The Evolutionary significance of range disjunctions in South Australian eucalypts.

Eucalyptus dalrympleana near Mount Torrens, South Australia

Ed Biffin [1], Kor-Jent van Dijk[1,2] and Michelle Waycott[1,2]

1. Botanic Gardens and State Herbarium of South Australia, Hackney Rd, Adelaide, 5001
2. School of Biological Sciences, The University of Adelaide, Adelaide, 5005

Final report on a project supported by a grant provided by the Australian Flora Foundation in 2016

Study completed August 2020

The Evolutionary significance of range disjunctions in South Australian eucalypts.

Abstract

Many temperate Australian plant species are widespread on the continents' eastern seaboard but extend to the Flinders-Mount Lofty Ranges, showing a range disjunction across the lower Murray Basin. In this study, we use DNA sequencing to investigate the age and origin of this range disjunction in three eucalypt species: *Eucalyptus dalrympleana* (mountain white gum), *E. albens* (white box) and *E. macrorhyncha* (red stringybark). Each is widespread in temperate eastern Australia but within South Australia, have limited distributions within the mesic regions of the Flinders and Mount Lofty Ranges and are listed as rare within that State. By interrogating phylogenetic patterns and divergence times in each of the species, we contrast two competing hypotheses: first, range disjunctions might arise through the action of large scale extrinsic processes such as geological or climatic change, leading to contraction of a formerly widespread range (vicariance); alternatively, range disjunctions might reflect long-distance dispersal across a barrier to gene flow into areas of previously unoccupied suitable habitat. While long-distance dispersal might occur at any time in the history of a species, vicariance would be expected to preserve a signal from the process(es) that generated the disjunction.

We used a hybrid capture approach to develop nuclear (nDNA) and chloroplast (cpDNA) DNA sequence data sets for a range-wide sample of each of the three eucalypt species. The nDNA data, comprising c. 20 low copy nuclear gene regions, was analysed using a Bayesian clustering approach (Structure) and a phylogenetic model that simultaneously estimates evolutionary relationships under a molecular clock while accounting for incomplete lineage sorting and inter-population gene flow. These analyses were largely consistent in the identification of distinct genetic diversity in the South Australian populations of each species and the early divergence of these lineages from the eastern 'core' range. Divergence time estimates generally pre-dated the Last Glacial Maximum, suggesting refugial persistence in

the Flinders-Mount Lofty ranges. Phylogenetic analyses of the cpDNA revealed a strong geographic signal in each of our data sets and strong discordance with morphological species concepts suggesting widespread introgressive hybridisation among geographically proximate individuals. In two of our cpDNA data sets, a split across the Murray Basin separated the eastern populations (*E. dalrympleana*), or those from western Victoria (*E. macrohyncha*) from the Mount Lofty Ranges. Molecular clock estimates suggest these disjunctions coincide with the Oligocene-Miocene marine inundation of the Lower Murray Basin. A more complex phylogenetic pattern was recovered for *E. albens* and related species that suggests both ancient vicariance and possible reconnection of geographically separated lineages in the Flinders-Mount Lofty region.

Taken together, our molecular analyses strongly support the hypothesis of long-term persistence of eucalypt species in the Flinders-Mount Lofty ranges and the importance of the region as a refugium for the Australian southern temperate biota. Significantly, our findings highlight the dynamic history of the region, suggesting range expansion and contraction in response to glacial-interglacial cycling as well as signatures of ancient vicariance that are preserved among hybridising species.

Introduction

Disjunct populations of species can provide important insights into the evolution and biogeography of those species and the ecological communities in which they occur. Disjunction may arise due to contraction of a formerly wider range or through founder events following long distance dispersal from a source area into a suitable new area (Karanth, 2003). In the case of the former, populations may be divided by large scale climatic or geological events and subsequently persist in geographically separated regions, whereas long distance dispersal can occur at any time through the history of the species. While each of these processes generates identical geographic patterns, alternative explanations for species range disjunctions can be untangled using phylogeographic approaches based upon molecular markers (e.g. Larcombe et al. 2011; French et al. 2016; Worth et al. 2018) and interrogating the evolutionary relationships, patterns of genetic diversity and divergence times among lineages (e.g. Zelmer et al. 2012). Understanding the origins of range disjunctions has important implications, and for instance disjunct populations may be indicative of long-term

climatic refugia, incipient species formation and/or distinct genetic lineages worthy of conservation attention (Larcombe et al. 2011; Bradbury et al. 2019).

The Flinders and Mount Lofty Ranges are, respectively, the northern and southern sections of an ancient uplifted marine sedimentary basin that extends c. 900 km in a north-south direction in near-coastal southern South Australia (Figure 1). These ranges are topographically diverse and experience relatively high rainfall, harbouring plant communities with strong affinities to the sclerophyll forests and woodlands found more broadly in south-eastern Australia. A number of species in the region are at the western and/or northernmost extreme of their distribution and are disjunct from populations of the same species in south eastern Australia across the Murray basin (Costermans, 1981), a region of low relief that has experienced frequent marine inundation and may present both an edaphic and climatic barrier to species migration (Figure 1). The Flinders-Mount Lofty region has been recognised as an important centre of species richness and endemism on a continental scale (Crisp et al. 2001) and has been postulated as a refugium for mesic adapted species through the climatic oscillations of the Pleistocene (Byrne et al, 2008; Guerin and Lowe, 2013) and for species that were more widespread during the Miocene (Byrne, 2008). Anthropogenic climate change and habitat fragmentation are current and ongoing threats to species diversity within these regions (Dore et al. 2000; Guerin and Lowe, 2013; McCallum et al. 2014).

The eastern Australian temperate zone extends from south-eastern Queensland to Tasmania and westwards into south eastern South Australia. The region has been largely unglaciated throughout the Pleistocene but has nevertheless experienced periods of significantly reduced temperatures, increased aridity and lowered sea levels during glacial intervals. Fossil pollen records indicate that at the Last Glacial Maximum (LGM: c. 22-18 Ka) temperate eastern Australia was dominated by treeless steppe vegetation with no modern analogue (Hope, 1994), and mesic adapted species are predicted to have persisted within refugia close to their current range (Byrne, 2008). A handful of studies have considered the origins of range disjunctions in widespread, mesic adapted species in southern temperate Australia (e.g. *Eucalyptus bicostata:* Freeman et al. 2001; *Hardenbergia violacea:* Larcombe et al. 2011; *Correa*: French et al. 2016; *Callitris:* Worth et al. 2018) and have identified divergent genetic lineages consistent with refugial persistence, including those in the Flinders-Mount Lofty Ranges. However, there have so far been no attempts to estimate the timing of divergence of the disjunct lineages limiting the biogeographic inferences that can be drawn.

The genus *Eucalyptus* is a dominant element of the Australian flora and is extremely diverse across the eastern temperate region of Australia. As foundation species in many ecosystems, eucalypts strongly influence the microenvironment and ecological processes that determine the habitat for a suite of co-occurring species (Bennett, 2006). Understanding how historical processes have shaped the modern distributions of eucalypt species may therefore give insight into the history of their ecological communities. In this study, we focus upon three eucalypt species that are widespread across southern Australia, with 'core' distributions in the forests and woodlands of south-eastern Australia, and disjunct distributions in the Flinders-Mount Lofty region of South Australia. We use molecular (DNA sequence) data generated using a hybridisation-enrichment approach and high-throughput sequencing (Hyb-Seq: Weitemier et al. 2014) to explore alternate phylogeographic hypotheses for the origins of these range disjunctions. Specifically, a vicariance-refugial hypothesis posits that the pattern of disjunction was formed by the fragmentation of a once continuous range, predicting that the disjunct lineages will be recovered as sister groups and the age of disjunction predates the barrier that caused range division . On the other hand, recent long distance dispersal implies that populations from the geographically separated regions will cluster together in a phylogenetic tree and the age of disjunction post-dates barrier formation.

Methods

Study species

We intensively sampled populations of three *Eucalyptus* species (and closely related taxa; Figure 1; Table 1) which are widely distributed in south-eastern Australia and disjunctly distributed in the Flinders-Mount Lofty Ranges of South Australia. With respect to the disjunct populations, all three species are listed as *rare* under the South Australian *National Parks and Wildlife Act 1972* :

- *Eucalyptus dalrympleana* (subgenus *Symphyomyrtus,* Section *Maidenaria,* Series *Viminales*) is a medium to tall tree of forested mountain country ranging from south eastern Queensland through New South Wales, Victoria and into Tasmania. In South Australia, *E. dalrympleana* is almost entirely restricted to the upper catchment of the Onkaparinga River in the southern Mount Lofty Ranges. The South Australian populations have been previously referred to *E. rubida*, which is closely related to *E. dalrympleana* and shares a similar distributional range along the eastern seaboard. In the present study, we included populations of *E. rubida* (Candlebark) in order to test

the distinctiveness of these species, along with *E. viminalis* (Manna gum), another closely related and co-occurring species that, at least anecdotally, is thought to hybridise with *E. dalrympleana.*

- *Eucalyptus albens* (subgenus *Symphyomyrtus,* section *Adnataria,* series *Buxales*) is a distinctive and widespread species occurring predominantly in woodland communities on the Western Slopes of the Great Dividing Range from southern Queensland to central Victoria. A disjunct population occurs near Melrose, South Australia, on the eastern side of the southern Flinders Ranges. *E. albens* co-occurs with several closely related species including *E. microcarpa* (Grey box) and intergrades are commonly reported (e.g. Nicolle, 2013). We have included samples of several related, and potentially hybridising species, in this study.
- *Eucalyptus macrorhyncha* (subgenus *Eucalyptus,* section *Capillulus* series *Pachyphloiae*) is a medium sized tree of relatively dry hills and tablelands occurring from south eastern Queensland to Stawell in western Victoria. In South Australia, *E. macrorhyncha* is restricted to the Clare region in the northern Mount Lofty Ranges where it is dominant component of woodland communities. We sampled several populations of the widespread subsp. *macrorhyncha,* including those in South Australia, as well as subsp. *cannonii*, which is restricted to a relatively small region of the Central Tablelands in New South Wales.

Bait design

We used hybrid capture and high throughput sequencing (e.g. Weitemier et al. 2014) to generate a DNA sequence data set. The hybrid capture approach enables us to develop data that is comparable across divergent taxa and can be readily extended to include new samples (and loci). In addition, hybrid capture is well suited to data generation from degraded samples such as herbarium material (Hart et al. 2016).

