The genetic and demographic impacts of contemporary disturbance regimes in mountain ash forests

by

Brenton von Takach Dukai

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy
of the Australian National University
September 2019
Candidate's Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the author’s knowledge, it contains no material previously published or written by another person, except where due reference is made in the text.

Brenton von Takach Dukai  
Date: 09 September 2019
Acknowledgements

The aptly-titled Outstanding Professor Sam Banks has a wonderfully scientific mind. So wonderful, in fact, that his effortlessly casual approach of highlighting the flaws in my logic must be very amusing to witness. Throughout this project he would regularly and gently illustrate how my deepest thinking was just some small piece of jetsam, foundering feebly in the shallow waters of intellectual capacity. And without this constant improvement, or rather complete overhaul, of my thoughts and ideas, this work would definitely be but a shadow of what it is now.

Professors David Lindenmayer, Justin Borevitz, and Rod Peakall have also helped greatly to manage my loosely assembled plans and ideas. I don’t know if David actually goes to sleep at night, but one day I will make a histogram of his email times to prove that he never stops working. Justin, with his wonderful ability to see the (eucalypt) forest for the trees, would infallibly place some small hang-up of mine into perspective and push me into understanding, and more importantly writing, about the bigger implications of my data. And Rod, who is possibly the kindest and wisest of all professors, often bestowed upon me his form of casual yet philosophically revealing advice. This, along with his attention to detail and absolute knowledge of genetic analysis have certainly helped to guide and improve my research outcomes.

Many people gave up their valuable time to provide early insight into ecological processes that my inferior mind had (=has) yet to properly comprehend, and I cannot list them all. However, Geoff Cary, Annabel Smith, Malcolm Gill, Carsten Külheim and Shannon Dillon deserve mention in this regard. Geoff also managed to make me laugh by feigning to drink a beer every time he saw me rummaging through a freezer for misplaced tissue samples.

Wade Blanchard got me out of (quite a few!) statistical conundrums, and his calm, understanding demeanour and unsurpassed advice regarding all aspects of ecological modelling was extremely helpful when complex datasets were doing their best to irritate me.

In the field, I am grateful to have been assisted by an array of fantastic people, including Oliver King, Bec Taylor, Joe Salmona, Donna Belder, Lachie McBurney, and Dave Blair. You not only helped greatly with data collection, but also allowed me to see the humour in unforeseen circumstances.

In our charmingly old Forestry Building laboratory, with its highly skewed sex-ratio, I was fortunate enough to spend time filling pipette-tip boxes and autoclaving things with a number of brilliant women, including Robyn Shaw, Katie Moon, Georgeanna Story, Laura Wilson, Amanda Edworthy, Elle Bowd, Kristen Abicair, April Suen, and Xenia Weber. For
their assistance with mindless and repetitive preparation of my samples and lab equipment, Bec Taylor, Damian Lettoof, and Yuzun Guo deserve a big thankyou. You saved me weeks of depressingly solitary labwork. And in other labs around ANU, Niccy Aitken, Stephanie McCaffery, and Ursula Hurley were all remarkably patient in providing me with access to and assistance with equipment and protocols.

Aside from my research, there are a number of people who involved me in their lives over the last four years, for which I am thankful. Phil Gibbons provided a much-needed distraction and some extra income by taking me on as a demonstrator for a couple of his extremely enjoyable and enlightening coursework subjects. Kelly Dixon, Ross Crates, and Mark Allen need mention, for their raw dedication to both birds and beers, and the many enjoyable times we spent appreciating the finer details of those subjects. Kelly, forever the stalwart colleague, even followed me to Darwin to maintain my supply of cold schooners and listen to my assorted ravings…

For providing me with love and support during my time in Canberra, a special thankyou must go to Bec Taylor – you will always have a place in my heart. The friendship of Damian Lettoof was also a lifesaver, and were it not for his regular company I may well have gone insane. Similarly, for their generosity, friendship, and disarmingly crude humour during my time in Darwin, Chris Jolly and Alana de Laive will forever have my endless gratitude.
Preface

This thesis is presented as six chapters, including a general introduction, four data-driven chapters, and a synthesis. Figures and tables have been incorporated into each chapter after the appropriate section, and literature cited is provided at the end of each chapter (with a consolidated list of references provided at the end of the thesis). Chapters 2, 3, 4 and 5 are intended for publishing in peer-reviewed journals, and as a result there is a small amount of unavoidable repetition across chapters with respect to background information, study sites and laboratory methods. An additional appendix chapter is written in the style of a manuscript paper but may not be submitted for publishing. At the time of thesis submission, Chapter 2 has been published in *Forest Ecology and Management*, Chapter 3 has been published in *Evolutionary Applications*, Chapter 4 is in press at *Conservation Genetics*, and Chapter 5 is in review at *Molecular Ecology*. To maintain consistency between published articles and their corresponding thesis chapters, I have retained some minor formatting differences that have resulted from particular journal requirements (e.g. prioritisation of scientific names in Chapter 3). While I designed the research, conducted the fieldwork, performed the statistical analyses and authored all chapters, the co-authors on each manuscript have made a variety of contributions to the research. Thus, I have used the pronoun “we” in some chapters, and detail here the specifics of the contributions made by all co-authors.

**Chapter 2: Environmental influences on growth and reproductive maturation of a keystone forest tree: Implications for obligate seeder susceptibility to frequent fire**

I designed the study with advice from Sam Banks. I collected all field data. I analysed the data with assistance from Sam Banks. I authored the chapter, with input from David Lindenmayer and Sam Banks regarding structure and framing. All authors contributed to editing this chapter.

**Chapter 3: Pervasive admixture between eucalypt species has consequences for conservation and assisted migration**

I designed the study with advice from Sam Banks and Justin Borevitz. I collected all field data and performed all sample preparation for sequencing. Cameron Jack assisted with the bioinformatic pipeline, and I analysed the sequencing data with assistance from Sam Banks and Justin Borevitz. I authored the chapter, with input from Justin Borevitz, Sam Banks and David Lindenmayer regarding structure and framing. All authors contributed to editing this chapter.
Chapter 4: The influence of fire and silvicultural practices on the landscape-scale genetic structure of an Australian foundation tree species

I designed the study with advice from Sam Banks. I collected all field data and performed all sample preparation for microsatellite sizing and genotyping-by-sequencing, with sequencing machines run by laboratory technicians. I analysed the genetic data with assistance from Sam Banks and Rod Peakall. I authored the chapter, with input from Rod Peakall, Sam Banks and David Lindenmayer regarding structure and framing. All authors contributed to editing this chapter.

Chapter 5: Scale-dependent signatures of local adaptation in a widespread foundation tree species

I designed the study with advice from Sam Banks. I collected all field data and performed all sample preparation for genotyping-by-sequencing, with sequencing conducted by laboratory technicians. I analysed the genetic data with assistance from Sam Banks. I authored the chapter, with input from Sam Banks and David Lindenmayer regarding intellectual content, structure and framing. Sam Banks and David Lindenmayer contributed to editing this chapter.

Appendix 1: Investigating patterns of seed dispersal in mountain ash (Eucalyptus regnans) by genotyping fire-killed adults and regenerating saplings

I designed the study with advice from Sam Banks. I collected all field data with assistance from Sam Banks. I performed all sample preparation for next-generation sequencing and microsatellite sizing, with sequencing machines run by laboratory technicians. I analysed the genetic data with advice from Sam Banks. I authored the section, with contributions from Sam Banks to the editing process.
Thesis Abstract

Anthropogenic influences are modifying the natural disturbance regimes of forest ecosystems globally. Timber harvesting, frequent wildfires and a changing climate are influencing ecosystem composition, structure and function, with resulting losses to biodiversity and economic indicators. In south-eastern Australia, these factors are causing the rapid ecosystem collapse of montane forests, where foundation species rely on particular fire regimes for regeneration and maintenance of ecosystem integrity. Here, I characterise and quantify the demographic and genetic impacts that changing environments are having on mountain ash (*Eucalyptus regnans*), a foundation species and one of the world’s tallest trees.

Plant species that rely on particular fire regimes are at risk of population decline when fire frequency and intensity changes. To test whether mountain ash populations exhibit variation in susceptibility to increasing fire frequency, I investigated the response of growth rates and seed production to stand age under different environmental conditions. My results show that environmental factors determine the age of maturation, in turn affecting the time taken for populations to develop reproductively viable amounts of seed. This suggests that reduced fire return intervals may result in niche contractions of obligate seeders such as mountain ash.

Accurate information on population genetic structure is required to develop strategies for ecological restoration and reintroduction programs. I conducted a range-wide analysis of mountain ash population genetic structure to determine the extent of hybridisation with messmate stringybark (*Eucalyptus obliqua*), and investigate how genetic diversity parameters are influenced by hybridisation. Hybrid occurrence was not distributed evenly across environmental gradients or populations, and, after accounting for admixture, mountain ash showed very little population genetic structure, with a small effect of isolation-by-distance and low global $F_{ST}$ (0.03). This suggests that decisions around provenancing for restoration may depend on knowledge of how admixture influences population genetic structure, and that for some species there may be little benefit in planning conservation strategies around environmental adaptation of seed sources.

Next, I examined whether fire and silvicultural practices are modifying patterns of within- and among-population genetic diversity and fine-scale spatial genetic structure of mountain ash in the Central Highlands of Victoria. As chloroplast DNA and nuclear DNA are dispersed via different mechanisms, manual sowing of logged sites using non-local seed is likely to have differing effects on these two genomes. To test this, I utilised chloroplast microsatellites and genome-wide single-nucleotide polymorphisms to compare genetic
parameters between undisturbed, burnt, and logged stands. The patterns and extent of genetic diversity and genetic differentiation among stands at nuclear loci were largely similar among disturbance histories, but chloroplast microsatellites revealed significantly higher levels of genetic diversity in logged stands. This suggests that logging is having minor impacts on the nuclear genome but large impacts on the chloroplast genome, with haplotypes entering the system via the use of non-local seed in the regeneration process.

Understanding the patterns and drivers of local adaptation is important for the conservation management of foundation species in rapidly-changing environments. I investigated the genome of mountain ash for signatures of local adaptation at a regional and range-wide spatial scale, using three methods at each scale to identify outlier loci. I found 30 loci in the range-wide dataset and 10 loci in the regional dataset that were significantly associated with environmental variables, with very little overlap between spatial scales. The results demonstrate that investigation of multiple spatial scales provides a greater understanding of adaptive genetic variation, and are consistent with literature suggesting that local adaptation in forest trees occurs primarily through small-effect loci acting on polygenic traits.

This thesis provides novel insights into the demographic and genetic impacts that modified disturbance regimes have on an ecologically and economically important foundation tree species. I found geographic variation in vital rates; population genetic patterns suggestive of high pollen flow and limited seed dispersal (indicating low recolonisation ability); high levels of within-population genetic variation (implying high adaptive capacity); frequent hybridisation with a fire-tolerant congener (suggesting potential for adaptive introgression); minor impacts of silviculture on nuclear genetic diversity; and different genomic architectures of local adaptation at different spatial scales. Importantly, furthering our understanding of all these aspects will assist with the conservation management of foundation tree species and forest ecosystems, and contribute to the maintenance and/or maximisation of adaptive capacity, allowing forests to persist into a changing and uncertain future.
# Table of Contents

Candidate's Declaration .................................................................................................................... ii  
Acknowledgements .......................................................................................................................... iii  
Preface ................................................................................................................................................... v  
Thesis Abstract ....................................................................................................................................... vii  
Table of Contents .................................................................................................................................... ix  

## Chapter 1: Introduction .................................................................................................................. 1  
1.1 Forest systems in the Anthropocene ........................................................................................... 1  
1.2 Disturbance and genetic diversity ............................................................................................... 3  
1.3 The eucalypts ............................................................................................................................. 5  
1.4 Mountain ash .............................................................................................................................. 6  
1.5 Disturbance in mountain ash forests ............................................................................................ 17  
1.6 Thesis aims ..................................................................................................................................... 22  
1.7 References ................................................................................................................................. 27  

## Chapter 2: Environmental influences on growth and reproductive maturation of a keystone forest tree: implications for obligate seeder susceptibility to frequent fire... 34  
2.1 Abstract ......................................................................................................................................... 35  
2.2 Introduction ................................................................................................................................... 36  
2.3 Materials and methods ............................................................................................................... 40  
2.4 Results .......................................................................................................................................... 46  
2.5 Discussion ..................................................................................................................................... 58  
2.6 Conclusions ................................................................................................................................. 62  
2.7 Acknowledgements ..................................................................................................................... 63  
2.8 References ..................................................................................................................................... 64  

## Chapter 3: Pervasive admixture between eucalypt species has consequences for conservation and assisted migration ........................................................................ 69  
3.1 Abstract ......................................................................................................................................... 70  
3.2 Introduction ................................................................................................................................... 71  
3.3 Materials and methods ............................................................................................................... 74  
3.4 Results .......................................................................................................................................... 83  
3.5 Discussion ..................................................................................................................................... 92  
3.6 Conclusion .................................................................................................................................... 98  
3.7 Acknowledgements ..................................................................................................................... 99  
3.8 Data archiving ............................................................................................................................. 100  
3.9 References ..................................................................................................................................... 101  
3.10 Supporting information .............................................................................................................. 110
Chapter 4: The influence of fire and silvicultural practices on the landscape-scale genetic structure of an Australian foundation tree species

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Abstract</td>
<td>116</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>117</td>
</tr>
<tr>
<td>4.3 Materials and methods</td>
<td>120</td>
</tr>
<tr>
<td>4.4 Results</td>
<td>127</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>139</td>
</tr>
<tr>
<td>4.6 Conclusions</td>
<td>144</td>
</tr>
<tr>
<td>4.7 Acknowledgements</td>
<td>145</td>
</tr>
<tr>
<td>4.8 References</td>
<td>146</td>
</tr>
<tr>
<td>4.9 Supplementary material</td>
<td>152</td>
</tr>
</tbody>
</table>

Chapter 5: Scale-dependent signatures of local adaptation in a widespread foundation tree species

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Abstract</td>
<td>157</td>
</tr>
<tr>
<td>5.2 Introduction</td>
<td>158</td>
</tr>
<tr>
<td>5.3 Materials and methods</td>
<td>161</td>
</tr>
<tr>
<td>5.4 Results</td>
<td>169</td>
</tr>
<tr>
<td>5.5 Discussion</td>
<td>182</td>
</tr>
<tr>
<td>5.6 Conclusion</td>
<td>186</td>
</tr>
<tr>
<td>5.7 Acknowledgements</td>
<td>187</td>
</tr>
<tr>
<td>5.8 References</td>
<td>188</td>
</tr>
<tr>
<td>5.9 Supplementary material</td>
<td>194</td>
</tr>
</tbody>
</table>

Chapter 6: Synthesis

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Findings and implications</td>
<td>202</td>
</tr>
<tr>
<td>6.2 Consequences of disturbance</td>
<td>206</td>
</tr>
<tr>
<td>6.3 Forest trees and climate change</td>
<td>208</td>
</tr>
<tr>
<td>6.4 Future directions</td>
<td>210</td>
</tr>
<tr>
<td>6.5 Concluding remarks</td>
<td>213</td>
</tr>
<tr>
<td>6.6 References</td>
<td>214</td>
</tr>
</tbody>
</table>

Consolidated references

Appendix 1: Testing contemporary genotyping methods in fire-killed eucalypts

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1.1 Abstract</td>
<td>244</td>
</tr>
<tr>
<td>A1.2 Introduction</td>
<td>245</td>
</tr>
<tr>
<td>A1.3 Materials and methods</td>
<td>248</td>
</tr>
<tr>
<td>A1.4 Results</td>
<td>253</td>
</tr>
<tr>
<td>A1.5 Discussion</td>
<td>259</td>
</tr>
<tr>
<td>A1.6 Conclusion</td>
<td>261</td>
</tr>
<tr>
<td>A1.7 Acknowledgements</td>
<td>262</td>
</tr>
<tr>
<td>A1.8 References</td>
<td>263</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Forest systems in the Anthropocene

While large disturbances in the Earth’s ecosystems can appear very destructive, many ecological communities have evolved with a particular frequency and intensity of historical disturbance events, called the disturbance regime. In forest ecosystems, some of the most common naturally-occurring disturbances are wildfires, wind storms, and insect attacks (Foster and Boose, 1992; Seidl et al., 2014; White and Pickett, 1985). Unfortunately, anthropogenic influences are modifying the natural disturbance regimes of forest ecosystems around the world at unprecedented rates (de Groot et al., 2013; Kelly et al., 2013; Liu et al., 2012; Michaelian et al., 2011; Seidl et al., 2016; Ayers and Lombardero, 2000). These influences are altering the spatial distribution, composition, structure and function of forests globally, with resulting losses to both biodiversity and economic value (Laurance et al., 2014; Hanewinkel et al., 2012; Bradshaw et al., 2009; Edwards et al., 2014; Foster et al., 1998; Peres et al., 2006).

On short timescales, human-induced changes to disturbance regimes can cause shifts from one ecosystem state to an alternative state, leading to what are known as abrupt or rapid shifts (Scheffer et al., 2001; Scheffer and Carpenter, 2003; Olszewski, 2015). Australian ecosystems are no exception to these shifts in ecosystem state, and there is evidence to suggest that ecological communities in Australia may be just as threatened by abrupt shifts in ecosystem state as those overseas (Lindenmayer and Sato, 2018). For example, the montane forests of Australia’s southeast may be susceptible to rapid and widespread ecosystem collapse – a situation that is being driven by anthropogenic modifications to the pre-European fire regime and is exacerbated by climate change (Lindenmayer et al., 2011; Bowman et al., 2014; Burns et al., 2015). Such examples of modified fire regimes are known from many other regions of the planet (Pausas and Keeley, 2014a), and are an example of how humans can cause permanent and large-scale alterations to exploited natural environments. This presents a real challenge to management agencies, who typically have a responsibility to maximise ecosystem viability and may not have the resources to address the drivers of modified disturbance regimes.

