Maynard 2004). Even though the revision was comprehensive, at least five putative new species have been reported from public herbaria in Australia since then. These are still awaiting in-depth study and description.

Furthermore, in Coode's revision, closely related taxa with highly variable morphologies associated with differences in habitat and ecology were documented (Coode, 1984). One example is the *E. obovatus* complex, which comprises *E. obovatus*, *E. arnhemicus*, an undescribed entity known informally as *Elaeocarpus*. sp. Mt. Bellenden Ker (L.J. Brass 18336), and five variants and intermediates. This diversity has not been systematically studied to date thus this taxonomic confusion remains unresolved. Moreover, genetic structure and relationships within the group remain unknown.

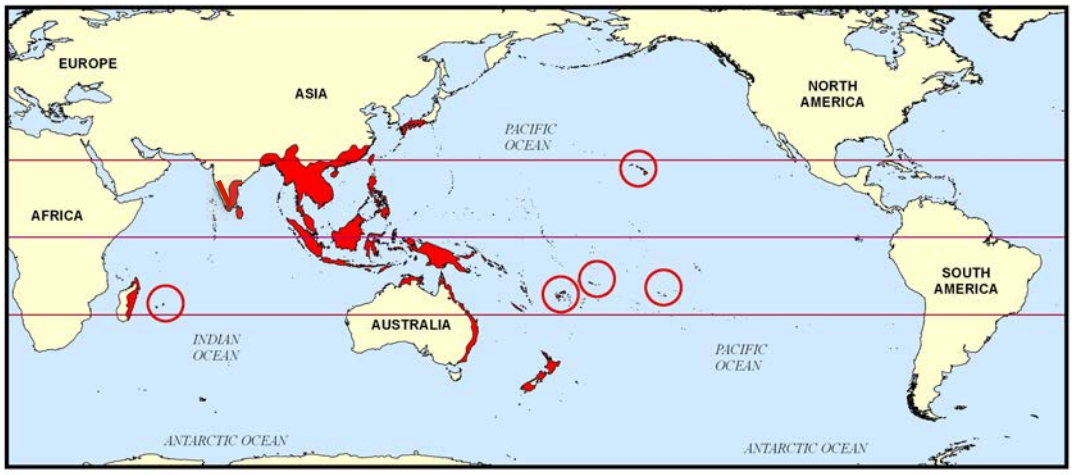


Figure 1.1. Worldwide distribution of *Elaeocarpus*. The circles indicate small islands in the Pacific and Indian Oceans (Hawaii, Fiji, Tonga, Samoa and Rarotonga in the Pacific Ocean and Mauritius in the Indian Ocean) where *Elaeocarpus* is distributed.

1.5 Approaches and molecular and morphological tools

1.5.1.1 Species Concepts

As a science, systematics is hypothesis driven. In order to erect taxa such as species, the criteria by which these should be defined - the species concepts - must be clear. Species concepts have been the subject of ongoing philosophical debate. Judd et al. (2007) summarised the numerous

species concepts in the literature and reduced those relevant to plants to seven. They are the biological (Mayr, 1942; 1992), recognition (Paterson, 1985), phenetic (Sokal & Crovello, 1970), evolutionary (Wiley, 1978), autapomorphy (Donoghue, 1985; Mishler & Brandon, 1987), diagnosability (Nixon & Wheeler, 1990) and genealogical (Baum & Donoghue, 1995) species concepts. The focus of both the biological and recognition concepts is gene flow but the difference between them lies in the role of gene flow in species circumscription. In the biological species concept, gene flow is interpreted as a diversifying force, creating genetic discontinuities between 'species', whereas in the recognition species concept, it is a cohesive force, maintaining similarity between individuals within a species (Judd et al., 2007). Gaps in morphological variation are important in the phenetic species concept (Sokal & Crovello, 1970). Unique combinations of character states found in the smallest aggregation of populations or lineages will circumscribe species in the diagnosability species concept (Nixon & Wheeler, 1990). The evolutionary species concept recognises unique evolutionary lineages (Wiley, 1978). Basal exclusivity is emphasised by the genealogical species concept, in which species are defined by gene coalescence between individuals (Baum & Donoghue, 1995). The autapomorphy species concept defines a species as a clade containing all the descendants of one ancestral population, identifiable by one or more autapomorphies (Donoghue, 1985; Mishler & Brandon, 1987).

However, biology lacks a universally accepted species concept, and there is no consensus on which of the above concepts to apply (Judd et al. 2007). All of the species concepts have in common the acknowledgement that species are separately evolving metapopulation lineages: the underlying common property of species and the only necessary property of species. De Queiroz (2007), however, argues that biologists have long confused the philosophical question of what is the nature of species with the practical question of how we recognise that speciation has occurred and that species concepts should simply be seen as operational criteria by which species may be recognised (De Queiroz, 2007). The choice of which criterion or criteria to use will depend on the data available and the stage of the speciation process observed in the study group.

In the present study, species are defined as groups which are clearly separated on both morphometric and population genetic data.

1.6 Data and analysis

Use of cladistic, biogeographical and molecular information combined with traditional taxonomy can result in a much more robust and reliable classification (Bickford et al., 2007;

Diniz-filho et al. 2008). I have utilised several techniques (see 2. 2, 3.2 and 4.2; materials and methods in the data chapters) to arrive at a hierachical classification for the Australian *Elaeocarpus* species.

1.6.1 Phylogenetic markers

While phylogenetic inference can be undertaken with various data sources, molecular and morphological data are the most sought after in plant phylogenetic research.

Of the three genomes (chloroplast, nuclear and mitochondria) in plant cells, the nuclear and chloroplast (hereafter plastid) genomes are most commonly used in plant systematics. Mitochondrial markers are seldom used because they undergo frequent structural rearrangements and often these mutations are not characterised in the species under study and therefore are not suited to inferring interspecific relationships (Judd et al., 2007). Each gene region accumulates mutations at different rates (Judd et al., 2007). For instance, plastid DNA coding regions are relatively conserved whereas noncoding regions evolve faster, presumably due to there being fewer functional constraints (Small et al., 2005). When the mutation rate of the marker is too low more data must be accumulated, whereas if the rate is too fast the signal will be lost because parallelisms and reversals will accumulate simultaneously (Judd et al., 2007), and therefore choosing markers appropriate to the type of phylogenetic enquiry is pivotal.

1.6.1.1 Nuclear DNA

The nuclear genome (nDNA) is the largest genome (Judd et al., 2007; Lowe et al., 2004). A frequently used part of the nDNA genome for phylogenetic analysis is nuclear ribosomal DNA (nrDNA) (Alvarez & Wendel, 2003). This region comprises three coding regions (18S, 5.8S and 26S rDNA) and the non-coding spacer regions (Intergenic spacer (IGS), External transcribed spacer (ETS) and Internal transcribed spacer (ITS)) separating them. The nrDNA comprises several hundreds to several thousands of tandemly aligned copies of this gene cassette (Judd et al., 2008). Because of this ubiquitous existence of copies, and semi-universal primers, this region has been popular for phylogenetic reconstruction (Alvarez & Wendel, 2003).

Concerted evolution - the parallel homogenization of mutations in the repetitive regions - is one of the unique characteristics of this part of the nuclear genome (Alvarez & Wendel, 2003; Elder & Turner, 1995; Feliner & Rossello, 2007; Hills & Dixon, 1991; Schlötterer & Tautz, 1994). A disadvantage of this phenomenon is that often homogenization is not complete, which results in

the presence of polymorphic loci and pseudogenes when multiple mutations occur in a short period of time, such as with rapid speciation. The high levels of variation within the non-coding parts of nrDNA is, nonetheless, advantageous for phylogenetic studies even in population level studies of genetic diversity (Besnard et al., 2007; Butcher et al., 2007; Judd et al., 2007). The ITS has been used in previous phylogenetic studies of Elaeocarpaceae (Maynard, 2004; Crayn, 2006) but as yet has not provided a satisfactory resolution of some of the clades.

1.6.1.2 Plastid DNA

The plastid genome (cpDNA) encodes approximately 100 protein coding genes including 30 tRNA and four rRNA loci (Sugiura, 1989) and an inverted repeat region. The protein-coding region of plastid DNA is typically conservative in size, gene content and structure (Downie & Palmer, 1992) and major variations are in non-coding region, which are derived from insertion and deletion events (Downie & Palmer, 1992; McCauley, 1995). In previous phylogenetic analyses of Elaeocarpaceae, one non-coding region (the *trnL-trnF* intergenic spacer) was employed (Maynard, 2004; Crayn, 2006), with little success in resolving relationships. Shaw et al. (2007) and Aoki et al. (2004) explored non-coding regions of plastid DNA and found variation that may be useful for the inference of phylogenetic relationships at lower taxonomic levels in *Elaeocarpus*. It will be crucial to screen more rapidly evolving plastid markers such as these for their potential utility for resolving infra-generic phylogenetic relationships in Elaeocarpaceae.

One of the advantages of plastid DNA in phylogenetic analysis is that it allows assessment of the geographical patterns; plastid DNA is uni-parentally inherited through organelles which are generally inherited maternally in angiosperms (Milligan, 1992), are haploid, and do not undergo recombination. Because of its inheritance mechanisms, plastid DNA is often more indicative than is nuclear DNA of geographical patterns such as migration. Nuclear DNA, on the other hand, shows bi-parental inheritance, and is subject to recombination, allowing evolution to occur at a much faster rate compared to plastid DNA, and is often suitable in assessing inter-specific relationships and detecting hybridisation events. In order to exploit the complementary advantages of these genomes, both nDNA and plastid DNA are routinely employed in inferring phylogenetic relationships in plants.

1.6.1.3 Morphology

Prior to the broad availability of molecular markers, morphology was the main source of characters for phylogenetic analysis. Most phylogenetic studies however, use morphology-

based classifications to help in forming hypotheses of relationships that are subsequently tested using analysis of molecular data (Buerki et al., 2009; Manos et al., 2001; McDade et al., 2008). Often, morphological characters are then mapped onto the molecular-based phylogenies in order to test hypotheses about the evolution of those characters (Blanke et al., 2012; Horn et al., 2012; Nürk et al., 2013; Oakley & Cunningham, 2002).

Previous researchers attempted to use morphological characters for establishing an infrageneric classification of *Elaeocarpus* (Coode & Weibel, 1994; Masters, 1874; Schuman, 1890; Smith, 1944, 1953; Tirel, 1983; Weibel, 1968). The resulting infra-generic groups were established on the basis of characters such as number of carpels per ovary and ovules per carpel, embryo shape, and endosperm features. In recent years the reliability of morphology-based classifications has been tested against molecular analyses using plastid DNA and nrDNA (Crayn et al., 2006; Maynard, 2004). A molecular phylogeny based on nuclear *ITS* and plastid *trnL-F* data (Crayn et al., 2006) is generally concordant with the relationships of the genera proposed from the cladistic analysis based on morphological characters (Coode, 1987). On the other hand, within the genus *Elaeocarpus*, only limited resolution was achieved using *ITS* and *trnL-F* markers and congruence with previous morphology-based classifications was limited to only a few infra-generic groupings (Maynard, 2004).

1.6.2 Phylogenetic Analysis Methods

Approaches to phylogenetic analysis may be categorised into two major types: parsimony criterion and model-based criterion.

1.6.2.1 Parsimony method

Parsimony is a non-parametric statistical (Spencer et al., 2005) method. This method seeks the tree(s) that represents the fewest evolutionary events (or steps) needed to explain the observed data (Fitch, 1975).

An advantage of parsimony analysis is that it imposes the minimum assumptions on the data (Yang & Rannala, 2012). One of the disadvantages is that analysis of large datasets, and datasets with weak phylogenetic signal, may take excessive time (Blair & Murphy, 2011). Another major drawback is the lack of explicit models to incorporate knowledge of sequence evolution (Yang & Rannala, 2012) and to correct for substitution saturation.

1.6.2.2 Model-based methods

Model-based methods are the parametric counterpart of parsimony (Blair & Murphy, 2011). Model-based phylogenetic inference is commonly undertaken using Maximum likelihood (ML) and Bayesian inference methods. Both methods explore "tree space" but the two approaches differ fundamentally in how they perform this (Bromham, 2008).

The ML approach seeks the set of parameter values that maximises the probability of the data given the models applied, to find the tree topology that best explains the sequence data observed (Bromham, 2008). Bayesian inference calculates the probability of trees (hypotheses) given prior belief (parameters), and the dataset (Bromham, 2008). Both ML and Bayesian inference methods can be applied to various data types provided an appropriate, explicit model of evolution can be formulated (Bromham, 2008; Yang & Rannala, 2012). Bayesian posterior probability values are more straightforward to interpret as a measure of clade support than are ML bootstrap values (Yang & Rannala; 2012) however posterior probability values tend to be inflated (Lewis et al., 2005; Suzuki et al., 2002; Yang & Rannala, 2005; 2012).

Because of the flexibility of the evolutionary model, model-based methods become computationally intensive when the dataset dimensions, number of parameters and partitions increase (Yang & Rannala; 2012). However, recent advances in computational algorithms and software, increased access to multicore processors and web-accessible supercomputers, have partially overcome the computational disadvantages (Yang & Rannala, 2012).

1.6.2.3 Methods for assessing group support

Group support on parsimony and ML trees is usually assessed using resampling methods such as bootstrap and jackknife. Bootstrap resampling involves creating a new alignment of the same dimensions by sampling with replacement from the original dataset, and reconstructing a tree independently from each new alignment. This process is replicated a specified number of times (Felsenstein, 1985). The bootstrap values for a group represents the percentage of trees from the population of best trees from all replicates, in which that group appears. Jackknife resampling is similar but samples without replacement, creating alignments of smaller dimensions than the original (Farris et al., 1996). Recent comparative studies proved some evidence that favours the Jackknife resampling method over bootstrap (Freudenstein & Davis, 2010; Simmons, 2011). This is because unlike bootstrap analyses, jackknife analyses resample the original dataset without replacement therefore generating replicate datasets that are more closely related to the original dataset (Freudenstein & Davis, 2010).

1.6.3 Population genetics

DNA sequence data has demonstrated its utility for phylogenetic reconstruction at higher levels of the taxonomic hierarchy in many major plant and animal lineages. With the discovery of fast evolving sequence markers (Shaw et al., 2007), together with advent of sophisticated computer software and tree search algorithms that can handle large multi-locus data sets, the resolution of relationships at lower taxonomic levels may be facilitated (Blair & Murphy, 2011). However, it remains difficult to resolve relationships among closely related, recently derived groups. In recent years various PCR based techniques such as Single Nucleotide Polymorphisms (SNPs), Random Amplification of Polymorphic DNA (RAPD), microsatellites and Amplified Fragment Length Polymorphisms (AFLP) have been used to generate datasets suitable for investigation of relationships among samples within closely related groups (Schlötterer, 2004; Islam et al., 2005; Shaffer & Thomson, 2007; Brito & Edwards, 2009; Schulte et al., 2010; Španiel et al., 2011; Sampson & Byrne, 2012; Thurlby et al., 2012).

Microsatellite and AFLP markers are the most commonly used alternatives to sequence data to assess intraspecific taxonomic boundaries because of their tractability, repeatability and high levels of variation.

Microsatellites are DNA regions containing short tandem repeats of mono- to tetra-nucleotides (Goldstein & Pollock, 1997). They are relatively abundant, codominant, polymorphic in both nDNA and cpDNA, give uniform coverage across the genomes and are regarded as hyper-variable markers (Morgante & Olivieri, 1993). The markers allow powerful insight into gene flow, fitness and genetic distance among individuals within species. One of the limitations is the initial investment required to develop the often taxon-specific markers. Despite this limitation, microsatellite markers are now widely used in the fields of systematics, ecology, and conservation biology (Catania et al., 2008; Cerón-Souza et al., 2012; Clarke et al., 2012; Harata et al., 2012; Jørgensen et al., 2008; Piotti et al., 2012; Sampson & Byrne, 2012; Thurlby et al., 2012). In *Elaeocarpus*, nuclear microsatellite markers have been characterised and used successfully to assess inter - and intra-specific relationships across a range of Australian species (Jones et al., 2002; Rossetto et al., 2004; 2007; 2008; 2009).

In the past AFLP markers were often utilised for rapid assessments of the genetic diversity and structure within populations. Advances in methods for generating and analysing AFLP marker data (FAMD: Schlülter & Harris, 2006: GeneAlex Peakall & Smouse, 2006; Peakall & Smouse, 2012 and STRUCTURE: Falush et al., 2003, 2007; Pritchard et al., 2000) has increased their precision and cost and time efficiency. Consequently AFLPs are becoming increasingly utilised in systematics research (Flanagan et al., 2006; Paun et al., 2008; Schulte et al., 2010; Španiel et

al., 2011). AFLP markers may also be used for reconstruction of phylogenetic relationships, owing to the sheer number of alleles generated from across the entire genome (Belaj et al., 2003).

1.6.3.1 Types of analyses

Delimitation of closely related groups is often assessed by estimating genetic diversity, relationships between individuals, genetic differentiation between individuals and populations, genetic isolation by distance and population structure (Bensch & Åkesson, 2005; Bonin et al., 2007; Sites & Marshall, 2003). For instance, genetic diversity can be assessed by estimation of allelic diversity and heterozygosity calculated from the alleles or allele frequencies. It can be visualised using principal coordinate analysis, Neighbour Joining and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) using a specified dissimilarity/similarity coefficient. Visual representation of a distance metric derived by the similarity/dissimilarity coefficient is a very useful initial approach, but it is important to combine such visual appraisals of patterns in the data with statistical measures, as interpretation of ordination plots can be subjective.

1.6.4 Morphometrics

Comparisons of independent data of molecular and morphology can lead to a more complete understanding of the nature of biological variation (Lowe et al., 2004; Wood & Nakazato, 2009). Amongst available approaches addressing morphological data, morphometric approach is most commonly used in systematics, especially when addressing morphologically closely related species.

Morphometrics is defined as a quantitative description, analysis and interpretation of shape, its variation, size of morphological characters (Rohlf, 1990) and any other source of information such as geography, ontogeny and environmental effects (Rohlf & Marcus, 1993). Morphometric analyses include both quantitative (numerical) and qualitative (categorical) measurements, depending on the analyses (Table 1.1). The approach can be uni-, multi-variate or a combination of both, which makes the analyses versatile. Because multivariate analyses offer insight into the relationships between variables, this analysis approach is use in various fields of science (James & McCulloch, 1990). One of the features of morphometrics is that it treats continuous numerical measurements to observe natural patterns that may exist in the measurements without breaking them into arbitrary categorical characters (multistate characters) and limit subjectivity (Stuessy, 2009). Further more, because of the diversity of measurements that morphometrics can accommodate, and its powerful statistical approach to the data, it is proven to be a useful

approach to address species' level taxonomic questions (Allen, 2001; Andrew et al., 2003; Fatemi et al., 2007; Flann, 2003; Gabrielson et al., 2011; Gibson et al., 2012; Glaw et al., 2012; Grossi et al., 2011; Lihová et al., 2010; Morrison & Weston, 1985; Páez-Moscoso & Guayasamin, 2012).

1.6.4.1 Types of analyses

Univariate analyses deal with one variable at a time, whereas multivariate analyses allow for the analysis of multiple variables simultaneously to determine levels of correlation between them (James & McCulloch, 1990). James and McCulloch (1990) evaluated multivariate analysis methods commonly used in ecology and systematics. The objectives of multivariate analysis can be categorised as: data reduction, ordination, description, prediction, inference, allocation, and classification (James & McCulloch, 1990). With the aim of delimiting a complex morphological species, analyses involved in ordination and classification methods were reviewed. The analysis objectives and limitations are given in Table 1.1. The data type dictates the choice of analysis, therefore the applicable analyses will be dependent upon the data available.

Table 1.1. Types of multivariate analysis. *LDFA: Linear discriminant function analysis, PCA: Principal component analysis, PCoA: Principal coordinate analysis, FA: Factor analysis, COA: Correspondence analysis, NMDS: Nonmetric multidimensional scaling, MLR: Multiple logistic regression, CT: Classification tree analysis. The table was summarised based on James & McCulloch, 1990, except *#1* Belbin, 1991 **#2** Tan et al. (2006)**.**

Categories	Analyses *	Analysis objectives	Limitation
Ordination	LDFA	Find linear relationships Discriminate groups of objects Reduce the dimensions of data (= Canonical variates analysis)	Analysis is intended mainly for continuous data. Equal variance-covariance are assumed
	PCA	Find linear relationships and dimension reduction	Analysis is mainly intended for continuous data Non-linear relationships may not be discovered
	PCO	Distance calculation	Result depends on the distance measure selected
	FA	Reproduce a correlation matrix among original variables Discover underlying structure in a data set	Interpretation is subjective Not ideal for non-linear relationships or categorical data.
	COA	Describe data consisting of counts. Dimension reduction. Find linear relationships.	Not suitable for data that are not counts, or non-linear data
	NMDS[#1]	Discover non-linear relationships	The analysis uses rank order information.
Classification	LDFA	Find linear relationships Discriminate groups of objects Reduce the dimensions of data (= Canonical variates analysis)	Analysis is intended mainly for continuous data. Equal variance-covariance are assumed
	MLR	Model dichotomous variables as a function of other categorical or continuous variables. Alternative to LDFA.	The procedure considers only linear functions
	CT[#2]	Build a classification model. Combination of data type dictates the decision-making algorithms.	Overfitting or underfitting of models can occur, cross validation is required.

Chapter 2 Molecular phylogenetics of the genus *Elaeocarpus* (Elaeocarpaceae)

2.1 Introduction

2.1.1 Systematics of the family Elaeocarpaceae

The Elaeocarpaceae Juss., including the former Tremandraceae R.Br. ex DC., comprises 12 genera. Members are found in all vegetated continents except Africa and North America, chiefly in tropical and subtropical environments with some species extending to the temperate zone.

Recent molecular genetic studies in Elaeocarpaceae have revealed some aspects of the evolutionary history that had not been discovered through studies of morphology alone. Placement of the family, including the Tremandraceae, in Oxalidales was robustly supported (Bradford & Barnes, 2001; Crayn et al., 2006), which is contrary to the previous morphology-based classification and placement in the Malvales (Cronquist, 1981; Takhtajan, 1997). The monophyly of all traditionally recognised genera was supported with moderate to high bootstrap values except for *Elaeocarpus*, which appeared to be paraphyletic with respect to *Aceratium* (Crayn et al., 2006). Phylogenetic relationships among the remaining genera in the family were generally robustly resolved (Crayn et al., 2006).

2.1.2 Systematics of the genus *Elaeocarpus*

Elaeocarpus is the largest genus in the Elaeocarpaceae. It is predominantly distributed in the Old World tropics and subtropics, from south India throughout Southeast Asia to Australia, with outliers in Madagascar, Mauritius, Japan, Hawaii and other Pacific Islands, and New Zealand (Coode, 2004). The highest species diversity is found in Papuasia (Zmarzty, 2001). The ecology of *Elaeocarpus*, including fruit dispersal mechanisms and pollination vectors, is largely unknown, due in part to the size and the diversity of the genus. Continuing discoveries of new species and propositions for new combinations of taxa have created a degree of taxonomic uncertainty concerning the total number of species within the genus, i.e. at least 250 spp. by Smith (1981), 300 spp. by Tirel (1985), 460 spp. by Merrill (1951) and 500 spp. by Tirel (1982). A general consensus regarding the number of species – c. 350-360 spp. – has been reached in more recent years (Coode 2002, 2005; Tang and Phengklai, 2007; Zmarzty, 2001). Revisions at the regional level have been published for the Malay Peninsula (Ridley, 1992), Australia and New Zealand (Coode, 1984), Papuasia (Coode, 1978, 1980a, 1980b, 1984, 1995, 1996a, 1996b, 1996c, 1998, 2001a, 2001b, 2001c, 2001d, 2001e, 2001f, 2002, 2003, 2010;

Coode & Weibel, 1994), Southern India and Sri Lanka (Zmarzty, 2001; Section *Elaeocarpus* only), China (Tang & Phengklai, 2008), Java (Backer et al., 1963), Fiji (Smith, 1981), New Caledonia (Tirel, 1982), Samoa (Christophersen, 1935) and Madagascar (Tirel, 1985). Despite the existence of these regional floras, the lack of a recent monograph for the genus means that there is considerable uncertainty about the number of taxa and the relationships between them.

2.1.3 Infrageneric classification of *Elaeocarpus*

The present taxonomic conundrum within the genus appears to be the result of three main factors. The first one is the large size of the genus, its wide distribution and the existence of morphologically complex species groups. The second is that different infrageneric classifications have been used in different local revisions. The third problem, especially with respect to SE Asian *Elaeocarpus*, is that many species published prior to World War II were described without mentioning important characters used in the infrageneric classification, because insufficient material was available at the time (Merrill, 1951; Smith, 1957). This problem was exacerbated by the destruction of two herbaria (Botanic Garden and Botanical Museum Berlin-Dahlem (B) and Philippine National Herbarium (PNH)) during the war, which housed major *Elaeocarpus* collections, including many holotypes. These problems have somewhat delayed progress toward a cohesive infrageneric classification and monograph of the genus.

The first practical (*sensu* Smith, 1944) infrageneric classification was proposed by Schlechter (1916). He erected and assigned taxa of *Elaeocarpus* to nine sections, four of which are sections incorporated from previous studies. The sections were proposed using Papuasian species as examples on the basis of ovary number, ovule number per carpel, pubescence on ovary, petal divisions, petal thickening and raceme arrangement (horizontal or erect). Smith (1944) subsequently extrapolated Schlechter's sectional classification with slight modifications, such as anther tip morphology and whether or not the putamen (hereafter called the stone) is flattened. Weibel (1968) discovered that variation in embryo shape provides useful characters for infrageneric classification. His classifications were developed and expanded on by himself and Coode (Coode, 1978, 1980a, 1980b, 1984, 1995, 1996a, 1996b, 1996c, 1998, 2001a, 2001b, 2001c, 2001d, 2001e, 2001f, 2002, 2003, 2010; Coode & Weibel, 1994) and included additional morphological characters and new combinations of characters such as indumentum of petals, endocarp texture and endosperm rumination (Table 2.2). As more characters were included, the circumscription of the sections/groups changed (Coode & Weibel, 1994; Coode, 1978, 1984; C. Tirel, 1983; Weibel, 1968). Presently, 14 groups are recognised (Table 2.1) based mainly on ten morphological characters (Table 2.2).

The 28 currently recognised Australian species represent nine groups (Coode, 1984; Maynard et al., 2008), some of which also occur in Malesia. Three of these groups that occurs in Australia (Group V subgroups A, B and D, Group VI subgroups B and D, and Group VII) extend into New Guinea, Indonesia, Malaysia and the Philippines (Table 2.1).

In Malesia, Fiji and Samoa, the group system described above is not used. Instead, a non-numerical/alphabetical system is used to designate the groups that often, but not always, correspond with the numberical/alphabetical system above.

Table 2.1. Comparisons of groups between Australian and Malesian/Pacific *Elaeocarpus* in Coode's classification scheme.

Note: *E. sedentarius* Maynard & Crayn is not assigned to a group.

Groups	Subgroups	Species in Australia	Location of groups and their alternative group name *no alternative group name, numerical group name used
I		None	**Malesia** *Lobopetalum* Schltr. **New Caledonia***
II		None	**Malesia** *Dactylosphaera* Schltr. **New Caledonia**
III	A&B	None	**Malesia** *Elaeocarpus* Coode **Samoa** *Chascanthus Schltre sensu lato* **New Caledonia***
IV		*E. johnsonii* F.Muell. ex C.T.White	**Malesia** *Blepharoceras* Schltr. s*ensu stricto* **Fiji** *Blepharoceras* Schltr. s*ensu stricto* **Samoa** *Blepharoceras* Schltr. s*ensu stricto* **New Caledonia***
V	A&B	*E. grandis* F.Muell.	**Malesia, Fiji, New Caledonia and Samoa**

Groups	Subgroups	Species in Australia	Location of groups and their alternative group name *no alternative group name, numerical group name used
			Ganitrus Brongn. & Gris *pro parte*
	C	None	**Malesia***
	D	*E. arnhemicus* F.Muell. *E. obovatus* G.Don	**Malesia, Fiji** *Fissipetalum* Schltr
	E	None	*Fissipetalum* sensu Smith *pro parte*
VI	A	None	**Malesia and Fiji** *Monocera* Smith *auct. non* Brongn. & Gris, *pro parte* **New Caledonia***
	B	*E. bancroftii* F.Muell. & F.M.Bailey *E. stellaris* L.S.Sm.	**Malesia***
	C	None	**Malesia** *Papuanthus* Schltr.
	D	None	**Malesia** *Blepharoceras* Schltr *pro parte*
	E	*E. coorangooloo* J.F.Bailey & C.T.White *E. miegei* Weibel	**Malesia** *Oreocarpus* A.C.Sm. *auct. non* Schltr., *pro parte* **Vanuatu** *Oreocarpus* A.C.Sm
	F	*E. williamsianus* Guymer	None
VII		*E. carolinae* B.Hyland & Coode, *E. culminicola* Warb. *E. eumundii* F.M.Bailey, *E. grahamii* F.Muell. *E. kirtonii* F.Muell. ex F.M.Bailey, *E. linsmithii* Guymer	**Malesia** *Oreocarpus* Schltr *sensu stricto* **New Caledonia ***

Groups	Subgroups	Species in Australia	Location of groups and their alternative group name *no alternative group name, numerical group name used
		E. reticulatus Sm.	
VIII	A to D	None	**Malesia, Vanuatu, Samoa** *Coilopetalum* Schltr
IX		None	**Malesia***
X		*E. holopetalus* F.Muell.	None
XI	A	*E. ruminatus* F.Muell.	None
	B	*E. elliffii* B.Hyland & Coode, *E. ferruginiflorus* C.T.White *E. foveolatus* F.Muell., *E. largiflorens* C.T.White *E. sericopetalus* F.Muell, *E. thelmae* B.Hyland & Coode	None
XII		*E. costatus* M.Taylor	None
Informal group		None	**Malesia, possibly related to XI** *Polystachyus* group
Informal group		None	**Malesia** *Acronodia* group

Table 2.2. Key characters used in infrageneric classifications summarised from Coode (1978, 1984, 2010). Only subgroups containing samples included in this study are included in the table. Groups that further divided into subgroups but not summarised in this table are indicated with*. The Polystachyus and Acronodia groups include species having unisexual flowers, the rest of the groups are bisexual. #: following endosperm ornamentation defined by Coode (1984)

Infrageneric groupings	Ovules /loculus	Loculus /ovary	Ovary Indu-mentum	Petal apex	Petal character and indumentum	Stamen numbers	Anther tip	Embryo	Endosperm[#]	Stone shape ornamentation
I	2	2	Glabrous	Entire or 3-notched	Thin and minutely hairy	8-10	No awns	Straight	Entire	Smooth; ellipsoid
II	2	3	Minutely hairy	Divided	Thickened at apex, and hooked inwards, minutely hairy	10-16	No awns	Straight	Entire	Smooth to weakly rugose, ellipsoid
III*	2	3	Minutely hairy	Divided	Thin and hairy	10-40	Awned, apiculate, or group of bristles	Straight	Entire	Rugose; ellipsoid, ovoid or globose
IV	4	2	Hairy	Divided	Thin and hairy	13-35	Awned or not	Straight	Entire	Smooth; ellipsoid
V A, B	4-6	5-7	Minutely hairy	Divided	Thin and hairy/glabrous	More than 30	Not awned	Straight	Entire	Sculptured or rugose; globose or ellipsoid or obovoid
V C	4-8	4-5				20-30				
V D	4-8	2				20 or less rarely 25				
V E	5	2				15-20				
VI A	6	2	Hairy or glabrous	Divided	Thin and hairy/glabrous	c.30	Awned	Straight	Entire	Smooth; globose, ellipsoid or ovoid; winged, flattened, no flattened or and 4-sided
VI B	9-10	5		With round teeth		30-50				
VI C	9-12	2		With teeth		25-50				
VI D	6-8	2		Divided		15-40				
VI E	6-8	2		Divided		c.25	Not awned			Rugose or sculptured; ellipsoid or spherical
VI F	6	3-4		Divided		30-38				
VII	6-8	2-3	Glabrous	Divided	Thin and hairy	15-40	Awned	Curved	Entire	Sculptured; ellipsoid or ovoid

Infrageneric groupings	Ovules /loculus	Loculus /ovary	Ovary Indu-mentum	Petal apex	Petal character and indumentum	Stamen numbers	Anther tip	Embryo	Endosperm[#]	Stone shape ornamentation
VIII*	6-12	2-4	Hairy or glabrous	Divided or with or without teeth	Thickened, hairy on the back	15-90	Awned or not	Curved	Ruminate	Rugose; ellipsoid, ovoid, ovoid-ellipsoid
IX	8	2	Hairy	With blunt teeth	Thin and hairy,	c.15	Not awned	Unknown	Unknown	Unknown
X	4	2	Hairy	Undivided	Thin and glabrous	c.15	Not awned	Curved	Entire	Rugose; oblong-globose
XI A	7-8	2	Hairy	With 3-5 irregular teeth or entire	Thickened and hairy	16-22	Awned	Curved	Ruminated	Sculptured; ellipsoid
XI B	6	3				30-70	Not awned			Smooth; ellipsoid
XII	6-10	3-4	Hairy	With 5-6 irregular teeth	Thin and hairy	c.20	Awned	Curved	Entire	Rugose-tuberculate; ellipsoid
Polystachyus	(4-) 6 (-12)	2-3	Hairy	With < 10 teeth	Thickened and hairy	35-80	Not awned	Curved	Ruminate	Rugose and 6-sided; ellipsoid
Acronodia	4	2	Hairy or glabrous	Scarcely divided or with teeth	Thin and hairy/glabrous	8- 12	Not awned	Curved	Ruminate	Rugose or smooth; ellipsoid or ovoid; flattened or not

2.1.4 Toward a phylogenetic infrageneric classification of *Elaeocarpus*

To date infrageneric classifications of *Elaeocarpus* were based on intuitive assessment of morphological characters. However, molecular-based studies, which seek to reconstruct evolutionary relationships as a basis for classification, have been initiated in recent years. Phylogenetic analyses of sequence data derived from nuclear and plastid markers have supported the monophyly of Group V Subgroups A, B & D (Crayn et al., 2006; Maynard, 2004). Sampling of *Elaeocarpus* species in these studies, however, was low and the dataset contained only a few non-Australian representatives of Coode's (1984) groups. Furthermore, the markers used (nuclear internal transcribed spacer [ITS] region, and plastid *trnL-trnF* region) proved inadequate to resolve the deep nodes within the genus. To more thoroughly test the current classification, increased sampling of taxa from across the geographic range of the genus including representatives from all of the groups, and new and more informative molecular markers are required.

Coode (1984) highlighted several species-level problems including the *E. obovatus* species complex and the existence of undescribed taxa requiring further investigation. In the *E. obovatus* species complex there are three putative entities that show high morphological diversity. The first approach for resolving this species complex is to investigate if this morphological group is a natural or artificial assemblage.

Five putative new species have been documented in Australia under informal phrase names (CHAH, 2011). Each of these taxa shows an affinity to a known species but also maintains some degree of difference. It is important to ascertain which groups these taxa fit in to and to test the taxonomic boundaries between these species by correlating the morphological classification against molecular phylogenetics, prior to formal description, so as to avoid the duplication or misapplication of names.

Infrageneric classifications are more reliable asinformation storage and retrieval devices when they are based on evolutionary relationships. In this study, I test and re-evaluate the classification of Coode, the most comprehensive and most widely accepted classification, against a molecular phylogeny.

The specific questions this chapter addresses are:

1. Is *Elaeocarpus* monophyletic?
2. Are the existing infrageneric classifications of *Elaeocarpus* accurate with respect to evolutionary relationships?

3. Is the *Elaeocarpus obovatus* species complex a monophyletic group?
4. To which infrageneric groups do the putative new Australian *Elaeocarpus* taxa belong?

2.2 Materials and Methods

2.2.1 Taxon sampling

The taxon set included seventy-three species of *Elaeocarpus* and seventy-one outgroups representative of all the genera in Elaeocarpaceae, for a total of 144 terminals. Because comprehensive sampling from all the sections from across their geographical range was beyond the scope of this study, the sectional relationships were assessed only where suitable samples were available. All species (including undescribed taxa) found in Australia, except *E. miegei* from offshore islands of the Northern Territory, were sampled.

Forty-two sequences of the Internal Transcribed Spacer region of 18S–5.8S–26S nuclear ribosomal cistron (hereafter referred to as nrITS) and 60 sequences from the plastid *trnL-trnF* region from across all twelve genera of Elaeocarpaceae were obtained from GenBank (http://www.ncbi.nlm.nih.gov/). Additionally 12 nrITS sequences from Maynard (2004) were included.

Genomic DNA for some Elaeocarpaceae was obtained from previous studies. For most Australian species, however, new collections were required. For these, fresh leaves from field collections were dried in silica gel. DNA was extracted using DNEasy Plant Mini kits (Qiagen, Hilden, Germany) following the manufacturer's protocol. Details of all accessions used in this study are provided in Table 2.3.

Table 2.3. Samples used in this study. Groups assigned here are provisional, taxa of unknown groups was left unassigned. Subgroups are in the brackets. The herbarium/herbaria at which voucher is housed follow collector ID in parentheses. AU: Australia, NC: New Caledonia, NZ: New Zealand and PNG: Papua New Guinea. + New sequences generated from the present study. * Sequences donated from the study of Maynard (2004). # Sequences donated from the study of Boucher et al. (unpublished). Any symbols (+ and *) that are not accompanied by a GenBank accession number are suspected to be a paralogous gene. N/A: not available

Taxon	Groups	Collector IDs	*Localities*	*trnV-ndhC*	*trnL-F*	nrITS
Elaeocarpus						
E. alaternoides Brongn. & Gris	I	D. M. Crayn 749 (NSW)	NC	KJ658421^{+}	KJ631296^{+}	KJ675644^{+}
E. angustifolius Blume_ PNG	V (A)	D. M. Crayn 572 (NSW)	PNG	KJ658422^{+}	KJ631297^{+}	KJ675645^{+}
E. angustifolius Blume_ India	V (A)	NSW710750 (NSW)	India	KJ658423^{+}	KJ631298^{+}	KJ675646^{+}
E. sphaericus K.Schum._ India	V (A)	NSW710753 (NSW)	India	N/A	KJ631299^{+}	KJ675679*
E. sphaericus K.Schum._ Hawaii	V (A)	Flynn 7277 (NSW)	Hawaii	KJ658424^{+}	N/A	KJ675647^{+}
E. arnhemicus F.Muell.	V (D)	Y. Baba 341 (CNS)	AU	KJ658425^{+}	KJ631300^{+}	KJ675648^{+}
E. bancroftii F.Muell. 2	VI (B)	Y. Baba 351 (CNS)	AU	KJ658426^{+}	KJ631301^{+}	KJ675649^{+}
E. bancroftii F.Muell. 1	VI (B)	D. M. Crayn 502 (NSW)	AU	N/A	DQ444685	DQ448687
E. bifidus Hook. & Arn.	Unassigned	Trauernicht 649 (PTBG)	Hawaii	KJ658427^{+}	KJ631302^{+}	KJ675650^{+}
E. brachypodus Guillaumin	VI	Y. Pillon 71 (NOU)	NC	KJ658428^{+}	KJ631303^{+}	KJ675651^{+}
E. bullatus Tirel	Unassigned	J. Munzinger 2906 (NOU)	NC	KJ658429^{+}	KJ631304^{+}	N/A
E. carolinae B.Hyland & Coode	VII	A. Ford 4444 (CNS)	AU	KJ658430^{+}	KJ631305^{+}	N/A
E. carolinensis Koidz.	Unassigned	Lorence 10004 (PTBG)	Caroline Island, Federated States of	KJ658431^{+}	KJ631306^{+}	N/A

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
			Micronesia			
E. coumbouiensis Guillaumin	Unassigned	Y. Pillon 388, M. Gaudeul, E. A. Brown & G. McPherson (NOU)	NC	KJ658432+	KJ631307+	N/A
E. coorangooloo J.F.Bailey & C.T.White	VI (E)	Y. Baba 695 (CNS)	AU	KJ658433+	KJ631308+	KJ675652+
E. costatus M.Taylor	XII	E. A. Brown 2003/55 (NSW)	AU	N/A	N/A	*
E. crenulatus R.Knuth	III (E)	D. M. Crayn 539 (NSW)	PNG	KJ658434+	KJ631309+	N/A
E. michaelii C.T.White 1	VII	Y. Baba 350 (CNS)	AU	KJ658435+	KJ631310+	+
E. michaelii C.T.White 2	VII	D. M. Crayn 499 (NSW)	AU	N/A	DQ444688	N/A
E. dentatus (j. R. & G. Forst.) Vahl	V (D)	N/A	NZ	KJ658436+	KJ675689+	N/A
E. dongnaiensis Pierre	Unassigned	NHK 1118 (K)	Vietnam	KJ658437+	KJ675690+	KJ675653+
E. elliffii B.Hyland & Coode	XI (B)	D. M. Crayn 884 (NSW, CNS, BRI)	AU	KJ658438+	KJ675691+	+
E. eumundii F.M.Bailey 2	VII	A. Ford 4459 (NSW, CNS)	AU	KJ658439+	KJ675692+	+
E. eumundii F.M.Bailey 1	VII	D. M. Crayn 505 (NSW)	AU	N/A	DQ444682	N/A
E. ferruginiflorus C.T.White 2	XI (B)	D. M. Crayn 882 (NSW)	AU	KJ658440+	KJ675693+	+
E. ferruginiflorus C.T.White 1	XI (B)	G. Fensom 401 (NSW)	AU	N/A	DQ444692	N/A
E. foveolatus F.Muell. 1	XI (B)	D. M. Crayn 856 (NSW)	AU	KJ658441+	KJ675694+	+
E. foveolatus F.Muell. 2	XI (B)	P. D. Hind 6265 (NSW)	AU	N/A	DQ444691	N/A
E. geminiflorus Brongn. & Gris	Unassigned	J. Munzinger 2866 (NOU)	NC	N/A	KJ675695+	N/A
E. glaber Blume	III	Living collection VI.C.179A. Kebun	Java,	KJ658442+	KJ675696+	KJ675654+

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
		Raya Bogor	Indonesia			
E. gordonii Tirel	Unassigned	Y. Pillon 300 & C. Grignon (NOU, P)	NC	KJ658443[+]	KJ675697[+]	N/A
E. grahamii F.Muell.	VII	A. Ford 4313 & G. Sankowsky (CNS)	AU	KJ658444[+]	KJ675698[+]	*
E. grandiflorus Sm.	Unassigned	NSW710751	Malesia	KJ658445[+]	KJ675699[+]	N/A
E. grandis F.Muell.	V (A)	P. I. Forster 27569 (BRI)	AU	KJ658446[+]	KJ675700[+]	KJ675655[+]
E. gummatus Guillaumin	Unassigned	Y. Pillon 260 & A.Vergnes (NOU)	NC	KJ658447[+]	KJ675701[+]	+
E. holopetalus F.Muell.	X	J. M. Allen s.n. (NSW605470)	AU	KJ658448[+]	KJ675702[+]	KJ675686[+]
E. hookerianus Raoul 1	V (D)	J. M. Allen s.n. (NSW605721)	NZ	N/A	DQ444686	DQ448688
E. hookerianus Raoul 2	V (D)	C. D. Kilgour 787 (AK)	NZ	KJ658449[+]	KJ675703[+]	KJ675656[+]
E. hortensis Guillaumin	VI	J. Munzinger 2968 (NOU)	NC	KJ658450[+]	KJ675704[+]	N/A
E. hylobroma Y.Baba & Crayn	Unassigned	D. M. Crayn 838 (NSW)	AU	KJ658451[+]	KJ675705[+]	KJ675680[*]
E. japonicus Siebold	Unassigned	K. Aoki 010974 (KYO)	Japan	KJ658452[+]	N/A	KJ675687[+]
E. johnsonii F.Muell. ex C.T.White	IV	W. W. Cooper 2122 (CNS, BRI)	AU	KJ658453[+]	KJ675706[+]	KJ675657[+]
E. kerstingianus Schltr.	Unassigned	Perlman 21433 (PTBG)	Federated States of Micronesia	KJ658454[+]	KJ675706[+]	N/A
E. kirtonii F.Muell. ex F.M.Bailey	VII	D. M. Crayn 501 (NSW)	AU	KJ658455[+]	DQ444687	*
E. largiflorens subsp. *largiflorens* C.T.White	XI (B)	D. M. Crayn 796 (NSW)	AU	KJ658456[+]	KJ675708[+]	*
E. largiflorens subsp. *retinervis* B.Hyland	XI (B)	D. M. Crayn 503 (NSW)	AU	N/A	DQ444684	DQ448686

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
& Coode 1						
E. largiflorens subsp. *retinervis* B.Hyland & Coode 2	XI (B)	Y. Baba 357 (CNS)	AU	KJ658457+	KJ675709+	KJ675658+
E. linsmithii Guymer	VII	Y. Baba 833, W.W. Cooper, R. Jensen & S. McKenna (CNS)	AU	KJ658458+	KJ675710+	N/A
E. multisectus Schltr.	III (A)	D. M. Crayn 561 (NSW)	PNG	KJ658459+	KJ675711+	N/A
E. nouhuysii Koord. 1	VI (B)	D. M. Crayn 530 (NSW)	PNG	KJ658460+	KJ675712+	KJ675659+
E. nouhuysii Koord. 2	VI (B)	D. M. Crayn 533 (NSW)	PNG	N/A	KJ675713+	KJ675660+
E. obovatus G.Don	V (D)	N/A	AU	KJ658461+	KJ675714+	KJ675661+
E. ovigerus Brongn. & Gris	VI	D. M. Crayn 763 (NSW)	NC	KJ658462+	N/A	N/A
E. polydactylus Schltr.	V (D)	D. M. Crayn 577 (NSW)	PNG	KJ658463+	KJ675715+	KJ675682*
E. ptilanthus Schltr.	V (A)	D. M. Crayn 554 (NSW)	PNG	KJ658464+	KJ675716+	KJ675683*
E. pulchellus Brongn. & Gris	I	D. M. Crayn 758 (NSW)	NC	KJ658465+	KJ675717+	KJ675684*
E. reticulatus Sm.	VII	Allen s. n. (NSW605722)	AU	KJ658466+	DQ444683	*
E. rotundifolius Brongn. & Gris	I	D. M. Crayn 761 (NSW)	NC	N/A	KJ675718+	N/A
E. ruminatus F.Muell.	XI (A)	Y. Baba 446 (CNS)	AU	KJ658467+	KJ675719+	KJ675662+
E. sedentarius Maynard & Crayn 1	Unassigned	D. Maynard 2 (NSW)	AU	N/A	DQ444676	DQ448682
E. sedentarius Maynard & Crayn 2	Unassigned	Y. Baba 408, D. M. Crayn, M. Harrington, C. Puente & R. Kooyman	AU	KJ658468+	KJ675720+	KJ675663+

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
		(CNS)				
E. sericopetalus F.Muell.	XI (B)	D. M. Crayn 823 (NSW)	AU	KJ658469[+]	DQ444692	N/A
E. seringii Montrouz.	I	J. Munzinger 2852 (NOU)	NC	KJ658470[+]	KJ675721[+]	+
E. speciosus Brongn. & Gris	Unassigned	Y. Pillon 115 (NOU)	NC	KJ658471[+]	KJ675725[+]	KJ675664[+]
E. stellaris L.S.Sm.	VI (B)	C. Costion 3531 (CNS)	AU	KJ658472[+]	KJ675726[+]	KJ675665[+]
E. stipularis Blume_ Java	III	NSW710749	Java, Indonesia	KJ658473[+]	KJ675727[+]	KJ675666[+]
E. sylvestris Poir.	Unassigned	K. Aoki 011355 (KYO)	Japan	KJ658474[+]	KJ675728[+]	KJ675667[+]
E. sp. Mt Bellenden Ker (L.J.Brass 18336)	Unassigned	Y. Baba 443 (CNS)	AU	KJ658475[+]	KJ675722[+]	KJ675668[+]
E. sp. Mt Misery (L.J.Webb+ 10905)	Unassigned	A. Ford 4312 (NSW)	AU	KJ658476[+]	KJ675723[+]	*
E. sp. Windsor Tableland (L.W.Jessup & GJM 1378)	Unassigned	Y. Baba 397 & C. D. Kilgour (CNS)	AU	KJ658477[+]	KJ675724[+]	KJ675669[+]
E. thelmae B.Hyland & Coode	XI (B)	Y. Baba 792, R. Jensen, S.-N. Phoon & T. Roberts (CNS)	AU	KJ658478[+]	KJ675729[+]	KJ675670[+]
E. weibelianus Tirel	IV	J. Munzinger 2833 (NOU)	NC	KJ658479[+]	KJ675730[+]	N/A
E. williamsianus Guymer	VI (F)	D. M. Crayn 513 (NSW)	AU	N/A	DQ444693	DQ448691
Aceratium						
Ac. concinnum (S.Moore) C.T.White	-	D. M. Crayn 858 (NSW)	AU	KJ658480[+]	DQ444678	DQ448684
Ac. doggrellii C.T.White	-	M. Bradford 4 (CNS, NSW)	PNG	KJ658481[+]	N/A	N/A
Ac. ferrugineum C.T.White	-	M. Harrington 296 (CNS)	AU	N/A	DQ444681	DQ448685

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
Ac. ledermannii Schltr.	-	D. M. Crayn 534 (NSW)	PNG	N/A	DQ444677	DQ448683
Ac. megalospermum (F.Muell.) van Balg.	-	D. M. Crayn 523 (NSW)	AU	KJ658482[+]	DQ444679	KJ675671[+]
Ac. sericoleopsis van Balg.	-	D. M. Crayn 779 (NSW)	AU	KJ658483[+]	DQ444680	N/A
Sericolea						
Se. calophylla subsp. *grossiserrata* Coode	-	D. M. Crayn 550 (NSW)	PNG	KJ658484[+]	DQ444675	DQ448681
Se. gaultheria Schltr.	-	D. M. Crayn 553 (NSW)	PNG	N/A	DQ444674	DQ448680
Se. micans var. *micans* Schltr.	-	D. M. Crayn 536 (NSW)	PNG	KJ658485[+]	DQ444673	KJ675672[+]
Tetratheca						
Te. affinis Endl.	-	Cranfield & Ward 126 (PERTH)	AU	KJ658486[+]	N/A	N/A
Te. aphylla subsp. *aphylla* F.Muell.	-	N/A	AU	N/A	AY237265	N/A
Te. aphylla subsp. *megacarpa* R.Butcher	-	R. Butcher 908 (PERTH)	AU	N/A	AY237268	EF150670
Te. bauerifolia F.Muell.	-	T. Downing 38 (MEL)	AU	KJ658487[+]	EF095748	EF095741
Te. ciliata Lindl.	-	T. Downing 33 (MEL)	AU	KJ658488[+]	DQ444698	DQ448669
Te. confertifolia Steetz	-	D. M. Crayn 722 (NSW)	AU	KJ658489[+]	KJ675731[+]	KJ675673[+]
Te. efoliata F.Muell.	-	R. Davis 10496 (PERTH)	AU	KJ658490[+]	KJ675732[+]	N/A
Te. ericifolia Sm.	-	J. Howell s.n. (NSW619997)	AU	KJ658491[+]	EF095746	EF095738
Te. filiformis Benth.	-	R. Butcher 966 (MEL, NSW, PERTH)	AU	KJ658492[+]	DQ444695	DQ448666
Te. harperi F.Muell.	-	N/A	AU	N/A	AY237277	N/A
Te. hirsuta Lindl.	-	R. Butcher 915 (PERTH, MEL, NSW)	AU	KJ658493[+]	EF095742	EF095732

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
Te. hispidissima Steetz	-	R. Butcher 964 (PERTH, MEL, NSW)	AU	KJ658494[+]	KJ675733[+]	KJ675674[+]
Te. juncea Sm.	-	M. Rossetto 524 & D. M. Crayn (NSW)	AU	KJ658495[+]	DQ444696	DQ448667
Te. nephelioides R.Butcher	-	R. Butcher 909 & J.A. Wege (PERTH)	AU	N/A	AY237271	N/A
Te. nuda Lindl.	-	D. M. Crayn 731, K. Kron & A. Perkins (NSW)	AU	KJ658496[+]	KJ675734[+]	EF095739
Te. parvifolia Joy Thomps.	-	R. Butcher 916 (MEL, NSW, PERTH, UW)	AU	KJ658497[+]	DQ444697	DQ448668
Te. paynterae subsp. *cremnobata* R.Butcher	-	R. Butcher 902, J. Bull, E. Alacs & D. Seivers (PERTH)	AU	N/A	AY237273	N/A
Te. pilifera Lindl.	-	R. Butcher 922 (PERTH, MEL, NSW)	AU	KJ658498[+]	EF095745	EF095736
Te. pubescens Turcz.	-	T. Downing 39 (MEL, NSW)	AU	KJ658499[+]	DQ444699	DQ448670
Te. retrorsa Joy Thomps.	-	R. Butcher 929 (MEL, NSW, PERTH)	AU	KJ658500[+]	EF095743	EF095733
Te. rupicola Joy Thomps.	-	J. Bradford 871	AU	N/A	AF299192	N/A
Te. shiressii Blakely	-	D. M. Crayn 604 & M. Rossetto (NSW, PERTH)	AU	N/A	EF095747	EF095740
Te. stenocarpa J.H.Willis	-	T. Downing 53 (MEL)	AU	KJ658501[+]	DQ444700	DQ448671
Te. thymifolia Sm.	-	D. M. Crayn 602 (NSW)	AU	KJ658502[+]	KJ675735[+]	N/A
Te. virgata Steetz	-	R. Butcher 928 (MEL, NSW, PERTH)	AU	KJ658503[+]	EF095749	N/A

Taxon	Groups	Collector IDs	*Localities*	*trnV-ndhC*	*trnL-F*	nrITS
Platytheca						
Pl. galioides Steetz	-	A. N. Rodd 4973 & G. Fensom (NSW)	AU	KJ658504^{+}	DQ444694	EF095731
Tremandra						
Tr. diffusa subsp. *diffusa* DC.	-	R. Butcher 961 (PERTH, MEL, NSW)	AU	N/A	DQ444701	DQ448672
Tr. diffusa subsp. *stelligera* R.Br.	-	D. M. Crayn 706 (NSW)	AU	KJ658505^{+}	DQ444702	KJ675675^{+}
Dubouzetia						
D. campanulata Pancher ex Brongn. & Gris	-	D. M. Crayn 745 (NSW)	NC	KJ658506^{+}	DQ444667	KJ675676^{+}
D. caudiculata Sprague	-	G. McPherson 3305 (NSW)	NC	N/A	DQ444668	DQ448675
D. confusa Guillaumin & Virot	-	T. J. Entwistle s.n. (NSW615320)	NC	KJ658507^{+}	KJ675736^{+}	KJ675677^{+}
D. elegans Brongn. & Gris	-	J. Munzinger 2928 (NOU)	NC	KJ658508^{+}	N/A	+
D. guillauminii Virot	-	MCP 19401 (NOU)	NC	KJ658509^{+}	KJ675737^{+}	N/A
D. kairoi Coode	-	D. M. Crayn 578 (NSW)	PNG	N/A	DQ444670	N/A
D. saxatilis A.R.Bean & Jessup	-	D. Silke s.n. (CNS: QRS127132)	AU	KJ658510^{+}	DQ444669	DQ448676
Crinodendron						
C. hookerianum Gay	-	J. M. Allen s.n. (NSW605484)	Chile and Bolivia	KJ658511^{+}	DQ444666	DQ448674
C. patagua Molina	-	J. M. Allen s.n. (NSW605483)	Chile	KJ658512^{+}	DQ444665	DQ448673

Taxon	Groups	Collector IDs	*Localities*	*trnV-ndhC*	*trnL-F*	nrITS
Peripentadenia						
Pe. mearsii (C.T.White) L.S.Sm.	-	P. I. Forster 29760 (BRI)	AU	KJ658513^{+}	DQ444672	DQ448679
Pe. phelpsii B.Hyland & Coode	-	D. M. Crayn 887 (NSW)	AU	KJ658514^{+}	DQ444671	DQ448678
Aristotelia						
Ar. australasica F.Muell.	-	J. M. Allen s.n. (NSW605725)	AU	N/A	DQ444661	DQ448661
Ar. chilensis Stuntz	-	J. M. Allen s.n. (NSW605486)	Chile	N/A	DQ444660	DQ448660
Ar. fruticosa Hook.f.	-	M.W. Chase 781 (K)	NZ	KJ658515^{+}	DQ444662	DQ448662
Ar. peduncularis Hook.f.	-	L. Mulcahy s.n. (NSW606884)	AU	KJ658516^{+}	DQ444659	DQ448664
Ar. serrata Oliv.	-	J. M. Allen s.n. (NSW605729)	NZ	KJ658517^{+}	DQ444663	DQ448663
Sloanea						
Sl. australis (Benth.) F.Muell.	-	J. Bradford 862	AU	N/A	AF299191	AF299244
Sl. berteroana Choisy ex DC.	-	N/A	Caribbean	N/A	#	#
Sl. caribaea Krug & Urb. ex Duss	-	N/A	Caribbean	N/A	#	#
Sl. dentata L.	-	N/A	Caribbean	N/A	#	#
Sl. dussii Urb.	-	N/A	Caribbean	N/A	#	#
Sl. forbesii F.Muell.	-	B. J. Conn 5029, K. Damas, K. Fazang, O. Paul & T. Kuria (NSW709402)	PNG	KJ658518^{+}	N/A	N/A
Sl. guianensis Benth.	-	N/A	Caribbean	N/A	#	#
Sl. langii F.Muell.	-	P. I. Forster 30070 (BRI)	AU	KJ658519^{+}	DQ444655	+

Taxon	**Groups**	**Collector IDs**	***Localities***	***trnV-ndhC***	***trnL-F***	**nrITS**
Sl. lepida Tirel	-	J. Munzinger 3778 (NOU)	AU	KJ658520^{+}	KJ675738^{+}	+
Sl. macbrydei F.Muell.	-	A. Ford 4295 & B. Hewett (CNS, NSW)	AU	KJ658521^{+}	KJ675739^{+}	+
Sl. massonii Sw.	-	N/A	Caribbean	N/A	#	#
Sl. montana (Labill.) A.C.Sm.	-	D. M. Crayn 765 (NSW)	NC	KJ658522^{+}	KJ675740^{+}	KJ675678^{+}
Sl. sogerensis Baker f.	-	D. M. Crayn 532 (NSW)	PNG	KJ658523^{+}	DQ444657	DQ448658
Sl. woollsii F.Muell.	-	D. M. Crayn 780 (NSW)	PNG	KJ658524^{+}	DQ444654	DQ448657
Sl. GUY_142	-	N/A	Caribbean	N/A	#	#
Sl. GUY_155	-	N/A	Caribbean	N/A	#	#
Vallea						
V. stipularis L.f.	-	M.W. Chase 654 (K)	South America	KJ658525^{+}	DQ444664	DQ448665

2.2.2 Amplification and sequencing

Initially, a number of chloroplast regions were selected (based on comparable studies: e.g. Taberlet et al. 1991; Shaw et al. 2007) for assessment of their utility for phylogenetic reconstruction at the infra-generic level across *Elaeocarpus*. Although Aoki et al (2004) demonstrated the potential utility of several noncoding chloroplast markers for phylogeographic study of Japanese forest trees, including one *Elaeocarpus* species, the variation displayed was generally not parsimony-informative. Thus only the markers with highest potential for informative variation with *Elaeocarpus* suggested by the wider studies (Taberlet et al. 1991; Shaw et al. 2007) were employed in a pilot study.

These markers were 3' *trnV-ndhC, trnL-F, trnQ-5-rps16, rpl32-trnL* and *trnH-psbA*. Additionally, *matK, rpoC* and *psbK-psbI* were tested (Table 2.6). These regions were sequenced, aligned and assessed for their level of sequence variation across a limited sample of species that included members of four infrageneric groups that occur in Australia. For *matK* and *trnQ, rpoC, psbK-psbI* there was little to no variation across the samples whereas for *trnL-F, trnV-ndhC* and *trnH-psbA* up to c. 10 % of characters were parsimony informative.

Two plastid (*trnL-F, trnV- ndhC*) and one nuclear (nrITS) marker were sequenced across 12 genera of Elaeocarpaceae. Details of the taxa sampled (including voucher details) are given in Table 2.3 and the primers used to amplify and sequence the plastid and nuclear markers are provided in Table 2.4.

For the nrITS region, two primer pairs, *ITS 4* & *5* and *AB101* & *AB102,* were tested on the each sample and the products appeared to be a 'single band' on the electrophoresis gel images. In order to detect potential pseudogene sequences, further identification using conserved short sequence motifs was employed (see section 2.2.4).

Table 2.4. Primer set and sequences tested and then employed (*) in this study. # indicates the marker failed to amplify.

Region	Primer name		Primer sequences	References
trn L-F*	5’	*tab f*	AATTGAACTGGTGACACGAG	Taberlet et al. (1991)
	3’	*tab c*	CGAAATCGGTAGACGCTACG	
trn V-ndhC*	5’	*trnV$^{(UAC)}$x2*	GTC TAC GGT TCG ART CCG TA	Shaw et al. (2007)
	3’	*ndhC*	TAT TAT TAG AAA TGY CCA RAA AAT ATC ATA TTC	
trnQ-5 - rps16	5’	*trnQ $^{(UUG)}$*	GCG TGG CCA AGY GGT AAG GC	Shaw et al. (2007)
	3’	*rpS16x1*	GTT GCT TTY TAC CAC ATC GTT T	
rpl32-trnL$^{\#}$	5’	*trnL$^{(UAG)}$*	CTG CTT CCT AAG AGC AGC GT	Shaw et al. (2007)
	3’	*rpL32-F*	CAG TTC CAA AA A AAC GTA CTT C	
matK	5’	*3F_KIM_f*	CGTACAGTACTTTTGTGTTTACGAG	Kim Unpublished
	3’	*1R_KIM_r*	ACCCAGTCCATCTGGAAATCTTGGTTC	
rpoC	5’	*2f-rpoC1*	GGCAAAGAGGGAAGATTTCG	PWG Phase 2, Royal Botanic Gardens, Kew (2007)
	3’	*4r-rpoC1*	CCATAAGCATATCTTGAGTTGG	
trnH-psbA	5’	*psbA3_f*	GTTATGCATGAACGTAATGCTC	Sang et al. 1997
	3’	*trnHf_05*	CGCGCATGGTGGATTCACAATCC	Tate & Simpson, 2003
psbK-psbI	5’		TTAGCCTTTGTTTGGCAAG	Kim Unpublished
	3’		AGAGTTTGAGAGTAAGCAT	
nrITS 1 and 2*	5’	*AB101*	TAGAATTCCCCGGTTCGCCGTTAC	Modified based on Sun et al., 1994
	3’	*AB102*	TAGAATTCCCCGGTTCGCTCGCCGTTA	
	5’	*ITS 5*	GGAAGTAAAAGTCGTAACAAGG	White et al, 1990
	3’	*ITS 4*	TCCTCCGCTTATTGATATGC	

PCR reactions were performed in a Mastercycler® ep Realplex (Eppendorf, Hamburg, Germany). Each 20 μL reaction for all regions contained: 2.5 μL of 10x PCR buffer (Kapa Biosystems), 3.13 mM of MgCl, (Kapa Biosystems), 0.25 mM of dNTP (Kapa Biosystems), 0.63 mM of primers each, (Genworks), 0.40 mM of Dimethyl sulfoxide, 0.25 μL of 0.4% Bovine Serum Albumin, 0.2 μL of DNA polymerase (5U/μL) (Kapa Biosystems), 1 μL of template DNA, and distilled H_2O to 20 μL. The reaction mix was initially incubated at 95^{o}C for

2 min; then subjected to 35 cycles of the following profile: denaturation at 95°C for 30 s, annealing at 55°C for 30s and extension at 72°C for 1 minute; and completed with a final incubation at 72°C for 2 minutes. PCR products were purified with ExoSAP-IT (USB Inc., Cleveland, OH, USA) in accordance with the manufacturer's protocol. PCR products were sequenced in both directions. Each 10 μL sequence reaction mixture contained approximately 10-20 ng of purified PCR product, 1 μL of BigDye Terminator v 3.1 (Applied Biosystems, Foster City, CA, USA), 1.5 μL of BDT buffer (ABI) and sterile water to 10 μL and was subjected to 25 cycles of 10 s at 96 °C, 5 s at 50 °C and 4 min at 60 °C, after which a final 1 min incubation at 60 °C was carried out. The sequencing products were purified using a magnesium sulphate clean-up protocol, separated and visualised on an ABI 3730 x I automated sequencer at the Australian Genome Research Facility (AGRF).

2.2.3 Sequence alignment

The chromatograms were checked and contigs built in Chromas Pro v 1.7.5 (http://www.technelysium.com.au/ ChromasPro.html). The contigs were then assembled into separate matrices using BIOedit v 7.1.3 (Hall, 2011). Each data partition was aligned using MAFFT version 6 (Katoh & Toh, 2010) with the following settings: Q-ins-i for nrITS and G-INS-I for the plastid partitions. Alignments were subsequently manually adjusted using BIOedit.

2.2.4 Checking homology of nrITS with conserved motifs

As a fast evolving marker, nrITS has the potential to provide data useful for resolving relationships among recently diversified groups (Alvarez & Wendel, 2003; Baldwin et al., 1995; Feliner & Rossello, 2007; Hughes et al., 2006). However, problems associated with incomplete concerted evolution in this multiple copy tandem repeat unit can violate the assumption of orthology of the nrITS sequences and therefore potentially limit the use of this marker in phylogenetic analyses (Feliner & Rossello, 2007; Small et al., 2004).

The existence of paralogous copies ('pseudogenes') is a confronting issue; however, these can be potentially detected and excluded using conserved motifs found in the ITS1 and the 5.8S gene (Liu & Schardl, 1994; Hershkoviz & Lewis, 1996; Jobes & Thien, 1997; Coleman, 2003; Feliner & Rossello, 2007; Harpke & Peterson, 2008). Motifs in 5.8S are highly conserved because nucleotide substitutions within them may disrupt the formation of secondary structure (Harpke & Peterson, 2008). The possession of these motifs is therefore a reliable indicator of orthologous gene copies.

As yet no thorough assessment of the practicality and applicability of the use of the nrITS marker for phylogenetic reconstruction in Elaeocarpaceae has been undertaken despite potential paralogous copies having been identified within a single *Elaeocarpus* (Crayn et al., 2006). Here one conservative motif from the ITS1 and three from 5.8s were utilised for the detection of putative pseudogene sequences. Motifs are provided in Table 2.5.

Table 2.5. nrITS motifs used for pseudogenic sequences detection in this study.

	Motif sequences	**Reference**
ITS1 motif	GGCry- (4-7n)-GyGyCAAGGAA	Liu and Schardle (1994)
5.8s motif 1	CGATGAAGAACGyAGC	Harpke and Peterson (2008)
5.8s motif 2	GAATTGCAGAAwyC	
5.8s motif 3	TTTGAAyGCA	

2.2.5 Indel coding

A body of evidence supports the benefit of inclusions of coded insertion/deletion (indels) events for phylogenetic reconstruction. Indel characters tend to be less homoplasious than nucleotide characters (Simmons et al., 2001) and including them can dramatically increase the support for the resulting phylogenetic hypotheses by providing additional highly informative characters (Blair & Murphy, 2011; Giribet & Wheeler, 1999; Kawakita et al., 2003; Ogden & Rosenberg, 2007).

A comparative study of approaches to indel coding supported the high reliability of Simple Indel Coding (SIC) and Modified Complex Indel Coding (MCIC) methods (Simmons et al., 2007; Belinky et al., 2010). The study of Belinky et al. (2010) indicated MCIC performed slightly better of than SIC; however, MCIC derives complex metrics from multiple indel events that makes it difficult to identify which individual indel event corresponds to which position of indel in the matrix. Furthermore, MCIC utilizes indels arising from single nucleotide strings, which cannot be distinguished from artifactual gaps generated by PCR error. Since the trnV-ndhC data set contained a number of these indels, the SIC method was used in this study.