We targeted low-copy nuclear genes (Table 2) as well as 15 regions of the chloroplast genome (Table 3) using 'universal' angiosperm bait sequence sets designed in-house from published genomic resources (e.g. Phytozome, https://phytozome.jgi.doe.gov/; 1KP, http://www.onekp.com/) and synthesised by Arbor Biosciences (formerly MYcroarray; Ann Arbor, Michigan, USA) (Waycott et al. in prep). Sample preparation was also performed in-house, including the extraction of genomic DNA (gDNA) from silica-dried or herbarium leaf

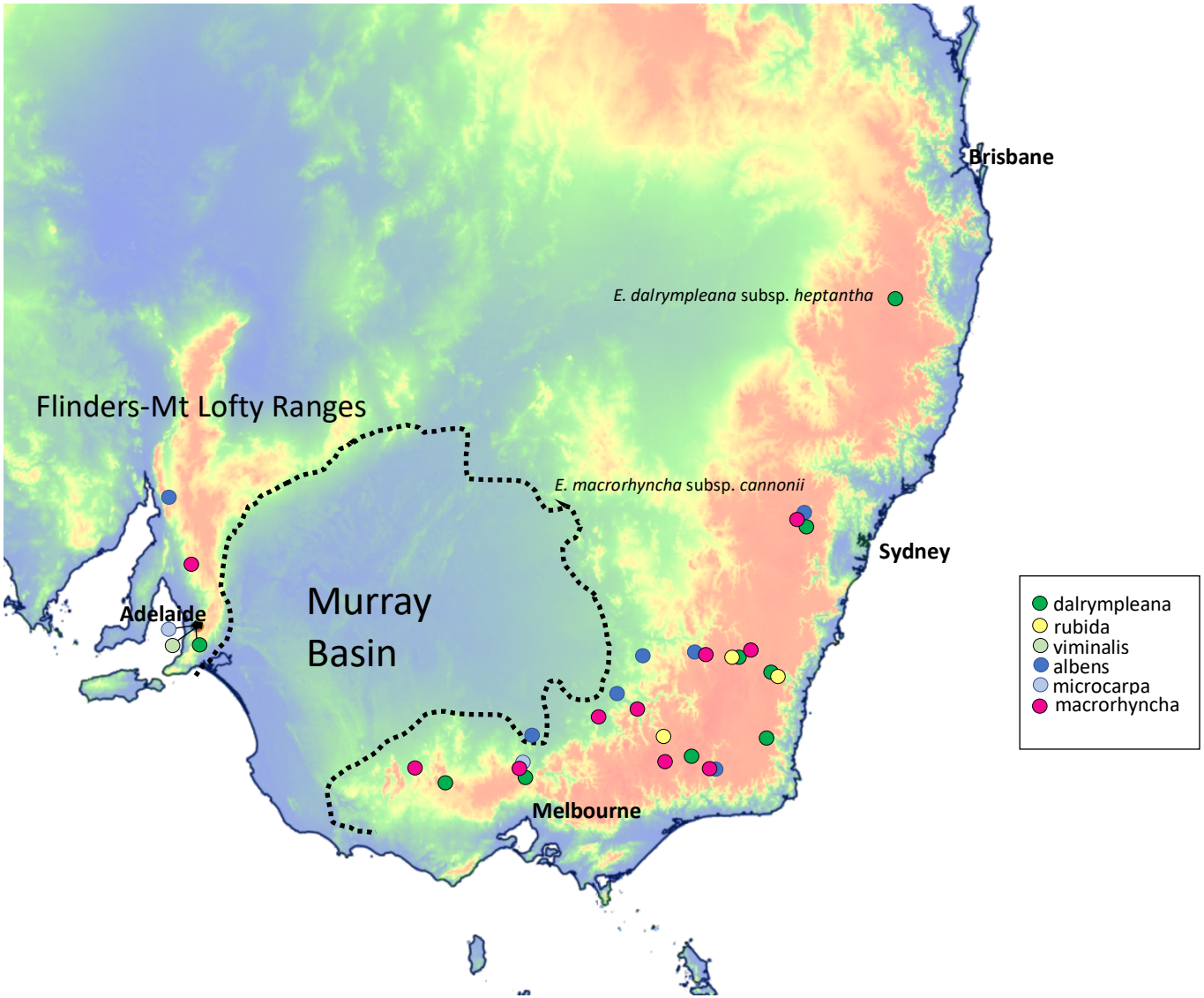

Figure 1: Relief map of South-eastern Australia showing the location of population sampling of eucalypt species. Also indicated is the approximate boundary of the Lower Murray Basin (dotted line) and the location of the Flinders-Mount Lofty Ranges.

Table 1: Sampling locations and number of samples (n) for *Eucalyptus* populations included in this study.

Taxon	Location	Lat.	Long	n
Eucalyptus albens Benth.	Howlong, NSW*	-35.934	146.581	8
	Adelong, NSW*	-35.308	148.0524	8
	The Rock, NSW	-35.3424	147.073	5
	Rushworth, Vic	-36.5786	144.9767	8
	Melrose, SA*	-32.8229	138.1843	10
	Capertee-Glen Davis, NSW	-33.13214	150.1542	8
	Tooborac, Vic	-37.0218	144.8415	7
	McKillops Bridge	-37.0817	148.4144	5
Eucalyptus bicostata Maiden, Blakely & J.H.Simmonds	Mount Bryan, SA*	-33.45	138.95	4
Eucalyptus dalrympleana Maiden subsp. *dalrympleana*	Beaufort, Vic*	-37.2907	143.383	4
	Forest Range*	-34.91194	138.7969	10
	Wolgan Valley, NSW	-33.32444	150.0978	8
	Daylesford, Vic*	-37.1508	144.7711	5
	Limestone Rd, Vic*	-36.8963	147.9869	8
	Brown Mt, NSW*	-36.6106	149.4055	7
	Tallaganda, NSW	-35.6177	149.4993	6
	Brindabella, NSW*	-35.359	148.7994	8
Eucalyptus dalrympleana subsp. *heptantha* L.A.S.Johnson	New England Tableland, NSW			[1]4
Eucalyptus macrorhyncha F.Muell. subsp. *macrorhyncha*	Spring Gully CP, SA*	-33.912544	138.60479	12
	Stawell, Vic*	-37.0842	142.7821	5
	Tooborac, Vic*	-37.0503	144.7923	5
	Baranduda, Vic*	-36.1769	146.9527	5
	Bundarra River, Vic*	-36.9791	147.4936	5
	Little River, Vic	-37.079	148.3239	5
	Canberra, ACT*	-35.2493	149.0982	5
	Tumut SF, NSW*	-35.3146	148.2076	8
	Warby Range, Vic*	-36.3108	146.2084	5
Eucalyptus macrorhyncha subsp. *cannonii* (R.T.Baker) L.A.S.Johnson & Blaxell	Baal Gap, NSW*	-33.21	150.021	8
Eucalyptus microcarpa (Maiden) Maiden	Castlemaine, Vic*	-37.0685	144.1775	5
	Mt Remarkable, SA*	-32.725607	138.081843	3
Eucalyptus odorata Behr in D.F.L. von Schlechtendal	Melrose	-32.8229	138.1843	5
Eucalyptus rubida H.Deane & Maiden subsp. *rubida*	Omeo Hwy, Vic*	-36.599	147.436	5
	Krawarree, NSW	-35.6638	149.6342	5
	Brindabella, NSW*	-35.361	148.7668	5
Eucalyptus viminalis Labill. subsp. *viminalis*	Blockers Rd, SA*	-34.92955	138.7679	5

[1] sampling was based upon 4 geographically proximate samples from herbarium specimens rather than a population sample per se.

*individual(s) from population included in whole chloroplast genome analyses (see Figs. 7-8)

tissues, fragmentation of the gDNA and addition of sample specific barcodes to enable downstream bioinformatics processing of the DNA sequences. Following hybrid-capture, which uses the bait sequences to enrich each sample for the targeted genomic regions, samples were pooled and sent for Illumina paired-end sequencing at the Australian Genome Research Facility.

Data processing

High-throughput 150 bp paired-end reads were processed using CLC Genomics Workbench v7.5.1 (https://www.qiagenbioinformatics.com). Following demultiplexed and quality trimming (Phred-score threshold of 20) , we used *de novo* assembly of pooled samples to generate a set of contigs for each of the target species. In order to recover the targeted nuclear loci, the *de novo* assembly contigs were converted to a BLAST database and we used reference genomic sequences in *Eucalyptus grandis* (downloaded from Phytozome v 12.1, https://phytozome.jgi.doe.gov\) as query sequences (Table 2). The *de novo* contigs matching the *Eucalyptus* genes were then used as a reference for each individual to generate a per sample mapping for each locus. The resultant mapping files were exported in BAM format and allele phasing was performed on the BAM files using SAMTools Phase (Li et al. 2009). Consenus sequences were generated for each allele per locus and individual using a coverage cut-off of 10. Allele sequences were imported into Geneious v1.11.5 (Kearse et al. 2012), aligned using the MUSCLE (Edgar et al. 2004) plugin with default parameters and each alignment was manually checked and adjusted.

In order to recover the chloroplast target regions we used the chloroplast genome sequence of *Eucalyptus globulus* (Genbank accession number AY780259) as a reference to query the BLAST databases (as above) and extracted *de novo* contigs matching the reference genes. The extracted contigs were used as a mapping reference for each sample and from each locus we extracted the majority rule consensus sequence inserting 'Ns' when coverage was lower than 10. Consenus sequences for each individual and locus were aligned, as above, and manually adjusted. In addition to the targeted cpDNA regions, we were able to assemble whole chloroplast genomes for some samples where there was sufficent capture of non-targeted regions (by-catch) of the chloroplast (Tables 1 and 4).

Table 2: Nuclear genes targeted in this study. The target gene name is according to the *Eucalyptus grandis* V. 2.0 genome sequence available on Phytozome v12.1 (https://phytozome-next.jgi.doe.gov/info/Egrandis_v2_0).

target gene name	protein product	average contig length
Eucgr.A01830	CHROMATIN MODIFICATION-RELATED PROTEIN EAF1	1212
Eucgr.J00560	HISTONE DEACETYLASE 8	1148
Eucgr.F04230	SPHINGOMYELIN SYNTHETASE	1043
Eucgr.I02060	REGULATOR OF CHROMOSOME CONDENSATION FAMILY PROTEIN	1334
Eucgr.K02435	U3 SMALL NUCLEOLAR RNA-ASSOCIATED PROTEIN 12	1569
Eucgr.J02388	FG-GAP REPEAT-CONTAINING PROTEIN	1965
Eucgr.E00067	chitinase domain-containing protein 1	1631
Eucgr.I02387	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase	902
Eucgr.H01265	TRANSCRIPTION INITIATION FACTOR IIB-RELATED	2585
Eucgr.G01740	tRNA-dihydrouridine(47) synthase (NAD(P)(+))	773
Eucgr.G02722	3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE-LIKE PROTEIN	1771
Eucgr.G01983	NO EXINE FORMATION 1	2236
Eucgr.K02331	ALCOHOL DEHYDROGENASE-LIKE 6	937
Eucgr.H04377*	peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase	1043
Eucgr.H04378*	peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase	953
Eucgr.B02258	DNA damage-binding protein 2	1124
Eucgr.F01438*	Peptide chain release factor N(5)-glutamine methyltransferase	1457
Eucgr.F01442*	Peptide chain release factor N(5)-glutamine methyltransferase	1276
Eucgr.F00122	Guanidinoacetate N-methyltransferase	1995
Eucgr.I00705	peptidyl-prolyl cis-trans isomerase-like 2	1747

Table 3: Chloroplast genes targeted in this study. The coordinates are relative to position on the *Eucalyptus globulus* chloroplast genome sequence (genbank accession number AY780259).

target gene name	protein product	ref. position(average contig length)
accD	acetyl-CoA carboxylase beta subunit	60854-63001 (2148)
atpB	ATP synthase CF1 beta subunit	57636-59020 (1385)
atpH	ATP synthase CF0 C subunit	14486-15088 (603)
atpI	ATP synthase CF0 A subunit	15732-17176 (1445)
matK	maturase K	2110-3826 (1717)
ndhF	NADH dehydrogenase subunit 5	116899-117691 (793)
ndhC	NADH dehydrogenase subunit 3	54006-54872 (867)
petD	cytochrome b6/f complex subunit IV	81676 -82585 (910)
psbA	photosystem II protein D1	308 -1851 (1544)
psbD	photosystem II protein D2	35644-36444 (801)
psbH	photosystem II protein H	78-930-79591 (662)
psbZ	photosystem II protein Z	38642-39268 (627)
rbcL	ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	59133-60034 (902)
rpl16	ribosomal protein L16	85578-86664 (1087)
rpoC1	RNA polymerase beta' subunit	24381-25169 (789)

nDNA analyses

Bayesian clustering analyses

We used the Structure v. 2.3.4 software (Pritchard et al. 2000) to determine how many genetic groups were present in each of our data sets. For the Structure analyses, each aligned locus was first converted to an appropriate format using the 'Mask Alignment' option in Geneious. For each locus, we stripped alignment columns containing identical bases and ambiguous sites (i.e. missing data represented by Ns). The reduced alignments were converted to numeric coding (i.e. A=1, C=2, etc.) and each was treated as an unlinked locus. In Structure, we used the admixture model with correlated allele frequencies and incorporating sample location information as a prior (LOCPRIOR option; Hubisz et al. 2009). We performed 10 independent runs for each value of K ranging from 1-8. For all analyses, the burn-in fraction was set to 200,000 iterations with 1,000,000 post-burnin MCMC iterations. The *Delta-K* method of Evanno et al. (2005) as implemented in Structure Harvester (Earl & vonHoldt, 2012) to estimate the most likely number of clusters, *K,* in the data.