Understanding species-level responses to changing disturbance regimes and climatic conditions is critical for the conservation of forest ecosystems. This is particularly true for
keystone and foundation species that have strong functional and structural roles. Foundation tree species in forests around the world are at risk from global change, with tree mortality, shifting species composition, shifting climatic niches, and invasions and/or attacks of pathogens observed (Bussotti et al., 2015; Peñuelas et al., 2013; Alberto et al., 2013). Developing our knowledge around the extent of these impacts and the adaptive responses made by species will help to maximise the efficiency of conservation and management actions, especially when limited resources are available.
1.2 Disturbance and genetic diversity

Disturbance regimes are critical components of many ecosystems, affecting both the patterns and extent of spatial and temporal heterogeneity in habitat suitability, which in turn control population and community dynamics (Sousa, 1984; Turner, 2010b). It is not surprising then that disturbance is also a major driver of the patterns and distribution of genetic diversity observed in natural populations (Banks et al., 2013). With changing frequencies and intensities of disturbances in forest ecosystems, there is not only flow-on effects on species abundance and community composition (Clarke, 2002b; Alfaro et al., 2014), but also on the spatial distribution of genetic diversity, through either selective processes or selectively-neutral demographic processes (Banks et al., 2013). Selective processes modify genetic structure by selecting for particular behavioural or physiological traits (Clarke, 2002a), whereas selectively-neutral processes modify genetic structure through direct effects (e.g. through population bottlenecks or redistribution of individuals) or indirect effects (e.g. manipulation of habitat suitability or permeability) (Apodaca et al., 2013; Brown et al., 2013; Templeton et al., 2011).

Life history traits and recolonisation pathways after disturbance events strongly influence the genetic structuring and diversity of populations. For example, some plant species maintain high levels of genetic diversity despite disturbances causing very high mortality rates, because recruitment of the new cohort is driven by soil-stored seed banks (Ayre et al., 2009; Dillon et al., 2015). Conversely, when disturbance-induced mortality is high and recruitment is driven by rare, stochastic colonisation events, such as in the Australian mallee emu-wren (*Stipiturus mallee*), allelic diversity can be severely reduced (Brown et al., 2013).

The level of standing genetic variation in a population is typically the primary source of adaptive potential in a changing environment, as opposed to novel allelic variations (Hoffmann and Sgrò, 2011; Alberto et al., 2013; Savolainen et al., 2013). Thus, it is critical that we determine how modified disturbance regimes influence the spatial distribution of genetic diversity, particularly for foundation species that have a strong role in structuring ecosystems. For species such as forest trees, loss of adaptive potential and resulting reductions in fitness under changing disturbance regimes and environmental conditions could lead to substantial changes in ecosystem structure and state (Lindenmayer and Sato, 2018).

Forest trees can mitigate the impacts of changing environments in three primary ways, including (1) acclimatisation and phenotypic plasticity, (2) local adaptation through changing
allele frequencies, and (3) migration to more suitable climates (Bussotti et al., 2015). To determine whether these species have the adaptive capacity to respond to rapid environmental changes, it is important that we develop our understanding of these adaptive pathways. This will allow us to identify which species are likely to persist in the long-term, and which species require anthropogenic intervention to prevent population declines and resulting losses to biodiversity. For species that do not have the capability to match the current rate of environmental change, it may then be possible to increase their adaptive capacity, for example through assisted migration, whereby species are translocated into new habitable locations (McLane and Aitken, 2012), or assisted gene flow, whereby the genetic diversity of populations is increased through the transfer of germplasm (Aitken and Bemmels, 2016; Aitken and Whitlock, 2013).
1.3 The eucalypts

“Trees about 300 feet high are known in Victoria, and huge in girth and straight in trunk they tower into the sky, affording little shade from their foliage because the scanty crown of leaves is so far removed from the earth.”

- Maiden (1905)

The family Myrtaceae is a diverse group of dicotyledonous woody plants that includes a number of well-known, speciose Australian genera such as the tea-trees (*Leptospermum*), paperbarks (*Melaleuca*), and bottlebrushes (*Callistemon*). Within the Myrtaceae is the tribe Eucalypteae, a group referred to as eucalypts. The taxonomy of the Eucalypteae has been variously circumscribed and debated in the past (Pryor and Johnson, 1971; Chippendale, 1988; Hill and Johnson, 1995; Brooker, 2000; Ladiges and Udovicic, 2000); however, until further clarification, it is thought to contain seven genera, consisting of four small rainforest genera *Allosyncarpia*, *Stockwellia*, *Eucalyptopsis*, and *Arillastrum*, and three larger sclerophyllous genera *Eucalyptus*, *Angophora* and *Corymbia* (Ochieng et al., 2007; Bayly et al., 2013). With around 700 species (Bayly, 2016), *Eucalyptus* (from the Greek roots *eu* and *calyptos* meaning ‘well’ and ‘covered’) is the most diverse of these genera. Species in this group are a fundamental component of the structure and composition of most Australian vegetation communities, providing an array of critically important ecosystem functions whilst dominating and typifying the Australian landscape.

With few exceptions, eucalypts are broad-leaved evergreen plants that have many similarities in growth, development, structure and function to woody dicotyledonous genera on other continents. However, the radiation of eucalypts, particularly the genus *Eucalyptus*, across Australia has resulted in a huge variety of growth habits and many unique specialisations in morphology, physiology and ecology (Pryor, 1976). One such species of high economic and ecological value is mountain ash (*Eucalyptus regnans*), a foundation species and tall forest tree that grows in mid-elevation wet forests of the south-eastern Australian states of Victoria and Tasmania. Forests dominated by mountain ash are considered critically endangered, with wildfire and timber harvesting contributing to the collapse of the ecosystem (Lindenmayer and Sato, 2018; Burns *et al.*, 2015). As the focal species of this thesis, the following section details our current understanding of mountain ash biology and ecology.
1.4 Mountain ash

1.4.1 Taxonomy

The genus *Eucalyptus*, *sensu* Brooker (2000), is broken into subgeneric taxa in the descending order of subgenus, section, and series. Mountain ash is placed within the subgenus *Eucalyptus* and the section *Eucalyptus*, and forms the series *Regnantes* with its sister species, brown barrel (*Eucalyptus fastigata*). Shared morphological traits between the two species include their tall, straight-trunked growth habit, thin glossy green leaves, and the distinctive feature of paired inflorescences in the axils of leaves (Brooker and Kleinig, 2006). The two species also grow in similar ecosystem types and on similar landforms.

Mountain ash was originally considered a form or variety of black peppermint (*Eucalyptus amygdalina*), a species first described in 1806 and now known to be restricted to Tasmania. Ferdinand von Mueller, the vice-president of the Victorian Acclimatisation Society at the time, first suggested that mountain ash might be a unique species in one of the society’s annual reports (von Mueller, 1871), writing:

“In our sheltered springy forest glens attaining not rarely a height of over 400 feet, there forming a smooth stem and broad leaves, producing also seedlings of a foliage different to the ordinary state of *E. amygdalina* as occur in more open country. This species or variety, which may be called *Eucalyptus regnans*, represents the loftiest tree in the British territory, and ranks next to the Sequoia Wellingtonia in size anywhere on the globe. …. Labillardière’s name [E. amygdalina] applies ill to any of the forms of this species.”

Mountain ash typically grows as a tall, single-stemmed, straight-trunked tree. Generally, there is a basal stocking of rough bark extending to a variable height on the stem, above which smooth or ‘gum’ bark is present to the crown branches. The rough basal bark may be variable in appearance but has been described as subfibrous to ‘peppermint’ type (Ashton, 1956).

In a number of locations throughout Victoria, trees displaying intermediate characteristics between mountain ash and other species have been recorded. These have been identified as hybrid individuals between either mountain ash and red stringybark (*E. regnans* x *macrorhyncha*), or mountain ash and messmate (*E. regnans* x *obliqua*) (Ashton and Sandiford, 1988; Ashton, 1958). At least two individuals have also been found that appear to be tri-hybrids – the result of a mountain ash-messmate hybrid individual mating with a red stringybark (Yorke and Ashton, 1982). Whilst at least one hybrid with narrow-leaved
peppermint (*Eucalyptus radiata*) is known (Brett Mifsud, pers. comm. 2015), mountain ash may also potentially hybridise with alpine ash (*E. delegatensis*) and silvertop ash (*E. sieberi*) (Ashton, 1956). It is also possible that intermediate forms with the closely related brown barrel are present where their ranges overlap to the south of Delegate (near the Victorian/New South Wales border). The extent of hybridisation with messmate and the conditions under which hybridisation occurs is examined in detail in Chapter 4.

### 1.4.2 Distribution and ecological niche

“Having reached the bitterly cold slopes of Mt Baw Baw, the group conducted one hasty search and managed to measure just one living tree before the appalling conditions made them turn tail and head for safety.”

- Beale (2007)

Mountain ash is patchily distributed in the south-eastern Australian states of Victoria and Tasmania (Brooker and Kleinig, 2006). In Victoria it is found on the southern side of the Great Dividing Range, from the mountainous areas of East Gippsland west to the Otway Ranges, with the main distribution across in the Central Highlands. In Tasmania, it is found across a large portion of the northern hinterland, as well as most of the south-eastern corner of the state from about Lake King William in the northwest to Southport and Bruny Island in the south.

Monotypic stands of mountain ash are typically found in areas that have high (mean value of 1280 mm) rainfall, deep fertile soils (e.g. red ferrosols), and are sheltered from cold or strong and hot winds (Cochrane, 1969; Lindenmayer *et al*., 1996). As the species is rather drought intolerant, it requires large amounts of available water to maintain water balance and associated shoot extension in summer (Florence, 1996). Mostly these conditions are thought to be met on slopes with southerly or easterly aspects, with mountain ash forest transitioning to a messmate (*E. obliqua*) or silvertop ash dominated vegetation community as the aspect changes to become more northerly or westerly (Lindenmayer, 2009a). However, stands of mountain ash may occur on any aspect so long as the above conditions are met (Cochrane, 1969). Because there are so many environmental variables that appear to influence the distribution of mountain ash (in stark contrast to the wide ecological amplitude of messmate, a species that is found across a far greater range of environmental variables), we can consider mountain ash to be stenovalent, i.e. a species that is strongly limited by abiotic factors.

Where environmental variables are marginal, and conditions allow mountain ash to persist but are not ideal, mixed species forest may extend 400 m or more from the boundary of the monotypic stand (Florence, 1996). In the Central Highlands of Victoria there are three
key eucalypt species that may be subdominant or codominant in mountain ash forest, including messmate, alpine ash, and shining gum (E. nitens). Alpine ash is ecologically similar to mountain ash, but occurs mostly at higher elevations, with mixed-species forest occurring at the elevational boundary between the two. Mountain ash thus experiences a milder climate, with higher mean annual temperatures and higher mean minimum temperatures in the cold months (Lindenmayer et al., 1996). Notably, while there may be heavy frosts in surrounding areas, temperatures within 2-5 cm of ground level in mountain ash forest rarely fall below zero, meaning frosts are uncommon on the forest floor (Ashton, 2000), and snow rarely lies on the ground (Cremer, 1975). In contrast to alpine ash, messmate mostly occupies mixed forest communities at lower elevations, but is also found on more weathered soils and exposed habitats across most of the altitudinal range of mountain ash. Shining gum, while not being as strongly delimited by elevational boundaries as alpine ash or messmate, is usually found in small patches that are characterised by low values of mean minimum temperature during the coldest months (Lindenmayer et al., 1996). Thus, the composition of mixed-species mountain ash forest varies according to local conditions and adjacent forest types across its distribution. Other species that are observed in ecotones of mountain ash forest include red stringybark (E. macrorhyncha), mountain grey gum (E. cypellocarpa or, in the Dandenong Ranges, E. cypellocarpa x goniocalyx), ribbon gum (E. viminalis) and Tasmanian blue gum (E. globulus).

1.4.3 Age, height and growth rates

“When Abel Tasman discovered Tasmania in 1642 it was a lusty tree with a diameter at 4’ 3” of about 3 feet and probably a height of 150-200 feet”.

“But it so happened that this huge tree become the property of Australian Newsprint Mills and with its neighbours was logged to provide [75 tons of] newsprint for the Australian market.”

– Helms (1945)

The tallest living trees in the world are not the tallest trees ever recorded. Unfortunately, the tallest trees of recent centuries have been destroyed by anthropogenic influences such as land clearing, timber harvesting, and frequent fires. There are, however, some spectacularly tall trees still extant in some areas. For example, the tallest known living tree is a coast redwood (Sequoia sempervirens) slightly over 115.5 m tall. Even so, the physiological height limit for this species has been estimated at between 122 and 130 m (Koch et al., 2004), and there is no reason to suggest that trees of such prodigious heights have not existed in the past. The tallest living mountain ash is an individual, known as Centurion, located in the Arve
Valley of Tasmania. This specimen has a stem diameter of 405 cm and is now thought to be over 100 m tall. But, while all of these living trees are extremely large, examination of the historical records identifies some astounding tree heights that have been reliably recorded from both North America and Australia. For example, there are historical records of Douglas-fir (Pseudotsuga menziesii), a northern hemisphere species, attaining heights of up to 126 m, a height which seems to be the physiological limit for that species (Koch et al., 2004). In Australia, there are some equally remarkable heights recorded for mountain ash, traditionally known as the world’s tallest flowering plant and hardwood (Beale, 2007). There are some widely accepted records of mountain ash reaching heights of 114.3 m (Hardy, 1918; Hardy, 1935), although these pale in comparison to some of the more dubious records. Whilst there is clearly some exaggeration in historical measurements (see Hardy (1918) and Mace (1996) for some improbable claims), one seemingly well-documented individual that had fallen across a tributary of the Watts River in Victoria was measured using a tape and found to be 132.6 m “from its roots to the top of the trunk” (Simpfendorfer (1982), quoted in Mace (1996)). Unfortunately, it is unlikely that such records will ever be confirmed.

Mountain ash individuals can reach at least 500 years of age (Wood et al., 2010), and only begin to develop hollows (a limiting habitat feature in regrowth forest) after about 120 years of age (Lindenmayer, 2009a). Mountain ash saplings grow extremely rapidly, with mean height growth in young stands exceeding 1.3 m y\(^{-1}\) (Figure 1.1). Growth rates do, however, slow with age, with trees generally about 50 m tall in 50-year-old stands, and about 60 m tall in 80-year-old stands. Having said this, dominant individuals within stands tend to grow much faster, with trees 50 m tall recorded in 35 year old stands (Ashton, 1975a) and trees 80 m tall recorded in 70 year old stands (Beale, 2007). Not surprisingly, it has also been found that nearly half the growth of the total stem diameter occurs in the first 90 years (Wood et al., 2010). Over longer periods, growth rates not only slow, but in the mature and overmature stages of growth (as categorised by Ashton (1975a)) they eventually become negative. This is due to the apical branches dying back with age, followed by the top of the crown breaking off in storms or high winds, inevitably causing old trees to experience a reduction in their total height.
Figure 1.1 Height of young mountain ash (*Eucalyptus regnans*) trees (n=358) in stands of different age classes. Grey shading represents the standard error of the mean. The equation for the linear relationship between age and height is $y = 1.3752x + 2.4868$. The $R^2$ for the model is 0.588.
1.4.4 Stand development

From the time of germination, and without a subsequent disturbance, mountain ash forests undergo a natural decline in stem density. Smith *et al.* (2014) found that the intensity of wildfire can influence the density of seedlings that regenerate, with ‘moderate’ intensity fires (where the understorey and midstorey layers are scorched but the canopy remains unburned) resulting in typical seedling densities of over 281000 per hectare. A rapid decline in density of stems occurs over the first 50 years of stand development, with a logarithmic decline in the density of living stems over the course of the following three to four centuries (Figure 1.2). Tree densities in 70-year-old stands are about 100 per hectare, dropping to around 40 stems per hectare at 300 years of age. In the absence of fire or other disturbance event, stands of such great age essentially consist of loosely spaced, emergent mountain ash trees that overlie a main canopy of rainforest species, including myrtle beech (*Nothofagus cunninghamii*), southern sassafras (*Atherosperma moschatum*), blackwood (*Acacia melanoxylon*) and tree ferns (*Dicksonia antarctica* and *Cyathea australis*).

Such a radical change in stem density results in substantial changes to the structure and function of mountain ash forests through time (Lindenmayer *et al.*, 1999; Lindenmayer *et al.*, 2000), with young, dense stands supporting different suites of bird and other faunal communities when compared with older stands (Serong and Lill, 2012).

Detailed long-term studies of vegetation structure and composition in more than 160 mountain ash-dominated sites around the Central Highlands of Victoria provided the means to develop a practical definition of ‘old-growth’ mountain ash stands (Lindenmayer *et al.*, 2000; Lindenmayer, 2009b). By characterising numerous structural features, it was found that a combination of characters could be used to identify forest stands in which over 70% of the living trees were older than 180 years (Lindenmayer *et al.*, 2000). Structural features associated with these old-growth mountain ash forests include (Lindenmayer *et al.*, 2015):

- Large diameter living overstorey trees typically greater than 1 m in diameter at breast height,
- Trees of markedly different ages within the same stand,
- Numerous trees with hollows, often readily observed from ground level,
- Large quantities of decorticating bark strips hanging from lateral branches of overstorey trees,
- Well-developed clumps of mistletoe in the crowns of overstorey trees,
- A number of distinct vegetation layers,
- An abundance of tree ferns and rainforest trees, and
- Numerous large diameter logs.
Whilst many of these characters are not limited to old-growth patches, one feature that was strongly associated with old-growth forest was the presence of multiple age cohorts. And because stem diameter and tree age are strongly correlated in mountain ash forests (Ashton, 1976), stands containing multiple age cohorts of mountain ash can be easily identified by the presence of two or more distinctly different stem diameters.