Single–residue indels may be highly homoplasious (Belinky et al., 2010). However, when the taxon sampling was increased and closer out-groups used, the true signals of the single-residue may increase (Belinky et al., 2010). To maximise the homologous indel information the number of outgroups was reduced; only *Dubouzetia* was used for the genus level phylogeny. Prior to

this, trees were inferred for the full and reduced datasets for each data partition and the combined data (including only *Elaeocarpus*, *Dubouzetia*, *Aceratium* and *Sericolea* but no indel characters) which confirmed that there is no obvious effect on the internal relationships of reducing the outgroup. Indels were scored using the SIC method in Seqstate (Müller, 2005), and subsequently manually checked. All indels associated with single nucleotide strings were excluded from the analyses.

2.2.6 Phylogenetic analysis

Phylogenetic analyses were undertaken using maximum parsimony and Bayesian inference methods. It is now well understood that gene tree heterogeneity is ubiquitous (Degnan and Rosenberg 2009). Gene tree congruence was assessed by analysing the five data partitions separately: *ndhC-trnV* data only, *trnL-F* data only, plastid data only, nrITS data only, and the combined data. For the family level phylogeny, all taxa of *Aristotelia*, *Sloanea* and *Vallea* included in this study were assigned to outgroups, following the results of Crayn et al. (2006).

2.2.6.1 Parsimony analysis

The data were analysed under un-weighted (Fitch) parsimony using the heuristic tree search algorithm in PAUP 4.08 (Swofford, 2002). The heuristic search parameters were as follows: Tree Bisection Recombination (TBR) branch swapping, MULPARS on, 1000 random-addition-sequence replicates (RAS) saving a maximum of 100 trees per replicate. The most parsimonious trees identified from the first search were used as starting trees for a second search in order to search for shorter trees, using the same parameters but saving a maximum of 20000 trees.

Recent comparative studies suggest that jackknife resampling (JK) methods perform better than bootstrap (BS) methods especially for the supermatrix data set (Freudenstein & Davis, 2010; Simmons, 2011). This is because unlike BS, JK resamples the original dataset without replacement therefore generating replicate datasets that are more closely related to the original dataset. Clade support was therefore assessed by Jackknife resampling. Thus, heuristic searches were performed with 1000 pseudo-replicates, 36.7879 % (following Farris et al. (1996), evaluated by Freudenstein and Davis (2010)) characters deleted per replicate, emulating Jac resampling, 100 RAS, TBR branch swapping, and 10 trees saved per replicate. Jackknife values of between 95% and 100% were herein interpreted as robust support, 80-95% as moderate, and below 80% was considered as not significant (Zander, 2004).

2.2.6.2 Bayesian Analysis

2.2.6.2.1 Model selection

The evolutionary model that best fits the data was determined using the Akaike Information Criterion (AIC) in Jmodeltest V. 0.1.1 (Guindon & Gascuel, 2003; Posada, 2008). Model tests were run on the each data partition. The best fit models under AIC are shown in Table 2.6.

2.2.6.2.1 Settings and parameters

The data were analysed using Mr Bayes 3.2.1 (Ronquist et al., 2011) run on the High Performance Computing unit through the Cyberinfrastructure for Phylogenetic Research (CIPRES) phylogeny portal (Miller, Pfeiffer, & Schwartz, 2010). Two separate runs of four concurrent chains (one cold, three heated each) were applied to each data set, with trees sampled every 1000 generations and other parameters and priors set to default. The length of analyses for each marker and the combined plastid and all-combined datasets are shown in Table 2.6.

It was assumed that stationarity had been reached when the split frequency of both runs fell below 0.01. Tracer v1.5 (Drummond and Rambaut, 2007) was used to inspect the likelihood scores of tree sampling parameters visually. As a result, the first 20% of trees were discarded as burn-in. Bayesian posterior probabilities were calculated from the remaining trees after the burn-in.

Bayesian posterior probability values are generally regarded as inflated (Simmons et al., 2004; Suzuki et al., 2002). Here, posterior probability values of 0.95 or higher were taken to indicate significant clade support (Larget & Simon, 1999).

2.3 Results

2.3.1 Matrix characteristics

This study adds 224 sequences from three markers for the family to those generated by Maynard (2004) and Crayn et al. (2006).

The combined data set consisted of 3113 aligned characters: 2242 cpDNA (*trnL-F*: 1394, *trnV-ndhC*: 848) and 2283 nrITS (871). Of these, 2033 were constant, 468 were variable but parsimony uninformative and 612 parsimony informative. To these totals, 1783, 226 and 263, respectively, were found in cpDNA; and 316, 181 and 374, respectively, were found in nrITS (Table 2.6).

Not all taxa were sequenced for all three markers. Missing taxa totalled 25%, 7% and 27% of the *trnV-ndhC, trnL-F* and nrITS datasets respectively. The overall average of missing taxa is given in Table 2.6.

2.3.2 Detection of potential paralogues (pseudogenes) in nrITS copies

Point mutations were found in the highly conserved motifs in some samples of *Elaeocarpus*, *Dubouzetia, Sericolea* and *Sloanea*. Samples with mutations in one motif are *E. michaelii* 1, *E. gummatus, E. kirtonii, E. reticulatus, Sl. langii* and *Sl. sogerensis* (Table 2.7), in two motifs are *E. foveolatus* 1, *E. costatus, E. elliffii, E. eumundii, E. ferruginiflorus* 2, in three motifs are *Sl. macbrydei and Se. gaultheria,* and in all four motifs are *E. grahamii* and *D. elegans*. Initially these candidate pseudogene sequences were included in the analyses and relationships, branch lengths and support values were observed (Figure 2.1 - 2.4).

All of the candidate pseudogene sequences of *Sloanea*, *Dubouzetia*, *Sericolea* were nested within their prospective genus; and no obviously erroneous relationship was observed (when referenced against previous estimates of relationship based on molecular data and against the plastid based phylogenetic trees obtained here) (Figure 2.4). Only *Sloanea langii* showed a long branch, which is likely to be explained by the effect of using data from a non-functional gene. Within *Elaeocarpus*, however, all the candidate pseudogene sequences except *E. gummatus* formed a strongly supported clade in the nrITS tree (Figure 2.4). This clade exhibits a long branch, was not observed in the plastid gene trees, and does not correspond to any of the infrageneric groups in Coode's (1984) classification. On this basis that these sequences are probable pseudogenes, they were excluded from the nrITS dataset, which totalled 89 terminal taxa (missing taxa = 38%; Table 2.6). Other dataset statistics are given in Table 2.6.

2.3.3 Parsimony analyses

The analysis statistics from the parsimony analysis of the various data partitions are given in Table 2.6. The 50% majority rule consensus of the shortest trees found for the *trnV-ndhC* data is shown in Figure 2.1 for the *trnL-F* data in Figure 2.2, for the nrITS data in Figure 2.5, for the combined data (no indels) in Figure 2.8 and Figure 2.9, and for the combined data with indels (including only *Aceratium*, *Dubouzetia*, *Elaeocarpus*, *Sericolea*) in Figure 2.10.

Table 2.6. Data set information, results of parsimony analysis, nucleotide evolutionary models, and number of generations used for Bayesian inference. PI: parsimony informative; MPT: most parsimonious trees found in the first run; CI: Consistency index (excluding uninformative characters); RI: retention index; RC: rescaled consistency index. Indel characters were excluded from the dataset in the analyses of each marker. Numbers in the brackets were obtained from the dataset when pseudogene sequences were included. GTR: General Time Reversible model. HKY: Hasegawa, Kishino and Yano model. +I: with a proportion of invariable sites. +G: a gamma-shaped distribution of rates across sites.

Data set	Chloroplast DNA			nrITS	Comb.1	Comb. 2
	trnV-ndhC	*trnL-F*	Comb.			Only D. E. Ac. Se.
Terminal taxa	107	134	141	89 (104)	144	89
Align.	848	1394	-	871	3113	3141
indels	22	8	-	20	-	50
Missing taxa	-	-	17%	-	23 % (20%)	22% (18%)
No. PI	95	149	238	312 (374)	550 (612)	280 (344)
% PI	11	11	11	36 (43)	18 (19)	9 (11)
MPT	90400	94800	33200	8700 (32500)	5700 (11700)	1000 (5800)
MPT length	283	380	713	1431 (1741)	2193 (2509)	1009 (1240)
CI	0.80	0.71	0.66	0.45 (0.44)	0.50 (0.55)	0.48 (0.49)
RI	0.92	0.96	0.94	0.81 (0.80)	0.86 (0.84)	0.75 (0.77)
RC	0.73	0.78	0.73	0.42 (0.41)	0.52 (0.57)	0.45 (0.47)
Nucleotide evolutionary model	GTR+G	GTR+G	-	GTR+G+I (HKY +G)	-	-
N of generations (millions)	2.5	2.5	3	7	10	8.5

Table 2.7. Summary of motif sequences. Point mutations found in the conserved motif sequences are highlighted in bold. *undetermined nucleotide bases (n) were also highlighted bold because those were assigned due the unreadable chromatogram caused by the polymorphic locus. Abbreviations for genera are given in Table 2.3.

	ITS 1	Motif 1	Motif 2	Motif 3
	GGCR(g)Y-5n-GYGYCAAGGAA	CGATGAAGAACGY(t)AGC	GAATTGCAGAAW(t)Y(c)C	TTTGAAY(c)GCA
D. elegans	**••n**gc-gatct **•n•••••n•••***	**•n••n••a••••••a•***	**••••••••a••**tc**•**	**•••n••nnnn***
E. costatus	**ta•**gt-gatct- **•g••••••t••**	**•a••••••••••••••**	**•••••a•••••**tc**•**	-
E. elliffii	**t••**gt-gatct- **••••••••t••**	-	-	**••••••**c**•t•**
E. eumundii 2	**t••**gt-gatct- **••••••••t••**	-	**••••••••a••**tc**•**	-
E. ferruginiflorus 2	**t••**gt-gatct **••••••••t••**	-	**•••••at••••**tc**•**	-
E. foveolatus 1	**t••**gt-gatct **n•••••••t••**	**t••••••••••**t**•••**	-	-
E. grahamii	**t••**gt-gatct **a•••••••t••**	**••••••••••t•t•••**	**•••••a•••••**tc**•**	**a•••••**c**••••**
E. gummatus	-	-	**•••••••••g•**tc**•**	-
E. kirtonii	**t••**at-gatct- **••••••••t••**	-	-	-
E. michaelii 1	**t••**gt-gatct- **••••••••t••**	-	-	-
E. reticulatus	**t••**gt-gatct- **••••••••t••**	-	-	-
Se. gaultheria	g**••**gc-gatcc- **•••••rr•••w**	**••••••••••••**t**•a•**	**•••••••a•••**tc**•**	-
Sl. langii	-	**••••••••••••**t**•k•**	-	-
Sl. macbrydei	g**••**gn-natnc- **••c•••n•n•c***	**n••nn••n••n•n••**n*	**n••nnnn•n••**tc**•***	-
Sl. sogrensis	-	**•a••••••••••**t**••••**	-	-

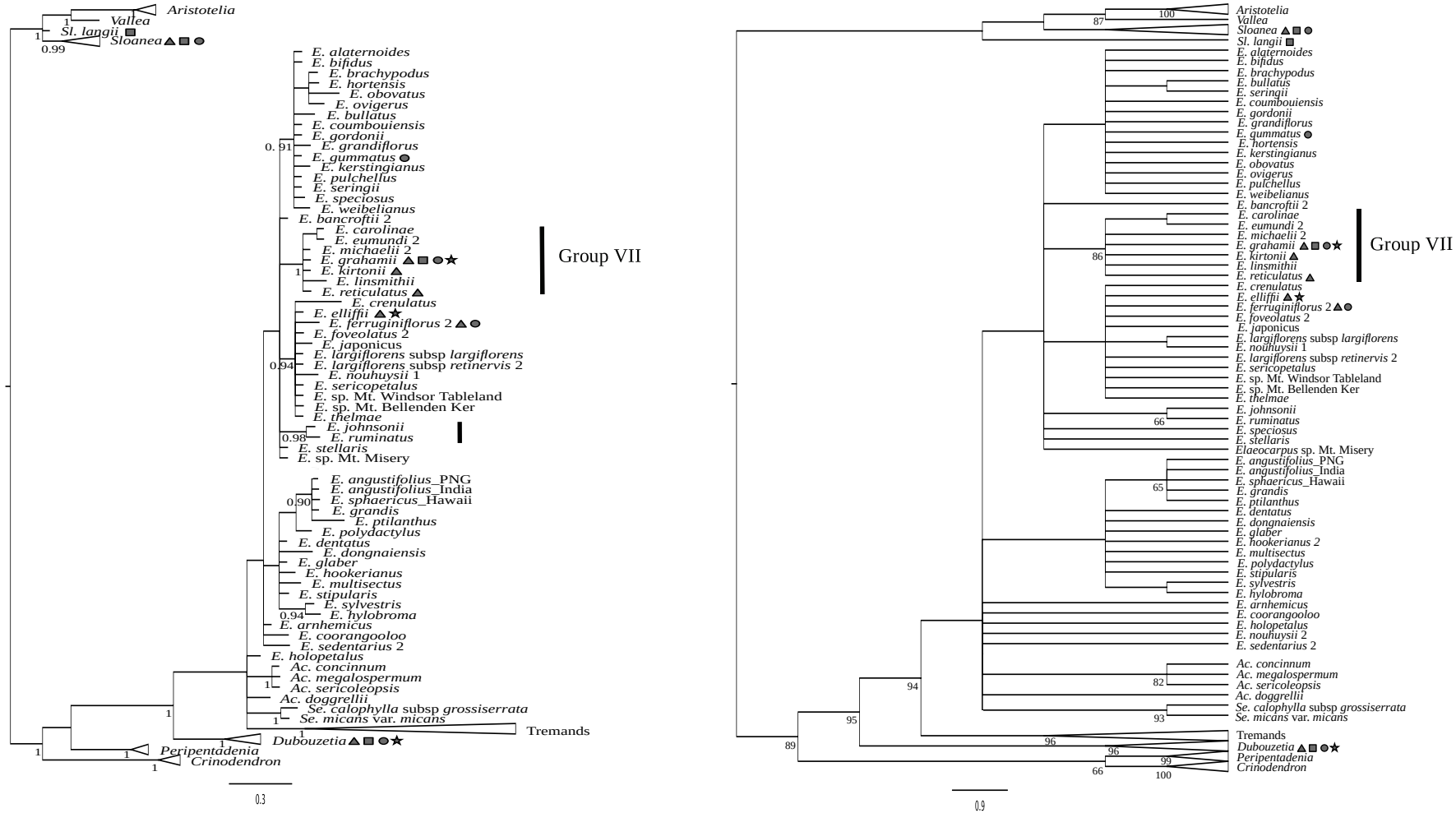


Figure 2.1. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis (right) of the *trnV-ndhC* data. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions (right) greater than 80% shown below the branches. See Table 2.3 for abbreviations for the genera. 'Tremands' consists of *Platytheca, Tetratheca* and *Tremandra*. Symbols indicate the position of the members with one or more point mutations in the nrITS motifs: ❑ 5.8s motif 1, ❍ 5.8s motif 2 ★ 5.8s motif 3 and ▲ ITS 1 motif.

Figure 2.2. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis (right) of the *trnL-F* data. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions (right) greater than 80% shown below the branches. See Table 2.3 for abbreviation for the genera. 'Tremands' consists of *Platytheca*, *Tetratheca* and *Tremandra*. Symbols indicate the position of the members with one or more point mutations in the nrITS motifs: ❑ 5.8s motif 1, ❍ 5.8s motif 2 ★ 5.8s motif 3 and ▲ ITS 1 motif.

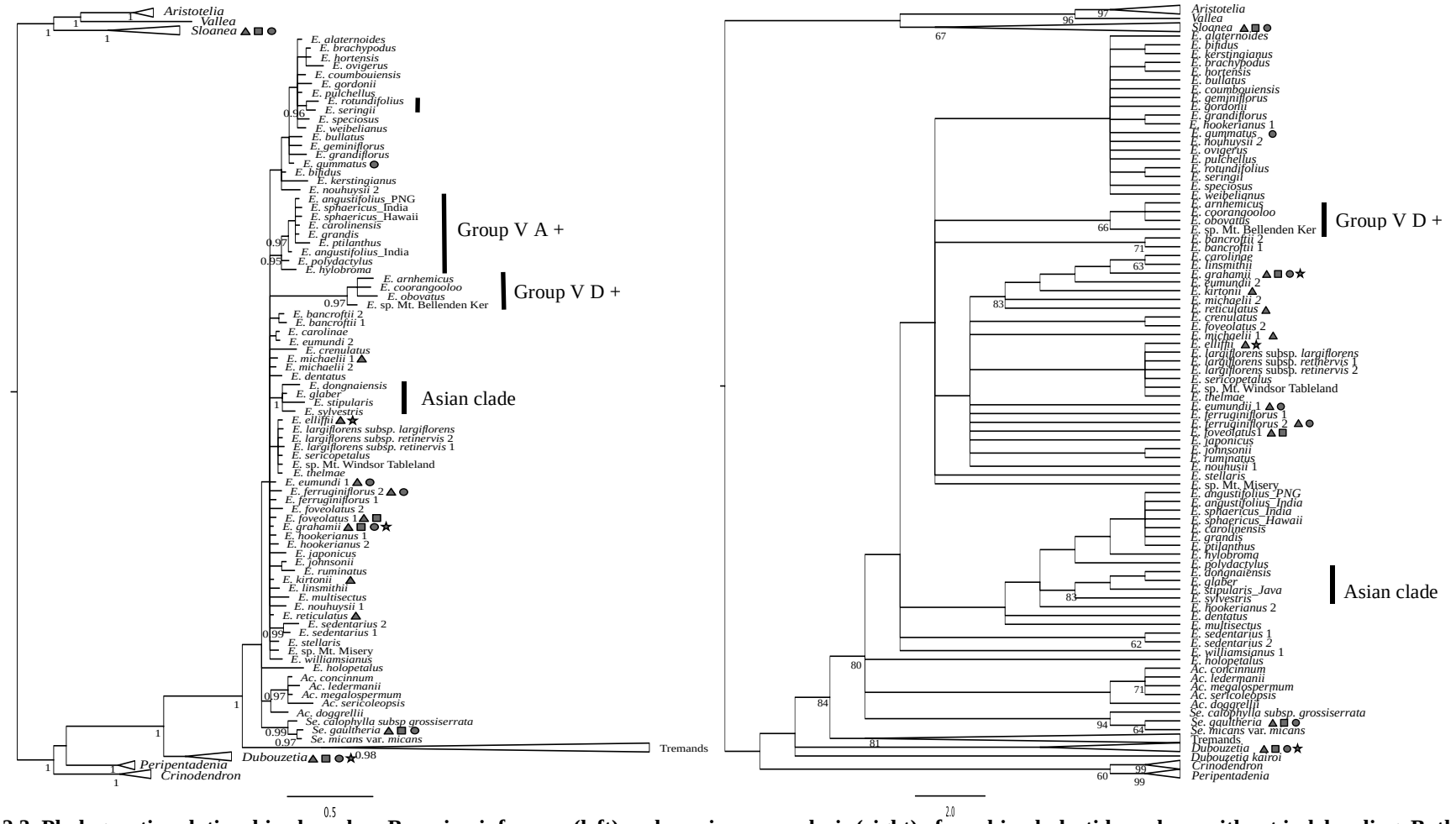


Figure 2.3. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis (right) of combined plastid markers without indel coding. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions (right) greater than 80% shown below the branches. See Table 2.3 for abbreviations for the genera. 'Tremands' consists of *Platytheca*, *Tetratheca* and *Tremandra*. Symbols indicate the position of the members with one or more point mutations in the nrITS motifs: ❑ 5.8s motif 1, ❍ 5.8s motif 2 ★ 5.8s motif 3 and ▲ ITS 1 motif.

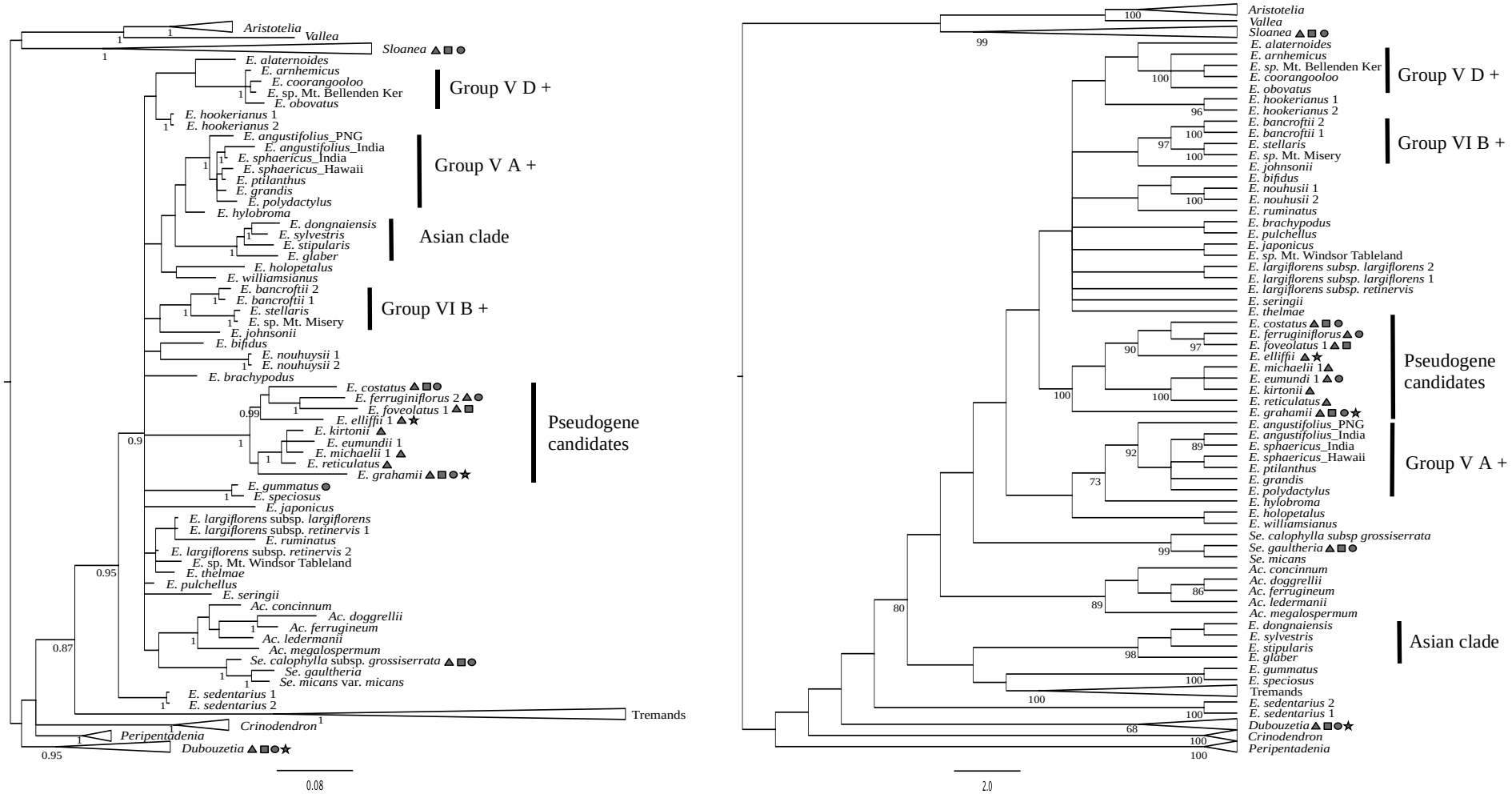


Figure 2.4. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis of the nrITS data with pseudogene candidate sequences. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions greater than 80% (right) shown below the branches. See Table 2.3 for abbreviations for the genera. 'Tremands' consists of *Platytheca*, *Tetratheca* and *Tremandra*. Symbols indicate the position of the members with one or more point mutations in the nrITS motifs: ❑ 5.8s motif 1, ❍ 5.8s motif 2 ★ 5.8s motif 3 and ▲ ITS 1 motif.

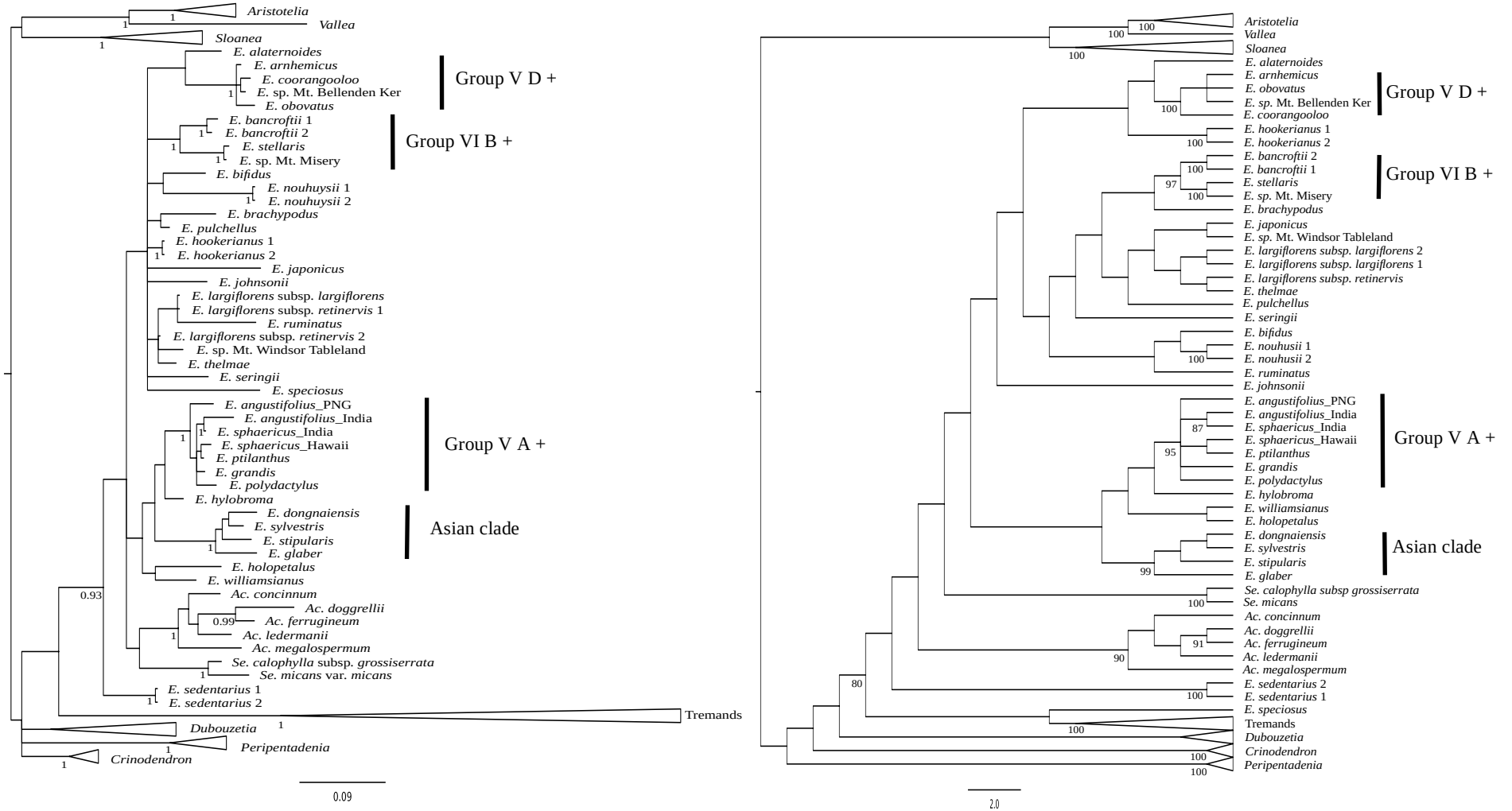


Figure 2.5. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis (right) of the nrITS data without pseudogene sequences. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions (right) greater than 80% shown below the branches. See Table 2.3 for abbreviations for the genera. 'Tremands' consists of *Platytheca*, *Tetratheca* and *Tremandra*.

2.3.4 Bayesian analyses

Runs with four chains each converged with satisfactory mixing and effective sample sizes greater than 200 for all parameters using Mr Bayes. After the burn in process, discarding 20% of saved trees, a topology was recovered in the 50% majority rule consensus tree from the Bayesian analysis.

2.3.5 Tree topology

In order to assess congruence between the results of the Bayesian inference and maximum parsimony analyses, the 50% majority rule consensus trees of each marker were compared. Branch supports of posterior probabilities (PP) resulting from Bayesian inference and Jackknife (JK) proportions obtained for the maximum parsimony analysis are presented in the format PP/JK. The term 'Tremands' is used to denote the group of three genera formerly assigned to the family Tremandraceae: *Platytheca*, *Tetratheca* and *Tremandra*.

2.3.5.1 Plastid dataset

For the *trnV-ndhC* data set, the tree topologies from the two analyses were largely congruent (Figure 2.1). Only one difference in the topology was found. The position of *Crinodendron* was sister to (*Peripentadenia* (*Dubouzetia* (*Tremands* (*Aceratium* + *Sericolea* + *Elaeocarpus*)))) in the Bayesian tree, whereas the relationship of (*Crinodendron* + *Peripentadenia* (Tremands (*Aceratium* + *Sericolea* + *Elaeocarpus*)))) was recovered by the maximum parsimony analysis. Most of the genera were monophyletic and strongly supported (1/100 - 1/87) in both analyses, except *Sloanea*, *Aceratium*, *Sericolea* and *Elaeocarpus*. *Sloanea langii* appeared to be segregated in an unresolved clade with the rest of *Sloanea*, *Aristotelia* and *Vallea*. The clade (Tremands + *Aceratium* + *Sericolea* + *Elaeocarpus*) was strongly supported by Bayesian inference (PP = 1) but only weakly by jackknife analysis (JK = 79); however, the Tremands, *Sericolea* and *Aceratium* were strongly supported by both analyses (1/96, 1/82, and 1/82 respectively). The internal resolution within *Elaeocarpus* was poor, yet two clades were supported. One of these clades corresponds to Group VII (1/86: *E. carolinae* + *E. eumundii* + *E. michaelii* + *E. grahamii* + *E. kirtonii* + *E. linsmithii* + *E. reticulatus*) and the other is a small clade comprising only *E. johnsonii* and *E. ruminatus* (0.98/66). Other clades which gained some but not significant support are Group XIB (0.94/-: *E. elliffii* + *E. ferruginiflorus* + *E. foveolatus* + *E. largiflorens* subsp. *largiflorens* + *E. largiflorens* subsp. *retinervis* + *E. sericopetalus* + *E.* sp. Mt. Windsor Tableland + *E. thelmae*, and samples from other morphological groups) and Group VA (0.90/65: *E. angustifolius* from PNG + *E. angustifolius* from India + *E. sphaericus* from India + *E. sphaericus* from Hawaii + *E. grandis* + *E. ptilanthus*).

In the *trnL-F* trees, the topology between the analyses was largely congruent (Figure 2.2). While the monophyly of *Sloanea*, *Aristotelia* and the Tremands were strongly supported by both analyses, the position of *Crinodendron* and *Peripentadenia* remained unsupported. A position of *Vallea* as sister to *Aristotelia* was strongly supported (1/98). Species of *Dubouzetia* formed a well supported (1/99) polytomy with (*Peripentadenia* + *Crinodendron*) (1/98), and (Tremands + *Aceratium* + *Sericolea* + *Elaeocarpus*). *Sericolea*, *Aceratium* and *Elaeocarpus* formed a large polytomy with the Tremands; the Tremands were strongly supported (1/100).

There is virtually no resolution recovered within the genus, only two clades obtained strong support. One constitutes Group VD + (1/100: part of group VD (*E. arnhemicus* + *E. obovatus*) + *E.* sp. Mt. Bellenden Ker + *E. coorangooloo*), the other comprises samples from Asia (1/86: *E. dongnaiensis* + *E. glaber* + *E. stipularis* + *E. sylvestris*). There are some non-supported clusters formed such as Group VA (0.90/64), Group XIB (-/61), which was also observed in the *trnV-ndhC* trees.

In the combined plastid tree, despite the lack of major conflict in the topology between the two data sets, there was no obvious improvement in the branch support for the clades already resolved in each marker, and no further resolution in the unresolved clades (Figure 2.3). On the other hand, the clade resolved in the *trnV-ndhC* tree, Group VII, was only weakly supported by the maximum parsimony analysis (JK = 83) and *E. holopetalus* did not group with the rest of *Elaeocarpus*, instead forming a polytomy with *Aceratium*, *Sericolea* and *Elaeocarpus*.

2.3.5.2 nrITS dataset

For the nrITS data, the tree topologies from the two analyses generally agreed (Figure 2.4). The major genera were monophyletic with moderate to strong support (1/100 -0.95/68 (*Dubouzetia*)), but the relationships between the clades were not supported. The main difference between the nrITS and combined plastid trees was the sister group of *Elaeocarpus*. The position of *E. sedentarius* as sister to (*Aceratium* + *Sericolea* + *Elaeocarpus*) but without significant support (0.9/-), made *Elaeocarpus* paraphyletic. While an indication of *E. sedentarius* being sister to (*E. gummatus* + *E. speciosus* + Tremands (*E. glaber*+ *E. stipularis* + *E. sylvestris* + *E. dongnaiensis* (*Aceratium* + *Sericolea* + *Elaeocarpus*))) was observed in the tree from parsimony analysis; however, this relationship is most likely explained by long-branch attraction caused by the Tremands clade (Figure 2.4)

While the internal resolution was generally poor in *Elaeocarpus*, four monophyletic groups that formed within the genus were well supported by both analyses. These were the clades of samples consistent with part of Group V D + (1/100), Group V A+ (1/95: Group V A, *E. carolinensis* + *E. polydactylus*), Group VI B+ (1/97: Group VI B (*E. bancroftii* + *E. stellaris*) + *E.* sp. Mt. Misery) and the Asian clade (1/99).

The topologies resulting from the nrITS and plastid datasets agreed but support was slightly higher for the deeper branches in the plastid dataset and for the shallow branches in the nrITS data set. Differences were found in the resolution of the (*Dubouzetia* + *Crinodendron* + *Peripentadenia*) clade and (*Aceratium* + *Sericolea* + *Elaeocarpus*). This difference is likely to be a result of both groups being unresolved in both analyses.

2.3.5.3 Combined

2.3.5.3.1 Family phylogeny

More extensive taxon sampling and a combined analysis of two plastid and nrITS markers resulted in a more robust phylogeny of the family (Figure 2.6) compared with previous studies (e.g. Crayn et al, 2006). Node support and the number of nodes supported by Bayesian inference were greater than that from maximum parsimony. The major genera were supported as monophyletic, generally with moderate to high PP and high to weak JK proportions: *Sloanea* (1/81), *Aristotelia* (1/98), *Crinodendron* (1/99), *Peripentadenia* (1/100), *Tremandra* (0.9/99), *Platytheca* (0.99/85) and *Tetratheca* (1/97*)*. A position of *Vallea* sister to *Aristotelia* was strongly supported (1/96). *Dubouzetia* was recovered in a polytomy (0.91/54*)* but its position sister to (1/83: *Peripentadenia* + *Crinodendron*) and (1/ 79: Tremands + *Aceratium* + *Elaeocarpus*+ *Sericolea*) was moderately supported. A clade comprising *Aristotelia, Sloanea* and *Vallea* was sister to the rest of the family with strong PP and weak JK (81) support (Figure 2.7 and 2.9). The clade was moderately resolved, recovering all species of *Aristotelia* with strong to moderate support (1/100- 1/85). Within the *Sloanea* clade internal resolution was poor; only part of the Mesoamerican samples formed a clade (0.99/ 95).

A Tremands clade was resolved with *Tetratheca*, *Tremandra* and *Platytheca* being monophyletic, and was placed sister with strong support (1/99) to a large clade comprising *Elaeocarpus*, *Aceratium*, *Sericolea* (0.99/-). Within the Tremands clade, all *Tetratheca* (1/85), *Tremandra (*1/84*)* and *Platytheca* (0.97/73) were resolved with high/weak support from the Bayesian/maximum parsimony analyses. Internal clades of *Tetratheca* were well resolved but

supported in only the Bayesian analysis (Figure 2.7); these were, however, congruent with the results of the recent extensive study of *Tetratheca* (McPherson, 2008).

The same combination of clades obtained from the nrITS marker within *Elaeocarpus* were resolved in the combined tree except for Group VI B. These are Group V A + (0.98/66), Group V D + (1/80) and the Asian clade (1/98). A Pacific clade comprising New Caledonian and Hawaiian samples and Group VI B were only recovered in Bayesian (0.95/-), and maximum parsimony analysis (-/84).

2.3.5.3.2 Genus phylogeny

The topology obtained from analysis using the reduced outgroup was identical to the topology obtained using the full outgroup sampling, therefore only the reduced outgroup analysis (with indel characters) is shown (Figure 2.10).

The combined (including indel characters) trees mainly from the Bayesian inference showed increased resolution of some relationships; the monophyly of *Sericolea* was supported with both PP and JK (0.89/93) as sister to *Aceratium* and *Elaeocarpus*. *Aceratium* was placed in the large polytomy with *Elaeocarpus* but the crown node was strongly supported by both analyses (0.95/93). Of the 50 scored indels, one in *trnV-ndhC* indel was found only in *Elaeocarpus*, and two in *trnL-F*, and five in *trnV-ndhC* were shared between *Aceratium, Elaeocarpus and Sericolea*.

Within *Elaeocarpus* there are six clades with support from both or Bayesian analyses only: Group V A + (0.96/87: Group V A + *E. polydactylus* + *E. hylobroma*), Group VI B + (1/95), Group V D+ (1/89), an Asian clade (1.0/97), a small clade comprising *E. rotundifolius* and *E. speciosus* (0.95/-) and Group XI part + (0.95/-: part of Group XI B (*E. elliffii* + *E. largiflorens* subsp. *largiflorens* + *E. largiflorens* subsp. *retinervis* + *E. sericopetalus* + *E. thelmae*) + *E.* sp. Mt Windsor Tableland). Two unique indel characters, one each from *trnL-F* and nrITS, supported Group V subgroup A. Group V D was supported by three nrITS and one *trnL-F* indels. The Asian clade possesses three unique indels: two in nrITS and one in *trnV-ndhC*.

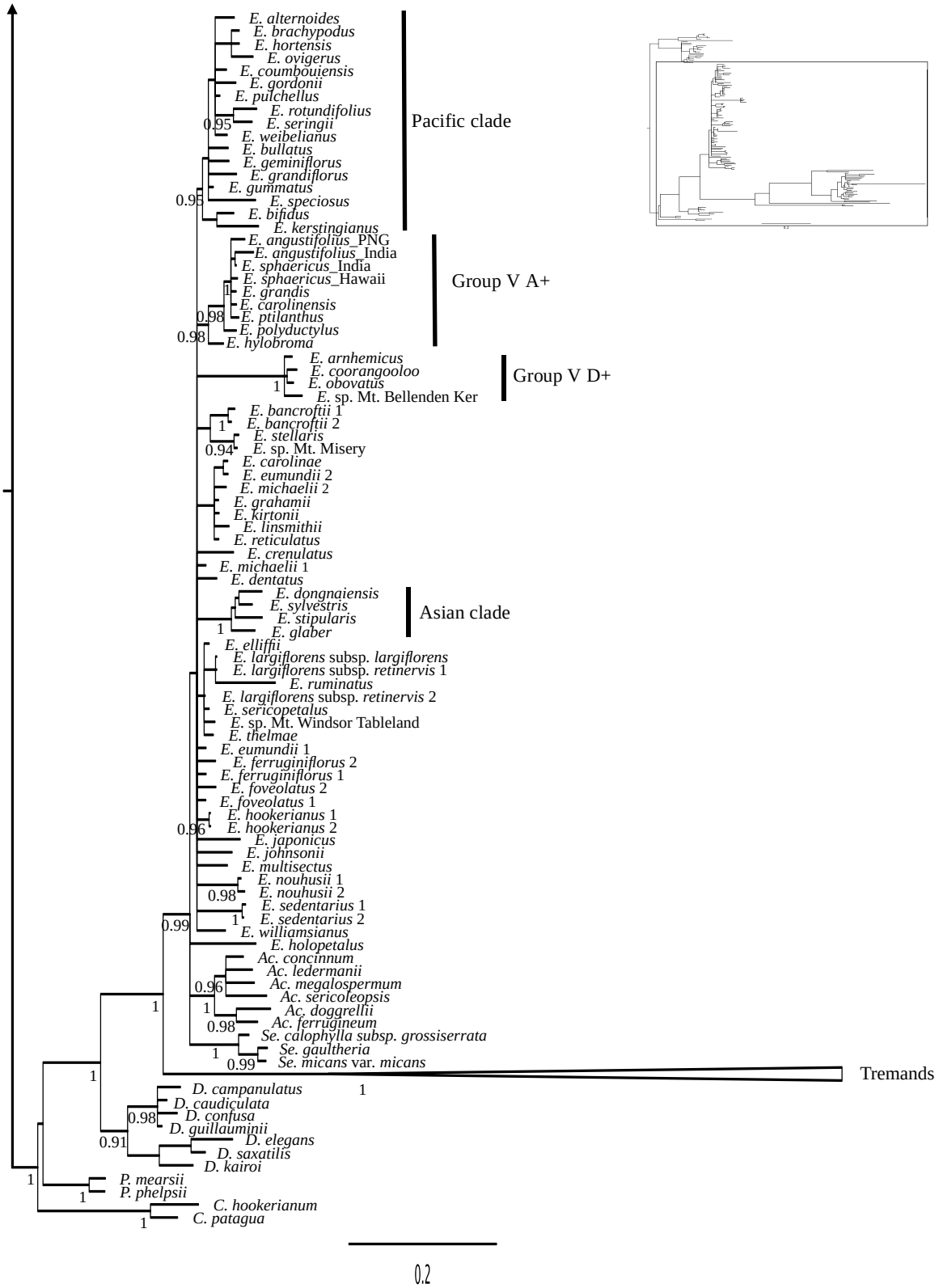


Figure 2.6. Bayesian 50% majority rule consensus phylogram of the combined data set (*trnV-ndhC, trnL-F* and nrITS) without pseudogene sequences and indel characters. Bayesian posterior probabilities are shown below the branches. The positions of the tree are indicated in the inset. See Table 2.3 for abbreviation for the genera. Tremands consists of *Platytheca*, *Tetratheca* and *Tremandra*.

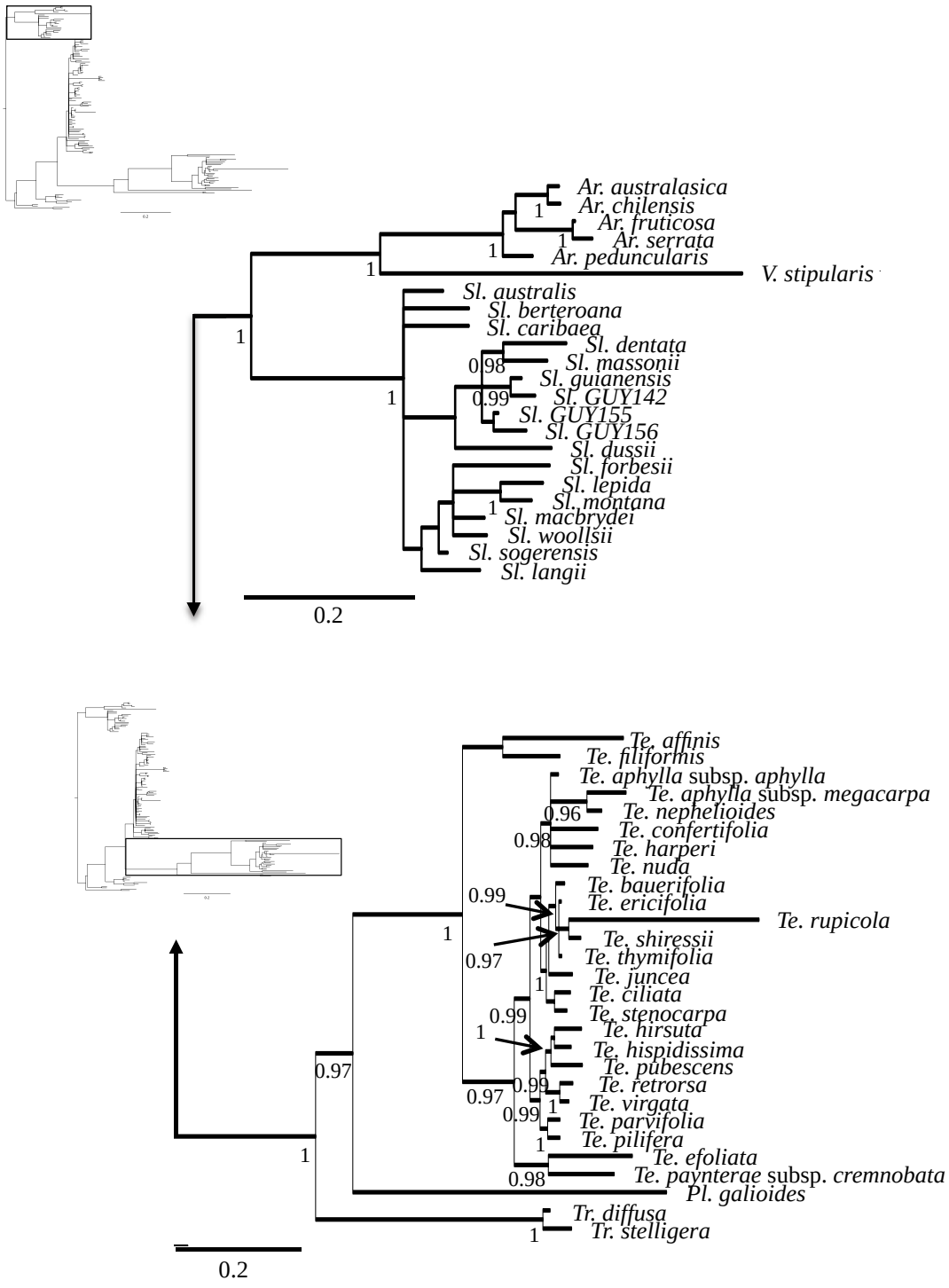


Figure 2.7. Bayesian 50% majority rule consensus phylogram of the combined data set (*trnV-ndhC, trnL-F* and *nrITS*) without pseudogene sequences and indel characters, continued from Figure 2.6. The positions of the tree are indicated in the each inset. Above, the top corner of the tree showing the *Sloanea*, *Aristotelia* and *Vallea* clade. Below, the lower part of the tree describes the Tremands clade. Bayesian posterior probabilities are shown below the branches. See Table 2.3 for abbreviation for the genera.

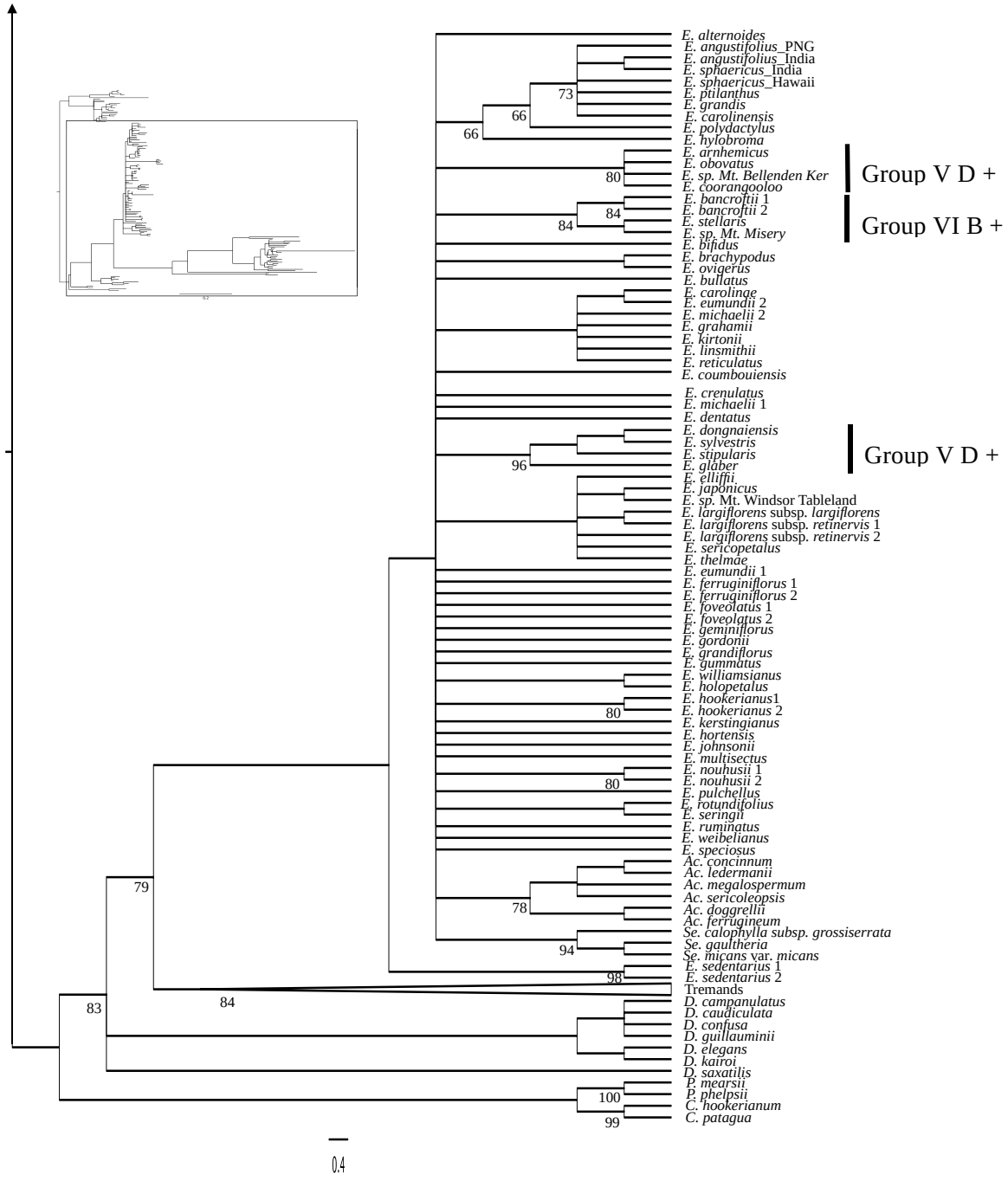


Figure 2.8. Maximum parsimony 50% majority rule consensus cladogram of the combined data set (*trnV-ndhC, trnL-F* and nrITS) without pseudogene sequences and indel characters. Jackknife proportions are shown on the lower branch. The positions of the tree are indicated in the inset. See Table 2.3 for abbreviation for the genera. Tremands consists of *Platytheca, Tetratheca* and *Tremandra*.

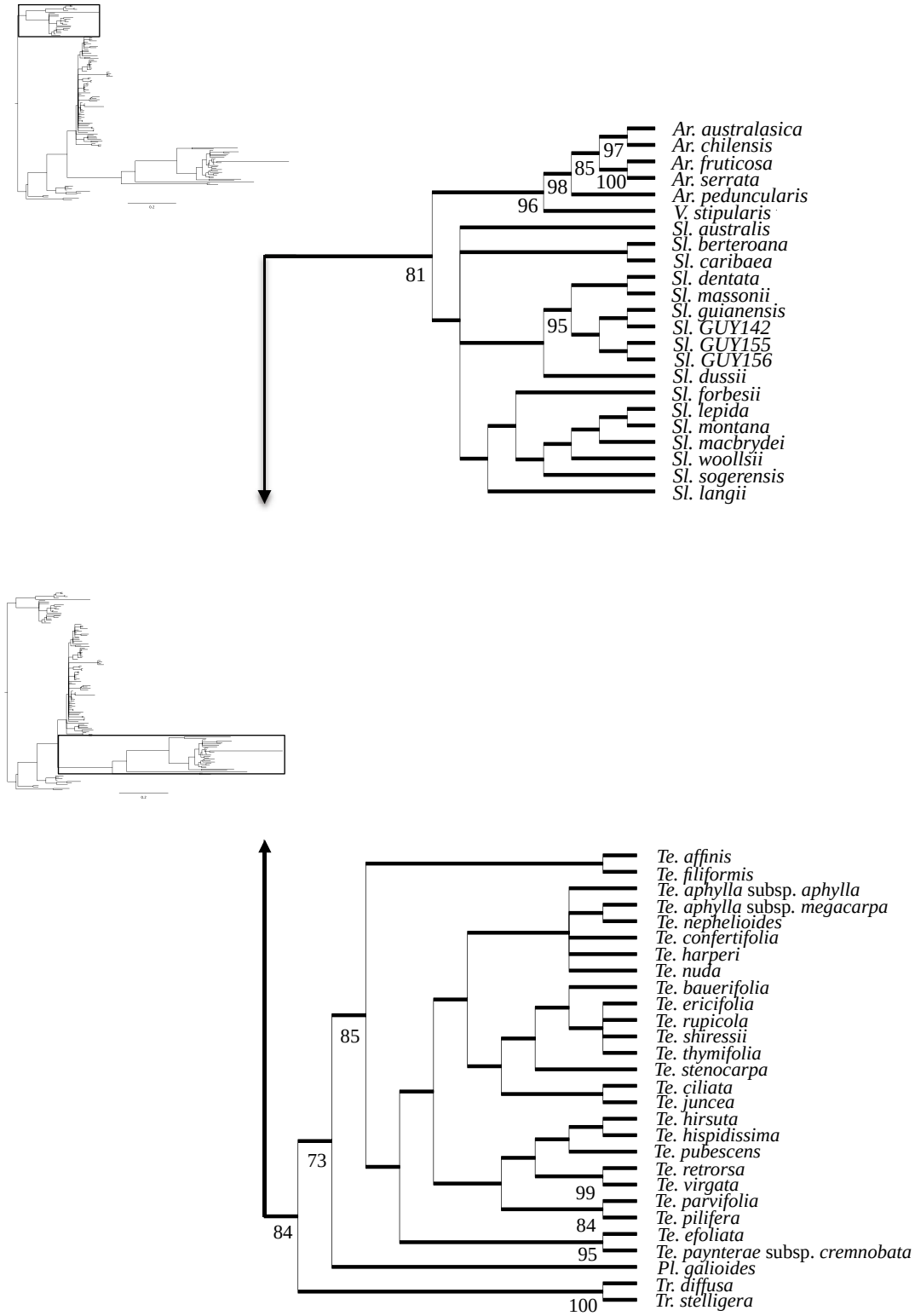


Figure 2.9. Maximum parsimony 50% majority rule consensus cladogram of the combined data set (*trnV-ndhC, trnL-F* and nrITS) without pseudogene sequences and indel characters, continued from Figure 2.8. The positions of the tree are indicated in the each inset. Above, the top corner of the tree showing the *Sloanea*, *Aristotelia* and *Vallea* clade. Below, the lower part of the tree describes the Tremands clade. Jackknife proportions are shown on the lower branch. See Table 2.3 for abbreviation for the genera.

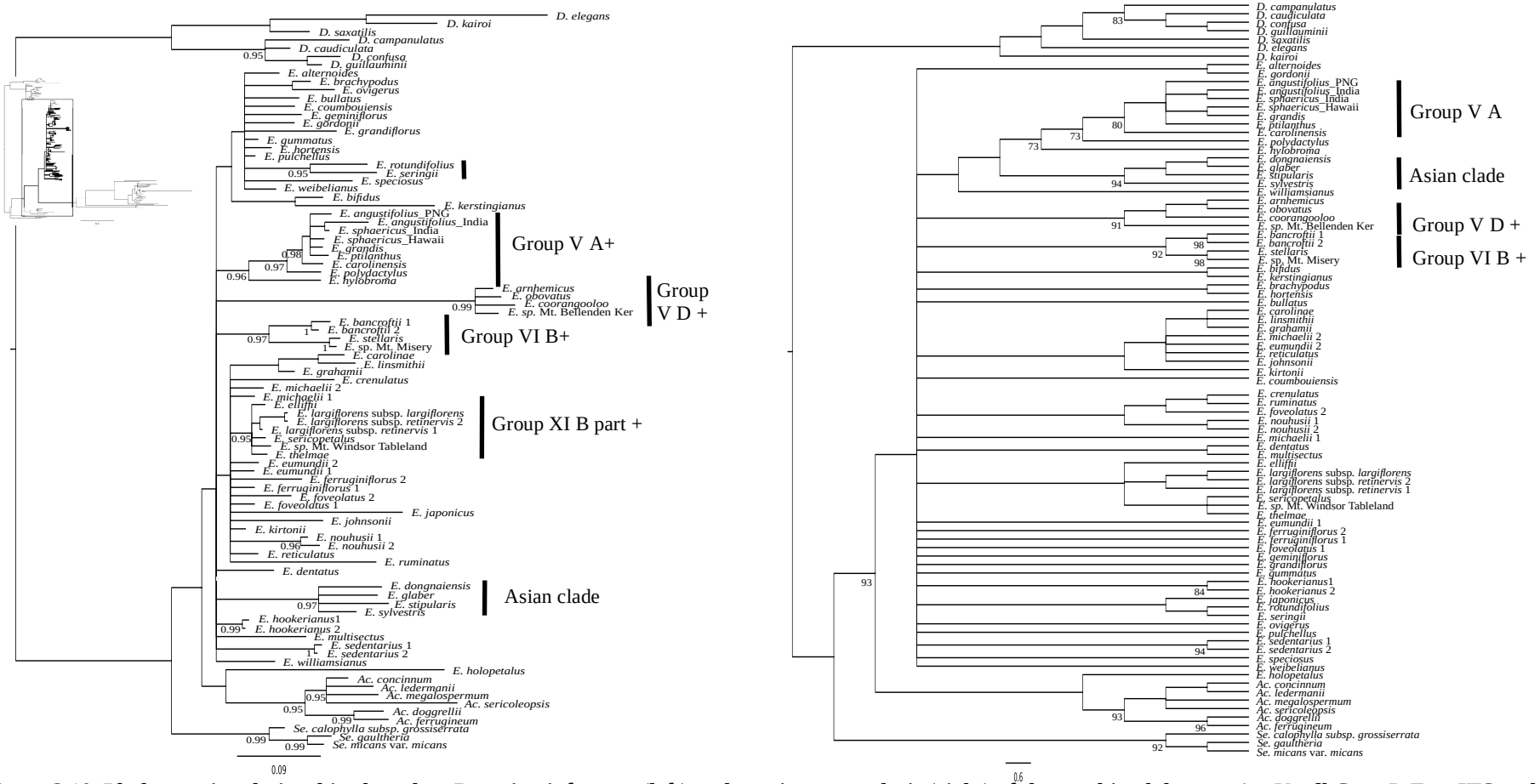


Figure 2.10. Phylogenetic relationships based on Bayesian inference (left) and parsimony analysis (right) of the combined data set (*trnV-ndhC*, *trnL-F*, nrITS and indel characters) without pseudogene sequences. Both trees are 50% majority rule consensus trees with Bayesian posterior probabilities (left) and jackknife proportions (right) greater than 80% are shown below the branches. See Table 2.3 for abbreviations for the genera. The positions of the tree are indicated in the inset.

2.4 Discussion

2.4.1 Marker evolution and their effect on resolution

Each marker used in this study showed a different degree of resolution at different levels of the phylogeny. Chloroplast markers were generally inadequate to resolve shallow nodes but contributed significantly in resolving the deeper nodes.

One of the chloroplast non-coding markers, *trnV-ndhC*, revealed phylogenetic utility for resolving genus level relationships and showed interesting geographical patterns not resolved by analysis of the nrITS data. This may be due to the plastid genome being inherited maternally.

Another plastid marker, *trnL-trnF*, was the least variable marker among the three employed in this study but of all markers it showed most clearly the disparity in molecular evolutionary rate between the Tremands and the other members of the family. This is discussed in more detailed in 2.4.6.

The nrITS, the only nuclear marker included in this study, contributed to the resolution of the tree topology. While deeper nodes at the infrageneric level were more resolved with this marker alone, shallower nodes at the generic level were not successfully resolved.

These markers together were useful in eliciting the tempo of evolution in Elaeocarpaceae lineages in the *Elaeocarpus*+ *Aceratium*+ *Sericolea* clade. There is also a mixed signal between plastid and nuclear markers in regards to species that appear to be sister to the rest of *Elaeocarpus*: *E. holopetalus,* and *E. sedentarius* and *E. speciosus* in plastid and nrITS respectively. Weak signals of each marker in addition to a conflicting dataset, which may be explained as incomplete lineage sorting, contributed to the inconclusive phylogenetic placement of the internal *Elaeocarpus*+ *Aceratium* + *Sericolea* node.

2.4.2 Effect of pseudogenes on phylogeny reconstruction

Attempts at detecting pseudogenes within nrITS sequences using highly conserved ITS1 and 5.8S motives were carried out and 15 candidates from four genera were identified (Table 2.7).
The candidate pseudogenes and PCR failure were prevalent in the Australian Group VII and Group XI, which may indicate the rapid radiation of these groups resulted in failure of homogenization of the nrITS region.

In order to test the utility of the nrITS region in the phylogenetic analysis, detection of orthologous genes is essential. Identified pseudogene candidates in this study may require validation with a different type of analysis, such as analysis of the minimum free energy of ITS2 secondary structures (Feliner & Rossello, 2007; Harpke & Peterson, 2008a, 2008b), because it is not possible to cross reference the phylogeny attained with the data from the ITS region with that from the chloroplast region due to the low resolution of the chloroplast phylogeny from this study.

2.4.3 Bayesian inference vs maximum parsimony analyses

Significant differences in clade support and the number of nodes supported between Bayesian inference and maximum parsimony analysis of the nrITS and the combined dataset are noteworthy. The consistency index for the nrITS and combined data sets is low at 0.42-0.52. This fact highlights that the nucleotide sequences contain highly homoplasious characters that cannot be explained with the assumption that the fewest possible evolutionary steps represent the 'true' evolutionary relationships. Bayesian inference with the general time reversible model, the best nucleotide evolutionary model identified for the nrITS regions, provided better-resolved trees than maximum parsimony analysis. It can be postulated that there may be significant saturation of mutations, which violates the assumptions of maximum parsimony analyses. Therefore, in this section the results from the Bayesian inferences were considered more robust than those from the maximum parsimony analyses, and clades that received significant PP were mainly discussed.

2.4.4 Relationships of genera within the family

Within the family, the monophyly of the major genera was supported and hypotheses of relationships between them suggested by previous studies (Crayn et al., 2006; Maynard, 2004) were supported. The monophyly of *Elaeocarpus* was not supported by either maximum parsimony or Bayesian analysis of any data set. The nrITS and *trnV-ndhC* and combined trees, and both family and genus phylogeny independently indicated segregations of different combinations of species from the rest of *Elaeocarpus*. However there is no significant support for the position of *E. holopetalus* sister to the rest of *Elaeocarpus* + *Aceratium* + *Sericolea*, which was persistently resolved in the *trnV-ndhC* and combined (including indels) tree by Bayesian analysis. Moreover, even though there was no statistical support, both analyses of nrITS data indicated *E. sedentarius* and the other strongly supported clade (the Asian clade) were together sister to the rest of the genus. These suggest the relationships between *Elaeocarpus*, *Aceratium* and *Sericolea* are close and the markers included in this study do not

show sufficient variation to resolve them. The evidence presented in this study leaves open the possibility that *Elaeocarpus* is not monophyletic.

2.4.5 Lineages within *Elaeocarpus*

The lineages supported in the phylogenetic analyses in this study show some congruence with morphology and geography. These lineages are discussed below.

2.4.5.1 Lineages congruent with morphological groups

2.4.5.1.1 Group V

The members of this group are characterised by flower buds narrowly ovoid, stamens lacking awns (but sometimes with setae), discs hairy and embryos broad and straight (Table 2.2).
The monophyly of Group V as a whole was not supported in any tree; the monophyly of some subgroups, however, was supported in the combined trees. Relationships between these subgroups were not clear from the current study.

Subgroup A - Ganitrus Group- exhibits a set of distinctive features in the genus, namely, petals highly divided, loculi (3-) 5-7, ovules 4-6 per locules, fruit spheroid to ellipsoid with a deeply sculptured woody stone (Table 2.2). The samples of this subgroup included in this study are *E. angustifolius* (India and PNG), *E. sphaericus* (India, Hawaii and PNG), *E. grandis* (Australia) and *E. ptilanthus* (PNG).

Coode (Coode, 2010) has provided a detailed classification of this group, however the boundaries of the subgroups of Group V are not yet well reflected in the phylogenetic trees. In this study, subgroup A formed a very strongly supported clade together with *E. polydactylus* (Subgroup D) and *E. hylobroma* (unassigned to Subgroup). The Hawaiian species, *E. carolinensis*, not yet classified in Coode's scheme, was nested in this group with strong support. The relationships of *E. polydactylus* are discussed under Subgroup D below.

Group V subgroup A formed a strongly supported clade whose species are distributed across various geographical areas from Australia to India. This group shares distinct morphological characters and is the best example within *Elaeocarpus* of a morphology-based infrageneric grouping supported by the molecular phylogenetic results.

Subgroup D- Fissipetalum group- This group is defined by plants with locules 2-5, ovules 4-8 per loculus, and fruit globose to bluntly ellipsoid. Coode (2010) distinguished subgroup D from

subgroup A on the basis of its smaller petals (up to 6mm long cf. > 9 mm) and fewer stamens (20- {-25} cf. 30 or more). Samples of this subgroup included in this study are *E. arnhemicus*, *E. obovatus*, *E. hookerianus*, *E. dentatus*, and *E. polydactylus*. Coode (1978) questioned the position of Australian members of subgroup D and in the present study these taxa were nested in two groups: *E. polydactylus* was placed sister to subgroup A, and the Australian taxa *E. arnhemicus* and *E. obovatus* formed a clade (with *E. coorangooloo* and *E.* sp. Mt Bellenden Ker) in *trnL-F*, nrITS and combined trees. The relationships *of E. dentatus* and *E. hookerianus* were not resolved by the analyses in this study and therefore are not discussed further. The entities grouped in the *E. obovatus* clade form a species complex of morphologically similar and poorly differentiated entities, investigation of which is the subject of the following chapters of this thesis. Given this clade was strongly supported in all of the analyses (except *trnV-ndhC* tree), exhibiting exceptionally long branch length and supported by four indel synapomorphies, it may be appropriate to assign it to its own subgroup. This clade is also distinguished by a combination of morphological characters: petals divided, ovary hairy, anther tips blunt, locules 2, 4- 6 ovules per locule, endosperm entire and embryo straight (Table 2.2). There needs to be further investigation of the morphology and a revision to expand the circumscription will be required. Coode (1984) placed *E. coorangooloo* in Group VI whereas the current study positions *E. coorangooloo* in the Australian Group V D.

2.4.5.1.2 Group VI Subgroup B

This group can be defined by members possessing the following characters: flowers large (exceeding 2.5 cm in diameter), anthers awned, locules 2-5, ovules (4-) 6-12 per loculus, and fruit large (c. 6 cm) ellipsoid, or plated (flat-sided) (Table 2.2). Samples from this group included in this study were *E. bancroftii*, *E. stellaris*, *E. williamsianus* and *E. coorangooloo* (Coode, 1984).

Elaeocarpus bancroftii and *E. stellaris* together with *E.* sp. Mt. Misery formed a strongly supported clade in both analyses of nrITS and combined with indel characters (Figure 2.10). The position of *E. williamsianus* was not resolved due to the low resolution within the genus, therefore it is not possible to infer relationships of this taxon in Group VI. *Elaeocarpus coorangooloo* was nested in a stable clade, Group V D +, that persisted in almost all the dataset in both analyses. These results suggest *E. coorangooloo* should be removed from Group VI.

2.4.5.1.3 Group XI Subgroup B

This Australian endemic group is defined by members possessing the following characters: petals thickened and undivided, with 5-6 irregular teeth, anthers blunt, locules 3, ovules 6 per loculus, and, endosperm ruminated, stone ellipsoid smooth (except *E. elliffii*) (Table 2.2). Additionally, they share two further characters: not strongly differentiated petals and sepals and formation of a 'tight ball of short filamented anthers' (Coode, 1984). Samples from this group included in this study were *E. elliffii, E. ferruginiflorus, E. foveolatus, E. largiflorens* subsp. *largiflorens, E. largiflorens* subsp. *retinervis, E. sericopetalus* and *E. thelmae* (Coode, 1984).

The Bayesian analysis of the combined data with indel characters supported a clade (PP = 0.95) consisting of *E. elliffii, E. largiflorens, E. sericopetalus, E. thelmae* and an undescribed taxon, *E*. sp. Mt. Windsor Tableland. Although not well supported, a cluster containing these members was recurrent in analyses of all the data sets. Exclusion of *E. foveolatus* and *E. ferruginiflorus* from this clade may be due to the weakness of phylogenetic signal in the current study, or may reflect a natural segregation from the rest of the group. Further investigation to include additional and faster evolving molecular markers is crucial to further understanding the relationships of this group.