Phylogeographic analyses

We used the AIM (Approximate Isolation-with-Migration) model in *BEAST (Müller et al. 2019) to infer phylogenetic relationships between sampled populations. This approach uses the multi-species coalescent (MSC) model (Lui et al., 2009; Heled and Drummond, 2010) to infer a species tree while accounting for discordance among gene trees due to incomplete lineage sorting. However, standard implementations of the MSC assume complete isolation of lineage after they have split, while ongoing gene flow can lead to biased estimates of the evolutionary history of the species (Leaché et al. 2014). The AIM model allows the inference of the species tree under the MSC while explicitly modelling gene flow between lineages (Müller et al. 2019).

We used BEAST v. 2.6 (Bouckaert et al., 2014) to estimate the pattern and timing of diversification in each of our eucalypt samples using the AIM model as implemented in *BEAST2 (Ogilvie et al. 2017). For each locus, we used the HKY+I+G model of sequence evolution and a strict clock model that were linked across loci and parameters of the AIM model (effective population size, Ne, and migration rate, M) were estimated from the data.

For each data set, we conducted two independent analyses 2.0 $\times 10^8$ steps and convergence between runs and the appropriate burn-in fraction was assessed using Tracer v1.6.0 (Rambaut et al., 2015). Output tree files, excluding the burn-in fraction, were combined using LogCombiner v2.5.1 and summarised in TreeAnnotator v2.5.1. Trees were visualised using FigTree v1.4.3 (Rambaut, 2012).

cpDNA analyses

Chloroplast haplotype analysis

For phylogenetic and network analyses of the cpDNA data, we generated a concatenated alignment of the targeted chloroplast genes for each of our target species groups. Because network analysis is sensitive to missing or undetermined values, we stripped all ambiguous sites and gaps to generate an alignment of haplotypes. For network analyses, we used Popart (http://popart.otago.ac.nz) to generate a median joining network (Bandelt et al. 2014). We constructed phylogenies from the same alignments using PhyML (Guindon et al. 2010) and the GTR+I+G model of sequence evolution. Branch support was assessed using 200 non-parametric bootstrap replicates.

To explore chloroplast diversity in *E. albens*, we developed a phylogenetic hypothesis based upon the alignment of whole chloroplast genomes for 14 species from Section *Adnataria* along with outgroups (hereafter, referred to as the *Adnataria* data set). By-and-large, these data were sourced from Alwadani et al. (2019) but were supplemented with South Australian samples (Table 4). Alignment of the cpDNA genome data was performed using the MAFFT v7.45 (Katoh and Standley, 2013) plugin for Geneious. Phylogenetic analysis was performed using RAxML v. 8 (Stamatakis, 2014) and the GTR GAMMA I model of sequence evolution. Statistical support was assessed using 1000 rapid bootstrapping replicates.

Molecular dating of chloroplast genome data

Molecular dating was performed using BEAST v.2.5.1 (Bouckaert et al., 2014) for an alignment of whole chloroplast genomes and including key lineages identified in the analyses of chloroplast data outlined above. The alignment includes the majority of samples from Bayly et al. (2013) but supplemented with representatives of the Section *Adnataria* group (including *E. albens*), the Section *Maidernaria* group (including *E. dalrympleana*), and *E .macrorhyncha.* With respect to Bayly et al.'s data, we excluded samples that were of unknown geographic origin but on the basis of preliminarly analyses were found to be

Table 4: South Australian samples included in the whole chloroplast genome analysis of the *Adnataria* group

Taxon		Location	Lat	Long
Eucalyptus albens	EB526	Mt Remarkable	-32.82	138.184
Eucalyptus albens	EB527	Mt Remarkable	-32.82	138.184
Eucalyptus leucoxylon	EB519	Spring Gully	-33.91	138.605
Eucalyptus leucoxylon	EB522	Beetaloo	-33.2	138.252
Eucalyptus leucoxylon	EB524	Pitchi-Ritchi	-32.38	137.984
Eucalyptus leucoxylon	EB528	Mt Remarkable	-32.73	138.082
Eucalyptus leucoxylon	EB515	Callington	-35.1	138.963
Eucalyptus leucoxylon	EB535	Mt Torrens	-34.883	138.981
Eucalyptus microcarpa	EB529	Mt Remarkable	-32.73	138.083
Eucalyptus microcarpa	EB534	Norton Summit	-34.91	138.704
Eucalyptus odorata	EB521	Laura	-33.19	138.259
Eucalyptus odorata	EB523	Pitchi-Ritchi	-32.45	137.978
Eucalyptus odorata	EB530	Mt Remarkable	-32.85	138.18
Eucalyptus porosa	EB525	Dutchmans Stern	-32.3	137.986
Eucalyptus porosa	EB531	Port Germaine Gorge	-32.99	138.067
Eucalyptus porosa	EB534	Tepko	-34.94	139.077

but on the resolved amongst our target groups, as these may confound the interpretation of geographic signal in the data.

We used fossil data to calibrate the molecular clock and estimate lineage divergence times. We used the same fossil calibrations as Thornhill et al. (2019) including: (1), a Patagonian *Eucalyptus* fossil of Gandolfo et al. (2011) from the Early Eocene (51.7–52.1 million years ago) as a crown *Eucalyptus* calibration; (2), an Australian *Myrtaceidites tenuis* fossil pollen (Thornhill and Macphail 2012) from the Eocene (45–47 million years) to calibrate the *Angophora+Corymbia* crown group; and (3), a Paleocene pollen *Myrtaceidites mesonesus* from New Zealand (61.7–65 million years ago; Thornhill et al. 2012) to provide a plausible upper age for the tribe Eucalypteae (Figure 8). In each instance, we used log-normal calibration priors to constrain the age of the relevant node with a zero offset approximately 10% younger than the age of the fossil and a median value equivalent to the fossil age.

DNA sequence alignment was performed using MAFFT, as outlined above. We used a GTR+I+G substitution model and a relaxed lognormal clock model. We performed two

parallel BEAST runs, each of 200 million steps sampling trees and parameter values every 20,000 steps. Convergence and the appropriate burn-in fraction was assessed using Tracer v1.6.0, Output tree files, excluding the burn-in fraction, were combined using LogCombiner v2.5.1 and summarised in TreeAnnotator v2.5.1. Trees were visualised using FigTree v1.4.

Results

nDNA

Genetic Structure

For each of the three included eucalypt species, population assignment analyses using Stucture support the genetic distinctiveness of populations of populations sampled from the Flinders-Mount Lofty Ranges (Figs. 2-4).

For the *E. dalrympleana* data set, the optimal number of clusters estimated using the Evanno et al. (2005) *Delta-K* method was 4. At *K*=4, the genetic groupings largely align with morphological species concepts, with *E. viminalis, E. rubida* and eastern samples of *E. dalrympleana* subsp. *dalrympleana* forming distinct genetic groups. Samples of *E. dalrympleana* collected in the Mount Lofty Ranges are distinguished from the eastern populations of the same species (Figure 2a).

For *E. macrorhyncha,* the optimal number of genetic groups was 2, with the South Australian (Spring Gully) population forming one cluster, while the majority of eastern populations were unambiguously assigned to the second cluster. However, the Stawell population of *E. macrorhyncha* is an exception, showing evidence of admixture between the two main genetic groups (Figure 3a)

The Structure analyses of the *E. albens* data using the *Delta-K* method suggests the most likely number of clusters in the data is 3. At K=3, the Melrose samples of *E. albens* form a distinct cluster. Cluster 2 includes the majority of *E. albens* individuals sampled from the eastern part of the range, along with populations of *E. microcarpa* sampled from South Australia and Western Victoria. The third cluster includes 2 individuals sampled from The Rock, while samples from Adelong show admixture between cluster 2 and cluster 3. This third cluster is not strongly geographically or taxonomically aligned, but might potentially indicate misidentification of samples and/or the influence of interspecific gene flow