Examination of the abundance of old-growth forest in the Central Highlands of Victoria shows that less than 2% (about 1886 ha) of the 157,000 ha of mountain ash forest is old-growth, and this fraction is represented in 147 separate patches (Burns et al., 2015; Lindenmayer et al., 2012a; Lindenmayer, 2009a). Whilst such detailed data is lacking for the remainder of the species distribution, the situation is not likely to be substantially different in Tasmania or other regions of Victoria. Such a low representation of old-growth forest is particularly concerning, given that estimates of pre-European forest age structure suggest that 60-80% of the total area of mountain ash forest would typically have been old-growth (Lindenmayer 2009a; Lindenmayer et al. 2011), with the decline since European settlement thought to be the combined result of frequent fires and timber harvesting (Lindenmayer et al. 2011).
Figure 1.2 Stem densities in natural stands of mountain ash (*Eucalyptus regnans*) of different age classes. Grey shading represents the standard error. Data collected during Chapter 3 of this thesis and collated from other sources (Smith et al., 2014; Ashton, 1976).
1.4.5 Reproduction

1.4.5.1 Ontogeny

Inflorescences in mountain ash are mostly paired umbels in the axils of leaves, usually with between 9 and 15 flowers on each umbel (Brooker and Kleinig, 2006). From the time of appearance of macroscopic buds in leaf axils there is about a year of development before bud burst, when umbellate buds are revealed from beneath bracts that house the inflorescence (Ashton, 1975b; House, 1997). After about a year and a half, the umbellate buds reach anthesis, when flowers are fully open and functional (Ashton, 1975b). It is generally another 8-10 months until the fruit capsules are fully matured and seeds are viable (Ashton, 1975b; Flint and Fagg, 2007). Rather than shedding seed immediately, it is held firmly in the capsules until they die and dry out, allowing the loculi to widen and release the seed (Cremer, 1965b).

Pollination of flowers in mountain ash is thought to be predominantly effected by insects, with medium- and large-bodied flies in the families Calliphoridae (blow flies), Syrphidae (hoverflies) and Tachinidae (tachinids), thought to be the major pollinators (Griffin, 1980; House, 1997). Whilst there is some anecdotal evidence to suggest that honeyeaters (Aves: Meliphagidae) and small arboreal mammals (e.g. sugar gliders Petaurus breviceps) are also pollen vectors (Ashton, 1975b), Griffin (1980) states that nectarivorous birds were rarely sighted in a 2 ha seed orchard over several seasons of flowering. Furthermore, the small, white flowers of mountain ash are indicative of insect-mediated pollination, whereas the large, red flowers of some other species (e.g. E. sideroxylon) suggest pollination by birds (House, 1997). Thus, the contribution of vertebrate fauna to the pollination process in mountain ash is not expected to be a substantial factor, although it remains largely unquantified.

As one of just nine east-Australian eucalypts known to be obligate seeders (Nicolle, 2006), mountain ash is among those species that do not produce vegetative growth from either a lignotuber or epicormic buds after disturbance. As a result, they are, ceteris paribus, more susceptible to be killed by fire than species that are able to resprout from such structures. Interestingly, research has shown that there are few, if any, relevant morphological differences between mountain ash and related species that commonly produce epicormic growth after fire (Waters et al., 2010). Thus it is generally suggested (Ashton, 1981b; Waters et al., 2010) that the epicormic meristems of this species are probably left vulnerable to fire damage as a result of rapid juvenile height growth precluding development of thick protective bark. Of note is a statement by Ashton (1975a) that individuals of 220 years age have greater crown coverage down the stem than 150 year old trees, suggesting that this is due to “the development of epicormic shoots on the trunk”, indicating that epicormic buds may be more
functional in older trees. It may also be true that the thicker bark of older individuals makes them less susceptible to fires, thus contributing to the production of multi-aged stands and even lower severity fires in old-growth forest (Ashton, 1981b; McCarthy et al., 1999; Taylor et al., 2014).

Despite the susceptibility of individual mountain ash trees to high-severity fires (where the canopy of the forest is fully scorched), the ontogeny of the species is strongly tied to fire, because, “unless accompanied by exceptional circumstances” (Ashton and Chinner, 1999), it is only through fire that the species is naturally able to reproduce. This is because the only source of viable seed in mountain ash forests is the seeds stored in the canopy of the trees (Wang, 1997). Typical levels of annual seed fall in mature mountain ash forests result in very low levels of seed germination and subsequent seedling survival, due to predation of seeds by ants (Ashton, 1979; O'Dowd and Gill, 1984), browsing of seedlings by herbivores, and fungal infection of seedlings (Ashton and Macauley, 1972). After a moderate- or high-severity fire, the fruit capsules retained within the canopy dry out as they die, resulting in the locules widening and allowing the seed to fall free (Cremer, 1965a; Cremer, 1965b). This process of capsule death and locule widening is thought to take place mainly in the hours to weeks after a fire (Cremer, 1965a; Ashton, 1979). Flint and Fagg (2007) state that if there is sufficient scorching at the level of the canopy, then mature capsules will shed most of their seed within two to three weeks, with capsules being shed one to eight weeks after seedfall.

1.4.5.2 Time to reproductive maturity

Despite nearly half a century of study on the biology and ecology of the montane ash forests of Victoria, there is still confusion in the published literature regarding that factors that influence the time to reproductive (=sexual) maturity of mountain ash. I cover this subject in more depth in Chapter 2, but, briefly, many authors appear to extrapolate the age of seed production from the seminal work of David Ashton (1956), who states:

“Observations of flowering in the younger forests indicate that, although flowering may commence in stands only 7-8 years old, it is very sparse and patchy (Wallaby Creek and Dandenong plantations and Mt. St. Leonard). At 10-15 years the flowering is well established, but is still patchy and poor on a stand basis. As the forest develops, the flower production becomes heavier and more uniform until maturity is reached.”

While this is probably accurate with respect to flowering, it does not adequately quantify the age at which seed production across a stand is high enough to maintain ecosystem integrity after a wildfire. Further, despite time being an important component of the maturation
process, it is likely that other factors are also important. For example, environmental variables such as rainfall and solar radiation may influence the growth rates of the species and lengthen or reduce the primary juvenile period as a result.

1.4.5.3 Dispersal

The upper tail of the dispersal curve for most plant species is very difficult to quantify, due because with increasing distance increases it becomes exponentially harder to identify the source of pollen or seed (Silvertown and Charlesworth, 2001). In eucalypts, it is likely that the influence of pollen flow versus seed dispersal on a population’s genetic structure depends on the method of pollination. For example, a highly mobile bird pollinator may contribute to pollination over a landscape scale (Hingston et al., 2004a; Hingston et al., 2004b), whereas a less mobile insect pollinator may only reach a neighbouring branch or tree. In contrast, the distinct lack of specialised dispersal mechanisms in eucalypt seeds suggest there is less interspecific variability in the dispersal distances of seeds (Cremer, 1965b).

Seed dispersal from mountain ash is often considered to occur at only small distances, up to the equivalent of one tree height (Gill, 1994; Lamb and Smyth, 2006). Despite this, (Grose, 1960) states that seed dispersal of the ecologically similar alpine ash may, during wind storms, extend up to about 121 m from the maternal parent, although he considers such events unimportant for adequately restocking disturbed areas. Cremer (1966) provides further evidence that seed dispersal distances depend on multiple factors, such as stand density, wind speed and direction, and topography of the landscape, noting that under certain conditions seeds could be distributed up to 150 m from the maternal parent.

While there have been a number of studies into seed dispersal patterns, there is very limited data on pollination distances in natural stands of mountain ash. One study in a seed orchard found a mean dispersal distance of 42 m (Potts and Wiltshire, 1997), but this may not be reflective of patterns in more contiguous landscapes.

As there are a number of limitations in our existing knowledge of dispersal patterns in mountain ash (detailed in Appendix 1), it is possible that dispersal kernels may actually be less leptokurtic than currently thought, due to unusual climatic conditions and stand structure during and immediately after natural regeneration (wildfire) events (Potts and Wiltshire, 1997; Kirkpatrick, 1977).
1.5 Disturbance in mountain ash forests

1.5.1 Wildfire

The history of fire events in a region is known as the fire regime (Gill and Allan 2008). Fire regimes consist of component variables that include the fire intensity, fire frequency, seasonality of fire, and type of fire (Gill, 1975; Gill, 1978). Some authors, possibly fallaciously, include a size or area component (Gill and Allan, 2008).

Cheney (1981) describes three types of fires: those that occur in the lower layers of the forest (surface fires), those that burn through the canopy of the forest (crown fires), and those that burn under the soil surface (peat/humus/ground fires). While peat fires can occur in Australia from the northern tropics to the most southerly temperate zones, surface and crown fires are of more relevance to mountain ash forests and these will be the primary focus in this discussion. The literature relating to Australia’s ash forests often considers crown fires as high-intensity or high-severity fires, with surface fires being described as having low- to moderate-intensities and severities.

More specifically, fire intensity is defined as the product of rate of forward spread, amount of fuel consumed, and the heat of combustion (i.e. energy content of the fuel) (Byram, 1959). This product reaches some of the highest levels worldwide in mountain ash forests, where intensities can be 100,000 kW m⁻¹ (Gill and Moore, 1990), a figure 25 times higher than that considered to be the upper limit for forest-fire control. During such conflagrations, maximum temperatures can reach 1000°C (with temperature declining with increasing height above ground), and temperatures over 100°C persisting for several minutes near the ground (Gill, 1997). Even if flames do not reach the canopy of the forest, it is still possible for the canopy to experience 100% foliar death due to scorching of leaves. This is because tissue-damaging temperatures (e.g. above about 60°C) can occur up to six times the height of the flames (Luke and McArthur, 1978; Cheney, 1981).

Historically, fire regimes in Australia are likely to have varied substantially according to ecosystem types, with some parts of the country being burnt almost annually and others not experiencing substantial fire events for more than a century at a time (Bradstock, 2010). In mountain ash forests (and similar ecosystem types), there are a number of particular characteristics that assist with the ignition and spread of wildfire. Such characters include the presence of volatile oils in the foliage and twigs, pendulous leaves (that encourage updraughts), heavy annual leaf fall (which contributes to surface fuel), and the development of bark streamers that not only encourage spot fires to form ahead of the fire front, but also draw flames into the canopy to create crown fires (Ashton, 1981b). However, these the
presence of these characters does not imply that mountain ash forests have traditionally burnt frequently. In fact, it is thought that high-intensity, stand-replacing wildfires occurred at intervals of between 75 and 150 years (McCarthy et al., 1999), and much of the landscape was in a late successional stage (Lindenmayer, 2009a; Lindenmayer et al., 2015).

Since European colonisation of south-east Australia in the 18th century, anthropogenic disturbances to mountain ash forests have led to the occurrence of more frequent high-intensity wildfires (Table 1.1). Frequent fire, along with other stressors, has led to the idea that mountain ash forests are in a landscape-scale ecological trap (Burns et al., 2015; Lindenmayer and Sato, 2018; Lindenmayer et al., 2011). Frequent disturbance ensures that the landscape remains in a young age-category, which primes the system for further disturbance and impairs ecological function. To understand how this occurs, the accumulation of fuel must be examined. The build-up of litter in forests is a critically important component of ground fuel accumulation, contributing strongly to the intensity of wildfires (Gill, 1978). As litter accumulation in young (e.g. 25 year old) and old (e.g. 200 year old) mountain ash forests is more or less equivalent (Ashton, 1975c), even relatively young stands have large volumes (20 tonnes per ha) of litter being amassed. In the absence of anthropogenic activity, this is unlikely to have resulted in frequent fire due to the low likelihood of ignition. However, contemporary anthropogenic activities have provided a suite of available ignition sources, with well-trafficked roads, recreational and commercial activities, and arson all possible nucleation points. In addition, climate change is causing substantial drying of south-eastern Australia (Cai and Cowan, 2008), resulting in drier fuels and increased chance of ignition and propagation of fire events. These factors are all contributing to the conclusion that young forests are more flammable than old forests (Taylor et al., 2014). Adding to this, recent work has shown that mature ash-type forests (including mountain ash) are more than eight times less flammable than young (< 21 years old) forests (Zylstra, 2018). This finding also has substantial implications for our understanding of silvicultural practices and their influences on mountain ash forest.
<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Total extent (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Thursday</td>
<td>1851</td>
<td>76000</td>
</tr>
<tr>
<td>Red Tuesday</td>
<td>1898</td>
<td>2600</td>
</tr>
<tr>
<td>Black Sunday</td>
<td>1926</td>
<td>3940</td>
</tr>
<tr>
<td>-</td>
<td>1932</td>
<td>2040</td>
</tr>
<tr>
<td>Black Friday</td>
<td>1939</td>
<td>13800</td>
</tr>
<tr>
<td>Ash Wednesday</td>
<td>1983</td>
<td>870</td>
</tr>
<tr>
<td>Black Saturday</td>
<td>2009</td>
<td>4500</td>
</tr>
</tbody>
</table>
1.5.2 Silviculture

“Greatly is to be feared in ages hence, when much of the woods will have sunk under ruthless axes, the deductions of advanced knowledge thereon will have to be based solely on evidence early placed on record.”


Eucalypts are typically an inefficient source of high quality timber, due to the frequency of defects such as decayed heart, gum pockets and veins, and pinholes (Ferguson, 1987). Mountain ash is no exception to this, and received early condemnation as being “altogether inferior in quality” (Maiden, 1905). Despite this, the relatively small amount of sawn timber that is produced is of high quality, with von Mueller (1871) considering it “fissile, well adapted for shingles, rails, for housebuilding, for the keelson and planking of ships”. As a result, mountain ash has long been harvested in Victoria and Tasmania for various purposes. In the latter half of the 19th century the forests were selectively logged by paling splitters using hand tools, and by the 1930s machinery such as steam winches and sawmills were established, with tramways used to convey timber to railway stations (Flint and Fagg, 2007).

Today, clear-cutting is the dominant form of silvicultural technique in mountain ash forests, with 95% of the harvested area in Victoria subjected to this method (Lutze et al., 1999). Clear-cutting is conducted in a number of stages; initially almost all trees are felled on sites (of variable size, often 15-40 ha) using mechanical harvesting machines. After trees have been de-barked and de-limbed, they are loaded onto trucks and taken to a timber mill. The biomass remaining on site, consisting of non-commercial species and mountain ash bark and branches, is left to dry on-site for about six months, after which it is subjected to high-intensity fire to promote regeneration (Fagg et al., 2013; Blair et al., 2016). After a period of one to six months, the site is sown with mountain ash seed, usually broadcast from a helicopter.

As commercial methods of harvesting have become more efficient, the use of mountain ash in building materials has become much less common, with just 10% of the merchantable biomass removed from the site being converted into sawn timber products (Keith et al., 2014). Of the remainder, about half is converted to paper pulp, and the rest is removed as waste products (Keith et al., 2014).

Some authors (e.g. Attiwill, 1994) have suggested that current silvicultural practice, of clear-felling coupes with subsequent treatment of logging debris with a high-intensity fire, is largely equivalent to the natural disturbance regime in mountain ash forests. However, more recent research has shown that there are substantial differences between logged stands and
natural stands, with logged stands having reduced species richness and abundance in particular functional groups (e.g. midstorey trees, shrubs and ferns) and taxa (Blair et al., 2016; Bowd et al., 2018). While it has been suggested that the clear-cutting promotes ignition and propagation of wildfires, by modifying stand structure, microclimate, and shifting species composition towards that of drier forests (Lindenmayer et al., 2009; Mueck and Peacock, 1992), there is some debate around this idea (Attiwill et al., 2013). However, as timber harvesting ensures that much of the landscape is maintained in a young age category, and younger forests are more than eight times more flammable than older forests (Zylstra, 2018), it can be concluded that current management of mountain ash forests is increasing their flammability.
1.6 Thesis aims

Broadly, this thesis aims to contribute to our knowledge of how contemporary disturbance regimes are impacting the population demographics and genetic structure of forest trees. Developing our understanding of this topic is critical for efficient and effective conservation management of forest trees and forest ecosystems. With contemporary disturbance regimes having similar impacts on many of the world’s forests, the results and conclusions drawn from this thesis will ideally have relevance across many forest types.

Due to our good understanding of this system and its underlying ecological processes, as well as its history of utilisation by Europeans, mountain ash forests are a good model system to develop an understanding of the ways in which modified fire regimes and climate change can influence the structure, function and longevity of forest communities. By using mountain ash as a study species, the data presented here are particularly relevant to this species, and in this respect my aims are to evaluate the impacts of contemporary disturbances in the context of mountain ash as a foundation species in a critically-endangered ecosystem type. Thus, the results presented here should be informative for land managers, allowing better conservation outcomes for mountain ash forests, other obligate seeder-dominated forests, and forest trees in general.

To achieve these aims, I undertake a number of studies that characterise aspects of the species life history, gene flow and dispersal patterns, adaptation to local environmental conditions, and adaptive potential, and place these in the context of current disturbance regimes and environmental changes. Each chapter of the thesis thus has one or more somewhat independent aims. In short, the thesis will (1) examine the likely impacts of contemporary disturbance regimes on the maturation rate, and stand persistence under different topographic conditions, (2) investigate range-wide population genetic structure and hybridisation, and determine the consequences of these factors for conservation of adaptive potential and assisted gene flow, (3) explore the impacts that disturbances (wildfire and logging) have on genetic diversity, and (4) investigate the patterns and drivers of local adaptation and the extent to which these vary across spatial scales. There is also an appendix in which I evaluate different methods of genotyping fire-killed trees for the purposes of genetically determining dispersal distances under natural regeneration scenarios.

1.6.1 Chapter 2

While the timing of reproductive development is a key life history trait in many plants, the transition to a reproductive state is particularly important for obligate seeders, as these
species may be susceptible to population declines and extinctions if they depend on a particular fire regime for regeneration. In some cases, increasing fire frequencies threaten population persistence, with the age of reproductive maturity placing a critical lower threshold on the interval between fires that can be tolerated by a species.

In this context, my aims here are to determine the extent to which relevant environmental variables influence the growth of tree height and stem diameter, and investigate how these size components interact with the environment to influence the age to seed production in mountain ash. I hypothesise that environmental factors such as rainfall and solar radiation will substantially influence growth rates of height and stem diameter, and thus also impact the speed with which reproductive maturity is reached. To test this, I investigate whether mountain ash stands exhibit variation in susceptibility to increasing fire frequency by characterising the response of key vital rates to stand age (time since fire) under different environmental conditions.