2.4.5.2 Lineages congruent with geographical areas

2.4.5.2.1 Pacific clade

A Pacific clade comprising samples from New Caledonia and Hawaii (with the exception of *E. grandiflorus* from Asia) was strongly supported only in the combined data set by Bayesian inference, but relationships within the clade were not resolved. Members of this clade exhibit diverse morphologies and represent at least three infrageneric groups: Group I (*E. rotundifolius, E. seringii, E. pulchellus* and *E. alaternoides*), Group IV (*E. weibelianus*) and Group VI (*E. ovigerus*, *E. hortensis* and *E. brachypodus*) and species that are unassigned to any existing group (*E. kerstingianus, E. bifidus, E. bullatus, E. coumbouiensis*, *E. gordonii, E. geminiflorus*, *E. speciosus*, and *E. grandiflorus*) (Tirel, 1982). The current level of resolution recovered from this study does not allow inferring whether or not the New Caledonian and Australian Group VI and Group IV are monophyletic. Further research in this clade is also required.

2.4.5.2.2 Asian clade

The Asian clade, comprising *E. stipularis* (Group III), *E. glaber* (Group III), *E. dongnaiensis* (group unassigned), and *E. sylvestris* (group unassigned) is strongly supported. Taxon sampling

of Asian species in this study is sparse and therefore it is premature to draw conclusions on the classification of this clade from the phylogeny.

2.4.6 Molecular evolutionary rate variation

In Elaeocarpaceae, there is a stark difference in the molecular evolutionary rates (branch lengths) between the tremands, which are exclusively small shrubs, and remainder of the family which are trees or large shrubs (see Appendix 1). This is most obvious in the *trnL-trnF* data, a marker which is widely used for resolving phylogenetic relationships at generic and lower taxon levels (Eriksson et al., 2003; Guo et al., 2004; Wang et al., 2005; Wurdack, et al., 2005; Soejima & Wen, 2006; Pirie et al., 2007 and many more) perhaps more successfully in shrub and herb groups than tree groups (Eriksson et al., 2003; Guo et al., 2004; Wang et al., 2005; Wurdack, et al., 2005; Soejima & Wen, 2006). Long-lived trees are deemed to have experienced low nucleotide substitution rates and low speciation rates per unit time (Petit & Hampe, 2006) compared with other life forms. This is true of Elaeocarpaceae: some tremands may reach maturity within 9 months (Crayn et al., 2004), whereas it is unlikely that any of the tree species in the family would exhibit a generation time even remotely approaching this duration. However, the underlying causes of this observed correlation between molecular evolutionary rate and generation time are still incompletely known (Whittle & Johnston, 2003; Petit & Hampe, 2006; Smith & Donoghue, 2008).

2.5 Further research

There is a strong need for further data to help resolve shallow phylogenetic relationships in and among the closely related genera *Elaeocarpus, Aceratium* and *Sericolea*. Data from a low copy nuclear marker would be beneficial to augment the nrITS data. Additionally, more plastid markers, including the potentially informative *petB-rpoA* (Aoki et al. 2004) and *trnH-psbA* (this study, see materials and methods) should be explored. More thorough taxon sampling from the other geographical regions and groups that are not represented by Australian species will help advance knowledge of relationships.

In this study, nrITS contributed to the highest resolution in the shallow parts of the phylogeny. However, the success rate of amplification of this marker is low due to the presence of multiple PCR products. Cloning those multiple copies of nrITS may be one way to circumvent this problem. Employing conservative motifs for detecting potential paralogous copies proved to be a useful tool in this study. This screening method is straightforward and effective, and therefore would be easily applied to cloned sequences to identify gene orthologs.

The phylogenetic trees were generally poorly resolved with the markers used. Once a greater internal resolution is achieved, mapping of morphological characters will be possible to investigate their evolution and to determine possible synapomorphies for the molecular groups.

2.6 Conclusions

Molecular phylogenetic studies have revealed some significant relationships within Elaeocarpaceae and *Elaeocarpus.*

1. Results from this study, based on a significantly larger taxon sampling compared with previous studies, provide a better resolution within the family and genus and largely agree with previously hypothesised phylogenetic relationships. This study also provides the first evidence that *Sericolea* is well-supported monophyletic lineage and should be maintained at genus level. *Elaeocarpus* appears to be paraphyletic, with *E. holopetalus* sister to *Aceratium*.

2. Four highly supported clades were congruent with the existing infrageneric classification. These are: Group V Subgroup A + (but including two species previously unassigned), Australian Group V subgroup D + (the *E. obovatus* species complex and *E. coorangooloo*), Australian Group VI B excluding *E. williamsianus* and Group XI B (part). While each of these groups is congruent with the existing infrageneric classification, a more thorough study incorporating additional markers and broader sampling is required to establish the boundaries of the groups and the relationships between them. Two clades, each reflecting geographic groupings, were strongly supported. These are the Pacific and the Asian clades. The Pacific clade contains rich morphological diversity. For the Asian clade, limited taxon sampling prevented the determination of groupings within.

3. Australian Group V subgroup D +, consisting of the *E. obovatus* species complex and *E. coorangooloo,* was recovered as a monophyletic group. The results of population genetic and morphometric studies of the *E. obovatus* species complex are reported in Chapters 3 and 4 respectively.

4. The positions of two putative new Australian species were confirmed. *E.* sp. Mt. Misery was nested in Group VI Subgroup B and *E.* sp. Mt. Windsor Tableland was placed in

Group XI Subgroup B. A third species, the recently described *E. hylobroma*, formed a clade with members of Group V Subgroup A.

Chapter 3 Resolving taxon limits in the *Elaeocarpus obovatus* species complex (Elaeocarpaceae) – a population genetic approach.

3.1 Introduction

Species are shaped by their evolutionary and geographical history (James & Moritz, 2000). Understanding the genetic structure, diversity and demography is the first step to understanding the nature of contemporary species and thus the most appropriate taxonomic circumscription.

The interaction of geomorphological features with climatic oscillations has been well documented in the Wet Tropics of North Queensland (Graham et al., 2010; VanDerWal et al., 2009). These periods of rainforest expansion and contraction coincided with the glacial and inter-glacial cycles as they alternated in North Queensland and resulted in the frequent extirpation of lowland rainforest, whereas upland rainforests remained stable and likely functioned as refugia (VanDerWal et al., 2009). These refugia may have acted as reservoirs, influencing regional biodiversity through repeated interaction during the alternating glacial cycles (VanDerWal et al., 2009). These climate fluctuations are likely to have played a key role in shaping the genetic patterns found within rainforest species and in the speciation processes of contemporary species. Such genetic divergence induced by climatic fluctuations has been observed in rainforest skinks (Dolman & Moritz, 2006) and some *Elaeocarpus* taxa (Rossetto et al, 2009).

3.1.1 The *E. obovatus* species complex

A ‘species complex’ for the purpose of this study is defined as a morphologically closely related group of individuals that involves two or more species. The *Elaeocarpus obovatus* species complex includes three morphological entities that can usually be distinguished on floral characters and leaf shape (Coode, 1984): *E. obovatus*, *E. arnhemicus* and an as yet undescribed taxon, *E.* sp. Mt. Bellenden Ker (L.J.Brass 18336). However, this group was last revised in 1984 and many new collections have been obtained since then, and these blur the boundaries between the entities.

One of the major problems in the group is that the taxonomic status of *Elaeocarpus* sp. Mt. Bellenden Ker has never been thoroughly assessed. This entity occurs mainly in upland and rarely lowland rainforests in the Australian Wet Tropics bioregion (Guymer, 1997; 2002; 2007;

2010), including in areas identified by palaeoclimatic modelling as past refugia for rainforest taxa (VanDerWal et al., 2009). The rare occurrence of this taxon in the lowlands of the Daintree area also coincides with the location of an identified refugium (Graham et al., 2010; VanDerWal et al., 2009). While *E.* sp. Mt. Bellenden Ker occurs in typical *Elaeocarpus* habitats (i.e. well developed, often upland rainforests), *E. arnhemicus* and *E. obovatus* occur in atypical habitats for the genus. *Elaeocarpus arnhemicus* occurs in monsoon forest in the seasonally arid zone of Cape York Peninsula in Australia, savannas in Papua New Guinea and West Papua, and the ecotone between limestone forest and open natural grassland in East Timor, Flores Island, Sumba Island and Central and West Java (Coode, 1978). *Elaeocarpus obovatus* probably exhibits the most diverse habitat preferences, ranging from wet sclerophyll forests to lowland/mid altitude drier rainforests, and littoral rainforests in the eastern coast of Australia.

Phylogenetic analyses confirmed that the *E. obovatus* species complex is monophyletic (Chapter 2). However, relationships within the group are yet to be resolved. The degree of mutation and support for the clade is exceptionally high for the genus (the clade is strongly supported in the multigene phylogeny with a stem branch length at least twice that of the other clades in the genus). Comparing the genetic diversity and distance between rainforest and dry adapted taxa provides an opportunity to investigate whether adaptation to novel environments influenced the diversification of this species.

The role of taxonomy in conservation biology is paramount, as studies in taxonomy and ecology provide fundamental knowledge of biodiversity (Kim & Byrne, 2006). Inadequate knowledge of the number and the delineation of species can hamper effective conservation. An interesting example is the North Queensland taxon *Elaeocarpus* sp. Mt. Bellenden Ker. This taxon is represented by a small number of populations, each of them containing often a small number of individuals. Understanding the taxon limits, genetic structure and relationship to other populations and taxa, is fundamental information required for the effective management of this taxon.

There is growing realization that resolving controversial taxonomic groups based only on morphological characters that may result from convergent evolution risks misrepresenting evolutionary scenarios for the species, therefore the morphological characters are often tested in the light of molecular evidence (Bardy et al., 2010; Schönswetter et al., 2009; Stefanović et al., 2008). It is also noteworthy that the increase in the number of re-circumscriptions and discovery of new species and varieties in the last three decades is the result of enhanced and developing concepts of multidisciplinary approaches in taxonomy (Domínguez Lozano et al., 2007).

Microsatellite markers are widely used to assess intraspecific taxonomic boundaries due to their high levels of polymorphism. Microsatellite markers are co-dominant, and can provide deep insights into gene flow, fitness and genetic distance within species. In addition to the taxonomic applications, these markers are now widely applied in ecology and conservation biology (e.g. Catania et al., 2008; Cerón-Souza et al., 2012; Clarke et al., 2012; Harata et al., 2012; Jørgensen et al., 2008; Piotti et al., 2012; Sampson & Byrne, 2012; Thurlby et al., 2012).

The aim of this chapter is to investigate the patterns of genetic differentiation within the *E. obovatus* species complex based on microsatellite data. In particular, I investigated:

1. the genetic diversity within and amongst morphological entities; and
2. the correlation between genetic clusters and conventional species boundaries within the *E. obovatus* species complex.

3.2 Materials and Methods

3.2.1 Microsatellite analysis

3.2.1.1 Plant material

One hundred and fifty samples representing three putative entities: *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker and *E. coorangooloo*, a taxon newly identified as closely related to the group (Chapter 2) were collected from 30 sites in Cape York Peninsula, and Central and South-East Queensland (Table 3.1). Additional samples from northern New South Wales and northern Queensland were donated by the Royal Botanic Gardens and Domain Trust (NSW). In total 181 samples were included in the study (Table 3.1). Two populations collected from Davies Creek National Park, which exhibit morphological affinity to *E. arnhemicus* and *E.* sp. Mt. Bellenden Ker are coded as ADC and BDC. They are treated separately from *E. arnhemicus* and *E.* sp. Mt. Bellenden Ker because of their uncertain identity. The sister of the group is unknown because phylogenetic relationships among the Australian species are still poorly understood (Chapter 2). The sampling strategy was based on distributions estimated from herbarium records and records from previous literature, and aimed to capture as much of the natural range of the entities as possible, and to sample intensively in areas of sympatry. The range of *Elaeocarpus arnhemicus* extends beyond northern Queensland, the Northern Territory, and Torres Strait, into New Guinea and eastern Indonesia (Figure 3.1). However material from outside of Australia was unavailable. Plant material was collected from 1-20 individuals per population, with each individual vouchered at the Australian Tropical Herbarium (CNS).

3.2.1.2 DNA isolation

Genomic DNA was extracted at the Australian Genomic Research Facility (AGRF) and quantified using either Picogreen (AGRF, Adelaide, South Australia) or a NanoDrop 2000/2000c Spectrophotometer (Thermo Scientific, Waltham, USA) and normalised to c. 5 ng /μL prior to PCR.

3.2.1.3 Microsatellite assay

Nuclear microsatellite markers were used to assess species boundaries within the *E. obovatus* species complex. Previous studies developed a set of markers for *E. grandis* and applied these to phylogeographical and conservation genetics questions (Jones et al., 2002; Rossetto et al., 2009; Rossetto et al., 2004; Rossetto et al., 2007; Rossetto et al., 2008). Of these markers, ten (*Scu01Eg*, *Scu20Eg*, *Scu21Eg*, *Scu22Eg*, *Scu25Eg*, *Scu27Eg*, *Scu31Eg*, *Scu32Eg*, *Scu33Eg* and *Scu34Eg*) which are known to amplify in *E.* sp. Mt. Bellenden Ker (M. Rossetto pers. comm.) were assessed further for amplification success and variation within the group. The primer sets used in this study are provided in Table 3.2.

The M-13 linkers were attached to each of the forward primers following the method provided by Schuelke (2000). The M-13 linkers and the fluorescence labels FAM were acquired from Geneworks. The fluorescence labels NED and VIC were obtained from Applied Biosystems.

Two samples of each entity were selected for a pilot study. The aim of this pilot study was to establish and optimise PCR conditions and primer sets and to check fluorescence combinations for multiplexing and pooling. PCR conditions established from the pilot study were: 1 cycle of 95 ^{o}C for 10 m; followed by 30 cycles of 94 ^{o}C for 30 s, optimised temperature for 45 s, 72 ^{o}C for 45 s; 8 cycles of 94 ^{o}C for 30 s, 53 ^{o}C for 90 s, and 72 ^{o}C for 60 s, followed by 1 cycle of 60 ^{o}C for 30 m and 25 ^{o}C for 60 s. Optimised PCR conditions with the specific annealing temperature for each marker are given in Table 3.2. PCR reactions (10 μL volume) contained: c. 5ng of template DNA, 1 μL of 10x ImmoBuffer (Bioline, Immolase), 3mM of MgCL2 (Bioline, Immolase), 0.5mM of DNTPs (Kappa), 0.5 μL of 0.4% Bovine Serum Albumin (BSA), 0.05 μM of forward primer (Geneworks), 0.2 μM of reverse primer (Geneworks), 0.1μM of M13 linker with fluorescence dye and 0.2 unit of Polymerase Taq (Bioline, Immolase). PCR success was determined by agrose gel electrophoresis (Cleaver Scientific Ltd., 2% agarose, 80V, 350A for 30 minutes). PCR products were visualised using EZ-VISION (AMRESCO) and sized using Hyperladder II (Bioline).

After the combination of markers and PCR conditions were established, the rest of the samples were assayed. Because multiplexing of the PCR failed, all the PCR products were assayed as single reactions in this study. The PCR products were pooled (0.33-0.5 μL each in a total 10 μL) and the capillary separation was performed by the Australian Genome Research Facility (AGRF, Melbourne) on 3730 DNA Analyzer (Applied Biosystems, USA) machines. Fragment analysis was carried out using Genemarker V1.90 software (Softgenetics, USA).

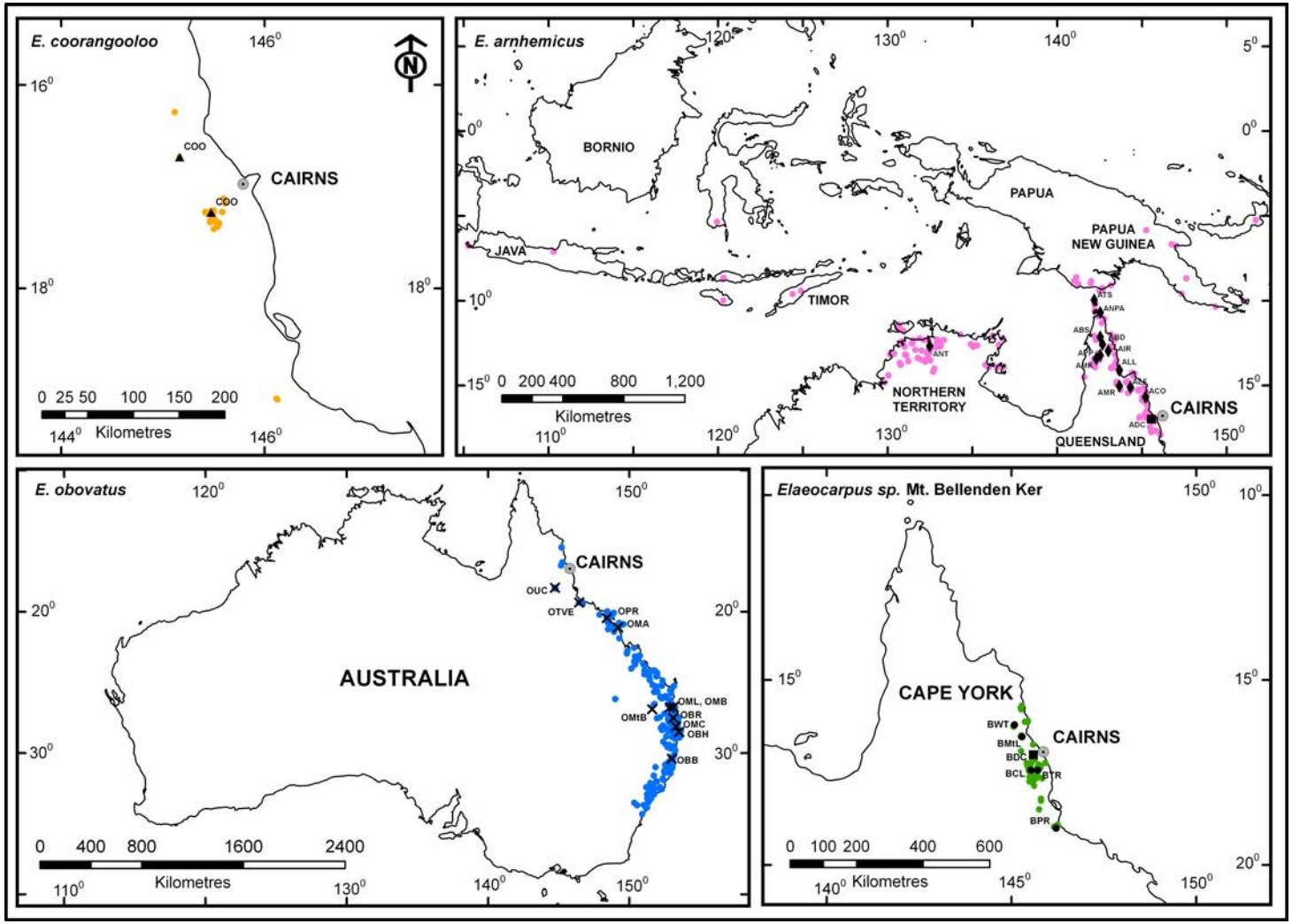


Figure 3.1. Distribution maps of each entity and the sampling sites. The distributional ranges are based on herbarium records downloaded from Australia's Virtual Herbarium and the literature (Coode, 1978; 2001f). Sampling sites were denoted by symbols as follows: ▲ *E. coorangooloo*, ◆ *E. arnhemicus*, ● *E.* sp. Mt. Bellenden Ker and ✕ *E. obovatus*. Population codes for the sampling sites are provided in Table 3.1

Table 3.1. Sampling site details for the nuclear microsatellite analysis of *E. arnhemicus, E. obovatus* and *E.* sp. Mt. Bellenden Ker. * No voucher specimen from the population of *E. obovatus* from Bongil Bongil National Park is available. The geocode for the DNA samples from Bongil Bongil National Park was generated using Google Earth.

Code	N	Locality	Vouchers	Latitude	Longitude	Altitude
E. arnhemicus						
ACO	5	Cooktown vicinity, Cape York Peninsula, QLD	Baba 476-480 (CNS)	-15 40 55	145 12 23.6	133 m
ALF	17	Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 484, 486-487, 489-493, 495-498, 501, 502 & 505-507 (CNS)	-15 6 23.6	144 18 53.5	44 m
AIR	3	Iron Range National Park, Cape York Peninsula, QLD	Baba 546-548 (CNS)	-12 57 30.8	143 0 53.2	150 m
APP	9	Australian Wildlife Conservancy Picaninny Plains Sanctuary, Cape York Peninsula, QLD	Baba 549, 551, 552, 558 & 561-565 (CNS)	-13 12 55.9	142 44 24.3	93 m
ANPA	10	Northern Peninsula Area, Cape York Peninsula, QLD	Baba 586-589, 593, 597, 600 & 601, (CNS) Kilgour 468 &469 (CNS)	-10 42 4.1	142 32 13.5	3 m
ABS	6	Bromwell Station, Cape York Peninsula, QLD	Baba 603-607, 610 (CNS)	-12 8 25.3	142 37 22.9	72 m
ABD	8	Batavia Downs, Cape York Peninsula, QLD	Baba 611-614 & 616-619 (CNS)	-12 34 34.1	142 18 57.4	70 m
AMK	20	Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 623-625, 627-631, 633-635, 637, 638, 641, 642 & 648-652 (CNS)	-13 24 39	142 18 57.6	55 m

Code	N	Locality	Vouchers	Latitude	Longitude	Altitude
ALL	7	Lama Lama National Park, Cape York Peninsula, QLD	Baba 653-658 & 660 (CNS)	-14 5 38.3	143 40 15.1	8 m
AMR	4	Morehead River, Cape York Peninsula, QLD	Baba 661-665 (CNS)	-15 1 21	143 39 56	61 m
ATS	1	Mabuyag Island, Torres Strait, QLD	Fell 10502 (BRI)	-9 57 27.6	142 11 8.6	13m
ANT	2	Kakadu National Park, NT	Cooper 2132 & 2133 (CNS)	-12 40 22	132 28 52	12 m
E. obovatus						
North						
OUC	9	Undara Crater National Park, Einasleigh Uplands, QLD	Baba 680-682, 685, 687, 689, 690 & 692 (CNS)	-18 18 35.3	144 44 20.4	1021 m
OTVE	2	Townsville, Brigalow Belt, QLD	Baba 969 & 967 (CNS)	-19 20 13.3	146 27 26.2	463 m
OMA	5	Mackay, Central Queensland Coast, QLD	Baba 698, 699 & 701-703 (CNS)	-21 7 4.9	149 12 34.7	51 m
OPR	5	Proserpine, Central Queensland Coast, QLD	Baba 706-708, 710 & 711 (CNS)	-20 26 46.3	148 24 42.5	101 m
South						
OBR	6	Brisbane region, South East Queensland, QLD	Baba 726-729, 758A & 759 (CNS)	-27 30 46.4	153 7 12.1	64 m
OMC	5	Mudgeeraba Creek, South East Queensland, QLD	Baba 730, 731 & 734-736 (CNS)	-28 6 46.1	153 19 4.5	55 m
OML	4	Maleny, South East Queensland, QLD	Baba 740-743 & 752 (CNS)	-26 45 34	152 52 15.3	400 m
OMtB	8	Road to Bunya Mountains, South East Queensland, QLD	Baba 744-751 (CNS)	-26 53 39.2	151 37 0.1	1026 m
OMB	4	Mooloolaba Beach, South East Queensland, QLD	Baba 754-757 (CNS)	-26 44 23.4	153 7 59.6	17 m
OBH	5	Brunswick Head, North Coast, NSW	Crayn 510 (NSW)	-28 31 56	153 32 30	15 m

Code	N	Locality	Vouchers	Latitude	Longitude	Altitude
OBB	4	Bongil Bongil National Park, North Coast, NSW*	N/A	-30 24 00	153 02 00	18 m
***E.* sp. Mt. Bellenden Ker**						
BPR	5	Paluma Range, Wet Tropics, QLD	Jackes 20102 & 20103 (CNS), Crayn 800 (NSW)	-19 0 30.2	146 12 26.31	889 m
BCL	4	Cloudland Nature Reserve, Wet Tropics, QLD	Baba 768, 440, 444 & 445 (CNS)	-17 26 10	145 31 48	950 m
BWT	5	Windsor Tableland National Park, Wet Tropics, QLD	Ford 4466 (CNS)	-16 13 00	145 05 00	900 m
BMtL	4	Mount Lewis Forest Reserve, Wet Tropics, QLD	Crayn 868 (NSW)	-16 32 18	145 17 16	1020 m
BTR	5	Topaz road, Wet Tropics, QLD	Crayn 846 (NSW)	-17 26 43	145 42 47	700 m
E. coorangooloo						
COO	2	Bakers Blue Mountain, Einasleigh Uplands, QLD Atherton, Atherton Tableland, Wet Tropics, QLD	Ford 5752 (CNS) Baba 695 (CNS)	27 31 11 17 15 10	153 31 48 145 26 35	940 m 752 m
Davies Creek populations						
ADC (cf. *E. arnhemicus*)	3	Davies Creek National Park, Wet Tropics, QLD	Baba 821-823 (CNS)	16 58 46.4	145 33 14.6	427 m
BDC (cf. *E.* sp. *Mt. Bellenden Ker)*	3	Davies Creek National Park, Wet Tropics, QLD	Crayn 808 (NSW)	17 1 37	145 35 24	650 m

Table 3.2. Nuclear microsatellite markers, repeats, sequences and annealing temperature employed in the PCR assays in this study. * size given in the length excluding the 5'M13 universal primer sequences.

Marker	*Repeat*	5'	3'	Size *	Annealing Temperature (C°)
scu01Eg	*(AG)n*	CCAAATGAAGAATACCTCCA	AGTGGCTTGGTCAGAGATTA	292-332	56
scu20Eg	(AG)n	TACGCCATCACTGTCTTCACCACC	ACACTCACCATCCTGTGCTCTATCC	144-168	61
scu21Eg	(AG)n	CCAAAATACCCTTCAAACCCACCT	TGATCGAAACAGCTCAAGAGTCTCC	357-351	54
scu22Eg	(AG)n	CGCTTCTTACGGTTCTTCTTGAAA	TTCGCTGCTCCCCCTGATAA	99-130	61
scu25Eg	(TGG)n	TTTGAGTAGCTCACTCTGCTCTTGG	CGTTGGATTGCCTCCCGATT	282-343	55
scu27Eg	(AG)n	TGTTTGCTGTGTCTTTCTCAGGAGG	GCCTTGTGTGCGAGTGGTATTTG	330-411	61
scu31Eg	(GA)n...(GT)n	GCAAAGCAAGGGCAAGTTCTCTT	CGGCTTCCTAAATTCACTGTATGGA	316-406	61
scu32Eg	*(GAA)n*	TGGAGAGCAAGGGACCGACTTA	CCACACAAGAGCACTAACAGCAGC	256-278	61
scu33Eg	(AG)n	GCTTTACACCAAGTAGGAACTACCA	CTAGCTTCTCGGTCGTTATCATTT	287-307	55
scu34Eg	(AG)n	TGGGAGATCAATAGGATTCAACAAT	TGCTTGCTCCATTTTCAAGATATG	150-174	60

3.2.1.4 Error Checking

Genotyping error can distort the results and lead to erroneous conclusions (see Pompanon et al., 2005). Causes of such errors include low quantity or quality of template DNA leading to allelic dropout (null alleles) or false alleles, and PCR artefacts leading to stutter bands and sizing errors. To help determine the impact of such errors on results, Pompanon et al. (2005) suggested performing blind repeats of 5-10% of the samples. This approach reduces the likelihood of ambiguous genotyping. In this study, blind sample repeats of 29% of samples were carried out before the subsequent analyses.

Null alleles, which are highly reproducible PCR artefacts, are difficult to detect in polyploid taxa. There is no currently available software to test for null alleles in polyploid taxa, and it was therefore not possible to assess null alleles with the available software for groups potentially containing more than two different cytotypes. An alternative approach for null allele detection in polyploid species suggested by Palop-Estaban et al. (2011) is checking the allele dosage using Microsatellite DNA Allele Counting-Peak Ratio (MAC_PR) methods (Esselink et al., 2004). However, this approach is not appropriate to use in this study, as homozygotes are not distinguishable from partial heterozygotes with null alleles without prior knowledge of the ploidy level in individual samples. In this study, there is no robust test for null alleles that can be applied to the present data. The rates of PCR failure, as well as microsatellite profiles were observed before the analyses.

3.2.1.5 Data preparation

While there is potential cytotypic variation found within the *E. obovatus* complex, the knowledge of potential polyploidy was delivered from microsatellite profiles only and cytotypes of all the individual samples and the patterns of inheritance -poly- and disomic inheritance of chromosomes assessed is absent. To reduce the influence of polyploidy in the analyses, the data were converted to binary format, where each allele for each locus was scored as present (1) or absent (0) across all individuals, except in one analysis, where the dataset was used in an allelic format (see 3.2.1.7).

3.2.1.6 Microsatellite profile

To assess the microsatellite profile, allelic diversity was measured for each species. For some populations fewer than the target 20 individuals per population were available. Because the assessment of allele diversity is not meaningful in small populations the results should be treated with caution.

The following allele diversity measures were calculated:

- **Allelic richness (A_I)**: mean number of alleles per locus
- **Allelic diversity (A_R):** mean number of unique alleles per locus
- **Unbiased allelic richness (uA_I)**: = A_I/N - unbiased mean number of alleles per locus (mean number of alleles per locus, averaged by the number of samples in the population)
- **Unbiased allelic diversity (uA_R)**: = A_R/N - unbiased mean number of unique alleles per locus (mean number of unique alleles per locus averaged by the number of samples in the population)
- **Proportion of individuals which possess two or more alleles ($P_{\geq 2A}$)**: proportion of individuals with two or more alleles per locus at the population level
- **Proportion of polymorphic loci (P)**: proportion of polymorphic loci at the population level

The proportion of individuals that possess two or more alleles ($P_{\geq 2A}$) was calculated to determine the frequency of heterozygous loci, since it was not possible to calculate genetic diversity (i.e. observed heterozygosity and expected heterozygosity) for potentially polyploid taxa.

For the taxa that were tentatively determined to be diploid based on their microsatellite profiles, expected and observed heterozygosity (*He* and *Ho* respectively) and the fixation index (*F*) were calculated. Tests for departure from Hardy-Weinberg Equilibrium (HWE) were performed using GeneAlex v.6.5 (Peakall & Smouse, 2006; Peakall & Smouse, 2012).

Only the reports on the microsatellite profile are discussed in the main text, the allele diversity assessment and the test for HWE are presented in the appendix (Appendix 3).

3.2.1.7 Analyses

Population genetic approaches for dealing with polyploid species are still developing (Kloda et al., 2008; Sampson & Byrne, 2012), especially for species showing variable cytotypes (Kloda et al., 2008; Španiel et al., 2011). In spite of these disadvantages, researchers have been successful in resolving the taxonomy of species complexes exhibiting heterogeneous cytotypes (Kloda et al., 2008; Španiel et al., 2011; Assoumane et al., 2013) using combined approaches. There are several approaches available to test the connectivity, relationships, genetic distance and taxon boundaries in mixed or polyploid groups, however, there is no consensus on the choice of test.

While there exist some models to elucidate relationships between taxa of the same ploidy (Thrall & Young, 2000; Van Puyvelde et al., 2010), analysis methods that can be applied across different ploidy levels remain underdeveloped. Furthermore, where there is ambiguity in the ploidy level or where the level of ploidy has been determined from microsatellite profiles only, genotyping can be problematic. Based on the likely existence of two different cytotypes in the *E. obovatus* complex, three analytical approaches were employed: (1) a similarity based approach using the binary and allelic datasets; (2) a model based approach using the binary dataset with ploidy levels tentatively assigned (2- and 4- ploidy); and (3) a hypothesis testing approach using the groups identified from the previous approaches.

Table 3.3. Summary of data types and software used in each analysis

Analysis type	Data type	Parameters	Software
Similarity-based	Binary	Jaccard's similarity coefficient (PCoA and Neighbour Joining)	FAMD GeneAlex
		Φ_{PT} and Nei's genetic distances (PCoA)	GeneAlex
	Allelic	Bruvo distance (PCoA)	POLYSAT (on the R platform)
Model-based	Binary	Bayesian cluster analysis	STRUCTURE
Hypotheses testing	Binary	Analyses of Molecular Variance (AMOVA)	GeneAlex
	Binary	Multigroup Discriminant Function Analyses	SPSS

3.2.1.7.1 Similarity based analyses

Principal Coordinate Analysis (PCoA) is an appropriate tool to explore the structure, distances and relationships amongst populations of different ploidy levels (Kloda et al., 2008; Španiel et al., 2011). Sampson & Byrne (2012) and Kloda et al. (2008) utilised Euclidian distance because it does not treat the character state "0" as a shared, common state. Therefore they suggested this approach is the most appropriate for assessing the diversity for groups with mixed ploidy, whereas Španiel et al (2011) employed Jaccard's and Sørensen-Dice's similarity coefficients to explore the diversity and relationships of their group of interest.

Several authors have suggested that Jaccard's', Sørensen-Dice's, and Cosine coefficients are the most suitable for binary variables (Da Silva Meyer et al., 2004; Duarte et al., 1999). A study by Jackson (1989) showed Jaccard's and Sørensen-Dice's coefficients produced identical results. Furthermore, Španiel et al. (2011) demonstrated the interchangeability between Jaccard's and Sørensen-Dice coefficient measures to assess the diversity of binary alleles. In this study Jaccard's similarity coefficient was chosen. The genetic distance and relationships were assessed by calculating a distance matrix from the binary data using Jaccard's similarity coefficient (Jaccard, 1901), which was visualised using two different methods, multidimensional scaling (PCoA) and clustering (Neighbour Joining). The robustness of the groupings on the Neighbour Joining tree was evaluated using a bootstrap analysis with 10000 replicates. All calculations were computed in the software FAMD 1. 25 (Schlülter & Harris, 2006).

Nei's pairwise genetic distance, and the pairwise population distance (Φ_{PT}, an analogue to F_{ST}, an appropriate measure for the binary dataset) were also calculated for the binary data and visualized using PCoA in the software GeneAlex v.6.5 (Peakall & Smouse, 2006; Peakall & Smouse, 2012).

Bruvo's genetic distance (Bruvo et al., 2004) was calculated using the allelic data and visualised using PCoA. This measure was developed to calculate the pairwise genetic distance of taxa irrespective of ploidy level, however the measure is not appropriate for allopolyploid organisms (Bruvo et al., 2004). Although allopolyploidy cannot be ruled out for the *E. obovatus* complex based on available data, the Bruvo distance was nonetheless calculated to assess the congruence with the results of the other distance measures. This analysis was performed using POLYSAT software implemented in R (Clark & Jasieniuk, 2011). Tentative ploidy levels were assigned for each sample based on ploidy estimates produced by POLYSAT (Table 3.6). This software determines the number of alleles present in each sample and considers the maximum allele number as the ploidy level. There were many samples showing a maximum allele number of three; however, whether these are triploid cannot be confirmed with the available data. In this context there are two interpretations of "more than two alleles": triploidy; and tetraploidy with one duplicated allele. To test the effect of these two interpretations on the results, separate analyses were undertaken. Samples containing missing data for more than two loci were excluded.

3.2.1.7.2 Model-based analysis

To obtain another perspective into group boundaries, population subdivision within and between populations of the group, and patterns of gene flow, Bayesian Structure analysis was conducted using STRUCTURE 2.3.2 (Falush et al., 2003, 2007; Pritchard et al., 2000). The modelling assumptions of STRUCTURE are Hardy-Weinberg equilibrium within populations and complete linkage equilibrium between loci within populations (Pritchard et al., 2000). Because there are loci that show departure from HWE in *E. sp*. Mt. Bellenden Ker (Appendix 3), the results should be interpreted with caution.

Membership coefficients (*q*) were calculated using the binary data set. The configuration followed the study by Falush et al. (2003): 10 runs for each of 10^5 iterations after 10^5 burn-in, under recessive model on the effect, with the admixture model for correlating allele frequencies. Based on a pilot run which tested a wide range of values for *K* (number of clusters), up to the number of collection sites (31), it was ascertained that more than five clusters are neither biologically meaningful nor necessary. Those up to 5 *K* were included for further analysis. The analyses were repeated separately on each group identified in the first run to test for further undetected substructure within the populations.

The *K* value was obtained by calculating the *ad hoc* statistic, ΔK, using the Evanno Method (Evanno et al., 2005) implemented in Structure Harvester (Dent & VonHoldt, 2012). ΔK quantity (= mean (|L”K|)/ sd L (K)) is based on the rate change in the log probability of the data between successive *K* values.

The individual runs at the optimum *K* were checked to ensure that the likelihood of each run was similar, and the highest probability plot was chosen. The plots were edited using Microsoft ®Excel ® 2011 v.14.2.5.

3.2.1.7.3 Hypotheses testing for evolutionary units

To assess the significance of the putative evolutionary units identified from the previous approaches, Multigroup Discriminant Function Analysis (MDFA) and Analysis of Molecular Variance (AMOVA: Excoffier et al., 1992) were employed on the binary dataset.

AMOVA analysis was undertaken to test whether genetic distance is hierarchically structured within the dataset, and was performed by calculating Φ_{PT} with 999 permutations on the binary data using the software GeneAlex v.6.5 (Peakall & Smouse, 2006; Peakall & Smouse, 2012). Data partitions tested include different cytotypes, clusters identified in the STRUCTURE

analyses, and hypothetical groups of interest (i.e. between and among genetic clusters resolved by the similarity-based analyses).

MDFA tests the significance of a set of discriminant functions and then assigns the cases to the reference groups. MDFA was performed in SPSS Statistics (IBM, v.19). MDFA. The significance test used in MDFA is identical to that of MANOVA (Hill & Lewicki, 2006). Like MANOVA, MDFA also requires certain assumptions to be met: 1) that the data have equal sample sizes; 2) that there is homogeneity of covariance; and 3) that the variables are normally distributed. It was ensured that all the requirements except homogeneous sample size were met before the results were interpreted. 'Leave-one-out' cross validation was used to test the accuracy of the classification model. This cross validation procedure reclassifies each sample depending on the functionality of all the other cases, excluding one case (Stone, 1974; Arlot & Celisse, 2010). A priori groups to be tested were chosen from the groups obtained in the previous analyses.

3.3 Results

3.3.1 Microsatellite locus screening and error checking

Initial screening identified three problematic loci, namely *scu27Eg*, *scu32Eg* and *scu21Eg*. These three loci were discarded on the basis of not being successfully optimised and repeatedly failing to amplify (*scu27eg*), containing ambiguous peaks (*scu32eg*), or being insufficiently variable within the study group (*scu21eg*).

A total of 29% of samples were repeated to assess the amplification error. Error rates are given in Table 3.4. In each of *scu25Eg* and *scu20Eg* up to 8 peaks were visualised. Because systematic screening with flow-cytometry was not available the samples possessing 8 peaks were treated as PCR amplification failures, which resulted in an artificially high proportion of PCR failures for locus *scu20Eg. W*hen these peaks were excluded the proportion of PCR failures was < 1%, the lowest of all the loci, therefore *scu20Eg* was included in the analysis. The high error rate shown in *scu34Eg* was probably biased by the low number of repeats available for the locus, and a low amplification success rate.

Suggested indicators for the presence of null alleles such as high rates of PCR failure (Palop-Esteban et al., 2011) or excess homozygosity (Kloda et al, 2005), were generally not observed in the present study. The only exception is *scu34Eg* for which a significant number of samples were homozygous. The frequency of null alleles may increase with cross-species amplifications.

However, in *Elaeocarpus*, Rossetto et al. (2009) successfully performed a wide range of cross-species amplifications and no mutation was found in the flanking regions. Furthermore, there is an empirical basis to reject the presence of null alleles in this marker: there was no difference observed in the frequency of homozygous alleles between freshly collected samples (where DNA quality is likely to be high) and silica-gel preserved samples previously collected where the lower DNA quality would be expected to increase the likelihood of observing null alleles. Nonetheless, a conservative approach was taken and *Scu34Eg* was excluded from the analyses. This, together with the lack of evidence for null alleles in the remaining markers engenders confidence that null alleles are unlikely to influence the results.

Table 3.4. Error rate for each locus across all the entities. PCR amplification failure is out of 181 samples. The numbers in the parentheses indicates the numbers of samples that showed more than 4 peaks.

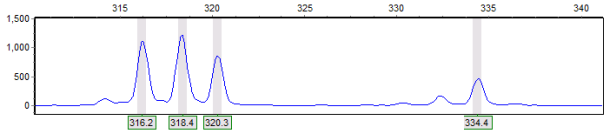
Locus	Success number /Repeat samples	Error rate	PCR failure	PCR failure rate
scu01Eg	13/13	0%	4	2%
scu20Eg	91/93	2%	17 (12)	17%
scu22Eg	18/19	5%	2	1%
scu25Eg	15/15	0%	14 (2)	9%
scu31Eg	95/100	5%	6	3%
scu33Eg	107/111	4%	13	7%
scu34Eg	10/12	17%	14	8%
Total	349/363	4%	84/1267	7%

3.3.2 Microsatellite profiles

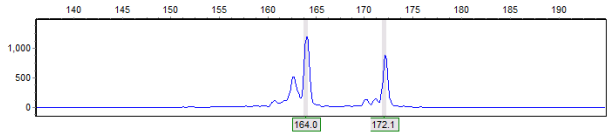
Microsatellite profiles revealed potential polyploidy within *E. obovatus* and *E. arnhemicus*, with some samples having more than two alleles expected for diploids. Polyploidy detection using microsatellite markers needs to be conducted with caution because higher numbers of alleles than the expected ploidy level may be caused by cross-contamination, PCR artefacts resulting from insufficiently optimised PCR conditions or genotyping errors caused by stutter peaks (e.g. Pompanon et al., 2005) or presence of plus-A artefacts (e.g. Davison & Chiba, 2003; Morin et al., 2010). To eliminate the possibility of cross-contamination and to minimise artefacts, PCR conditions were optimised and repeats were performed. The results confirmed that the high number of peaks observed (ranging from one to four: Figure 3.2) was not due to cross-contamination or PCR artefacts.

When combining all loci and all populations, the number of private alleles per locus was *11 in* scu*33Eg and* scu*25Eg* 13 in scu*01Eg*, 14 in scu*20Eg*, 16 in scu*22Eg*, and 19 in scu*31Eg*, with a mean value of 13 alleles per locus. Individual genotypes consisted of one to four alleles per locus for *E. arnhemicus* and *E. obovatus*, one to three alleles per locus for *E.* sp. Mt. Bellenden Ker (Table 3.5) and one to two alleles for *E. coorangooloo*, ADC and BDC. For *E.* sp. Mt. Bellenden Ker, the three alleles were found in loci *scu20Eg* and *scu25Eg* of the BWT (Mt. Windsor Tableland) population only.

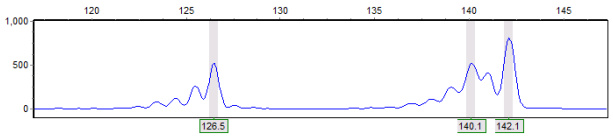
In total, 1287 alleles corresponding to 58 different allele sizes for *E. arnhemicus* (N=93), 707 alleles corresponding to 69 distinct alleles for *E. obovatus* (N=57) and 176 alleles corresponding to 31 distinct alleles for *E.* sp. Mt. Bellenden Ker (N=23, see also Table 3.5) were observed. The number of alleles shared by the species were: 15 for *E. arnhemicus* and *E. obovatus*; six for *E. obovatus* and *E.* sp. Mt. Bellenden Ker, and one for *E. arnhemicus* and *E.* sp. Mt. Bellenden Ker. In total, seven alleles were shared by all these three taxa. Six private alleles were found in *E. arnhemicus*, 14 in *E. obovatus* and three in *E.* sp. Mt. Bellenden Ker respectively.
The number of alleles found in *E. coorangooloo* (N=2), BDC (N=3) and ADC (N=3) was 17, 29 and 24 respectively, which corresponds to 16, 22 and 18 different allele sizes. One private allele for *E. coorangooloo* and BDC and none for ADC were found.



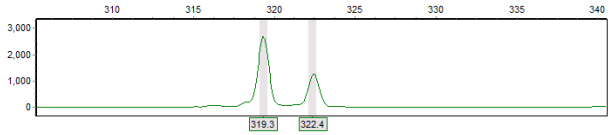
scu01Eg



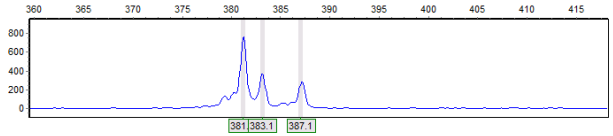
scu20Eg



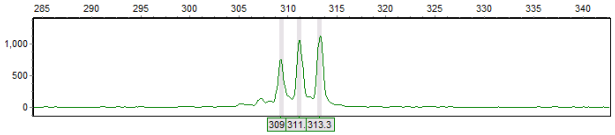
scu22Eg



scu25Eg

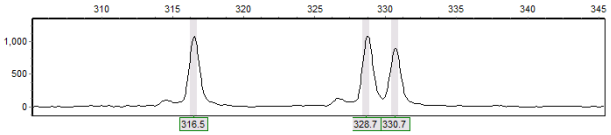


scu31Eg

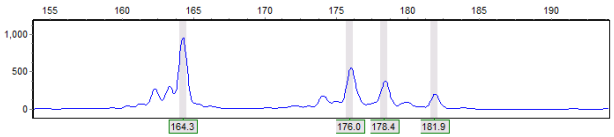


scu33Eg

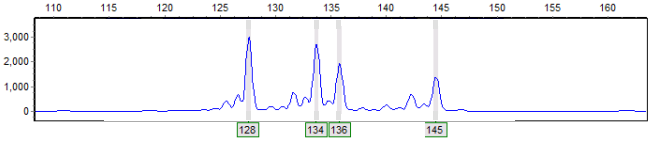
ABD5



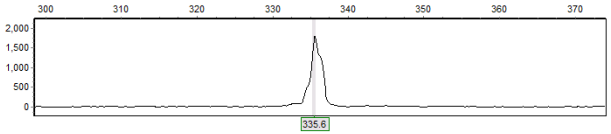
scu01Eg



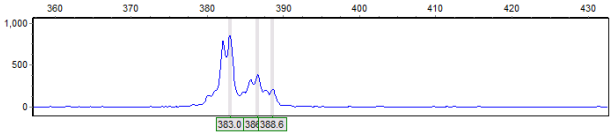
scu20Eg



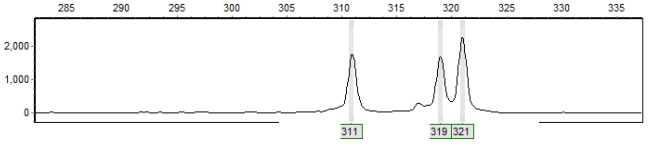
scu22Eg



scu25Eg



scu31Eg



scu33Eg

OMC1

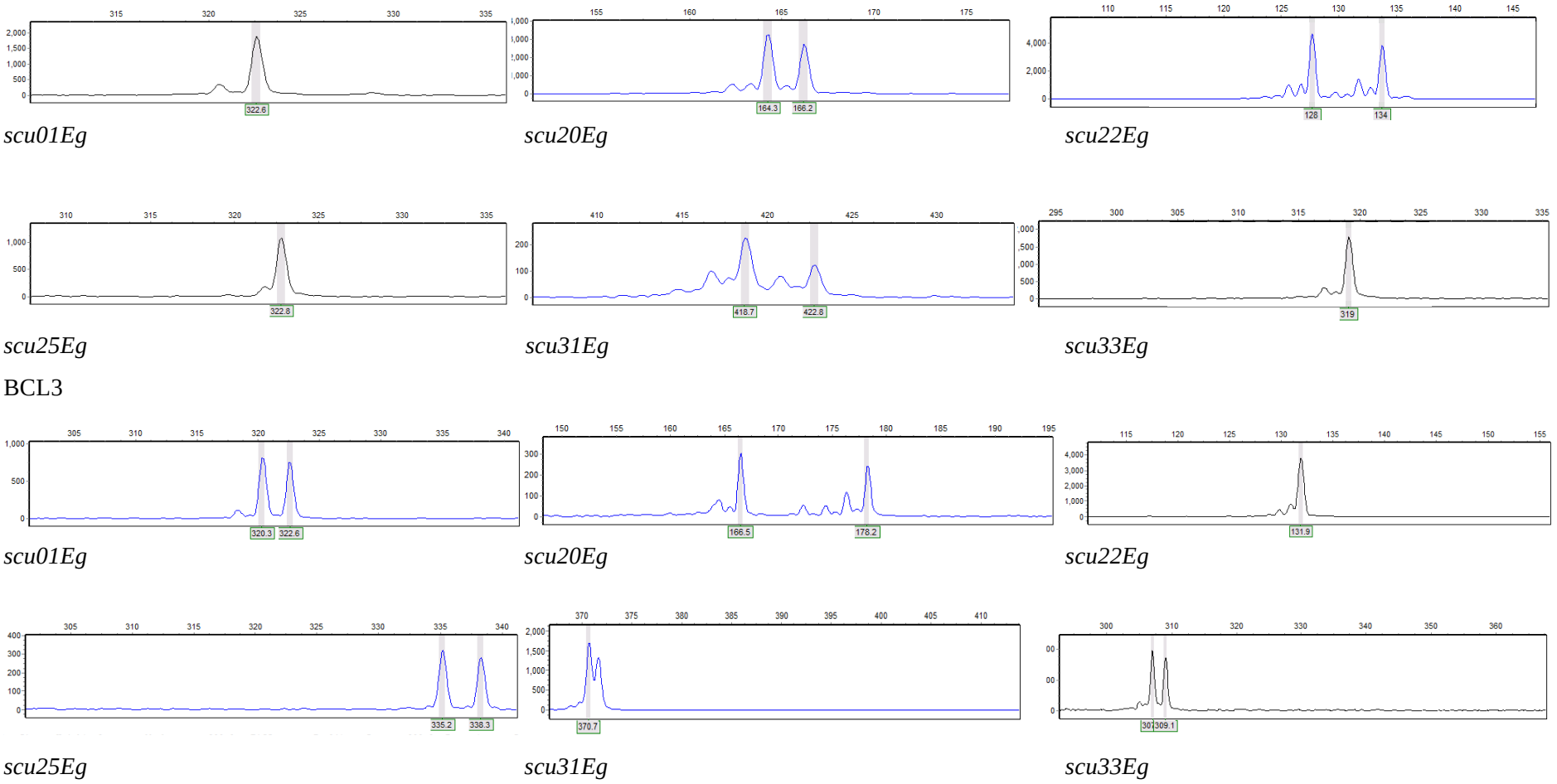


Figure 3.2. Example electropherograms for each marker used in this study. Some markers (e.g. ABD5: *scu01Eg*, *scu22Eg*, *scu31Eg* and *scu33Eg* and OMC1: *scu01Eg*, *scu20Eg*, *scu22Eg*, *scu31Eg* and *scu33Eg*) showed more than two alleles, which may indicate polyploidy. See table 3.1. for the population codes.

Table 3.5. Characteristics of the six nuclear microsatellite loci used in this study for *E. arnhemicus, E. obovatus,* and *E.* sp. Mt. Bellenden Ker. Results for *E. coorangooloo,* ADC and BDC are not shown, as samples sizes are less than five. SD: standard deviation.

Locus	*E. arnhemicus*					*E. obovatus*					*E.* sp. Mt. Bellenden Ker				
	N	***S***	***uA$_I$ (P≥2A)***	***uA$_R$***	***R***	***N***	***S***	***uA$_I$ (P≥2A)***	***uA$_R$***	***R***	***N***	***S***	***uA$_I$ (P≥2A)***	***uA$_R$***	***R***
scu01Eg	92	314-335	2.0 (0.67)	8	1-4	56	310-341	2.3 (0.70)	11	1-4	21	316-335	1.3 (0.33)	5	1-2
scu20Eg	92	160-184	2.3 (0.83)	12	1-4	37	160-186	2.4 (0.97)	13	1-4	13	164-172	1.7 (0.46*)	6	1-3*
scu22Eg	92	120-144	2.7 (0.98)	13	1-4	55	116-148	2.6 (0.80)	15	1-4	23	128-136	2.5 (0.82)	6	1-2
scu25Eg	87	300-340	2.2 (0.80)	9	1-4	54	300-343	1.7 (0.52)	7	1-4	16	300-340	1.6 (0.38*)	6	1-3*
scu31Eg	93	372-390	2.5 (0.88)	7	1-4	57	374-420	2.0 (0.65)	15	1-4	22	372-424	2.0 (0.77)	9	1-2
scu33Eg	92	305-325	2.8 (0.98)	9	1-4	52	309-325	1.9 (0.62)	7	1-4	19	319-325	1.1 (0.05)	3	1
Total	93	-	2.4 (SD = 0.31) (0.86; SD = 0.12)	58	-	57	-	2.2 (SD = 0.34) (0.71; SD = 0.16)	68		23	-	1.7 (SD = 0.50) (0.47; SD = 0.29)	35	-

N: sample size

S: range of allele sizes (bp)

uA$_I$: *A$_I$/N,* unbiased mean number of alleles per locus (mean number of alleles per locus, averaged by the number of samples in the population)

uA$_R$: *A$_R$/N,* unbiased mean number of unique alleles per locus (mean number of unique alleles per locus averaged by the number of samples in the population)

P≥$_{2A}$: proportion of individuals with 2 or more alleles per locus

R: range of alleles per individual

* three alleles were found in individuals from BWT at *scu20Eg* & *scu25Eg*.

Table 3.6. Evidence of polyploidy based on allele peak counts of the microsatellites for each locus and POLYSAT ploidy assessment. * manipulated to adjust to the minimum ploidy level of two. '-9' indicates missing data. See table 3.1. for the population codes.

Sample	Numbers of allele peaks per locus						POLYSAT Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
ACO1	2	3	3	2	3	2	3	2.5	3
ACO2	3	3	4	1	3	2	4	2.7	4
ACO3	2	3	3	2	3	2	3	2.5	3
ACO4	3	1	3	-9	2	3	3	2.4	3
ACO5	2	2	2	-9	3	2	3	2.2	3
ALF1	1	1	3	-9	1	3	3	1.8	3
ALF2	2	1	2	-9	4	2	4	2.2	4
ALF3	2	3	4	-9	2	2	4	2.6	4
ALF4	1	1	4	1	2	4	4	2.2	4
ALF5	2	1	2	1	2	4	4	2.0	4
ALF6	2	1	-9	3	2	3	3	2.2	3
ALF7	2	3	3	3	2	3	3	2.7	3
ALF8	3	2	3	1	3	2	3	2.3	3
ALF9	3	3	2	2	2	4	4	2.7	4
ALF10	2	2	3	1	2	4	4	2.3	4
ALF11	2	2	3	-9	2	2	3	2.2	3
ALF12	3	2	3	2	2	4	4	2.7	4
ALF13	3	1	3	2	4	2	4	2.5	4
ALF14	2	1	2	3	3	3	3	2.3	3
ALF15	2	2	2	2	1	2	2	1.8	2
ALF16	3	2	3	2	2	3	3	2.5	3
ALF17	3	2	3	2	2	4	4	2.7	4
AIR1	3	2	3	2	1	3	3	2.3	3
AIR2	3	2	3	2	1	3	3	2.3	3
AIR3	3	2	3	2	1	3	3	2.3	3
APP1	2	2	2	2	3	3	3	2.3	3

	Manual allele count						POLYSAT		
	Locus						Allele		
Sample	***scu01Eg***	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
APP2	1	4	2	3	2	4	4	2.7	4
APP3	1	4	3	3	2	2	4	2.5	4
APP4	2	1	2	1	2	4	4	2.0	4
APP5	1	1	2	3	4	4	4	2.3	4
APP6	2	3	3	2	4	2	3	2.5	3
APP7	2	3	3	2	4	2	3	2.5	3
APP8	1	2	2	3	4	4	4	2.5	4
APP9	3	2	3	2	4	3	3	2.7	3
ANPA1	2	1	2	1	2	3	3	1.8	3
ANPA2	3	2	3	2	3	3	3	2.5	3
ANPA3	3	2	2	1	2	3	3	2.0	3
ANPA4	1	2	2	2	1	3	3	1.8	3
ANPA5	3	4	2	2	2	3	4	2.7	4
ANPA6	1	2	3	2	3	3	3	2.3	3
ANPA7	2	2	3	2	3	3	3	2.5	3
ANPA8	1	2	2	2	2	3	3	2.0	3
ANPA9	1	1	3	2	2	3	3	2.0	3
ANPA10	1	1	1	2	1	1	2	1.2	2
ABS1	2	3	2	1	3	3	3	2.3	3
ABS2	1	3	2	1	3	4	4	2.3	4
ABS3	3	3	2	2	3	1	3	2.3	3
ABS4	1	3	2	2	2	3	3	2.2	3
ABS5	1	-9	3	3	2	3	3	2.4	3
ABS6	3	3	2	3	2	3	3	2.7	3
ABD1	4	3	3	2	3	3	4	3.0	4
ABD2	1	4	3	2	3	3	4	2.7	4
ABD3	2	4	3	3	3	3	4	3.0	4
ABD4	1	4	4	3	1	3	4	2.7	4
ABD5	4	2	3	2	3	3	4	2.8	4

Sample	Manual allele count						POLYSAT		
	Locus						Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
ABD6	2	2	2	3	3	4	4	2.7	4
ABD7	2	4	2	3	3	4	4	3.0	4
ABD8	1	3	3	2	2	2	3	2.2	3
AMK1	2	3	3	2	3	2	3	2.5	3
AMK2	-9	2	3	2	2	4	4	2.6	4
AMK3	3	1	2	2	4	4	4	2.7	4
AMK4	3	1	2	2	2	2	3	2.0	3
AMK5	1	2	3	2	2	2	3	2.0	3
AMK6	2	3	2	2	1	3	3	2.2	3
AMK7	3	3	3	2	2	4	4	2.8	4
AMK8	1	2	3	2	2	3	3	2.2	3
AMK9	2	3	4	3	3	2	4	2.8	4
AMK10	1	2	4	2	4	3	4	2.7	4
AMK11	2	2	4	2	1	3	4	2.3	4
AMK12	1	2	2	2	2	4	4	2.2	4
AMK13	3	2	3	2	2	3	3	2.5	3
AMK14	2	2	4	2	2	2	4	2.3	4
AMK15	1	3	4	2	2	3	4	2.5	4
AMK16	1	4	3	4	2	2	4	2.7	4
AMK17	1	2	3	1	3	2	3	2.0	3
AMK18	2	3	3	2	4	3	4	2.8	4
AMK19	1	2	1	3	3	2	3	2.0	3
AMK20	2	3	3	3	3	2	3	2.7	3
ALL1	3	2	3	3	3	3	3	2.8	3
ALL2	1	3	2	2	3	4	4	2.5	4
ALL3	1	3	3	2	3	2	3	2.3	3
ALL4	2	3	2	2	2	2	3	2.2	3
ALL5	2	3	3	1	2	-9	3	2.2	3
ALL6	1	4	2	1	2	3	4	2.2	4

Sample	Manual allele count						POLYSAT		
	Locus						Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
ALL7	1	3	2	2	2	3	3	2.2	3
AMR1	2	2	2	2	3	2	3	2.2	3
AMR2	2	1	2	1	2	3	3	1.8	3
AMR3	3	2	3	3	2	2	3	2.5	3
AMR4	2	2	3	2	4	2	4	2.5	4
AMR5	2	2	4	1	3	3	4	2.5	4
ANT1	2	2	3	1	3	3	3	2.3	3
ANT2	1	2	3	1	1	2	3	1.7	3
ATS1	2	2	3	2	3	4	4	2.7	4
OUC1	2	3	1	1	2	2	3	1.8	3
OUC2	1	2	1	1	1	2	2	1.3	2
OUC3	1	2	1	1	1	2	2	1.3	2
OUC4	2	3	1	1	1	1	3	1.5	3
OUC5	1	4	1	1	2	1	4	1.7	4
OUC6	2	4	2	1	1	1	4	1.8	4
OUC7	-9	4	-9	-9	1	2	4	2.3	4
OUC8	1	3	2	1	1	2	3	1.7	3
OUC9	1	4	2	1	1	2	4	1.8	4
OTVE1	2	3	2	1	1	2	3	1.8	3
OTVE2	1	4	1	1	2	1	4	1.7	4
OMA1	1	3	1	2	1	1	3	1.5	3
OMA2	1	4	2	1	2	1	4	1.8	4
OMA3	1	4	2	1	2	1	4	1.8	4
OMA4	1	3	3	2	4	1	4	2.3	4
OMA5	1	-9	1	1	3	1	3	1.4	3
OPR1	1	4	2	2	2	2	4	2.2	4
OPR2	2	4	2	2	2	2	4	2.3	4
OPR3	1	4	1	1	1	1	4	1.5	4
OPR4	1	-9	1	1	1	1	1	1.0	2*

Sample	Manual allele count						POLYSAT		
	Locus						Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
OPR5	1	-9	2	1	2	2	2	1.6	2
OBR1	4	4	4	4	2	2	4	3.3	4
OBR2	2	-9	4	1	2	1	4	2.0	4
OBR3	1	-9	2	1	2	1	2	1.4	2
OBR4	3	-9	3	2	1	1	3	2.0	3
OBR5	3	4	1	2	2	2	4	2.3	4
OBR6	3	-9	3	2	2	1	3	2.2	3
OMC1	3	4	4	1	2	3	4	2.8	4
OMC2	3	4	4	1	3	2	4	2.8	4
OMC3	3	4	3	3	3	3	4	3.2	4
OMC4	2	-9	3	3	3	2	3	2.6	3
OMC5	2	-9	4	2	3	3	4	2.8	4
OML1	3	4	3	2	1	3	4	2.7	4
OML2	3	-9	4	-9	2	3	4	3.0	4
OML3	4	4	3	2	4	2	4	3.2	4
OML4	3	-9	4	2	3	3	4	3.0	4
OMtB1	3	3	3	-9	2	3	3	2.8	3
OMtB2	4	4	4	3	3	3	4	3.5	4
OMtB3	4	4	4	3	2	4	4	3.5	4
OMtB4	4	3	4	3	1	3	4	3.0	4
OMtB5	4	4	4	3	2	3	4	3.3	4
OMtB6	4	4	4	3	2	3	4	3.3	4
OMtB7	4	-9	3	1	3	2	4	2.6	4
OMtB8	4	4	4	3	1	2	4	3.0	4
OMB1	2	3	2	2	3	3	3	2.5	3
OMB2	3	-9	3	3	1	3	3	2.6	3
OMB3	4	4	4	1	3	1	4	2.8	4
OMB4	3	-9	2	2	1	2	3	2.0	3
OBB1	2	2	3	2	1	-9	3	2.0	3

Sample	Manual allele count						POLYSAT		
	Locus						Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
OBB2	3	-9	4	1	3	1	4	2.4	4
OBB3	3	-9	-9	1	4	1	4	2.3	4
OBB4	2	-9	3	2	2	1	3	2.0	3
OBH1	2	4	2	2	2	-9	4	2.4	4
OBH2	1	1	3	1	1	-9	3	1.4	3
OBH3	2	-9	3	2	1	-9	3	2.0	3
OBH4	3	-9	4	1	2	-9	4	2.5	4
OBH5	2	-9	3	2	2	1	3	2.0	3
BCL1	2	2	2	2	2	1	2	1.8	2
BCL2	1	1	2	2	2	1	2	1.5	2
BCL3	1	2	2	1	2	1	2	1.5	2
BCL4	1	1	2	1	2	1	2	1.3	2
BPR1	1	1	2	1	2	1	2	1.3	2
BPR2	2	1	2	2	2	1	2	1.7	2
BPR3	2	1	2	2	2	1	2	1.7	2
BPR4	2	-9	2	-9	2	1	2	1.8	2
BPR5	1	-9	2	1	2	1	2	1.4	2
BWT1	-9	3	1	-9	-9	-9	3	2.0	3
BWT2	2	3	2	1	1	1	3	1.7	3
BWT3	1	4	2	3	2	1	4	2.2	4
BWT4	2	-9	1	-9	2	-9	2	1.7	2
BWT5	1	-9	2	2	2	1	2	1.6	2
BMtL1	2	2	2	1	1	2	2	1.7	2
BMtL2	1	1	2	1	2	1	2	1.3	2
BMtL3	1	1	2	1	1	1	2	1.2	2
BMtL4	1	-9	2	-9	2	-9	2	1.7	2
BTR1	1	-9	1	1	2	-9	2	1.3	2
BTR2	1	-9	2	1	1	1	2	1.2	2
BTR3	1	-9	2	-9	2	1	2	1.5	2

Sample	Manual allele count						POLYSAT		
	Locus						Allele		
	scu01Eg	***scu20Eg***	***scu22Eg***	***scu25Eg***	***scu31Eg***	***scu33Eg***	max.	mean	ploidy
BTR4	1	-9	2	-9	2	1	2	1.5	2
BTR5	-9	-9	2	-9	1	1	2	1.3	2
BDC1	1	2	2	2	2	1	2	1.7	2
BDC2	1	1	2	2	2	2	2	1.7	2
BDC3	2	1	2	1	2	1	2	1.5	2
ADC1	2	1	2	1	1	1	2	1.3	2
ADC2	2	2	1	2	-9	2	2	1.8	2
ADC3	1	1	1	2	1	1	2	1.2	2
COO1	2	3	3	1	-9	-9	3	2.3	3
COO2	2	1	2	2	-9	2	2	1.8	2

3.3.3 Similarity based analyses

3.3.3.1 PCoA on Jaccard's similarity coefficient

An ordination plot was obtained by PCoA using Jaccard's similarity coefficients calculated from binary data (Figure 3.3). There are two major groupings that were obtained: one is a homogenous cluster consisting of all the samples of *E. arnhemicus*, and the other corresponds to the samples of *E. obovatus, E. sp.* Mt. Bellenden Ker, *E. coorangooloo,* ADC and BDC only. Although the boundary was somewhat indistinct, the latter group was further divided into subgroups, each corresponding to the samples of *E.* sp. Mt. Bellenden Ker, BDC, ADC, *E. coorangooloo*, *E. obovatus* from the northern (OUC, OTVE, OMA and OPP, Figure 3.3) and the southern parts of the species' range (OMtB, OML, OBR, OMB, OMC, OBH, OBB, Figure 3.3) respectively. Hereafter the northern and the southern populations of *E. obovatus* are called *E. obovatus* North *E. obovatus* South respectively.

3.3.3.2 Neighbour Joining on Jaccard's similarity coefficient

An unrooted tree based on Jaccard's similarity coefficient (Figure 3.4) showed clear separation of *E. arnhemicus* from the other entities although there is no bootstrap support. There is one cluster within this large cluster of *E. arnhemicus* with bootstrap support of 89%; however, there is no correlation with known populations. Similarly, although there is no support, a large cluster formed with samples of *E. obovatus* North, *E. obovatus* South, *E. sp*. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC. Four clusters that gained moderate to high bootstrap values were: the Undara National Park (OUN) samples (89%), a couple of Davies Creek samples (ADC 1 and 3: 89%), one *E. coorangooloo* (COO2) and one Davies Creek (ADC2) sample (90%), and another sample of *E. coorangooloo* (COO1) and OBH1, OBH2, OBB1 (60%) (Figure 3.4).

3.3.3.3 PCoA on Pairwise individual on Nei's genetic distance

An ordination plot of PCoA on pairwise individual distance using Nei's genetic distance metric showed no strong congruence with the priori groups, except in the samples of *E. arnhemicus* (Figure 3.5 a to c). On the other hand, an ordination plot of PCoA on pairwise population distance revealed two major clusterings (Figure 3.6 a to c). The first cluster corresponds to all the samples of *E. arnhemicus*, the second corresponds to the rest of the samples. The second cluster can be divided into populations of *E. obovatus* North, *E. obovatus* South and the samples of *E. sp*. Mt. Bellenden Ker, COO, ADC and BDC. A boundary within the subcluster of *E. sp.* Mt. Bellenden Ker, COO, ADC and BDC appears to be somewhat distinct.

3.3.3.4 PCoA on Pairwise population distance on Φ_{PT}

A PCoA ordination plot of on pairwise population distance Φ_{PT} (Figure 3.7 a to c) revealed identical patterns to the ordination of PCoA using Nei's genetic distance on populations (Figure 3.6 a to c).