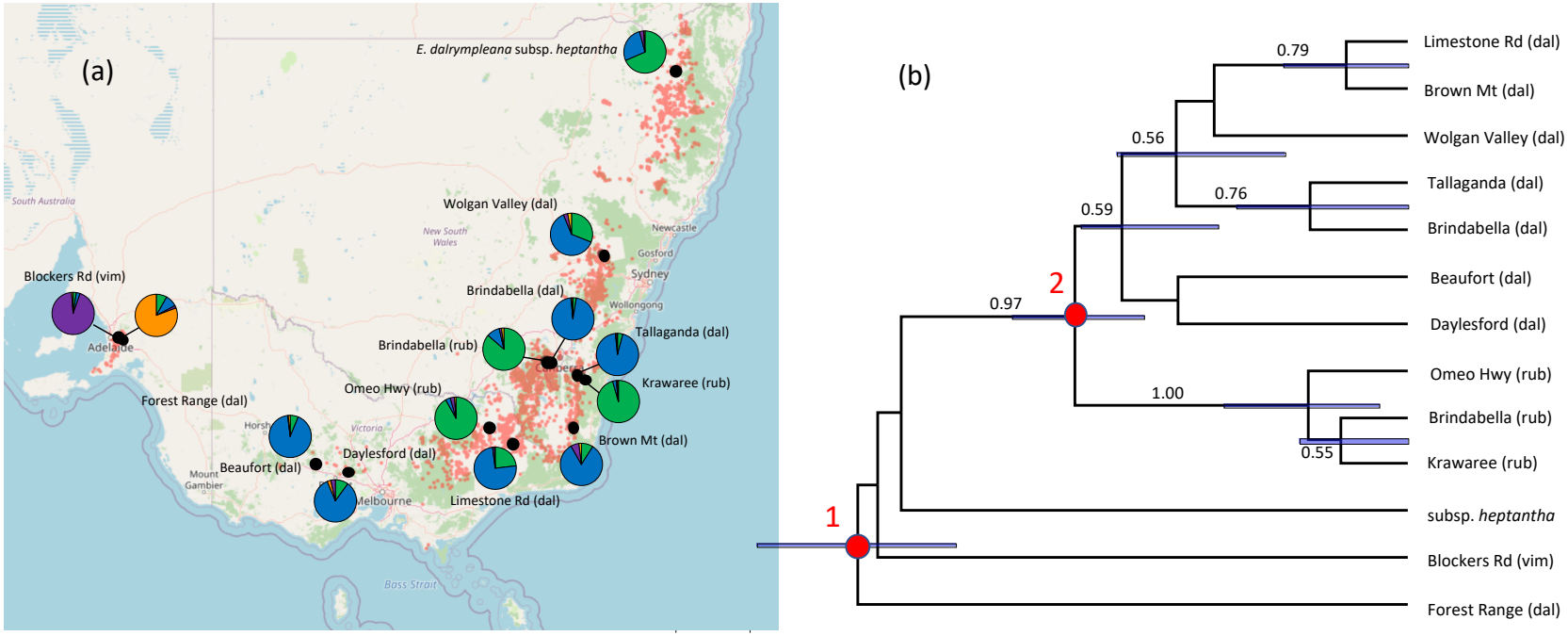

Figure 2: (a) Structure analysis of the *E. dalrympleana* data set. Population pie charts indicate the proportion of individuals assigned to each of four genetic clusters with each cluster represented by a different colour. The distribution (red points) of *E. dalrympleana* (data sourced from the Australian Virtual Herbarium: https://avh.ala.org.au/ accessed March 2020) is also indicated; (b) Maximum clade credibility tree for sampled populations inferred using the AIM model in *BEAST. Posterior probabilities ≥ 0.5 are indicated adjacent to the branch and branch lengths (substitutions/site, scale bar indicated) are proportional to time. Blue bars indicate the variation in branch length estimates (95% highest posterior density, HPD) around that node. The divergence of South Australian *E. dalrympleana* from eastern populations is inferred to have occurred between (1), the age of the root and (2), the crown group of eastern *E. dalrympleana* and *E. rubida,* indicated by the red circles. Population codes: vim, *E. viminalis;* dal, *E. dalrympleana* subsp. *dalrympleana;* rub, *E. rubida.*

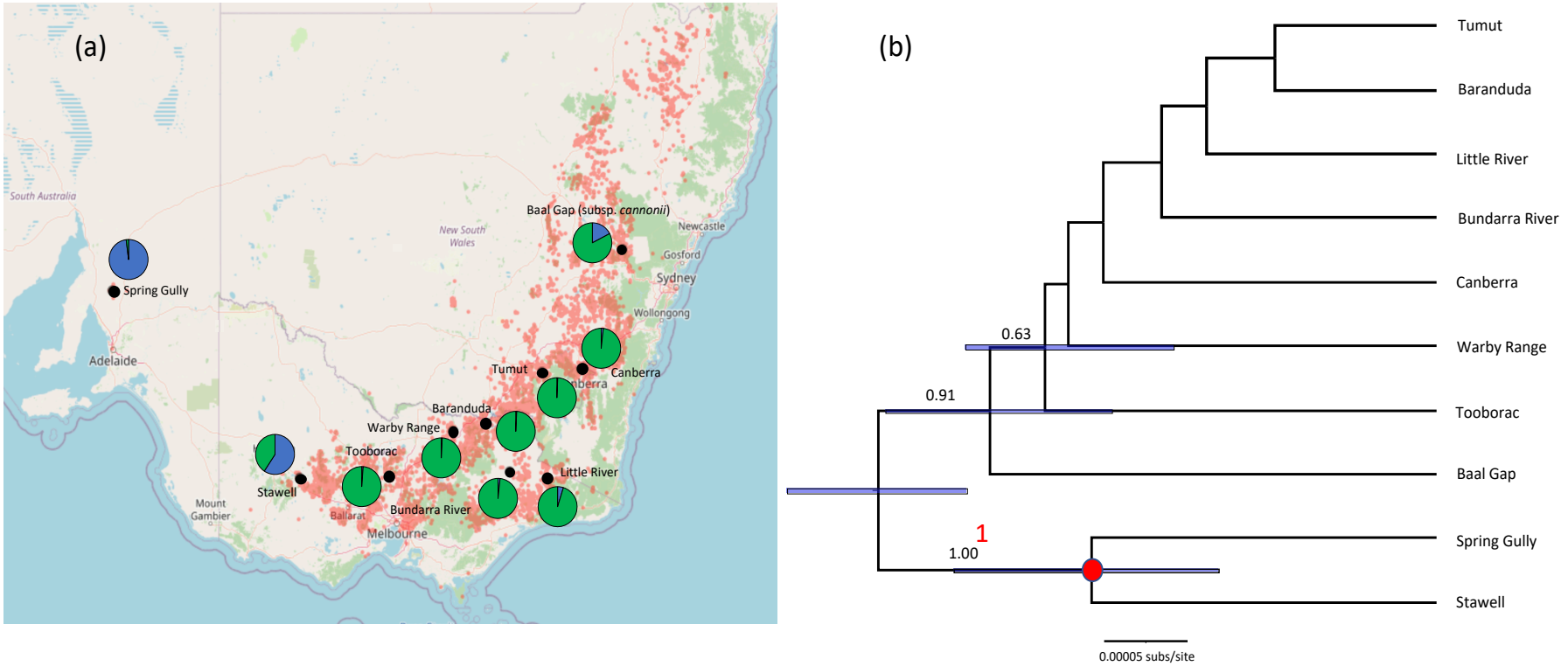

Figure 3: (a) Stucture analysis of the *E. macrorhyncha* data set. (b) Maximum clade credibility tree for sampled populations inferred using the AIM model in *BEAST. The red circle indicates the divergence of the Spring Gully population from the Stawell population in western Victoria.

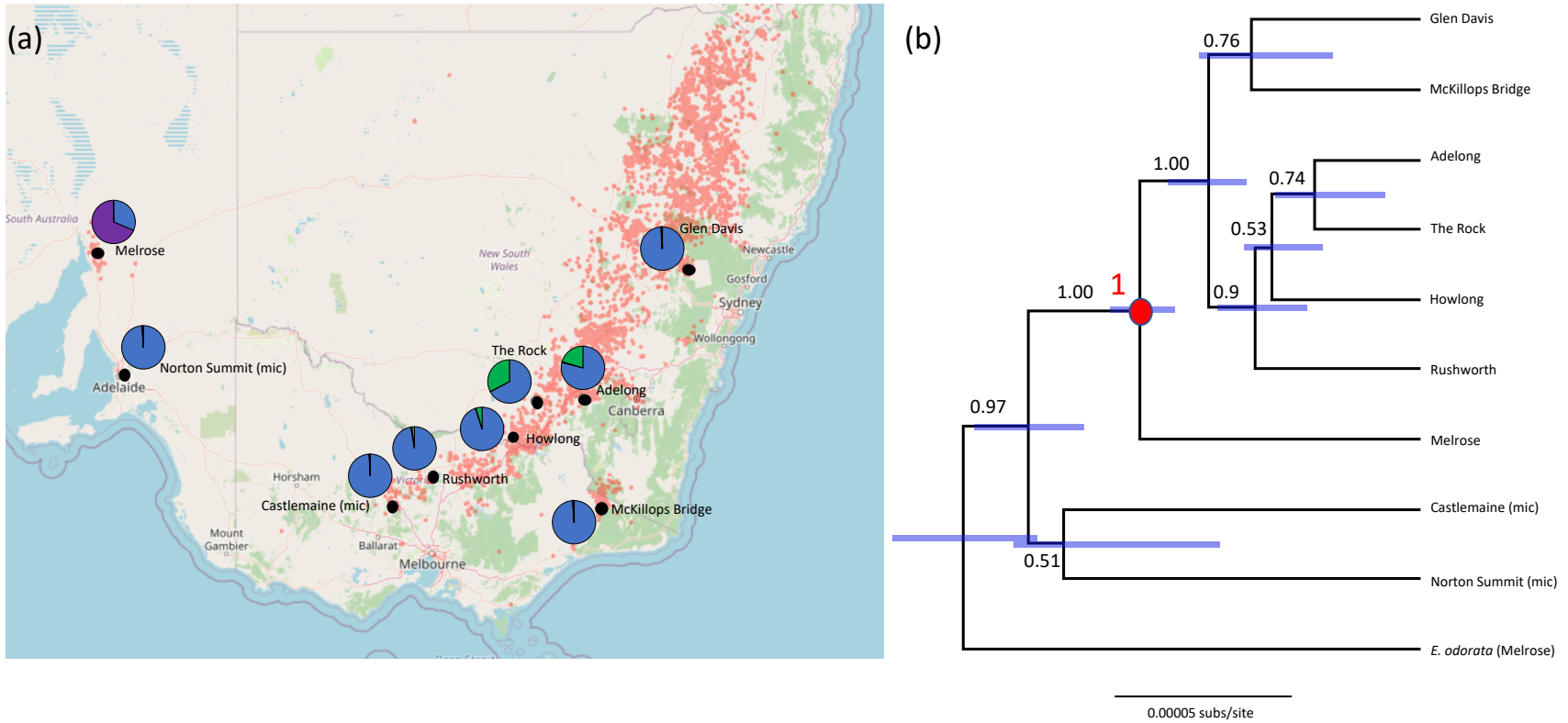

Figure 4: a) Structure analysis of the *E. albens* data set; (b) Maximum clade credibility tree for sampled populations inferred using the AIM model in *BEAST. The red circle indicates the divergence of the Melrose population of *E. albens* from conspecifics to the east of the Murray Basin. mic, *E. microcarpa*

involving *E. albens* and an unsampled, hybridising species (Figure 4a).

Phylogeographic analysis

In each of our study species, phylogenetic analyses using *BEAST resolved the South Australian populations as sister to most, if not all, of the populations sample from the eastern part of their range. The phylogenetic relationships also show a reasonable correspondence with the main genetic groupings recovered in the Structure analyses, above.

For the *E. dalrympleana* data (Figure 2b), an 'eastern' clade was strongly supported, including a monophyletic group comprising *E. dalrympleana* that is sister to populations assigned to *E. rubida.* The Mount Lofty Ranges populations of *E. viminalis* and *E. dalrympleana* and *E. dalrympleana* subsp. *heptantha,* from Northern NSW, fall outside of the 'eastern' clade but relationships among these are not well-resolved.

In the *E. albens* data set, we recovered a strongly supported clade comprising the eastern populations of *E. albens,* while the Flinders Ranges (Melrose) population was sister to these. Overall, *E. albens* was found to be monophyletic and sister to populations of *E. microcarpa,* while *E. odorata,* sample from Melrose, was sister to these (Figure 3b)

The South Australian population of *E. macrorhyncha* was resolved with strong support in a clade including the Stawell population, from western Victoria. These are sister to a lineage including *E. macrorhyncha* subsp. *cannonii* from the NSW Central Tablelands, along with all other populations sampled from the eastern part of the species range (Figure 4b).