1.6.2 Chapter 3

Conservation management often uses information on genetic population structure to assess the importance of local provenancing for ecological restoration and reintroduction programs. For species that do not exhibit complete reproductive isolation, the estimation of population genetic parameters may be influenced by the extent of hybridisation. To avoid perverse outcomes for conservation, genetically-informed management strategies must determine whether hybridisation between species is driving population genetic structure and patterns, and identify the proportion of the total genetic diversity present within populations. If most genetic variation is already present within populations, there may be little benefit in planning for climate-adjusted provenancing or similar conservation strategies.

Given the importance of understanding population genetic structure for conservation, and the knowledge that hybridisation in eucalypts is a widespread and common phenomenon, I investigated these two aspects in mountain ash. The existence of hybrids between mountain ash and messmate stringybark has long been known (Ashton, 1956), however the extent of hybridisation across the range of the species has never been examined. Similarly, while the chloroplast genetic structure of mountain ash has been studied, the structure of the nuclear genome has not. To address these knowledge gaps, my aims were to (1) identify the extent and possible drivers of hybridisation across the geographic distribution of mountain ash, (2) describe how identification of hybridisation using genomic data may influence our understanding of population genetic structure, and (2) consider how these factors would
influence current management strategies in eucalypts. To obtain data, I use genotyping-by-sequencing to obtain large numbers of genome-wide genetic markers on individual samples. I predict that (1) some individuals and populations will show greater levels of admixture with messmate stringybark, (2) levels of hybridisation will be driven in part by local environmental variables, and (3) the inclusion or exclusion of hybrid individuals in population genetic analyses will lead to different strategic outcomes for management. If these predictions are true, there are implications for future studies of population genetic structure, as well as the planning of restoration plantings and assisted migration.

1.6.3 Chapter 4

Forests dominated by mountain ash are considered critically endangered, with frequent wildfires and widespread logging causing a multitude of changes to the structural attributes of the forest. However, changes to disturbance regimes not only have flow-on effects on species abundance and community composition, but also on the spatial distribution of genetic diversity. In this context, my primary aims in this chapter are to identify the effects of logging and fire history on measures of genetic diversity, fine-scale genetic structure, and population genetic differentiation in mountain ash stands, as well as investigate whether maternally-inherited (seed dispersed) genomic components are more susceptible to disturbance than biparentally-inherited (seed and pollen dispersed) genomic components. To achieve this, I investigate the patterns and extent of genetic diversity and genetic differentiation within and among populations at both chloroplast and nuclear DNA loci in forest stands subjected to timber harvesting, and compare these to (1) stands that had regenerated naturally after fires prior to 1940, had burnt in the 2009 fire, and had not experienced any known timber harvesting, and (2) stands that had regenerated naturally after fires prior to 1940 and had not experienced any major disturbance since.

I expect that silvicultural practices will lead to different patterns of genetic diversity in logged stands compared to burnt and naturally-regenerated stands. This is because logged sites cannot regenerate from canopy-stored seedbanks and need to artificially sown with seed (typically via use of a helicopter) to ensure regeneration. Knowing this, there are a number of possibilities as to how genetic structure in logged stands could be changing, including (1) the invasion of non-local alleles due to seed for artificial regeneration being sourced from distant localities, leading to altered patterns of diversity and structure among logged stands, or (2) a reduction in genetic diversity as a result of selective regeneration of a small number of (local or non-local) genotypes, due to either the use of few maternal trees for seed collection or the effects of local selective processes. I also predict that the effects of logging will differ between
nuclear and chloroplast genomes, with the naturally higher levels of structure and lower within-site diversity of chloroplast DNA making it more susceptible to seed provenancing strategies.

1.6.4 Chapter 5

Understanding the patterns and drivers of local adaptation is important for the conservation and management of ecologically and economically important species under rapidly-changing environmental conditions. Even where gene flow is very high and genetic differentiation across the landscape is very low, strong selective filtering processes can still produce local adaptation. As there has been little study into the effect of spatial scale of sampling on detection of loci involved in local adaptation, I aim to search the genome of mountain ash for signatures of local adaptation, with sampling conducted at two spatial scales; (1) across the full range of the species, in relation to coarse-scale climatic and topographic variation, and (2) within the geographically-restricted region of the Victorian Central Highlands, in relation to fine-scale environmental drivers of climatic and edaphic conditions.

To achieve this aim, I utilise three methods to identify loci under putative selection; (1) genetic differentiation ($F_{ST}$) outlier analysis, (2) latent-factor mixed models, and (3) redundancy analysis. I predict that low levels of allelic isolation between mountain ash populations ensures that a high level of standing genetic variation is available in most populations, suggesting that many of same alleles will be available for selection to act on across both large- and fine-scale environmental gradients. In this context, we hypothesise that many of the loci under putative selection at the range-wide spatial scale will also be identified at the small spatial scale. However, we also suggest that fundamental differences between spatial scales (such as the extent of variation in environmental covariates), as well as the large array of potential genetic pathways available for adaptive change, will lead to the identification of more putatively adaptive loci than would be found if a single spatial scale was investigated.

1.6.5 Appendix 1

Next-generation sequencing methods typically rely on high quality samples that produce high quantities and qualities of DNA. However, there are there are a range of situations in which high quality samples may not available. For example, to conduct a parent-offspring
analysis (in which dispersal distance is directly measured using genetic relatedness) on obligate seeders (which are typically killed by fire), would require either pre-fire knowledge of parent genotypes or genotyping of fire-killed parents. As it is difficult to predict where fires will occur, such a study is likely to require genotyping of fire-killed trees.

Here, I aim to determine whether it is possible to genotype fire-killed mountain ash trees using three methods, on the premise that this would allow us, with appropriate sampling, to directly measure dispersal distances under natural conditions using parent-offspring analyses. I utilise two methods of genome complexity reduction followed by next-generation sequencing, and one target amplification method whereby nuclear microsatellites are subjected to polymerase chain reaction and then sized on a capillary sequencer. As the parent trees had been dead for at least six years at the time of sampling, I predict that the quantity of DNA present in the samples will be low and/or degraded and that this could prevent accurate genotyping using next-generation sequencing methods. I also suggest that microsatellite amplification and sizing will be the most viable way of genotyping mountain ash, to allow for direct measurement of dispersal distances using genetic methods.

As dispersal is such a fundamental component of species biology, determining the rate at which species can recolonise areas of local extinction and allowing them to modify their range in response to changing environmental factors, I consider this an important first step towards obtaining accurate estimates of dispersal parameters.
Chapter 1: Introduction

1.7 References


Hingston, A. B., Potts, B. M. and McQuillan, P. B., 2004b. Pollination services provided by various size classes of flower visitors to Eucalyptus globulus ssp. globulus (Myrtaceae), *Australian Journal of Botany*, 52(3): 353-369. Available at: 10.1071/BT03002

Hoffmann, A. A. and Sgro, C. M., 2011. Climate change and evolutionary adaptation, *Nature*, 470: 479. Available at: 10.1038/nature09670


32

Chapter 1: Introduction


Chapter 2: Environmental influences on growth and reproductive maturation of a keystone forest tree: implications for obligate seeder susceptibility to frequent fire


Key-words: Eucalyptus regnans, fire regime, fire return interval, niche shift, obligate seeder, primary juvenile period, serotiny, vital rate
2.1 Abstract

Anthropogenic modifications to climate and natural fire regimes are occurring globally, leading to the production of environments that may be unsuitable for some species. Fire-intolerant plant species that rely on specific fire regimes for reproduction are at risk of population decline when successive fires occur in less than the time taken to produce seed. Quantifying key fire-related life history traits in such species is therefore critical for developing models of population viability, species distributions and ecosystem persistence.

We studied the Australian mountain ash (*Eucalyptus regnans*), the world’s tallest angiosperm and an ecologically and economically important keystone species. We tested whether mountain ash populations exhibit variation in susceptibility to increasing fire frequency by characterising the response of key vital rates to stand age (time since fire) under different environmental conditions.

We found that the time taken to produce seed varied geographically. Mean growth rates were greater in areas receiving higher levels of solar radiation, a trend that became stronger with tree age. Tree size and age had the strongest influence on the production of fruit capsules. Mature fruit capsules were found in trees as young as 11 years old, but stands may not contain reproductively viable seed crops until they are more than 21 years old.

Our results show that environmental factors influence the primary juvenile period of a keystone obligate seeder, in turn affecting the time taken for a population to develop a reproductively viable amount of seed. Reduced fire return intervals may therefore constrain the species’ realised niche (and geographic distribution) to areas where it can tolerate shorter fire return intervals due to faster growth and maturation. We suggest that populations of obligate seeders that reach reproductive viability faster are thus more likely to persist when exposed to multiple fires in short succession. Intra-stand variation in seed crops suggests that selection could also act on rapidly-maturing individuals, resulting in some populations exhibiting high levels of precocious reproductive activity.
2.2 Introduction

Fire influences vegetation communities in multifarious ways, driving ecosystem structure (Bond and Keeley, 2005) and the evolution of plant traits (Dantas et al., 2013; Keeley et al., 2011; Pausas et al., 2004; Pausas et al., 2006). However, anthropogenic modifications to natural fire regimes are occurring in ecosystems worldwide (Abatzoglou and Williams, 2016; Pausas and Keeley, 2014a), with important consequences for species that have strong successional preferences or traits adapted to specific fire regimes (Bowman et al., 2014; Morrison et al., 1995). Altered fire regimes can therefore lead to geographic shifts in environmental suitability for species. This can influence population viability and therefore the distribution of individual species. However, in the case of ecologically important ‘foundation’ or keystone species (sensu Tewksbury and Lloyd, 2001; Lindenmayer and Laurance, 2016) such as dominant overstorey forest trees, there are potential major implications for many other dependent species. If we are to predict the responses of plant species to changing fire regimes, we need to quantify the key life history traits that are likely to determine plant sensitivity to changing fire regimes (Enright et al., 2015).

One functional group of plants that may be impacted by modified fire regimes is the obligate seeders (Bradstock, 2008; Keith, 1996). Such species do not produce vegetative regrowth from lignotubers or epicormic buds, and thus regenerate only from seed, with adults usually being killed by crown fire (Pausas and Keeley, 2014b). For many plants, one of the key life history traits for maximising fitness is the age of reproductive activity (Amasino, 2010; Huijser and Schmid, 2011), with seed production a key component of regeneration and community assembly (Larson and Funk, 2016). But for obligate seeders, the timing of the transition (phase change) to a reproductive state is particularly relevant, as these species may be susceptible to population declines and extinctions if they depend on a particular fire regime for regeneration (Stephens et al., 2013). In some cases, increasing fire frequencies threaten population persistence, with the age of reproductive maturity placing a critical lower threshold on the interval between fires that can be tolerated by a species (Bassett et al., 2015; Bowman et al., 2014; Keeley et al., 1999; Lindenmayer et al., 2011; Syphard et al., 2006).

In the context of global changes to disturbance regimes and climate, it is critical that we are able to link plant vital rates such as growth and reproduction to environmental variables (Ehrlén et al., 2016; Enright et al., 2015), particularly for predicting how the distribution and viability of a population is likely to change. For obligate seeders, clarification of parameters relating to the age of reproductive maturity is an essential part of this process (de Gouvenain and Ansary, 2006). In particular, two parameters are critical for predicting species responses to shifting fire regimes:
- The primary juvenile period of an individual, defined here as the age at which the plant has first produced viable seed, ready for dissemination. This is the best definition in the context of obligate seeders, as some species do not produce flowers (thus negating references to flowering) (Trauernicht et al., 2012), and flowering species may not produce viable seed for some years after the initiation of inflorescence buds (Ashton, 1975b).

- Stand-level reproductive viability, which we define as a state in which a population is capable of producing an equivalent population of offspring plants, without human intervention. This state is determined by a combination of factors. For example, the amount and distribution of seed available for regeneration is likely to be a critical aspect for many obligate seeders (Bassett, 2011). For serotinous species (which release seed in response to an environmental trigger), and species that rely on ash-beds or gaps in vegetation for successful regeneration, the timing and intensity of a fire event is likely to influence germination and survival (Ooi et al., 2006). For species with canopy-stored seedbanks, stand characteristics (e.g. canopy height) may also influence how much seed is available for dissemination after fire (Taylor et al., 2014).

To accurately quantify the two above parameters involves understanding the degree to which they are influenced by environmental variation (Ehrlén et al., 2016). This is because, while the primary juvenile period and time to stand-level reproductive viability define thresholds for tolerance of fire return intervals, environmentally-driven variation in these parameters may drive niche shifts in obligate seeders under increasing fire frequency, whereby some populations have increased tolerance for short fire return intervals under particular ecological conditions (Scheele et al., 2017; Swab et al., 2012).

In this study, we investigate the influence of environmental variables on individual and stand-level maturation patterns in the world’s tallest flowering plant, the obligate seeder mountain ash (Eucalyptus regnans). These forests are a good system to develop an understanding of the ways in which modified fire regimes can influence life history traits such as reproductive maturation, as research from the past 60 years has led to a good understanding of the ecology of the ecosystem (Ashton, 1956; Cremer, 1975; Lindenmayer, 2009b; Lindenmayer et al., 1996; Lindenmayer et al., 1991; Loyn, 1985; McCarthy et al., 1999; Smith et al., 2016).

In mountain ash forests, intervals between high-intensity stand-replacing fire events may be reduced from historical levels of 75-150 years (McCarthy et al., 1999) to less than the time to reproductive viability of a stand. As all seed is stored in the canopy of mature mountain ash trees, the occurrence of tree-killing fires before the production of viable canopy seedbanks
could cause the widespread collapse of mountain ash forest ecosystems (Figure 2.1) (Bowman et al., 2014; Gilbert, 1959; Lindenmayer, 2009a). Indeed, we have already begun to see large areas of forest dominated by obligate seeding species require artificial reseeding to maintain ecosystem persistence (Bassett et al., 2015; Fagg et al., 2013).

In the context of reduced fire return intervals, the objectives of this study were to determine the extent to which relevant environmental variables influence the growth of tree height and stem diameter, and investigate how these size components interact with the environment to influence the age to seed production in mountain ash. We hypothesise that environmental factors such as rainfall and solar radiation will substantially influence growth rates of height and stem diameter, and thus also impact the speed with which reproductive maturity is reached. While a good conceptual model of the factors involved in determining the extent of a species’ susceptibility to fire regime shifts exists (Enright et al., 2015), there are relatively few studies that have empirically examined the relevant life history traits of a species and used these to predict potential changes to population dynamics under contemporary fire regimes. In light of this, we further suggest that understanding these relationships for mountain ash will also be informative in predicting how contemporary fire regimes will affect landscape-scale distribution of other obligate seeding species.
Figure 2.1 Conceptual model of the lifecycle of an obligate seeder, showing some of the factors that influence vital rates. While some obligate seeders have soil-stored seed banks, others, such as the *Eucalyptus regnans* depicted here, rely solely on canopy-stored seed for reproduction. Clarification of factors influencing the length of the primary juvenile period is critical for understanding the influence of changing disturbance regimes on population viability.
2.3 Materials and methods

2.3.1 Study species

Mountain ash is a dominant overstorey tree in tall wet forests of south-eastern Australia. These forests contain some of the highest known carbon biomass densities in the world (Keith et al., 2009), and are heavily utilised for pulpwood and timber resources across their range in Victoria and Tasmania. The combined effects of logging and numerous large wildfires over the past 80 years has left much of the mountain ash forest in Australia in a younger age class than would be typical historically (Lindenmayer, 2009a; Lindenmayer et al., 2011; Lindenmayer et al., 2012a).

Floral development in mountain ash typically includes a year of growth between the time that macroscopic buds are visible in leaf axils to the time of bud burst, when bracts enclosing the developing inflorescence are shed to reveal umbellate buds (Ashton, 1975b; House, 1997). Following bud burst, the umbellate buds take another 13-17 months to reach anthesis, when flowers are fully open and functional (Ashton, 1975b; Bassett, 2011). After flowering (a period of weeks, typically between March and May), it is another 8 to 12 months before the fruit capsules are fully developed (Ashton, 1975b; Flint and Fagg, 2007). As true reproductive capacity is reached only at the end of this developmental period (when fruit capsules have fully matured and seeds become viable), nearly 3.5 years after inflorescence buds become visible, it is the timing of this final stage that determines the ability of an individual to produce offspring. It is also of note that fruit capsules mature during the austral summer, with seeds becoming ready for dispersal around the same time (late January or early February) that wildfires frequently occur, and that shifts in either the timing of wildfires or the maturation of fruit capsules can influence the level of seedfall and resultant germination (Fagg et al., 2013).

Much of what has been written on the age of reproductive activity in mountain ash is found in the seminal work of David Ashton (1956), which states that (1) stands of 7-8 years old may contain sparsely distributed flowering plants, (2) by 10-15 years of age flowering is common but still patchy, and (3) flowering becomes heavier and more uniform until maturity is reached. Most authors appear to use a basic extrapolation of these statements or other aspects of Ashton’s research, and either implicitly or explicitly suggested that the overriding determinant of reproductive maturity in mountain ash is time. Various publications over the past few decades have suggested that individuals or stands either begin reproducing or are reproductively viable at 15-20 years {e.g. Gill (1975), Fagg et al. (2013), Smith et al. (2014) referencing Gill (1981), and Fairman et al. (2015) referencing Ashton (1976)}, 20 years {e.g.}
McCarthy et al. (1999) referencing Ashton and Attiwill (1994), or 20-30 years {e.g. Lindenmayer et al. (2011) referencing Ashton (1975b)}. This confusion results from a lack of explicit published data on the reproductive state of stands of known ages. Most of the literature to date also tends to ignore the potential importance of environmental influences on vital rates such as tree growth and stand maturation.

2.3.2 Study sites

Yarra State Forest, southeast of Warburton, and Toolangi State Forest, to the west of Marysville, were the primary regions used in the study. Both of these state forests (SFs) contain large areas of mountain ash forest, much of which is in various stages of regrowth after fires in the 20th and 21st centuries. Both SFs have also been subjected to timber harvesting operations for over 100 years, and contain logged stands of varying shapes and sizes.