3.3.3.5 PCoA on Pairwise individual of Bruvo distance

The two POLYSAT analyses with different ploidy assumptions for the samples with more than two alleles showed identical results, and therefore only one output is presented here.

A PCoA ordination plot based on Bruvo's pairwise individual distance showed clear segregation of two clusters of *E. arnhemicus* and *E.* sp. Mt. Bellenden Ker samples respectively from another cluster of *E. obovatus* samples (Figure 3.8). While *E. obovatus* North and South group separately, the boundary of these clusters is not clear. *Elaeocarpus coorangooloo*, ADC and BDC are positioned adjacent to the *E. obovatus* cluster but cannot be clearly distinguished from it. It is notable that two samples of BDC (BDC 1 and 3) are part of the *E.* sp. Mt. Bellenden Ker cluster. OBH2 was found to be an outlier and BMtL1 was found to be part of the *E. obovatus* cluster.

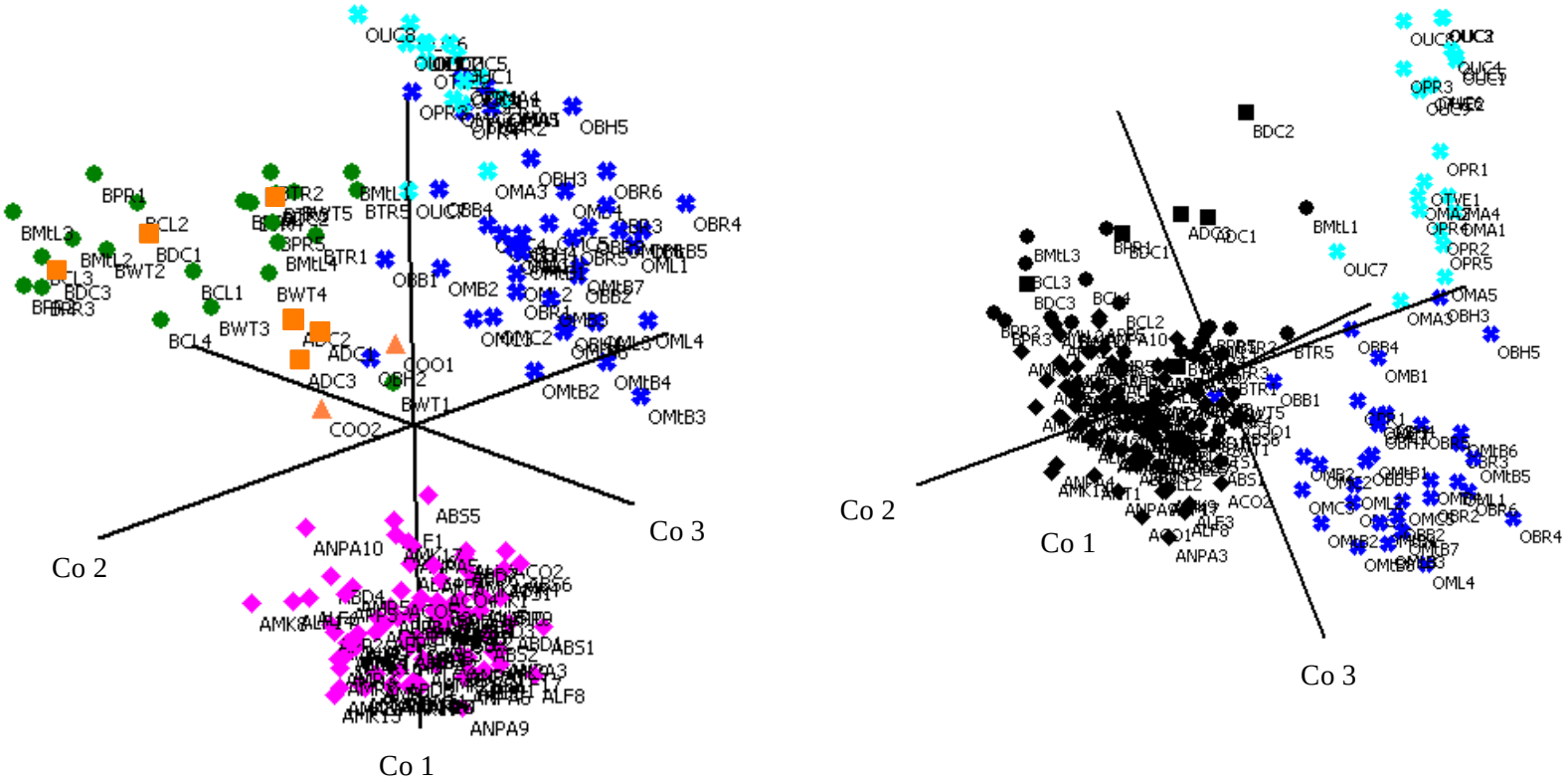


Figure 3.3. PCoA three-dimensional ordination plots based on Jaccard's similarity coefficient from two different angles. Three coordinates explain 21.4% (coordinate 1), 8.4% (coordinate 2) and 7.3% (coordinate 3) of the variation. Left: general view. Right: view focusing on *E. obovatus* samples only (sky blue and dark blue), all other samples are in black. ◆ *E. arnhemicus* (pink), ✕ *E. obovatus* (sky blue - samples of *E. obovatus* North, dark blue - samples of *E. obovatus* South) ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC and BDC (orange). See table 3.1. for the population codes.

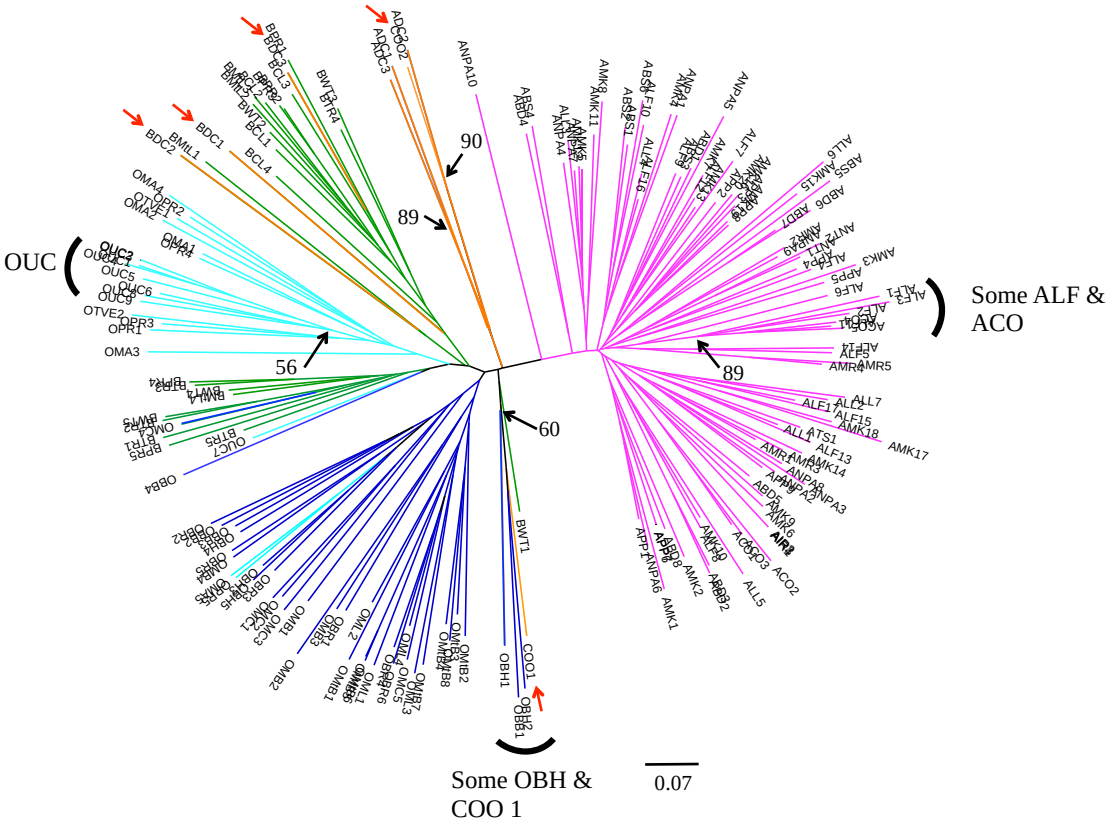


Figure 3.4. An unrooted neighbour-joining tree constructed from the binary data set using Jaccard's similarity coefficient with all samples included. Each branch represents an individual sample. Pink - *E. arnhemicus*, sky blue - *E. obovatus* North, dark blue - *E. obovatus* South, green - *E.* sp. Mt. Bellenden Ker and orange - *E. coorangooloo*, ADC and BDC. Numbers indicate branches receiving bootstrap support above 50%. Red arrows indicate the placements of the samples of BDC and COO. Brackets indicate small clusters that were supported by bootstrap values. See table 3.1. for the population codes.

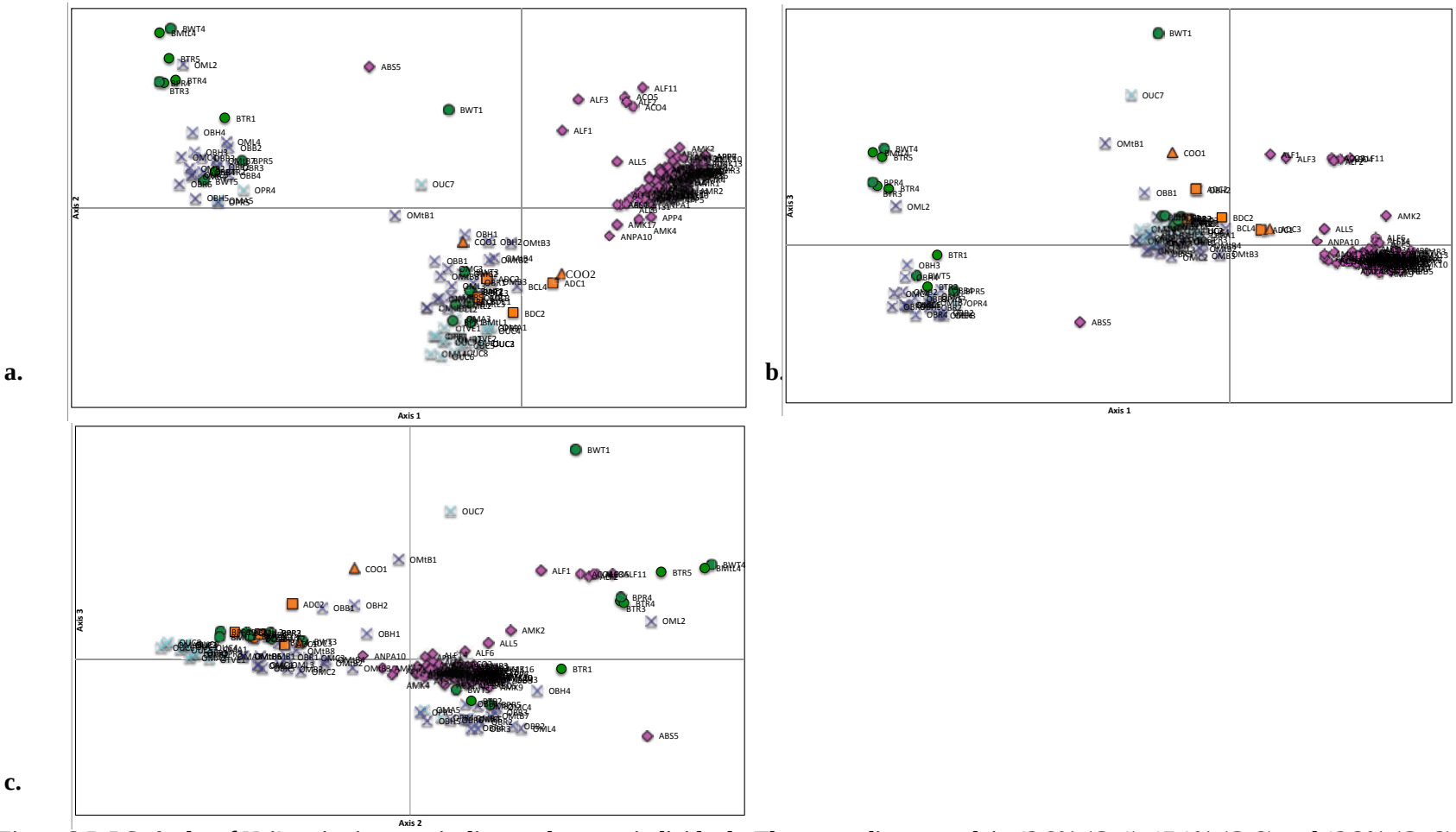


Figure 3.5. PCoA plot of Nei's pairwise genetic distance between individuals. Three coordinates explain 43.6% (Co1), 17.1% (Co2) and 13.3% (Co 3) of the variation for the pairwise individual distance. A: plot of axes 1 and 2, b: axes 1 and 3 and c: axes 2 and 3. Each point represents an individual sample. ◆ *E. arnhemicus* (pink), ✕ *E. obovatus (*sky blue: samples of *E. obovatus* North, dark blue: samples of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC and BDC (orange). See table 3.1. for the population codes.

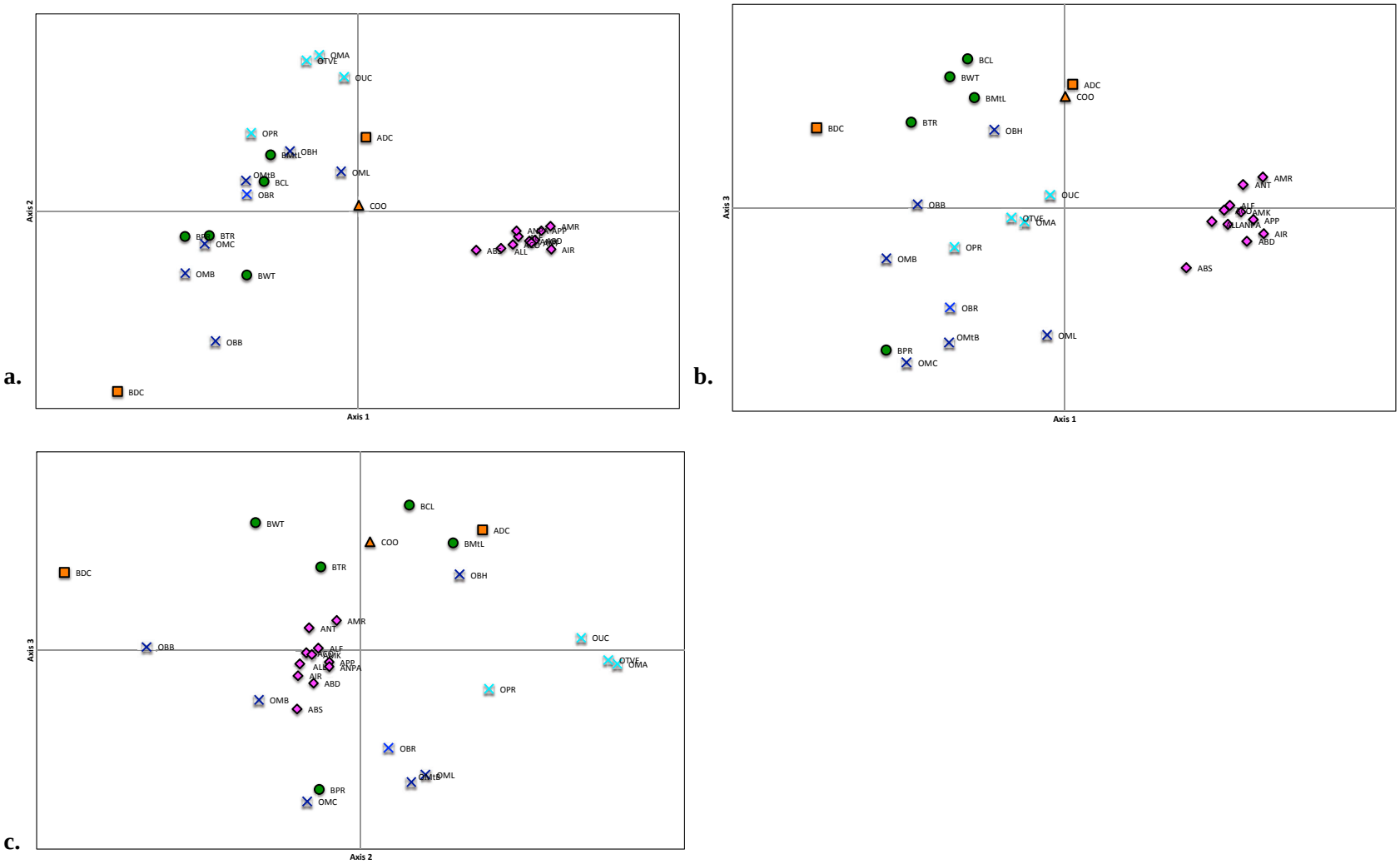


Figure 3.6. PCoA plot of Nei's pairwise genetic distance between populations. Three coordinates explain 50.6% (Co1), 16.6% (Co2) and 12.7% (Co 3) of the variation for the pairwise population distances respectively. A: plot of axes 1 and 2, b: axes 1 and 3 and c: axes 2 and 3. Each point represents an individual sample. ◆ *E. arnhemicus* (pink), ✕ *E. obovatus* (sky blue: *E. obovatus* North, dark blue: *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC and BDC (orange). See table 3.1. for the population codes.

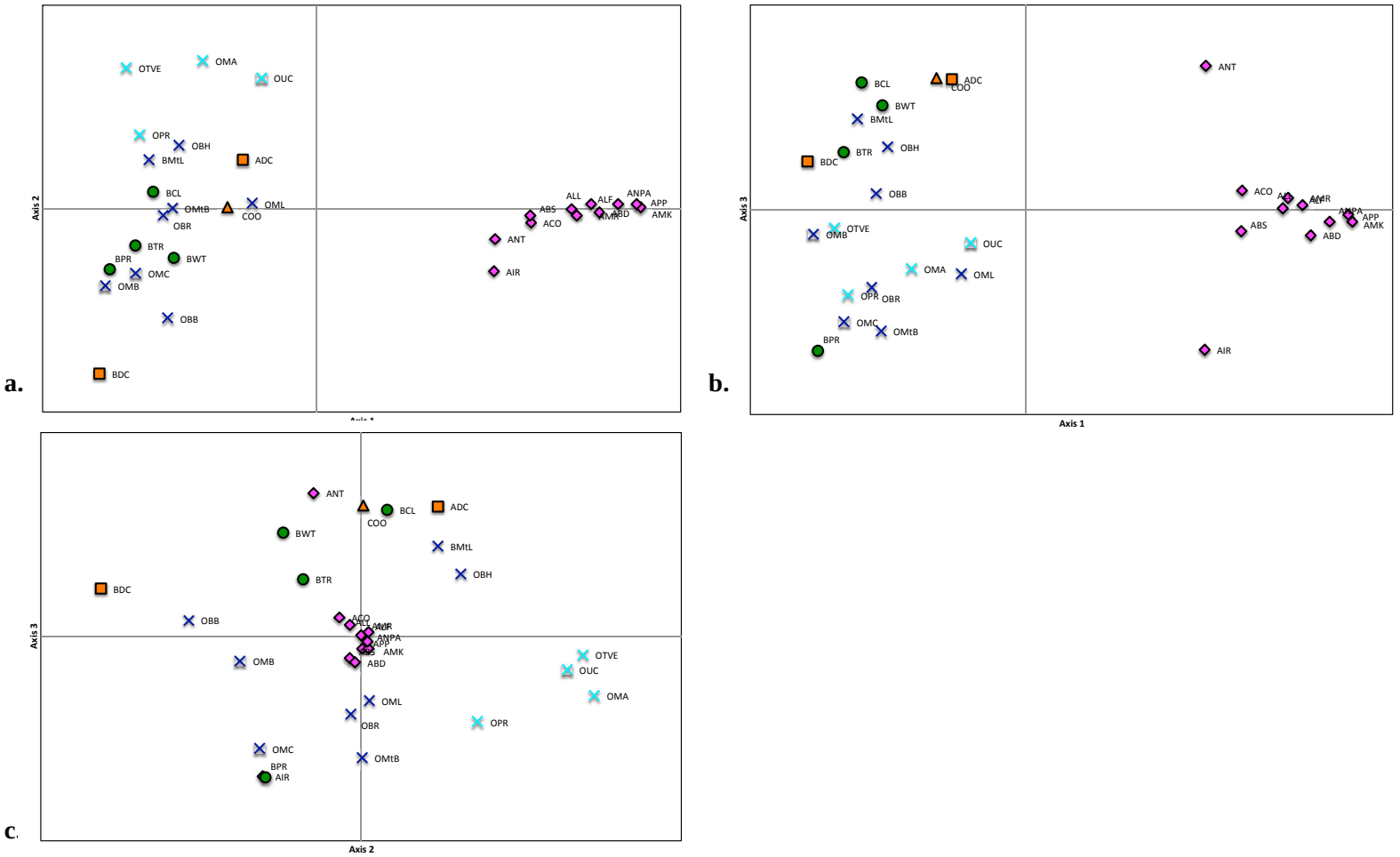


Figure 3.7. PCoA ordination plot of pairwise Φ_{PT} on populations. Three coordinates explain 44.7% (coordinate 1), 16.2% (coordinate 2) and 12.8% (coordinate 3) of the variation. A: plot of axes 1 and 2, b: axes 1 and 3 and c: axes 2 and 3. Each point represents an individual sample. ◆ *E. arnhemicus* (pink), ✕ *E. obovatus* (sky blue: *E. obovatus* North, dark blue: *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC and BDC (orange). See table 3.1. for the population codes.

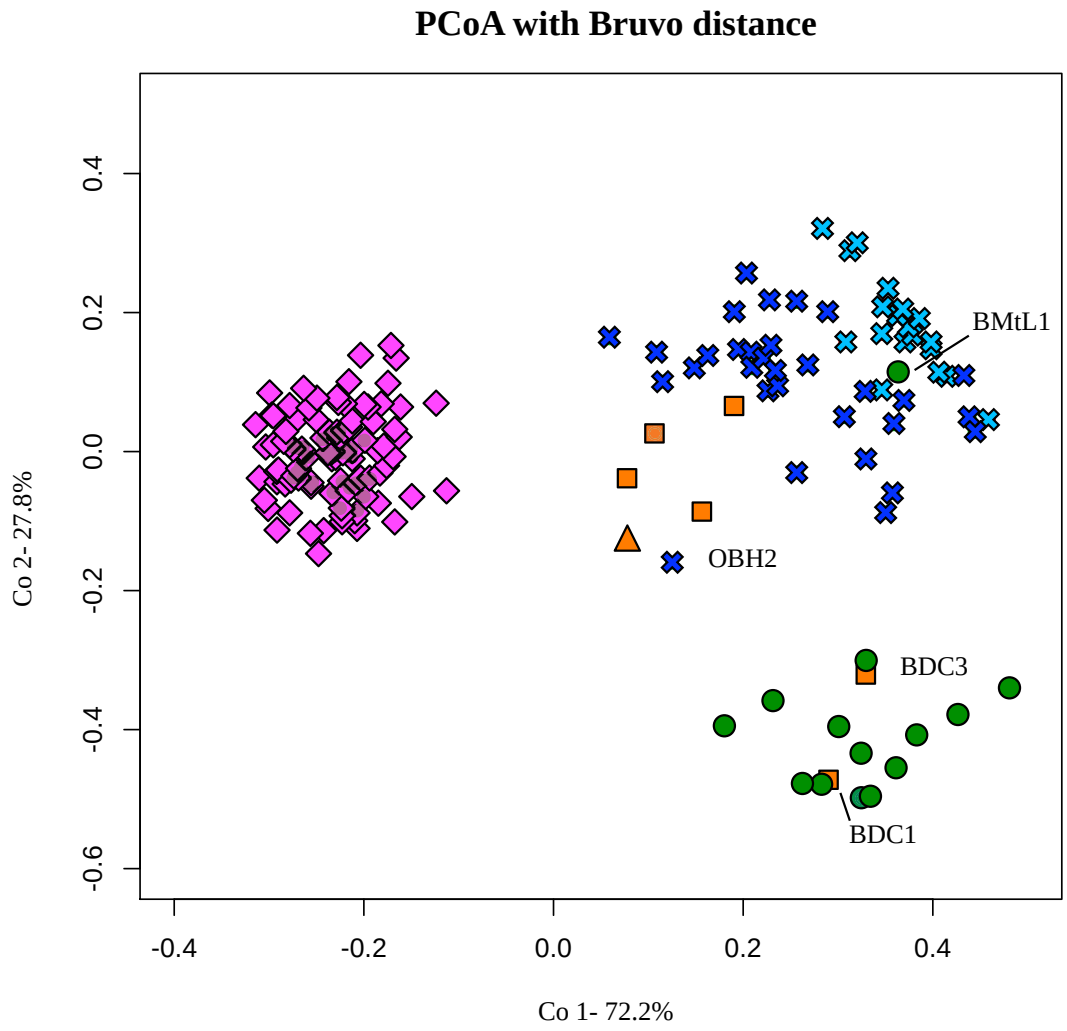


Figure 3.8. PCoA ordination plot of Bruvo distances calculated on the codominant data set of all the samples. ◆ *E. arnhemicus* (pink), ✕ *E. obovatus* (sky blue: *E. obovatus* North and dark blue: *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC and BDC (orange). 72.2% and 27.8% of the variation is accounted for by axes 1 and 2 respectively. See table 3.1. for the population codes.

3.3.4 Model-based analysis

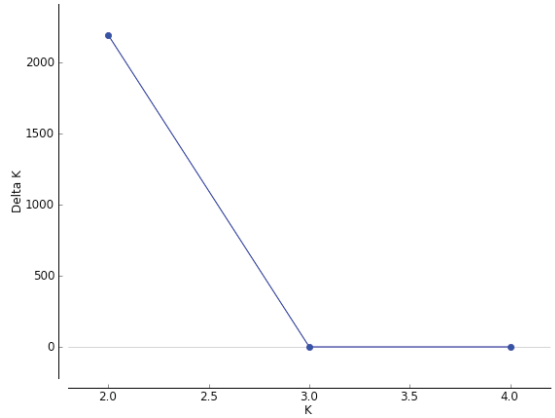
Bayesian cluster analysis of the binary data set was conducted using STRUCTURE, on (1) all of the samples, (2) *E. arnhemicus*, and (3) *E. obovatus*, *E.* sp. Mt. Bellenden Ker, ADC, BDC and *E. coorangooloo*. These subsets of data were generated based on the pattern of genetic clustering that resulted from the first run of STRUCTURE. The results obtained when tetraploid and diploid settings were used were identical (Appendix 4), and therefore only the results from the tetraploid setting are presented.

There was a clear drop in the value of ΔK, determined by the Evanno Method, at K=2 (Figure 3.9**a**) for the dataset containing all of the samples. The STRUCTURE output distinguished two clusters corresponding to *E. arnhemicus* (hereafter Q1) and the other entities (hereafter Q2) respectively, with little mixture of the two (Figure 3.10a), except some samples ($q1$ < 0.95: ACO2, ALF7, ABS4 AMK11, AMK8, AMK17, ABS5, ABS6 and ATS1, and $q2$< 0.95: OMtB 2 & 3, OMC 2, OMtB4, OML4, BCL4, ADC3 and COO2, $q1 \approx q2$: BWT1) and, indicating that there is also genetic elements derived from Q1 or Q2. The membership coefficient (q) of Q1 ranged from 0.746 to 0.997 with an average of 0.977 and was greater than 0.926 for 95% of *E. arnhemicus* individuals. The values of q (BWT excluded) for Q2 ranged from 0.763 to 0.998 with an average of 0.979 and greater than 0.917 for 95 % of individuals in the respective cluster.

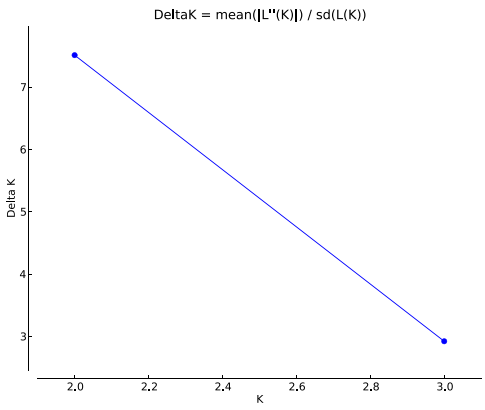
Further analysis of Q1 samples revealed two clusters with peaks of ΔK at K=2 (Figure 3.9b); however, all of the individuals in both clusters were admixed (Figure 3.10b). The q values for each cluster (pale and dark pink in Figure 3.10b) were around 0.5.

The separate analysis of the last subset of Q2 data revealed clusters with the peaks of ΔK at K=3 Figure 3.9c). When the threshold of each cluster proportion was considered >60% (exception of OBB1 and OBB4), the three clusters correspond chiefly to North (OUC, OTVE, OMA and OPP, OBH3, OBH5, BMtL1, BDC2, ADC1, ADC3) and South (OMtB, OMC, OMB, OML, OBR, OBB and OBH, ADC2, BWT1, COO2) populations of *E. obovatus*, and *E.* sp. Mt. Bellenden Ker (BCL, BPR, BWT, BMtL, BTR, BDC, OBB1, OBH2, and COO2) (Figure 3.10c). The *E. obovatus* North cluster exhibits little to no admixture. The q values of this cluster ranged from 0.501 to 0.992 (average 0.933) and were greater than 0.836 for 95% of individuals. In contrast, the *E. obovatus* South cluster showed little (0.95 < OML, OMtB, OMC2, OMC3 OBB2-3, OBR4, OBH4 and COO2), moderate and high genetic admixture. The q values ranged between 0.686 (0.585) and 0.989 (average 0.915) and were greater than 0.746 for 95% of individuals. The *E.* sp. Mt. Bellenden Ker cluster also showed little (> 0.95: BPR, BWT2,

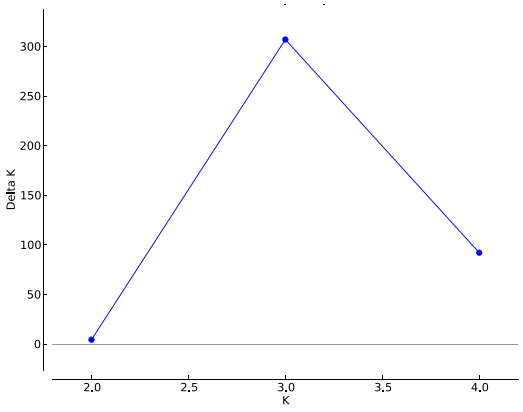
BWT4, BWT5, BMtL2-4 and BTR), moderate and high genetic admixture. The *q* value ranged from 0.633 (0.57) to 0.994 (average 0.947) and were greater than 0.713 for 95% of individuals.



a.



b.

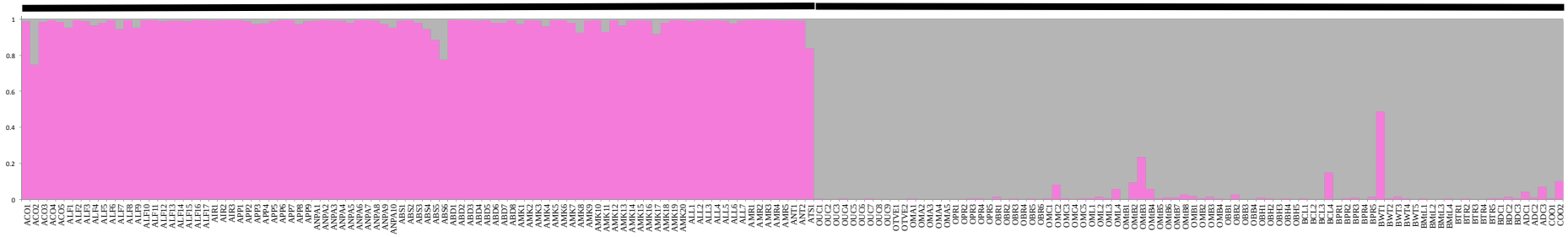


c.

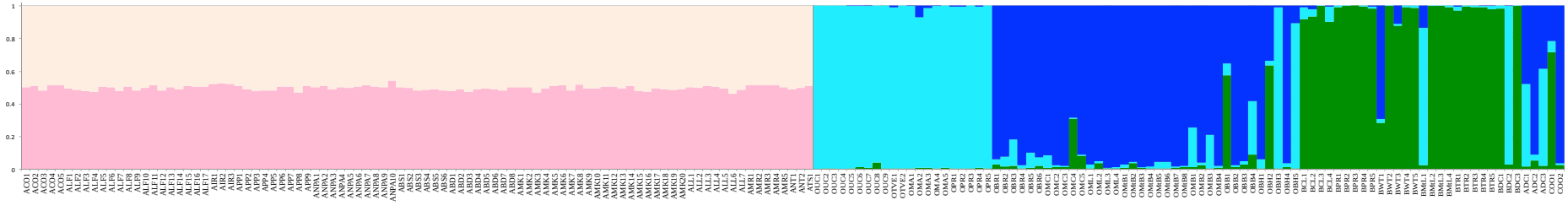
Figure 3.9. Plot of ΔK following the Evanno method. a. all the samples combined. b. samples of *E. arnhemicus*, c. samples of *E. obovatus* , *E.* sp. Mt. Bellenden Ker *E. coorangooloo*, ADC and BDC. The peaks are found at K=2, K=2 and K=3 for all the samples combined, samples of *E. arnhemicus* and samples of *E. obovatus*, *E.* sp. Mt. Bellenden Ker, *E. coorangooloo* ADC and BDC, respectively.

Q1: *E. arnhemicus*

Q2: *E. obovatus*, *E.* sp. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC



a.



b.

c.

Figure 3.10. STRUCTURE plots of a. all samples combined (K=2), b. *E. arnhemicus* only (K=2), c. *E. obovatus*, *E.* sp. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC (K=3) analysed separately. The 'tetraploid' ploidy level setting was used. Population codes for the sampling sites are provided in Table 3.1. All the plots depict the lowest Ln probability.

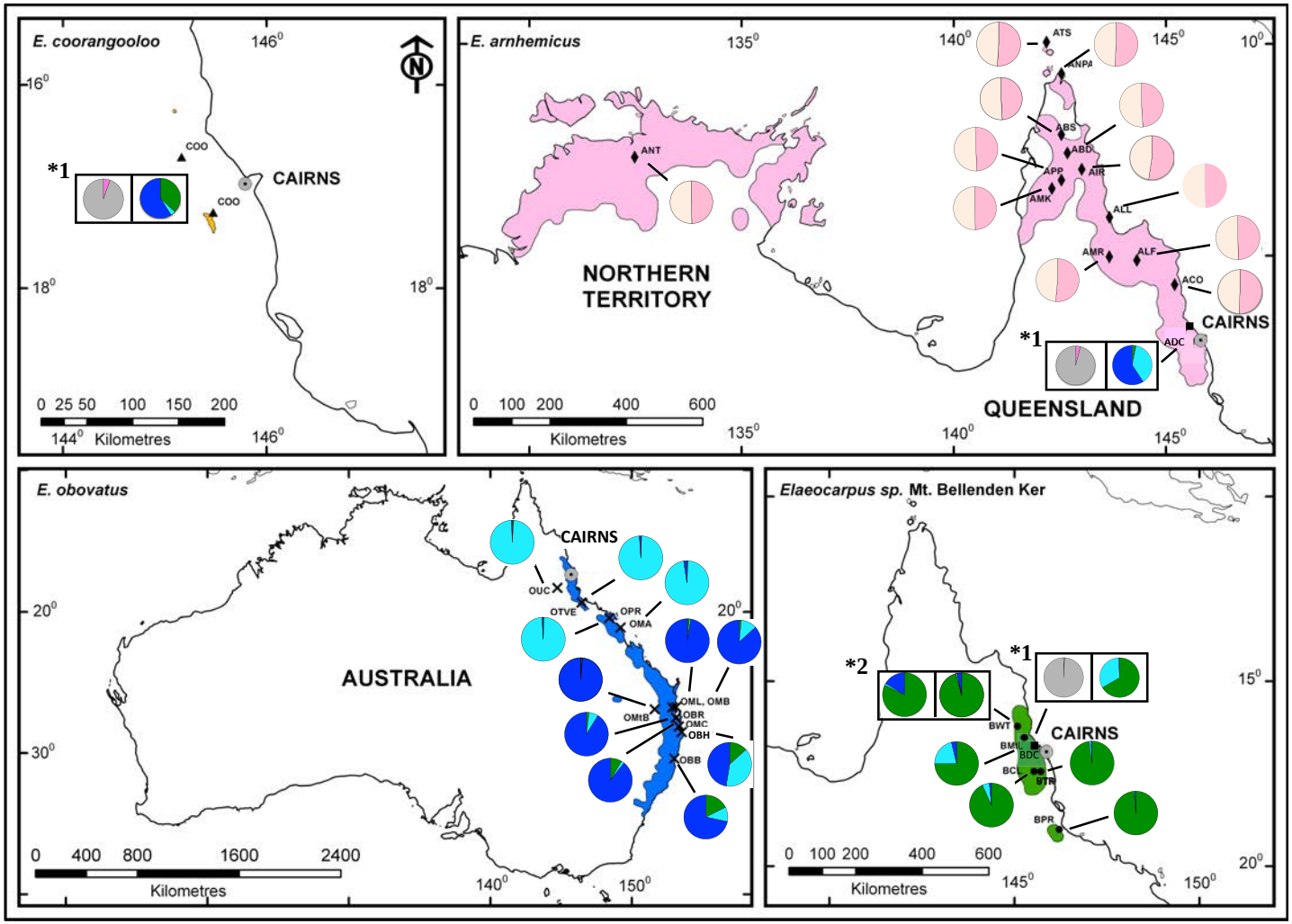


Figure 3.11. Genetic admixture averaged over populations. The colours correspond to the STRUCTURE plots in Figure 3.9 a, b and c. *1 left; average *q* of Figure 3.9. a; Right, average *q* of Figure 3.9.c. *2: BWT population, Right, all samples included; Left, BWT1 samples excluded. See table 3.1. for the population codes.

3.3.5 Hypothesis testing for evolutionary units

3.3.5.1 Analyses of Molecular Variance

Results of AMOVA analyses performed on all hypothesised groups including clusters proposed by PCoA and STRUCTURE analyses are shown in Table 3.7. All values were significant ($P < 0.5$) indicating that genetic variation in the data set is hierarchically structured (Table 3.7). For all entities combined, approximately 50 - 90 % of the diversity was within populations and 10 - 30 % between populations (Table 3.7).

3.3.5.1.1 Between cytotypes

The results of the AMOVA analysis of diploids and samples showing ploidy levels greater than 2 indicated strong divergence between the two cytotypes: only 15% of the observed variation was shared ($\Phi_{PT} = 0.438$).

3.3.5.1.2 Between *E. arnhemicus* (Q1) and others (Q2)

The results showed strong divergence between clusters Q1 and Q2 ($\Phi_{PT} = 0.425$) with 23% of variation shared.

3.3.5.1.3 Clusters within Q2

The divergence within Q2 (*E. obovatus* North, *E. obovatus* South, and *E.* sp. Mt Bellenden Ker) was high ($\Phi_{PT} = 0.311$) with only 16.7 % of variation shared amongst the subclusters (Table 3.7).

The AMOVA analyses between *E. obovatus* North and 1) *E. obovatus* South, 2) *E.* sp. Mt. Bellenden Ker 3) *E. obovatus* South and *E.* sp. Mt. Bellenden Ker produced Φ_{PT} values of 0.308, 0.306 and 0.325 respectively (Table 3.7). The amount of variation shared among each pair was 17.18%, 20.42% and 10.28% respectively.

Similarly, the AMOVA between *E. obovatus* South and 1) *E.* sp. Mt. Bellenden Ker 2) *E. obovatus* North and *E.* sp. Mt. Bellenden Ker produced Φ_{PT} values of 0.292 and 0.321 respectively (Table 3.7). The amount of variation shared among each pair was 11.78% and 8.63% respectively.

3.3.5.1.4 Between ADC and BDC, and all other genetic clusters

Although the sample size for each population of ADC and BDC was small, the significant genetic differentiation between them, and the other entities (*E. obovatus* North *E. obovatus* South, and *E.* sp. Mt. Bellenden Ker) was supported by the AMOVA (Table 3.8).

The AMOVA between ADC and *E. obovatus* North, *E. obovatus* South and *E.* sp. Mt. Bellenden Ker produced high values of Φ_{PT} (0.349, 0.332 and 0.286 respectively, Table 3.8) and no significant differences were found between ADC, BDC, and *E. coorangooloo*.

Similarly, the analyses show that BDC is divergent from *E. obovatus* North and South (Table 3.8: Φ_{PT}= 0.298 and 0.242 respectively). On the contrary, the results do not support differentiation between BDC and *E.* sp. Bellenden Ker.

Table 3.7. Analysis of molecular variance (AMOVA) of microsatellite data performed with different groupings.

Source of variation	Df	Est. var.	Percent	Φ_{PT}
Diploids vs tetraploids				
Among cytotypes	1	2.130	15.221	0.438*
Among populations	28	4.000	28.550	
Within populations	150	7.870	56.230	
Tetraploids (*E. arnhemicus and E. obovatus*)				
Among tetraploid taxa	1	4.512	33.043	0.464**
Among populations	20	1.824	13.355	
Within population	127	7.319	53.602	
Diploids (*E.* sp. Mt. Bellenden Ker, ADC, BDC vs *E. coorangooloo*)				
Among populations	7	2.969	20.885	0.209**
Within populations	22	11.246	79.114	
Between clusters of *E. arnhemicus* (Q1) vs others (Q2)				
Among the clusters	1	3.764	23.036	0.425**
Among the populations within each cluster	28	3.182	19.479	
Within the populations	150	9.391	57.484	
Between subclusters within Q2				
Among the clusters	2	2.421	16.766	0.311**
Among populations within each subcluster	16	12.017	17.053	
Within the population	68	9.949	68.948	
***E. obovatus* vs *E.* sp. Mt. Bellenden Ker**				
Among entities	1	1.761	12.369	0.305**
Among populations within each entity	15	2.575	18.089	
Within the populations	65	9.898	69.542	
***E. obovatus* North vs *E. obovatus* South**				
Among regions	1	2.361	17.179	0.308**
Among the population within each region	9	1.866	13.580	
Within the populations	45	9.515	69.241	
***E. obovatus* North vs *E.* sp. Mt. Bellenden Ker**				
Between clusters	1	2.702	20.428	0.306**
Among populations within each entity	8	1.347	10.187	

Source of variation	Df	Est. var.	Percent	Φ_{PT}
Within population	36	9.177	69.385	
***E. obovatus* South vs *E.* sp. Mt. Bellenden Ker**				
Between entities	1	1.817	11.766	0.292**
Among populations within each entity	13	2.688	17.408	
Within population	52	10.936	70.826	
***E. obovatus* North vs *E. obovatus* South + *E.* sp. Mt Bellenden Ker**				
Among clusters	1	1.513	10.281	0.325**
Among populations within each cluster	17	3.256	22.123	
Within population	68	9.949	67.598	
***E. obovatus* South vs *E. obovatus* North + *E.* sp. Mt Bellenden Ker**				
Among clusters	1	1.237	8.632	0.321**
Among populations within each cluster	17	3.363	23.474	
Within population	68	9.728	67.894	

df - degree of freedom, *Est. var.* - estimate of variance, percent - percentage of total variation,

*P< 0.01, ** P< 0.001, the probabilities of obtaining by chance an Φ_{PT} value equal or greater than the observed value, estimated from 999 permutation

Table 3.8. Pairwise Φ_{PT} value between ADC and BDC populations against other units within Q2.

	E. obovatus North	*E. obovatus* South	*E.* sp. Mt. Bellenden Ker	ADC	BDC	E. *coorangooloo*
ADC	0.349**	0.332**	0.286**		Not significant	Not significant
BDC	0.242**	0.298**	Not significant	Not significant		Not significant

** P< 0.001, the probabilities of obtaining by chance an Φ_{PT} value equal or greater than the observed value, estimated from 999 permutation

3.3.5.2 Multigroup Discriminant Function Analyses

Multigroup Discriminant Function Analyses were undertaken to test the significance of hypotheses of evolutionary unit membership based on the results of the previous analyses: *E. arnhemicus*, *E. obovatus* North, *E. obovatus* South, *E.* sp. Mt. Bellenden Ker, BDC, ADC and *E. coorangooloo*. MDFA omits samples containing missing data, which resulted in the complete exclusion of *E. coorangooloo*.

The analyses most strongly supported five groups: *E. arnhemicus*, *E. obovatus*, *E.* sp. Mt. Bellenden Ker (BDC included) and ADC (Figure 3.12). All samples were classified into the same groups as the previous analyses, and 92.6 % of samples were correctly classified in the

subsequent cross validation process. Misclassifications in the cross validation were found mainly between *E. obovatus* South and *E. obovatus* North populations and samples of ADC (Table 3.9). In the preliminary analyses, which included BDC in a preassigned group with ADC, samples of BDC were classified as *E.* sp. Bellenden Ker (results not shown).

Table 3.9. Summary of incorrect classification found in the cross validation process.

Samples	Original (priori) classification	Incorrect classification in the cross validation
ATS1	*E. arnhemicus*	*E. obovatus* South
OTVE1	*E. obovatus* North	*E. obovatus* South
OMB3	*E. obovatus* South	*E. obovatus* North
OBB1	*E. obovatus* South	*E. obovatus* North
BCL4	*E.* sp. Mt. Bellenden Ker	*E. arnhemicus*
ADC1	ADC	*E. arnhemicus*
ADC2	ADC	*E.* sp. Mt. Bellenden Ker

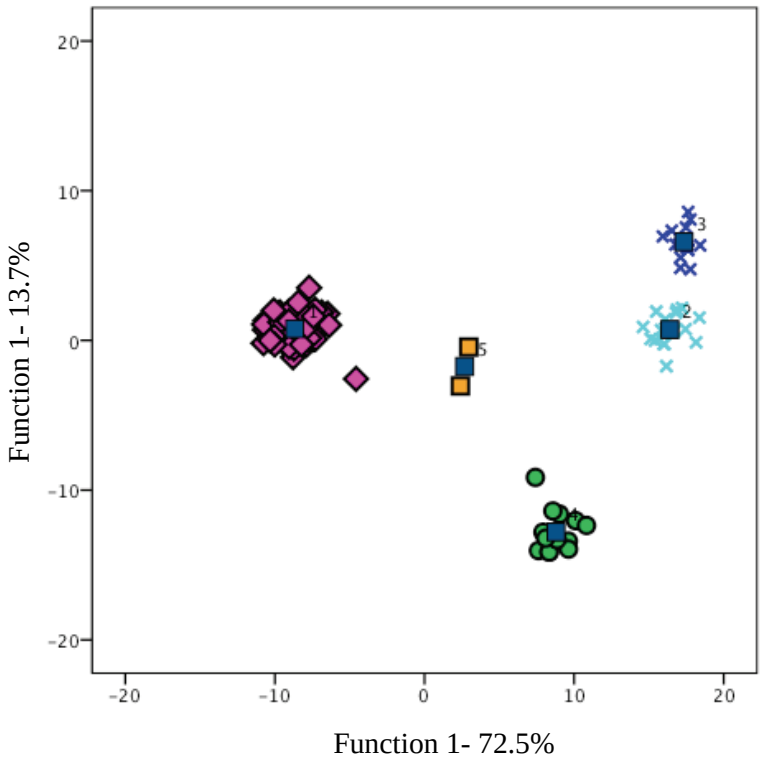


Figure 3.12. Ordination plot of the results of the Multigroup Discriminant Function Analysis on the binary dataset. A total of 86.2% of the observed variation was accounted for in function 1 (72.5%) and function 2 (13.7%). ◆ *Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue, *E. obovatus* North and dark blue, *E. obovatus* South), ● *E. sp.* Mt. Bellenden Ker (green) and ■ population ADC (orange). Closed squares are group centroids.

3.4 Discussion

3.4.1 Ploidy assessment

While there are a number of cytogenetic studies of *Elaeocarpus* (Rattenburry, 1947; Allan, 1961;; Carr & McPherson, 1986; de Lange et al., 2004; Mehara & Sareen, 1973; Ono, 1975), in only two cases were base chromosome numbers inferred from haploid cells (pollen grains in metaphase II) (Carr & McPherson, 1986; Mehara & Sareen, 1973). In these studies, the base chromosome number was x=15. In contrast, the studies which utilised somatic cells showed that 2n=30 (*E. hookerianus* (Rattenburry, 1947; Allan, 1961 and de Lange et al., 2004), *E. lancaefolius* (Mehara & Sareen, 1973) and *E. speciosus* (Carr & McPherson, 1986)), 2n=28 (*E. photiniaefolius* (Ono, 1975)) and 2n=32 (*E. angustifolius* (Gamage & Schmidt, 2009)).

This level of variation in only three studies illustrates the importance of determining ploidy from haploid cells and of this information for ongoing studies such as inheritance patterns of tetraploidy. As haploid cells are only available during the reproductive phase of the plant, while not desirable, it is therefore often necessary to infer ploidy using a combination of techniques such as chromosome counts in somatic cells, flow cytometry, and allele counts from microsatellite data. While inferring ploidy from microsatellite profiles only is also risky, the consistent occurrence of more than two allele peaks (Figure 3.2 and Table 3.6) indicates *E. arnhemicus* and *E. obovatus* may be polyploid. Over all loci, *E. arnhemicus* showed more frequent occurrence of three alleles than *E. obovatus*. This result poses the question: is *E. arnhemicus* triploid, or tetraploid with one of the allele being homozygous. It is noteworthy that many *E. arnhemicus* samples used in this study were collected from trees bearing fruit containing morphologically normal seed, and the species is known for its abundant crops of fruit. Although the viability of these apparently normal seed was not determined, the abundant production suggests the plants are likely to be tetraploid rather than triploid.

As it was not possible to re-collect and repeat this study of ploidy, while it would be preferable, extensive ploidy screening using chromosome counting technique in somatic cells combined by flow cytometry using new samples to determine the frequency of polyploidy would be beneficial.

3.4.2 General patterns of genetic clustering and genetic admixture

Several groups were consistently resolved in most analyses namely *E. arnhemicus, E. obovatus* and *E.* sp. Mt. Bellenden Ker. The hypothesis of *E. obovatus* being further divided into Northern and Southern groups was supported by some analyses (see Figure 3.3, Figure 3.6,

Figure 3.7 and Figure 3.10) and challenged by others (see Figure 3.8 and Figure 3.12). Sometimes loose groupings were found between *E. sp*. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC, and some analyses also supported genetic affinity between samples (Figure 3.3, Figure 3.6 and Figure 3.7).

3.4.2.1 *E. arnhemicus*

The genetic cluster Q1 identified by the STRUCTURE analysis contains all the samples of *E. arnhemicus* and shows little to no admixture with Q2, which contains all other samples in the analysis. This genetic cluster was the most consistently supported in all the analyses (Figure 3.3, Figure 3.5, Figure 3.6, Figure 3.7, Figure 3.10 and Figure 3.12). Differences in flowering phenology may be one of the factors responsible for maintaining genetic isolation between *E. arnhemicus* and the other entities within the group. The range of floral phenology recorded in herbarium specimens of *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker are mid April to mid August (January onwards in the Northern Territory), early September to early December, and late October to early December respectively. Flowering times of *E. coorangooloo* and *E. arnhemicus* overlap in March to August, however the geographical and ecological differences between the two species is likely to contribute strongly in limiting gene exchange between the two species.

The STRUCTURE analysis identified two subclusters within *E. arnhemicus* (Q1) (Figure 3.11b). However, the AMOVA analysis did not support genetic differentiation between these clusters (Φ_{PT} = 0.049). This admixture of the two genetic clusters could be explained by fruit dispersal mechanisms. The dispersal of the seeds is possibly associated with the Pied Imperial pigeon (*Ducula bicolor*) which migrates from New Guinea to Australia in August - October and returns in March - April (Storr, 1973, 1977) coinciding with fruiting of *E. arnhemicus*. Birds of this species have been observed feeding on *E. arnhemicus* fruit in Cooktown, QLD (F. Venter pers. comm). The birds' daily movement is approximately 14 km and up to 80 km when they commence migration (Price, 2006). Provided the pigeon retains the fruit for up to 540 minutes, which is considerably longer than for many frugivores, and defecates the seeds intact (summarized in Corlett, 1998), long distance dispersal across the Torres Strait of fruit for both clusters of this species is entirely plausible. There were no samples from New Guinea available for inclusion in this genetic study. The morphological study described in Chapter 4 however, confirms that there are no differences in fruit size or floral characters between the samples from Queensland, Northern Territory and Papua New Guinea, based on herbarium specimens. It is crucial that future genetic research includes more intensive sampling throughout the range of this species to test the hypothesis of genetic admixture mediated by seed dispersal.

Elaeocarpus arnhemicus also often occurs at the edge of banks of river systems in Cape York Peninsula, which support a more mesic environment all year round and are subject to inundation by annual floods. It is possible that such floods may disperse seeds of *E. arnhemicus*. While there is no evidence that *Elaeocarpus* seeds float (conversely, *E. bifidus* seeds are non-buoyant; Carlquist, 1966), violent floodwaters might move submerged seeds considerable distances. Seed viability might be maintained due to the stones remaining imbibed (Khan et al., 2003).

3.4.2.2 *E. obovatus*

The separation of samples of *E. obovatus* North from those of *E. obovatus* South was indicated by the STRUCTURE analysis. This pattern was supported by the genetic distance measure Φ_{PT}, confirming significant difference between *E. obovatus* North and *E. obovatus* South. However, a single *Elaeocarpus obovatus* group was supported by PCoA using the Jaccard's similarity coefficient (Figure 3.3) Bruvo distance (Figure 3.3), and Multigroup Discriminant Function Analysis (Figure 3.12). The possibility of the STRUCTURE pattern (north-south differentiation) being an artefact of a major sampling gap of about 800 km across coastal Central Queensland (Figure 3.1) cannot be ruled out. As highlighted in the STRUCTURE results, the high rate of genetic admixture with genetic components derived from other clusters within the samples of *E. obovatus* North does not support the existence of two discrete evolutionary units within *E. obovatus*. Further investigation is required with finer sampling in Central Queensland to confirm this finding. Until such time, it would be premature to conclude the existence of distinct northern and southern entities within this species.

3.4.2.3 *E.* sp. Mt. Bellenden Ker

Elaeocarpus sp. Mt. Bellenden Ker forms a distinct genetic entity, well separated from *E. arnhemicus*, *E. obovatus* North, *E. obovatus* South and *E. obovatus*. The cluster was supported by the PCoA using Jaccard's similarity coefficient (Figure 3.3), Nei's pairwise population distance (Figure 3.6), Φ_{PT} pairwise population (Figure 3.7), STRUCTURE (Figure 3.10). The results of AMOVA (Table 3.7) and MDFA (Figure 3.12) further support *E. sp.* Mt. Bellenden Ker as being a distinct evolutionary unit.

STRUCTURE analysis revealed that the level of genetic admixture with the other entities in the *E. obovatus* North population is either low or negligible. A high proportion of genes derived from both *E. obovatus* North and South were prominent in BMtL and BWT respectively, the populations geographically most distant from *E. obovatus* sampled in this study. Moreover, the

STRUCTURE results showed little or no evidence of genetic admixture in BPR, the nearest population of *E.* sp. Mt. Bellenden Ker from *E. obovatus* North (Figure 3.11). This implies a lack of genetic exchange in a potential contact zone and suggests that the genetic admixing found in *E.* sp. Mt. Bellenden Ker is not the result of genetic exchange between populations of these entities but rather shared ancestral alleles.

Low genetic admixture in the populations of *E.* sp. Mt. Bellenden Ker was observed in the wet highlands of BTP (Topaz and Mt. Bartle Frere) and BPR (Paluma Range), areas identified by palaeo-climatic modelling as 'stable' strongholds for rainforest during climatic oscillations since the last glacial maximum (VanDerWal et al., 2009). In contrast, the populations showing high genetic admixture are from localities such as BWT (Mt. Windsor Tableland) and BMtL (Mt. Lewis), areas that are historically less stable (VanDerWal et al., 2009) and have been disturbed or invaded by drier rainforest types. The effects of range expansion and range contraction may have favoured stronger selective pressures and/or genetic exchanges in these areas (Rieseberg & Willis, 2007).

3.4.2.4 ADC and BDC

Populations from Davies Creek exhibit morphologies that are somewhat intermediate between the three entities but more closely resemble those of *E. arnhemicus* (ADC) or *E.* sp. Mt. Bellenden Ker (BDC), and therefore defy identification within the current classification. Two populations of Davies Creek from two different altitudes were sampled in this study (ACD and BDC) and showed different relationships.

The results of some cluster analyses (Figure 3.3,Figure 3.4 and Figure 3.8) and STRUCTURE (Figure 3.10) suggest the existence of genetic admixture between BDC and other populations of *E.* sp. Mt. Bellenden Ker. This population occurs at mid altitude (approx. 650m abs.), which is noteworthy considering there are no collections of *E.* sp. Mt. Bellenden Ker from more suitable habitat at higher altitudes in this part of the Lamb Range. The genetic similarity between BDC and *E. obovatus* North was also supported by the results of STRUCTURE analyses (only BDC2); however, the cause of the admixture between these populations is not clear. A possible explanation may be stochastic shared ancestral polymorphism. Given that there is some indication from the microsatellite profiles that populations of *E. obovatus* North are tetraploid and BDC is diploid, hybridization between these entities is unlikely. The retention of ancestral alleles is plausible for tree species where the generation time is relatively long compared to herbs or shrubs (Halverson et al., 2008; Münzbergová et al., 2013). The results presented here nonetheless reinforce the ambiguous status of BDC.

ADC, the other population from Davies Creek found in a different habitat at a lower altitude (approx. 400m abs.), is distinct from the other entities. However, the STRUCTURE results indicated that the samples of ADC show the greatest degree of genetic admixture of all samples in this study (Figure 3.10a and c). ADC shares a substantial proportion of its genetics with *E. obovatus* North and South, but not with the geographically close *E.* sp. Mt. Bellenden Ker or *E. arnhemicus*. This result raises more questions about the origin and identity of this community. Finer sampling from this locality and a thorough study of the ecology and habitat preferences of the entities in the vicinity will aid in drawing further conclusions about the origin and identity of these populations.

3.4.2.5 *E. coorangooloo*

The results of this study also highlight the ambiguous genetic identity of *E. coorangooloo.* The results of clusters (Figure 3.3, Figure 3.4, Figure 3.8), STRUCTURE (Figure 3.11) and AMOVA (Table 3.8) indicate that samples of *E. coorangooloo* are genetically close to those of *E. obovatus, E. sp*. Mt. Bellenden Ker, ADC and BDC. While this taxon is morphologically distinct from the other entities in the group it also requires a more thorough sampling strategy to obtain a comprehensive and more accurate understanding of its genetic relationships to the other members of the group.

3.5 Future research

Determination of the ploidy level was attempted in this study from microsatellite profiles. There are some indications from constant >2 peaks of microsatellite alleles in *E. arnhemicus* and *E. obovatus* samples and two alleles in *E. sp*. Mt. Bellenden Ker. Detailed screening of cytotypes is crucial for the mixed polyploidy groups to thoroughly characterise the exact cytotypes across the group – this will require extensive screening, e.g. by flow cytometry methods, of sufficient samples from across the geographical range will be required to confidently ascertain the ploidy levels of each species. This information will contribute to the accurate assignment of allele dosage analysis, which will reduce the ambiguity in the genotyping for polyploid entities. Furthermore, cytomorphological study is required to ascertain the inheritance pattern of the polyploid species.

Previous work has suggested that the origins of this group may date to c. 10 - 15 million years ago when the group separated from its sister (Crayn et al., 2006; estimated using *E. arnhemicus*). However, this analysis included only a few *Elaeocarpus* species and therefore the

divergence data estimates should be taken to be preliminary. Therefore it is not plausible to postulate the evolutionary scenario of the entities of this group, but it is likely that the drying of the continent in the Tertiary may have allowed the evolution of dry-adapted species such as *Elaeocarpus arnhemicus*. Further investigation of the origins and timing of evolutionary events in the genus, including polyploidisation events, can only be addressed when an accurately dated phylogeny, thoroughly sampled for the study group, is achieved.

Further molecular and ecological analysis with more intensive sampling for BDC and ADC populations will give a clearer understanding of genetic structure, gene flow, and migration patterns of entities within contact zones with other populations and the other entities in the broader Atherton Tableland Area. Furthermore, finer sampling of *E. obovatus* from Central Queensland is required to determine whether there are populations that show genetic admixture between *E. obovatus* North and *E. obovatus* South.

3.6 Conclusions

Analysis of the population genetic structure within the *Elaeocarpus obovatus* species complex revealed a complex evolutionary history. The major findings of this chapter include:

1. Patterns of genetic structure within and among entities broadly corresponds to the current taxonomic boundaries.
2. Working hypotheses regarding appropriate taxonomic ranks for the entities were determined: three equally distinctive genetic entities, *E. arnhemicus* and *E. obovatus* and *E.* sp. Mt. Bellenden Ker.

In the following chapter, patterns of morphological diversity are explored and discussed. The working entities will be used as hypotheses to test if the morphological groupings are congruent with the genetically based groupings. Taxonomic decisions regarding ranks and circumscriptions are made in the light of the results of the following chapter.

Chapter 4 Resolving taxon limits in the *Elaeocarpus obovatus* species complex (Elaeocarpaceae) - a morphometric approach

4.1 Introduction

Elaeocarpus L. is the largest genus in the family Elaeocarpaceae (Oxalidales). It is predominantly distributed in the Old World tropics and subtropics, from southern India throughout Southeast Asia to Australia, with outliers in Madagascar, Mauritius, Japan, Hawaii and other Pacific Islands, and New Zealand. The highest species diversity is found in Papuasia (Coode, 2004; Zmarzty, 2001). Recent discoveries of putative new species in both Australia and New Guinea contribute to a growing number of species in the genus, estimated at over 350 spp. (Zmarzty, 2001; Coode, 2002, 2005; Tang and Phengklai, 2007).

The first Australian species of *Elaeocarpus* to be described was *E. reticulatus* Sm., by the British botanist J.E. Smith (1809). In the last regional account in *Flora Australiensis* (1863), Bentham recognised four more. A further 24 species and one subspecies were described over the next 150 years. Coode (1984) completely revised the genus in Australia, adding four species, and since then only two additional species have been described (Maynard et al., 2008; Baba & Crayn, 2012). Despite Coode's thorough morphological examination, several taxonomic problems among the Australian species remain to be addressed. Among these is the *Elaeocarpus obovatus* complex, comprising *E. obovatus* G.Don, *E. arnhemicus* F.Muell., and various intermediate forms.

Elaeocarpus obovatus is distributed along the east coast of Australia from 18° 17' (Undara, QLD) to 34°19' (Bulli, NSW), and *E. arnhemicus* is distributed from Java eastwards to southern Papua New Guinea (PNG) and southward to 17° 55' (Mission Beach, QLD). *Elaeocarpus obovatus* is found in dry rainforest and wet sclerophyll forest at low to mid altitudinal ranges, whereas herbarium records suggest *E. arnhemicus* occurs in relatively drier environments such as gallery forests, monsoon scrub and consolidated beach dunes. These ecological differences between the two species are notable, however, there are no studies to confirm whether species distribution corresponds strictly to these niches. Additionally, intermediate forms are found where the distributions overlap (Coode, 1984).

Coode (1984) identified seven 'variants' and 'intermediates' within the *E. obovatus* group based on the number of petal divisions, ratio of leaf length to width, petiole length, fruit size, and geographical distribution (Table 4.1). One of these variants is now referred to as *Elaeocarpus*

sp. Mt Bellenden Ker (L.J.Brass 18336) Qld Herbarium. These characters, however, were not systematically documented across the variants and numerous specimens have been collected since Coode's (1984) revision. Therefore a thorough, systematic investigation and assessment of the taxonomic status of these variants and intermediates is now required. Moreover, *E. coorangooloo* J.F.Bailey & C.T.White was identified as a member of the clade by molecular phylogenetic analysis (Chapter 2) and shows close morphological affinities to the *E. obovatus* group, casting a question over its species status.

Population genetic analysis (Chapter 3) revealed that three entities - *E. arnhemicus* (excluding the populations from Davis Creek, Dinden National Park), *E. obovatus* and *E.* sp. Mt Bellenden Ker - are genetically distinct. Two populations from Davies Creek are morphologically similar to *E. arnhemicus* and *E.* sp. Mt Bellenden Ker. One of these (coded as ADC in Chapter 3) was shown to be genetically closest to *E. obovatus*, whereas the other (BDC) is genetically indistinguishable from *E.* sp. Mt. Bellenden Ker. However, broader sampling is needed before firm conclusions regarding their genetic distinctiveness and relationships can be drawn. In the present chapter, code ADC is retained to identify this morphologically and genetically different group and these specimens are treated separately from the other entities.

While genetic information may provide evidence of species boundaries, a practically useful classification requires that taxa are characterised by morphological features. Systematic investigation of morphological variation is therefore crucial to help understand taxonomically complex species groups.

Morphometric approaches, and more specifically, multivariate analyses are often used to examine joint relationships of characters measured (James & McCulloch, 1990), and are therefore useful in assessing diverse morphological characters. Principal Component Analysis (PCA), cluster analysis and multidimensional scaling (MDS) are the most commonly used multivariate approaches in taxonomy (Stuessy, 2009) and are routinely used as a first step to assess species limits (examples for plants include Fatemi et al., 2007; Henderson & Ferreira, 2002; Lihová et al., 2010; invertebrates: Lattig et al., 2007; and vertebrates: Glaw et al., 2012, Victoriano et al., 2010). PCA uses linear correlation to reduce the multidimensionality of the original dataset, and displays the overall pattern of variation in ordination space (Lihová et al., 2010). MDS utilises various distance metrics between variables and projects the groupings in multidimensional planes. Rohlf (1972) concluded, from an empirical and comparative study between three ordination techniques, PCA, MDS and Principal Coordinate Analysis (PCo), that MDS outperformed PCA or PCo in 2- or 3- dimensional summary of phenetic relationships.

Hartman (1988) recommended examining both MDS and PCA or PCo to improve accuracy in the representation of variables.

In this chapter, I present the results of a morphometric analysis of the variation in the *Elaeocarpus obovatus* species complex, with a specific focus on the following questions:

1. Are morphological groupings discernible?
2. Are these groupings congruent with the molecular groupings? and
3. Which morphological variables are taxonomically informative, in that they correlate with the groupings?

Table 4.1. Comparison of morphological characters of *E. obovatus*, *E. arnhemicus*, their variants and intermediates, and the related species *E. coorangooloo*. Information and remarks were extracted from Coode (1984). N/A: information not available in Coode (1984).