Divergence times

To estimate divergence times, we used the genome-wide per generation mutation rate for *E. grandis* (4.93×10^{-8} substitutions/site/generation; Silva-Junior and Grattapaglia, 2015) to scale mutation rates on our trees (Table 5). Allowing for uncertainty in phylogenetic resolution and in branch length estimates, the number of generations since the divergence of the South Australian from 'eastern' populations range from c. 4.5-16.6 thousand generations for *E. dalrympleana;* c. 9.5-12 thousand generations for *E. albens;* and c. 2.6-4.3 thousand generations for the divergence of the Melrose and Stawell populations of *E. macrorhyncha.*

In their study, Silva-Junior and Grattapaglia (2015) cite a generation time for *E. grandis* as 10 years, this being the time to first flowering in natural populations. However, there are other definitions of generation time, and for instance, Lande et al. (2003: p. 55) defines generation time as "the mean age of mothers of newborn individuals when the population has achieved a stable age distribution". Under this definition, time to first flowering is weighted by annual adult survivorship, which in trees is typically high (Petit and Hampe, 2006). In Table 5, we present the estimated years since divergence for the South Australian lineages, assuming either a generation time of 10 or 25 years with the latter estimate derived from Lande et al.'s (2003) definition of generation time and assuming a conservative annual adult survivorship rate of 0.95. Under either scenario, the time since divergence exceeds the timing of the LGM for each of our species, while the upper estimates suggest that the South Australian eucalypt lineages diverged more than 100 Ka.

Chloroplast data

Chloroplast haplotype analyses

E. dalrympleana

The cpDNA data included 79 individuals from 5 species and 3 subspecies from sect. *Maidenaria*. The phylogenetic tree, based upon 65 variable positions, provides strong support for an geographically 'eastern' versus a 'South Australian' clade, which are also found in the haplotype network analysis (Figure 5). *E. dalrympleana* is resolved in both groups and within each, is not monophyletic, as is the case for *E. viminalis* (subsp. *viminalis* and subsp. *cygnetensis,* 'South Australian' clade) and *E. rubida* ('eastern' clade).

E. macrorhyncha

The chloroplast haplotype data for *E. macrorhyncha* included 77 individuals and 149 variable positions. Phylogenetic analyses of these data recovered a well-supported topology including an 'eastern' clade that is sister to a group including all South Australian samples, along with individuals from western Victorian (Stawell and Tooborac, Figure 6) populations. Within the latter, we resolved two distinct groups, one including individuals from the Mount Lofty Ranges, the second including individuals from South Australia's south eastern region

Table 5: Divergence time estimates for eucalypt species based upon population analyses of the nDNA using the AIM model in *Beast

	node[2]	node height (subs./site) : mean (lower, upper)	generations[3]: mean (lower, upper)	div. time (yrs): g[4]=10 years: mean (lower, upper)	div. time (yrs): g=25 years: mean (lower, upper)
Eucalyptus dalrympleana	1	0.0006 (0.00041, 0.00082)	12170 (8316, 16632)	121704 (83164, 166329)	304260 (207911, 415822)
	2	0.00032 (0.00022, 0.00043)	6491 (4462, 8722)	64909 (44625, 87221)	162272 (111562, 218053)
Eucalyptus albens	1	0.00055 (0.00048, 0.00061)	11156 (9736, 12373)	111561 (97363, 123732)	278905 (243408, 309331)
Eucalyptus macrorhyncha	1	0.00021 (0.00013, 0.00029)	4260 (2637, 5882)	42596 (26369, 58824)	106491 (65923, 147059)

[2] node numbers are as per Figures 2-4.
[3] calculated by dividing node height (subs./site) by per generation mutation rate for *E. grandis* (refer text)
[4] g = generation time

(i.e. east of the Murray Darling Depression) along with the Stawell and Tooborac populations. Haplotypes show geographic relationship regardless of their morphological species assignment and for example, the Mount Lofty Ranges clade includes individuals of *E. macrorhyncha, E. baxteri, E. diversifolia* and *E. obliqua.*

E. albens

The chloroplast haplotype data included 122 individuals and 61 variable positions. Phylogenetic analyses of these data produced a poorly resolved topology with strong statistical support largely restricted to individuals sampled from the same population and little resolution of the backbone (results not shown).

The analysis of whole chloroplast data included fewer (68) individuals from 14 species within section *Adnataria* including representatives of *E. albens, E. microcarpa, E. odorata, E. leucoxylon* and *E. porosa* sampled from the Flinders-Mount Lofty region (Table 1 and 4). The maximum likelihood topology resolved two major well-supported clades (clades A and B in Figure 7) and several well-supported sub-clades. With respect to the morphologically defined species, all were resolved as polyphyletic and in most cases, individuals from the same morphological species were resolved in each of the major clades. Geographic origin of the sample may better predict its phylogenetic relationships (see Alwadani et al. 2019) and for instance, Clade A includes two well-supported sub-clades with individuals from (predominantly) north-eastern and central NSW (clades A_1 and A_2, respectively) while Clade B has a largely 'southern' (Victoria and south-eastern NSW) membership. The South Australian individuals are resolved in both major clades and in some instances, occur in close geographic proximity. For example, at Mount Remarkable individuals of *E. albens* are resolved with both major clades along with *E. odorata* from Clade A (Figure 7).

Molecular dating of chloroplast genome data

Phylogenetic relationships inferred from the alignment of whole chloroplast DNA sequences produced a well-supported topology that is mostly consistent with accepted relationships amongst eucalypts (e.g. Thornhill et al. 2019; see Bayly et al. 2013 for chloroplast genome analysis) and with those inferred for each our target groups based upon cpDNA and a denser sample of individuals (Figures 5-7). With respect to divergence times, our estimates are comparable to Thornhill et al. for deeper nodes (e.g. sub-genera) although for each of our target groups, the divergence time estimates among samples generally exceeds the age of the

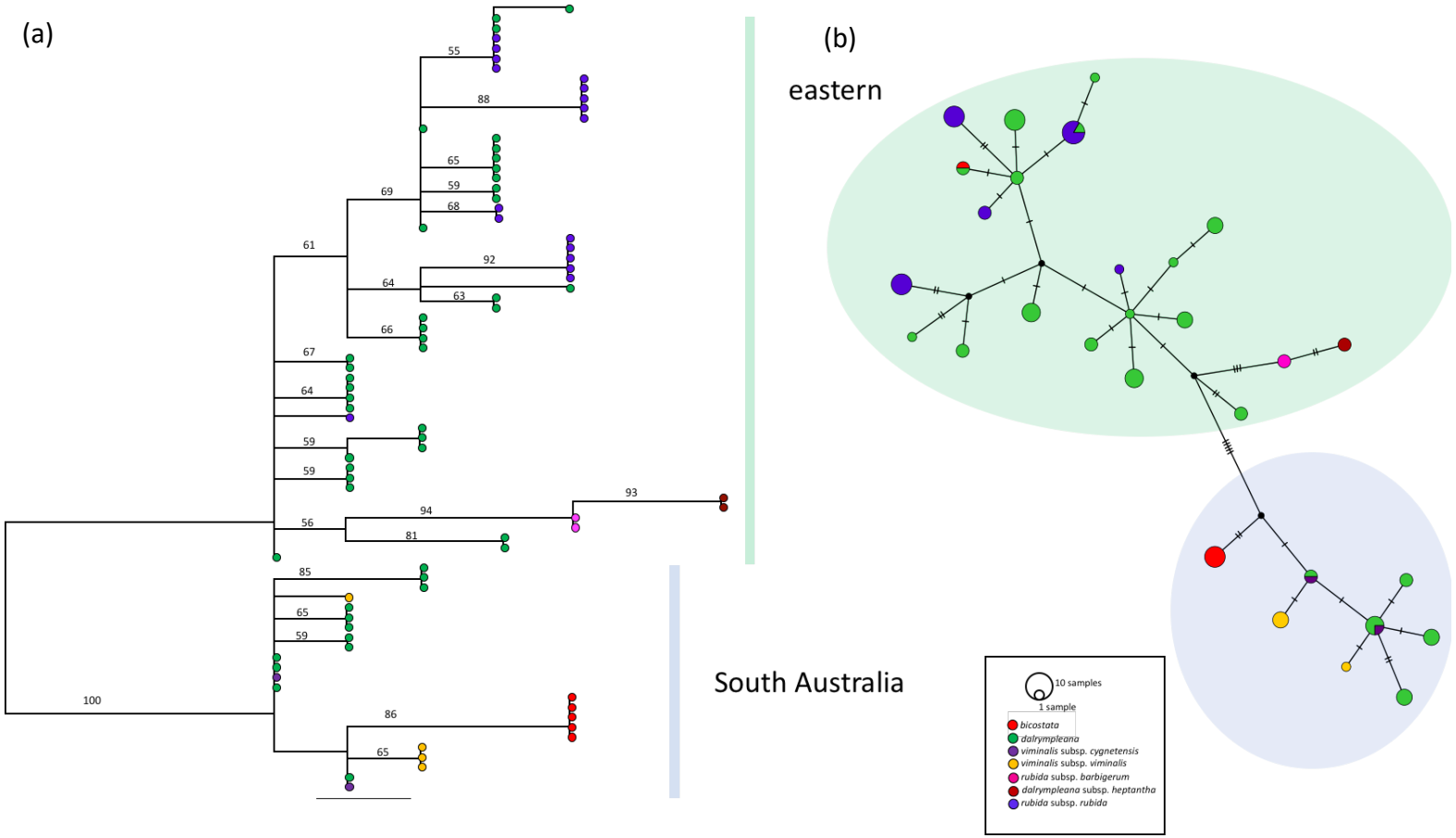

Figure 5: (a) Maximum likelihood phylogeny of chloroplast haplotypes and (b) minimum spanning haplotype network for the *E. dalrympleana* data set. In (a), branch lengths are proportional to the number of changes along that branch, numbers adjacent to the branch indicate support from 200 non-parametric bootstrap replicates and terminal branches are coloured according to the morphological species sampled. In (b), the size of the pie chart is proportional to the number of individuals sharing that haplotype, hash marks indicate mutational changes occurring along that branch, black dots represent unsampled haplotypes and haplotype pies are coloured as in (a). Samples from the east ('eastern') and the west ('South Australia') of the Lower Murray Basin are indicated.