We used a forestry GIS layer (DELWP, 2014) to select stands based on multiple criteria, including size (no stands <8 ha) and shape (circular rather than long and thin), overstorey composition (mountain ash as the dominant tree species), time since logging, and the absence of wildfire between the date of seeding and the time of surveys. This produced a total of 36 stands, of which 20 were in Toolangi SF and 16 were in Yarra SF, where stands were sown with seed between about July 1993 and July 2007 (i.e. that were aged between about 8 and 23 years of age at the time of surveys). Thus all ages were surveyed multiple times, with each age class visited a mean of 2.25 times, and a total of 360 trees measured. As data on the precise date of regeneration burns and subsequent seeding were not available, the time between the end of harvesting and reseeding was assumed to be less than two years.

2.3.3 Field surveys

Field surveys were undertaken in the 2015-2016 summer (December-March), to coincide with the maturation time of fruit capsules and seed (Bassett, 2011). Each stand was surveyed by walking a 90 m transect, starting about 50 m from the stand boundary. Ten trees were selected by choosing the closest living mountain ash tree at every ten metre increment along the transect, and a waypoint for each was taken using a handheld Garmin GPS MAP 64s. Measurements taken on each tree included (1) the stem diameter at breast height (1.5 m) over bark (DBH), (2) height (by either cutting down the tree and measuring it with a tape, or where this was not possible, by using a Suunto clinometer and trigonometry), and (3) seed crop information. The presence of buds and/or flowers was also recorded.
To estimate the number of viable seeds available in the canopy of each tree, we used a modified version of the updated seed-crop assessment technique presented in Bassett et al. (2006), Bassett (2009), and Bassett (2011). This technique builds upon the original work of Harrison et al. (1990) in estimating seed abundance in mountain ash forests, and essentially relies upon knowledge of three factors: (1) the number of capsule-bearing branchlets in the crown (NB), (2) the average number of capsules per capsule-bearing branchlet (CPB), and (3) the average number of viable seeds per capsule (VS). Multiplying these three factors together allows one to estimate the number of viable seeds in the canopy of a standing tree, i.e.

\[ \text{Seed-crop per tree} = \text{NB} \times \text{CPB} \times \text{VS} \]

Due to our interest in stands of young trees, we were unable to use the regression models for estimating the number of capsule-bearing branchlets per tree, as suggested in Harrison et al. (1990) and Bassett (2009). Instead, we directly counted the total number of capsule-bearing branchlets per tree, taking care to distinguish between mature capsules (indicated by the presence of clearly-defined valve slits) and immature capsules. Only mature capsules were used in the calculations, to prevent overestimation of the number of viable seeds.

To calculate the number of capsules per branchlet, we randomly selected six capsule-bearing branchlets with diameters of approximately 2 cm and used a pair of high-quality binoculars (Carl Zeiss Conquest HD 10 x 42) to visually compare their appearance to a series of photographs showing known capsule densities. Where necessary, a spotting scope was used to obtain acceptable views of the branchlets. Where the number of capsule-bearing branchlets on a tree was less than six, the total number of capsules was calculated by summing the branchlet densities rather than averaging.

Estimates for the numbers of viable seeds per capsule can be far from the true value without detailed examination of large numbers of capsules from multiple trees within each stand. However, the estimate of one viable seed per capsule suggested for recently matured capsules by Bassett (2011) was used to obtain rough estimates of the variability in stand level maturation periods.

To calculate the number of viable seeds per hectare (ha) in each stand, we multiplied the seed-crop per tree by the density of living mountain ash stems per ha. To obtain the number of stems per ha, all mountain ash individuals within 2 m of the transect from the start to the 50 m mark were counted, and this number was multiplied by 50.

Fagg et al. (2013) state that approximately 40% of canopy-stored seed in mountain ash forests is unlikely to be available for germination after a moderate-intensity fire event,
suggesting that to obtain commercial levels of seedfall \((2 \times 10^5\) viable seeds ha\(^{-1}\)) there would need to be \(3.33 \times 10^5\) viable seeds ha\(^{-1}\) in the canopy. For what they term ‘ecological stocking’, i.e. where the stand is not intended to be harvested in the future, Fagg et al. (2013) also suggest that just 10-40\% of the commercial seed crop densities are required, implying capsule densities of between \(3.33 \times 10^4\) ha\(^{-1}\) and \(1.33 \times 10^5\) ha\(^{-1}\). We took the higher end of these estimates (40\% of commercial densities) and estimated the time to which this benchmark would be reached in a regenerating stand of uniformly-aged trees.

### 2.3.4 Statistical analysis

We used a model selection approach to determine which of a suite of predictor variables most influenced the growth and maturation of mountain ash trees (Johnson and Omland, 2004). Response variables for each tree included the presence or absence of fruit capsules (FRUITYN), estimated number of fruit capsules (FRUIT), tree height (HEIGHT), and stem diameter (DBH).

Using geographic information system (GIS) datasets for the region, we extracted values of predictor variables for each surveyed tree. These variables were derived from a 20 m digital elevation model (DEM) developed by John Stein at the Fenner School of Environment and Society, with source data from the Victorian Department of Environment, Land, Water and Planning spatial datasets ‘VICMAP Elevation Contour’ and ‘VICMAP Hydro Watercourse’. This DEM was used to determine topographic wetness index (TWI), a measure of potential soil moisture calculated as \(\log[\text{specific catchment/slope}]\) (Moore and Hutchinson, 1991), as well as north-south (Q) and east-west (P) aspect-adjusted slopes (-1 - 1) (Hutchinson, 1998). The DEM was also used with the ANUCLIM software (Xu and Hutchinson, 2013) to calculate spatially-explicit values of mean annual solar radiation (13.5 - 14.3 MJ m\(^{-2}\) day\(^{-1}\)), mean annual temperature (10 - 13 °C), and mean annual precipitation (1254 - 1627 mm y\(^{-1}\)), labelled RAD, TEMP and RAIN respectively (Table 2.1). Stand age in years (AGE), as well as interactions between AGE and all environmental and climatic variables, were included as predictor variables in global models.

Prior to fitting global models, we checked for multicollinearity between all predictor variables using variance inflation factors. We corrected for multicollinearity in two ways. First, we removed mean annual precipitation (which had a strong negative correlation with solar radiation) and mean annual temperature (which had a strong positive correlation with elevation) from the analyses. Second, as HEIGHT and DBH did not vary independently, we used these measurements to create two predictor variables, called SIZE and SHAPE (Table 2.1), for use in models where the fitted response was FRUITYN or FRUIT.
For the response variables of DBH and HEIGHT, we fitted linear mixed-effects models. To assess the presence or absence of fruit capsules in trees, we fitted a generalised linear mixed-effects binomial model. To account for over-dispersion that was present in the response variable FRUIT, we fitted it with a generalized linear mixed-effects negative binomial model. A random effect of stand was fitted in all models, to account for repeated sampling at each location.

To find the best subset of models for each response variable, we fitted all combinations of fixed effects and ranked models using Akaike’s Information Criterion corrected for finite sample sizes (\(\text{AIC}_c\)) (Burnham and Anderson, 2004; Burnham et al., 2011). Top-ranked models (i.e. models with a \(\Delta\text{AIC}_c \leq 2\)) were considered useful for inference.

To estimate the time to viable seed crop production at the level of the population, we used mean values for each stand and fitted a generalized linear model of fruit per ha (calculated as mean fruit per tree multiplied by the number of living stems per ha), with all relevant predictor variables.

All model construction and selection was completed in R version 3.2.1 (R Core Team, 2017), using the ‘lme4’ (Bates et al., 2015), ‘glmmADMB’ (Fournier et al., 2012; Skaug et al., 2016), and ‘MuMIn’ (Barton, 2016) packages for model fitting and selection.

After fitting the global models, we investigated spatial autocorrelation by visually examining variograms of the model residuals. Variograms were created using the variog function in the ‘geoR’ package (Ribeiro Jr. and Diggle, 2016), with the smooth option selected. As no evidence of a nugget or sill was observed in any of the variograms, spatial autocorrelation was not considered any further.
Table 2.1 Response and predictor variables considered in analyses of mountain ash (*Eucalyptus regnans*) height growth, stem diameter growth and fruit capsule production models. Interactions between AGE and all environmental and climatic variables were included in global models. Mean annual temperature and mean annual precipitation were removed from global models to correct for multicollinearity.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Response/Predictor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRUIT</td>
<td>Response</td>
<td>Number of mature fruit capsules in canopy</td>
</tr>
<tr>
<td>FRUITYN</td>
<td>Response</td>
<td>Presence or absence of fruit (binomial)</td>
</tr>
<tr>
<td>HEIGHT</td>
<td>Response</td>
<td>Tree height (m)</td>
</tr>
<tr>
<td>DBH</td>
<td>Response</td>
<td>Stem diameter (cm) at breast height (1.5 m above ground) over bark.</td>
</tr>
<tr>
<td>SIZE</td>
<td>Predictor</td>
<td>Measure of tree size (0.01 - 7.67), calculated as ( \text{HEIGHT} \times (\pi \times (\text{DBH}/2)^2) ), where DBH is in metres.</td>
</tr>
<tr>
<td>SHAPE</td>
<td>Predictor</td>
<td>Measure of tree shape (63 - 230), calculated as ( \text{HEIGHT}/\text{DBH} ), where DBH is in metres.</td>
</tr>
<tr>
<td>P</td>
<td>Predictor</td>
<td>East-west aspect-adjusted slope (-1 to 1). Negative values represent westward facing slopes; positive values represent eastward facing slopes; and values near zero represent ridgetops or valley bottoms.</td>
</tr>
<tr>
<td>Q</td>
<td>Predictor</td>
<td>North-south aspect-adjusted slope (-1 to 1), Negative values represent southward facing slopes; positive values represent northward facing slopes; and values near zero represent ridgetops or valley bottoms.</td>
</tr>
<tr>
<td>DEM</td>
<td>Predictor</td>
<td>Elevation (300 - 891 m above sea level)</td>
</tr>
<tr>
<td>RAD</td>
<td>Predictor</td>
<td>Mean annual solar radiation (13.5 - 14.3 MJ m(^{-2}) day(^{-1}))</td>
</tr>
<tr>
<td>TEMP</td>
<td>Predictor</td>
<td>Mean annual temperature (10 - 13 °C)</td>
</tr>
<tr>
<td>RAIN</td>
<td>Predictor</td>
<td>Mean annual precipitation (1254 - 1627 mm y(^{-1}))</td>
</tr>
<tr>
<td>TWI</td>
<td>Predictor</td>
<td>Topographic wetness index (-3.8 – 2.0), calculated as ( \log(\text{specific catchment}/\text{slope}) ).</td>
</tr>
<tr>
<td>AGE</td>
<td>Predictor</td>
<td>Time since stand was sowed (years)</td>
</tr>
</tbody>
</table>
2.4 Results

2.4.1 Seed crop development

We found mature fruit capsules in trees as young as 11 years of age, but fruiting trees were uncommon within these young stands. In contrast, all stands that were 20 or more years old had at least 40%, and up to 70%, of trees containing mature fruit capsules. The number of capsules observed also increased with tree age, with no capsules formally recorded in trees younger than 11 years of age, compared to a mean of 208 capsules in 20-year-old trees. Unsurprisingly, larger trees tended to have more capsules than smaller trees (Figure 2.2).

The variables AGE, SIZE, RAD, Q and two interaction terms, AGE:Q and AGE:RAD were identified in all of the top models for the presence/absence of fruit (Table 2.2 and Table 2.3). Other variables that were observed in fewer top models included P, SHAPE, TWI, and the interaction of P with AGE. Tree size and age had the strongest effects on the probability of a tree having mature fruits, with larger and older trees having a greater likelihood of fruiting (Figure 2.3 and Figure 2.4). The interaction of AGE and Q indicates that trees are more likely to begin fruiting at a younger age on north-facing aspects than on south-facing aspects (Figure 2.3), whereas the interaction between AGE and RAD indicates that older trees show a reduction in the probability of fruiting with increasing levels of solar radiation.

Unsurprisingly, the response variable of FRUIT had a similar set of top models to FRUITYN, with predictor variables found in all top models including AGE, SIZE, P, Q, and the interaction of AGE with Q (Table 2.2 and Table 2.3). Larger trees, and trees on north- and east-facing aspects, were positively associated with the abundance of fruit capsules. Variables present in fewer top models included RAD, DEM, TWI, and various interaction terms with AGE.
Figure 2.2 Number of fruit capsules observed in mountain ash (*Eucalyptus regnans*) trees (n=358), aged between eight and 23 years old, plotted against stem diameter and tree height.
Table 2.2 Mixed effects models for presence/absence of fruit capsules and number of fruit capsules in mountain ash (*Eucalyptus regnans*) trees, ranked by Akaike’s Information Criterion adjusted for small sample sizes (AIC<sub>c</sub>). All models that differed from the top model (ΔAIC<sub>c</sub> ≤ 2) are shown, as well as the intercept-only model. Model terms include tree age (AGE), tree size (SIZE), tree shape (SHAPE), topographic wetness index (TWI), east-west aspect-adjusted slope (P), north-south aspect adjusted slope (Q), elevation (DEM), and mean annual solar radiation (RAD). Also shown are the degrees of freedom (d.f.), log-likelihood (log(L)), and relative likelihood of each model (AIC<sub>w</sub>).

<table>
<thead>
<tr>
<th>Presence/absence of fruit capsules</th>
<th>d.f.</th>
<th>log(L)</th>
<th>AIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>ΔAIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>AIC&lt;sub&gt;w&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE + SIZE + P + Q + RAD + AGE:RAD + AGE:P + AGE:Q</td>
<td>10</td>
<td>-123.68</td>
<td>268.00</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + AGE:RAD + AGE:Q</td>
<td>9</td>
<td>-125.35</td>
<td>269.21</td>
<td>1.22</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + AGE:RAD + AGE:Q</td>
<td>8</td>
<td>-126.45</td>
<td>269.31</td>
<td>1.32</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + SHAPE + P + Q + RAD + AGE:RAD + AGE:P + AGE:Q</td>
<td>11</td>
<td>-123.42</td>
<td>269.60</td>
<td>1.61</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + TWI + AGE:RAD + AGE:P + AGE:Q</td>
<td>11</td>
<td>-123.43</td>
<td>269.62</td>
<td>1.63</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + AGE:P + AGE:Q</td>
<td>9</td>
<td>-125.65</td>
<td>269.82</td>
<td>1.83</td>
<td>0.02</td>
</tr>
<tr>
<td>Intercept only</td>
<td>2</td>
<td>-171.73</td>
<td>347.49</td>
<td>79.49</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of fruit capsules</th>
<th>d.f.</th>
<th>log(L)</th>
<th>AIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>ΔAIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>AIC&lt;sub&gt;w&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE + SIZE + P + Q + AGE:Q</td>
<td>8</td>
<td>-698.35</td>
<td>1413.11</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + DEM + TWI + AGE:DEM + AGE:P + AGE:Q + AGE:TWI</td>
<td>14</td>
<td>-692.08</td>
<td>1413.38</td>
<td>0.27</td>
<td>0.04</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + TWI + AGE:Q + AGE:TWI</td>
<td>10</td>
<td>-696.54</td>
<td>1413.72</td>
<td>0.61</td>
<td>0.04</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + DEM + TWI + AGE:P + AGE:Q + AGE:TWI</td>
<td>13</td>
<td>-693.46</td>
<td>1413.98</td>
<td>0.87</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + AGE:P + AGE:Q</td>
<td>9</td>
<td>-697.76</td>
<td>1414.03</td>
<td>0.92</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + TWI + AGE:P + AGE:Q + AGE:TWI</td>
<td>12</td>
<td>-694.65</td>
<td>1414.21</td>
<td>1.10</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + TWI + AGE:P + AGE:Q + AGE:TWI</td>
<td>11</td>
<td>-695.84</td>
<td>1414.43</td>
<td>1.32</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + TWI + AGE:Q + AGE:TWI</td>
<td>11</td>
<td>-695.88</td>
<td>1414.52</td>
<td>1.41</td>
<td>0.02</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + DEM + AGE:Q</td>
<td>9</td>
<td>-698.01</td>
<td>1414.54</td>
<td>1.43</td>
<td>0.02</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + TWI + AGE:RAD + AGE:Q + AGE:TWI</td>
<td>12</td>
<td>-695.06</td>
<td>1415.02</td>
<td>1.91</td>
<td>0.02</td>
</tr>
<tr>
<td>AGE + SIZE + P + Q + RAD + DEM + TWI + AGE:Q + AGE:TWI</td>
<td>12</td>
<td>-695.09</td>
<td>1415.09</td>
<td>1.98</td>
<td>0.02</td>
</tr>
<tr>
<td>Intercept only</td>
<td>3</td>
<td>-733.36</td>
<td>1472.78</td>
<td>59.68</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Chapter 2: Environment and vital rates
Table 2.3 Parameter estimates and standard errors of the top five models identified in Table 2, for two fruiting response variables in mountain ash (*Eucalyptus regnans*). Models were ranked by Akaike’s Information Criterion adjusted for small sample sizes (AICc). Also shown is the relative likelihood of each model (w). Model terms include tree age (AGE), tree size (SIZE), tree shape (SHAPE), east-west aspect-adjusted slope (P), north-south aspect adjusted slope (Q), mean annual solar radiation (RAD), topographic wetness index (TWI), elevation (DEM), and interactions of AGE with topographic and climatic variables.