	Remarks	Plant height (m)	Number of petal divisions	Ovules/ carpel	Leaf length (cm)	Leaf width (cm)	Ratio of leaf length to width	Petiole length (cm)	Fruit size length x width (mm)	Distribution
E. obovatus	Glabrous leaves, petioles and flowers	3-36	4-9	4	(3.5-) 4.8-5 (-9)	(0.8-) 1.4-2.2 (-2.6)	3 times longer than wide	2-6 (-12)	10 x 8	NSW (Bulli to the northern border), QLD (southern border to Townsville; Undara National Park; Bakers Blue Mountain)
E. obovatus 'var'	Broad leaved variant, glabrous petals	3-12	(3?-4-) 5-8 (-9)	N/A	4.5-8.5	1.6-3.4	2.4-3.7 times longer than wide	(2-)3-5 (-6)	c. 9 x 8	Qld and NSW (Sympatric with the typical *E. obovatus* north of Brisbane)
E. arnhemicus	Petiole more or less hairy, finely pubescent on inflorescence at least when young.	3-10	10-16	4	6-12-5	(2.2-) 2.6-5	Rarely more than 2.8x longer than wide	(3-) 5-12 (-14)	12-16 X 10-12	QLD (Cook district), NT
E. obovatus-N. Queensland 'intermediates' 1	Broad leaved intermediate. Indumentum on midribs, petioles, young twigs, and inflorescence axes	N/A	7-10	N/A	N/A	N/A	Broad	N/A	N/A	QLD (Strathdicke near Proserpine)
E. obovatus -N. Queensland 'intermediates' 2	Smaller leafed intermediate. Largish fruit.	Small	(7-) 9-10 6-12	N/A	Smaller	Smaller	N/A	N/A	Largish	QLD (Annan Gorge S of Cooktown, and NW Laura)

	Remarks	Plant height (m)	Number of petal divisions	Ovules/ carpel	Leaf length (cm)	Leaf width (cm)	Ratio of leaf length to width	Petiole length (cm)	Fruit size length x width (mm)	Distribution
E. obovatus-N. Queensland 'intermediates'3	Mid- large tree, narrow leaved intermediate. Midrib of leaves and petiole glabrous.	≤ 25	5-8 (-9)	4 <	N/A	N/A	Moderately narrow	N/A	N/A	QLD (Inland of Innisfail)
E. obovatus-N. Queensland 'intermediates'4	High numbered ovules, maybe a hybrid with *E. coorangooloo.* Petals pubescent.	N/A	N/A	(7-) 8	N/A (Hyland 5955: 4-6.3)	N/A (Hyland 5955: 1.7-2.5)	N/A	N/A/ (Davies Creek specimens: petioles relatively long)	N/A	QLD (Near Mareeba and Davies Creek)
E. obovatus-N. Queensland 'intermediates'5	Mid-large tree, long-acuminate papery leaved intermediate.	≤ 40	N/A	N/A	N/A	N/A	N/A	5-10	c. 9	QLD (Mt. Lewis, Mt. Spec and Davies Creek)
E. coorangooloo	Overall appearance is more robust than all of the entities.	15 ≤	c.7, fringes are arranged into blunt lobes.	(3-) 6	(6-) 7-11(-14)	(2.5-) 3-5.5	C.2 x longer than wide	1-3 cm	c. 12 x 9	QLD (Tolga, Atherton, and Baker's Blue)

4.2 Materials and methods

4.2.1 Sampling

Specimens of *E. obovatus* complex were collected from 32 sites in north-east Queensland (Cape York Peninsula and the Cairns area), and 13 sites in south-east Queensland (Table 4.2). The sampling scheme was designed to capture as much of the geographical and ecological range as possible with a focus on areas where two or more putative entities were sympatric. To compare the variation within the group across its distribution, supplemental herbarium specimens from Papua New Guinea and the Northern Territory were included. In total, 94 specimens from the *E. obovatus* complex were examined. Most of the specimens used for morphometric analyses were from the same individuals as the samples used for the population genetic study (Chapter 3). The samples used in the previous chapter are indicated in Table 4.2.

Collections were made under permits WISP05391710 and WITK04740210, issued by the Queensland Department of Science, Information Technology, Innovation and the Arts. All specimens including spirit material have been lodged at CNS.

4.2.1.1 Character selection and dataset preparation

Measurements of vegetative characters were made on dried specimens and those of reproductive characters (flowers and fruit) were made on rehydrated material. When spirit material was not available, reproductive parts, inflorescences and fruit, were sampled from the herbarium sheets and prepared by boiling them in a small quantity of water using a commercially available microwave for up to 2 minutes or until parts are well rehydrated, and stored in 70% ethanol.

Nine vegetative, 19 floral and four fruit characters were selected for analysis based on extensive observations on both field and herbarium specimens (Table 4.2). These characters were scored for 102 specimens: 46 of *E. arnhemicus*, 34 *E. obovatus*, 14 *E. sp*. Mt. Bellenden Ker, six *E. coorangooloo* and two ADC. For each character five measurements, where possible, per specimen were made and the average calculated. Character states and definitions are given in Table 4.3

Table 4.2. Plant specimens used in the present chapter. *1 FL: Flower, FR: Fruit. Vegetative characters were measured from all the specimens listed below. *2 DNA: samples used in population genetics study in Chapter 3.

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
E. arnhemicus	Annan River, South of Cooktown, Cape York Peninsula, QLD	Baba 477 (CNS)	FL, FR	-15 40 53	145 12 21	DNA
E. arnhemicus	Archer Point, South of Cooktown, Cape York Peninsula, QLD	Baba 479 (CNS)	FR	-15 35 25	145 16 32	DNA
E. arnhemicus	Archer Point, South of Cooktown, Cape York Peninsula, QLD	Baba 480 (CNS)	FR	-15 35 25	145 16 32	DNA
E. arnhemicus	Kennedy Bent, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 484 (CNS)	FR	-15 6 23	144 18 53	DNA
E. arnhemicus	Hann Crossing Campsite 10, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 489 (CNS)	FR	-14 45 42	144 4 39	DNA
E. arnhemicus	Hann Crossing Campsite 10, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 491 (CNS)	FL, FR	-14 45 43	144 4 40	DNA
E. arnhemicus	Midway Waterhole, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 497 (CNS)	FL, FR	-14 53 15	144 12 23	DNA
E. arnhemicus	Breeza Plains Outstation, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 494 (CNS)	FR	-14 49 35	144 6 45	DNA
E. arnhemicus	Breeza Plains Outstation, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 496 (CNS)	FL	-14 49 34	144 6 47	DNA
E. arnhemicus	Saltwater Creek, Rinyirru (Lakefield) National Park, Cape York Peninsula, QLD	Baba 507 (CNS)	FR	-14 37 2	143 53 56	DNA
E. arnhemicus	Portland Road, c. 94km West of Chilli Beach, near Iron Range National Park, Cape York Peninsula, QLD	Baba 547 (CNS)	FL, FR	-12 57 30	143 0 53	DNA
E. arnhemicus	Picaninny Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary, Cape York Peninsula, QLD	Baba 549 (CNS)	FL, FR	-13 12 55	142 44 24	DNA
E. arnhemicus	Picaninny Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary. Cape York Peninsula, QLD	Baba 551 (CNS)	FL	-13 12 55	142 44 24	DNA
E. arnhemicus	Picaninny Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary, Cape York Peninsula, QLD	Baba 552 (CNS)	FL, FR	-13 11 22	142 41 31	DNA
E. arnhemicus	Scrubby Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary, Cape York Peninsula, QLD	Baba 565 (CNS)	FL	-13 16	142 34	DNA
E. arnhemicus	Pajinka Walk, Northern Peninsula Area, Cape York	Baba 586 (CNS)	FL, FR	-10 41 44	412 31 56	DNA

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
	Peninsula, QLD					
E. arnhemicus	Muddy Bay, Northern Peninsula Area. Cape York Peninsula, QLD	Baba 587 (CNS)	FR	-10 41 44	412 31 56	DNA
E. arnhemicus	Muddy Bay, Northern Peninsula Area, Cape York Peninsula, QLD	Baba 588 (CNS)	FL	-10 41 44	412 31 56	DNA
E. arnhemicus	Muddy Bay, Northern Peninsula Area, Cape York Peninsula, QLD	Baba 589 (CNS)	FL, FR	-10 42 4	142 32 13	DNA
E. arnhemicus	Mutee Head, Northern Peninsula Area, Cape York Peninsula, QLD	Baba 600 (CNS)	FL	-10 55 44	142 17 35	DNA
E. arnhemicus	Rocky Creek, Bramwell Station, Cape York Peninsula, QLD	Baba 603 (CNS)	FL	-12 8 25	142 37 22	DNA
E. arnhemicus	Rocky Creek, Bramwell Station Cape York Peninsula, QLD	Baba 606 (CNS)	FR	-12 8 25	142 37 22	DNA
E. arnhemicus	Rocky Creek, Bramwell Station, Cape York Peninsula, QLD	Baba 610 (CNS)	FR	-12 10 16	142 34 29	DNA
E. arnhemicus	9 km South of Batavia Downs, Cape York Peninsula, QLD	Baba 611 (CNS)	FL, FR	-12 34 34	142 40 39	DNA
E. arnhemicus	9 km South of Batavia Downs, Cape York Peninsula, QLD	Baba 612 (CNS)	FL	-12 34 34	142 40 39	DNA
E. arnhemicus	Rocky Creek, Cape York Peninsula, QLD	Baba 619 (CNS)	FL	-12 54 11	142 44 42	DNA
E. arnhemicus	Coen River, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 624 (CNS)	FL	-13 37 29	142 36 13	DNA
E. arnhemicus	Horsetailers Waterhole, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 630 (CNS)	FL, FR	-13 24 36	142 19 4	DNA
E. arnhemicus	Edge of small billabongs, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 633 (CNS)	FL	-13 26 17	142 18 48	DNA
E. arnhemicus	Governors Waterhole, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 634 (CNS)	FR	-13 26 19	142 18 46	DNA
E. arnhemicus	Governors Waterhole, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 635 (CNS)	FR	-13 26 19	142 18 46	DNA
E. arnhemicus	Coen River, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 642 (CNS)	FR	-13 30 52	142 26 18	DNA
E. arnhemicus	Archer River, Mungun Kaanju National Park, Cape York Peninsula, QLD	Baba 651 (CNS)	FL	-13 27 2	142 42 14	DNA
E. arnhemicus	Stewart River, Lama Lama National Park, Cape York	Baba 655 (CNS)	FL, FR	-14 6 15	143 25 18	DNA

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
	Peninsula, QLD					
E. arnhemicus	Stewart River, Lama Lama National Park, Cape York Peninsula, QLD	Baba 657 (CNS)	FL	-14 5 38	143 40 15	DNA
E. arnhemicus	Stewart River, Lama Lama National Park, Cape York Peninsula, QLD	Baba 660 (CNS)	FL, FR	-14 54 42	143 40 10	DNA
E. arnhemicus	Morehead River, Cape York Peninsula, QLD	Baba 662 (CNS)	FL	-15 1 18	143 39 59	DNA
E. arnhemicus	Morehead River, Cape York Peninsula, QLD	Baba 664 (CNS)	FL FR	-15 1 9	143 40 5	DNA
E. arnhemicus	Cultivation Sankowsky Arboretum, ex Big Mitchel river	Baba 847 (CNS)	FL	N/A	N/A	
E. arnhemicus	Arnhem Land, Darwin & Gulf, NT	Cowie 8876 (CNS)	FL	-12 28 22	134 27 16	
E. arnhemicus	Fitzmaurice River, Victoria Bonaparte, NT	F. Mueller s.n (MEL)	FR	-14 50	129 46	Lectotype of *E. arnhemicus*
E. arnhemicus	Ramingining, Darwin & Gulf, NT	Russel-Smith 3922 (CNS)	FR	-12 34 30	134 55 30	
E. arnhemicus	Mary River National Park, 1 km downstream from Arnhem Highway on Mary River, Darwin & Gulf, NT	Cooper 2133 (CNS)	FR	-12 53 29	131 38 21	
E. arnhemicus	Mabaduan, Western, PNG	Henty E.E, NGF 38621 (CNS)	FL, FR	-9 15 30	142 40 30	
E. arnhemicus	Near Weam, Western, PNG	Ridsdale & Galre NGF 33638 (CNS)	FL, FR	-8 38 29	141 7 30	
E. obovatus	Undara Crater, Undara Crater National Park, Einasleigh Uplands, QLD	Baba 685 (CNS)	FL	-18 18 36	144 44 21	DNA
E. obovatus	Undara Crater, Undara Crater National Park, Einasleigh Uplands, QLD	Baba 687 (CNS)	FL, FR	-18 18 39	144 44 22	DNA
E. obovatus	Hansen Cave, Undara Crater National Park, Einasleigh Uplands, QLD	Baba 692 (CNS)	FL	-18 18 20	144 44 8	DNA
E. obovatus	Palm Tree Creek, off Thornton's Gap Road, Harvey Range, West of Townsville, Brigalow Belt, QLD	Baba 696 (CNS)	FL	-19 20 13	146 27 26	DNA
E. obovatus	Mackay Botanic Garden- ex Mt Basset, Mackay, South East Queensland, QLD	Baba 698(CNS)	FL, FR	-21 7 4	149 12 34	DNA
E. obovatus	Mt. Basset Cemetery, Mackay, South East Queensland, QLD	Baba 703 (CNS)	FL	-21 7 5	149 12 38	DNA

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
E. obovatus	Kelsey Creek, Kelsey Creek Road, Proserpine, South East Queensland, QLD	Baba 706 (CNS)	FL	-20 26 46	148 24 42	DNA
E. obovatus	Kelsey Creek, Kelsey Creek Road, Proserpine, South East Queensland, QLD	Baba 708 (CNS)	FL	-20 27 1	148 24 6	DNA
E. obovatus	Kelsey Creek, Kelsey Creek Road, Proserpine, South East Queensland, QLD	Baba 711 (CNS)	FL	-20 27 1	148 24 6	DNA
E. obovatus	Yapoon hill, Port Curtis, QLD	Batianof 9261 (BRI)	FL	-23 00 00	150 24 20	
E. obovatus	Yapoon, Port Curtis, QLD	Sloan s.n.-AQ185115 (BRI)	FR	-23 04 54	150 45 30	
E. obovatus	Blackmans Gap, 21 km from the Bruce Highway near Miriam Vale, Port Curtis QLD	Forster 12281 (BRI)	FL	-24 25 59	151 25 00	
E. obovatus	Five Mile Creek, 'Booringa', SE of Kalpowar, Widebay, QLD	Bean 9343 (BRI)	FL	-24 49 24	151 27 33	
E. obovatus	3.5 km E of Childers, Conlons Road, Widebay, QLD	Forster 25188 (BRI)	FL	-25 13 59	152 19 00	
E. obovatus	Stony Creek, near Didcot, Widebay, QLD	Forster 260B (BRI)	FR	-25 28 45	151 54 08	
E. obovatus	Tinnanbar, Widebay, QLD	Brooks 90 (BRI)	FR	-25 46 24	151 27 33	
E. obovatus	Mt Coonowrin, Glass House Mountains, South East Queensland QLD	Hubbard 4118 (BRI)	FL			
E. obovatus	Belmont Hill Reserve, Brisbane, South East Queensland, QLD	Baba 726 (CNS)	FL	-27 30 46	153 7 12	DNA
E. obovatus	Enoggera Creek, Brisbane, South East Queensland, QLD	Baba 727 (CNS)	FL, FR	-27 26 45	152 57 49	DNA
E. obovatus	Mary Road Reserve, Brisbane, South East Queensland, QLD	Baba 728 (CNS)	FL, FR	-27 39 46	153 9 40	DNA
E. obovatus	Oxley Creek Brisbane, South East Queensland, QLD	Baba 758A (CNS)	FR	-27 33 2	152 59 34	DNA
E. obovatus	Mudgeeraba CK, South East Queensland, QLD	Baba 730 (CNS)	FL	-28 6 46	153 19 4	DNA
E. obovatus	Road to Mt Bunya, South East Queensland, QLD	Baba 745 (CNS)	FL, FR	-26 53 39	151 37 0	DNA
E. obovatus	Road to Mt Bunya, South East Queensland, QLD	Baba 747 (CNS)	FL	-26 53 39	151 37 0	DNA
E. obovatus	Haly Creek Road, near Mt Bunya, South East Queensland, QLD	Baba 748 (CNS)	FL	-26 40 10	151 48 23	DNA

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
E. obovatus	Haly Creek Road, near Mt Bunya, South East Queensland, QLD	Baba 750 (CNS)	FL, FR	-26 40 10	151 48 31	DNA
E. obovatus	Obi Obi Creek, near Maleny township, South East Queensland, QLD	Baba 752 (CNS)	FL	-26 45 43	152 50 46	DNA
E. obovatus	Mooloolaba Beach, South East Queensland, QLD	Baba 754 (CNS)	FL, FR	-26 44 23	153 7 59	DNA
E. obovatus	Mooloolaba Beach, South East Queensland, QLD	Baba 756 (CNS)	FR	-26 44 3	153 7 59	DNA
E. obovatus	Mooloolaba Beach, South East Queensland, QLD	Baba 757 (CNS)	FR	-26 45 19	153 8 2	DNA
E. obovatus	Fernleigh Rd, Fernleigh, North Coast, NSW	Wilson 7667 (BRI)	FR	-28 46 59	153 28 59	
E. obovatus	Tweed River, Opposite Scott's Island, North Coast, NSW	McDonald 1602 (BRI)	FL	-28 15	158 25	
E. obovatus	Hogan Park, Tweed River, Opposite Scott's Island, North Coast, NSW	Williams 75005 (BRI)	FR	- 28 9 36	153 18 04	
E. obovatus	c. 13 km S of Urunga, on the Pacific Highway, North Coast, NSW	Coveny 2152 (BRI)	FL	- 30 43 11	153 01 12	
E. sp. Mt. Bellenden Ker	Cloudland Nature Reserve, Wet Tropics, QLD	Baba 768 (CNS)	FL, FR	-17 26 13	145 31 52	DNA
E. sp. Mt. Bellenden Ker	Topaz, Atherton Tableland, Wet Tropics, QLD	Baba 760 (CNS)	FL, FR	-17 26 32	145 42 47	DNA was vouchered from the same tree in a different occasion.
E. sp. Mt. Bellenden Ker	Topaz, Atherton Tableland, Wet Tropics, QLD	Baba 761 (CNS)	FL	-17 26 32	145 42 47	DNA was vouchered from the same tree in a different occasion.
E. sp. Mt. Bellenden Ker	Paluma Range, Wet Tropics, QLD	Jackes 20102 (CNS)	FL part	-19 0 30	146 12 26	DNA
E. sp. Mt. Bellenden Ker	Koombooloomba Forest Reserve, Tully, Wet Tropics, QLD	Ford 4478 (CNS)	FL	-17 51 55	145 36 40	
E. sp. Mt. Bellenden Ker	East Downey Logging Area, Wet Tropics, QLD	Hyland 5618 (CNS)	FL	-17 40 00	145 50 0	

Putative Taxon	Locality	Vouchers	Material scored*1	Lat.	Long.	Notes *2
E. sp. Mt. Bellenden Ker	Mount Lewis Forest Reserve, Wet Tropics, QLD	Forster 18119 (CNS)	FL	-16 31 00	145 16 00	DNA
E. sp. Mt. Bellenden Ker	Ongera Logging Area, Wet Tropics, QLD	Gray 3691 (CNS)	FL	-17 44 00	145 33 00	
E. sp. Mt. Bellenden Ker	Windsor Tableland National Park, Wet Tropics, QLD	Ford 4466 (CNS)	FL part	-16 13 00	145 05 00	DNA
E. sp. Mt. Bellenden Ker	Tully Falls Road, Atherton Tableland, Wet Tropics, QLD	Costion 2092 (CNS)	FR	-17 44 56	145 29 6	
E. sp. Mt. Bellenden Ker	Mt Bellenden Ker, Wooroonooran National Park, Wet Tropics, QLD	Brass 18336 (CNS)	FR	-17 15 00	145 55 00	
E. sp. Mt. Bellenden Ker	Topaz, Atherton Tableland, Wet Tropics, QLD	Cooper 1916 (CNS)	FR	-17 26 49	145 42 44	
E. sp. Mt. Bellenden Ker	Mt Spec, NE of Townsville, Brigalow Belt, QLD	Webb 8192 (CNS)	FR	-18 55 00	146 15 00	
E. coorangooloo	Atherton, Atherton Tableland, Wet Tropics, QLD	Baba 695 (CNS)	FL, FR	-17 15 10	145 28 35	DNA
E. coorangooloo	Hallorans Hill, Atherton Tableland, Wet Tropics, QLD	Irvine 1941 (BRI)	FL	- 17 16 0	145 28 59	
E. coorangooloo	Wongabel, Atherton Tableland, Wet Tropics, QLD	Smith 3791	FL, FR	- 17 19 24	145 30 33	
E. coorangooloo	Carson Road, Malanda, Atherton Tableland, Wet Tropics, QLD	Forster 29540 (BRI)	FR	- 17 21 33	145 33 6	
E. coorangooloo	Mazlin Creek, Atherton Tableland, Wet Tropics, QLD	Volck 770 (BRI)	FL	- 17 15 10	145 28 36	
E. coorangooloo	Phillips Avenue, Atherton, Atherton Tableland, Wet Tropics, QLD	Hyland 12637 (CNS)	FR	- 17 16 0	145 28 0	
ADC	Davis Creek National Park, Wet Tropics, QLD	Baba 823 (CNS)	FL, FR	-16 58 46	145 33 14	DNA
ADC	Davis Creek National Park, Wet Tropics, QLD	Baba 821 (CNS)	FR	-16 58 46	145 33 14	DNA
Potentially ADC	Foot of Walsh's Pyramid, Wet Tropics, QLD	Lyons 158 (BRI)	FR	-17 7 24	145 47 34	

4.2.2 Data analyses

Because specimens generally do not bear both flowers and fruit concurrently, the data set was divided into two parts: floral plus vegetative characters, and fruit plus vegetative characters. All variables were normalised to Z-score before PCA and MDS analysis because the units of measurement varied among some characters (Davis, 2002). Z-scores were calculated by $z = (x - \mu) / \sigma$ when x is a raw value, μ is the mean of a population, and σ is the standard deviation of the population. A strong correlation between the length of inflorescence axis and flower numbers (0.897), and stone length and seed length (0.910) were found, therefore only the length of inflorescence axis and stone length was included.

PCA (Pearson, 1901) was undertaken to examine the relationships among the groups and correlation between the variables. Prior to analysis the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were calculated. For the KMO measure of sampling adequacy, a cut off 60% was used (Kaiser, 1974). Bartlett's test of sphericity was employed to test the null hypothesis that the variables are uncorrelated in the population and statistical significance ($p < 0.05$) was ensured for all the analysis. In order to obtain and more clearly visualise patterns for each component, factor loadings were orthogonally rotated using Varimax with Kaiser Normalization.

In order to assess the groupings with different methods, multidimensional scaling and cluster analysis were employed. Only variables identified as influential in the PCA were included in the analysis. Semi-strong hybrid multidimensional scaling (SSH MDS), an extension of multidimensional scaling (MDS), and flexible unweighted pair group method of arithmetic averages (UPGMA) (Sneath & Sokal, 1973) were performed using the software package PATN (Belbin 1993). The dissimilarity matrix was calculated using the Gower metric association measure (Gower, 1971), as it best handles mixed datasets and missing values (Crisp & Weston, 1993). The same association matrix was used in the both analyses. The ordination plot was produced with a cut-off value of 0.9, 100 random starts and a maximum of 50 iterations and the lowest stress values were used. PCC, values of Kruskal-Wallis and Monte-Carlo characters in ordination (MCAO) (Manly, 1991) were calculated to elucidate the variables that most strongly discriminate the groups. The larger Kruskal-Wallis values, the higher the significance of contribution the variables have to the separation of the object groups (Belbin, 1993)

Once the clusters were assigned to the hypothetical reference entities, refined analyses were run. Multigroup Discriminant Function Analysis (MDFA) with stepwise methods was performed to test the significance of the refined groups. MDFA tests the significance of a set of discriminant

functions and then classifies the cases to the reference groups. The significance test used in MDFA is identical to that of MANOVA (Hill & Lewicki, 2006). Like MANOVA, MDFA also requires certain assumptions to be met: 1) that the data have equal sample sizes; 2) that there is homogeneity of covariance; and 3) that the variables are normally distributed. Box's test for equality of covariance was calculated to test the homogeneity of covariance. Even though Box's test of equality resulted in $p>0.001$, the sample size of the reference groups were different from each other, therefore the results were interpreted with caution. Fruit and vegetative characters were lognormally transformed prior to analysis to meet the assumption of equality of covariance. All the variables included in MDFA were checked as to whether the data were normally distributed or not prior to analysis. 'Leave-one-out' cross validation was employed to test the accuracy of the classification model. This cross validation reclassifies each sample depending on the functionality of all the other cases, excluding one case (Stone, 1974; Arlot & Celisse, 2010).

Lastly, to identify the influential characters that discriminate the reference groups, classification tree analysis using the classification and regression tree algorithm (CART) was employed. CART is a non-parametric approach which makes a decision to split the data set dichotomously by fitting 'if-not' models (a regression model) to a single variable at each step in a recursive process (Saraswati & Sabnis, 2006). Through CART analysis, a tree-like diagram is produced to help visualise the binary splits amongst the groups. The resulting groups do not necessarily correspond to the putative entities at the earlier layers as the model simply splits the data set into two groups to make each group as homogenous as possible. Child nodes are then treated as the parent nodes in the subsequent layers, as the process is recursive. CART analysis can identify specific characters (and their states or ranges) by which groups can be defined, information which is useful for constructing identification keys.

Analyses were performed using SPSS Statistics (IBM, v.19) for descriptive statistics, Principal Component Analysis, Multivariate discriminant function analysis and CART. Semi-strong hybrid multidimensional scaling and flexible UPGMA were performed in the pattern analysis package PATN (Belbin, 1993).

4.3 Results

4.3.1 Principal Component Analysis

4.3.1.1 Vegetative characters

The first factor loading, which explains 41.05 % of the variation, was heavily influenced (high loading) by the ratio of leaf width to length (character 4), leaf width (character 3), and secondary vein angle (character 7) (Table 4.4). The second loading explains 19.94% of the variation and was most strongly influenced by leaf length (character 2) and leaf shape (character 1). Factor three explains 12.96 % of the variation and was most strongly influenced by number of secondary veins (character 6). With all the components together, each entity formed a broad cluster but was not clearly separated from all other clusters (Figure 4.2). There was no clear separation of *E. obovatus* North and *E. obovatus* South.

The KMO measure of sampling adequacy was 0.661 and a Bartlett test of sphericity was significant ($C2 = 786.692$, $df = 45$, $p < 0.001$).

Table 4.3. Characters and their definitions. Basic statistical parameters of the morphological characters of *E. obovatus, E. arnhemicus*, and *E. sp*. Mt. Bellenden Ker are also presented: range and mean value (top row) and standard deviation (bottom row). All measurements are in mm. A specimen, Lyons 158 was included in *E. arnhemicus*.

Character		Definition	*E. obovatus*	*E. arnhemicus*	*E.* sp Mt. Bellenden Ker	*E. coorangooloo*	ADC population
1	Leaf: shape	Distance along the midrib from the widest point of the lamina to the junction of the lamina and petiole.	**28.6 – 64.8 (39.3)**	**31.6 – 56.0 (43.0)**	**25.5 – 38.3 (30.7)**	**48.8 – 66.5 (55.9)**	**41.8 – 45.6 (43.7)**
			± 7.8	**±6.3**	**±3.6**	**±6.6**	**±2.7**
2	Leaf: length	Distance along the midrib from the tip to the junction of the lamina and petiole.	**52.3 – 97.7 (70.3)**	**56.8 – 110.0 (82.7)**	**55.4 – 81.0 (63.8)**	**88.5 – 114.7 (100.1)**	**95.2 – 96.4 (95.8)**
			± 11.1	**± 10.6**	**± 6.6**	**±11.5**	**±0.9**
3	Leaf: width	Maximum distance between the lateral edges of the lamina.	**10.6 – 29.0 (20.6)**	**24.2 – 49.8 (35.7)**	**22.3 – 30.8 (24.5)**	**42.0 – 50.5 (46.7)**	**33.0 – 36.4 (34.7)**
			± 3.3	**± 6.4**	**± 2.2**	**±3.6**	**±2.4**
4	Leaf: ratio	Ratio of leaf length to width.	**2.5 – 5.3 (3.5)**	**1.7 – 3.2 (2.4)**	**2.4 – 2.8 (2.6)**	**1.8 – 2.7 (2.2)**	**2.7 –2.9 (2.8)**
			±0.7	**±0.4**	**±0.2**	**±0.3**	**±0.2**
5	Leaf: petiole length	Distance between the point of attachment of the leaf to the stem and to the leaf blade.	**2.6– 8.0 (5.1)**	**5.0– 18.0 (10.2)**	**6.0 – 13.4 (10.0)**	**24.7– 32.6 (27.4)**	**21.2– 22.6 (21.9)**
			± 1.4	**±3.0**	**±2.4**	**±3.2**	**±1.0**
6	Leaf: secondary veins	Number of secondary veins. Veins are regarded as secondary when they arise from the midrib and end in a loop inside the margin.	**6 – 11 (9.0)**	**6 – 11 (8.3)**	**9 – 13 (11.0)**	**7 – 10 (8.4)**	**12 – 13 (12.5)**
			± 1.1	**± 1.1**	**±1.1**	**±0.7**	**±0.7**
7	Leaf: secondary vein angle	Angle between the midrib and the line from the base of the fourth vein to the intersection of the fourth vein and a line from the base of the fifth vein drawn perpendicular to the midrib (Figure 4.1).	**20 – 43 (32.0)**	**32 – 52 (40.3)**	**42 – 54 (46.3)**	**30 – 43 (37.0)**	**39 – 42 (40.6)**
			± 5.2	**± 5.4**	**±3.2**	**±4.9**	**±2.0**

Character		Definition	*E. obovatus*	*E. arnhemicus*	*E.* sp Mt. Bellenden Ker	*E. coorangooloo*	ADC population
8	Leaf: domatia	Number of domatia occurring in the axils of secondary veins.	**0 – 4 (0.71)**	**0 – 14 (4.3)**	**0 – 6 (1.6)**	**2 – 6 (3.6)**	**0**
			± 1.0	**± 2.9**	**±1.9**	**±1.6**	**-**
9	Leaf: teeth	Number of teeth on the left (when viewed from above) margin.	**3 – 11 (7.0)**	**6 – 15 (10.5)**	**6 – 11 (7.6)**	**12 – 15 (13.6)**	**9 – 12 (10.9)**
			± 1.9	**± 2.3**	**±1.3**	**±0.7**	**±2.1**
10	Leaf: acumen	Length of the tip of the leaf measured from the point of inflexion of the margin curvature (convex to concave) (Figure 4.1)	**2.8 – 10.8 (6.6)**	**3.0 – 25.4 (8.8)**	**5.5 – 9.8 (6.9)**	**1.3 – 14.0 (5.8)**	**7.8 – 10.4 (9.1)**
			± 2.1	**± 4.3**	**±1.2**	**±4.8**	**±1.8**
11	Inflorescence: axis length	Length of inflorescence axis measured from the point of attachment to the branch, to the tip.	**15.3 – 60.8 (37.8)**	**11.5 – 45.8 (32.5)**	**26.3 – 48.6 (39.3)**	**70.3 – 115.0 (93.0)**	**60.0 – 73.4 (66.7)**
			±12.2	**± 9.7**	**± 7.3**	**±16.9**	**±15.0**
12	Bracteole: length	Distance between the point of attachment of the bracteole and its tip.	**0.4 – 0.8 (0.6)**	**0.5 – 1.5 (1.1)**	**0.3 – 0.7 (0.5)**	**1.0 – 1.9 (1.4)**	**1.7 (-)**
			±0.1	**± 0.3**	**± 0.1**	**±0.4**	**Not applicable**
13	Flower: number	Number of flowers per inflorescence (including pedicel scars which indicate fallen flowers).	**7 – 38 (20.0)**	**8 – 21 (15.5)**	**17 – 46 (26.1)**	**34 – 48 (42.1)**	**21 – 51 (37.0)**
			±6.5	**±2.9**	**± 8.4**	**±5.8**	**±14.7**
14	Pedicel: length	Length of flower stalk measured from the point of attachment to the inflorescence axis, to the point of attachment of the first sepal.	**2.0 – 7.5 (3.5)**	**1.3 – 6.6 (2.7)**	**3.4 – 5.6 (4.1)**	**5.0 – 9.2 (6.9)**	**5.0 – 6.6 (5.6)**
			±1.0	**± 1.0**	**± 0.7**	**±1.4**	**±0.8**
15	Sepal: length	Distance between the point of attachment of the sepal and its tip.	**2.6 – 4.5 (3.4)**	**2.3– 3.6 (2.9)**	**3.0 – 3.5 (3.3)**	**3.6 – 5.0 (4.4)**	**3.3 (-)**
			±0.5	**± 0.3**	**± 0.2**	**±0.5**	**Not applicable**
16	Sepal: width	Maximum distance between the edges of the sepals.	**0.8 – 1.6 (1.3)**	**1.1 – 1.8 (1.4)**	**1.0 – 1.4 (1.1)**	**1.5 – 1.7 (1.6)**	**1.1 (-)**
			± 0.2	**± 0.2**	**± 0.1**	**±0.9**	**Not applicable**

Character		Definition	*E. obovatus*	*E. arnhemicus*	*E.* sp Mt. Bellenden Ker	*E. coorangooloo*	ADC population
17	Sepal: shape	Distance between the widest point and the point of attachment of the sepals.	**0.5 – 1.5 (1.1)**	**0.5 – 1.9 (0.9)**	**0.6 – 1.3 (0.9)**	**0.9 – 1.8 (1.4)**	**0.9 (-)**
			±0.3	**± 0.3**	**± 0.2**	**±0.3**	**Not applicable**
18	Petal: length	Distance between the attachment of the petal and its tip.	**2.5 – 4.4 (3.5)**	**2.3 – 4.1 (3.1)**	**3.2 – 3.9 (3.6)**	**4.2 – 5.5 (4.7)**	**3.7 (-)**
			± 0.6	**±0.4**	**± 0.2**	**±0.6**	**Not applicable**
19	Petal: width	Maximum distance between the edges of the petal.	**0.8 – 2.0 (1.5)**	**1.1– 2.2 (1.8)**	**1.1 – 1.5 (1.4)**	**1.5 – 2.2 (1.8)**	**1.8 (-)**
			± 0.3	**± 0.3**	**± 0.1**	**±0.3**	**Not applicable**
20	Petal: shape	Distance between the widest point and the point of attachment of the petals**.**	**1.5 – 2.8(1.98)**	**1.3 – 2.8 (1.84)**	**1.8 – 2.1 (1.94)**	**2.1 – 3.0 (2.52)**	**2.0 (-)**
			± 0.4	**±0.3**	**±0.1**	**±0.5**	**Not applicable**
21	Petal: length of petal divisions	Length of the middle petal division.	**0.5 – 1.9 (1.04)**	**0.5 – 1.1 (0.83)**	**1.1 – 1.5 (1. 38)**	**1.5 – 2.2 (1.78)**	**0.6 (-)**
			± 0.4	**± 0.3**	**± 0.1**	**±0.3**	**Not applicable**
22	Petal: number of divisions	Number of divisions comprising the petal fringe.	**5-12 (8)**	**7-13 (9.3)**	**6-7 (7.1)**	**6 – 10 (7.8)**	**8 (-)**
			±1.7	**± 1.2**	**±0.2**	**±1.5**	**Not applicable**
23	Style: length	Distance from the top of the ovary to the style tip.	**0.7 – 2.3 (1.7)**	**0.8 – 2.4 (1.6)**	**1.2 – 2.3 (1.8)**	**1.9– 2.6 (2.2)**	**2.2 (-)**
			± 0.4	**± 0.4**	**± 0.4**	**±0.3**	**Not applicable**
24	Disk: height	Distance from the base to the top of the disk.	**0.3 – 0.7(0.5)**	**0.4 – 0.7 (0.5)**	**0.4 – 0.6 (0.5)**	**0.5 – 0.8 (0.6)**	**0.6 (-)**
			±0.9	**± 0.1**	**± 0.1**	**±0.1**	**Not applicable**
25	Ovary: height	Distance from the base of the ovary to the attachment of the style.	**0.7 – 1.7 (1.1)**	**0.8– 1.4 (1.1)**	**0.9 – 1.2 (1.1)**	**1.4 – 2.0 (1.5)**	**1.4 (-)**
			±0.2	**± 0.2**	**± 0.1**	**±0.3**	**Not applicable**
26	Stamen: number	Number of stamens**.**	**13 – 21 (16.2)**	**11 – 26 (17.6)**	**10 – 12 (11.1)**	**22 – 27 (24.8)**	**14 (-)**
			±2.1	**± 3.3**	**± 0.7**	**± 1.8**	**Not applicable**
27	Stamen: filament	Distance between the point of	**0.2 – 0.7 (0.4)**	**0.3 – 0.5 (0.8)**	**0.4 – 0.6 (0.5)**	**0.5 – 0.7 (0.8)**	**0.6 (-)**

Character		Definition	*E. obovatus*	*E. arnhemicus*	*E.* sp Mt. Bellenden Ker	*E. coorangooloo*	ADC population
	length	attachment of filament to the receptacle and to the anther.	**±0.1**	**± 0.1**	**±0.1**	**± 0.1**	**Not applicable**
28	Stamen: anther length	Length of anther from the bottom to the top.	**0.9 – 2.0 (1.4)**	**0.7 – 1.8 (1.2)**	**1.1 – 1.2 (1.2)**	**1.3 – 2.6 (2.1)**	**1.0 (-)**
			±0.1	**± 0.2**	**± 0.1**	**±0.5**	**Not applicable**
29	Ovules per locule	Number of ovules per locule.	**4-8 (5.6)**	**4-5 (4.2)**	**6**	**6- 8 (6.5)**	**6 (-)**
			±0.9	**± 0.5**	**±0.0**	**±1.0**	**Not applicable**
30	Fruit: stone length	Length of longest axis of mericarp (pulp removed).	**5.4 – 8.3 (6.9)**	**8.5 – 15.9 (12.1)**	**7.1 – 10. 0 (8.5)**	**9.7 – 10.5 (10.1)**	**9.7 – 12.5 (11.1)**
			±1.0	**± 1.6**	**± 1.2**	**±0.4**	**±2.0**
31	Fruit: stone width	Length of shortest axis of mericarp (pulp removed).	**4.3 – 6.0 (5.0)**	**5.9 – 10.1 (7.7)**	**5.0– 6.8 (6.0)**	**6.6 – 7.6 (7.1)**	**6.3– 8.0 (7.2)**
			± 0.6	**±1.0**	**± 0.2**	**±0.5**	**±1.2**
32	Fruit: stone wall thickness	Distance between inner and outer wall of mericarp.	**1.0 – 1.4 (1.2)**	**1.7 – 3.4 (2.3)**	**1.3 – 1.8 (1.6)**	**1.8 – 2.1 (2.0)**	**2.0 (2.0)**
			± 0.1	**± 0.4**	**± 0.2**	**±0.2**	**±0.3**
33	Fruit: seed length	Length of longest axis of seed.	**3.6 – 5.6 (4.8)**	**5.6 – 10.1 (7.7)**	**5.0 – 6.5 (5.7)**	**5.9 – 7.4 (6.5)**	**5.9 – 8.0 (6.9)**
			±0.9	**± 1.0**	**± 0.7**	**±0.8**	**±1.4**

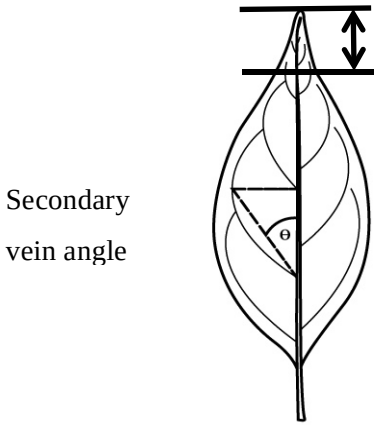


Figure 4.1. Illustrated guide to secondary vein angle, and the length of acumen. The image was used with the permission from the Center for Australian National Biodiversity Research in Australian Tropical Rainforest Plants Edition 6 (2010).

Table 4.4. Components extracted from PCA and their loadings on selected vegetative characters only. All variables were normalised prior to the analysis. Components were rotated by Varimax with Kaiser normalisation.

	Component		
	1	2	3
Leaf: length	0.24	0.94	-0.02
Leaf: width	0.81	0.49	-0.15
Leaf: ratio	-0.91	0.11	0.12
Leaf: shape	0.11	0.90	-0.21
Leaf: petiole length	0.55	0.47	0.00
Leaf: teeth	0.59	0.57	0.16
Leaf: secondary vein	-0.03	0.02	0.91
Leaf: secondary angle	0.73	-0.22	0.53
Leaf: acumen	-0.17	0.57	0.34

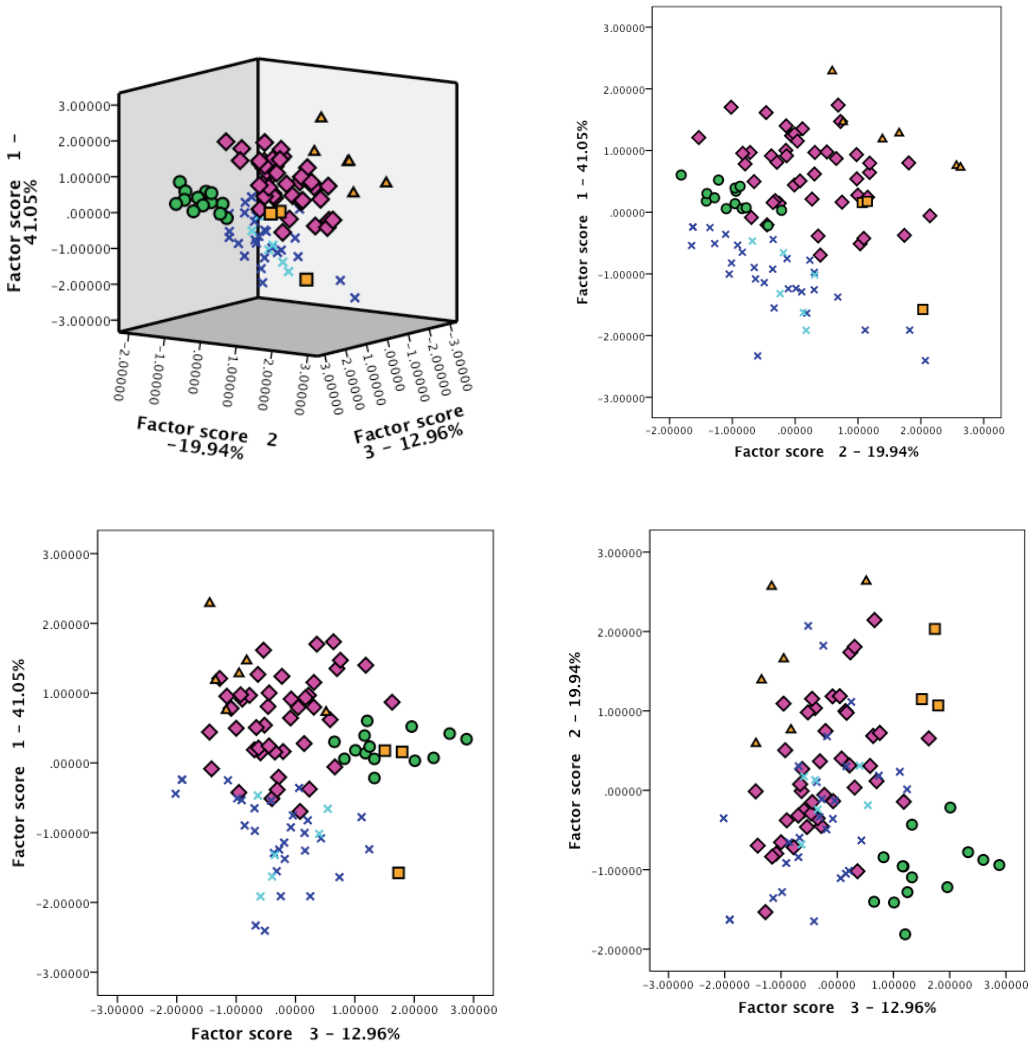


Figure 4.2. Ordination plot of Principal Component Analysis on vegetative characters. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange)

4.3.1.2 Flower characters

PCA revealed that the first component explains 41.44% of the variation when flower characters only are analysed. The first factor loading was heavily influenced by pedicel length (character 14), petal length (character 18), petal division length (character 21), petal shape (character 20), and sepal length (character 15) (Table 4.5.). The second loading explains 14.55% of the variation and was most strongly influenced by petal width (character 19). Factor three explains 8.69% of the variation and was most strongly influenced by inflorescence axis length (character 11) and bracteole length (character 12). The KMO measure of sampling adequacy was 0.766 and a Bartlett test of sphericity was significant ($C2 = 649.619$, $df = 153$, $p < 0.001$).

The scatter plot of all the components exhibited loose clustering of specimens into groups that broadly correspond to each putative entity (Figure 4.3). The ordination plot of the first and the third components shows strong isolation of samples of ADC and *E. coorangooloo* from the other three entities (Figure 4.3). This pattern was supported by the larger size of floral parts, which was greatly influenced by the heavy loading in component one. The other three entities formed one cluster. There was no evidence of separation of *E. obovatus* specimens into North and South groups (Figure 4.3).

Table 4.5. Components extracted from PCA and their loadings on selected floral characters only. All the variables were normalised prior to the analysis. Components were rotated by Varimax with Kaiser normalization.

	Component		
	1	2	3
Inflorescence axis length	0.68	-0.09	0.58
Pedicel length	0.84	-0.12	0.23
Ovules per locule	0.59	-0.25	-0.10
Bracteole length	-0.03	0.32	0.75
Stamen: number	0.02	0.18	0.74
Stamen: filament length	0.64	0.03	0.03
Stamen: anther length	0.49	0.24	-0.17
Ovary height	0.61	-0.002	0.04
Disc high	0.33	0.57	0.08
Style: length	0.77	0.24	0.03
Petal: length	0.80	0.09	-0.05
Petal: width	0.15	0.85	0.18
Petal: number of divisions	-0.14	0.75	-0.08
Petal: length of petal divisions	0.82	0.09	-0.05
Sepal: length	0.88	0.16	0.06
Sepal: width	0.14	0.75	0.35
Sepal: shape	0.27	-0.09	0.25

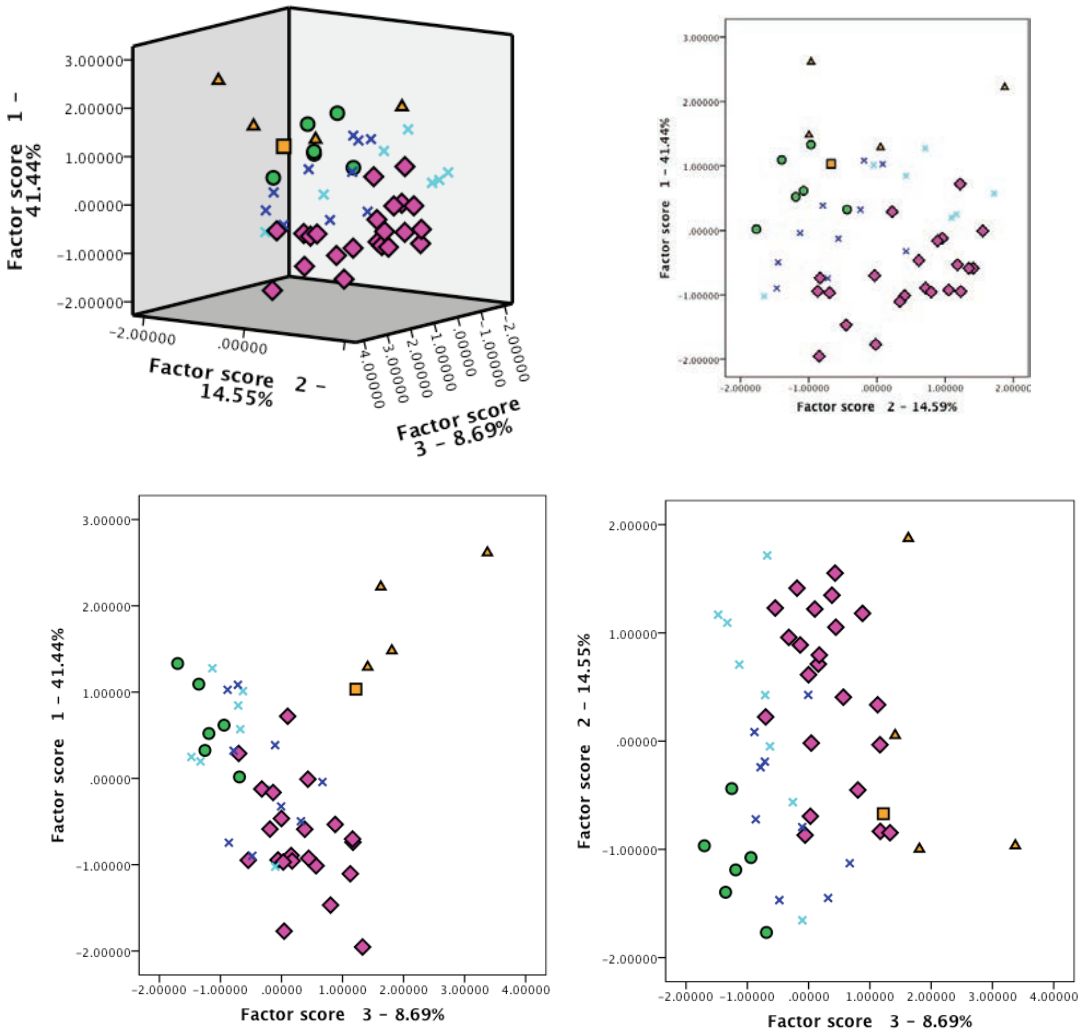


Figure 4.3. Ordination plot of Principal Component Analysis on floral characters only. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.1.3 Floral and vegetative characters

PCA analysis of the floral and vegetative characters combined revealed that 30.24% of the variation is explained by the first component. The first factor loading was heavily influenced by the pedicel length (character 14), petal length (character 18), sepal length (character 15), length of petal division (character 21), and petal shape (character 20) (Table 4.6). The second loading explains 18.16% of the variation and was most strongly influenced by leaf width (character 3) and petiole length (character 5). Factor three explains only 9.12% of the variation and was influenced by petal width (character 19). The KMO measure of sampling adequacy was 0.646 and a Bartlett's test of sphericity was significant ($C2 = 1207.256$, $df = 351$, $p < 0.001$).

With all the components together, only *E. coorangooloo* was clearly isolated from the other entities. The remainder formed clusters broadly corresponding to the entities but which were not separated from other entities (Figure 4.4).

While there is some evidence of subclustering within the *E. arnhemicus* cluster in the plot of first with the third component (Figure 4.4), both clusters include specimens from the same locations and there is no geographical explanation for these subclusters. *Elaeocarpus coorangooloo* separated from the others in the scatter plot with the first two components and ADC was positioned close to the *E. coorangooloo* cluster (Figure 4.4). Similarly to the analysis of floral characters alone (Figure 4.3), there was no evidence of separation of *E. obovatus* specimens into North and South groups (Figure 4.4).

Table 4.6. Components extracted from PCA and their loadings on selected floral and vegetative characters. All variables were normalised prior to analysis. Components were rotated by Varimax with Kaiser normalization.

	Component		
	1	2	3
Inflorescence axis length	0.65	0.52	-0.11
Bracteole length	-0.03	0.79	0.25
Pedicel length	0.81	0.30	-0.15
Ovules per locule	0.61	-0.08	-0.24
Stamen: number	0.03	0.60	0.14
Stamen: filament length	0.66	0.13	-0.01
Stamen: anther length	0.59	-0.09	0.18
Style length	0.79	-0.001	0.20
Ovary height	0.62	0.25	-0.04
Disk high	0.39	0.10	0.51
Petal: length	0.93	0.05	0.25
Petal: width	0.17	0.26	0.83
Petal: shape	0.85	0.02	0.03
Petal: number of divisions	-0.04	0.02	0.72
Petal: length of divisions	0.85	-0.02	0.03
Sepal: length	0.91	0.01	0.11
Sepal: width	0.19	0.42	0.72
Sepal: shape	0.42	0.16	-0.17
Leaf: ratio	0.17	-0.47	0.01
Leaf: shape	-0.11	0.49	0.32
Leaf: length	0.13	0.77	0.20
Leaf: width	-0.03	0.85	0.22
Leaf: petiole length	0.29	0.84	-0.11
Leaf: teeth	0.14	0.73	0.11
Leaf: secondary vein	0.43	-0.09	-0.47
Leaf: secondary vein angle	-0.19	0.11	-0.20
Leaf: acumen	-0.31	0.14	0.09

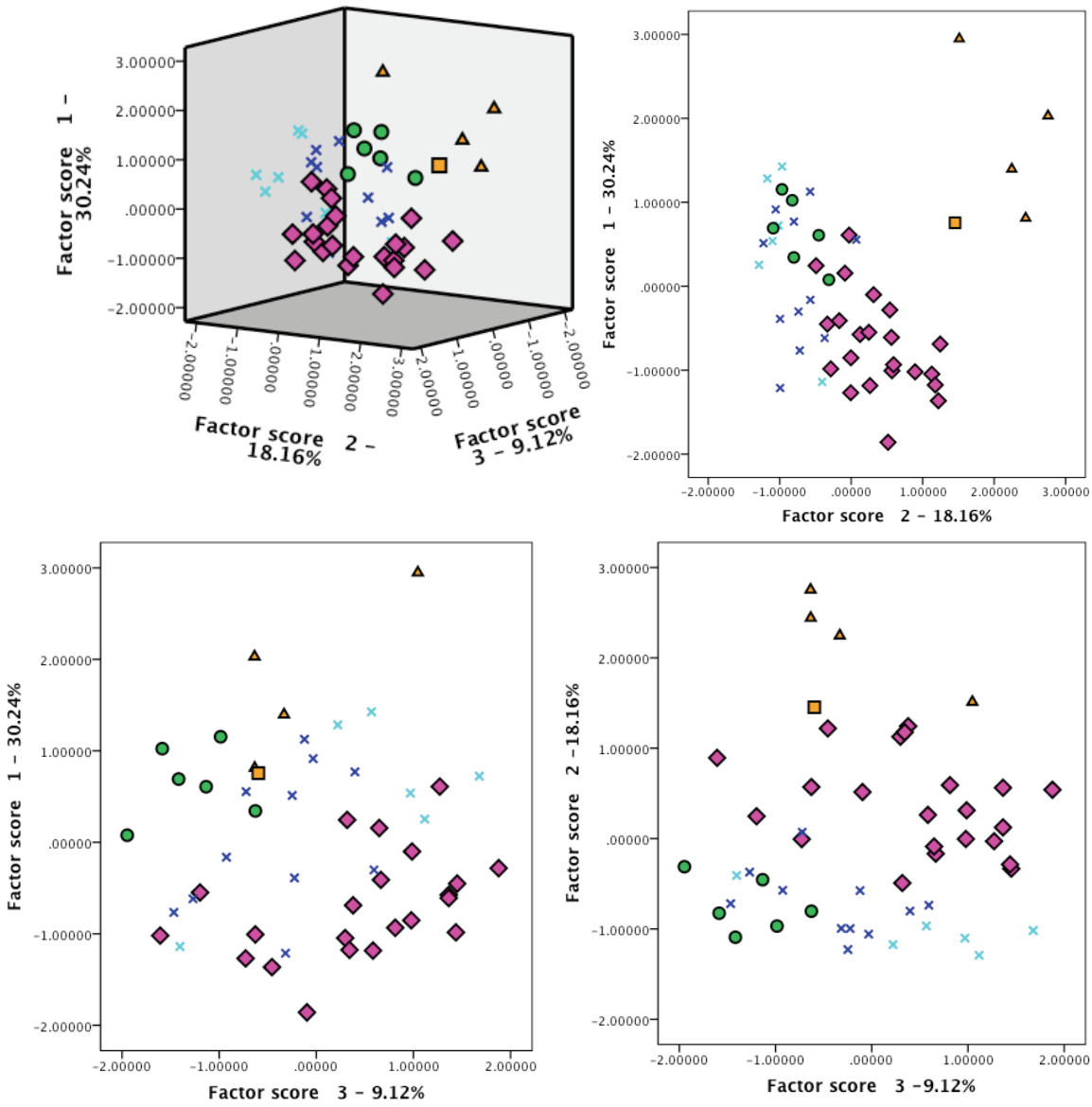


Figure 4.4. Ordination plot of Principal Component Analysis of floral and vegetative characters. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E.* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.1.4 Fruit and vegetative characters

PCA analysis revealed that 44.10% of the variation is explained by the first component. The first factor loading was heavily influenced by stone length (character 30), stone width (character 31), stonewall thickness (character 32), leaf width (character 3) and number of teeth (character 9) (Table 4.7). The second loading explains 19.04% of the variation and was most strongly influenced by the leaf shape (character 1) and leaf length (character 2). Factor three explains

only 13.06% of the variation, and was influenced by the number of secondary veins (character 6). The KMO measure of sampling adequacy was 0.727 and a Bartlett test of sphericity was significant ($C2 = 626.32$, $df = 66$, $p < 0.001$).

In the scatter plot with the all components included, specimens of *E. arnhemicus, E. obovatus* and *E.* sp. Mt. Bellenden Ker formed relatively distinct clusters, respectively. The specimens of *E. coorangooloo* and ADC were not clearly differentiated from the *E. arnhemicus* cluster (Figure 4.5). The segregation of *E. obovatus* and *E.* sp. Mt. Bellenden Ker clusters was strongly influenced by the characters that were heavily loaded in component 3 such as the number of secondary veins and vein angle (Table 4.7). Again, there was no evidence of separation of *E. obovatus* specimens into North and South groups (Figure 4.5).

Table 4.7. Components extracted from PCA and their loadings on selected fruit and vegetative characters. All the variables were normalised prior to the analysis. Components were rotated by the Varimax with Kaiser normalization.

	Component		
	1	2	3
Fruit: stone length	0.88	0.21	0.02
Fruit: stone width	0.90	0.08	0.05
Fruit: stone wall thickness	0.89	0.15	-0.03
Leaf: length	0.28	0.93	-0.01
Leaf: width	0.80	0.47	0.03
Leaf: ratio	-0.88	0.21	-0.09
Leaf: shape	0.13	0.88	-0.16
Leaf: petiole length	0.27	0.52	0.32
Leaf: teeth	0.65	0.48	0.36
Leaf: secondary vein	-0.24	0.10	0.90
Leaf: secondary vein angles	0.41	-0.22	0.75
Leaf: acumen	-0.10	0.61	0.08

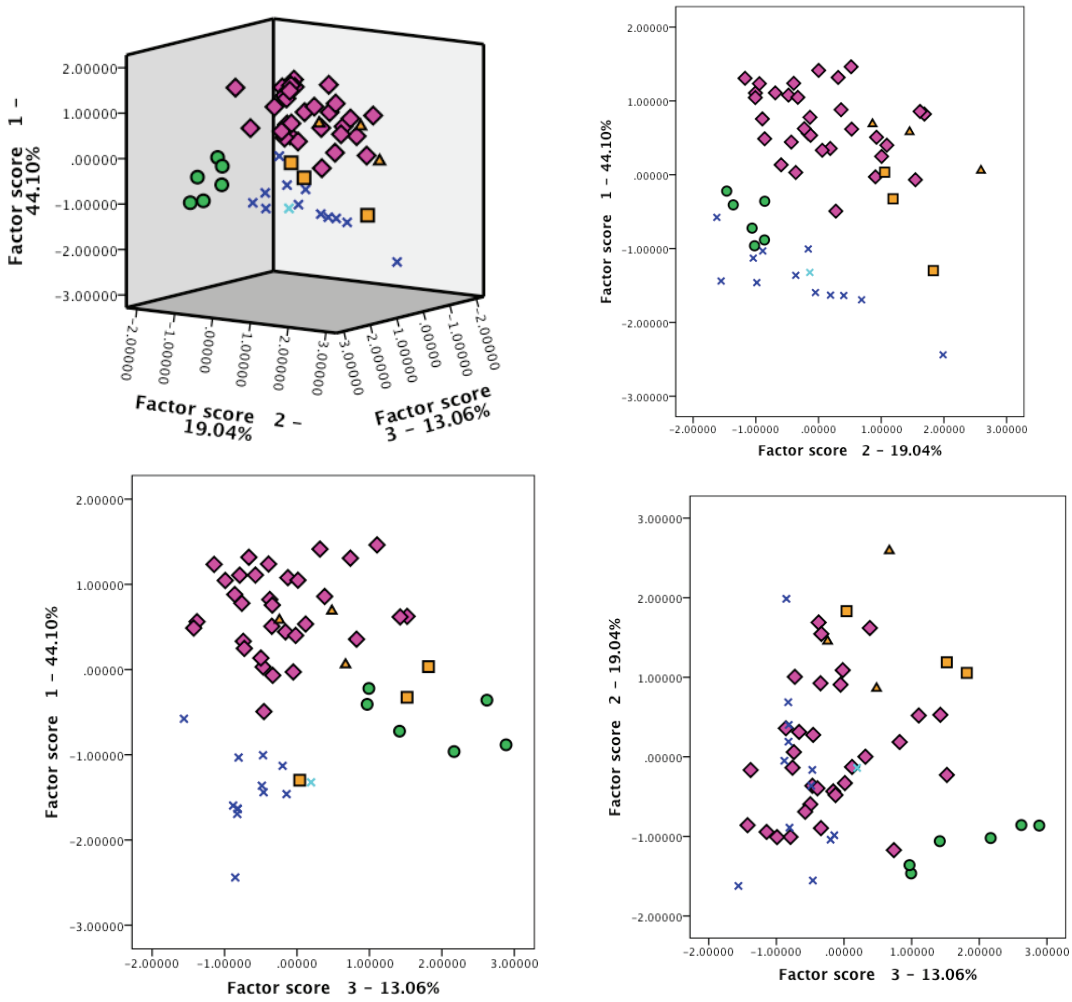


Figure 4.5. Ordination plot of Principal Component Analysis of fruit and vegetative characters. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E. sp.* Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.2 Hierarchical Clustering Analysis

To examine groupings to contrast to the PCA, an hierarchical clustering analysis was conducted.

4.3.2.1 Floral and vegetative characters

The hierarchical clustering analysis retrieved three distinct clusters comprising: (1) specimens of *E. arnhemicus* (Group 1); (2) specimens of *E. obovatus* and *E.* sp. Mt. Bellenden Ker (Group

2) and; (3) specimens of *E. coorangooloo* and ADC (Group 3) (Figure 4.6: indicated by the solid line). Subgroups were resolved within the last two clusters (Figure 4.6: indicated by the dotted line). There were three subgroups within the *E. obovatus—E.* sp. Mt. Bellenden Ker cluster which resulted in two subclusters (subgroup 1 and 2: one cluster includes a specimen of *E. arnhemicus*) and one corresponding to *E.* sp. Mt. Bellenden Ker specimens (subgroup 3), and two subgroups in the *E. coorangooloo* and ADC cluster (subgroup 4 and ADC). While two subclusters of *E. obovatus* reflect weak geographical patterns the clustering pattern was supported by petal size (character 18 &19). A subcluster of *E. obovatus* that is a sister cluster of *E.* sp. Mt. Bellenden Ker was an assembly of the specimens with the character of larger petals compared to the other subcluster of *E. obovatus*.

4.3.2.2 Fruit and vegetative characters

Analysis of the combined fruit and vegetative characters identified two major clusters: (1) specimens of *E. obovatus* and *E.* sp. Mt. Bellenden Ker (Group 1), and; (2) specimens of *E. arnhemicus*, *E. coorangooloo*, and ADC (Group 2) (Figure 4.7: indicated by the solid line). The morphological similarity between *E. arnhemicus, E. coorangooloo* and ADC was evident in the PCA ordinations and it appears strongly in the UPGMA dendrogram. Within a cluster of *E. obovatus-E.* sp. Mt. Bellenden Ker, subgroups were identified (Figure 4.7: indicated by the dotted line); two subgroups in one *E.* sp. Mt. Bellenden Ker clusters (subgroup 1) and one *E. obovatus* (subgroup 2). There were no assemblages of the groups that corresponding to *E. obovatus* North and South (Figure 4.7).

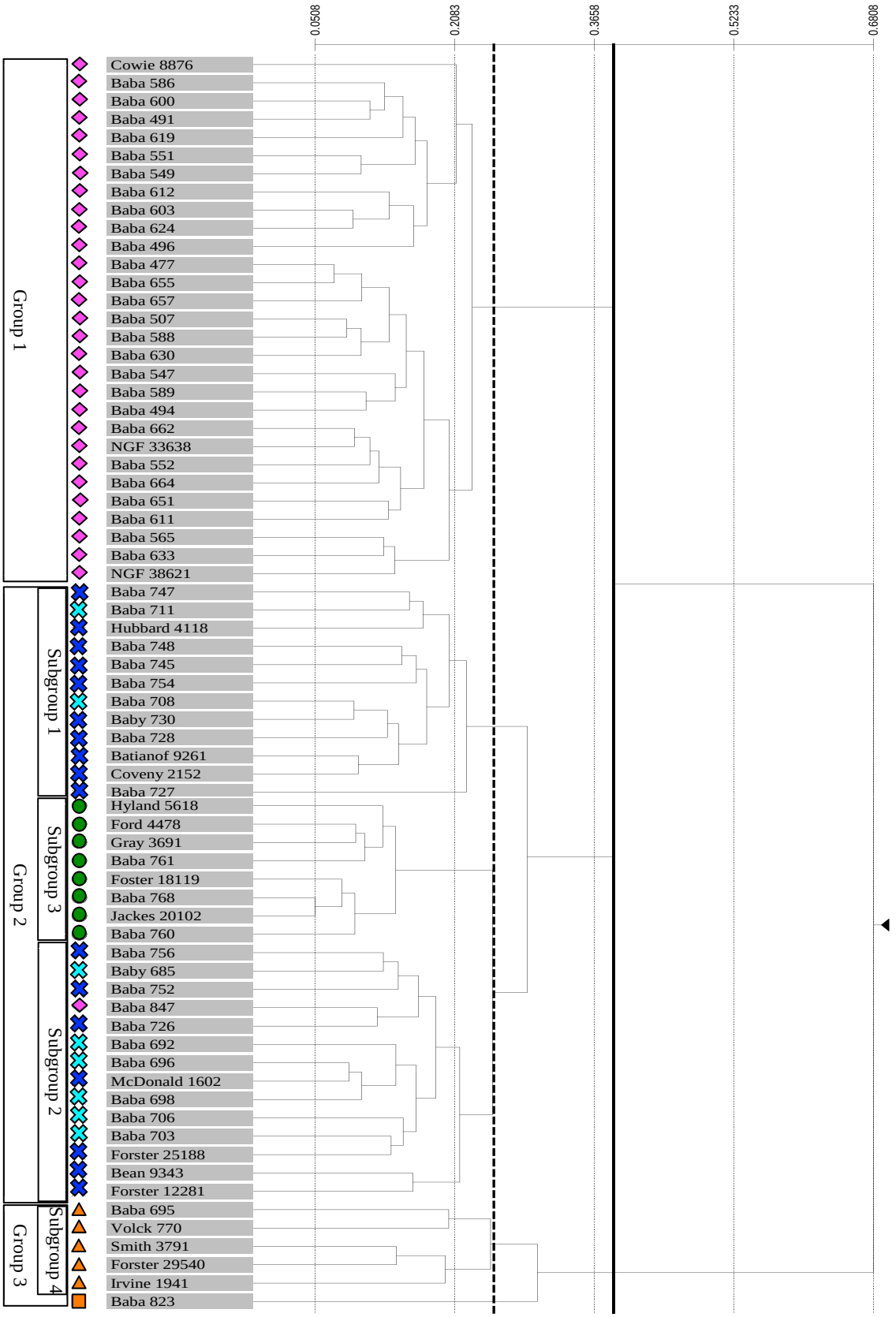


Figure 4.6. Dendrogram of specimens of the *E. obovatus* group based on an hierarchical clustering analysis of floral and vegetative characters using the flexible UPGMA algorithm. Solid and dotted lines indicate where the major- and sub- clusters are separated respectively. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

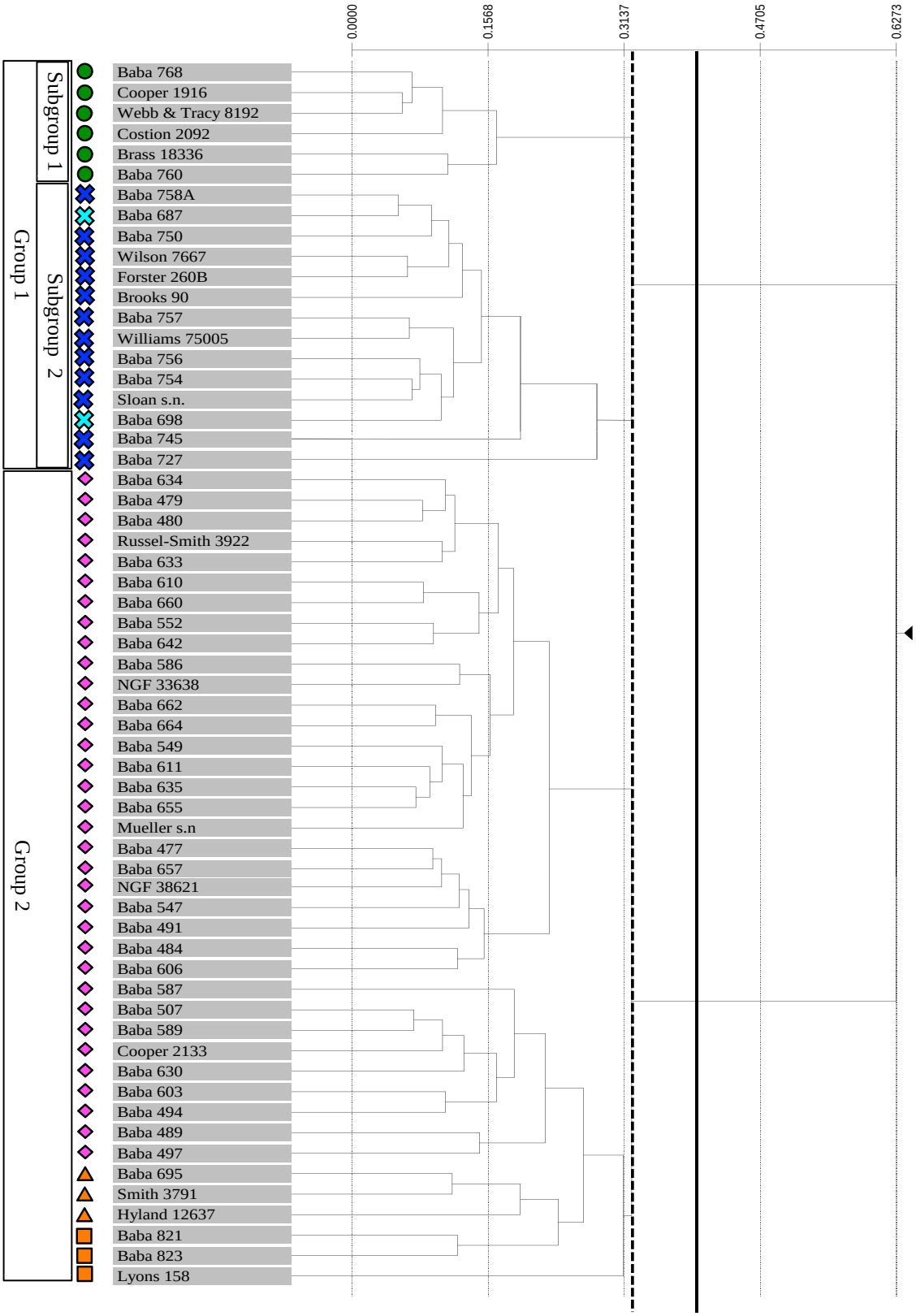


Figure 4.7. Dendrogram of specimens of *E. obovatus* based on a hierarchical clustering analysis of fruit and vegetative characters using the flexible UPGMA algorithm. Solid and dotted lines indicate where the major- and sub- clusters are separated respectively. ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.3 Semi-strong hybrid multidimensional Scaling (SSH MDS)

4.3.3.1 Flower and vegetative characters

Analysis of the flower and vegetative characters combined with SSH MDS using standardised characters showed that specimens of *E. coorangooloo* and ADC cluster separately from the rest of the specimens (Figure 4.8). Amongst the rest, a few subclusterings were observed: *E. arnhemicus* formed the clearest subcluster and *E.* sp Mt. Bellenden Ker formed a tight subcluster within the specimens of *E. obovatus* (Figure 4.8: a plot of axis 1 and 2). There is no obvious clustering of *E. obovatus* North and South.