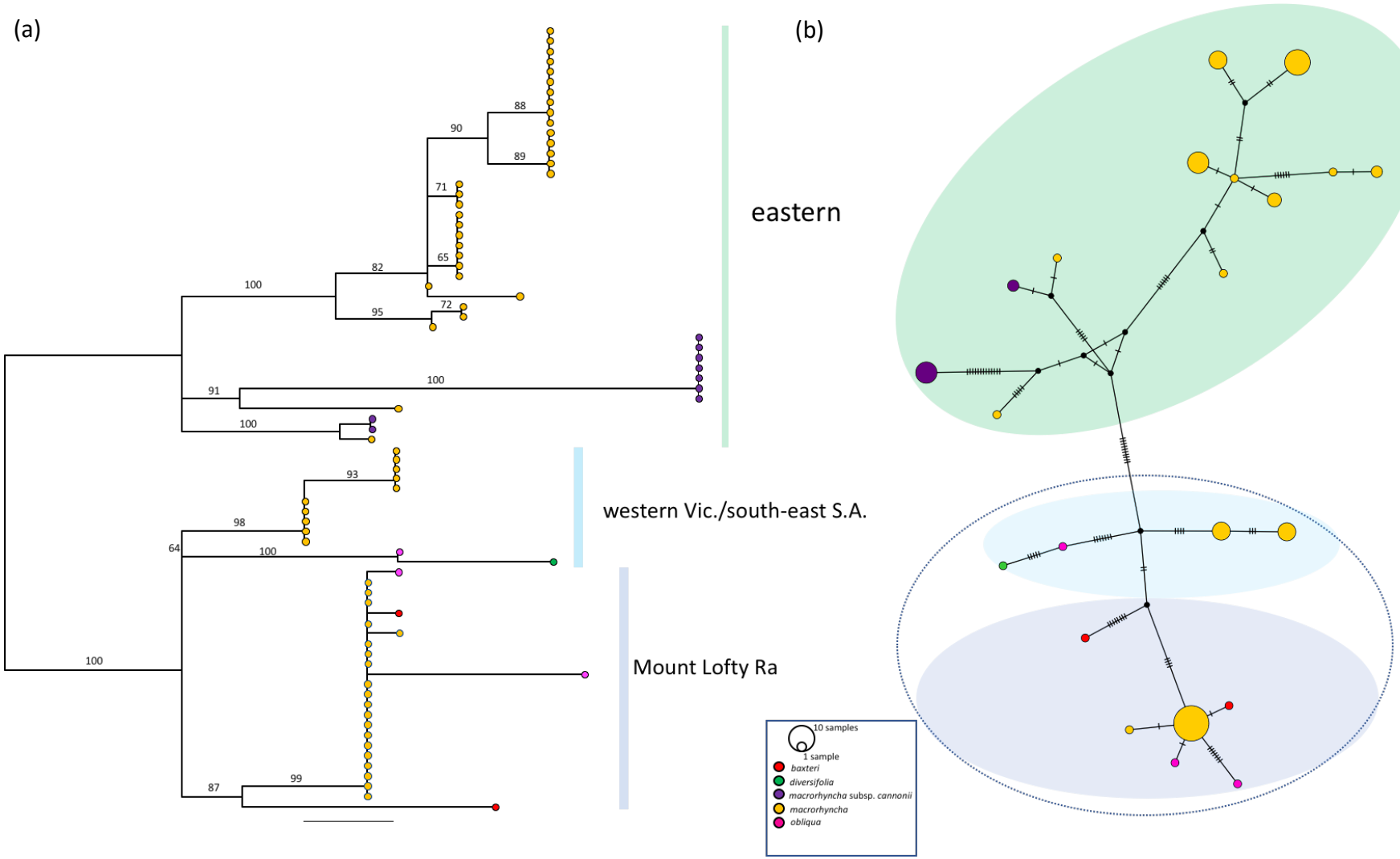

Figure 6: (a) Maximum likelihood phylogeny of chloroplast haplotypes and (b) minimum spanning haplotype network for the *E. macrorhyncha* data set. Three main groups are highlighted: an 'eastern' group that is distinct from 'Mount Lofty Ranges' + 'western Victoria/south eastern South Australia'.

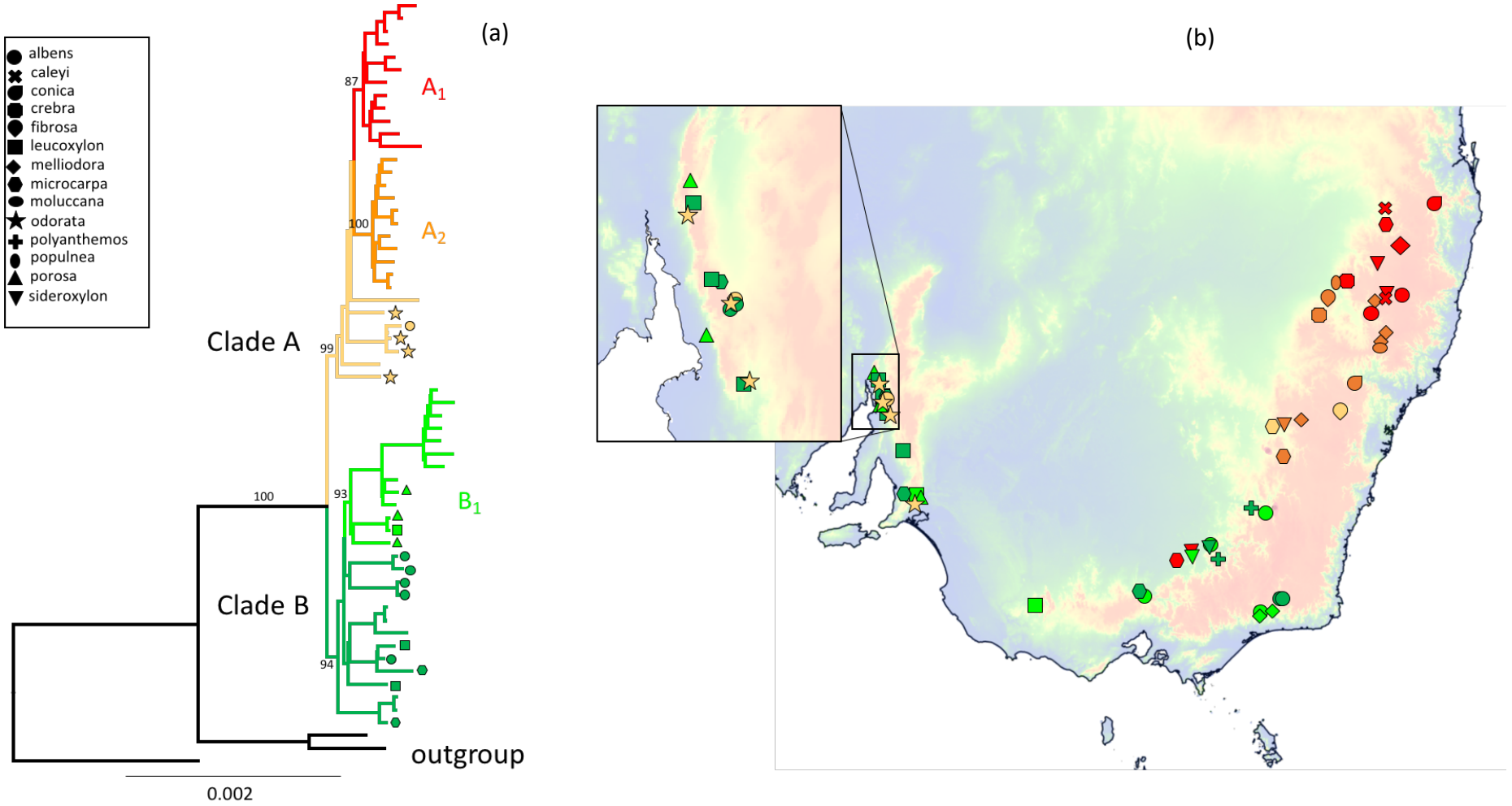

Figure 7: (a) Maximum likelihood phylogeny based upon whole chloroplast genomes for species of *Eucalyptus* section *Adnataria.* Bootstrap support for key groups is indicated adjacent to the branch and clades discussed in the text are highlighted. Symbols on terminal branches correspond to the morphological species assignment of individuals sampled from South Australia. In (b), the sampling location of individuals is indicated, along with the morphological species assignment with the colouring of symbols according to phylogenetic relationships indicated in 7(a).

relevant taxonomic section according to Thornhill et al.'s estimates. For Section *Maidenaria* (*E. dalrympleana*), the crown group is estimated to be of Miocene age (14 [10-19] Mya) which also marks the split between a Mount Lofty Ranges and an eastern Australian clade. For Section *Adnataria* (*E. albens*), the crown group age is estimated at 13 (10-17) Mya, including two main clades that are consistent with Clades A and B, above, both with a mixture of 'eastern' and 'South Australian' individuals (Figure 8). The crown group age of *E. macrorhyncha* is 22 (16-28) Mya, with an 'eastern' clade that is sister to a well-supported group including individuals sampled from Western Victoria (Stawell, Tooborac) and South Australia (Spring Gully) (Figure 8). The divergence of the western Victorian and South Australian clades is estimated at 20 (14-26) Mya.

Discussion

Patterns of genetic diversity among our three study species strongly support the hypothesis of long-term persistence for the populations occurring disjunctly in the Flinders-Mount Lofty region of South Australia. Our phylogenetic reconstructions using nDNA suggest that the South Australian populations diverged early in the evolution of each of the three species, and are resolved as sister to most, if not all of the conspecific populations sampled on the east coast (Figures 2-4). These phylogenetic patterns are consistent with vicariance whereas populations founded by long distance dispersal from a core eastern range would be more likely nested within an 'eastern' clade. With the respect to the plastid data, we recovered patterns of phylogenetic relationship that are strongly at odds with morphological species concepts (Figures 5-7), but more closely reflect the geographical origin of the samples, a pattern that has been widely reported in eucalypts (e.g. McKinnon et al. 2001, 2004; Flores-Rentería et al. 2017; Alwadani et al. 2019) and other plant lineages (e.g. *Nothofagus*; Premoli et al. 2011) from analyses of cpDNA. Nevertheless, these data are largely consistent with the evidence from nDNA in supporting the deep divergence of eucalypts in the Flinders-Mount Lofty Ranges from eastern populations but the estimated splitting times from cpDNA are up to orders of magnitude older than those inferred from the nDNA analyses (Figure 8).

Phylogeography from nDNA

Our divergence time estimates from nDNA (Table 5) are based upon several assumptions including the veracity of the mutation rate estimates used to scale our phylogenetic trees.