<table>
<thead>
<tr>
<th>Presence/absence of fruit capsules</th>
<th>Rank 1 (w= 0.06)</th>
<th>Rank 2 (w= 0.03)</th>
<th>Rank 3 (w= 0.03)</th>
<th>Rank 4 (w= 0.03)</th>
<th>Rank 5 (w= 0.03)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.16 (0.35)</td>
<td>-2.06 (0.33)</td>
<td>-2.04 (0.33)</td>
<td>-2.13 (0.35)</td>
<td>-2.19 (0.36)</td>
</tr>
<tr>
<td>AGE</td>
<td>1.15 (0.33)</td>
<td>1.09 (0.32)</td>
<td>1.09 (0.33)</td>
<td>1.13 (0.33)</td>
<td>1.19 (0.34)</td>
</tr>
<tr>
<td>SIZE</td>
<td>1.29 (0.23)</td>
<td>1.30 (0.23)</td>
<td>1.28 (0.24)</td>
<td>1.41 (0.29)</td>
<td>1.30 (0.23)</td>
</tr>
<tr>
<td>SHAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.47 (0.21)</td>
<td>0.28 (0.19)</td>
<td></td>
<td>0.48 (0.21)</td>
<td>0.50 (0.22)</td>
</tr>
<tr>
<td>Q</td>
<td>0.80 (0.34)</td>
<td>0.73 (0.33)</td>
<td>0.73 (0.33)</td>
<td>0.79 (0.34)</td>
<td>0.83 (0.35)</td>
</tr>
<tr>
<td>RAD</td>
<td>-0.13 (0.24)</td>
<td>-0.12 (0.24)</td>
<td>-0.13 (0.24)</td>
<td>-0.15 (0.24)</td>
<td>-0.14 (0.25)</td>
</tr>
<tr>
<td>TWI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.13 (0.18)</td>
</tr>
<tr>
<td>AGE:RAD</td>
<td>-0.55 (0.27)</td>
<td>-0.53 (0.26)</td>
<td>-0.54 (0.27)</td>
<td>-0.58 (0.27)</td>
<td>-0.56 (0.27)</td>
</tr>
<tr>
<td>AGE:P</td>
<td>-0.39 (0.21)</td>
<td></td>
<td></td>
<td>-0.40 (0.21)</td>
<td>-0.42 (0.22)</td>
</tr>
<tr>
<td>AGE:Q</td>
<td>-1.00 (0.34)</td>
<td>-0.89 (0.34)</td>
<td>-0.90 (0.34)</td>
<td>-0.99 (0.34)</td>
<td>-1.01 (0.35)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of fruit capsules</th>
<th>Rank 1 (w= 0.05)</th>
<th>Rank 2 (w= 0.04)</th>
<th>Rank 3 (w= 0.04)</th>
<th>Rank 4 (w= 0.03)</th>
<th>Rank 5 (w= 0.03)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.63 (0.63)</td>
<td>1.28 (0.38)</td>
<td>1.03 (0.54)</td>
<td>1.09 (0.38)</td>
<td>0.61 (0.63)</td>
</tr>
<tr>
<td>AGE</td>
<td>2.88 (0.68)</td>
<td>2.56 (0.51)</td>
<td>2.75 (0.6)</td>
<td>2.65 (0.53)</td>
<td>2.80 (0.67)</td>
</tr>
<tr>
<td>SIZE</td>
<td>1.46 (0.37)</td>
<td>1.60 (0.31)</td>
<td>1.44 (0.36)</td>
<td>1.53 (0.28)</td>
<td>1.46 (0.37)</td>
</tr>
<tr>
<td>P</td>
<td>1.10 (0.43)</td>
<td>1.07 (0.28)</td>
<td>1.02 (0.36)</td>
<td>1.01 (0.27)</td>
<td>1.26 (0.46)</td>
</tr>
<tr>
<td>Q</td>
<td>1.37 (0.57)</td>
<td>2.06 (0.45)</td>
<td>1.57 (0.51)</td>
<td>1.88 (0.44)</td>
<td>1.42 (0.56)</td>
</tr>
<tr>
<td>RAD</td>
<td>-0.75 (0.29)</td>
<td></td>
<td></td>
<td>-0.75 (0.3)</td>
<td></td>
</tr>
<tr>
<td>DEM</td>
<td>-0.90 (0.46)</td>
<td></td>
<td></td>
<td>-0.45 (0.33)</td>
<td></td>
</tr>
<tr>
<td>TWI</td>
<td>1.13 (0.43)</td>
<td>0.61 (0.56)</td>
<td>1.13 (0.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE:P</td>
<td>-0.87 (0.37)</td>
<td></td>
<td></td>
<td>-0.68 (0.4)</td>
<td>-0.55 (0.5)</td>
</tr>
<tr>
<td>AGE:Q</td>
<td>-1.27 (0.61)</td>
<td>-2.09 (0.5)</td>
<td>-1.46 (0.6)</td>
<td>-1.91 (0.5)</td>
<td>-1.33 (0.6)</td>
</tr>
<tr>
<td>AGE:DEM</td>
<td>1.07 (0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE:TWI</td>
<td>-1.89 (0.46)</td>
<td>-1.25 (0.64)</td>
<td>-1.78 (0.46)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.3 Effect plots showing how (a) tree age, (b) tree size, (c) annual solar radiation, and (d) the north-south aspect-adjusted slope (Q) influence the probability of mature fruit capsules being present in young mountain ash (*Eucalyptus regnans*) trees. The variable Q shows the changes occurring as landscape position varies from southward facing slopes (negative values) to northward facing slopes (positive values). Grey areas indicate 95% confidence intervals.
Figure 2.4 Modelled probability of mature fruit capsules being present in mountain ash (*Eucalyptus regnans*), plotted onto real values of tree height and stem diameter. Trees of similar size with differing probabilities of fruiting indicate the influence of other variables.
2.4.2 Height and stem diameter growth

At the level of the stand, mean height growth varied from 1 - 2 m y⁻¹, and mean stem diameter varied from 0.8 – 2 cm y⁻¹. Mean tree height at 22 years of age was 33.2 m, compared to 15.2 m at 8 years old. The mean stem diameter at breast height for these two ages was 30.7 cm and 11.3 cm respectively. Dominant predictor variables in the top models for DBH were AGE and RAD, with Q, TWI, DEM, and interactions of AGE with RAD or TWI present in some models (Table 2.4 and Table 2.5). Trees in areas receiving higher levels of solar radiation had higher growth rates of height and stem diameter than trees in areas receiving less sunlight. For the response variable of HEIGHT, AGE and RAD were similarly dominant predictor variables, with the interaction between AGE and RAD present in three out of the four models with ΔAICc < 2. Thus, it appears that tree height and stem diameter are predominantly determined by tree age and the degree of solar radiation (Figure 2.5). The interaction of AGE and RAD suggests that, for young trees, annual solar radiation does not substantially affect tree height, whereas older trees tend to grow faster in topographic locations that receive more sunlight (Figure 2.5).

The strong negative correlation \( r = -0.95, p < 0.001 \) of mean annual rainfall with solar radiation indicates that areas of high rainfall receive lower levels of solar radiation, and vice versa. Thus, somewhat counterintuitively, trees in areas of high annual rainfall (e.g. 1600 mm) experience reduced height and stem diameter growth when compared to trees in areas of lower annual rainfall (e.g. 1250 mm). Whether this trend holds true for areas receiving less than 1250 mm per annum is uncertain.

Some predictor variables, such as DEM and TWI, found in highly ranked models appear to have only weak explanatory power and do not lead to major improvements in model fits. As elevation is strongly negatively correlated \( r = -0.94, p < 0.001 \) with mean annual temperature, neither elevation nor temperature is considered to be strongly influential on the growth rate or maturation of mountain ash.
Table 2.4 Mixed effects models for tree height and stem diameter in mountain ash (*Eucalyptus regnans*), ranked by Akaike’s Information Criterion adjusted for small sample sizes (AIC<sub>c</sub>). All models that differed from the top model (\(\Delta\text{AIC}_c\)) by ≤ 2 are shown, as well as the intercept-only model. Model terms include tree age (AGE), tree height (HEIGHT), stem diameter at breast height (DBH), topographic wetness index (TWI), north-south aspect adjusted slope (Q), elevation (DEM), and mean annual solar radiation (RAD). Also shown are the degrees of freedom (d.f.), log-likelihood (log(L)), and relative likelihood of each model (AIC<sub>w</sub>).

<table>
<thead>
<tr>
<th>Stem diameter</th>
<th>d.f.</th>
<th>log(L)</th>
<th>AIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>(\Delta\text{AIC}_c)</th>
<th>AIC&lt;sub&gt;w&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE + RAD</td>
<td>5</td>
<td>-1177.87</td>
<td>2368.88</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>AGE + RAD + Q</td>
<td>6</td>
<td>-1176.66</td>
<td>2369.39</td>
<td>0.51</td>
<td>0.04</td>
</tr>
<tr>
<td>AGE</td>
<td>4</td>
<td>-1179.80</td>
<td>2369.73</td>
<td>0.85</td>
<td>0.04</td>
</tr>
<tr>
<td>AGE + RAD + AGE:RAD</td>
<td>6</td>
<td>-1176.89</td>
<td>2370.08</td>
<td>1.20</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + RAD + Q + AGE:RAD</td>
<td>7</td>
<td>-1175.48</td>
<td>2370.16</td>
<td>1.28</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + RAD + TWI + AGE:RAD</td>
<td>7</td>
<td>-1176.47</td>
<td>2370.41</td>
<td>1.53</td>
<td>0.03</td>
</tr>
<tr>
<td>AGE + RAD + DEM</td>
<td>6</td>
<td>-1177.01</td>
<td>2370.57</td>
<td>1.69</td>
<td>0.02</td>
</tr>
<tr>
<td>AGE + Q</td>
<td>5</td>
<td>-1178.79</td>
<td>2370.71</td>
<td>1.83</td>
<td>0.02</td>
</tr>
<tr>
<td>Intercept only</td>
<td>3</td>
<td>-1204.28</td>
<td>2416.98</td>
<td>48.10</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tree height</th>
<th>d.f.</th>
<th>log(L)</th>
<th>AIC&lt;sub&gt;c&lt;/sub&gt;</th>
<th>(\Delta\text{AIC}_c)</th>
<th>AIC&lt;sub&gt;w&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE + RAD + AGE:RAD</td>
<td>6</td>
<td>-1004.29</td>
<td>2023.49</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>AGE + RAD</td>
<td>5</td>
<td>-1006.36</td>
<td>2025.05</td>
<td>1.56</td>
<td>0.05</td>
</tr>
<tr>
<td>AGE + RAD + TWI + AGE:RAD</td>
<td>7</td>
<td>-1004.55</td>
<td>2025.36</td>
<td>1.87</td>
<td>0.04</td>
</tr>
<tr>
<td>AGE + RAD + TWI + AGE:RAD + AGE:TWI</td>
<td>8</td>
<td>-1004.01</td>
<td>2025.43</td>
<td>1.93</td>
<td>0.04</td>
</tr>
<tr>
<td>Intercept only</td>
<td>3</td>
<td>-1035.31</td>
<td>2078.93</td>
<td>55.44</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 2.5 Parameter estimates and standard errors of the top five models identified in Table
4, for two size-related response variables in mountain ash (Eucalyptus regnans).
Models were ranked by Akaike’s Information Criterion adjusted for small
sample sizes (AICc). Also shown is the relative likelihood of each model (w).
Model terms include tree age (AGE), east-west aspect-adjusted slope (P), north-
south aspect adjusted slope (Q), mean annual solar radiation (RAD), topographic
wetness index (TWI), elevation (DEM), and interactions of AGE with
topographic and climatic variables.

<table>
<thead>
<tr>
<th>Stem diameter</th>
<th>Rank 1 (w= 0.05)</th>
<th>Rank 2 (w= 0.04)</th>
<th>Rank 3 (w= 0.04)</th>
<th>Rank 4 (w= 0.03)</th>
<th>Rank 5 (w= 0.03)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td></td>
<td>20.27 (0.64)</td>
<td>20.26 (0.64)</td>
<td>20.26 (0.66)</td>
<td>20.05 (0.68)</td>
<td>20 (0.68)</td>
</tr>
<tr>
<td>AGE</td>
<td>6.22 (0.68)</td>
<td>6.22 (0.68)</td>
<td>6.61 (0.66)</td>
<td>6.11 (0.7)</td>
<td>6.09 (0.69)</td>
</tr>
<tr>
<td>RAD</td>
<td>1.15 (0.68)</td>
<td>1.23 (0.68)</td>
<td>1.20 (0.69)</td>
<td>1.3 (0.68)</td>
<td>0.87 (0.63)</td>
</tr>
<tr>
<td>Q</td>
<td>0.77 (0.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE:RAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.64 (0.68)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tree height</th>
<th>Rank 1 (w= 0.11)</th>
<th>Rank 2 (w= 0.05)</th>
<th>Rank 3 (w= 0.04)</th>
<th>Rank 4 (w= 0.04)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td></td>
<td>22.95 (0.58)</td>
<td>23.31 (0.56)</td>
<td>22.96 (0.58)</td>
<td>22.93 (0.57)</td>
</tr>
<tr>
<td>AGE</td>
<td>5.78 (0.59)</td>
<td>5.97 (0.6)</td>
<td>5.8 (0.59)</td>
<td>5.78 (0.58)</td>
</tr>
<tr>
<td>RAD</td>
<td>1.35 (0.58)</td>
<td>1.26 (0.59)</td>
<td>1.34 (0.58)</td>
<td>1.4 (0.57)</td>
</tr>
<tr>
<td>TWI</td>
<td></td>
<td></td>
<td>-0.13 (0.27)</td>
<td>-0.11 (0.27)</td>
</tr>
<tr>
<td>AGE:RAD</td>
<td>1.07 (0.57)</td>
<td>1.04 (0.58)</td>
<td>1.02 (0.57)</td>
<td></td>
</tr>
<tr>
<td>AGE:TWI</td>
<td></td>
<td></td>
<td>0.37 (0.27)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.5 Effect plots showing the influence of tree age and annual solar radiation on the growth of (a,b) height and (c,d) stem diameter in mountain ash (*Eucalyptus regnans*) trees. The differing effects of annual solar radiation on tree height in young (10 years) and older (20 years) trees is also shown. Grey areas indicate 95% confidence intervals.
2.4.3 Stand-level reproductive viability

Surveyed stands had stem densities of between 150 and 4650 stems ha\(^{-1}\), and there was a moderate negative correlation \((r = -0.63, \ p < 0.001)\) between stem density and stand age. Estimates for the number of fruit capsules per ha ranged from zero (in young stands) to 615,125 in one 19-year-old stand. Modelling the number of fruit capsules per ha as a function of stand age suggests that the mean age at which stands would contain 40% of the commercial stocking density of seeds (about 133,000 viable seeds ha\(^{-1}\)) is approximately 21 years. However, there are numerous factors that influence the length of time taken to reach this particular threshold (Figure 2.6), implying that some stands will reach this level when they are much older (our lower 95% confidence intervals at 21 years was about 69,000 seeds ha\(^{-1}\)), or younger (upper 95% confidence interval at 21 years was about 237,000 seeds ha\(^{-1}\)).
Figure 2.6 Predicted values of the number of viable seeds per hectare in mountain ash (Eucalyptus regnans) stands of different ages. Four different combinations of hypothetical variables were used for predictions, including (a) large tree size and mean north-south aspect-adjusted slope, (b) small tree size and mean north-south aspect-adjusted slope, (c) mean tree size and north-facing aspect-adjusted slope, and (d) mean tree size and south-facing aspect-adjusted slope. Grey areas indicate 95% confidence intervals. The horizontal dotted lines represent the amount of viable seed required to produce germination equivalent to 40% of that produced by the commercial stocking density. Note that a cube root scale is used on the y-axis.
2.5 Discussion

We have quantified some of the factors that influence the growth rates and primary juvenile period in an obligate seeder, and shown that this affects the time taken for a population to develop a reproductively viable amount of seed. Importantly, tree size was found to be influenced by environmental factors, and tree size was a strong determining factor of the probability that an individual would contain fruit capsules. This has a number of implications for obligate seeding species, including the possibility that reduced fire return intervals may be associated with local niche shifts towards areas of faster growth and maturation.

2.5.1 Environmental influences of growth and maturation

Environmental conditions influence the time to seed production in mountain ash. In particular, the level of solar radiation received by a tree has a substantial effect on the growth rates of height and stem diameter, which in turn help to determine the likelihood that a tree of a given age will produce viable seed. In addition, the production of seed crops in young stands appears to depend partially on topographic variables, with solar radiation and aspect found to be important factors. This is consistent with a previous study of eucalypt growth patterns, which found, in productive environments, there was strong competition for sunlight in developing stands (Prior and Bowman, 2014).

It has previously been found that the largest trees of a given age will typically flower first, and that a shortening of the juvenile phase is associated with environmental conditions that promote vegetative growth (Harper and White, 1974). Visser (1964), who used DBH as a measure of ‘vigour’, found that environmental factors influence the time taken to reach a minimum size for flowering in apple and pear trees. In Australia, Benson (1985) found that, in Hawkesbury sandstone plant species, it was often the more advanced plants that were the first of an age cohort to flower. Studies such as these have led to the acceptance that there is a clear size-related component marking the transition from a juvenile plant to a reproductively competent plant (i.e. a plant with the ability to flower) (Taiz and Zeiger, 2002). Our data suggests that this holds true for mountain ash, with tree size being a reasonable predictor of whether individuals are reproductively mature.

Age- and size-related changes in developmental processes are often influenced by both intrinsic (internal) and extrinsic (environmental) factors (Day et al., 2002). For example, changes in environmental conditions (such as access to light) occur as trees grow taller and stem density declines, and these may act on developing or developed organs in a stimulus-response manner (Day et al., 2002; Taiz and Zeiger, 2002; Wareing, 1959). However, there
are also autonomous, gibberellin, and age-dependent pathways, which are not only able to influence the timing of the phase change from juvenile to reproductive competency but may be independent of environmental cues (Poethig, 2013; Wellmer and Riechmann, 2010). The importance of the various intrinsic and extrinsic factors are species specific and remain largely unexplored for many plant species.

The influence of environmental conditions on the speed of growth and the timing of maturation has led to the idea that, for some species, it may be more reasonable to use size categories, rather than age categories, to classify life-history attributes (Werner and Caswell, 1977). Indeed, categories based partly on size and partly on age have already been used by researchers when discussing mountain ash life histories (Ashton, 1976; Ashton, 1956; Lindenmayer et al., 2016), and it is essential to clarify how size and age interact if we are to fully understand life history patterns in such species (Watkinson and White, 1985).