The clusters resulting from the SSH MDS analysis were mostly concordant with the groups in the PCA ordination plots and UPGMA cluster analysis (Figure 4.8). Overall, clusters appeared to be close to each other.

Kruskal-Wallis values indicated that the top five characters that contributed to the segregation of clusters were leaf width (character 3), number of teeth on leaf margins (character 9), petal length (character 18), sepal length (character 15) and length of petal divisions (character 21).

Table 4.8. MDS statistics on selected floral and vegetative characters. The five characters with the highest Kruskal –Wallis values are indicated by stars.

	PCC (r^2)	Kruskal-Wallis	MCAO
Inflorescence axis length	0.534	19.504	0
Pedicel length	0.664	36.103	0
Bracteole length	0.711	36.442	0
Ovules per locule	0.490	28.837	0
Stamen: number	0.276	23.773	0
Stamen: filament length	0.545	22.464	0
Stamen: anther length	0.399	14.800	0
Style length	0.514	20.816	0
Ovary height	0.412	17.446	0
Disc high	0.342	7.8457	0
Petal: length ★	0.889	44.005	0
Petal: width	0.540	18.792	0
Petal: number of divisions	0.337	19.471	0
Petal: length of petal divisions★	0.681	42.788	0
Sepal: length ★	0.847	42.572	0
Sepal: width	0.450	20.838	0
Leaf: length	0.635	26.448	0
Leaf: width ★	0.865	52.134	0
Leaf: shape	0.478	19.299	0
Leaf: ratio	0.545	34.185	
Leaf: petiole length	0.714	37.877	0
Leaf: teeth ★	0.587	45.055	0

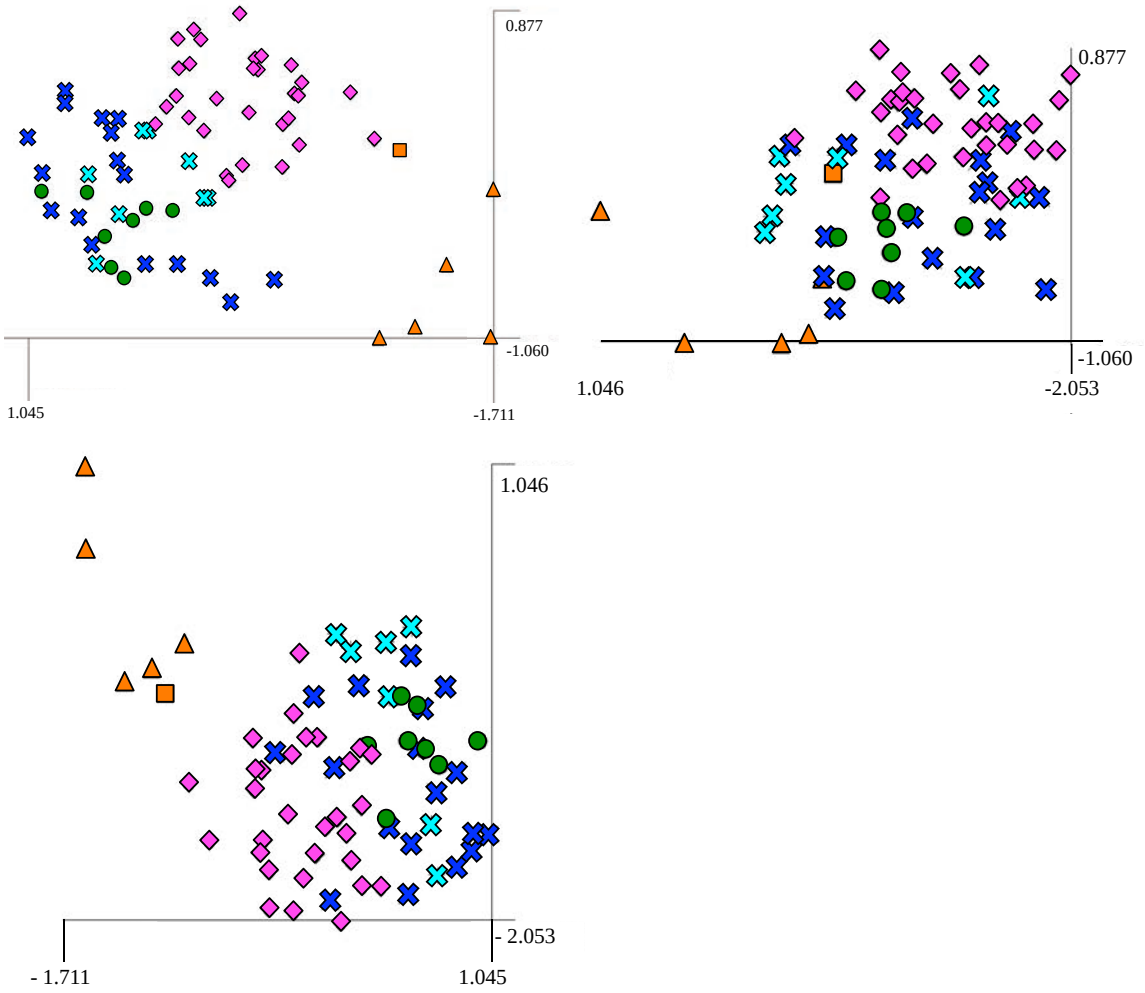


Figure 4.8. Ordination plots of the results of SSH MDS analysis of flower and vegetative characters (stress = 0.1535). Top left: axes 1 (X) and 2 (Y), top right: axes 1 (X) and 3 (Y) and bottom left: axes 2 (X) and 3 (Y). ◆ *Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E. obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.3.2 Fruit and vegetative characters

Analysis of the fruit and vegetative characters combined with SSH MDS using standardised characters revealed three separate clusters comprising specimens of: (1) *E. arnhemicus* and *ADC*; (2) *E. obovatus*, and; (3) *E.* sp Mt. Bellenden Ker (Figure 4.9). The specimens of *E. coorangooloo* appeared to be adjacent to or part of the *E. arnhemicus* cluster in all the plots. Differentiation of the groups between the *E. arnhemicus* and *E.* sp. Mt. Bellenden Ker or *E. obovatus* in the plot of axes 1 and 2 was low while they were segregated clearly in the plot of axes 2 and 3. In contrast, a group constituting all the specimens of *E. obovatus* (except one specimen of *E.* sp Mt. Bellenden Ker) was segregated from others clearly in the plots of axes 1 and 3, and axes 2 and 3. These clustering patterns were congruent to the groupings of PCA and

UPGMA analysis. Based on Kruskal-Wallis values, the top five characters that contributed to the segregation of clusters were leaf width (character 3), stone length, (character 30) stone width (character 31), stone wall thickness (character 32) and numbers of teeth on leaf margins (character 9).

Table 4.9. MDS statistics on selected fruit and vegetative characters. The five characters with the highest Kruskal–Wallis values are indicated by the stars.

	PCC (r^2)	**Kruskal-Wallis**	**MCAO**
Fruit: stone length ★	0.82	40.137	0
Fruit: stone width ★	0.83	37.372	0
Fruit: stone wall thickness ★	0.78	37.760	0
Leaf: length	0.79	23.588	0
Leaf: width ★	0.83	40.097	0
Leaf: shape	0.69	20.576	0
Leaf: petiole length	0.34	26.471	0
Leaf: teeth★	0.75	32.082	0
Leaf: secondary vein number	0.12	13.141	7
Leaf: secondary vein angle	0.35	26.167	0
Leaf: acumen	0.13	2.4132	4

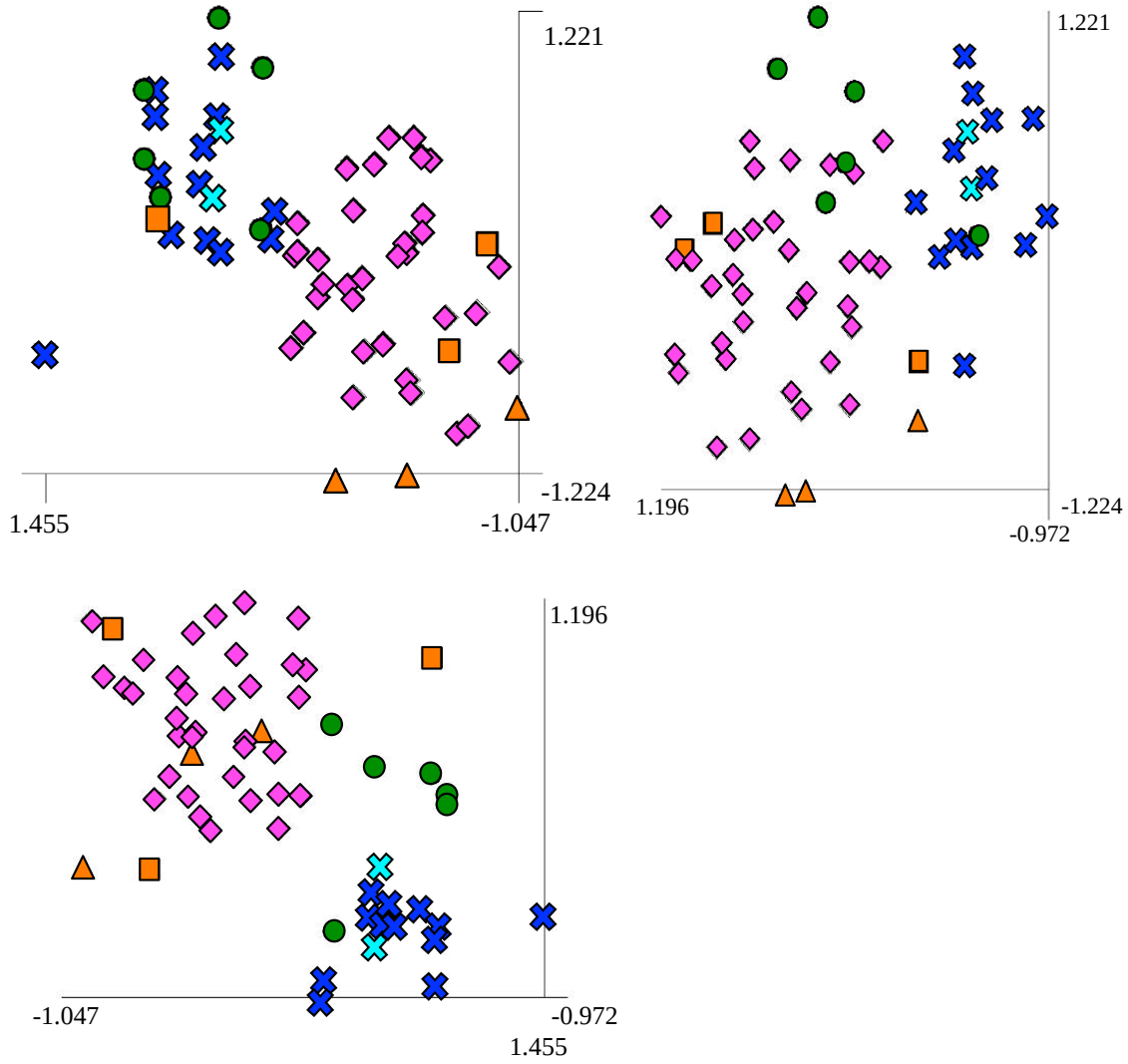


Figure 4.9. Ordination plots of the results of SSH MDS analysis of fruit and vegetative characters (stress = 0.0982). Top left: axes 1 (X) and 2 (Y), top right: axes 1 (X) and 3 (Y) and bottom left axes 2 (Y) and axes 3(X). ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (sky blue - specimens of *E.obovatus* North, dark blue - specimens of *E. obovatus* South), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange).

4.3.4 Multigroup Discriminant Function Analysis (MDFA)

In order to assess the reliability of clustering from the previous analyses, MDFA was employed. The floral data set included only one ADC specimen and the fruit data set included only two *E. coorangooloo* specimens. Because most of the other analyses in this study indicated that these two entities are morphologically very similar, they were pooled for MDFA analysis.

4.3.4.1 Floral and vegetative characters

When four reference entities (*E. arnhemicus*, *E. obovatus*, *E.* sp. Mt. Bellenden Ker, and *E. coorangooloo*+ADC), which were identified from the UPGMA cluster analyses, were tested, MDFA separated only one predefined group, *E. coorangooloo*+ADC, 100% correctly from the rest of the groups (94.4% accuracy within the *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker group) by an eigenvalue (or canonical root) of 9.234 (Figure 4.10). There was also a strong association between the discriminant function and the original variables (canonical correlation coefficient = 0.95 for function 1, and 2.738 and 0.86 respectively for function 2). The cross validation results showed that 7.8% of the remaining specimens were misclassified; all the groups contained at least one misclassified specimen (Table 4.10).

When the *E. obovatus* North and *E. obovatus* South specimens were included as separate reference entities classification accuracy decreased to 82.4% due to specimens being misclassified among these two groups.

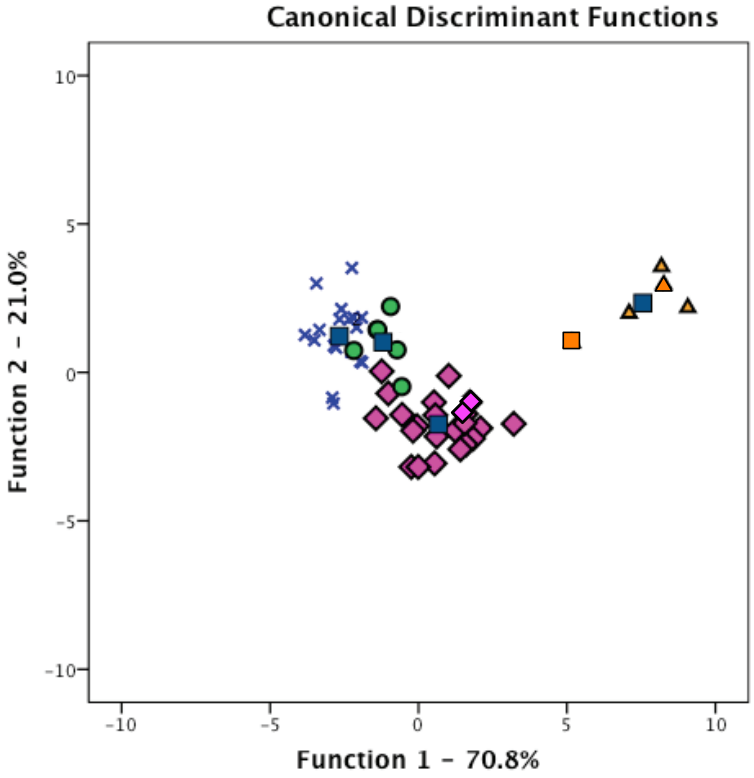


Figure 4.10. Ordination plot of Multigroup Discriminant Function Analysis of floral and vegetative characters. A total of 91.9% of variation is explained by function 1 (70.8%) and function 2 (21.0%). ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (blue), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange). Closed squares (dark blue) are group centroids.

Table 4.10. Predicted classification (top row within each entity) and cross validation (bottom row within each entity) results of Multigroup Discriminant Function Analysis of floral and vegetative characters.

	Predicted group membership				Total
	E. arnhemicus	*E. obovatus*	*E.* sp. Mt Bellenden Ker	*E. coorangooloo* and ADC	
E. arnhemicus	21	1	0	0	22
	20	2	0	0	22
E. obovatus	0	17	1	0	18
	0	17	1	1	18
E. sp. Mt Bellenden Ker	0	0	6	0	6
	1	0	5	0	6
E. coorangooloo and ADC	0	0	0	5	6
	0	0	0	5	6

4.3.4.2 Fruit and vegetative characters

The four reference entities were clearly delimited by MDFS analysis of selected fruit and vegetative characters (Figure 4.11). The separation was supported by an eigenvalue (or canonical root) of 6.245, and a strong association between the discriminant function and the original variables (canonical correlation coefficient = 0.93 for the first discriminant function and 2.90 and 0.86 respectively for the second function). All cases were assigned to the correct reference group (*E. obovatus*, *E. arnhemicus*, *E.* sp. Mt. Bellenden Ker, and *E. coorangooloo*–ADC), except one ADC specimen (Lyons 158) (Table 4.11). When this specimen was excluded from the *E. coorangooloo*–ADC reference entity and the analyses were rerun, the classification accuracy after cross validation was 100%. This result supports the existence of four entities, and inclusion of two of the ADC specimens (*Baba 821* and *823*) in *E. coorangooloo* with the third (*Lyons 158*) in *E. arnhemicus*.

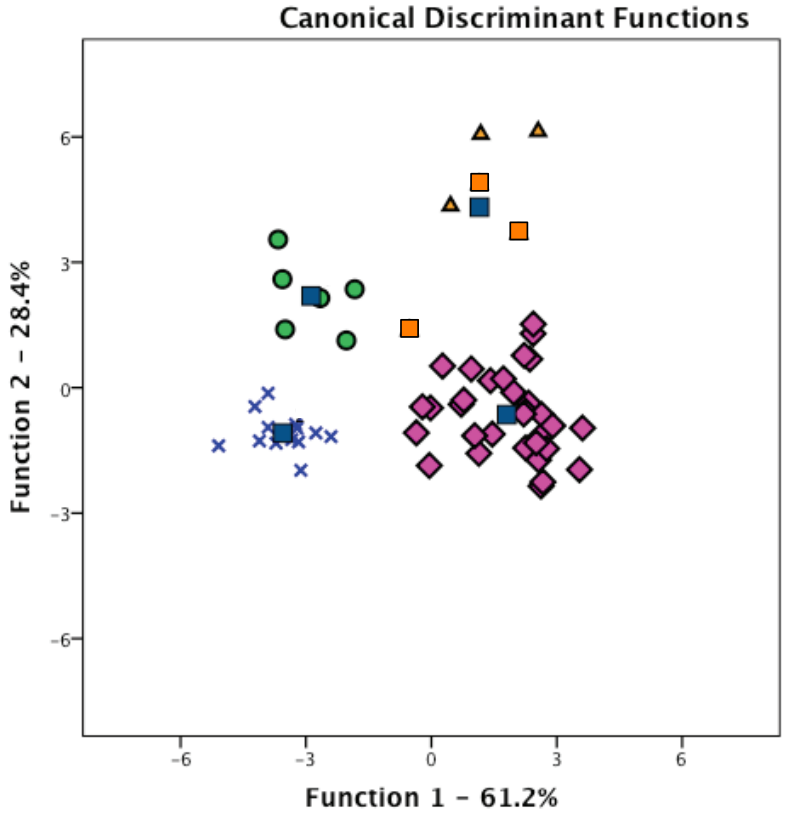


Figure 4.11. Ordination plot of Multigroup Discriminant Function Analysis of fruit and vegetative characters. A total of 89.6% of the variation is explained by function 1 (61.2%) and function 2 (28.4%). ◆*Elaeocarpus arnhemicus* (pink), ✕ *E. obovatus* (blue), ● *E.* sp. Mt. Bellenden Ker (green), ▲ *E. coorangooloo* (orange) and ■ ADC (orange). Closed squares (dark blue) are group centroids.

Table 4.11. Predicted classification (top row within each entity) and cross validation (bottom row within each entity) results of Multigroup Discriminant Function Analysis of fruit and vegetative characters.

	Predicted group membership				Total
	E. arnhemicus	*E. obovatus*	*E.* sp. Mt Bellenden Ker	*E. coorangooloo* and ADC	
E. arnhemicus	32	0	0	0	32
	32	0	0	0	32
E. obovatus	0	13	0	0	13
	0	13	0	0	13
E. sp. Mt Bellenden Ker	0	0	6	0	6
	0	0	6	0	6
E. coorangooloo and ADC	1	0	0	5	6
	1	0	0	5	6

4.3.5 Classification tree analysis

Specimens of ADC exhibit morphology that is somewhat intermediate between the other entities. While none of the analyses indicated the consistent placement for this morphological group, it is unlikely to cause difficulties in the identification of specimens of the *E. obovatus* complex generally because it is only known from a few sites on the lower slopes of the Lamb Range near Cairns, and. therefore, the specimen was excluded from the CART analyses. One of the ADC specimens (Lyons 158), collected from out side of the ADC site, was identified to be closest to *E. arnhemicus* based on MDFA analysis of fruit + vegetative data set and it was included as *E. arnhemicus* in the CART analysis.

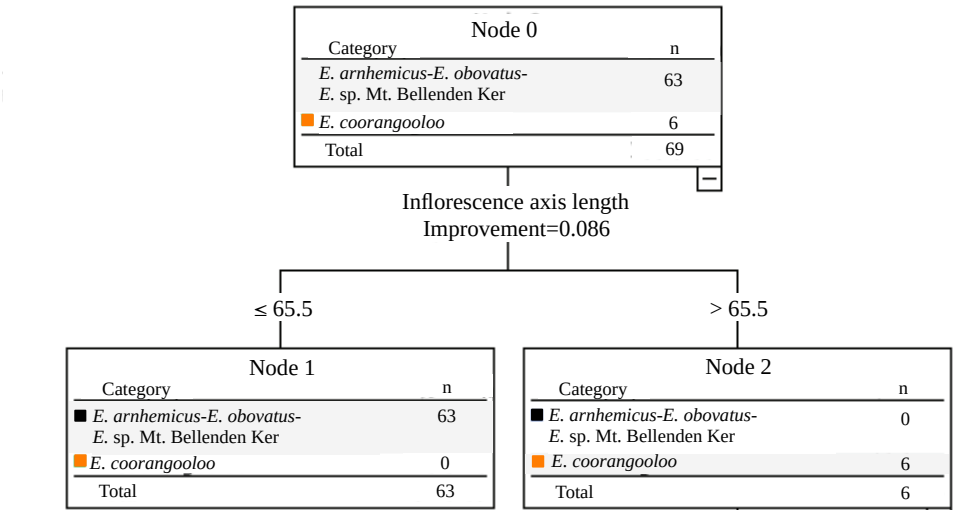
4.3.5.1 Flower and vegetative characters

The cluster and ordination analyses undertaken in this study clearly segregate specimens of *E. coorangooloo* from the *E. obovatus* complex (Figure 4.4, Figure 4.6, Figure 4.8 and Figure 4.10) based on floral +vegetative characters. On the other hand, there was not a clear segregation within the *E. obovatus* complex like there was between *E. coorangooloo* and the *E. obovatus* complex. This was well reflected in the MDFA cross validation analyses. Nearly 8% of specimens were misclassified within the *E. obovatus* complex. Preliminary CART analyses repeatedly failed to build an effective model to identify morphological groups within the *E. obovatus* group; therefore an analysis to identify characters that segregates only *E. coorangooloo* from the rest of the specimens was employed here.

CART analysis of 69 specimens using two reference groups (*E. coorangooloo*, and the rest) identified two variables to be influential and built two effective models separately on inflorescence axis length (character 11) and petiole length (character 5).

Based on inflorescence axis length, the split distinguishes *E. coorangooloo* (> 65.5 mm) from the *E. arnhemicus-E. obovatus-E.* sp. Mt. Bellenden Ker group (inflorescence axis length ≤ 65.5 mm (Figure 4.12 a). Similarly, the spilt, based on petiole length, segregates *E. coorangooloo* (> 21.3 mm) from the *E. arnhemicus-E. obovatus-E.* sp. Mt. Bellenden Ker group (≤ 21.3 mm) (Figure 4.12 b). Thus, *E. coorangooloo* can be clearly segregated from the rest of specimens based two features readily observable in the field: inflorescence axis length and petiole length.

a.



b.

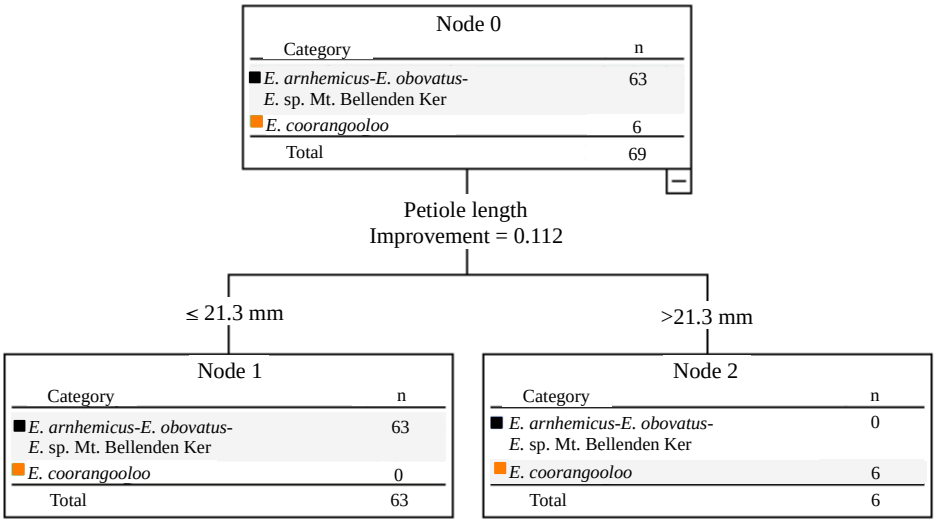


Figure 4.12. Classification tree analysis using the CART algorithm of selected floral and vegetative characters. CART model based on a. inflorescence axis length and b. petiole length. Black and orange denote the *E. arnhemicus*+*E. obovatus*+*E.* sp. Mt. Bellenden Ker, and the *E. coorangooloo* groups, respectively.

4.3.5.2 Fruit and vegetative characters

E. coorangooloo can be segregated clearly from the rest of the *E. obovatus* complex with inflorescence axis length and petiole length, and are available through fruiting time on the tree, The entity was excluded from the analyses.

A second CART analysis was undertaken of 58 specimens using three reference groups (*E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker) identified three variables (Figure 4.13): stone wall thickness (character 32), petiole length (character 5) and leaf length (character 2). The first split, on stone wall thickness, separates a group (Node 1) consisting of *E. obovatus* and four of the seven specimens of *E.* sp. Mt. Bellenden Ker (≤ 1.7 mm), from a group (Node 2) consisting of *E. arnhemicus* and the remaining three specimens of *E.* sp. Mt. Bellenden Ker (> 1.7 mm). Thus, this character is effective at distinguishing *E. arnhemicus* from *E. obovatus*, but by itself is not effective at distinguishing *E.* sp. Mt. Bellenden Ker from those two entities. At the second level the remaining two characters effectively distinguish *E.* sp. Mt. Bellenden Ker: petiole length > 7.7 mm separates the four specimens of *E.* sp. Mt Bellenden Ker which were grouped with *E. obovatus* at node 1, and leaf length (≤ 67.1 mm separates the three specimens that were grouped with *E. arnhemicus* at node 2.

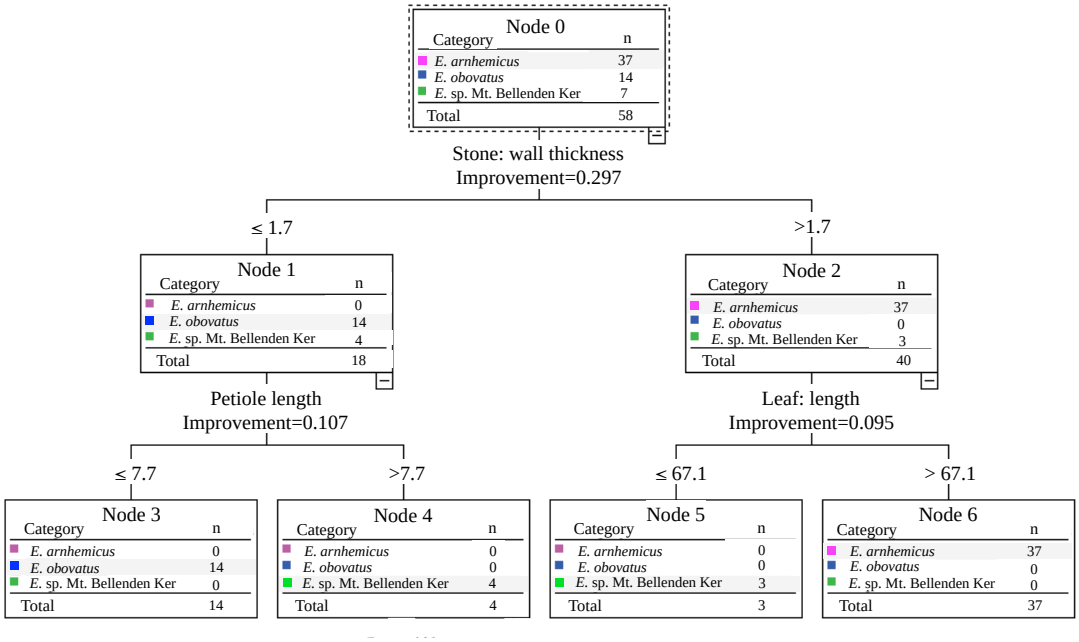


Figure 4.13. Classification tree analysis using the CART algorithm of selected fruit and vegetative characters. Pink, blue, and green denote *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker respectively.

4.4 Discussion

4.4.1 Character states

4.4.1.1 Vegetative characters

Some patterns were detected in the factor loadings of the second and the third PCA components whereas the first component did not show an obvious trend. The first component of the PCA was influenced by various types of leaf features, which do not seem to show any direct correlation, whereas the second and third components showed a strong trend in correlations in component factors. The second component of the PCA was influenced by variables that dictate leaf appearance (size, leaf shape and acumen length) while the third PCA component was influenced by venation characters; i.e. high density of secondary veins giving an appearance of congested venation as opposed to lesser number of veins.

It is interesting to note that the leaf length to width ratio was considered important by Coode (1984) for distinguishing *E. obovatus* and *E. arnhemicus*, but that it did not appear to be influential in segregating clusters in the present analysis compared to the other vegetative characters.

4.4.1.2 Floral characters

The factor loadings suggest that the size of floral characters were significant in the PCA. The first component, which had the strongest influence in segregating *E. coorangooloo* from the other specimens, was influenced by differences in size ranges in characters such as inflorescence axis, style, petal and sepal. However, no other obvious patterns were observed in the second and third factor loadings for each PCA.

The flexible UPGMA dendrograms showed a distinction between the entities while other analyses did not support distinct groups between *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker. Only *E. arnhemicus* appeared to be more segregated from the other entities but the boundaries between them were not clear in any of the ordination methods.

Coode (1984) suggested that petal division number and petal width might be useful for separating his variants, however the results presented here show variation in these characters to be taxonomically uninformative for the within-group segregations.

4.4.1.3 Fruit and vegetative characters

In all the factor and cluster analyses of fruit and vegetative characters, three stone characters (stone length, width and stone wall thickness) were highly influential. Discriminant function analysis separated all of the entities clearly on fruit and vegetative characters, with a high ratio for both axes. However, it is noteworthy that fruit characters alone do not have enough discriminatory power to resolve clear groups. CART analyses based solely on fruit characters also failed to segregate groups (result not shown).

Only one of the variants identified by Coode (1984), *E.* sp. Mt. Bellenden Ker, was retrieved in analyses of fruit and vegetative characters.

4.4.2 Methodological evaluation

PCA is a powerful tool for reducing the multi-dimensionality of data sets and is also useful for identifying the characters that are most influential in separating the groups. However, a disadvantage of PCA is that often as much as 40% or more of the variation is not accounted for. In contrast, SSH MDS accounts for all of the variation in the data set.

The results of the PCA and SSH MDS analyses of the floral and vegetative character set showed similar clustering patterns, however they differed in the sets of influential characters that were identified (similarly in the fruit and vegetative data set). Conflicting results may reflect a lack of strong signal in the data and therefore caution must be exercised in interpreting them.

It is also notable that the SSH MDS analyses revealed clearer clustering patterns when characters that were identified as influential in the PCA analyses were used, but not when all characters were used (result not shown). Therefore, it is recommended to employ PCA as a guide for detecting underlying groupings and to identify the most influential characters, which may then be used in SSH MDS analysis.

4.4.3 Taxonomic implications

The results of the morphometric analyses provide evidence for the existence of morphologically discrete groups within the *E. obovatus* complex corresponding to *E. arnhemicus*, *E. coorangooloo*, *E. obovatus* and *E.* sp. Mt Bellenden Ker. These groups are briefly discussed below.

4.4.3.1 *Elaeocarpus arnhemicus*

The *E. arnhemicus* group was clearly resolved by PCA, UPGMA and SHH MDS analyses of fruit and vegetative characters and somewhat resolved by UPGMA and SHH MDS analyses of floral and vegetative characters. The CART classification analysis also resolved *E. arnhemicus* as distinct and identified several features that clearly distinguish the entity from the others: thick (>1.7mm) fruit stone wall (character 32), and long (>67.1mm) leaves (character 2). Additional features identified through factor and cluster analyses that help to characterise *E. arnhemicus* are the large fruit stones (characters 30 and 31), long petioles (character 5), small number of secondary veins (character 6), large bracteoles (character 12) and 4-5 ovules per locule.

One of the putative morphological intermediates (*E. obovatus* – N. Queensland 'intermediates' 2, Table 4.1) represented in the analysis by two specimens with largish fruit from NW of Laura (*Baba 660* and *Baba 633*) and one specimen with small leaves (*Baba 651*) clustered with *E. arnhemicus* in this study. Although *Baba 651* possesses the smallest leaf length amongst *E. arnhemicus*, no distinct subclusters were identified within the *E. arnhemicus* group, therefore there are no grounds for recognising this intermediate as a separate morphological entity.

4.4.3.2 *Elaeocarpus obovatus*

Specimens of *E. obovatus* were clearly resolved as a group by PCA, UPGMA and SHH MDS analysis of fruit and vegetative characters, MDFA and CART analyses also supported *E. obovatus* as a distinct entity and identified the characters which separate the entity from the others: thin (<1.7mm) fruit stone wall (character 32), and short (<7.7mm) petiole (character 5). Other features identified through factor and cluster analyses as useful for discriminating this entity are the small fruit stone (character 30 and 31), narrow leaves (character 3), low to medium number of secondary veins (character 6), small secondary venation angle (character 7), low number of teeth on the leaf margins (character 9) and small bracteoles (character 12).

The SHH MDS analysis of floral and vegetative characters further divided the group into two clusters which loosely correspond to specimens collected from the north and the south of its range. However, this pattern was influenced greatly by only one character - petal size - and no other analysis supported this pattern. Therefore, I conclude that there is no strong evidence that *E. obovatus* North and *E. obovatus* South populations are morphologically distinct and that *E. obovatus* should be treated as a single entity.

Specimens of *Elaeocarpus obovatus* 'variant' (Table 4.1) from SE Queensland (*Baba 754, 756* and *727*) which show somewhat wider leaves than 'typical' *E. obovatus* did not segregate on any analysis. Therefore I conclude that there are no grounds for recognising this variant.

Specimens of *E. obovatus* N. Queensland 'intermediates' 1 (Table 4.1) from near Proserpine (*Baba 706, 708, 711*) show similarity in floral characters to *E.* sp. Mt. Bellenden Ker, and have longer petioles. However, no analysis segregated these specimens from *E. obovatus* therefore they should be treated as *E. obovatus*. While the long petioles which characterise these specimens may lead to difficulty in distinguishing some *E. obovatus* specimens from *E.* sp. Mt. Bellenden Ker, other characters such as the lamina base tapering into petiole can help to identify them.

4.4.3.3 *Elaeocarpus* sp. Mt. Bellenden Ker

Most of the analyses clearly resolve an *E.* sp. Mt. Bellenden Ker group, and CART analysis indicated that this group can be characterised by thin-medium stone wall (character 32), leaves ≤ 67.1 mm long (character 2) and petioles > 7.7 mm long (character 5). Other characters, based on factor and cluster analyses, which distinguish *E.* sp. Mt. Bellenden Ker are the large secondary vein angle (character 7), high number of secondary veins (character 6), small number of stamens (character 26), medium to long perianth segments (characters 18, 15, 20 and 21), long pedicels (character 14), and medium sized fruit (characters 30 and 31).

Elaeocarpus sp. Mt. Bellenden Ker can be further distinguished from *E. obovatus* and *E. arnhemicus* on the following characters: medium to large tree, long leaf acumen, the possession of six ovules per locule. Additionally, characters of taxonomic importance that were not included in the analysis are: pedicels which are recurved in the early bud stages (Figure 4.14) and a high frequency of axils bearing two inflorescences are useful to identify this entity. Variation in these characters is difficult to quantify therefore they were not included in this study.

Specimens identified by Coode (1984) as *E. obovatus* N. Queensland 'intermediates' 5 (Table 4.1) from Mt. Lewis (*Sanderson 546*) and Mt. Spec (*Webb and Tracy 8192, Francis s.n.*) are here determined as *E.* sp. Mt. Bellenden Ker. Furthermore, the one included specimen of *E. obovatus* N. Queensland 'intermediates' 3 (Table 4.1) – *Hyland 5618* – which possesses 4-6 ovules per carpel is shown to cluster with *E.* sp. Mt. Bellenden Ker. The occurrence of four ovulate carpels is rare for this entity and was only observed on one specimen in this study.

4.4.3.4 ADC

Specimens of *E. obovatus* N. Queensland 'intermediates' 4 from Davies Creek (west of Cairns, Qld), labelled in this study as ADC (*Baba 821* and *Baba 823*), exhibit features of several different entities: the long acumens, deeply fringed petals and leaf venation are similar to *E.* sp. Mt. Bellenden Ker, the large fruit is reminiscent of *E. arnhemicus*, and long petioles and long pedicels are similar to *E. coorangooloo*. Generally, the ADC specimens clustered with *E. coorangooloo* specimens on analyses of floral and vegetative characters, and with *E. arnhemicus* and *E. coorangooloo* specimens on analyses of fruit and vegetative characters. The MDFA analysis supported the close association of ADC with *E. coorangooloo* specimens using floral, and fruit and vegetative characters, whereas the CART analysis using flower and vegetative characters did not successfully separate *E. coorangooloo*-ADC from the rest of the specimen characters using floral and vegetative characters (result not shown). This may be an indication of inflation of the MDFA results caused by the heterogeneous sample size. This trend was also observed in the fruit and vegetative characters. Distinguishing specimens of ADC from the other entities is challenging, except on geography.

Similarly to complex nature of the morphology of these specimens, the results of the population genetic analyses (Chapter 3) showed that these specimens represent genetic admixtures. However these admixtures were of two *E. obovatus* genetic types, rather than *E. arnhemicus* or *E.* sp. Mt. Bellenden Ker, the morphologically closer entities. Given this discordance between the morphological and genetic results, I consider it premature to assign these specimens to an existing taxon or to erect a new one for them. Should further occurrences of this entity be discovered, the new samples may help to resolve the problem of the relationships of this entity and shed light on the question of whether it represents an incipient lineage or a hybrid population.

4.4.3.5 Placement of *Elaeocarpus coorangooloo*

Elaeocarpus coorangooloo was strongly segregated from the other entities included in this study based on floral and vegetative characters (Figure 4.4, Figure 4.6, Figure 4.8 and Figure 4.10). On the other hand, specimens of *E. coorangooloo* showed striking similarity to those of *E. arnhemicus* and ADC on the basis of fruit and vegetative characters. Based on CART analysis, the length of the inflorescence axis can be used to segregate individuals of *E. coorangooloo* (> 65.5 mm) from the others even in the non flowering season, because the inflorescence axes linger on the tree, with the length of them unchanged, well beyond the fruiting season. While the occasional individual of *E. arnhemicus* with a long inflorescence axis

can be found, the petiole length will clearly place it to species. Therefore *E. coorangooloo* can be comfortably distinguished from the other entities based on morphology.

Population genetic analyses revealed a mixed genetic identity for the samples of *E. coorangooloo*: one sample showed a strong genetic similarity to *E. obovatus* and *E.* sp. Mt. Bellenden Ker and the other to *E. obovatus*. As has been argued for the ADC population, further occurrences and samples are required to determine the nature of this entity and therefore provide a solid basis for any taxonomic change.

4.4.4 Notes on ecology and phenology

Each of the entities considered in this study seems to have specific habitat requirements. *Elaeocarpus arnhemicus* is a subcanopy tree, reaching 10-12 m in height, which occurs in seasonal monsoon, riparian forests and occasionally in forests on consolidated dunes. It is almost exclusively found in dry country, and is never encountered in wet rainforest.

Elaeocarpus sp. Mt. Bellenden Ker is a subcanopy (10 - 20 m) tree of wet upland rainforest (the only exception to date being a collection from lowland rainforest in the Daintree) and its distribution is typically narrow (see the map:
Figure 3.1), restricted to the Wet Tropics Bioregion.

Elaeocarpus obovatus is a subcanopy to canopy tree (20–35 m) found in dry rainforest, wet sclerophyll forest, gallery forest and littoral rainforest. Its altitudinal range is from sea level to 1000 m. There are some areas where heterogeneous rainforest comes into contact with dry and wet rainforest, especially the high altitude forests west of Cardwell. As the morphology of *E. obovatus* and *E.* sp. Mt. Bellenden Ker is similar, specimens collected around this area can be problematic. To aid identification of the two entities and to determine whether they occur in true sympatry, further studies in the area with a focus on fine scale distribution and possible correlation with environmental parameters such as rainfall, substrate, canopy cover and associated species and communities, would be beneficial.

The phenology of the species is also different. *Elaeocarpus arnhemicus* develops flowers coinciding with the end of the wet season when moisture in the soil decreases (from March to August) whereas *E. obovatus* and *E.* sp. Mt. Bellenden Ker flower earlier (towards the end of the dry season, October to December).

4.4.5 Taxonomic rank

Elaeocarpus arnhemicus and *E. coorangooloo* are each clearly distinct from all other entities on most of the morphometric and genetic analyses. *Elaeocarpus* sp. Mt. Bellenden Ker and *E. obovatus* were not clearly distinguished in many morphometric analyses indicating apparent continuity in variation in most characters, however there are two qualitative characters, which separate them: the number of racemes per axil and curvature of the pedicels in early bud. In *E.* sp. Mt. Bellenden Ker the flowering axils almost always bear two racemes (one in the other species) and the pedicels on early buds are tightly recurved. Furthermore, the genetic analyses clearly separate *Elaeocarpus* sp. Mt. Bellenden Ker and *E. obovatus*. Taken together, this evidence supports the recognition of *E.* sp. Mt. Bellenden Ker at species rank. The name *E. biracemosus* Y.Baba & Crayn sp. nov., in reference to the occurrence of two racemose inflorescences in each axil, is proposed.

4.5 Taxonomic account

This study has generated substantial new knowledge of morphological variation within and among *Elaeocarpus arnhemicus*, *E. obovatus* and *E. coorangooloo*, and has resolved the limits of and appropriate rank for the entity previously known as *E.* sp. Mt. Bellenden Ker. Therefore, provided here are revised accounts of the first three species, and a full taxonomic description of the last as *E. biracemosus* Y.Baba & Crayn sp. nov.. Keys to the species are provided.

4.5.1 Keys to the *E. obovatus* complex

Keys to the *E. obovatus* complex are provided below. These are based on Coode's (1984) keys to flowering and fruiting material in his treatment of Australian and New Zealand *Elaeocarpus* (Coode 1984 p. 519 and p. 521) and replace couplets 5 and 15 in the key to flowering and fruiting material respectively. Measurements appear in the key are an average over up to five measurements per individual. Users are encouraged to take more than one measurement per specimen as variations in the measurement exist even within the same individual.

Key to flowering material:

5. Ovary glabrous

    6. Anthers clearly awned, disk glabrous (Australia: Victoria, New South Wales, Queensland)……..…….Group VII

    6. Anthers without awns though sometimes the posterior tooth may be pointed; disc hairy (the hairs may be very short)

7. Leaves heterophyllous and heteroblastic - those of young plants variously smaller, or narrower, and usually deeply lobed, some mature leaves also lobed; petal apex lobed, not divided into linear divisions (New Zealand) ……..…...…… 4. *E. hookerianus*

7. Leaves not heterophyllous; petal apex divided into various numbered linear divisions (Australia: New South Wales, Queensland)

8. Inflorescence axis length > 65.5 mm, if shorter, petiole length > 21.3 mm (Australia: North Queensland) … ……..... 9. *E. coorangooloo*

8. Inflorescence axis length ≤ 65.5 mm, if longer, petiole length ≤ 21.3 mm (Australia: Queensland & New South Wales) .....…….……...… 5. *E. obovatus*

5. Ovary hairy (at least with very sparse and various length of hairs)

9. Anthers clearly awned

10. Petioles (1.5-)2-4(4.5) cm long; petals hairy inside (Australia: Queensland) ……..…….. 19. *E. ruminatus*

10. Petioles 0.9-1.4 cm long; petals glabrous inside (Australia: Lord Howe Is.) ……. 26. *E. costatus*

9. Anthers without awns, though bristles may be present

11. Petals glabrous

12. Petals obovate, apex rounded and lacking apical divisions (Australia: Victoria & New South Wales) ..... 18. *E. holopetalus*

12. Petals divided at apex

13. Leaves adpressed-hairy beneath; petals divided in usually 3 +triangular lobes at apex, lobes occasionally subdivided (New Zealand) ……..... 3. *E. dentatus*

13. Leaves tomentose or finally glabrous beneath; petals with more numerous narrower apical divisions

14. Leaves glabrous or virtually so; petals up to 4.5 mm long

14A. Petioles red, bracteole margin once toothed, locules 3 (Australia: North Queensland) ……...…. 27. *E. hylobroma*

14A. Petioles green, bracteole margin entire, locules 2 (Australia: North Queensland, Northern Territory)

14B. Two racemes born per leaf axis almost persistently, pedicels recurved in the early bud developing stage (Australia: North Queensland) ……..…….. *E. biracemosus* sp. nov.

14B. One raceme born per leaf axis almost persistently, pedicels straight in the early bud developing stage (Australia: North Queensland & Northern Territory……....….. 6. *E. arnhemicus*

14. Leaves persistently tomentose beneath (and also petioles); petals 11-13 mm long (Australia: North New South Wales) ……..... 10. *E. williamsianus*

Key to fruiting material:

15. Stone with 3 sutures (Australia: North Queensland) ……..…. 27. *E. hylobroma*
15. Stone with 2 sutures
16. Stone wall rather thin and weak, surface not very rugose; fruit dark purple; leaves heterophyllous, often irregularly deeply lobed even on twigs with fruit; buds gummy-resinous (New Zealand) ……...…. 4. *E. hookerianus*
16. Stone walls well developed strong, surface clearly rugose; fruit bright or dull blue; leaves serrate at most, not heterophyllous; buds not gummy-resinous (Australia: New South Wales, Queensland)
17. Inflorescence axis length ≤ 65.5 mm .....…….…… *E. obovatus* complex
17. Inflorescence axis length > 65.5 mm
18. Petiole length ≤ 21.3 mm …….……..…….. *E. obovatus* complex
18. Petiole length > 21.3 mm (Australia: North Queensland) .…… 9. *E. coorangooloo*

Key to the *Elaeocarpus obovatus* complex based on the fruit and vegetative characters

1. Stone wall thickness ≤ 1.7 mm ..……...……...……..……2
1. Stone wall thickness > 1.7 mm ...…….……..…..…..……..……....3

2. Petiole ≤ 7.7 mm long, if shorter the base of lamina tapering into minute wings which terminate in a small pulvinus (Australia: Queensland & New South Wales) ……..……..... 5. *E. obovatus*
2. Petiole > 7.7 mm long, if longer the base of lamina terminates abruptly in a distinct pulvinus (Australia: North Queensland) ……. *E. biracemosus* sp. nov.

3. Leaf ≤ 67.1 mm long, if shorter, specimen is from upland and/or wet rainforests (Australia: North Queensland) ..……...……...… *E. biracemosus* sp. nov.
3. Leaf > 67.1 mm long, if shorter, specimen is not from the above (Australia: North Queensland and Northern Territory) ……..…… 6. *E. arnhemicus*

4.5.2 *Elaeocarpus arnhemicus* F. Muell., Descr. Notes Pap. P1. 1: 6 (1875)

Type: (Australia, N. Territory) Rivulets of the Fitzmaurice River, *F. Muell.* (lectotype K [Coode 1978: 202]; isolectotype MEL!).

E. obovatus var(?). *foveolatus* Benth Fl. Austral. 1: 281 (1863). Syntypes: N. Australia. Islands of the N. coast (n.v.), R. Brown; Liverpool River (n.v.), A. Cunningham; Fitzmaurice River and Macadam Range, *F. Mueller* (n.v.)

Small to mid sized *tree* to 15 m tall. *Branchlets* light green, shortly sparsely hairy (hairs ≤ 0.2 mm) when young, glabrescent at maturity, lenticels present; vegetative buds covered in adpressed dense hairs (≤ 0.4 mm); stipules black, narrowly triangular, 1 – 1.5 mm long, hairy (< 0.15 mm), caducous. *Leaves* scattered along the branch, petioles light green, hairy (< 0.5mm), (3–) 5– 18 (–22) mm long, often swollen inconspicuously at distal end, curved or straight; lamina dark green adaxially, paler green abaxially, elliptic-obovate or obovate, (33–) 56.8 – 110.0 (–123) x (19–) 24.2 – 49.8 (–57) mm, adaxial surface scarcely hairy (hairs > 0.5 mm) on the midrib, abaxial surface sparsely hairy (hairs < 0.5 mm) apex acute to acuminate, (1–) 3 – 20 (–26) mm; base cuneate; serrations (3–) 6 – 15 (–19), each terminated by a minute black tooth, secondary veins (5–) 6 – 11 (–13) pairs, angled at (20–) 32 – 52 (–66) degrees to the midrib; domatia present or absent in secondary vein axils, when present pocket-shaped, 1 – 10 (–17) per leaf, sparsely hairy outside (< 0.5 mm), glabrous inside. *Inflorescences* axillary, racemose, borne amongst leaves, rarely two inflorescences per axis; axis (6.5–) 11.5 – 45.8 (–58) mm long, pale green, hairy (< 0.1 mm rarely exceeding 0.1 mm); pedicels (1.1–) 1.3 – 4.7 (–5.1) mm in flower elongating up to 14 mm in fruit, hairy (≤ 0.1mm); flowers (5–) 8 – 21 (–24) per inflorescence, ovoid to ellipsoid; floral bracteole one per flower, caducous, ovate-deltoid, (0.4–) 0.5 – 1.5 (–1.6) mm long, hairy abaxially (> 0.1mm), glabrous adaxially. *Buds* cream, narrowly ovoid, apex conical, sparsely hairy (< 0.1mm); pedicels straight. *Flowers* (4 or) 5 (or 6) merous, narrowly ovoid; sepals cream to white, narrow to broad deltoid or broad ovate, (2–) 2.3 – 3.6 (–3.7) x (0.7–) 1.1 – 1.8 (–2) mm, apex acute, hairy (< 0.1 mm) along the inner line of edge and tip adaxially, sparsely and adpressed-hairy or glabrous abaxially (≥ 0.1 mm); petals white, narrow obdeltoid, oblong or obovate, (2–) 2.3 – 4.1 (–4.5) mm x (0.9–) 1 – 2.2 (–2.5) mm, hairy (≥ 0.1 mm) at the bottom adaxially, glabrous abaxially, upper $^1/_3$ - ½ of petal fringed with (6–) 8 – 12 (–14) equal narrow linear segments; disk orange, 0.3 – 0.6 (– 0.9) mm high, densely hairy (c. 0.1 mm), lobed; stamens (9–) 12 – 25 (–27), cream, borne between disk and ovary; filaments (0.2–) 0.3 – 0.5 (–0.7) mm long, minutely hairy (< 0. 1 mm), sigmoid, curved or straight; anthers (0.6–) 0.7 – 1.8 mm, dehiscing via an apical transverse slit, outer tip longer by 0.1 mm, shortly and sparsely hairy with antrorse hairs (< 0.02 mm), apex acute with hairs (c. 0.1 mm) longer than the body; ovary pale green, hairy (> 0.1 mm), (0.6–) 0.8 – 1.4 (–2) x 0.6 – 1.2

mm, tapering to style, 2-locular, 4-5 ovules per locule; style filiform 0.8 – 2.4 (–2.5) mm long, sometimes sigmoid in the top $^{2}/_{3}$ – $^{1}/_{6}$; stigma indistinguishable from style. *Fruit* bright blue, globular to round ellipsoid, (11.2–) 11.6 – 18.2 mm long x (7.4–) 7.9 – 12.9 (– 13.6) mm in diameter, glabrous; outer mesocarp (flesh) c. 2.5 mm thick, stone wall (1.6–) 1.7 – 3.4 mm thick sometimes less, surface baculate, sutures 2, 8.5 – 15.9 x 5.9 – 10.1 (–10.9) mm. *Seed* one (– 2) per fruit, (4.9–) 5.6 – 10.1 mm long; endosperm entire, straight; cotyledons broadly elliptic, straight.

Note: Non-quantifiable character that is not included in this study but potentially informative is a straight pedicel in the developing inflorescence axis (Figure 4.14).

Specimens examined: PAPUA NEW GUINEA. Western District. Mabaduan, 13 Jul. 1968, Henty E.E NGF 38621 (CNS); Near Weam, 10 Aug. 1967, Ridsdale NGF 33638 & Galre (CNS).

AUSTRALIA**.** Northern Territory. Darwin & Gulf Distr.: Arnhem Land, 15 Apr. 2000, *Cowie* 8876 (BRI, CNS); Fitzmaurice River, Victoria Bonaparte, Oct. 1855, F. *Mueller* s.n (Lectotype MEL!); Ramingining, 2 Nov. 1987, *Russel-Smith* 3922 (BRI, CNS); Mary River National Park, 1 km downstream from Arnhem Highway on Mary River, 29 Dec. 2010, *Cooper* 2133, *Cooper*, *Morris* & *Dempster* (CNS). Queensland. Cook Distr.: Pajinka Walk, Northern Peninsula Area, 26 Jul. 2010, *Baba* 586, *Lifu*, *Eden* & *Kilgour* (CNS); Muddy Bay, Northern Peninsula Area, 26 Jul. 2010, *Baba* 587, *Lifu*, *Eden* & *Kilgour* (CNS); *ditto* 588 (CNS); *ditto* 589 (CNS); Mutee Head, Northern Peninsula Area, 1 Aug. 2010, *Baba* 600 & *Kilgour* (CNS); Rocky Creek, Bramwell Station, 2 Aug. 2010, *Baba* 603 & *Kilgour* (CNS); Rocky Creek, Bramwell Station, 3 Aug. 2010, *Baba* 606 & *Kilgour* (CNS); *ditto* 610 (CNS); Rocky Creek, 3 Aug. 2010, *Baba* 619 & *Kilgour* (CNS); 9 km South of Batavia Downs, 3 Aug. 2010, *Baba* 611 & *Kilgour* (CNS); *ditto* 612 (CNS); Portland Road, c. 94km West of Chilli Beach, near Iron Range National Park, 21 Jul. 2010, *Baba* 547 & *Kilgour* (CNS); Picaninny Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary, 22 Jul. 2010, *Baba* 549 & *Kilgour* (CNS); *ditto* 551 (CNS); *ditto* 552 (CNS); Scrubby Creek, Australian Wildlife Conservancy Picaninny Plains Sanctuary, 22 Jul. 2010, *Baba* 565 & *Kilgour* (CNS); Coen River, Mungun Kaanju National Park, 4 Aug. 2010, *Baba* 624 & *Kilgour* (CNS); Horsetailers Waterhole, Mungun Kaanju National Park, 5 Aug. 2010, *Baba* 630 & *Kilgour* (CNS); Edge of small billabong, Mungun Kaanju National Park, 5 Aug. 2010, *Baba* 633 & *Kilgour* (CNS); Governors Waterhole, Mungun Kaanju National Park, 5 Aug. 2010, *Baba* 634 & *Kilgour* (CNS); *ditto* 635 (CNS); Coen River, Mungun Kaanju National Park, 5 Aug. 2010, *Baba* 642 & *Kilgour* (CNS); Archer River, Mungun Kaanju National Park, 6 Aug. 2010, *Baba* 651 & *Kilgour* (CNS); Stewart River, Lama Lama National

Park, 7 Aug. 2010, *Baba* 655 & *Kilgour* (CNS); *ditto* 657 (CNS); *ditto* 660 (CNS); Morehead River, 8 Aug. 2010, *Baba* 662 & *Kilgour* (CNS); *ditto* 664 (CNS); Hann Crossing Campsite 10, Rinyirru (Lakefield) National Park, 14 Jul. 2010, *Baba* 489 & *Kilgour* (CNS); *ditto* 491 (CNS); Midway Waterhole, Rinyirru (Lakefield) National Park, 14 Jul. 2010, *Baba* 497 & *Kilgour* (CNS); Breeza Plains Outstation, Rinyirru (Lakefield) National Park, 14 Jul. 2010, *Baba* 494 & *Kilgour* (CNS); *ditto* 496 (CNS); Saltwater Creek, Rinyirru (Lakefield) National Park, 14 Jul. 2010, *Baba* 507 & *Kilgour* (CNS); Kennedy Bend, Rinyirru (Lakefield) National Park, 13 Jul. 2010, *Baba* 484 & *Kilgour* (CNS); Archer Point, South of Cooktown, 11 Jul. 2010, *Baba* 480 & *Kilgour* (CNS); *ditto* 479 (CNS); Annan River, South of Cooktown, 10 Jul. 2010, *Baba* 477 & *Kilgour* (CNS); Sankowsky Arboretum, ex Big Mitchell River, 1 Jul. 2011, *Baba* 847 (CNS); Foot of Walshs Pyramid, 3 Jan. 1996, *Lyons* 158 (BRI).

4.5.3 *Elaeocarpus obovatus* G.Don, Gen. Syst. 1:559 (1831)

Type: 'Native of New Holland'- no type designated or found (Coode, 1984:534)
Neotype (here designated): Australia, Queensland, Stony Creek, near Didcot, 210m, 31 Jan. 1982, *Forster 260B*

Don (1831) did not designate a type, nor gave a detailed description. Don's protologue was chiefly based on the specimens housed in the private herbarium of Lambert in London (Don, 1831: 4). Upon Lambert's death those specimens were auctioned and distributed to herbaria around the world, and only scattered records of this diaspora exist (Miller, 1970). Enquiries were made of the major herbaria who received specimens (Miller, 1970) but none were able to locate specimens of *E. obovatus* potentially seen by Don. Therefore I conclude that the specimens seen by Don are probably lost and a neotype must be selected.

Elaeocarpus donianus F.Muell. Catalogue of Plants under Cultivation in the Melbourne Botanic Gardens: 19 (1858). *nom. superfl.*

Elaeocarpus parviflorus A.Rich., A. Rich., Sert. Astrol.: 67, t.24 (1834).
Type: Crescit in Nova- Hollandia [Port Jackson, Sydney], collector unknown (holotype? P - a specimen of Richard's, labeled *'E. micranthus* nob.' Coode, 1984:534)

Elaeocarpus eucalyptifolius R.Knuth. Feddes Rep. 49: 73 (1940).
Type: Queensland, Brisbane R., 1863-5, Dietrich s.n. *(holotype: B=).* Mt. Coonowrin, Hubbard 4118 (Paratypes: K!, MEL (n.v.), BRI!); Pullen Creek, zwischen Riverview und Moggill, Hubbard 4822 (Paratype: K (n.v.))

Small to large *tree*, to 36 m tall. *Branchlets* light green, shortly sparsely hairy (≤ 0.2 mm) when young, glabrous at maturity; vegetative buds covered in sparse hairs (c. 0.1 mm); stipules black, narrowly triangular, 0.2 – 0.4 mm long, glabrescent (hairs < 0.1 mm), caducous. *Leaves* scattered along the branch, petioles light green, hairy (≥ 0.1 mm), (2–) 3.2 – 8 (–10) mm long, often with pulvinae at distal end, curved or straight; young leaves dark red, lamina dark green adaxially, paler green abaxially, narrow-elliptic to obovate, (35–) 52 – 98 (–111) x (9–) 11 – 29 (–37) mm wide, adaxial surface glabrous except at the very base of the midrib where they are pubescent (hairs c. 0.3 mm long), abaxial surface with scattered hairs (> 0.1 mm long); apex acute to acuminate, (1–) 2.8 – 10.2 (–16) mm long; base attenuate and decurrent into the petiole forming a minute wing; serrations (3–) 4 – 11 (–12), each terminated by a minute black tooth, secondary veins in (4–) 6 – 11 (–13) pairs angled at (14–) 20 – 43 (–56) degrees to the midrib; domatia present or absent in secondary vein axils, when present deltoid-shaped, 1 – 4 (–6) per leaf, with a few hairs outside (c. 0.4 mm long) and glabrous inside. *Inflorescences* axillary, racemose, borne amongst leaves, one (rarely two) per axil; axis (6.5–) 15 – 61 (–71) mm long, pale green, hairy (< 0.1 mm); pedicels (1.6–) 2 – 4.8 (–5.5) mm, hairy (< 0.1mm); flowers 6 – 38 (–45) per inflorescence; floral bracteole one per flower, caducous, deltoid, (0.3–) 0.4 – 0.8 (–1.2) mm long. *Buds* cream coloured, narrowly ovoid, apex conical, sparsely hairy (c. 0.05 mm), pedicels slightly curved at the top end when buds immature. *Flowers* 5 (or 6)–merous, narrowly ovoid to nearly obdeltoid; sepals cream to white, obovate-deltoid, (2–) 2.6 – 4.5 (–4.6) x (0.6–) 0.8 – 1.6 (–1.9) mm, apex acute, hairy along the inner line of edge and tip adaxially (≥ 0.1 mm), sparsely hairy (≤ 0.1 mm) to glabrous abaxially; petals white, obovate to oblong (2.3–) 2.5 – 4.4 (–4.7) mm x (0.7–) 0.8 – 1.98 (–2.5) mm, pubescent (hairs ≥ 0.1 mm) at the inner edge and the bottom adaxially, glabrous abaxially, upper $^{1}/_{3}$ - ½ of petal fringed with (4–) 6 – 12 (–13) equal narrow linear segments; disk orange, 0.3 – 0.6 (–0.8) mm high, densely hairy (c. 0.1 mm), lobed; stamens (10–) 13 – 20 (–25), cream, borne between disk and ovary; filaments (0.1–) 0.2 – 0.7 mm long, hairy (< 0.1 mm), sigmoid, curved or straight; anthers (0.7–) 0.9 – 2 (–2.3) mm, dehiscing via an apical transverse slit, outer tip longer by c. 0.05 mm, shortly and sparsely hairy with antrorse hairs (c. 0.05 mm), apex acute with hairs (c. 0.1 mm); ovary pale green, glabrous, 0.7 – 1.9 x 0.6 – 1.1 mm, tapering to style, 2-locular, 4 - 6 ovules per locule; style filiform 0.7 – 2.3 (–2.5) mm long, sometimes sigmoid in the top $^{2}/_{3}$ – $^{1}/_{6}$; stigma indistinguishable from style. *Fruit* bright blue, globular to round ellipsoid, 6.5 – 10.4 mm long x 5.4 – 8.1 mm in diameter, glabrous; outer mesocarp (flesh) c. 0.7 mm thick, stone wall 0.9 – 1.4 (–1.5) mm thick, surface bacculate, sutures 2, 5.3 – 8.3 (–9.3) x 4.0 – 6.0 mm. *Seed* one (– 2) per fruit, (2.0–) 3.2 – 5.6 (–5.8) mm long; endosperm entire, straight; cotyledons broadly elliptic, straight.

Notes: Specimens of *E. obovatus* with longer than usual petioles always have leaf bases which taper to form a 'wing', ending in the small pulvinus.

Specimens examined: AUSTRALIA**.** Queensland. North Kennedy Distr.: Undara Crater, Undara Crater National Park, 6 Oct. 2010, *Baba* 685, *Poll* & *Kilgour* (CNS); *ditto* 687 (CNS); Hansen Cave, Undara Crater National Park, 7 Oct. 2010, *Baba* 692 & *Kilgour* (CNS); Palm Tree Creek, off Thornton's Gap Road, Harvey Range, West of Townsville, 13 Oct. 2010, *Baba* 696 & *Kilgour* (CNS). South Kennedy: Mackay Botanic Garden- ex Mt Basset, Mackay, 14 Oct. 2010, *Baba* 698, *Champion*, *Davidson* & *Kilgour* (CNS); Mt. Basset Cemetery, Mackay, 14 Oct. 2010, *Baba* 703, *Champion* & *Kilgour* (CNS); Kelsey Creek, Kelsey Creek Road, Proserpine, 15 Oct. 2010, *Baba* 706 & *Kilgour* (CNS); *ditto* 708 (CNS); *ditto* 711 (CNS). Port Curtis Distr.: Yapoon hill, 13 Oct. 1987, *Batianof* 9261 (BRI); Yapoon, 30 Nov. 1939, *Sloan* s.n. (BRI); Blackmans Gap, 21 km from the Bruce Highway near Miriam Vale, 5 Nov. 1992, *Forster* 12281 (BRI). Wide Bay Distr.: Five Mile Creek, 'Booringa', SE of Kalpowar, 4 Dec. 1995, *Bean* 9343 (BRI); 3.5 km E of Childers, Conlons Road, 9 Nov. 1999, *Forster* 25188 (BRI); Stony Creek, near Didcot, 31 Jan. 1982, *Forster* 260B (Neotype: BRI!); Tinnanbar, 18 Mar. 1999, *Brooks* 90 (BRI). Moreton Distr.: Mt Coonowrin, Glass House Mountains, 21 Sep. 1930, *Hubbard* 4118 (BRI); Belmont Hill Reserve, Brisbane, 27 Oct. 2010, *Baba* 726 & *Kilgour* (CNS); Enoggera Creek, Brisbane, 27 Oct. 2010, *Baba* 727 & *Kilgour* (CNS); Mary Road Reserve, Brisbane, 28 Oct. 2010, *Baba* 728 & *Kilgour* (CNS); Oxley Creek, Brisbane, 5 Nov. 2010, *Baba* 758A & *Kilgour* (CNS); Mudgeeraba Creek, 28 Oct. 2010, *Baba* 730 & *Kilgour* (CNS); Road to Mt Bunya, 2 Nov. 2010, *Baba* 745 & *Kilgour* (CNS); *ditto* 747 (CNS); Haly Creek Road, near Mt Bunya, 3 Nov. 2010, *Baba* 748 & *Kilgour* (CNS); *ditto* 750 (CNS); Obi Obi Creek, near Maleny township, 4 Nov. 2010, *Baba* 752 & *Kilgour* (CNS); Mooloolaba Beach, 4 Nov. 2010, *Baba* 754 & *Kilgour* (CNS); *ditto* 756 (CNS); *ditto* 757 (CNS).