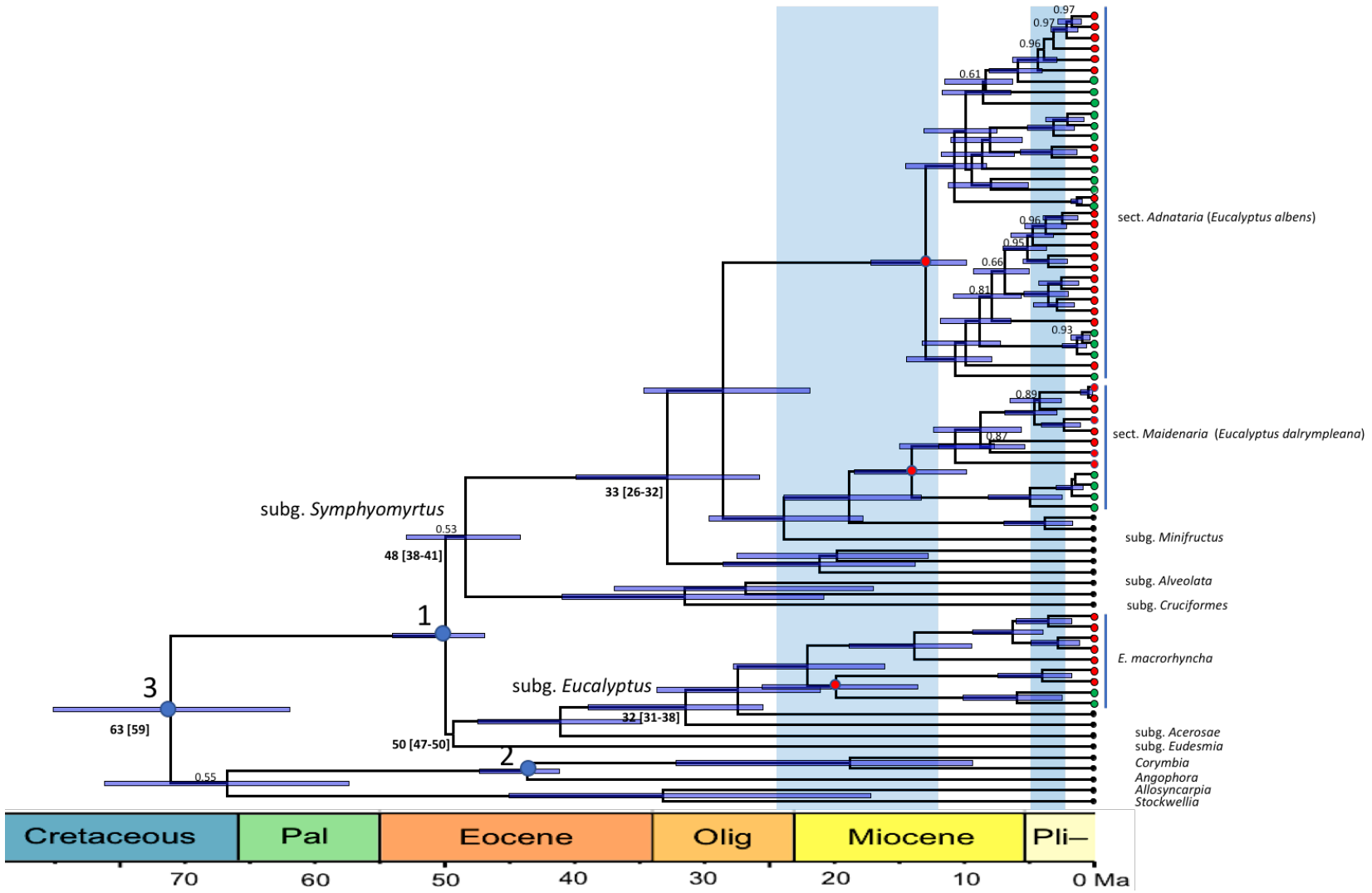
0.97
0.97
0.96
0.61
sect. Adnataria (Eucalyptus albens)
0.96
0.95
0.66
0.81
0.93
0.89
sect. Maidenaria (Eucalyptus dalrympleana)
0.87
33 [26-32]
subg. Symphyomyrtus
0.53
48 [38-41]
subg. Minifructus
subg. Alveolata
subg. Cruciformes
1
E. macrorhyncha
3
subg. Eucalyptus
32 [31-38]
63 [59]
50 [47-50]
subg. Acerosae
subg. Eudesmia
2
Corymbia
0.55
Angophora
Allosyncarpia
Stockwellia
Cretaceous
Pal
Eocene
Olig
Miocene
Pli–
70
60
50
40
30
20
10
0 Ma

Figure 8 (previous page): Maximum clade credibility tree estimated in BEAST using whole chloroplast genome sequences for representatives of tribe Eucalypteae. Branch lengths proportional to time (millions of years) and fossil calibration points are numbered (1-3, refer to text). Posterior probabilities < 1.0 are indicated adjacent to the branch, and branches receiving a posterior probability < 0.5 are dashed. Blue bars indicate the 95% HPD of divergence time estimates and for key nodes, divergence time estimates from Thornhill et al. (2019) (numbers in bold) are compared with the estimates in the present study (numbers in bold italics, bracketed). For each of the three eucalypt groups considered in this study, the terminal branches are coloured according to collection locality: green, west of the Murray Basin; red, east of the Murray Basin. For each group, the orange circle indicates the divergence event discussed in the text and the blue background shading indicates the approximate timing of major marine inundations of the Murray Basin (Stephenson and Brown, 1989).

While this is a source of uncertainty, we primarily aimed to determine whether the splitting times of the disjunct South Australian populations could be reasonably placed beyond the LGM. Given that this is likely the period of most inhospitable climate for mesic adapted plant species over the past 100 Ka, *in situ* survival through LGM would support the importance of the Flinders-Mount Lofty Ranges region as a refugium for those species. For each of our species, and accounting for phylogenetic uncertainty and a wide variation in molecular age estimates, divergence times exceed a minimum of c. 26 Ka, but are not inconsistent with the disjunct populations persisting through one or more full glacial cycles (Table 5), which have a periodicity of approximately 100 Ka with interglacials occurring at c. 120 Ka, 240 Ka and 340 Ka (Kawamura et al. 2007). Pollen records from south-western Victoria reveal cyclic shifts in vegetation composition through the region with distinct phases of forest and woodland expansion associated with interglacial periods, shifting to open grasslands, heath and herbfield communities through the driest glacial phases (Harle et al. 2004). More recent increases in eucalypt pollen have been documented from the Victorian Volcanic Plains, associated with increased precipitation and temperatures at c. 55 and 115 Ka (Cook, 2009). While there are no comparable fossil records from the Lower Murray river region, which is presently dominated by shrubland and discontinuous woodland communities, it is reasonable to hypothesise that migration of mesic adapted eucalypts across the lower Murray basin occurred during a warmer and wetter climatic phase pre-dating the LGM when forest vegetation was more widespread through the area.

cpDNA analyses

Incongruence between cpDNA and morphological species concepts has been ascribed to biological processes including incomplete lineage sorting, convergent evolution and introgressive hybridisation leading to chloroplast capture, i.e. repeated asymmetric gene flow resulting in species' capturing the chloroplast of other species (McKinnon et al. 2001). While these alternatives can be difficult to untangle, and may operate in concert, we suggest that chloroplast capture offers the most parsimonious explanation for the patterns recovered from our cpDNA data. In particular, lineage sorting is a stochastic process driven by genetic drift that would be expected to produce patterns of genetic diversity that are independent of geography (McKinnon et al. 2001, 2004; Alwadani et al. 2019) while are data show strong geographic structure. Convergent evolution, alone, also seems unlikely to explain the patterns of incongruence observed in the data given the number of species involved in each of our data sets and the large differences that separate chloroplast clades in different geographical regions (Figures 5-7). The propensity for hybridisation among eucalypt species is well documented, and both field observations and experimental crosses provide evidence for hybridisation potential between closely related species but extending to those placed in different taxonomic series and sections (Steane et al. 2015; Larcombe et al.). While incomplete lineage sorting and parallel evolution might explain aspects of our data, the observed pattern of geographically and phylogenetically related chloroplast types that cross morphological species boundaries suggests an important role for interspecific hybridisation and introgression.

Our molecular dating analyses of chloroplast genome data recovered a well-supported split between the South Australian and all (*E. dalrympleana*) or most (*E. macrorhyncha*) of the eastern samples (Figure 8), and this split was surprising deep given the divergence times estimated from nDNA (Table 5). In the case of the former, we estimated a divergence in the Miocene, while for the latter, a similar time-frame is inferred for the divergence of the Mount Lofty Ranges from individuals collected in western and south-western Victoria along with south-eastern South Australia, i.e. east of the Murray Darling Basin. Interestingly, our divergence time estimates for these splits may significantly exceed the age of the species concerned and for example, the most recent dated phylogeny of eucalypts (Thornhill et al. 2019) places the majority of species referred to *E.* subgenus *Symphyomyrtus* section *Maidenaria* (including *E. dalrympleana*) within a clade that is not more than c. 1 million years old, while *E. macrorhyncha* is resolved along with most species of *E.* subg. *Eucalyptus*

section *Capillulus* within a clade that radiated not more than c. 7 Mya. One the one hand, the discrepancies between analyses might reflect inaccuracies in molecular clock estimates, although our fossil constraints are consistent with those used by Thornhill et al. and generally returned comparable age estimates for deeper nodes (Figure 8). However, Thornhill et al. used penalised likelihood rate smoothing (r8s; Sanderson, 2003) to infer divergence times, which estimates branch lengths under a relaxed molecular clock from a pre-estimated (non-clock) tree topology, which was inferred from concatenated nuclear (internal and external transcribed spacers of 18S-26S rDNA) and 2 chloroplast loci. We suggest that the shallow divergences within sub-genera inferred by Thornhill et al. might be the result the higher variation in the nDNA relative to cpDNA with the former contributing disproportionately to the branch lengths of their non-clock trees and therefore to their divergence time estimates. In contrast, we used chloroplast genomes to co-estimate topologies and divergence times which may better reflect the evolutionary history of the chloroplast but not necessarily that of the included species.

Based upon simulations, it has been shown that during a range expansion into an occupied area, introgression occurs almost exclusively from the local to the invading species because the invading individuals will be uncommon relative to the local species, favouring matings with the latter. Furthermore, because the chloroplast is usually maternally inherited amongst angiosperms (i.e. via the seed parent), and seed dispersal is usually more limited than pollen dispersal (Booth, 2017), the extent of introgression is likely to be more pronounced for cpDNA than it is for nDNA (Currat et al. 2007). Therefore, when an expanding population encounters an established one, hybridisation and asymmetric introgression can lead to the replacement of the invaders chloroplast in favour of the local one. Indeed, it has been suggested that the earliest cpDNA haplotype to colonise an area could be inherited by subsequent invaders and persist over long evolutionary timescales (a 'ghost lineage' *sensu* Huang et al. 2014) contributing to strong geographic structuring in chloroplast diversity. In the context of our data, the clear geographic signal from the cpDNA along with the deep divergence between clades suggests a long term presence of eucalypt species in the Flinders-Mount Lofty Ranges and one or more cycles of range expansion and introgression among interbreeding species, favouring the persistence of established local haplotypes. In particular, divergence time estimated from the cpDNA are reasonably consistent with the timing of a major marine inundation across the Murray-Darling Basin (c. 24 -12 Ma; Stephenson and

Brown, 1989) (Figure 8) suggesting that an ancient vicariance is recorded in the chloroplast genes.

In the *Adnataria* data set, we recovered a more complex pattern than those outlined above. The analysis of chloroplast genome data recovered two major clades (clades A and B; Figure 7) including predominantly 'northern' and 'southern' sampled individuals, respectively, while in the Flinders-Mount Lofty region, individuals from both clades are represented. The divergence of these clades is estimated to have occurred during the Mid- to Late Miocene (Figure 8), consistent with the marine inundation of the Murray basin, which may have also effected the ancestral distribution of this lineage. Under this hypothesis, the initial divergence across the basin would be followed by a reconnection from the 'eastern' lineage, giving rise to a mixture of chloroplast lineages within the Flinders-Mount Lofty Ranges. Similarly, expansion from the west during periods of marine regression could generate a widespread distribution in the other major clade. In particular, present and presumably past climates would seem less of a barrier to the migration of chloroplast lineages in section *Adnataria* relative to our other eucalypt taxa given the broad range of ecological tolerances with several species (e.g. *E. leucoxylon, E. fasciculosa*) being more-or-less continuously distributed across the Murray Basin, while others (e.g. *E. populnea, E. largiflorens*) extend far onto the semi-arid riverine plains in western NSW (Costermans, 1981). However, alternative hypotheses could equally fit these data and for example, if lineages have migrated from the east as suggested above, we might expect the South Australian individuals to be nested within an eastern clade rather than resolved as sister lineages (Figures 7 and 8). More intensive sampling of the range of *Adnataria* is required to help distinguish several alternative scenarios (Alwadni et al. 2019).

Conservation implications

Understanding the phylogeographic patterns and current genetic connectivity of a species provides a greater understanding of evolutionary history and is important for conservation and management decisions (Moritz and Faith, 1998). Our findings indicate that in the *Eucalyptus* species studied, patterns of genetic diversity are likely tied to landscape level processes including long-term climatic change and variation in sea-level. In each of our study species, these processes have operated early in their evolution leading to disjunction and long-term isolation of South Australian lineages. It seems reasonable to assume that similar disjunctions, already reported (e.g. Freeman et al. 2001; Larcombe et al. 2011; French et al.