### 2.5.2 Stand-level reproductive viability

Our data suggest that mountain ash trees typically do not start producing mature fruit capsules containing viable seed until they are about 11 years old. While this is younger than any known published ages for seed production in this species, it is not entirely surprising, considering that flowering can begin at least as early as 6-8 years old (Ashton, 1956). There are also small numbers of individuals that produce fruit capsules earlier than 11 years, however these individuals are uncommon and not indicative of when a stand of mountain ash may be considered reproductively viable (although they may be of adaptive significance, as discussed in section 4.3). For example, Doherty et al. (2017) identify and discuss the ability of the ecologically similar alpine ash (*Eucalyptus delegatensis*) to produce fruit at young ages (e.g. six years after fire), likely as a stress response. During our fieldwork for the current study, one of the authors (BvTD) encountered a seven-year-old individual that contained 24 mature fruit capsules along with buds and flowers. Reproductive viability at the level of the stand, however, is determined by multiple factors, including the amount and distribution of seed available for dissemination, stand characteristics (e.g. canopy height), seasonal timing of fire, and fire intensity. If considering only the development of seed crops through time, our data suggest that stand-level reproductive viability is, on average, reached at 21 years of age, an age that agrees with information gathered from post-fire surveys of fire-killed ash species (Fagg et al., 2013).

We also found that as stand age increases, the proportion of trees containing fruit also increases. Whilst stands of mountain ash younger than 11 years of age effectively do not contain seed crops, the mean percentage of trees containing fruit in stands between 20 and 23 years old was 55%. It can therefore be seen that younger stands will have patchier
dissemination of seed after a fire, because fewer trees will be contributing to the regeneration process. Older stands, which have a more even distribution of seed in the canopy, will produce a more evenly distributed seed fall and subsequent regrowth. This is also likely to be the case for many other species that are potentially susceptible to ‘interval squeeze’ as a result of frequent fire (Enright et al., 2015). While we used age in our models, the strong correlation of stand age with stem density (for our data $r = -0.63, p < 0.001$) indicates that experimental manipulation of densities is probably needed to distinguish which of the two variables is more causal in relation to seed production (Ehrlén et al., 2016).

For serotinous species, which often require fire for successful seed dissemination, germination, and seedling survival, the influence of fire on the seedbank must also be considered. For example, previous research has shown that stands of mountain ash aged between 7 and 36 years have the greatest likelihood of experiencing canopy consumption or scorch during fires, with one reason being that stands in this age group have a limited canopy height (Taylor et al., 2014). Such canopy consumption may reduce seed availability by damaging fruit capsules (Judd, 1993), particularly in stands with shorter trees. However, fire weather has been found to be a more important factor than stand age in alpine ash forests (Bowman et al., 2016), with fire intensity and seasonal timing strongly influencing the availability of seed (Fagg et al., 2013). In addition, not all forest systems exhibit the same changes through time. For example, in Yellowstone National Park, older forests appear to be less prone to canopy consumption than younger forests (Turner and Romme, 1994). Such differences are likely to be the result of different successional processes amongst forest types, particularly with respect to fuel loads, fuel moisture content, and wind speed, all of which influence fire intensity and have been characterised for some ecosystems (Lindenmayer et al., 2000; Lindenmayer, 2009a).

2.5.3 Implications for obligate seeders under changing fire regimes

The data presented here highlight the importance of understanding how environmental conditions influence the time to reproductive viability in obligate seeders. This understanding becomes even more critical in light of ongoing impacts from climate change (Enright et al., 2014), which are likely to include altered rainfall and temperature patterns in many regions (Hewitson et al., 2014). In the context of reduced fire return intervals and spatially large wildfires, it is plausible that not all populations of obligate seeders will decline at the same rates. Populations that reach reproductive viability faster are more likely to persist when exposed to multiple fires in short succession, whereas populations experiencing conditions
that slow plant growth or otherwise extend the time to reproductive viability will be more susceptible to declines. Such changes to climatic niches are common in declining species and can plausibly result in extinctions (González-Orozco et al., 2016; Scheele et al., 2017). In the case of keystone species, these population declines will require costly human interventions, such as aerial seeding, to ensure that large areas of forest communities are not lost (Bassett et al., 2015).

The substantial variation in fruit capsule numbers between trees within stands suggests that there is potential for strong selection for individuals that reach reproductive viability quickly. In plants, precocious sexual maturity or reproduction is often caused by a cessation of organ growth at an earlier stage in the descendant organism, a process termed progenesis (Box and Glover, 2010; Wiltshire et al., 1994). Whilst this sort of progenetic response may at first seem unlikely, it is generally acknowledged that such changes in the relative timing or rate of developmental processes has been an important process in the evolution of numerous plant species (Conway and Poethig, 1993; Jones, 1992), including many eucalypts (reviewed in Potts and Wiltshire, 1997). Indeed, Wiltshire (1991) suggests that the progenetic features observed in Tasmanian populations of Risdon peppermint (Eucalyptus risdonii) are the result of frequent disturbance, with fire as a primary cause. As it is chiefly the neotenic and progenetic features that distinguish the threatened Risdon peppermint from the more widespread sister species silver peppermint (E. tenuiramis) (Brooker and Kleinig, 2006), such changes can be of substantial importance to both classification and conservation.
2.6 Conclusions

We have shown that mean growth rate and seed crop production of a keystone obligate seeder are substantially influenced by environmental variables. This, combined with a changing climate and modified fire regime, raises the possibility of shifts in life histories and realised niches, as well as population declines. If we are to fully understand and predict the way that fire-intolerant species will respond to contemporary environments, we need more empirical studies of the factors that influence relevant plant traits, and how these traits influence species persistence. It is only through detailed quantitative research into ecological processes that we will be able to predict changes to forest systems, and manage these changes most appropriately.
2.7 Acknowledgements

We thank John Stein for his efforts in compiling and making available the relevant GIS layers and associated information, and Wade Blanchard for statistical advice. Joe Salmona and Donna Belder provided assistance in the field. Drawings of mountain ash were kindly supplied by Ding Li Yong. This work was funded by a Lesslie Research Scholarship in Landscape Conservation and Ecology to B.v.T.D, and an Australian Research Council Future Fellowship (FT130100043) to S.B. We are also grateful to two anonymous reviewers for their constructive feedback, which has improved the manuscript substantially.
2.8 References


Bassett, O. D., 2009. Development of a seed-crop assessment model for Alpine Ash (Eucalyptus delegatensis), and including an updated model for Mountain Ash (E. regnans), Benalla, Victoria.


Chapter 2: Environment and vital rates

biodiversity threats under climate change, *Nature Climate Change*, 6(12): 1110-1114. Available at: 10.1038/nclimate3126


Poethig, R. S., 2013. *Vegetative phase change and shoot maturation in plants*, Elsevier Inc.,


Wiltshire, R. J. E., 1991. 'Heterochrony and heteroblasty in the Eucalyptus risdonii Hook.f./E. tenuiramis Miq. complex',


Chapter 3: Pervasive admixture between eucalypt species has consequences for conservation and assisted migration


Key-words: admixture, hybrid, mountain ash, Eucalyptus regnans, SNP, population genetics, gene flow, eucalypt
3.1 Abstract

Conservation management often uses information on genetic population structure to assess the importance of local provenancing for ecological restoration and reintroduction programs. For species that do not exhibit complete reproductive isolation, the estimation of population genetic parameters may be influenced by the extent of admixture. Therefore, to avoid perverse outcomes for conservation, genetically-informed management strategies must determine whether hybridisation between species is relevant, and the extent to which observed population genetic patterns are shaped by inter-specific versus intra-specific gene flow.

We used genotyping-by-sequencing to identify over 2400 informative single-nucleotide polymorphisms across 18 populations of Eucalyptus regnans F. Muell., a foundation tree species of montane forests in south-eastern Australia. We used this data to determine the extent of hybridisation with another species, Eucalyptus obliqua L’Hér., and investigate how admixture influences genetic diversity parameters, by estimating metrics of genetic diversity and examining population genetic structure in datasets with and without admixed individuals.

We found hybrid individuals at all sites, and two highly introgressed populations. Hybrid individuals were not distributed evenly across environmental gradients, with logistic regression identifying hybrids as being associated with temperature. Removal of hybrids resulted in increases in genetic differentiation ($F_{ST}$), expected heterozygosity, observed heterozygosity and the inbreeding coefficient, and different patterns of isolation-by-distance. After removal of hybrids and introgressed populations, mountain ash showed very little population genetic structure, with a small effect of isolation-by-distance, and very low global $F_{ST}$ (0.03).

Our study shows that, in plants, decisions around provenancing of individuals for restoration depend on knowledge of whether hybridisation is influencing population genetic structure. For species in which most genetic variation is held within populations, there may be little benefit in planning conservation strategies around environmental adaptation of seed sources. The possibility for adaptive introgression may also be relevant when species regularly hybridise.
3.2 Introduction

Substantial biodiversity declines are occurring in many regions of the world due to widespread land clearing, habitat degradation, introduced species and climate change (Evans et al., 2011; Pounds et al., 2006; Woinarski et al., 2015). Extensive and ongoing land clearing has led to major reductions in forest cover globally (Reside et al., 2017; Achard et al., 2014; Taubert et al., 2018), with synergistic interactions between stressors placing some ecosystems under high threat of rapid collapse or changes in ecosystem state (Brook et al., 2008; Lindenmayer et al., 2011; Lindenmayer and Sato, 2018). With such widespread changes facing ecosystems, it is critical to understand how these stressors interact with the fundamental ecological processes operating within and between foundation species, to adequately manage biodiversity across landscapes.

Using genetic approaches to inform management activities allows conservation efforts to be targeted towards sites of unique genetic composition or adaptive importance, making population genetic studies valuable in many taxa (Maunder et al., 2001; Reynolds et al., 2015; Ikeda et al., 2017; McCartney-Melstad and Shaffer, 2015). To maximise beneficial outcomes, it is vital that our understanding of the population genetic diversity and structure in target species is as accurate as possible, particularly in the context of a changing environment. For example, understanding patterns of local adaptation across the range of a species is important for developing methods of assisted gene flow to mitigate the impacts of climate change and other threatening processes (Supple et al., 2018; Kelly and Phillips, 2016).

Genetically-informed conservation requires a detailed understanding of the spatial distribution of genetic diversity, particularly as it relates to environmental adaptation. Spatial genetic structure and population genetic differentiation are typically considered to be driven by the influences of gene flow, genetic drift and local adaptation (Orsini et al., 2013). However, for species that do not exist in complete reproductive isolation, the estimation of population genetic parameters may be influenced by the extent of hybridisation and introgression with closely related species. This could have large implications for the application of genetic data to conservation management of species, e.g. by committing resources to putatively distinct populations, when they may actually contain highly admixed individuals.

With the advent of modern DNA genotyping techniques, studies investigating thousands of genetic markers from across the genome are becoming more common (Hudson et al., 2015; Gaughran et al., 2018; Harvey et al., 2017; Hand et al., 2015), and several studies have investigated patterns of nuclear genetic structure and gene flow across large geographic regions (Sampson et al., 2018; Hendricks et al., 2017; Shriver et al., 2005; Hecht et al., 2015).
Studies such as these provide critical information for the conservation of populations with unique genetic heritage, identification of areas of adaptive potential for assisted migration, and location of source populations or historical refugia (Supple et al., 2018; Hecht et al., 2015).

Australian natural vegetation communities are dominated by the hyper-diverse and commercially important tree genus *Eucalyptus* L’Hér. With about 700 species recognised (Bayly, 2016), eucalypts are an integral part of the Australian landscape and are foundation species in many ecological communities. For such an important component of Australia’s vegetation, there is still much to understand about gene flow, population dynamics, and genetic structure in eucalypts. Gene flow in plants is typically the result of both pollen and seed dispersal, with pollen typically playing a greater role in eucalypts because it tends to disperse further than seeds (Petit et al., 2005; Barber, 1965; Potts and Wiltshire, 1997). Comparisons of the maternally-inherited chloroplast and biparentally-inherited nuclear DNA have shown that pollen-mediated gene flow can be up to at least 200 times greater than seed-mediated gene flow in some species (Nevill et al., 2014; Bloomfield et al., 2011), although at least one study found that gene flow from seed dispersal is practically equivalent to that from pollen dispersal (Jones et al., 2006).

Given the importance of understanding population genetic structure for conservation, and the knowledge that hybridisation in eucalypts is a widespread and common phenomenon (Griffin et al., 1988), we investigated these two aspects in *Eucalyptus regnans* F. Muell. (mountain ash) one of Australia’s most well-known and economically important trees. The existence of hybrids between *E. regnans* and the frequently co-occurring *Eucalyptus obliqua* L’Hér. (messmate stringybark) has long been known (Ashton, 1956), however the extent of hybridisation across the range of the species has never been investigated. Similarly, while the chloroplast genetic structure of *E. regnans* has been studied (Nevill et al., 2010), the structure of the nuclear genome has not. To address these knowledge gaps our aims were to (i) identify the extent and possible drivers of hybridisation across the geographic distribution of *E. regnans*, (ii) describe how identification of admixture using genomic data may influence our understanding of population genetic structure, and (iii) consider how these factors would influence current management strategies in eucalypts. We address these aims using genotyping-by-sequencing to obtain large numbers of genome-wide genetic markers on individual samples across the natural geographic range of the species. We predict that (i) some individuals and populations will show greater levels of admixture with *E. obliqua*, (ii) levels of admixture will be driven in part by local environmental variables, and (iii) the inclusion or exclusion of hybrid individuals in population genetic analyses will lead to different strategic outcomes for management. If these predictions are true, there are
implications for future studies of population genetic structure and the planning of restoration plantings and assisted gene flow.
3.3 Materials and methods

3.3.1 Study area and species

_Eucalyptus regnans_ grows in wet forests of the south-east Australian states of Victoria and Tasmania. It is the tallest angiosperm in the world, with reliable records of individuals exceeding 100 m (Beale, 2007; Hardy, 1918; Hardy, 1935). It is also a serotinous obligate seeder, requiring high-intensity fires to open the understorey, create fertile ash beds, and stimulate the mass release of seeds from the forest canopy (Ashton, 1981b; Ashton and Chinner, 1999). Without fire, trees are typically unable to produce offspring that survive to maturity, primarily due to predation of seeds by ants (Ashton, 1979; O'Dowd and Gill, 1984), low availability of light (Gilbert, 1959), browsing of seedlings by herbivores, and fungal infection of seedlings (Ashton and Macauley, 1972).

_Eucalyptus regnans_ is patchily distributed through a 700 km by 500 km area, growing only where climatic conditions are suitable (Cochrane, 1969). It reaches its highest elevations (>1100 m ASL) in the northernmost part of its range, on the Errinundra Plateau, and grows near to sea-level in some southern parts of its Tasmanian distribution. As the island of Tasmania has been separated from the Australian mainland by more than 200 km for over ten thousand years (Lambeck _et al._, 2014; Duncan _et al._, 2016), it is assumed that there has been very little or no gene flow between _E. regnans_ stands in these two regions for at least that length of time.

In a number of locations throughout Victoria and Tasmania, trees displaying intermediate characteristics between _E. regnans_ and other species have been recorded. These specimens have been identified mostly as hybrid individuals between _E. regnans_ and _E. obliqua_, and, less commonly, _E. regnans_ and _E. macrorhyncha_ (Ashton and Sandiford, 1988; Ashton, 1958; Ashton, 1981a). At least two individuals have also been found that appear to be tri-hybrids – the result of a _E. regnans_ x _obliqua_ hybrid individual mating with a _E. macrorhyncha_ (Yorke and Ashton, 1982). As red stringybark does not occur naturally in Tasmania, _E. regnans_ x _macrorhyncha_ hybrids do not occur there.

3.3.2 Sample collection

We collected 387 _E. regnans_ tissue samples from across its geographic distribution (Figure 3.1). At each of 16 sites, we walked a transect collecting tissue from trees spaced at least 20 m apart, until we had sampled 20 trees. We targeted trees with diameters at breast height of more than 60 cm, to avoid sampling younger trees that were propagated after the practice of reseeding logged coupes using seed of non-local provenance became common.
practice (Flint and Fagg, 2007). As the combined effects of logging and wildfires have caused a reduction in the size and frequency of old undisturbed patches of trees (Lindenmayer et al., 2016), a linear transect of fixed length was sometimes impossible. At seven of the sites we collected a second sample from the 20th tree, to serve as a technical replicate from the field. We were also able to incorporate an extra 42 *E. regnans* samples collected during fieldwork for other studies into some analyses, taken from various locations (Supporting Information Table S4.1). Twenty-one *E. obliqua* samples, taken from the Cathedral Range region in Victoria, were also sequenced to allow us to determine the extent of hybridisation between the two species. All samples were putatively identified as *E. regnans* or *E. obliqua* using purported diagnostic morphological characters (Brooker and Kleinig, 2006).

Tissue collected was mostly cambium, obtained by cutting through the rough and smooth bark using a machete, and slicing off a 10 x 5 x 0.2 cm strip. A small number of samples were leaf tissue, collected by climbing trees using standard arborist techniques and picking two fresh, growing leaves. All samples were air-dried in individual brown paper bags and then stored at 4°C prior to DNA extraction.
Figure 3.1 Map showing the overall distribution of *Eucalyptus regnans* (green shading) and *Eucalyptus obliqua* (grey shading) in the Australian states of Victoria and Tasmania, and the locations visited for collection of tissue samples for genotyping-by-sequencing. Species distributions are derived from records of each species found on the Atlas of Living Australia. The blue circles represent sites where *E. regnans* samples were collected, and the red circle represents the site where *E. obliqua* samples were collected. The number of samples collected at each site is indicated by the size of the circle.
3.3.3 Sample preparation and sequencing

Approximately 600 mg of each tissue sample was chilled to -65°C and homogenised using an automated tissue grinding machine (Labman Max Planck Cryogenic Grinder Dispenser, Labman Automation & Custom Robotics), before storage at -18°C to await DNA extraction.