New South Wales. North Coast Distr.: Fernleigh Rd, Fernleigh, 26 Dec. 1990, *Wilson* 7667 (BRI); Tweed River, Opposite Scott's Island, 20 Sep. 1976, *McDonald* 1602 (BRI); Hogan Park, Tweed River, Opposite Scott's Island, 19 Feb. 1975, *Williams* 75005 (BRI); c. 13 km S of Urunga, on the Pacific Highway29 Sep. 1969, *Coveny* 2152 (BRI).

4.5.4 *Elaeocarpus biracemosus* Y.Baba & Crayn, sp. nov.

Elaeocarpus sp. (=BG/2287) Hyland & Whiffin in Australian Tropical Rainforest Trees: 107 (1993)

Elaeocarpus sp. Mt Bellenden Ker (L.J.Brass 18336) Guymer in Names and Distribution of Queensland Plants, Algae and Lichens: 66 (1997); Guymer in Census of the Queensland Flora: 61 (2002, 2007, 2010)

Elaeocarpus sp. (Bellenden Ker) Cooper & Cooper in Fruits of the Australian Tropical Rainforest: 162 (2004)

Elaeocarpus sp. 'Bellenden Ker' Crayn & Kupsch in Austral. Pl. 23: 362 (2006)

Elaeocarpus sp. 'BK' (Rossetto et al., 2009)

Type: Australia, Queensland, Topaz Road, Topaz, Atherton Tableland, *725m, 12 Feb. 2012 Baba 860, Cooper, Bransgrove, Jensen* (holotype CNS!; isotype to be distribute to BRI!).

Diagnosis: Similar to E. obovatus *G.Don and* E. arnhemicus *F.Muell. but differing from both in having usually two racemes per axil and pedicels that are reflexed in the early stage of bud development. The taxon also differs from* E. arnhemicus *in having smaller fruit and thinner fruit stone walls, and from* E. obovatus *G.Don in having cuneate lamina bases and longer petioles.*

Small *tree* to 20 m tall. *Branchlets* light red or green, shortly sparsely hairy (≤ 0.1 mm) when young, glabrescent; vegetative buds densely covered in adpressed hairs (c. 0.1 mm long); stipules black, narrowly triangular, 0.25 – 0.35 mm long, hairy (< 0.1 mm), caducous. *Leaves* scattered along the branches, petioles light green, hairy (≥ 0.1 mm), short (3–) on sterile branches, 6 – 13 (–19) mm long on fertile branches, often with pulvinae at both ends, straight; lamina dark green adaxially, paler green abaxially, elliptic, (47–) 55.5 – 81 (–93) x (17–) 22.3 – 30.8 (–32) mm, adaxial surfaces glabrous except very base of the midrib being hairy (> 0.2 mm), abaxial surfaces with scattered hairs (> 0.2 mm), apex acute to acuminate, (3–) 5.5 – 12.0 (–15) mm; base cuneate; margin sometimes with light red coloration, serrations (4–) 6 – 14 (–16) each terminated by a minute black tooth, often undulate; secondary veins (8–) 10-13 (–17), pairs angled at (30–) 43 – 54 (–63) degrees to the midrib; domatia present or absent in secondary and sometimes tertiary vein axils, when present pocket-shaped, 0 – 6 (–13) per leaf, glabrous inside. *Inflorescences* axillary, racemose, borne amongst leaves, usually two inflorescences per axil; axis 26.5 – 48.5 mm long, pale green, hairy (< 0.1 mm rarely exceeding 0.1 mm); pedicels 3.5 – 4.5 mm in flower elongating up to 5.5 mm in fruit, hairy (< 0.1mm); flowers 17 – 46 per inflorescence; floral bracteole one per flower, caducous, deltoid, 0.3 – 0.7 mm long. *Buds* cream coloured, narrowly ovoid, apex conical, hairy (c. 0.1mm), bud pedicels recurved when immature. *Flowers* 5 (or 6)-merous, narrowly ovoid to obdeltoid; sepals cream to white, deltoid, (2.85–) 3 – 3.5 (–4.1) x (0.8–) 1 – 1.4 mm, apex acute, sparse long hairs (≥ 0.1 mm) along the midline and surface, and short hairs (<0.05 mm) along the inner line of edge and tip adaxially, sparsely hairy (≥ 0.1 mm) abaxially; petals white, narrow obdeltoid to oblong (3.2–) 3.2 – 4.2 mm x (1.0–) 1.1 – 1.5 (–1.7) mm, hairy (≥ 0.1 mm) at the bottom adaxially,

glabrous abaxially, upper $^1/_3$ - ½ of petal fringed with 6 – 8 equal narrow linear segments; disk orange, c. 0.3 – 0.6 mm high, densely hairy (c. 0.1 mm), lobed; stamens 10 – 12, cream, borne between disk and ovary; filaments 0.3 – 0.6 mm long, hairy (< 0.05 mm), sigmoid, curved or straight; anthers 1.1 – 1.2 (–1.4) mm, dehiscing via an apical transverse slit, outer tip longer by 0.1 mm, shortly and sparsely hairy with antrorse hairs (< 0.05 mm), apex acute with hairs (c. 0.1 mm) longer than the body; ovary pale green, hairy (> 0.1 mm), 0.9 – 1.3 x 0.6 – 1 mm, tapering to style, 2-locular, 6 ovules per locule; style filiform (0.7–) 1.2 – 2.3 mm long, sometimes sigmoid in the top ⅕ –$^1/_6$; stigma indistinguishable from style. *Fruit* bright blue, globular to round ellipsoid, (8–) 12 – 14 x mm long x (6.4–) 9 – 10 mm in diameter, glabrous; outer mesocarp (flesh) thin, stone c 1.5 mm thick sometimes less, surface baculate, sutures 2, (6.8–) 7.1 – 10.2 x 5 – 6.8 mm. *Seed* one (– 2) per fruit, 4.5 - 6.7 mm; endosperm entire, straight; cotyledons broadly elliptic, straight.

Note: A qualitative character not included in this study but which is potentially definitive of this taxon is the presence of tightly recurved pedicels in the developing inflorescence axis (Figure 4.14).

Specimens examined: AUSTRALIA**.** Queensland. Cook Distr.: Cloudland Nature Reserve, 14 Nov. 2010, *Baba* 768, *Barrabe* & *Rigault* (CNS); Topaz, Atherton Tableland, 12 Nov. 2010, *Baba* 760, *Cooper*, *Barrabe* & *Rigault* (CNS); *ditto* 761 (CNS); Topaz, Atherton Tableland, *12 Feb. 2012*, *Baba* 860, *Cooper*, *Bransgrove & Jensen* (holotype CNS!; isotype to be distribute to BRI!); Koombooloomba Forest Reserve, Tully, 9 Nov. 2004, *Ford* 4478 & *Hewett* (CNS); East Downey Logging Area, 2 Nov. 1971, *Hyland* 5618 (CNS); Mount Lewis Forest Reserve, 16 Nov. 1995, *Forster* 18119 (CNS); Ongera Logging Area, 16 Nov. 1984, *Gray* 3691 (CNS); Windsor Tableland National Park, 8 Oct. 2008, *Ford* 4466 (CNS); Tully Falls Road, Atherton Tableland, 18 Mar. 2009, *Costion* 2092 (CNS); Mt. Bellenden Ker, Wooroonooran National Park, 7 Apr. 1948, *Brass* 18336 (CNS); Topaz, Atherton Tableland, 9 Feb. 2005, *Cooper* 1916 (CNS). North Kennedy Distr.: Mt Spec, NE of Townsville, 31 May 1965, *Webb* 8192 & *Tracy* (CNS); Paluma Range, Jackes 20102 (CNS).

4.5.5 *Elaeocarpus coorangooloo* F. Bailey & C. T. White, Dept. Agric., Bot. Bull. 19: 4 (1917).

Type: Queensland, Atherton District, H. W. Mocatta, (Rec'd through Director of Forests Feb. 1917) (lectotype BRI!; isolectotype K).

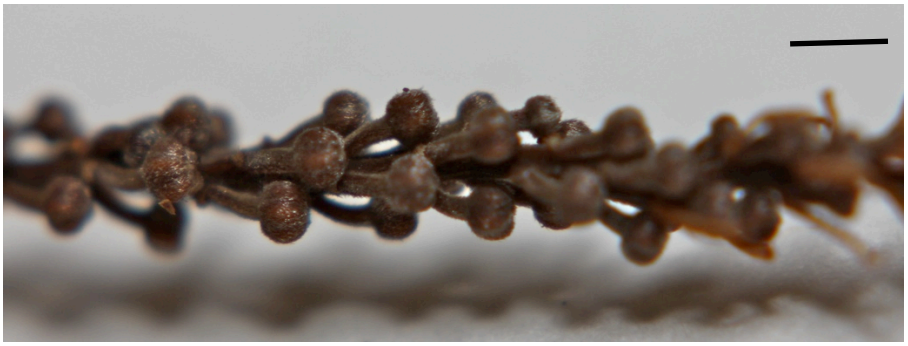
Tree small to large to 35m. *Branchlets* light green, glabrous, covered with lenticels; vegetative buds covered in sparse hairs (c. 0.1 mm); stipules black, narrowly triangular, 0.2 – 0.4 mm long, glabrous, caducous. *Leaves* scattered along the branch, petioles light green, glabrous, (16.0–) 24.7 – 32.6 (–37.0) mm long, often with pulvinae at distal end, straight; young leaves dark red, lamina dark green adaxially, paler green abaxially, elliptic to elliptic-obovate, (60.0–) 88.5 – 114.7 (–140.0) x (25.0–) 42.0 – 55.0 (–57.0) mm wide, both surfaces glabrous, apex round, acute or acuminate 0 – 14.0 (–21.0) mm; base cuneate or broadly so; margins crenate–serrate or sometimes merely sinuate, serrations (9–) 12 – 14 (–19), secondary veins not strongly developed, 7 – 9 (–12) pairs angled at (25–) 30 – 45 (–52) degrees to the midrib; domatia present or absent in secondary vein axils, when present foveolate, 1 – 6 (–9) per leaf, glabrous. *Inflorescences* axillary, racemose, borne amongst or behind leaves, one per axil; axis 30 – 130 (–140) mm long, pale green, sparsely hairy (< 0.1 mm); pedicels (4–) 5 – 9 (10) mm, sparsely hairy (< 0.1mm); flowers (17–) 34 – 50 (–56) per inflorescence, perfumed; floral bracteole one per flower, caducous, deltoid, (0.5–) 1 – 2 mm long, with dense hairs (c. 0.1 mm). *Buds* cream coloured, narrowly ovoid to ellipsoid, apex conical, sparsely hairy (c. 0.05 mm), pedicels straight, 4 – 9 (–10). *Flowers* 5 –merous, ovoid; sepals cream to white initially, becoming pinkish, obovate-deltoid, 3.2 – 5 x (1.3–) 1.5– 1.7 mm, apex acute, sometimes acutely recurved, hairy along the inner line of edge and tip adaxially (≤ 0.1 mm), sparsely hairy (≤ 0.1 mm) to glabrous abaxially; petals white, obovate to oblong-obovate, 4.2 – 5.5 mm x 1.5 – 2.2 mm, sparsely pubescent (hairs ≤ 0.1 mm) adaxially, pubescent at the bottom abaxially, ciliate on the lower part of the margins (hair c. 0.1 mm), upper $^{1}/_{3}$ of petal fringed with 6 – 10 narrow linear segments, often grouped into irregularly lobes; disk orange, 0.5 – 0.8 mm high, densely hairy (c. 0.1 mm), lobed; stamens c.25 (–30), cream, borne between disk and ovary; filaments (0.3–) 0.5 – 0.7 (–0.9) mm long, shortly and sparsely hairy with antrorse hairs (c. 0.05 mm), curved or straight; anthers 1.1 – 2.6 mm, dehiscing via an apical transverse slit, outer tip longer by c. 0.1 mm, apex acute, glabrous; ovary pale green, glabrous, 1.2 – 2 x 0.6 – 1.1 mm, tapering to style, 2-locular, usually 6 ovules per locule, with some variations of 4 – 8; style filiform 1.6 – 2.7 mm long, sometimes sigmoid in the top $^{2}/_{3}$ – $^{1}/_{6}$; stigma indistinguishable from style. *Fruit* bright blue, ellipsoid, c. 12 mm long x 9 mm in diameter, glabrous; outer mesocarp (flesh) c.2 mm thick, stone wall 1.7 – 2.5 mm thick, surface bacculate, sutures 2, (8.8–) 9.7 – 10.5 (–11.5) x 6 – – 7.8 mm. *Seed* one per fruit, 5.8 – 8 mm long; endosperm entire, straight; cotyledons broadly elliptic, straight.

Note: A qualitative character not included in this study but which is potentially useful for distinguishing this taxon is the presence of straight pedicels in the developing inflorescence axis (Figure 4.14).

Specimens examined: AUSTRALIA**.** Queensland. Cook Distr.: Atherton, Atherton Tableland, 8 Oct. 2010, *Baba* 695 (CNS); Hallorans Hill, Atherton Tableland, 5 Aug. 1979, *Irvine* 1941 (BRI); Wongabel, Atherton Tableland, 18 Aug. 1948, *Smith* 3791 (BRI); Carson Road, Malanda, Atherton Tableland, 18 Aug. 2003, *Forster* 29540 (BRI); Mazlin Creek, Atherton Tableland, 4 Jul. 1954, *Volck* 770 (BRI); Phillips Avenue, Atherton, Atherton Tableland, 17 Mar. 1983, *Hyland* 12637 (CNS).

a)

b)



c)

d)

Figure 4.14. Inflorescence with developing buds. a) *E. biracemosus* (Risley 124) showing reflexed pedicels, b) *E. obovatus* (Batianoff 5060) showing bent or straight pedicels c) *E. arnhemicus* (Hyland 2242RFK) showing straight pedicels, d) *E. coorangooloo* (Sanderson 1666) showing straight pedicels. Photos: a), b) and c) Andrea Lim and d) Queensland Herbarium. Bars represent 1mm.

4.6 Further work

To date, no study has systematically documented the ecology of members of the *E. obovatus* complex therefore it remains to be seen if there are fine scale habitat and ecological requirements that differentiate the groups identified in this study. This information is useful to understand niche differentiation specifically in relation to the ecotonal wet sclerophyll-wet rainforest vegetation where *E. biracemosus* and *E. obovatus* occur in broad sympatry.

Another avenue of research is to quantify the three-dimensional shape of flower, buds and fruit or two-dimensional shape of the each characters. The information of the geometry would provide extra data to assist more robust interpretation of species boundaries.

It came to my attention during fieldwork that *E. arnhemicus* was absent from many places where seemingly suitable habitat existed. Because ecological parameters, which may be important for the germination and establishment for the species, are largely unknown, superficially similar habitat may not have had any indicative value in terms of identifying suitability of habitat. Such ecological information is fundamental and one of the most crucial pieces of knowledge to understand the entities. Biological niche analyses using climatic and Regional Ecosystem data may assist in elucidating in more detail the ecological requirements of these species.

Lastly, this study included dense sampling in areas where putative entities were sympatric, because these are the areas where taxon boundaries were least distinct based on herbarium specimens and field observations. Sampling in other parts of the ranges of the species was sparse. This is especially true of the widespread *E. arnhemicus*. While it is very unlikely that increasing sampling from the western (Northern Territory and Indonesia) and extreme northern (Papua New Guinea) parts of the range would alter the results seen in this study, a deeper understanding of the geographical patterns of morphological variation may help shed further light on evolution in this group.

4.7 Conclusions

1. The specimens included in this study segregated into four main groups. PCA and MDHS analyses of the combined fruit and vegetative characters resolved three groups - *E. obovatus*, *E.* sp. Mt. Bellenden Ker, and *E. arnhemicus*/*E. coorangooloo* - whereas analysis of flower and vegetative characters resolved two groups - *E. coorangooloo*, and the rest.

2. The main morphological groupings are congruent with the molecular groupings (Chapter 3).

3. Of the fruit and vegetative characters, leaf length, petiole length, and fruit stone wall thickness were identified as having the greatest taxonomic value. Of the floral characters, inflorescence axis length and petiole length were identified as most useful for distinguishing *E. coorangooloo* from the others.

4. Revised taxonomic accounts of *E. arnhemicus*, *E. coorangooloo* and *E. obovatus* are provided, based on the abundance of new material unavailable to previous workers. A new species, *E. biracemosus* Y.Baba & Crayn was erected for material formerly included under the phrase name *Elaeocarpus* sp. Mt. Bellenden Ker, and a full description provided. Dichotomous keys to all four species based on flowering and fruiting material are included.

Chapter 5 *Elaeocarpus hylobroma* (Elaeocarpaceae): a new species endemic to mountain tops in northeast Queensland, Australia

5.1 Introduction

The family Elaeocarpaceae Juss. (including Tremandraceae R.Br. ex DC.) consists of 12 genera of mainly trees and shrubs (Coode 2004; Crayn et al. 2006). The family is widely distributed in the tropics extending into the sub tropics and temperate zone in almost all the continents, absent only from continental Africa and North America. There is a centre of diversity in Australia where nine genera are found, four of these being endemic: *Peripentadenia* L.S.Sm, *Platytheca* Steetz, *Tetratheca* Sm. and *Tremandra* R.Br. ex DC.

Elaeocarpus L. is the largest genus in the family and is distributed from Southern India along the Himalayas to South China, Japan, SE Asia, Malesia, Australia, New Zealand, islands in the Pacific and Indian Oceans, and Madagascar. Approximately 350 species are known worldwide (Coode 2004) with 34 taxa (CHAH 2011) occurring in Australia (30 endemic) including five phrase name taxa (Guymer 1997, 2007, 2010; Centre for Australian National Biodiversity Research 2010). The majority of the taxa are found along the east coast and ranges with a few extending northwest to the Northern Territory.

Infrageneric classifications of *Elaeocarpus* have been based on floral and fruit morphology (Weibel 1968; Coode 1978, 1984; Zmarzty 2001). Features used to diagnose infrageneric groups include the number of petal divisions, anther morphology, ovary indumentum, numbers of locules and ovules, stone [equivalent to putamen, comprising a woody inner mesocarp containing endocarps sensu Dettman & Clifford (2001)] ornamentation, and seed and embryo morphology. In a revision of Australian and New Zealand *Elaeocarpus*, Coode (1984) divided the then recognized 26 species among eight groups based on a scheme previously developed for Papuasian taxa (Coode 1978).

Here we describe a new species endemic to mountaintops in the Wet Tropics bioregion of northeastern Queensland, Australia. Material of this species was first collected by C.T. White in Sept. 1936 (White, C.T s.n.) from Mt Spurgeon, and Coode (1984) listed this entity as "*E.* sp. nov. 1" and subsequently it has most often been treated under *E.* sp. Mossman Bluff (D.G.Fell 1666). Despite this novelty having long been recognised, sufficient fertile material to enable its

description and an assessment of its affinities has only recently become available. All cited specimens have been seen by the first author.

5.2 Materials and methods

This study was based upon morphological observations on herbarium sheets and spirit material from CNS and BRI, and field observations by the first author. Where available, spirit materials were used in preference to dried. Characters used in the taxonomic account were as parallel as possible to the ones in the revision of Australia and New Zealand (Coode 1984). Dried flowers were rehydrated in diluted detergent and softened in a microwave oven. Rehydrated flower parts e.g. sepals, petals and stamens, were measured by removing them from the flower, imaging on 1mm grid paper using a USB digital MicroCapture version 2.0 (Publisher unknown), and viewing on a computer screen. Fruit measurements were done on spirit material sectioned with a hacksaw (16 teeth/cm). In Specimens Examined the abbreviation LA refers to Logging Area, SFR refers to State Forest Reserve, and TR refers to Timber Reserve.

5.3 Results and Discussion

5.3.1 Taxonomy

Elaeocarpus hylobroma *Y.Baba & Crayn, sp. nov.* Superficially similar to *Elaeocarpus elliffii* B.Hyland & Coode but differing in having petioles red with a pink tinge when young, leaves turning glossy bright red before falling, leaf margins minutely recurved and undulating; corolla cupiform, stamens 11 to 14, anthers with a terminal tuft of minute setae; fruit obovoid to ellipsoid, stones with bastionate ornamentation, three basally prominent sutures, seeds/embryos straight and endosperm entire (Table 5.1).

Typus: Australia, Queensland, Dinden Forest Reserve, Kahlpahlim Rock area near tower, east of Mareeba. [Lamb Range, Dinden National Park], 1290 m, Ford, Jensen & Cooper 4483, 12 Nov. 2004 (holotype: CNS!; isotypes to be distributed to BO!, BRI!, CANB!, E!, K!, KEP!, KYO!, L!, LAE!, MEL!, MO!, NSW!, P!)

http://www.ipni.org/urn:lsid:ipni.org:names:77121630-1

Elaeocarpus sp. *nov.* 1 (Coode, 1984: 582*)*
Elaeocarpus sp. (=KS/6) (Hyland and Whiffin 1993)
Elaeocarpus sp. Mossman Bluff (D.G.Fell 1666) (Guymer 1997, 2002; Hyland et al. 2003; Guymer 2007, 2010; Center for Australian National Biodiversity Research 2010)

Elaeocarpus sp. (Mossman Bluff) (Cooper & Cooper 2004)
Elaeocarpus sp. 'Mossman Bluff' (Crayn & Kupsch 2006)
Elaeocarpus sp. MB (Rossetto et al. 2009)

Small *tree* to 18 m tall, usually very poorly formed, coppicing, adventitious roots pinkish red with yellow tips. *Branchlets* pinkish red, shortly pubescent (hairs < 0.1 mm) when young, glabrescent; vegetative buds covered in long (> 0.1 mm) appressed hairs; stipules black, narrow triangular, 0.5 – 1.3 mm long, pubescent (hairs > 0.1 mm) at base, caducous.
Leaves clustered towards branch tips, petioles pinkish red, gradually turning green, often leaving only pulvinae red, both surfaces pubescent (hairs > 0.1 mm) when young, glabrescent, 5 – 19 mm long, often with pulvinae at both ends, straight; leaf blade glossy green adaxially, paler green abaxially, elliptic, narrow-elliptic to oblong, 37 – 106 (– 112) x 10 – 30 (– 37) mm, both surfaces glabrous except pubescent (hairs > 0.5 mm) on midrib, apex acute to acuminate and minutely emarginate with a very short black tooth on abaxial side; base attenuate-cuneate; margin serrate in upper 1/2 to 1/3 (rarely 2/3), serrations 2 – 6 each terminated by a minute black tooth associated with a vein ending, slightly recurved, often undulate; secondary veins (3 –) 4 – 8 pairs angled 45 – 65 degrees to midrib; domatia present in secondary vein axils, pocket-shaped, glabrous inside, 2 – 8 (– 10) per leaf, rarely absent.
Inflorescences racemose, borne amongst and below leaves; rachis (8 –) 18 – 51 mm long, pale green, sparsely hairy (c. 0.1 mm); pedicels 2.4 – 5 mm in flower elongating to 4.5 – 7 mm in fruit; flowers (3 –) 5 – 20 per inflorescence; bracteole one per flower, caducous, deltoid, once toothed, ciliate, 0.7 – 1.5 mm long. *Buds* cream with pinkish tinge, ellipsoid to ovoid, apex conical, pubescent (> 0.1 mm). *Flowers* 5-merous, cupiform; sepals cream to white sometimes with slightly pinkish tinge abaxially, deltoid, 2.5 – 3 x c. 1 mm, slightly keeled in the lower half, apex abruptly incurved and bearing a tuft of minute hairs (< 0.1 mm), sparsely hairy (c. 0.1 mm) along the midline otherwise glabrous adaxially, sparsely hairy (c. 0.1 mm) abaxially; petals white, oblanceolate, 2 – 3 mm x 0.5 – 1 mm, slightly keeled inside in lower half, glabrous adaxially, sparsely hairy along the midline abaxially, margin sparsely and shortly ciliate (≤ 0.1 mm), upper 1/3 fringed with 6 – 9 equal narrow triangular segments; disk orange, c. 0.5 mm high, densely hairy (<0.1 mm), lobed; stamens cream, borne between disk and ovary, 11 – 14; filaments (0.4 –) 0.5 – 0.7 mm long, with minute antrose hairs (< 0.1 mm), sigmoid, curved or straight; anthers 1.2 – 1.4 mm, dehiscing via an apical transverse slit, outer tip longer by 0.05 – 0.1 mm, shortly hairy with antrorse hairs (< 0.1 mm), apex acute with a cluster of short setae (< 0.1 mm) on the outer tip; ovary pale green, pubescent (> 0.1 mm), 0.7 – 1.1 x 0.8 – 1.1 mm, tapering to style, 3-locular, 4 ovules per locule; style filiform 1.4 – 1.7 mm long, sometimes sigmoid in the upper 1/4; stigma indistinguishable from style.

Fruit dull blue with pale speckles, obovoid to ellipsoid, 17.5 – 20 mm long x 11 – 13 mm in diameter, glabrous; outer mesocarp (flesh) 1.7 – 2.2 mm thick, slightly gritty, stone c. 2 – 4 mm thick, surface bastionate, sutures 3 becoming deep grooves distally, (11 –) 13.0 – 16 (– 17) x (7.5 –) 8.0 – 10.5 (– 11) mm.
Seed one (– 2) per fruit, 6.9 – 9.3 x 3.6 – 4.8; endosperm entire, straight; cotyledons broadly elliptic, straight (Figure 5.1).

DISTRIBUTION. AUSTRALIA. Queensland. Cook Distr. (Figure 5.2). The species is restricted to the Wet Tropics Bioregion of NE Queensland. It has been collected from Mount Windsor Tableland National Park, Daintree National Park (Thornton Peak, Pinnacle Rock, Mossman Bluff, Mt Lewis and Mt Spurgeon) Dinden National Park (Kahlpahlim Rock and Lambs Head), Mt Williams and Danbulla National Park and State Forest.

SPECIMENS EXAMINED. AUSTRALIA**.** Queensland. Cook Distr.:SFR 144, Whypalla, Bowerbird LA, about 6 km past West Spencer Creek crossing 26 Sept. 1997, *Ford* 2013, (CNS); Mount Windsor Forest Reserve, ca. 5.7 km from Spencer Creek Forestry barracks, site 105, 03 June 2004, *Ford* 4345 *& Hewett* (CNS); Mount Windsor Forest Reserve Road, 24 Aug. 2004, *Crayn* 888, *Rossetto, Ford & Hewett* (CNS); Mt Windsor Forest Reserve, ca. 5.7 km from Spencer Creek Forestry barracks, 22 Nov. 2004, *Ford* 4486 *& Hewett* (BRI, CNS); Mt Windsor Tableland, 22 Nov. 2004, *Sankowsky* 2544, *Ford & Sankowsky* (BRI); On old logging road, about 3 km from Spencer Ck Forest Barrack, Mt Windsor Tableland National Park, 04 Dec. 2010, *Baba* 793, *Jensen, Phoon & Roberts* (CNS); Thornton Peak, 24 Sept. 1984, Irvine 2244 (CNS); Thornton Peak, 19 Sept. 1937, *Brass & White* 221 (BRI); Thornton Peak, 14 Nov. 1973, *Hyland* 7081 (CNS), ditto, *Hyland* 7082 (CNS); Mt Spurgeon, Sept. 1936, C.T. *White s.n.* (BRI); Pinnacle Rock Track, NW of Mossman, just before the Gleichenia area, 13 Oct. 2005, *Ford* 4749, *Cooper & Russell* (BRI, CNS); Pinnacle Rock Track, 4.5 km W of Karnak, 22 June.1992, *Forster* 10709, *Sankowsky, & Tucker* (BRI); Pinnacle Rock Track, Karnak, 23 Jan. 2010, Baba 371, *Worboys & Skladaniec* (CNS); Pinnacle Rock Track, Karnak, 24 Jan. 2011, *Baba* 809, *Kilgour & Renner* (CNS), *ditto* 815; Mossman Bluff, ca 10 km W of Mossman, 30 Dec. 1988, *Fell* 1666 (BRI); Mossman Bluff track, W of Mossman Gorge, 14 June. 2010, *Baba* 456 *& Kilgour* (CNS), *ditto* 459 (CNS) & *ditto* 461 (CNS); Mount Lewis Road, 22.9 km from Rex Highway, 20 Aug. 2004, *Crayn* 867 *& Rossetto* (CNS); State Forest 143, North Mary Logging Area, 27 km along Mt Lewis road, ridge crest to East of road, 19 March 2001, *Forster* 27064, (BRI); Riflemead, SFR 143, Riflemead, North Mary LA, 05 Aug. 1988, *Hyland* 25564RFK (CNS); SFR 143, North Mary LA, 02 Feb. 1977, Irvine 1821 (CNS); TR 143, Zarda LA, near Zarda clearing, 27 Sept. 1973, *Hyland* 02909RFK (CNS); Lambs Head, Lamb Range,

30 June 1996, *Ford* 1747 (CNS); Dinden Forest Reserve, Kahlpahlim Rock area near tower, east of Mareeba, 12 Nov. 2004, *Ford* 4483, *Jensen & Cooper* (holotype: CNS!; isotypes: BO!, BRI!, CANB!, E!, K!, KEP!, KYO!, L!, LAE!, MEL!, MO!, NSW!, P!); Dinden Forest Reserve, Kahlpahlim Rock area, Lamb Range, 15 May 2004, *Ford* 4315*& McJannet* (BRI, CNS); ESE of helipad near Lambs Head, Lamb Range, site 202, 13 Nov. 2008, Ford 5413, Murphy & McKeown (CNS); Near Kahlpahlim Rock, on track, 12 Aug. 2004, *Crayn* 835, *Rossetto, Ford & Hewett* (CNS); Kahlpahlim Rock, Forest Reserve 607, Lamb Range, 12 Nov. 2004, *Cooper* 1907, *Ford, & Jensen* (BRI); Rocky outcrops vicinity of Kahlpahlim Rock, Lamb Range, 14 Aug. 2011, *Baba* 835 *& Kilgour* (CNS); SFR 185, Kauri LA, 11 May 1971, *Sanderson* 6 (CNS).

HABITAT. Mountain rainforest (complex notophyll vine forest sensu Webb, 1959), sometimes windswept and stunted, between (900 –) 1000 – 1330 m in the Wet Tropics Bioregion of northeast Queensland. Found on soils derived from granite and metasediments (A. Ford, personal communication), often growing among and over rocks.

CONSERVATION STATUS. Although the all known localities are within protected areas, the species is known from relatively few individuals and is restricted to mountain rainforest in northern Queensland. Population genetic analysis (Rossetto et al. 2009) (as *E.* sp. MB) revealed low genetic diversity within and among populations, likely resulting from contraction to refugia during Quaternary climatic oscillations, and climate change may pose a threat to this tropical mountaintop endemic. Assessment against the IUCN criteria (IUCN 2001) results in a status of 'VU D1' and against the Queensland Nature Conservation Act 1992 as 'Vulnerable, D1', based on an estimated population size of less than 1,000 mature individuals.

PHENOLOGY. Buds are recorded in late Oct. and flowers mid to late Nov. Fruit may remain on the tree until the following May.

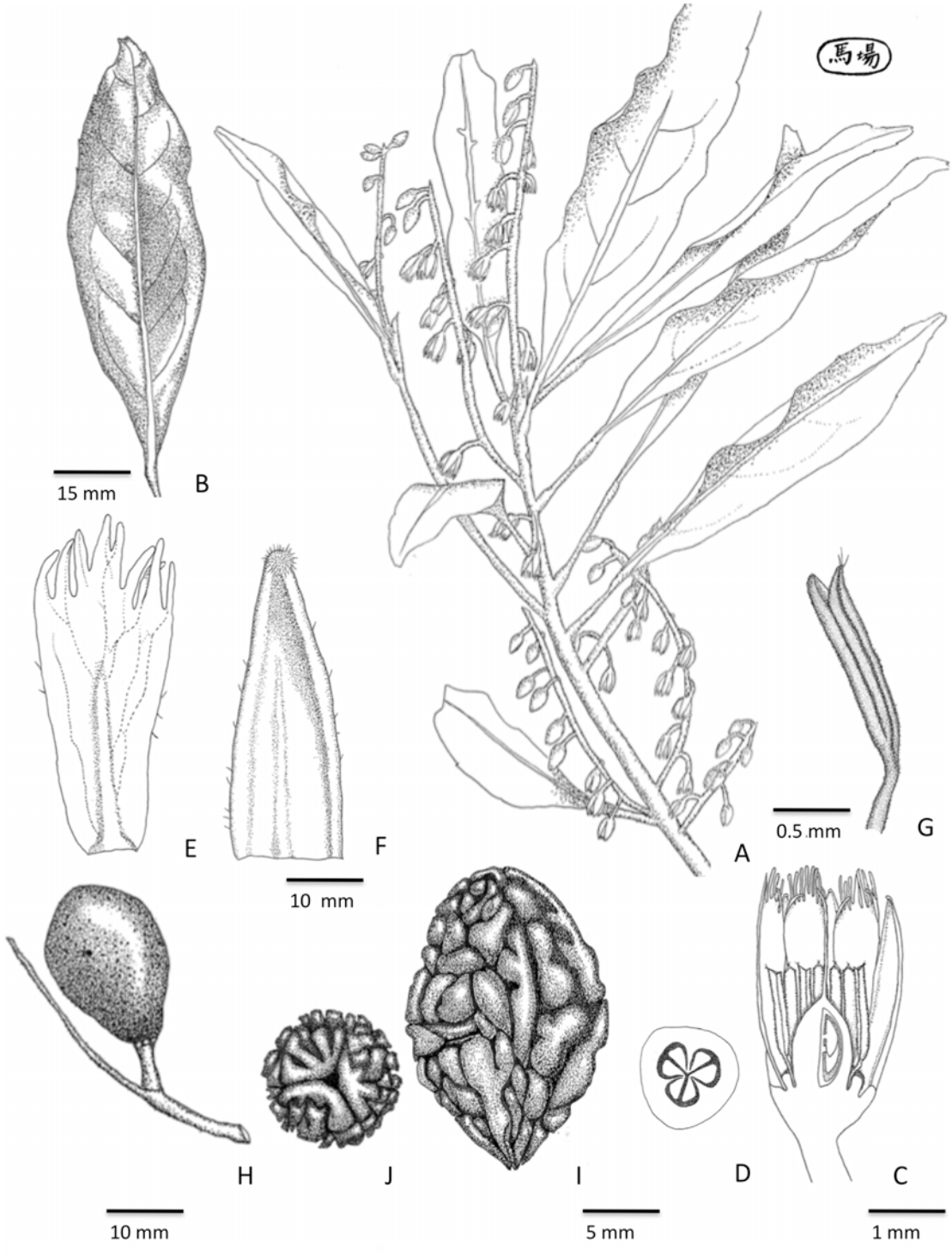


Figure 5.1. ***Elaeocarpus hylobroma*****. A Flowering twig; B Leaf enlarged to show undulation and reflexed margin; C half flower showing arrangement of floral segments and positioning of ovules in the ovary; D Placenta and ovule arrangement; E Petal adaxial surface; F Sepal adaxial surface; G stamen with minute setae; H Fruit habit; I Stone lateral view; J Stone dorsal view. A from** ***Ford*** **4483; B from** ***Baba*** **793; C & D from** ***Sankowsky*** **2544; E & G from** ***Ford*** **4486; F from** ***Ford*** **4483; H from** ***Ford*** **5413; I & J from** ***Baba*** **835. Del. Y. Baba**

Table 5.1. Morphological comparison between *Elaeocarpus hylobroma* and putatively superficially similar taxa. For the stone ornamentation, refer Rozefelds and Christophel (1996). * Information taken from Coode (1984).

	E. hylobroma	*E. elliffii*	*E. eumundi*	*E. carolinae*	*E. coorangooloo*
Distribution/ altitude range	Above 900 m.	350-1300 m	Sea level to 1000 m	1000-1200	700-1050 m
Leaf blade shape*	Elliptic, narrowly elliptic to oblong	Narrowly obovate	Obovate to elliptic	Obovate to elliptic	Elliptic or ellipticobovate
Petioles and branchlet colour	Pinkish red	Red	Green	Purple	Green
Leaf colour shed	Glossy bright red	Bright red	Red	Red	Red
Leaf teeth	Usually only upper half	Throughout	Only upper half	Throughout	Throughout
Leaf margin	Minutely recurved and undulate	Flat	Flat	Flat	Flat
Flower diameter*	c. 5 mm	c. 5 mm	9-9.5 mm	10-11 mm	c. 5 mm
Stamen number *	11-14	30-35	c. 35	15-16	c. 25
Locules *	3	3	2	2	2
Stone shape *	Ellipsoid obovoid	Ovoid	+- Ovoid or ellipsoid	Ellipsoid-fusiform	Rounded-ellipsoid
Stone size *	(11 –) 13.0 – 16 (– 17) x (7.5 –) 8.0 – 10.5 (– 11) mm	10 x 7 mm	c. 13 x 8 mm	c.12 x 8 mm,	8-9 x c.7 mm
Stone ornamentation	Bastionate	Granulose	Echinate	Echinate	Verrucate
Ovules/locule *	4	(5-) 6	c. 10	c. 8	(3-) 6
Endosperm *	Straight & entire	Hooked & ruminate	Hooked & entire	Curved & entire	Straight & entire
Embryo *	Straight	Hooked	Hooked	Curved	Straight

ETYMOLOGY. The specific epithet derives from the Greek hylo- (forest) and bromo (food). The name refers to the seed containing rich endosperm that is available as a food source for forest animals. We observed almost all fallen stones to have been gnawed through to the locule and the seed missing, suggesting seed predation by ground feeding animals, presumably native rodents.

KEY. The following replaces couplet 14 in the key to flowering material in the treatment of Australian and New Zealand *Elaeocarpus* (Coode 1984 p. 519). The species (as *E.* sp. nov. 1) was included in the key to fruiting material in the same publication (p. 520).

14. Leaves glabrous or virtually so; petals up to 4.5 mm long

14A. Petioles red, bracteole margin once toothed, locules 3 (Australia: North Queensland) ……..…….……...…. 27. *E. hylobroma* sp. nov

14A. Petioles green, bracteole margin entire, locules 2 (Australia: North Queensland, Northern Territory)

14B. Two racemes born per leaf axis almost persistently, pedicels recurved in the early bud developing stage (Australia: North Queensland) ……..….... *E. biracemosus* sp. nov.

14B. One raceme born per leaf axis almost persistently, pedicels straight in the early bud developing stage (Australia: North Queensland & Northern Territory…….....….. 6. *E. arnhemicus*

14. Leaves persistently tomentose beneath (and also petioles); petals 11-13 mm long (Australia: North New South Wales) ……..... 10. *E. williamsianus*

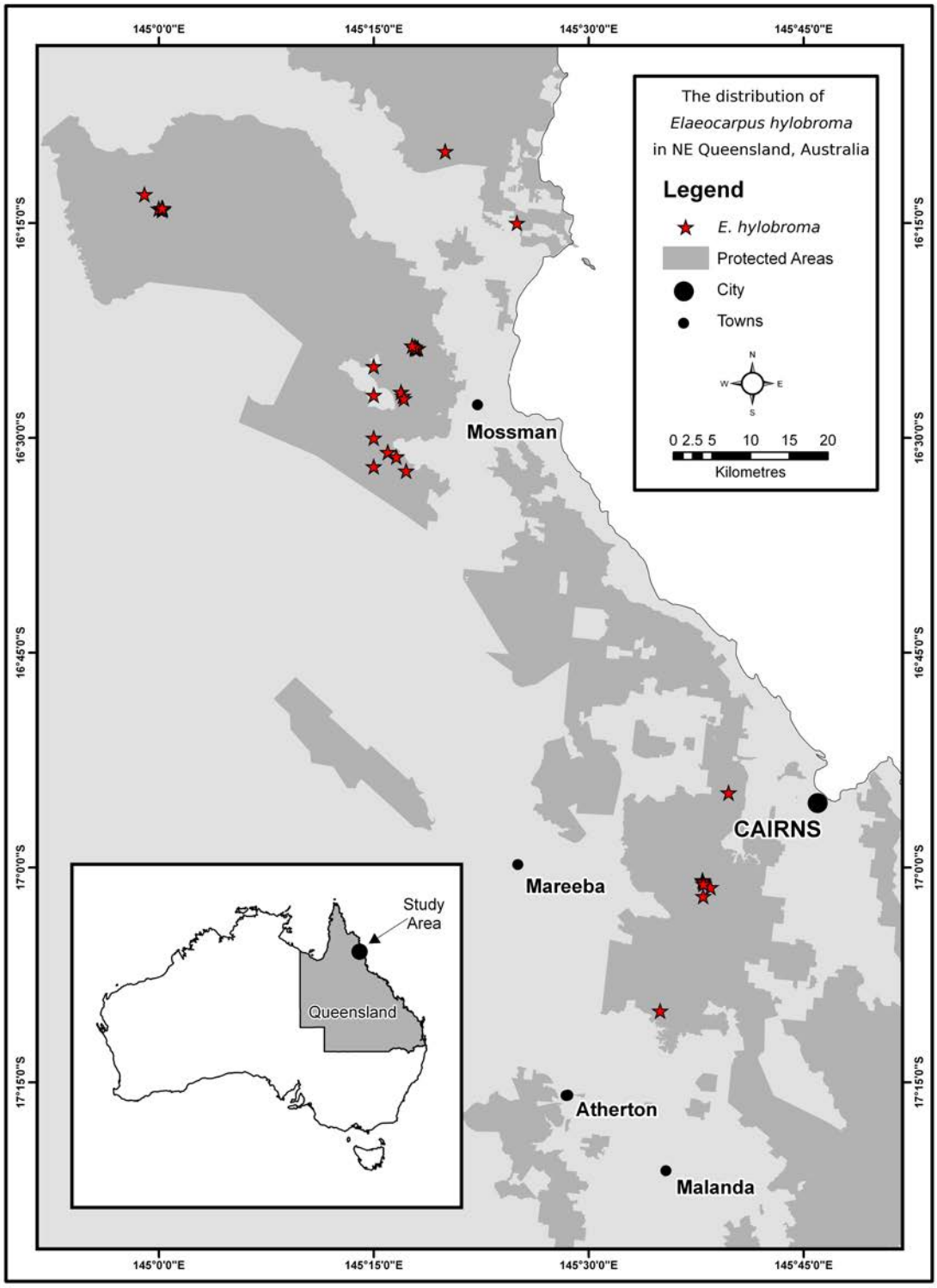


Figure 5.2. Distribution map of *Elaeocarpus hylobroma*.

RELATIONSHIPS. The morphological affinities of *E. hylobroma* (as E. sp. nov. 1) were discussed by Coode (1984). Stone ornamentation and vegetative characters suggest relationship to *E. carolinae* Hyland & Coode and some species in Group VII, and to *E. coorangooloo* J.F.Bailey & C.T.White in Group VI (Coode 1984). The species exhibits some similarity to *E. tariensis* Weibel from Papua New Guinea (M. Coode, personal communication), but *E. tariensis* differs in having hairy petioles and inflorescences, stamens with pronounced awns, and 2-locular ovaries. *Elaeocarpus hylobroma* can be distinguished from other Australian taxa by having leaves with undulate margins that are toothed in the upper half, small (c. 5 mm diam.) flowers with fewer than 15 stamens, anthers bearing a terminal group of setae, ovary with three locules each containing 4 ovules, stone with bastionate ornamentation, non-ruminate endosperm and straight embryo (Table 5.1).

Preliminary molecular phylogenetic analysis suggests *E. hylobroma* is related to members of Group V (Baba and Crayn, unpublished data): *E. angustifolius* Blume (India to Pacific), *E. grandis* F.Muell. (Australia), *E. polydactylus* Schltr. (New Guinea), *E. buderi* Coode (New Guinea), *E. dolichostylus* Schltr. (New Guinea) and *E. ptilanthus* Schltr. (New Guinea), which is supported by similarities in the number of locules and ovules, and seed morphology. *Elaeocarpus hylobroma* differs from members of this group in having fewer stamens (c. 15 rather than >20), inflorescences amongst the leaves (rather than behind), and smaller flowers (petals to 3.5 mm long rather than to 9 mm) (Coode 2010). Resolution within the clade is low in the current analysis, therefore we suggest this species to be placed in Group V but not to be assigned to subgroups until solid evidence for a placement within the group is obtained.

NOTES. Individual trees are often encountered with poorly formed trunks with large branches and crowns broken off (presumably by cyclonic and storm winds) and always with extensive basal coppicing. Red adventitious roots are frequently observed emerging from the buttress roots (Figure 5.3). Leaves turn glossy bright red before falling. Crystal-blisters appear on the aerial parts of the plants upon drying. Compared with other NE Queensland species in the genus, buds of *E. hylobroma* reach anthesis exceptionally quickly (A. Ford, personal communication).

It is noteworthy that record of this species is absent from the uplands in the southern part of the Wet Tropics Bioregion: Mt. Bellenden Ker, Mt. Bartle Frere, Mt. Elliot and the Paluma Range. These upland areas are ecologically and climatically similar to those in which *E. hylobroma* occurs and are thought to have functioned as refugia during Quaternary climate oscillations (Vanderwal *et al.* 2009; Graham *et al.* 2010).

Rozefelds and Christophel (1996) erected *E. cerebriformis* Rozefelds and Christophel for fossil endocarps from the Early to Middle Miocene Yallorn Formation in Victoria. These endocarps were described as being most similar to *E. hylobroma* (as *E.* sp. nov. 1) among extant taxa in having ellipsoid stones with bastionate ornamentation and three sutures.

Figure 5.3. Adventitious roots. Photo: C. D. Kilgour

Chapter 6 General conclusions

The primary purpose of this thesis was to assess the hierarchical relationships within the rainforest tree genus *Elaeocarpus* (including the taxonomic assessment of several putative undescribed Australian species), as the basis for a better understanding of the evolutionary processes that have given rise to the extant diversity. An hierarchical approach using different methods and datasets for different taxonomic questions was employed: phylogenetics, population genetics, morphometrics and descriptive taxonomy.

6.1 Aim one: Infer phylogenetic relationships within *Elaeocarpus* as a basis for evaluation of the existing classification and assessment of putatively new Australian taxa

6.1.1 Molecular phylogeny of Elaeocarpaceae

With the aim of establishing a reliable, robust and uniform morphological classification that reflects evolutionary relationships, the molecular phylogenetic relationships of Elaeocarpaceae were investigated using plastid *trnL-trnF* and *trnV-ndhC*, and nuclear Internal Transcribed Spacer (nrITS) sequences. Compared with previous studies (Maynard 2004; Crayn et al. 2006), this study included greater taxon sampling and additional molecular markers and provided an improved knowledge of the relationships within Elaeocarpaceae and *Elaeocarpus*.

Maximum parsimony and Bayesian analyses of the combined plastid and nuclear data provided improved estimates of relationships within *Elaeocarpaceae*. Monophyly of all the genera of Elaeocarpaceae except *Elaeocarpus*, *Aceratium* and *Sericolea* was confirmed. *Elaeocarpus*, *Aceratium* and *Sericolea* formed a strongly supported clade in the multigene tree in the Bayesian analysis but the determination of the appropriate taxonomic rank for each group still requires further investigation.

6.1.2 Relationships within *Elaeocarpus*, with special reference to the *E. obovatus* group

Some strongly supported clades with *Elaeocarpus* were concordant with morphological groups, confirming that the current classification reflects, at least in part, evolutionary relationships. Additionally, samples of the *E. obovatus* species complex (Group V D + *E. coorangooloo*) formed a monophyletic group strongly supported by most analyses. Some clades showed correlation with geographical patterns, namely New Caledonia- Pacific, and Asia.

In order to further test and evaluate the existing morphological classification a well resolved molecular phylogenetic framework is required. There is still substantial research required to obtain satisfactory resolution within the genus and between closely related genera, however, the improved understanding of the phylogeny obtained here provides insights into which aspects of the infrageneric classification are congruent with the evolutionary relationships.

6.2 Aim two: Address species-level problems (species complexes and putative new species) in Australian *Elaeocarpus* and describe new taxa where appropriate.

6.2.1 Species complex - assessment of genetic variation

With the aim to solve long-standing problems of taxonomic delimitation within the *E. obovatus* species complex, genetic variation, diversity and relatedness were assessed using a population genetics approach. Existing species circumscriptions within this morphologically closely related species group were reassessed utilising microsatellite markers.

Microsatellite profiles suggested that *E. arnhemicus* and *E. obovatus* may be tetraploid. If confirmed by other methods such as chromosome countings and flow cytometry, this would represent the first record of polyploidy within the genus. Because the appropriate methodology for analysing polyploid and diploid species together in population genetics is yet to be standardised, a synthesis of three different approaches were utilised: similarity based analysis, model based analysis, and hypothesis testing of evolutionary units.

The first approach using Principal Coordinate Analyses with various similarity/ distance measure revealed *E. arnhemicus* as a discreet and tightly clustered unit, segregated from all other entities, whereas *E. obovatus* could be further segregated into North and South populations on the basis of some analyses. Populations of *E.* sp. Mt. Bellenden Ker form a distinct cluster but also show close relationships with both populations of *E. obovatus*.

The model-based approach using STRUCTURE indicated two distinct clusters, one of which corresponds to all the samples of *E. arnhemicus* and the other consists of all the other samples. These clusters could be further segregated. The *E. arnhemicus* cluster contained two subclusters, coexisting in all of the individuals. The consistently coexisting genetic components are probably best explained by the means of seed dispersal: endozoochory and hydrochory. STRUCTURE analysis indicated that the other cluster identified in the first analysis could be further divided into three subclusters: *E. obovatus* North populations, *E. obovatus* South

populations and *E.* sp. Mt. Bellenden Ker populations. Levels of admixture within each subcluster differed; *E. obovatus* North had little or no admixture, *E. obovatus* South retaining ancestral alleles with *E.* sp. Mt. Bellenden Ker, *E.* sp. Mt. Bellenden Ker had admixture derived from both *E. obovatus* North and South populations. There is no current contact zone between *E. obovatus* North and *E.* sp. Mt. Bellenden Ker and therefore any genetic admixture is most likely due to the retention of ancestral polymorphism.

The clusters and subclusters identified from STRUCTURE and some hypothetical groupings that were not supported by STRUCTURE (such as *E. obovatus* and *E.* sp. Mt. Bellenden Ker forming distinct groups) were verified using a genetic differentiation measure approach: Analysis of Molecular Variance (AMOVA) and Multigroup Discriminant Function Analysis (MDFA). These results confirmed that there was divergence within the clusters and sub clusters and between *E. obovatus* as a whole and *E.* sp. Mt. Bellenden Ker. The result of AMOVA provided evidence for lineage isolation of these genetic clusters. Although some differentiation was found between *E. obovatus* North and South populations, it was decided that as a working hypothesis *E. obovatus* should be regarded as a single genetic entity because there exists a large sampling gap between them.

STRUCTURE analysis detected some genetic admixture between *E. arnhemicus* and the other entities, and *E. obovatus* and *E.* sp. Mt. Bellenden Ker. AMOVA indicated only up to 23 % of the variation between each pair with the likely explanation being retained ancestral polymorphism in both cases.
Taken together the results suggest that there are three distinct genetic groups: *E. arnhemicus*, *E. obovatus*, and *E.* sp. Mt. Bellenden Ker.

6.2.2 Species complex - assessment of morphological variation

Morphological variation was evaluated against the working hypothesis (existence of three genetic entities) that resulted from the population genetic study. The results of combined analysis of PCA, PCoA, Cluster analysis, Discriminant Analysis and Classification Tree analysis revealed that *E. obovatus*, *E. arnhemicus*, and *E.* sp. Mt. Bellenden Ker are morphologically discrete. The results of the morphometric study using floral and vegetative characters show some segregation of *E. obovatus* into Northern and Southern populations but these results are not supported statistically. Taken together the results of both the genetic and morphological analyses supported *E. obovatus* as a single entity. A combination of morphological characters distinguish *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker, and therefore recognising *E.* sp. Mt. Bellenden Ker at species rank is justified. Hence the two

named species, *E. obovatus* and *E. arnhemicus* are maintained and *E.* sp. Mt. Bellenden Ker is newly described as *E. biracemosus* Y.Baba & Crayn. Additionally, revised accounts of *E. arnhemicus* and *E. obovatus*, *E. coorangooloo* and keys to the species are provided.

6.2.3 Species concepts

Delimiting separately evolving metapopulation lineages (the hallmark of species: De Queiroz 2007) within the *E. obovatus* complex was a challenge due to the close relationships on both genetics and morphology. There is some genetic admixture (probably retention of ancestral polymorphism), which partly obscures the boundary between these entities. Even *E. arnhemicus*, which exhibits distinct genetic segregation from the other entities shares up to 20% of its genetic variation with them.

Three entities within the *E. obovatus* complex, corresponding to *E. arnhemicus*, *E. obovatus* and *E. biracemosus* were clearly separable on two operational criteria: diverging lineages and having a gap in morphological variation. Therefore they were deemed to be separately evolving metapopulation lineages and concluded to be distinct species.

Elaeocarpus hylobroma was identified as a divergent lineage strongly supported by analyses of the combined dataset as sister to Group V. The taxon also exhibits distinct morphology among extant *Elaeocarpus* in Australia. This taxon satisfies two operational criteria: complete lineage sorting and having a gap in morphological variation; and is therefore recognised at species rank.

6.2.4 Taxonomic implications

The phylogenetic study provided an evolutionary framework within which to place the undescribed taxa in Australia. These undescribed taxa sampled for this study now have their positions confirmed: *E.* sp. Mt Misery was nested in Group VI B; and *E.* Mt. Windsor Tableland was placed in Group XI Subgroup B. Further investigation of these entities was beyond the scope of this study, partly because insufficient material was available for a thorough analysis of variation. Preliminary assessment of *E.* sp. Windsor Tableland identified the need for thorough systematic investigation of morphological variation (using a morphometric approach) within and among this entity and its morphologically similar congeners: *E. largiflorens* and *E. thelmae*. Anecdotal field observations indicate the occurrence of *E.* sp. Mt. Windsor Tableland, *E. largiflorens* subsp. *retinervis* and *E. thelmae* in sympatry at least one site (R. Jensen pers. comm). While material of three flowering specimens (first records) and one fruiting specimen became available during the current study, these were not sufficient for thorough taxonomic

investigation. Similarly, flowering specimens of *E.* sp. Mt. Misery are only known from two populations and despite extensive field exploration by the author both at known sites and in similar habitat elsewhere, no further material was forthcoming.

In contrast to the other putative new species, detailed investigation of the long-standing putatively recognised taxon, *E.* sp. Mossman Bluff (D.G.Fell 1666) was undertaken and resulted in the description of a new species from the Australian Wet Tropics. Because preliminary assessment of the morphological diversity indicated that this entity was clearly distinct from all other Australian *Elaeocarpus*, morphometric analysis was deemed unnecessary and a descriptive taxonomical approach was employed. The taxon has been formally named *E. hylobroma* Y.Baba & Crayn, and a full description and illustrations published (Baba and Crayn 2012). The position of this species as sister to the morphologically distinct Group V was strongly supported by Bayesian analysis of the combined sequence data (including indels) and since the broader relationships of this clade are unclear the species is tentatively assigned to this group. Assignment to a subgroup is suspended until more evidence of its relationships becomes available.

The 28 currently recognised Australian species represent nine groups (Coode, 1984; Maynard et al, 2008). Through the present study two species were described and an additional lineage comprising *E. arnhemicus*, *E. biracemosus*, *E. coorangooloo* and *E. obovatus* (placed in Group V Subgroup D by Coode, 1984) was identified. The relationships of this lineage are still uncertain, therefore I recommend it be retained within Group V until new evidence shows otherwise.

6.3 Future directions

6.3.1 Further resolution of phylogenetic relationships and phylogeographical inference

The phylogenetic analysis presented in this thesis highlights some clades associated with geography yet the resolution was too poor to infer further geographical patterns. Increasing the taxon sampling in those geographical areas and increasing the resolution in the clades with data from additional markers may reveal some interesting geographical patterns. Additionally with the rich *Elaeocarpus* megafossil records available in Australia, internal and external calibration of divergence dates to assess the origin of the genus and historical biogeographical patterns of the genus may be achieved. This kind of analysis will enhance our knowledge of the evolution of the genus and may also provide insights into the origin and migration patterns of this widespread genus.

6.3.2 Assessment of other species complexes in *Elaeocarpus*

The population genetic study presented in this thesis is the first time that microsatellite loci have been used with the specific aim of delimiting taxa within a species complex in *Elaeocarpus*. This study has provided valuable insights into patterns of genetic variation that can be utilised to assess delimitation of the complex infra specific relationships. There are many *Elaeocarpus* species complexes documented in the literature that reflect the complex morphological diversity within *Elaeocarpus*. Population genetic studies using microsatellites or other markers such as amplified fragment length polymorphism (AFLP) will help in understanding these complexes. Especially some difficult species complexes that are widely distributed geographically such as the *E. angustifolius* - *E. sphaericus* - *E. grandis* clade which occurs from Australia to the Indian subcontinent, and the *E. culminicola* - *michaelii* complex which is distributed from Australia to the Philippines. The phylogenetic study in this thesis showed significant variation within the *E. angustifolius* clade, implying that *E. angustifolius* and *E. sphaericus*, synonymised under *E. angustifolius* in the past, might not be an homogeneous entity. The current study provides some evidence supporting the need for further investigation into this species complex.

References

Allen, G. A. (2001). Hybrid speciation in *Erythronium* (Liliaceae): a new allotetraploid species from Washington State. *Systematic Botany*, *26*, 263–272.

Alvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, *29*, 417–434.

Andrew, R. L., Miller, J. T., Peakall, R., Crisp, M. D., & Bayer, R. J. (2003). Genetic, cytogenetic and morphological patterns in a mixed mulga population: evidence for apomixis. *Australian Systematic Botany*, *16*, 69–80.

Aoki, K., Suzuki T., Hsu T.-W., Murakami, N. (2004). Phylogeography of the component species of broad-leaved evergreen forests in Japan, based on chloroplast DNA variation. *Journal of Plant Research, 117*, 77–94.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society, 161***,** 128–131.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*. *4*, 40-79.

Assoumane, A., Zoubeirou, A. M., Rodier-Goud, M., Favreau, B., Bezançon, G., & Verhaegen, D. (2013). Highlighting the occurrence of tetraploidy in *Acacia senegal* (L.) Willd. and genetic variation patterns in its natural range revealed by DNA microsatellite markers. *Tree Genetics and Genomes*, *9*, 93–106.

Baba, Y., & Crayn, D. (2012). *Elaeocarpus hylobroma* (Elaeocarpaceae): a new species endemic to mountain tops in north-east Queensland, Australia. *Kew Bulletin, 67*, 1–8.

Backer, C. A., & Bakhuizen van den Brink, R. C. (1963). *Elaeocarpaceae*. *Flora of Java*. Groningen, The Netherlands: N.V.P. Noordhoff.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., & Donoghue, M. J. (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*, *82*, 247–277.

Bardy, K. E., Albach, D. C., Schneeweiss, G. M., Fischer, M. A. & Schönswetter, P. (2010). Disentangling phylogeography, polyploid evolution and taxonomy of a woodland herb (*Veronica chamaedrys* group, Plantaginaceae s.l.) in southeastern Europe. *Molecular Phylogenetics and Evolution*, *57*, 771–786.

Barker, N. P., Clark, L. G., Davis, J. I., Duvall, M. R., Guala, G. F., Hsiao, C., Kellogg, E. A., Linder, H. P., Mason-Gamer, R. J., Mathews, S. Y., Simmons, M. P., Soreng, R. J., Spangler, R. E. (2001). Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden, 88*, 373–457.

Baum, D. A., & Donoghue, M. J. (1995). Choosing among alternative “Phylogenetic” species concepts. *Systematic Botany*, *20*, 560–573.

Belaj, A., Satovic, Z., Cipriani, G., Baldoni, L., Testolin, R., Rallo, L., & Trujillo, I. (2003). Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theoretical and Applied Genetics*, *107*, 736–744.

Belbin, L. (1991). Semi-strong Hybrid Scaling, a new ordination algorithm. *Journal of Vegetation, 21*, 491–496.

Belbin, L. (1993). *PATN analysis package. Users' and technical reference manuals*. Canberra: CSIRO Division of Wildlife and Ecology.

Belinky, F., Cohen, O., & Huchon, F. (2010). Large-Scale Parsimony Analysis of Metazoan Indels in Protein- Coding Genes. *Molcular Biology and Evolution*, *27*, 441–451.

Bensch, S., & Åkesson, M. (2005). Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, *14*, 2899–2914.

Bentham, G., & Hooker, J. D. (1862). Tiliaceae. In *Genera plantarum: ad exemplaria imprimis in herbariis kewensibus servata definita* (pp. 228–240). Londini, Reeve & co.

Besnard, G., Rubio De Casas, R., & Vargas, P. (2007). Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea*). *Journal of Biogeography*, *34*, 736–752.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*, 148–155.

Blanke, A., Wipfler, B., Letsch, H., Koch, M., Beckmann, F., Beutel, R. and Misof, B. (2012). Revival of Palaeoptera—head characters support a monophyletic origin of *Odonata* and *Ephemeroptera* (Insecta). *Cladistics*, 28, 560–581.

Blair, C., & Murphy, R. W. (2011). Recent trends in molecular phylogenetic analysis: where to next ? Journal of Heredity, 102, 130–138.

Bonin, A., Ehrich, D., & Manel, S. (2007). Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, *16*, 3737–3758.

Bradford, J. C., & Barnes, R. W. (2001). Phylogenetics and classification of Cunoniaceae (Oxalidales) using chloroplast DNA sequences and morphology. *Systematic Botany*, *26*, 354–385

Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, 135, 439–55.

Bromham, L. (2008). Tree of life. In *Reading the story in DNA* (pp. 224–278). New York: Oxford University Press.

Brown, E. D., & Hopkins, M. J. G. (1995). A test of pollinator specificity and morphological convergence between nectarivorous birds and rainforest tree flowers in New Guinea. *Oecologia*, *103*, 89–100.

Bruvo, R., Michiels, N. K., D'Souza, T. G., & Schulenburg, H. (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, *13*, 2101–2106.

Buerki, S., Forest, F., Acevedo-Rodriguez, P., Callmander, M. W., Nylander, J. A. A., Harrington, M., Sanmartin, I., Kupfer, P., Alvarez, N. (2009). Plastid and nuclear DNA markers reveal intricate relationshops at subfamilial and tribal levels in the soapberry family (Sapindaceae). *Molecular Phylogenetics and Evolution*, *51*, 238–258.

Butcher, R., Byrne, M., & Crayn, D. M. (2007). Evidence for convergent evolution among phylogenetically distant rare species of *Tetratheca* (Elaeocarpaceae, formerly Tremandraceae) from Western Australia. *Australian Systematic Botany*, *20*, 126–138.

Carlquist, S. (1966). The biota of long-distance dispersal. III. Los of dispersability in the Hawaiian flora. *Brittonia*, *18*, 310–335.

Carr, G. D., & McPherson, G. (1986). Chromosome numbers of New Caledonian plants. *Annals of the Missouri Botanical Garden*, *73*, 486–489.

Castro, I., & Robertson, A. W. (1997). Honeyeaters and the New Zealand forest flora: The utilisation and profitability of small flowers. *New Zealand Journal of Ecology*, *21*, 169–179.

Catania, F., Wurmser, F., Potekhin, A. A., Przyboś, E., & Lynch, M. (2008). Genetic diversity in the *Paramecium aurelia* species complex. *Molecular Biology and Evolution*, *26*, 421–431.

Center for Australian National Biodiversity Research. (2010). *Elaeocarpus* profile. *Australian Tropical Rainforest Plants Edition 6*, viewed 19 January 2012, <://www.anbg.gov.au/cpbr/cd-keys/rfk/index.html>.

Cerón-Souza, I., Bermingham, E., McMillan, W. O., & Jones, F. A. (2012). Comparative genetic structure of two mangrove species in Caribbean and Pacific estuaries of Panama. *BMC Evolutionary Biology*, *12*, 205–220.

Christophersen, E. (1935). Family 56. Elaeocarpaceae. In *Flowering Plants of Samoa* (pp. 135–140). New York: Kraus Reprint Co.

Council of Head of Australasian Herbaria (2011). Australian Plant Census. *Elaeocarpaceae*. Australian Government Initiative, viewed 3 April 2013, <http://www.anbg.gov.au/chah/apc/>.

Clark, L. V, & Jasieniuk, M. (2011). POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources*, *11*, 562–566.

Clarke, L. J., Jardine, D. I., Byrne, M., Shepherd, K., & Lowe, A. J. (2012). Significant population genetic structure detected for a new and highly restricted species of *Atriplex* (Chenopodiaceae) from Western Australia, and implications for conservation management. *Australian Journal of Botany*, *60*, 32–41.

Clout, M. N., & Hay, J. R. (1986). The importance of birds as browsers, pollinators and seed dispersers in New Zealand forests. *New Zealand Journal of Ecology*, *12*, 27–33.

Coleman, A. W. (2003). ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trens in Genetics*, *19*, 370–375.

Coode, M. J. E. (1978). A conspectus of Elaeocarpaceae in Papuasia. *Brunonia*, *1*, 131–302.

Coode, M. J. E. (1980a). An obscure new species of *Elaeocarpus* from central New Guinea. *Kew Bulletin*, *35*, 247–252.

Coode, M. J. E. (1980b). Notes on Papuasian Elaeocarpaceae. *Kew Bulletin*, *34*, 740.

Coode, M. J. E. (1984). *Elaeocarpus* in Australia and New Zealand. *Kew Bulletin*, *39*, 509–586.

Coode, M. J. E. (1987). *Crinodendron*, *Dubouzetia* and *Peripentadenia*, closely related in Elaeocarpaceae. *Kew Bulletin*, *42*, 777–814.

Coode, M. J. E. (1995). *Elaeocarpus* in the flora Malesiana area - *E. kraengensis* and ten new species from Sulawesi. *Kew Bulletin*, *50*, 267–294.

Coode, M. J. E. (1996a). *Elaeocarpus* for Flora Malesiana: Notes, new taxa and combinations in the Acronodia group. *Kew Bulletin, 51,* 267–300.

Coode, M. J. E. (1996b). *Elaeocarpus* for Flora Malesiana - notes, new taxa and combinations in sect. Elaeocarpus: 2. *Kew Bulletin, 51,* 83–101.

Coode, M. J. E. (1996c). *Elaeocarpus* for Flora Malesiana: The "Polystachyus group." *Kew Bulletin, 51,* 649–666.

Coode, M. J. E. (1998). *Elaeocarpus* for Flora Malesiana: Notes, new taxa, names and combinations for Borneo. *Kew Bulletin, 53,* 83–128.

Coode, M. J. E. (2001a). *Elaeocarpus* for Flora Malesiana: the Coilopetalum group in Sulawesi & Maluku. *Kew Bulletin, 56,* 837–874.

Coode, M. J. E. (2001b). *Elaeocarpus* for Flora Malesiana: the Verticellatae subgroup of the Monocera group and a new Philippines species. *Kew Bulletin, 56,* 885–901.

Coode, M. J. E. (2001c). *Elaeocarpus* for Flora Malesiana - the *E. stipularis* complex, *E. nitidus* group and *E. barbulatus. Kew Bulletin, 56,* 513–565.

Coode, M. J. E. (2001d). *Elaeocarpus* in New Guinea: New taxa in the Debruynii subgroup of the Monocera group. Contributions to the Flora of Mt Jaya, V. *Kew Bulletin, 56,* 449–460.

Coode, M. J. E. (2001e). *Elaeocarpus* for Flora Malesiana: the Coilopetalum group in the Lesser Sunda Islands. *Kew Bulletin, 56,* 875–883.

Coode, M. J. E. (2001f). *Elaeocarpus* for Flora Malesiana: The Fissipetalum group in central Malesia. *Kew Bulletin, 56,* 461–463.

Coode, M. J. E. (2002). *Elaeocarpus* in New Guinea: *E. sericoloides* (Fissipetalum group), *E. royenii* and *E. multisectus* (sect. Elaeocarpus). Contributions to the Flora of Mt Jaya, IX. *Kew Bulletin, 57,* 925–935.