2016) or as yet unstudied, may be the result of the same set of processes highlighting the significance of the Flinders-Mount Lofty Ranges as a refugium for mesic adapted species.

While each of the eucalypt species included in our study has a broad geographic range, they are geographically restricted in South Australia and are listed as rare in that State (*NPWA 1972*). The results of our study unambiguously support the notion that the South Australian populations of each of these species harbour distinct genetic diversity and are worthy of conservation consideration, particularly in light of threats arising from factors such as historical and ongoing habitat fragmentation and the influence of anthropogenic climate change (Guerin and Lowe, 2013; Guerin et al. 2016). In fact, the level of habitat fragmentation is perhaps most severe in South Australia when compared to other parts of these species distributional ranges (Dore et al. 2000). A related issue, which may be worthy of further study, concerns the taxonomic status of our study species. In particular, the South Australian populations of *E. dalrympleana* are genetically distinct from populations on the east coast, as they are from *E. rubida*, suggesting this might be a new South Australian endemic taxon. Similarly, *E. macrorhyncha* subsp. *cannonii* appears to be more closely related to *E. macrorhyncha* subsp. *macrorhyncha* on the east coast than are the Spring Gully and Stawell populations of *E. macrorhyncha* suggesting the need for taxonomic realignment in this group.

Acknowledgements

We wish to thank the Australian Flora Foundation for financial support for this project. We also acknowledge the curators and staff of South Australian Herbarium (AD) for their assistance and for providing access to herbarium specimens. Thanks to Francis Nge for assistance with field work.

References

Alwadani, KG, Janes, JK and Andrew, RL. 2019. Chloroplast genome analysis of box-ironbark Eucalyptus Molecular Phylogenetics and Evolution 136: 76-86.

Bandelt H, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48.

Bayly MJ, Rigault P, Spokevicius A, Ladiges PY, Ades PK, Anderson C, Bossinger G, Merchant A, Udovicic F, Woodrow IE, Tibbits J (2013). Chloroplast genome analysis of Australian eucalypts – *Eucalyptus, Corymbia, Angophora, Allosyncarpia* and *Stockwellia* (Myrtaceae). *Molecular Phylogenetics and Evolution* 69: 704–716.

Bradbury B, Binks RM, Coates BJ, Byrne, M (2019) Conservation genomics of range disjunction in a global biodiversity hotspot: a case study of *Banksia biterax* (Proteaceae) in southwestern Australia, *Biological Journal of the Linnean Society* 127: 390–406.

Bennett, AF (2006) Eucalypts, wildlife and nature conservation: from individual trees to landscape patterns. The Royal Society of Victoria 128: 71–86

Booth TH (2017). Going nowhere fast: a review of seed dispersal in eucalypts. *Australian Journal of Botany* 65: 401-410.

Byrne M (2008) Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews* 27: 2576–2585.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D., Suchard MA, Rambaut ., Drummond AJ (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10: e1003537.

Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ, Cooper S, Donnellan SC, Keogh JS, Leys R, Melville J, Murphy DJ, Porch N, Wyrwoll KH (2008) Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology* 17: 4398–4417.

Cook EJ (2009) A record of late Quaternary environments at lunette-lakes Bolac and Turangmoroke, Western Victoria, Australia, based on pollen and a range of non-pollen palynomorphs. *Review of Palaeobotany and Palynology* 153:185–224.

Costermans L (1981) *Native trees and shrubs of south-eastern Australia*. Weldon Publishing, Sydney

Crisp MD, Laffan S, Linder HP, Monro A (2001), Endemism in the Australian flora. *Journal of Biogeography*, 28: 183-198

Currat M, Ruedi M, Petit R, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. Evolution 62:1908–1920.

Dore J, Binning C, Hayes G (2000) Native vegetation national overview: States/Territories/Commonwealth stocktake of native vegetation management. Department of Environment and Heritage, Commonwealth of Australia, Canberra.

Earl DA, VonHoldt BM (2012). SRTRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources* 4: 359-361.

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5: 113.

Evanno G, Ragnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.

Flores-Renteria L, Rymer PD, Riegler M (2017). Unpacking boxes: Integration of molecular, morphological and ecological approaches reveals extensive patterns of reticulate evolution in box eucalypts. *Molecular Phylogenetics and Evolution* 108, 70–87.

Freeman J, Jackson H, Steane D, McKinnon G, Dutkowski G, Potts B, Vaillancourt R (2001) Chloroplast DNA phylogeography of *Eucalyptus globulus. Australian Journal of Botany* 49: 585–596.

French PA, Brown GK, Bayly MJ 2016. Incongruent patterns of nuclear and chloroplast variation in *Correa* (Rutaceae): introgression and biogeography in south-eastern Australia. *Plant Systematics and Evolution* 302: 447–468.

Gandolfo MA, Hermsen EJ, Zamaloa MC, Nixon KC, González CC, Wilf P, Cúneo NR, Johnson KR (2011) Oldest known *Eucalyptus* macrofossils are from South America. PLoS One 6, e21084

Guerin GR, Lowe AJ (2013). Multi-species distribution modelling highlights the Adelaide Geosyncline, South Australia, as an important continental-scale arid-zone refugium. *Austral Ecology* 38: 427-435.

Guerin GR, Biffin E, Baruch Z, Lowe AJ (2016) Identifying centres of plant biodiversity in South Australia. *PLoS One* 11: e0144779.

Guindon S, Dufayard J, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307-321.

Harle KJ, Kershaw AP, Clayton E (2004). Patterns of vegetation change in southwest Victoria (Australia) over the last two glacial/interglacial cycles: evidence from Lake Wangoom. *Proceedings of the Royal Society of Victoria* 116: 107-139.

Hart ML, Forrest LL, Nicholls JA, Kidner CA (2016) Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65: 1081-1092.

Heled J, Drummond AJ (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27: 570–580.

Hope GS (1994) *Quaternary vegetation. History of the Australian vegetation: Cretaceous to Recent* (ed. by RS Hill), pp. 368–390. Cambridge University Press, Cambridge.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332

Huang DI, Hefer CA, Kolosova N, Douglas CJ, Cronk QC (2014) Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely

related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytologist* 204: 693-703.

Karanth KP (2003) Evolution of disjunct distributions among wet-zone species of the Indian subcontinent: testing various hypotheses using a phylogenetic approach. *Current Science* 85: 1276-1283.

Kawamura, K., Parrenin, F., Lisiecki, L. *et al.* Northern Hemisphere forcing of climatic cycles in Antarctica over the past 360,000 years. *Nature* **448,** 912–916 (2007). https://doi.org/10.1038/nature06015

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649.

Lande R, Engen S, Saether BE (2003). *Stochastic population dynamics in ecology and conservation*. Oxford University Press on Demand.

Larcombe MJ, McKinnon GE, Vaillancourt RE (2011) Genetic evidence for the origins of range disjunctions in the Australian dry sclerophyll plant *Hardenbergia violacea* . *Journal of Biogeography* 38: 125-136.

Larcombe MJ, Holland B, Steane DA, Jones RC, Nicolle D, Vaillancourt RE, Potts BM (2015) Patterns of reproductive isolation in *Eucalyptus*—a phylogenetic perspective. *Molecular Biology and Evolution* 32: 1833-1846.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014) Species delimitation using genome-wide SNP Data. *Systematic Biology* 63: 534– 542.

Li H, Handsaker B, Wysoker A, et al. (2009) The Sequence alignment/map format and SAMtools. *Bioinformatics*. 25: 2078-2079.

Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV (2009). Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics and Evolution* 53: 320–328.

McCallum KP, Guerin GR, Breed MF, Lowe AJ (2014), Combining population genetics, species distribution modelling and field assessments to understand a species vulnerability to climate change. *Austral Ecology* 39: 17-28.

McKinnon GE, Vaillancourt RE, Jackson HD, Potts BM (2001) Chloroplast sharing in the Tasmanian eucalypts. *Evolution* 55: 703–711.

McKinnon GE, Jordan GJ, Vaillancourt RE, Steane DA, Potts BM (2004) Glacial refugia and reticulate evolution: the case of the Tasmanian eucalypts. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359: 275–284

Moritz C, Faith DP. Comparative phylogeography and the identification of genetically divergent areas for conservation. Mol Ecol. 1998;7(4):419–29. WOS:000073797200006.

Müller NF, Ogilvie HA, Zhang C, Drummond AJ, Stadler T (2018) Inference of species histories in the presence of gene flow. *bioRxiv*, 348391.

Nicolle D (2013). *Native eucalypts of South Australia*. D. Nicolle, Adelaide.

Ogilvie, HA, Bouckaert, RR and Drummond, AJ. 2017. StarBEAST2 Brings Faster Species Tree Inference and Accurate Estimates of Substitution Rates, *Molecular Biology and Evolution*, 34: 2101–2114.

Petit RJ, Hampe A (2006). Some evolutionary consequences of being a tree. Annual Review of Ecology, Evolution and Systematics 37: 187-214.

Premoli AC, Mathiasen P, Cristina Acosta M, Ramos VA (2012) Phylogeographically concordant chloroplast DNA divergence in sympatric *Nothofagus s.s.* How deep can it be?. *New Phytologist*, 193: 261-275.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Rambaut A (2012). FigTree v1. 4. Available at http:// tree.bio.ed.ac.uk/software/figtree/.

Rambaut A, Suchard MA, Xie D, Drummond AJ (2015). Tracer v1. 6. 2014 MCMC Trace File Analyser. Available at http://beast.bio.ed.ac.uk/Tracer.

Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19: 301–302.

Silva-Junior OB, Grattapaglia D (2015). Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of *Eucalyptus grandis*. *New Phytologist* 208:830-845.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014 May 1;30(9):1312-3.

Steane DA, Nicolle D, Sansaloni CP, Petroli CD, Carling J, Kilian A, Myburg AA, Grattapaglia D, Vaillancourt R (2011). Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping. *Molecular Phylogenetics and Evolution* 59:206–224.

Stephenson AE, Brown CM (1989). The ancient Murray River system. *Journal of Australian Geology and Geophysics* 11: 397-395.

Thornhill AH, Macphail MK (2012) Fossil myrtaceous pollen as evidence for the evolutionary history of the Myrtaceae: a review of fossil *Myrtaceidites* species. *Review of Palaeobotany and Palynology* 176–177: 1–23.

Thornhill AH, Popple LW, Carter RJ, Ho SYW, Crisp MD (2012) Are pollen fossils useful for calibrating relaxed molecular clock dating of phylogenies? A comparative study using Myrtaceae. *Molecular Phylogenetics and Evolution* 63, 15–27.

Thornhill AH, Crisp MD, Külheim C, Lam KE, Nelson LA, Yeates DK, Miller JT (2019) A dated molecular perspective of eucalypt taxonomy, evolution and diversification. *Australian Systematic Botany* **32**, 29-48.

Worth JRP, Sakaguchi S, Harrison PA, Brüniche-Olsen A, Janes JK, Crisp MD, Bowman DMJS (2018). Pleistocene divergence of two disjunct conifers in the eastern Australian temperate zone, *Biological Journal of the Linnean Society* 125: 459–474

Weitemier K, Straub SC, Cronn RC, et al. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Science* 2:apps.1400042.

Zellmer AJ, Hanes MM, Hird SM, Carstens BC (2012) Deep phylogeographic structure and environmental differentiation in the carnivorous plant *Sarracenia alata*. *Systematic Biology* 61: 763–777