Coode, M. J. E. (2003). *Elaeocarpus* for Flora Malesiana: two new taxa in the Coloides subgroup of the Monocera group from New Guinea. *Kew Bulletin, 58,* 453–458.

Coode, M. J. E. (2004). Elaeocarpaceae. In Kubitzki, K. (Ed.), *Families and Genera of Vascular Plants* (Vol. 6, pp. 135–144). Berlin: Springer.

Coode, M. J. E. (2005). *Elaeocarpus* for Flora Malesiana: *E. crenulatus, E. myrtoides* and *E. amabilis* from New Guinea, reconsidered. *Kew Bulletin, 60,* 305–311.

Coode, M. J. E. (2010). *Elaeocarpus* for Flora Malesiana: new taxa and understanding in the Ganitrus group. *Kew Bulletin, 65,* 355–399.

Coode, M. J. E., & Weibel, R. (1994). *Elaeocarpus* for Flora Malesiana - notes, new taxa and combinations in sect. Elaeocarpus: 1. *Kew Bulletin, 49,* 235–259.

Cooper, W., & Cooper, W. T. (2004). *Fruits of the Australian Tropical Rainforest.* Melbourne, Victoria, Australia: Nokomis Editions Pty Ltd.

Corlett, R. T. (1990). Flora and reproductive phenology of the rain forest at Bukit Timah, Singapore. *Journal of Tropical Ecology, 6,* 55–63.

Corlett, R. T. (1998). Frugivory and seed dispersal by vertebrates in the Oriental (Indomalayan) Region. *Biological Reviews*, *73*, 413–448.

Crayn, D., & Kupsch, K. (2006). Elaeocarpaceae in Australia. *Australian Plants*, *23*, 339–373.

Crayn, D. M., Rossetto, M., & Maynard, D. J. (2006). Molecular phylogeny and dating reveals an Oligo-Miocene radiation of dry-adapted shrubs (Tremandraceae) from rainforest tree progenitors (Elaeocarpaceae) in Australia. *American Journal of Botany*, *93*, 1328–1342.

Crisp, M. D., & Weston, P. H. (1993). Geographic and ontogenetic variation in morphology of Australian waratahs (*Telopea*: Proteaceae). *Systematic Biology*, *42*, 49–76.

Cronquist, A. (1981). An Integrated System of Classification of Flowering Plants. New York: Columbia University Press.

Da Silva Meyer, A., Franco Garcia, A. A., Pereira de Souza, A., & Lopes de Souza, C. (2004). Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L.). *Genetics and Molecular Biology*, *27*, 83–91.

Davis, J. C. (2002). Analysis of Multivariate Data. In *Statistics and Data Analysis in Geology.* (Third Edit., pp. 461–594). New York: John Wiley & Sons.

Davison, A., & Chiba, S. (2003). Laboratory temperature variation is a previously unrecognized source of genotyping error during capillary electrophoresis. *Molecular Ecology Notes*, *3*, 321–323.

Degnan, J. H., & Rosenberg, N. a. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*, 332–340.

de Lange, P. J., Murray, B. G., & Datson, P. M. (2004). Contributions to a chromosome atlas of the New Zealand flora – 38 counts for 50 families. *New Zealand Journal of Botany*, *42*, 873–904.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*, 879–886.

Dent, A. E., & VonHoldt, B. M. (2012). STUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361.

Dettmann, M. E., & Clifford, H. T. (2001). The fossil record of *Elaeocarpus* L. fruits. *Memoirs of the Queensland Museum*, *46*, 461–497.

Devy, M. S., & Davidar, P. (2003). Pollination systems of trees in Kakachi, a Mid-elevation wet evergreen forest in Western Ghats, India. *American Journal of Botany*, *90*, 650–657.

Diniz-filho, J. A. F., Pires de Campos Telles, M., Bonatto, S. L., Eizirik, E., O. de Freitas, T. R., Santos, R., De Marco, P. Jr., Santos, F. R., Sole-Cava, A. Nascimento Soares, T. (2008). Mapping the evolutionary twilight zone: molecular markers, populations and geography. *Journal of Biogeography*, *35*, 753–763.

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*, 2233–2244.

Dolman, G., & Moritz, C. (2006). A multilocus perspective on refugial isolation and divergence in rainforest skinks (Carlia). *Evolution*, *60*, 573–582.

Domínguez Lozano, F., Moreno Saiz, J. C., Sainz Ollero, H., & Schwartz, M. W. (2007). Effects of dynamic taxonomy on rare species and conservation listing: insights from the Iberian vascular flora. *Biodiversity and Conservation, 16*, 4039–4050.

Don, G. (1831). *A General History of the Dichlamydeous Plants: comprising complete descriptions of the different orders; together with the characters of the genera and species, and an enumeration of the cultivated varieties; their places of growth, time of flowering, more of culture, and uses in medicine and domestic economy; scientific names accentuated, their etymologies explained, and the classes and orders illustrated by engravings, and a glossary of the terms used: the whole arranged according to the natural system.* vol. 1. London: J.G. and F. Rivington.

Donoghue, M. J. (1985). A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist, 88*, 172–181.

Downie, S. R., & Palmer, J. D. (1992). Use of chloroplast DNA rearrangements in reconstruction plant phylogeny. In Pamela S. Soltis, D. E. Soltis, & J. J. Doyle (Eds.), *Molecular Systematics of Plants* (pp. 14–35). New York, New York, USA: Chapman and Hall.

Drummond, A. J., Ho, S. Y. W., Rawlence, N., & Rambaut, A. (2007). BEAST-Bayesian evolutionary analysis sampling tree. Auckland, New Zealand Edinburgh, Scotland: The University of Auckland University of Edinburgh .

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214.

Duarte, J. M., dos Santos, J. B., & Melo, L. C. (1999). Comparison of similarity coefficients based on RAPD markers in the common bean. *Genetics and Molecular Biology, 22*, 427–432.

Elder, J. F., & Turner, B. J. (1995). Concerted evolution of repetitive DNA sequences in Eukaryotes. *The Quarterly Review of Biology*, *70*, 297–320.

Eriksson, T., Hibbs, M., & Yoder, A. (2003). The phylogeny of *Rosoideae* (Rosaceae) based on sequences of the internal transcribed spacers (*ITS*) of nuclear ribosomal DNA and the t*rnL/F* region of chloroplast DNA. *International Journal of Plant Sciences, 164*, 197–211.

Esselink, G. D., Nybom, H., & Vosman, B. (2004). Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting — peak ratios) method. *Theoretical and Applied Genetics*, *109*, 402–408.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*, 2611–20.

Excoffier, Laurent, Smouse, P. E., & Quattro, J. M. (1992). Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479–491.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–87.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, *7*, 574–578.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., & Kluge, A. (1996). Parsimony jackknifing outperforms neighbor-joining. *Cladistics*, *12*, 99–124.

Fatemi, M., Gross, C. L., & Bruhl, J. J. (2007). The first phenetic analysis of species limits in *Bertya* (Euphorbiaceae). *Australian Systematic Botany*, *20*, 448–463.

Fehlberg, S. D., & Ferguson, C. J. (2012). Intraspecific cytotypic variation and complicated genetic structure in the *Phlox amabilis-P. woodhousei* (Polemoniaceae) complex. *American Journal of Botany*, *99*, 865–874.

Feliner, G. N., & Rossello, J. A. (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution*, *44*, 911–919.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, *39*, 783–791.

Fitch, W. J. (1975). Toward finding the tree of maximum parsimony. In G. F. Estabrook (Ed.), *Proceedings of the Eighth International Conference on Numerical Taxonomy* (pp. 189–230). San Fransisco: W.H. Freeman.

Flanagan, N. S., Peakall, R., Clements, M. A., & Otero, J. T. (2006). Conservation of taxonomically difficult species: the case of the Australian orchid, Microtis angusii. *Conservation Genetics*, *7*, 847–859.

Flann, C. (2010). Morphometric study of *Euchiton* (Gnaphalieae: Asteraceae). *Australian Systematic Botany*, *23*, 285–305.

Freudenstein, J. V., & Davis, J. I. (2010). Branch support via resampling: an empirical study. *Cladistics*, *26*, 643–656.

Gabrielson, P. E., Miller, K. A., & Martone, P. T. (2011). Morphometric and molecular analyses confirm two distinct species of *Calliarthron* (Corallinales, Rhodophyta), a genus endemic to the northeast Pacific. *Phycologia*, *50*, 298–316.

Gamage, H. K., & Schmidt, S. (2009). Short Communication. A robust method for chromosome quantification and ploidy determination in woody species. *Australalian Journal of Botany*, *57*, 87–93.

Gibson, R., Conn, B. J., & Bruhl, J. J. (2012). Morphological evaluation of the *Drosera peltata* complex (Droseraceae). *Australian Systematic Botany*, *25*, 49–80.

Giribet, G., & Wheeler, W. C. (1999). On Gaps. *Molecular Phylogenetics and Evolution*, *13*, 132–143.

Glaw, F., Köhler, J., Townsend, T. M., & Vences, M. (2012). Rivaling the world's smallest reptiles: discovery of miniaturized and microendemic new species of leaf chameleons (Brookesia) from Northern Madagascar. *PLoS One*, *7*, 1–24.

Goldstein, D. G., & Pollock, D. D. (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, *88*, 335–342.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, *27*, 857–871.

Graham, C H, VanDerWal, J., Phillips, S. J., Morits, C., & Williams, S. E. (2010). Dynamic refugia and species persistence: tracking spatial shifts in habitat through time. *Ecography*, *33*, 1062–1069.

Graham, Catherine H, Moritz, C., & Williams, S. E. (2006). Habitat history improves prediction of biodiversity in rainforest fauna. *Proceeding National Academy of Science U.S.A.*, *103*, 632–636.

Grossi, M. A., Gutiérrez, D. G., Berrueta, P. C., & Martínez, J. J. (2011). *Acanthostyles* (Asteraceae, Eupatorieae): a revision with a multivariate analysis. *Australian Systematic Botany*, *24*, 87–103.

Guindon, S., & Gascuel, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, *52*, 696–704.

Guo, Y. P., Ehrendorfer, F., & Samuel, R. (2004). Phylogeny and systematics of *Achillea* (Asteraceae-Anthemideae) inferred from *nrITS* and plastid *trnL-F* DNA sequences. *Taxon*, *53*, 657–672.

Guymer, G. (1997). Elaeocarpaceae. In R. J. F. Henderson (Ed.), *Names and Distribution of Queensland Plants, Algae and Lichens* (p. 67). Brisbane: State of Queensland Environmental Protection Agency.

Guymer, G. (2002). Elaeocarpaceae. In R. J. F. Henderson (Ed.), *Names and Distribution of Queensland Plants, Algae and Lichens* (pp. 65–66). Brisbane: Desktop Publishing.

Guymer, G. (2007). Elaeocarpaceae. In P. D. Bostock & A. E. Holland (Eds.), *Census of the Queensland Flora* (p. 66). Brisbane: Queensland Herbarium, Biodiversity Science, Environmental Sciences Division, Environmental Protection Agency.

Guymer, G. (2010). Elaeocarpaceae. In P. D. Bostock & A. E. Holland (Eds.), *Census of the Queensland Flora* (pp. 61–62). Brisbane: Queensland Herbarium, Biodiversity and Ecosystem Science, Department of Environment and Resource Management.

Halverson, K., Heard, S. B., Nason, J. D., & Stireman, 3rd John O. (2008). Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago alitissima* (Asteraceae). *American Journal of Botany*, *95*, 50–58.

Harata, T., Nanami, S., Yamakura, T., Matsuyama, S., Chong, L., Diway, B. M., Tan, S., Itoh, A. (2012). Fine-scale spatial genetic structure of ten dipterocarp tree species in a Bornean rain forest. Biotropica, *44*, 586–594.

Hardy, O. J., & Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes, 2*, 618–620.

Harpke, D., & Peterson, A. (2008a). Extensive 5.8S nrDNA polymorphism in Mammillaria (Cactaceae) with special reference to the identification of pseudogenic internal transcribed spacer regions. *Journal of Plant Research*, *121*, 261–270.

Harpke, D., & Peterson, A. (2008b). 5.8S motifs for the identification of pseudogenic ITS regions. *Botany*, *86*, 300–305.

Hartman, S. E. (1988). Evaluation of some alternative procedures used in numerical systematics. *Systematic Zoology*, *37*, 1–18.

Hedrick, P. W. (2000). *Genetics of Populations* (2nd ed.). Boston: Jones and Bartlett.

Henderson, A., & Ferreira, E. (2002). A morphometric study of Synechanthus (Palmae). *Systematic Botany*, *27*, 693–702.

Hill, T., & Lewicki, P. (2006). Discriminant function analysis. In *Statistics: Methods and Applications. A comprehensive reference for science, industry, and data mining* (pp. 155–134). Tulsa, USA: StatSoft, Inc.

Hills, D. M., & Dixon, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology*, *66*, 411–453.

Horn, J. W., van Ee, B. W., Morawetz, J. J., Riina, R., Steinmann, V. W., Berry, P. E., Wurdack, K. J. (2012). Phylogenetics and the evolution of major structural characters in the giant genus *Euphorbia* L. (Euphorbiaceae), *Molecular Phylogenetics and Evolution*, 63, 305–326.

Hughes, C. E., Eastwood, R. J., & Bailey, C. D. (2006). From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, *361*, 211–225.

Hyland, B. P. M., & Whiffin, T. (1993). *Australian Tropical Rainforest Trees - An interactive Identification System. Vol 2.* (p. 107). East Melbourne, Victoria: CSIRO Australia.

Hyland, B. P. M., Whiffin, T., Christophel, D. C., Gray, B., & Elick, R. W. (2003). *Australian Tropical Rain Forest Plants - Trees, Shrubs and Vines.* Collingwood, Victoria: CSIRO Publishing.

IUCN (2001). IUCN Red List Categories and Criteria: Version 3.1. In. IUCN Species Survival Commission, IUCN, Gland, Switzerland and Cambrige, UK.

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles*, *37*, 547–579.

Jackson, D. A., Somers, K. M., & Harvey, H. H. (1989). Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence? *American Society of Naturalist*, *133*, 436–453.

James, C. H., & Moritz, C. (2000). Intraspecific phylogeography in the sedge frog *Litoria fallax* (Hylidae) indicates pre-Pleistocene vicariance of an open forest species from eastern Australia. *Molecular Ecology*, *9*, 349–358.

James, F. C., & McCulloch, C. E. (1990). Multivariate analysis in ecology and systematics: Panacea or Pandora's box? *Annual Review of Ecology and Systematics*, *21*, 129–166.

Jobes, D. V, & Thien, L. B. (1997). A conserved motif in the 5.8S Ribosomal RNA (rRNA) gene is a usful diagnostic marker for plant internal transcribed spacer (ITS) sequences. Plant Molecular Biology Reporter, 15, 326–334.

Jones, R. C., McNally, J., & Rossetto, M. (2002). Isolation of microsatellite loci from a rainforest tree, *Elaeocarpus grandis* (Elaeocarpaceae), and amplification across closely related taxa. *Molecular Ecology Notes*, *2*, 179–181.

Jørgensen, M. H., Carlsen, T., Skrede, I., & Elven, R. (2008). Microsatellites resolve the taxonomy of the polyploid *Cardamine digitata* aggregate (Brassicaceae). *Taxon*, *57*, 882–892.

Joseph, L., & Moritz, C. (1993). Phylogeny and historical aspects of the ecology of eastern Australian scrubwrens Seuicomis spp. - evidence from mitochondrial DNA. *Molecular Ecology*, *2*, 161–170.

Joseph, L., & Moritz, C. (1994). Mitochondrial DNA Phylogeography of birds in eastern Australian rainforests: first fragments. *Australian Journal of Zoology*, *42*, 385–403.

Judd, W. S., Campbell, C. S., Kellogg, E. A., Stevens, P. F., & Donoghue, M. J. (2007). *Plant Systematics: A Phylogenetic Appoach-3rd edition* (pp. 103–118). Massachusetts: Sinauer Associates inc.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36.

Katoh, K., & Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics (Oxford, England), 26, 1899–1900.

Kawakita, A., Sota, T., Ascher, J. S., Ito, M., Tanaka, H., & Kato, M. (2003). Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in Bumble Bees (*Bombus*). Molecular Biology and Evolution, 20, 87–92.

Khan, M. L., Bhuyan, P., & Tripathi, R. S. (2003). Regeneration of Rudraksh (*Elaeocarpus ganitrus* Roxb.)- a threatened tree species: germination strategies. *International Journal of Ecology and Environmental Sciences*, *29*, 255–260.

Kim, K. C., & Byrne, L. B. (2006). Biodiversity loss and the taxonomic bottleneck: emerging biodiversity science. *Ecological Research*, *21*, 794–810.

Kloda, J. M., Dean, P. D. G., Maddren, C., Macdonald, D. W., & Mayes, S. (2008). Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes : an example from British Restharrows (*Ononis spinosa* and *Ononis repens*). *Heredity*, *100*, 253–260.

Larget, B., & Simon, D. L. (1999). Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution*, 16, 750–759.

Lattig, P., San Martín, G., & Martin, D. (2007). Taxonomic and morphometric analyses of the *Haplosyllis spongicola* complex (Polychaeta: Syllidae: Syllinae) from Spanish seas, with re-description of the type species and descriptions of two new species. *Scientia Marina*, *71*, 551–570.

Lewis, P. O., Holder, M. T., & Holsinger, K. E. (2005). Polytomies and Bayesian phylogenetic inference. *Systematic Biology*, *54*, 241–253.

Lihová, J., Kudoh, H., & Marhold, K. (2010). Morphometric studies of polyploid *Cardamine* species (Brassicaceae) from Japan: solving a long-standing taxonomic and nomenclatural controversy. *Australian Systematic Botany*, *23*, 94–111.

Liu, J. S., & Schardl, C. L. (1994). A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. Plant Molecular Biology, 26, 775–778.

Lowe, A., Harris, S., & Ashton, P. (2004). Markers and sampling in ecological genetics. In *Ecological genetics: design, analysis, and application* (pp. 6–45). Victoria: Blackwell Publishing.

Manly, B. J. F. (1991). *Randomizatiom and Monte Carlo Methods in Biology*. London: Chapman and Hall.

Manos, P. S., Zhou, Z. ‐ K., & Cannon, C. H. (2001). Systematics of Fagaceae: phylogenetic tests of reproductive trait evolution. *International Journal of Plant Sciences*, 162, 1361- 1379.

Masters, M. T. (1874). *Elaeocarpus*. In J. D. Hooker (Ed.), *Flora of British India* (pp. 400–408). Covent Garden, London: Savil, Edward and Co.

Maynard, D., Crayn, D., Rossetto, M., Kooyman, R., & Coode, M. (2008). *Elaeocarpus sedentarius* sp. nov. (Elaeocarpaceae) - morphometric analysis of a new, rare species from eastern Australia. *Australian Systematic Botany*, *21*, 192–200.

Maynard, D. J. (2004). *A molecular phylogeny for Australian Elaeocarpus (Elaeocarpaceae) and the affinities of a putative new taxon*. Honors thesis. The University of New South Wales.

Mayr, E. (1942). *Systematics and the origin of species*. New York: Columbia University Press.

Mayr, E. (1992). A local flora and the biological species concept. *American Journal of Botany*, *7*, 222–238.

McCauley, D. E. (1995). The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution*, *10*, 198–202.

McDade, L., Daniel, T. F., & Kiel, C. (2008). Toward a comprehensive understanding of phylogenetic relationships among lineages of Acanthaceae s.l. (Lamiales). *American Journal of Botany*, *95*, 1136–1152.

McNeill, J., Barrie, F.R., Buck, W.R., Demoulin, V., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Marhold, K., Prado, J., Prud'homme van Reine, W.F., Smith, G.F. & Wiersema, J.H. (2012) International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). Koeltz Scientific Books.

McPherson, H. (2008). *Phylogenetics and evolutionary dynamics of Tetratheca (Elaeocarpaceae)*. The University of New England.

Mehara, P. N., & Sareen, T. S. (1973). Cytology of some Himalayan trees. *Silvae Genetica*, *22*, 66–70.

Merrill, E. D. (1951). Notes on *Elaeocarpus* Linnaeus. *Journal of the Arnold Arboretum*, *32*, 157– 200.

Miller, H. S. (1970). The herbarium of Aylmer Bourke Lambert: notes on its acquisition, dispersal, and present whereabouts. *Taxon*, *19*(4), 489–553.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE) (pp. 1–8). New Orleans, LA.

Milligan, B. G. (1992). Is organelle DNA strictly maternally inherited? Power analysis of a binominal distribution. *American Journal of Botany*, *79*, 1325–1328.

Mishler, B. D., & Brandon, R. N. (1987). Individuality, pluralism, and the phylogenetic species concept. *Biology and Philosophy*, *2*, 397–414.

Morgante, M., & Olivieri, A. M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*, *3*, 175–182.

Morin, P. A., Martien, K. K., Archer, F. I., Cipriano, F., Steel, D., Jackson, J., & Taylor, B. L. (2010). Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. *The Journal of Heredity*, *101*, 1–10.

Moritz, C., Joseph, L., Cunningham, M., & Schneider, C. (1997). Molecular Perspectives on Historical Fragmentation of Australian Tropical and Subtropical Rainforests: Implications for Conservation. In W. F. Laurance & R. O. Bierregaard (Eds.), *Tropical Forest Remnants* (pp. 442–454). Chicago: The University of Chicago Press.

Morrison, D. A., & Weston, P. H. (1985). Analysis of morphological variation in a field sample of *Caladenia catenata* (Smith) Druce (Orchidaceae). *Australian Journal of Botany*, *33*, 185–195.

Müller, K. (2005). SeqState: primer design and sequence statistics for phylogenetic DNA datasets. Applied Bioinformatics, 4, 65–9.

Münzbergová, Z., Surinová, M., & Castro, S. (2013). Absence of gene flow between diploids and hexaploids of *Aster amellus* at multiple spatial scales. *Heredity*, *110*, 123–30.

Nakamoto, A., Kinjo, K., & Izawa, M. (2009). The role of Orii's flying-fox (*Pteropus dasymallus inopinatus*) as a pollinator and a seed disperser on Okinawa-jima Island, the Ryukyu Archipelago, Japan. *Ecological Research*, *24*, 405–414.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*, 583–590.

Nix, H. A. (1991). Biogeography: Patterns and process. In H. A. Nix & M. Switzer (Eds.), *Rainforest animals: Atlas of vertebrates endemic to Australia's Wet Tropics* (pp. 11–39). Canberra: Australian Nature Conservation Agency.

Nixon, K. C., & Wheeler, Q. D. (1990). An amplification of the phylogenetic species concept. *Cladistics*, *6*, 211–223.

Nürk, N. M., Madriñán, S., Carine, M. A., Chase, M. W., Blattner, F. R. (2013). Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). *Molecular Phylogenetics and Evolution*, 66, 1–16.

Oakley, T. H. and Cunningham C. W. (2002). Molecular phylogenetic evidence for the independent evolutionary origin of an arthropod compound eye. *Proceeding National Academy of Science U.S.A.*, 99, 1426–1430.

Obbard, D. J., Harris, S. A., & Pannell, J. R. (2006). Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Heredity*, *97*, 296–303.

Ogden, T. H., & Rosenberg, M. S. (2007). Alignment and topological accuracy of the direct optimization approach via POY and traditional phylogenetics via ClustalW + PAUP *. Systematic Biology, 56(2), 182–193.

Ono, M. (1975). Chromosome numbers of some endemic species of the Bonin Islands 1. *Botanical Magazine Tokyo*, *88*, 323–328.

Páez-Moscoso, D. J., & Guayasamin, J. M. (2012). Species limits in the Andean toad genus *Osornophryne* (Bufonidae). *Molecular Phylogenetics and Evolution*, *65*, 805–822.

Palmer, J. D. (1985). Evolution of chloroplast and mitochondrial genome evolution in plant and algae. In R. J. MacIntyre (Ed.), *Molecular Evolutionary Genetics* (pp. 99–133). London: Plenum Press.

Palop-Esteban, M., Segarra-Moragues, J. G., & González-Candelas, F. (2011). Polyploid origin, genetic diversity and population structure in the tetraploid sea lavender *Limonium narbonense* Miller (Plumbaginaceae) from eastern Spain. *Genetica*, *139*, 1309–22.

Paterson, H. E. H. (1985). The recognition species concept. In E. S. Vrba (Ed.), *Species and Speciation* (pp. 21–29). Pretoria: Transvaal Museum.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295.

Peakall, Rod, & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. population genetic software for teaching and research — an update. *Bioinformatics*, *28*, 2537–2539.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, *2*, 559–572.

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 187–214.

Piotti, A., Leonardi, S., Buiteveld, J., Geburek, T., Gerber, S., Kramer, K., Vettori, C., Vendramin, G. G. (2012). Comparison of pollen gene flow among four European beech (*Fagus sylvatica* L.) populations characterized by different management regimes. *Heredity*, *108*, 322–31.

Pirie, M. D., Balcázar Vargas, M. P., Botermans, M., Bakker, F. T., & Chatrou, L. W. (2007). Ancient paralogy in the cpDNA *trnL-F* region in Annonaceae: implications for plant molecular systematics. *American Journal of Botany*, *94*, 1003–1016.

Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, *6*(November), 847–859.

Pons, O., & Petit, R. J. (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, *144*, 1237–1245.

Posada, D. (2008). jModelTest: phylogenetic model averaging. Molecular Biology and Evolution, 25, 1253–1256.

Premathilake, R., & Nilsson, S. (2001). Pollen morphology of endemic species of the Pollen morphology of endemic species of the Horton Plains National Park, Sri Lanka. *Grana*, *40*, 256–279.

Price, O. F. (2006). Movements of frugivorous birds among fragmented rainforests in the Northern Territory, Australia. *Wildelife Research*, *33*, 521–528.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–59.

Pruett, C. L., & Winker, K. (2008). The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. *Journal of Avian Biology*, *39*, 252–256.

Queensland Herbarium. (2013). Regional Ecosystem Description Database (REDD). Version 6.1., viewed 3 March 2013, <http://www.ehp.qld.gov.au/ecosystems/biodiversity/regional-ecosystems/conditions_of_use.html>

Raven, P. H., & Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden*, *61*, 539–673.

Ridley, H.N. 1992. *Elaeocarpus*. in, The Flora of the Malay Peninsula. pp 309-321. L. Reeve and Co. Ltd., London.

Rieseberg, L. H., & Willis, J. H. (2007). Plant speciation. *Science (New York, N.Y.)*, *317*, 910–914.

Rohlf, F. J. (1990). Morphometics. *Annual Review of Ecology and Systematics*, *21*, 299–316.

Rohlf, F. J., & Marcus, L. F. (1993). A revolution in morphometrics. *Trends in Ecology and Evolution*, *8*, 129–132.

Rohlf, J. F. (1972). An empirical comparison of three ordination techniques in numerical taxonomy. *Systematic Biology*, *21*, 271–280.

Ronfort, J., Jenczewski, E., Bataillon, T., & Rousset, F. (1998). Analysis of population structure in autotetraploid species. *Genetics*, *150*, 921–30.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*, 539–542.

Rossetto, M., Crayn, D., Ford, A., Mellick, R., & Sommerville, K. (2009). The influence of environment and life-history traits on the distribution of genes and individuals: a comparative study of 11 rainforest trees. *Molecular Ecology*, *18*, 1422–1438.

Rossetto, M., Crayn, D., Ford, A., Ridgeway, P., & Rymer, P. (2007). The comparative study of range-wide genetic structure across related, co-distributed rainforest trees reveals contrasting evolutionary histories. *Australian Journal Botany*, *55*, 416–424.

Rossetto, M., Gross, C. L., Jones, R., & Hunter, J. (2004). The impact of clonality on an endangered tree (*Elaeocarpus williamsianus*) in a fragmented rainforest. *Biological Conservation*, *117*, 33–39.

Rossetto, M., Jones, R., & Hunter, J. (2004). Genetic effects of rainforest fragmentation in an early successional tree (*Elaeocarpus grandis*). *Heredity*, *93*, 610–8.

Rossetto, M., Kooyman, R., Sherwin, W., & Jones, R. (2008). Dispersal limitations, rather than bottlenecks or habitat specificity, can restrict the distribution of rare and endemic rainforest trees. *American Journal of Botany*, *95*, 321–9.

Roubik, D. W. (2005). Honeybees in Borneo. In D. W. Roubik, S. Sakai, & H. A. Abdul (Eds.), *Pollination ecology and rainforest-Sarawak study* (pp. 89–102). New York: Springer.

Rozefelds, A.C. & Christophel, D.C. (1996). *Elaeocarpus* (Elaeocarpaceae) endocarps from the Early to Middle Miocene Yallourn formation of Eastern Australia. *Muelleria*. 9: 229 – 237

Sampson, J. F., & Byrne, M. (2012). Genetic diversity and multiple origins of polyploid *Atriplex* nummularia Lindl. (Chenopodiaceae). *Biological Journal of the Linnean Society*, *105*, 218–230.

Sang, T., Crawford, D. J., & Stuessy, T. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, *84*, 1120–1136.

Saraswati, P. K., & Sabnis, S. V. (2006). Comparison of CART and discriminant analysis of morphometric data in Foraminiferal taxonomy. *Anuário do Instituto de Geociências*, *29*, 153–162.

Schlötterer, C., & Tautz, D. (1994). Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Current Biology*, *4*, 777–783.

Schlötterer, C. (2004). The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics*, *5*, 63-69.

Schlülter, P. M., & Harris, S. A. (2006). Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes*, *6*, 569–572.

Schönswetter, P., Solstad, H., Escobar Garcia, P., & Eleven, R. (2009). A combined molecular and morphological approach to the taxonomically intricate European mountain plant *Papaver alpinum* s.l. (Papaveraceae) - Taxa or informal phylogeographical groups? *Taxon*, *58*, 1326–1343.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments-A poor man's approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology*, *18*, 233–234.

Schulte, K., Silvestro, D., Kiehlmann, E., Vesely, S., Novoa, P., & Zizka, G. (2010). Detection of recent hybridization between sympatric Chilean *Puya* species (Bromeliaceae) using AFLP markers and reconstruction of complex relationships. *Molecular Phylogenetics and Evolution*, *57*, 1105–1119.

Schumann, K. (1890). Elaeocarpaceae. In A. Engler & K. Prantl (Eds), *Die Natürlichen Pflanzenfamilien: nebst Ihren Gattungen und wichtigeren Arten, insbesondere den Nutzpflanzen* (Vol. 3, pp. 1-8). Leipzig: W. Engelmann.

Shaffer, H. B., & Thomson, R. C. (2007). Delimiting species in recent radiations. *Systematic Biology*, *56*, 896–906.

Shaw, J., Lickey, E. B., Schilling, E. E., & Small, R. L. (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. *American Journal of Botany*, *94*, 275–288.

Simmons, M. P., & Freudenstein, J. V. (2011). Spurious 99 % bootstrap and jackknife support for unsupported clades. Molecular Phylogenetics and Evolution, 61, 177–191.

Simmons, M. P., Pickett, K. M., & Miya, M. (2004). How meaningful are Bayesian support values? Molecular Biology and Evolution, 21, 188–99.

Simmons, M. P., Müller, K., & Norton, A. P. (2007). The relative performance of indel-coding methods in simulations. *Molecular Phylogenetics and Evolution*, *44*, 724–740.

Simmons, M. P., Ochoterena, H., & Carr, T. G. (2001). Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Systematic Biology*, *50*, 454–462.

Sites, J. W., & Marshall, J. C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution*, *18*, 462–470.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, *139*, 457–462.

Small, R., Cronn, R. C., & Wendel, J. F. (2004). L. A. S. Johnson review No. 2. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany*, *17*, 145–170.

Small, R. L., Lickey, E. B., Shaw, J., & Hauk, W. D. (2005). Amplification of noncoding chloroplast DNA for phylogenetic studies in lycophytes and monilophytes with a comparative example of relative phylogenetic utility from Ophioglossaceae. *Molecular Phylogenetics and Evolution*, *36*, 509–522.

Smith, A. C. (1944). Studies of Papuasian plants, VI. *Journal of the Arnold Arboretum*, *25*, 222–270.

Smith, A. C. (1953). Studies of Pacific Island plants. XV. The genus *Elaeocarpus* in the New Herbrides, Fiji, Samoa and Tonga. *Contributions from the United States National Herbarium*, *30*, 523–575.

Smith, A. C. (1981). Elaeocarpaceae. In *Flora Vitiensis Nova: a new Flora of Fiji (spermatophytes only). Vo.2.* (pp. 349–365). Lawai, Kauai, Hawaii: Pacific Tropical Botanical Garden.

Smith, J. E. (1809). *Elaeocarpus*. In A. Rees (Ed.), *The Cyclopædia; or, Universal Dictionary of Arts, Sciences*. London, United Kingdom.

Smith, S. A., & Donoghue, M. (2008). Rates of molecular evolution are linked to life history in flowering plants. *Science (New York, N.Y.)*, *322*, 86–89.

Sneath, P., & Sokal, R. (1973). *Numerical Taxonomy: The Principles and Practice of Numerical Classification. The Principles and Practice of Numerical Classification*. San Fransisco: Freeman.

Soejima, A., & Wen, J. (2006). Phylogenetic analysis of the Grape family (Vitaceae) based on three chloroplast markers. *American Journal of Botany*, *93*, 278–287.

Sokal, R. R., & Crovello, T. (1970). The biological species concept: a critical evaluation. *The American Naturalist*, *104*, 127–153.

Španiel, S., Marhold, K., Filová, B., & Zozomová-Lihová, J. (2011). Genetic and morphological variation in the diploid – polyploid *Alyssum montanum* in Central Europe: taxonomic and evolutionary considerations. *Plant Systematics and Evolution*, *294*, 1–25.

Spencer, M., Susko, E., & Roger, A. J. (2005). Likelihood, parsimony, and heterogeneous evolution. *Molecular Biology and Evolution*, *22*, 1161–1164.

Stefanović, S., Lakušić, D., Kuzmina, M., Međedović, S., Tan, K., & Stevanović, V. (2008). Molecular phylogeny of *Edraianthus* (Grassy Bells; Campanulaceae) based on non-coding plastid DNA sequences. *Taxon, 57*, 452–475.

Stevens, P. F. (2012). Angiosperm Phylogeny Website. Version 12., viewed 1 March 2013 < http://www.mobot.org/MOBOT/research/APweb/>

Stocker, G. C., & Irvine, A. K. (1983). Seed dispersal by Cassowaries (*Casuarius casuarius*) in North Queensland's rainforests. *Biotropica, 15*, 170–176.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B, 36:111–147.

Storr, G. M. (1973). List of Queensland birds. *Western Australia Museum Special Publication No. 4*, 90.

Storr, G. M. (1977). Birds of the Northern Territory. *Western Australia Museum*.

Stuessy, T. F. (2009). The Phenetic Approach. In *Plant Taxonomy - The Systematic Evaluation of Comparative Data* (2nd ed., pp. 51–68). New York: Columbia University Press.

Sugiura, M. (1989). The chloroplast chromosomes in land plants. *Annual Review of Cell Biology*, *5*, 51–70.

Sun, Y., Skinner, D. Z., Liang, G. H., & Hulbert, S. H. (1994). Phylogenetic analysis of Sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics*, *89*, 26–32.

Suzuki, Y., Glazko, G. V, & Nei, M. (2002). Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceeding National Academy of Science U.S.A.*, *99*, 16138–16143.

Swofford, D. L. (2002). PAUP* Phylogemetic Anakysis Using Parsimony (*and Other Methods). Sanderland, Massachusetts: Sinauer Associates.

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, *17*, 1105–1109.

Takhtajan, A. (1997). *Diversity and Classification of Flowering Plants.* New York: Columbia University Press.

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). Classification: basic concepts, decision trees, and model evaluation. In *Introduction to Data Mining* (pp. 145–205). Boston: Addison-Wesley.

Tang, Y., & Phengklai, C. (2008). Elaeocarpaceae. In Z. Wu, P. H. Raven, & D. Hong (Eds.), *Flora of China Vol. 11* (pp. 223–239). Beijing and St. Louis: Science Press and Missouri Botanical Garden.

Tate, J., & Simpson, B. B. (2003). Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany*, *28*, 723–737.

Thackway, R. & Cresswell, I. (1995). An Interim Biogeographic Regionalisation for Australia: a framework for setting priorities in the National Reserves System Cooperative Program Version 4, Australian Nature Conservation Agency, Canberra.

Thrall, P. H., & Young, A. (2000). AUTOTET: a program for analysis of autotetraploid genotypic data. *The Journal of Heredity*, *91*, 348–349.

Thurlby, K. A. G., Wilson, P. G., Sherwin, W. B., Connelly, C., & Rossetto, M. (2012). Reproductive bet-hedging in a rare yet widespread rainforest tree, *Syzygium paniculatum* (Myrtaceae). *Austral Ecology*, *37*, 936–944.

Tirel, C. (1982). Elaeocarpaceae. In A. Aubreville & J.-F. Leroy (Eds.), *Flore De La Nouvelle-Caledonie et Dependances* (pp. 3–124). Paris.

Tirel, C. (1983). Une nouvelle espéce *d'Elaeocarpus* (Elaeocarpaceae) de Nouvelle-Calédonie. *Bulletin du Museum National d' Histoire Naturelle Section B, Adansonia, no 3, botanique, phytochimie*, *5*, 283–287.

Tirel, C. (1985). Éléocarpacées. In H. Humbert & J.-F. Leroy (Eds.), *Flore de Madagascar et des Comores: plantes vasculaires / pubiliée sous les auspices du gouvernement général de Madagascar et sous la direction de H.Humbert.* (pp. 1–116). Paris: Muséum National D'Historire Naturelle.

VanDerWal, J., Shoo, L. P., & Williams, S. E. (2009). New approaches to understanding late Quaternary climate fluctuations and refugial dynamics in Australian wet tropical rain forests. *Journal of Biogeography*, *36*, 291–301.

Van Puyvelde, K., Van Geert, A., & Triest, L. (2010). ATETRA, a new software program to analyse tetraploid microsatellite data: comparison with TETRA and TETRASAT. *Molecular Ecology Resources*, *10*, 331–334.

Victoriano, P., Coronado, T., & Ortiz, J. (2010). A multivariate analysis of taxonomic limits in *Diplolaemus* Bell 1843. *Gayana*, *74*, 23–36.

Wang, A., Yang, M., & Liu, J. (2005). Molecular phylogeny, recent radiation and evolution of gross morphology of the rhubarb genus *Rheum* (Polygonaceae) inferred from chloroplast DNA *trnL-F* sequences. *Annals of Botany*, *96*, 489–98.

Webb, L. J., & Tracy, J. G. (1981). Australian rainforests: Patterns and change. In A. Keast (Ed.), *Ecological biogeography of Australia* (pp. 605–695). The Hague, Netherlands: Dr. W. Junk.

Weber, E. (1994). *The reproductive biology of Elaeocarpus largiflorens C. White subsp. largiflorens (Elaeocarpaceae)*. Honors thesis. James Cook University.

Weibel, R. (1968). Morphologie de l'embryon et de la graine des *Elaeocarpus*. *Candollea*, *23*, 101–108.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. A. Innis (Ed.), *PCR protocols a guide to methods and applications* (pp. 315–322). Academic Press.

Whittle, C.-A., & Johnston, M. O. (2003). Broad-scale analysis contradicts the theory that generation time affects molecular evolutionary rates in plants. *Journal of Molecular Evolution*, *56*, 223–33.

Wiley, E. O. (1978). The evolutionary species concept reconsidered. *Systematic Zoology*, *27*, 17–26.

Wood, T. E., & Nakazato, T. (2009). Investigating species boundaries in the Giliopsis group of Ipomopsis (Polemoniaceae): strong discordance among molecular and morphological markers. *American Journal of Botany*, *96*, 853–861.

Wurdack, K. J., Hoffmann, P., & Chase, M. (2005). Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid *rbcL* and *trnL-F* DNA sequences. *American Journal of Botany*, *92*, 1397–1420.

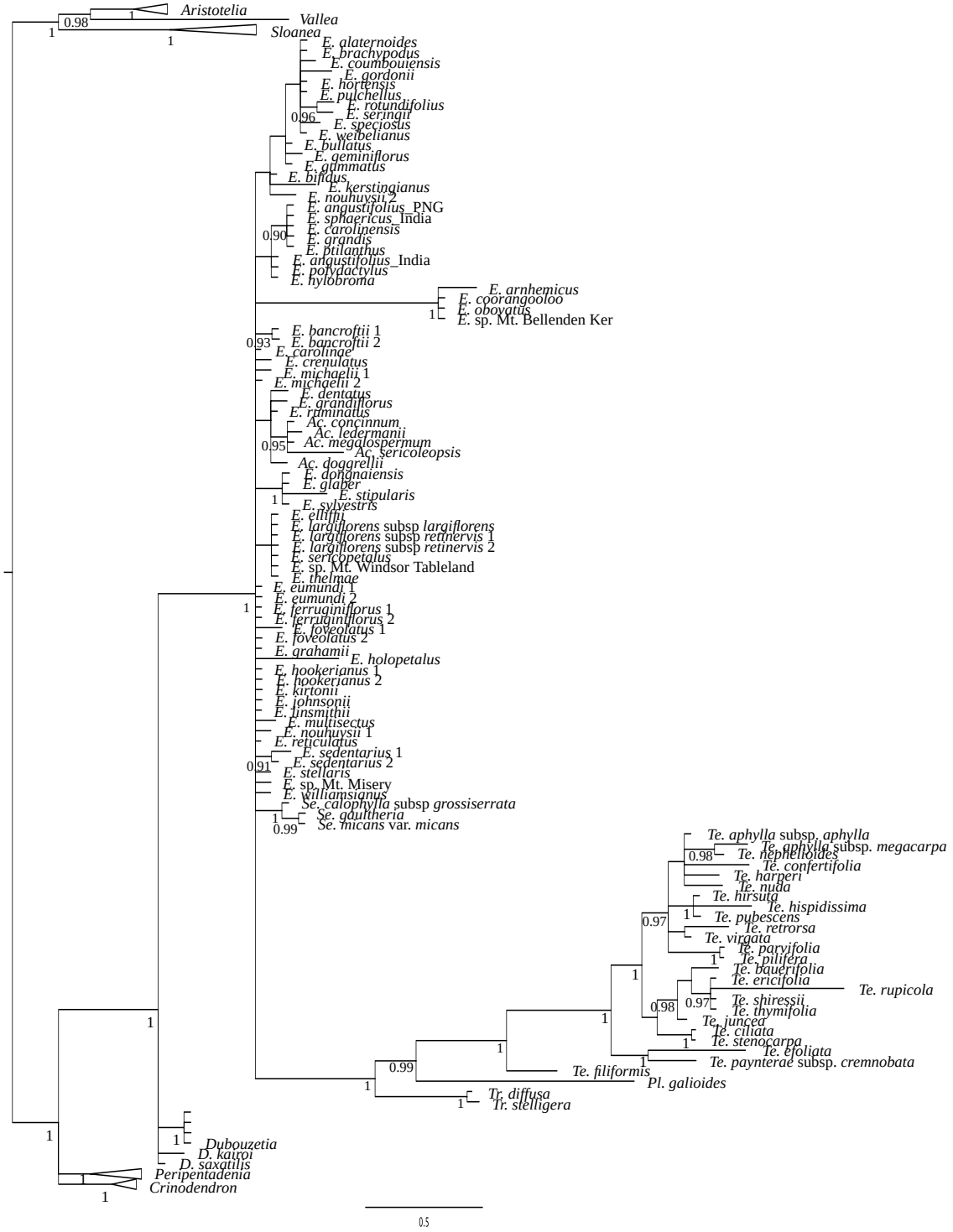
Yang, Z., & Rannala, B. (2005). Branch-length prior influences Bayesian posterior probability of phylogeny. *Systematic Biology*, *54*, 455–470.

Yang, Z., & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, *13*, 303–314.

Zander, R. H. (2004). Minimal values for reliability of Bootstrap and Jackknife proportions, decay index, and Bayesian Posterior Probability. Phyloinformatics, *2*, 1–13.

Zmarzty, S. (2001). Revision of *Elaeocarpus* (Elaeocarpaceae) section Elaeocarpus in Southern India and Sri Lanka. *Kew Bulletin*, *56*, 405–447.

Appendices



Appendix 1. Bayesian 50% majority rule consensus phylogram of *trnL-F* data set. Bayesian posterior probabilities are shown below the branches. See Table 2.3 for abbreviation for the genera.

Appendix 2. Microsatellite Alleles. '-9' indicates missing data. See table 3.1. for the population codes.

Samples	*Scu01Eg*			*Scu20Eg*			*Scu22Eg*				*Scu25Eg*			*Scu31Eg*				*Scu33Eg*			
ACO1	318	320		164	172	178	136	140	142		319	322		372	380	382		309	313		
ACO2	316	318	320	164	172	178	124	134	136	142	319			380	382	390		309	313		
ACO3	318	320		164	172	176	136	140	142		322	325		372	376	380		309	313		
ACO4	316	318	320	164			122	140	142		-9			376	380			309	311	315	
ACO5	318	320		164	172		122	140			-9			376	380	390		309	311		
ALF1	318			164			136	140	142		-9			380				305	309	319	
ALF2	318	320		164			140	142			-9			372	376	380	382	311	313		
ALF3	318	320		164	172	176	136	138	140	142	-9			372	380			313	321		
ALF4	318			164			136	138	140	142	319			380	382			305	311	313	319
ALF5	318	320		164			122	142			319			372	380			309	311	313	319
ALF6	318	320		164			-9				319	322	325	376	380			305	309	313	
ALF7	318	320		164	172	182	124	140	142		319	322	325	372	380			305	311	319	
ALF8	316	318	320	164	172		120	140	142		319			372	380	382		309	311		
ALF9	318	320	326	164	168	174	140	142			319	322		376	380			305	309	311	313
ALF10	316	318		164	168		120	140	142		319			376	380			305	309	313	315
ALF11	318	320		164	172		122	140	142		-9			376	380			309	313		
ALF12	318	320	334	164	172		136	140	142		319	322		376	380			305	311	313	319
ALF13	316	318	320	164			132	136	140		319	325		372	376	380	382	309	311		
ALF14	318	320		164			122	142			319	322	325	372	376	380		305	313	319	
ALF15	316	318		164	172		136	142			319	325		380				305	309		
ALF16	316	318	320	164	172		136	140	142		319	322		376	380			305	309	313	

Samples	*Scu01Eg*			*Scu20Eg*				*Scu22Eg*			*Scu25Eg*			*Scu31Eg*			*Scu33Eg*			
ALF17	316	318	320	164	172			120	136	142	319	325		380	382		305	309	311	313
AIR1	316	318	320	172	180			122	140	142	319	322		380			309	313	317	
AIR2	316	318	320	172	180			122	140	142	319	322		380			309	313	317	
AIR3	316	318	320	172	180			122	140	142	319	322		380			309	313	317	
APP1	316	318		164	172			122	140		319	325		376	380	382	309	311	313	
APP2	318			164	166	172	176	140	142		319	322	325	376	380		305	309	313	319
APP3	318			164	168	172	174	140	142	144	319	322	325	376	380		311	313		
APP4	316	318		164				140	142		319			376	380		311	313	319	321
APP5	318			164				140	142		319	322	334	372	376	384	309	311	313	321
APP6	318	320		164	172	180		122	140	142	319	325		376	380	384	309	311		
APP7	318	320		164	172	180		122	140	142	319	325		376	380	384	309	311		
APP8	318			164	172			140	142		319	322	334	376	380	382	311	313	319	321
APP9	316	318	320	164	172			122	136	142	319	334		376	380	382	305	309	311	
ANPA1	316	318		164				140	142		319			380	384		305	309	317	
ANPA2	318	320		164	172			122	136	144	319	322		376	380	382	309	311	321	
ANPA3	318	320		164	172			136	142		319			380	382		311	315	321	
ANPA4	318			164	172			136	140		319	322		380			305	309	311	
ANPA5	316	318	320	160	166	168	172	124	140		319	322		376	380		311	313	315	
ANPA6	318			164	172			122	124	140	319	325		376	380	382	309	311	315	
ANPA7	316	318		164	172			120	122	140	319	322		376	380	382	309	311	315	
ANPA8	318			164	172			136	142		319	322		376	382		309	311	315	
ANPA9	318			164				122	140	142	319	340		380	382		309	311	313	

Samples	*Scu01Eg*				*Scu20Eg*				*Scu22Eg*				*Scu25Eg*				*Scu31Eg*				*Scu33Eg*			
ANPA10	320				172				122				319	322			376				309			
ABS1	316	318			164	172	176		140	142			319				376	380	382		309	315	321	
ABS2	318				164	168	172		122	140			319				376	380	382		305	309	315	321
ABS3	316	318	326		164	168	172		140	142			319	322			376	380	382		311			
ABS4	318				164	168	176		136	140			319	322			376	382			307	309	311	
ABS5	318				-9				122	132	140		319	322	337		380	382			305	311	317	
ABS6	316	318	332		164	168	172		140	142			319	325	340		376	380			313	315	323	
ABD1	316	318	320	334	164	168	172		124	140	142		313	319			376	380	382		305	311	313	
ABD2	318				160	164	172	176	122	126	142		313	319			376	380	382		311	313	317	
ABD3	318	334			160	164	172	176	126	140	142		313	319	325		376	380	382		311	313	317	
ABD4	318				164	168	172	176	122	126	132	142	319	322	328		376				305	309	311	
ABD5	316	318	320	334	164	172			126	140	142		319	322			380	382	388		309	311	313	
ABD6	316	318			168	172			122	140			319	322	334		378	380	382		305	311	313	321
ABD7	316	318			164	168	170	172	122	140			319	322	334		376	380	382		305	311	313	321
ABD8	318				164	170	172		122	140	142		313	319			376	380			309	311		
AMK1	318	334			164	172	178		122	136	140		319	325			376	382	384		309	311		
AMK2	-9				164	172			122	140	142		319	325			376	380			309	311	313	321
AMK3	314	318	334		164				140	142			313	322			380	382	384	388	305	309	311	321
AMK4	316	318	334		164				140	142			319	340			376	382			305	309		
AMK5	318				164	172			122	130	140		319	322			376	380			309	311		
AMK6	316	318			164	172	180		136	142			319	322			380				309	311	317	
AMK7	318	320	334		164	166	172		122	140	142		319	322			376	380			305	311	321	325

Samples	*Scu01Eg*			*Scu20Eg*				*Scu22Eg*				*Scu25Eg*				*Scu31Eg*				*Scu33Eg*			
AMK8	318			166	172			134	142	144		319	322			376	380			309	311	317	
AMK9	316	318		164	172	180		130	138	140	142	319	322	325		380	382	384		309	311		
AMK10	318			164	172			120	132	140	142	313	319			372	376	380	382	309	311	313	
AMK11	316	318		164	166			122	134	140	142	319	322			380				309	311	315	
AMK12	318			164	172			122	140			319	322			376	380			305	309	311	315
AMK13	318	320	322	164	172			124	140	142		319	322			376	380			305	309	311	
AMK14	318	320		164	172			130	132	142	144	319	322			380	382			309	311		
AMK15	318			164	168	180		122	132	140	142	319	325			380	388			305	311	313	
AMK16	318			164	166	172	180	138	140	142		313	319	322	325	376	380			311	313		
AMK17	318			164	170			136	138	144		319				376	380	388		309	311		
AMK18	316	318		164	170	172		136	138	142		319	325			376	378	380	388	305	309	311	
AMK19	318			164	172			140				319	322	325		376	380	382		311	313		
AMK20	318	320		164	172	180		140	142	144		319	322	325		376	380	382		313	315		
ALL1	316	318	320	164	172			120	136	140		319	322	325		376	380	382		309	311	319	
ALL2	318			164	172	176		120	136			319	325			376	380	382		305	309	311	315
ALL3	318			164	172	184		122	136	140		319	322			376	380	382		309	315		
ALL4	316	318		164	172	184		140	142			319	322			376	380			305	309		
ALL5	318	320		164	172	176		120	140	142		319				376	382			-9			
ALL6	318			164	166	172	180	140	142			325				376	378			305	311	321	
ALL7	318			160	164	172		136	142			319	325			376	380			305	311	315	
AMR1	318	320		164	172			128	142			319	322			376	380	382		309	313		
AMR2	314	318		164				122	142			319				376	380			309	311	313	

Samples	*Scu01Eg*			*Scu20Eg*				*Scu22Eg*				*Scu25Eg*			*Scu31Eg*				*Scu33Eg*			
AMR3	314	318	320	164	172			122	128	142		319	322	325	376	380			309	311		
AMR4	314	318		164	172			122	128	142		319	322		372	376	380	382	309	313		
AMR5	314	318		164	172			122	128	132	142	319			372	376	382		305	309	313	
ANT1	318	320		164	166			140	142	144		319			378	380	382		309	311	313	
ANT2	318			164	166			140	142	144		319			380				309	313		
ATS1	316	320		164	172			136	140	142		301	319		376	380	382		305	309	311	319
OUC1	316	328		174	178	180		126				334			376	384			319	325		
OUC2	316			174	178			126				334			376				319	325		
OUC3	316			174	178			126				334			376				319	325		
OUC4	316	328		174	178	180		126				334			376				325			
OUC5	316			168	170	174	178	126				334			376	386			319			
OUC6	316	328		166	170	174	178	128	136			334			376				319			
OUC7	-9			168	174	178	180	-9				-9			376				319	325		
OUC8	316			170	174	178		128	134			334			376				319	325		
OUC9	328			162	168	174	178	128	136			334			376				319	325		
OTVE1	316	338		160	168	174		128	136			337			376				317	319		
OTVE2	316			160	170	174	184	128				334			376	388			319			
OMA1	316			164	168	174		124				334	340		376				319			
OMA2	316			160	166	168	174	124	128			334			382	386			319			
OMA3	316			164	170	174	184	124	136			340			378	390			319			
OMA4	316			168	174	178		124	128	136		334	340		390	392	400	402	319			
OMA5	316			-9				136				334			388	390	392		319			

Samples	*Scu01Eg*				*Scu20Eg*				*Scu22Eg*				*Scu25Eg*				*Scu31Eg*				*Scu33Eg*		
OPR1	316				160	170	178	180	128	136			334	337			376	388			319	321	
OPR2	316	338			164	168	170	174	128	136			334	340			388	392			317	319	
OPR3	316				164	170	174	180	128				334				376				319		
OPR4	316				-9				128				337				376				319		
OPR5	316				-9				128	136			334				384	388			317	319	
OBR1	316	320	328	332	160	164	170	174	132	136	140	142	301	322	334	337	374	382			315	319	
OBR2	320	338			-9				136	138	144	146	340				382	392			319		
OBR3	316				-9				136	140			337				382	384			319		
OBR4	316	328	338		-9				136	138	140		337	340			382				319		
OBR5	316	334	338		166	168	172	176	136				337	340			376	382			319	321	
OBR6	316	332	338		-9				116	134	136		337	340			382	390			319		
OMC1	316	328	330		164	176	178	182	128	134	136	144	337				382	386	388		311	319	321
OMC2	310	316	328		166	176	182	184	134	136	142	144	337				380	382	386		311	319	
OMC3	310	328	338		164	166	172	180	132	134	138		337	340	343		382	386	390		311	319	321
OMC4	322	338			-9				134	136	140		337	340	343		382	418	420		319	321	
OMC5	316	332			-9				130	132	134	136	337	340			382	386	418		311	315	319
OML1	310	316	332		160	166	174	178	136	138	140		337	340			382				309	315	319
OML2	316	322	328		-9				134	136	138	140	-9				382	416			309	315	319
OML3	316	320	328	332	164	166	172	176	124	136	138		337	340			374	382	386	416	315	319	
OML4	316	332	338		-9				134	136	138	140	337	340			380	382	416		309	315	319
OMtB1	316	322	328		160	174	186		122	136	138		-9				382	386			315	319	321
OMtB2	316	322	332	334	164	168	170	172	124	134	138	142	322	337	340		380	382	386		315	319	321

Samples	*Scu01Eg*				*Scu20Eg*				*Scu22Eg*				*Scu25Eg*				*Scu31Eg*				*Scu33Eg*			
OMtB3	316	320	326	332	164	170	172	174	130	136	138	140	325	337	340		380	382			309	315	319	321
OMtB4	316	320	322	332	164	172	186		122	136	138	140	334	337	340		382				309	319	321	
OMtB5	316	320	322	328	168	174	178	182	136	138	140	142	334	337	340		382	388			315	319	321	
OMtB6	316	320	322	328	168	174	178	182	136	138	140	142	334	337	340		382	388			315	319	321	
OMtB7	320	322	328	332	-9				136	138	142		337				374	382	388		315	319		
OMtB8	320	322	328	332	164	172	174	180	134	136	138	140	325	337	340		382				319	321		
OMB1	316	320			168	172	178		134	136			337	340			376	384	388		311	319	321	
OMB2	320	328	334		-9				132	134	136		301	337	340		382				317	319	321	
OMB3	316	328	334	338	164	170	172	174	132	134	136	140	340				376	382	384		317			
OMB4	328	334	338		-9				132	136			337	340			382				319	321		
OBB1	310	326			162	170			132	134	136		301	334			382				-9			
OBB2	310	328	338		-9				134	138	140	146	337				380	382	386		319			
OBB3	310	322	338		-9				-9				337				382	384	386	388	319			
OBB4	326	332			-9				136	144	148		340	343			374	376			319			
OBH1	316	338			164	172	180	182	122	136			334	337			382	386			-9			
OBH2	340				164				122	134	136		301				386				-9			
OBH3	316	338			-9				126	128	136		334	337			390				-9			
OBH4	316	328	338		-9				132	134	144	146	337				382	388			-9			
OBH5	316	338			-9				126	128	136		334	337			382	390			319			
BCL1	316	322			166	168			134	136			319	322			412	422			319			
BCL2	322				166				134	136			322	340			392	412			319			
BCL3	322				164	166			128	134			322				418	422			319			

Samples	*Scu01Eg*		*Scu20Eg*				*Scu22Eg*		*Scu25Eg*			*Scu31Eg*		*Scu33Eg*	
BCL4	322		164				134	136	319			372	376	319	
BPR1	322		166				132	134	334			418	424	319	
BPR2	322	326	166				134	136	319	322		412	418	319	
BPR3	322	326	166				134	136	319	322		412	418	319	
BPR4	322	326	-9				134	136	-9			412	422	319	
BPR5	326		-9				130	134	319			412	418	319	
BWT1	-9		164	166	172		136		-9			-9		-9	
BWT2	322	326	164	166	170		134	136	322			412		319	
BWT3	322		164	166	170	172	134	136	319	331	337	420	422	319	
BWT4	322	326	-9				134		-9			412	422	-9	
BWT5	322		-9				134	136	322	334		412	422	319	
BMtL1	328	334	174	182			134	136	334			376		323	325
BMtL2	322		166				134	136	322			412	422	319	
BMtL3	322		166				132	134	322			412		319	
BMtL4	322		-9				134	136	-9			412	418	-9	
BTR1	322		-9				136		322			416	418	-9	
BTR2	322		-9				134	136	322			418		319	
BTR3	322		-9				134	136	-9			412	424	319	
BTR4	322		-9				132	134	-9			420	422	319	
BTR5	-9		-9				134	136	-9			424		319	
BDC1	326		164	170			132	134	322	334		416	418	319	
BDC2	314		166				128	148	331	334		376	378	323	325

Samples	*Scu01Eg*		*Scu20Eg*			*Scu22Eg*			*Scu25Eg*		*Scu31Eg*		*Scu33Eg*	
BDC3	322	326	166			132	134		319		422	424	319	
ADC1	320	332	164			122	132		334		376		307	
ADC2	320	322	166	178		132			337	340	-9		307	309
ADC3	320		164			132			334	343	376		309	
COO1	340	350	164	168	170	134	136	140	301		-9		-9	
COO2	320	322	164			122	132		337	340	-9		309	315

Appendix 3. Allele diversity and Hardy-Weinberg equilibrium

Levels of interpopulational allelic richness were higher in the tentative tetraploid entities. The average unbiased allelic richness (uA_I) was around 1.93 (± 0.32) in *E. obovatus*, and 2.41 (± 0.54) in *E. arnhemicus*, while it was 1.32 (± 0.17) in *E.* sp. Mt Bellenden Ker. The values of presumed tetraploid *E. arnhemicus* and *E. obovatus* were nearly twofold higher compared to diploid *E.* sp. Mt. Bellenden Ker.

On the other hand, the levels of interpopulational allelic diversity did not show marked difference. The average unbiased allelic diversity (uA_R) was found to be 0.68 (± 0.24), 0.89 (± 0.28) and 0.63 (± 0.11) in the populations of *E. arnhemicus*, *E. obovatus* and *E.* sp. Mt. Bellenden Ker respectively. The average of unbiased allelic diversity (uA_R) was slightly higher in the populations of *E. obovatus* South (1.01 [± 0.18]) compared to the others.
For *E. coorangooloo* allelic richness (A_I) was found at 3.40 and allelic diversity (A_R) was found at 4.00. Values of unbiased allelic richness (uA_I) and unbiased allelic diversity (uA_R) were not obtained due to low samples size and they are not meaningful.

The genetic diversity estimates using *Ho*, *He* and *F*, and test for Hardy-Weinberg equilibrium were obtained only for a diploid taxon, *E.* sp. Mt. Bellenden Ker. For the populations of *E.* sp Mt. Bellenden Ker, observed heterozygosity levels varied considerably among entities ranging from *Ho* = 0.056 *in scuEg34*, to *Ho* = 0.850 in *scu22Eg.* Significant deviations from Hardy-Weinberg equilibrium were found at *scu01Eg*, *scu20Eg* and *scu33Eg.* All of the loci, *scu01Eg*, *scu20Eg and scu31Eg* were at excess of heterozygotes (*F* = 0.191 0.375 and 0.477 respectively). These reports were, however, treated with caution due to the sampling size < 50 (N=23), and the result may be biased (Hedrick, 2000). Other populations, BDC, ADC and COO were not tested due to the even smaller sampling size than *E.* sp. Mt. Bellenden Ker.

Characteristics of allele diversity at the population level. See population codes for table 3.1. Standard deviations were given for the averaged values.

	N	A_1	A_R	uA_1	uA_R	*P*
E. arnhemicus						
ACO	5	11.50±3.51	4.00±1.26	2.30	0.80	0.82
ALF	17	37.83±9.50	5.50±1.87	2.25	0.26	0.75
APP	9	21.14±4.49	5.00±1.41	3.14	1.66	0.81
AIR	3	7.00±2.45	2.33±0.82	2.33	0.71	0.71
ANPA	10	20.83±4.26	5.00±1.67	2.08	0.50	0.71
ABD	8	22.00±3.35	5.50±0.84	3.66	0.91	0.83
ABS	6	13.83±2.22	5.00±2.10	1.73	0.63	0.82
AMK	20	48.00±8.76	7.33±2.07	2.40	0.37	0.85
ALL	7	16.00±3.52	4.50±1.38	2.29	0.64	0.81
AMR	5	11.50±2.26	3.33±0.82	2.11	0.67	0.88
ANT	2	4.00±1.41	2.33±0.82	1.93	2.00	0.72
ATS	1	N/A	N/A	N/A	N/A	0.72
Mean	-	19.50±4.14	4.53±1.37	2.65±1.69	0.41±0.54	0.79±0.06
E. obovatus						
North						
OMA	5	8.67±3.72	4.00±3.58	1.73	0.80	0.41
OPR	5	8.00±2.10	3.33±1.97	1.60	0.67	0.47
OUC	9	14.17±7.60	3.17±2.13	1.57	0.35	0.38
OTVE	2	3.50±1.76	2.50±1.22	1.75	1.25	0.43
Mean*	-	8.58±3.80	3.25±2.27	1.66±0.09	0.77±0.37	0.42±0.04
South						
OMtB	8	24.00±5.80	6.50±2.26	3.00	0.87	0.94
OBH	5	7.83±4.71	4.00±2.28	1.57	0.80	0.64
OBR	6	12.00±3.84	6.71±2.14	2.00	1.03	0.64
OMB	5	9.17±1.94	4.33±1.03	1.83	0.83	0.81
OBB	4	6.83±3.71	4.83±2.79	1.71	1.21	0.68
OML	4	10.33±3.01	4.67±1.86	2.58	1.17	0.97
OMC	5	13.33±2.66	6.00±2.19	2.67	1.20	0.85
Mean	-	11.93±3.67	5.21±2.08	2.19±0.55	1.01±0.18	0.77±0.14
Mean (*E. obovatus* whole)	-	10.26±3.73	4.23±2.15	1.93±0.32	0.89±0.28	0.66±0.22
***E.* sp. Mt. Bellenden Ker**						
BMtL	4	5.00±1.79	3.00±0.63	1.25	0.75	1.00
BCL	4	6.17±1.60	3.00±1.68	1.54	0.75	0.83
BTR	5	5.40±2.97	3.00±2.17	1.08	0.60	0.17
BPR	5	7.00±2.83	2.50±1.38	1.40	0.50	0.50
BWT	5	6.67±2.34	2.83±1.47	1.33	0.57	0.50[#]
Mean	-	6.05±2.30	2.87±1.47	1.32±0.17	0.63±0.11	0.60±0.32
Davies Creek populations						
ADC	3	4.00±1.10	2.50±1.055	1.29	0.83	0.71
BDC	3	4.83±0.98	3.83±1.17	1.50	1.28	1.00
Mean	-	-	-	-	-	-

N: sample size

A_1: mean number of alleles per locus

A_R: mean number of unique alleles per locus

uA_1: A_1/N, unbiased mean number of alleles per locus (mean number of alleles per locus, averaged by the number of samples in the population)

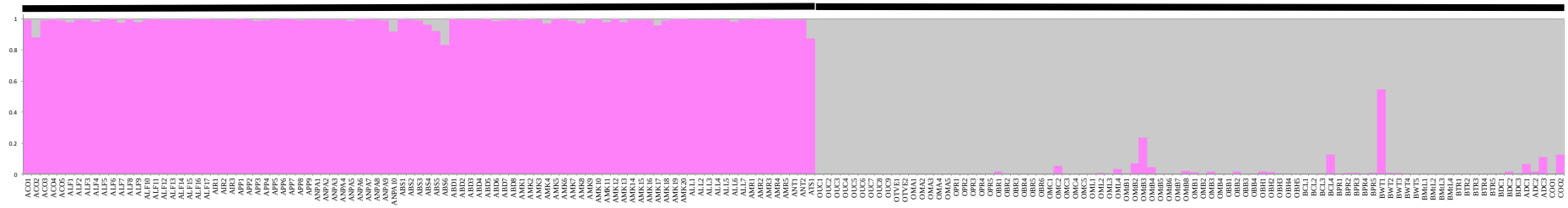
uA_R: A_R/N, unbiased mean number of unique alleles per locus (mean number of unique alleles per locus averaged by the number of samples in the population)

P: proportion of polymorphic loci. [#] locus *scu20Eg* and *scu25Eg* was excluded from the calculation.

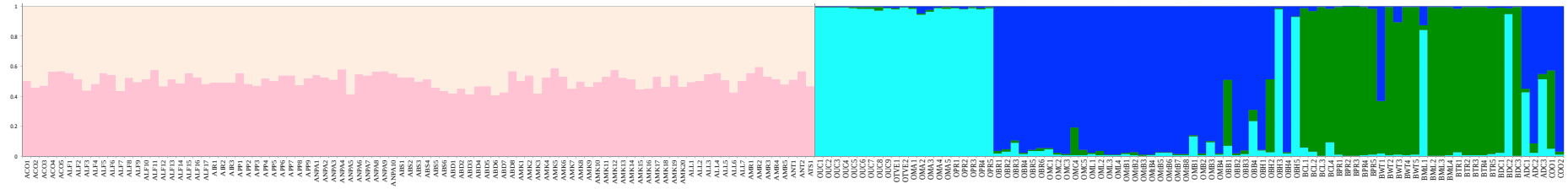
*TSVE was excluded from mean value due to the small samples size (N=2).

E. arnhemicus

E. obovatus, *E.* sp. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC



a.



b. c.

Appendix 4. STRUCTURE output of a. all samples combined (K=2), b. *E. arnhemicus* only (K=2), c. *E. obovatus*, *E.* sp. Mt. Bellenden Ker, *E. coorangooloo*, ADC and BDC (K=3) analysed separately. The 'diploid' ploidy level setting was used. Population codes for the sampling sites are provided in Table 3.1. All the plots depict the lowest Ln probability