Samples were ordered randomly and whole genomic DNA was extracted in plate format (Stratec Invisiorb DNA Plant HTS 96 Kit) by following the kit instructions. Library preparation for genotyping-by-sequencing included (1) digestion using PstI restriction enzyme (New England BioLabs Inc.), (2) ligation using T4 DNA ligase (New England BioLabs Inc.), (3) a purification step (Qiagen MinElute 96-well PCR purification kit), (4) PCR amplification using two GBS primers (Integrated DNA Technologies), (5) post-purification quantitation using microfluidic capillary electrophoresis (PerkinElmer LabChip GX II), (6) pooling of 12 ng DNA per sample using an automated robotic liquid handling machine (PerkinElmer NGS Express), and (7) a final purification step (Sigma-Aldrich Genelute PCR Clean-Up Kit).

Size fractionation, 250-450 base-pair gel cutout, and sequencing was conducted at the Australian Cancer Research Foundation Biomolecular Resource Facility (BRF) at the John Curtin School of Medical Research (Australian National University) on portions of four lanes (grouped with other E. regnans sequencing experiments) of an Illumina HiSeq 2500 machine using a 100-base paired-end read.

3.3.4 Demultiplexing and initial filters

Of the 408 samples (387 E. regnans and 21 E. obliqua), sequencing resulted in nearly 1.49 billion read pairs. We demultiplexed reads using exact matches and combinatorial index mode with Axe (Murray and Borevitz, 2018), and were unable to assign 7% of read pairs to a sample. We then used BBduk to remove adapters and quality-trim (Phred score Q=30) reads at both ends, and NextGenMap (Sedlazeck et al., 2013) to align reads to the E. grandis v2.0 reference genome (Bartholome et al., 2015; Myburg et al., 2014). We used SAMtools (Li et al., 2009) to convert the dataset into sample-specific Binary Alignment/Map (BAM) files and sort reads.

To create a sample-by-SNP matrix, we used the ‘ANGSD’ software package to first calculate genotype likelihoods (McKenna et al., 2010), and used these likelihoods to call genotypes. Loci were initially filtered based on (1) a probability of at least 99.999% that the site was variable, (2) the site was genotyped in at least 50 individuals, (3) the site had a minimum average sequencing depth per sample of 0.5, and (4) the site had a maximum
average sequencing depth per sample of 1000. Genotype likelihoods were retained and exported in BEAGLE file format for admixture analysis using ‘NGSAdmix’ (Skotte et al., 2013). Called genotypes were used for the remainder of the analyses, and were derived from the likelihoods based on a posterior genotype probability (≥ 0.95) and assuming a uniform prior. This produced a matrix containing 408 samples and 49622 SNPs, where the mean and median read depth per site per sample was 28.3 and 9.9 respectively.

Two separate filtering strategies were conducted on this dataset using the statistical software package R (R Core Team, 2017). The first of these developed a set of SNPs for investigating the extent of admixture with *E. obliqua*. The second method was used to investigate whether hybridisation influences population genetic structure and isolation-by-distance across the geographic distribution of *E. regnans*. Each of these filters has been discussed in the methods of the relevant analysis.

### 3.3.5 Extent of admixture

We investigated individual admixture proportions using three techniques, and averaged the results to improve accuracy and reliability. Firstly, the Bayesian clustering method in STRUCTURE v2.3.4 (Falush et al., 2003; Pritchard et al., 2000) was used, with a 50000 burn-in and 200000 Markov-Chain Monte-Carlo (MCMC) iterations, a *K* value of 2, using an admixture model and correlated allele frequencies. To obtain SNPs used in this analysis, we filtered on call rate (≥ 66% of samples genotyped) and minor allele frequency (MAF ≥ 0.01), retaining 2192 SNPs. Samples missing more than two thirds of these loci were removed from any further analysis, retaining 380 samples. Next, we used an eigen-analysis approach to investigate individual ancestries, using the `snpgdsAdmixProp` function of the ‘SNPRelate’ package (Zheng et al., 2012; Zheng and Weir, 2016), with the same 2192 SNPs and 380 samples used for the STRUCTURE analysis. Lastly, we used the expectation-maximisation algorithm in NGSadmix (Skotte et al., 2013), using the GATK genotype likelihoods (McKenna et al., 2010) produced by ANGSD, assuming two ancestral populations, and requiring the minor allele to be present in at least eight individuals, retaining 16634 loci. Barplots of admixture for all 380 samples allowed for visual comparison of each method. As the admixture proportions between the three methods were highly correlated (discussed in the Results section), we then averaged the admixture coefficients, and used these mean values to exclude or retain individuals for the remaining analyses. Samples with a *E. obliqua* ancestry coefficient greater than 0.1 were considered hybrids, with coefficients of 0.4-0.6 indicating intermediate levels of hybridisation, and 0.1-0.4 or 0.6-1 indicating closer affinity to *E. regnans* or *E. obliqua* respectively (Field et al., 2009; Melville et al., 2017). While this method of identifying hybrid individuals is unlikely to have completely removed admixture
from the study, we considered it sufficient to demonstrate how accounting for hybridisation and introgression in the analysis pipeline can influence the results of, and conclusions drawn from, genetic analyses. It is unfeasible and possibly inappropriate to try and completely remove all admixture, with gene flow between the two species possibly occurring throughout their recent evolutionary history.

As principal components analysis (PCA) is often used as an initial method of removing outlier samples (Jordan et al., 2017; Supple et al., 2018), we first checked whether PCA would be appropriate for the identification and removal of hybrid individuals from the dataset. We calculated pairwise Euclidean genetic distances using the 2192 SNPs and 380 samples and then performed a PCA using the `indpca` function of the ‘hierfstat’ (Goudet, 2005) package in R. The first two principal components for each sample were plotted using the ‘ggplot2’ (Wickham, 2009) package, with samples coloured by the mean level of admixture with *E. obliqua*, as determined above.

Next, we calculated mean admixture proportions for every population and mapped this across the landscape using the ‘ggplot2’ and ‘scatterpie’ (Yu, 2018) packages.

### 3.3.6 Environmental association with admixture

The location of all 359 putative *E. regnans* samples were uploaded into the Atlas of Living Australia’s (ALA) Spatial Portal (https://spatial.ala.org.au/#), and 15 environmental variables at each point were extracted. Environmental variables used in the ALA were collated or derived from various sources (Williams et al., 2010; De Vries, 2009; Xu and Hutchinson, 2013; Xu and Hutchinson, 2011; Williams et al., 2006). Variance inflation factors were used to remove variables that showed multicollinearity, and visual inspection of histograms for each variable allowed removal of two variables that showed very little variation across all individuals. This left eight variables remaining, including mean annual rainfall (RAIN), mean annual solar radiation (RAD), historical (pre-European) phosphorus availability (PHOS), topographic wetness index (TWI), mean maximum temperature of the hottest month (MAXTEMP), mean minimum temperature of the coldest month (MINTEMP), as well as topographic aspect, which was transformed into a north-south component (NORTH) and an east-west component (EAST). A binary response variable was also created that identified each sample as either a hybrid individual or a ‘pure’ *E. regnans*. All predictor variables were scaled (by subtracting the mean and dividing by the standard deviation) prior to fitting any models.

To identify whether any predictor variables showed non-linearity on the logistic scale, we fitted a binomial generalised additive model in R and plotted the component smooth
functions for each variable on the scale of the linear predictor. The MAXTEMP variable showed signs of non-linearity, and so was transformed with the inverse reciprocal, which improved linearity substantially.

A binomial generalised linear mixed effects model was then run in R, using the `glmer` function of the ‘lme4’ (Bates et al., 2015) package, using the binary response variable, the seven unmodified environmental predictor variables, and the new MAXTEMP predictor variable. The estimated tree age (AGE) for each sample was also included as a predictor to determine if this was relevant, and a random effect of stand was fitted to account for repeated sampling. Overdispersion was checked using the model residuals and degrees of freedom, and spatial autocorrelation was investigated in the model residuals using visual examination of a variogram created with the `variog` function of the ‘geor’ (Ribeiro Jr. and Diggle, 2016) package.

All possible sub-models of the global model were fitted and ranked by Akaike’s information criterion corrected for finite sample sizes (AICc) using the ‘MuMIn’ (Barton, 2016) package. Models with AIC values ≤ 2 above the top-ranked model were considered useful for inference (Hegyi and Garamszegi, 2011).

### 3.3.7 Influence of admixture on population structure

To determine the level of influence that unrealised hybridisation can have on population structure, we first filtered SNPs using call rate (≥ 0.4), minor allele frequency (MAF ≥ 0.01), and observed heterozygosity (≤ 0.5). The likelihood that each SNP does not deviate from Hardy-Weinberg equilibrium (HWE) was checked using the `HWChisqStats` function of the ‘HardyWeinberg’ (Graffelman, 2015) package. Any SNP out of HWE in more than three populations (where n ≥ 15) was removed from further analysis. In addition, the `snpgdsLDpruning` function in the ‘SNPRelate’ (Zheng et al., 2012) package was used to prune out SNPs using a linkage disequilibrium threshold of 0.5 and a sliding window of 5000 bp.

Filtering was done on two groups: (1) all samples including hybrids (but excluding the reference *E. obliqua* samples), and (2) pure *E. regnans* individuals (i.e. those with <10% admixture with *E. obliqua*). In the admixture-inclusive group, 2474 SNPs were retained. For the pure group, 2481 SNPs were retained. Any samples with more than 50% missing data were removed from each dataset, leaving 323 samples and 228 samples in the admixture-inclusive and admixture-free groups respectively. Visual inspection of hierarchical clustering dendrograms, created using the `hclust` function on a Euclidean genetic distance matrix, showed that all biological replicates were closely paired in both datasets (Supporting
Standard population genetic parameters were calculated for all loci in both groups, using the `basic.stats` function in the ‘hierfstat’ (Goudet, 2005) package. To compare the influence of including hybrid samples on these parameters, the means and standard errors of the inbreeding coefficient ($F_{IS}$), genetic differentiation ($F_{ST}$), expected heterozygosity ($H_E$), and observed heterozygosity ($H_O$) were calculated across all loci and plotted as barcharts.

Prior to further analysis, all sites containing less than 10 individuals were removed from both datasets, leaving the pure dataset with 15 sites and the admixture-inclusive dataset with 18 sites. Population genetic structure was assessed in both datasets using principal coordinates analysis (PCoA) of pairwise population genetic distances as well as through isolation-by-distance. Pairwise population Euclidean genetic distances were calculated by first converting genotypes to genpop objects using the `df2genind` and `genind2genpop` functions in the ‘adegenet’ (Jombart, 2008) package, and then using the `dist.genpop` function to produce the distance matrix. Principal components were created using the `cmdscale` function, with populations plotted using the ‘ggplot2’ package. To investigate isolation-by-distance, we calculated pairwise $F_{ST}$ between populations using the ‘Nei87’ method of the `genet.dist` function in the ‘hierfstat’ package, and regressed these distances against pairwise population geographic distance, calculated using the `earth.dist` function in the ‘fossil’ (Vavrek, 2011) package. Genetic differentiation and geographic distance were transformed to $F_{ST}/(1-F_{ST})$ and log(geographic distance) respectively, to allow for linear interpretation (Rousset, 1997).

### 3.3.8 Spatial structure and population genetics in *E. regnans*

Spatial structuring of genotypes was investigated using two methods; firstly, a Mantel test (Mantel, 1967) comparing pairwise genetic distance with the natural logarithm of pairwise geographic distance was performed using the `mantel.rtest` function in the R package ‘ade4’ (Dray and Dufour, 2007). We then used the spatial structure analysis function in GenAlEx v6.503 (Peakall and Smouse, 2006; Peakall and Smouse, 2012) to determine the maximum geographic distance at which genotypes show significant spatial autocorrelation. This was interpreted as the maximum distance at which the lower confidence interval of the spatial autocorrelation coefficient $r$ was greater than zero on the y-axis.

To investigate population genetics within *E. regnans*, we calculated means and standard errors of the number of alleles ($A$), number of effective alleles ($A_E$), $H_O$, $H_E$, and $F_{IS}$ for each of the 15 populations. We then calculated the pairwise genetic distance between all
individuals in these populations using the ‘Dch’ method of the `genet.dist` function. This distance matrix was then read into GenAlEx to conduct an analysis of molecular variance (AMOVA) with 999 permutations to determine the amount of genetic variation explained within and among sites.
3.4 Results

3.4.1 Extent of admixture

Despite NGSadmix using a different approach (likelihoods instead of called genotypes) and much larger number of loci, there was a very strong correlation of ancestry coefficients computed using that method and both STRUCTURE ($r = 0.94, p < 0.001$) and SNPRelate ($0.95, p < 0.001$), with barplots showing near-identical patterns between methods (Supporting Information Figure S4.1). There was an even stronger correlation between the results of STRUCTURE and SNPRelate ($0.98, p < 0.001$). Excluding the reference E. obliqua samples, SNPRelate, STRUCTURE and NGSadmix identified 59, 130, and 190 hybrid individuals (>10% admixture with E. obliqua) respectively. All hybrid samples identified in SNPRelate were also identified using STRUCTURE, but a small proportion of samples identified as hybrids in SNPRelate and STRUCTURE were not identified as such by NGSadmix. After calculating the mean value of the E. obliqua admixture coefficient for the three methods, 170 samples were identified as hybrids. Of these, 75 samples had E. obliqua admixture coefficients between 0.1 and 0.4, five had coefficients between 0.4 and 0.6, 11 had coefficients between 0.6 and 0.9, and 16 had coefficients of greater than 0.9.

Principal components analysis showed that highly admixed samples (e.g. those from Wilsons Promontory) were able to be easily identified due to their separation from the bulk of the E. regnans species cluster (Figure 3.2). Unfortunately, PCA methods were unable to clearly separate out individuals with low and moderate levels of admixture, which account for about two-thirds of the admixed samples. Pairwise genetic distances involving these individuals presumably fall within the natural variation of genetic distance within E. regnans, meaning that distance-based methods are more conservative in the identification and removal of hybrids.

The mean proportion of hybrids per sample (at sites where more than 10 samples were analysed) was 0.24. Not a single site was completely free of hybrids, but seven sites had just one hybrid. The mean E. obliqua ancestry coefficient of a site was 0.11 (±0.21), with sites varying considerably in the amount of admixture (Figure 3.3). Samples taken from Wilsons Promontory showed the highest degree of admixture, with a mean E. obliqua ancestry coefficient of 0.87. The Tasmanian sites on Bruny Island and the Tasman Peninsula also had mean E. obliqua ancestry coefficients greater than 0.1, at 0.42 and 0.15 respectively.
Figure 3.2 Principal components analysis of pairwise genetic distance between 380 putative *Eucalyptus regnans* and 20 putative *Eucalyptus obliqua* trees. *Eucalyptus regnans* was sampled from across the natural geographic range of the species. Euclidean genetic distances were calculated using 2192 single-nucleotide polymorphisms. Putative *E. regnans* samples are coloured by their proportion of admixture with *E. obliqua*. The reference *E. obliqua* samples are coloured black.
Figure 3.3 Map showing the mean proportion of admixture for 18 sites (with n>10) where *Eucalyptus regnans* was sampled. Pie charts show the mean proportion of admixture between *E. regnans* (yellow) and *Eucalyptus obliqua* (turquoise). Six geographically close sites sampled in the central region of Victoria were not significantly different in the amount of admixture (one-way ANOVA, F(5,108)=1.18, p=0.324), and so have been pooled here for clarity.
3.4.2 Environmental association with admixture

Model selection showed that hybrid individuals were not randomly distributed across all environmental variables. Of the 11 top models (ΔAICc < 2), MINTEMP and MAXTEMP were identified in all 11, suggesting that these two variables had the strongest effects on the probability of hybrid occurrence. Sites with a high MAXTEMP (hot summers) and sites with a low MINTEMP (cold winters) had lower probabilities of hybrid occurrence, whereas sites with a high MINTEMP and sites with a low MAXTEMP had higher probabilities of hybrid occurrence (Figure 3.4). Variables that occurred in fewer top models included NORTH, RAD, PHOS, RAIN, AGE and TWI, however none of these variables showed any clear trends with hybrid occurrence (Supporting Information Figure S4.3). EAST was not identified in any top models.
Figure 3.4 Effect plot showing the relationship of two predictor variables, the mean maximum temperature of the hottest month and the mean minimum temperature of the coldest month, with the probability that a *Eucalyptus regnans* individual will have greater than 10% genetic admixture with *Eucalyptus obliqua*. Grey areas indicate 95% confidence intervals.
3.4.3 Influence of admixture on population structure

The removal of hybrid individuals from the dataset modified the values of the genetic parameters investigated (Figure 3.5), with the parameters $H_E$, $H_O$, $F_{ST}$, and $F_{IS}$ increasing from 0.087, 0.092, 0.024, and -0.012 in the admixture-inclusive dataset, to 0.094, 0.097, 0.03, and -0.005 in the admixture-free dataset.

Removal of hybrids led to greater resolution of geographic population structure in the PCoA of pairwise population genetic distances, and a stronger pattern of isolation-by-distance from the Mantel test (Figure 3.6). In the admixture-inclusive dataset, the first principal coordinate was associated with the degree of admixture, with Wilsons Promontory and Bruny Island separated out from the rest of the sites. The high level of admixture at Wilsons Promontory resulted in comparatively much higher pairwise $F_{ST}$ values between this and other sites, leading to a reduction in fit of the isolation-by-distance model when this site was included in the analysis.
Figure 3.5 Comparison of genetic parameters, including the inbreeding coefficient (Fis), genetic differentiation (Fst), expected heterozygosity (He), and observed heterozygosity (Ho) between two datasets, one of which contains hybrid individuals (n=323) and another where hybrid individuals have been removed (n=228). The datasets were filtered using the same criteria, resulting in 2474 single-nucleotide polymorphisms when hybrids were included and 2481 when hybrids were removed. Error bars represent standard errors of the mean.
Figure 3.6 Plots of pairwise genetic distance and isolation by distance in *Eucalyptus regnans* populations across its natural geographic distribution. Plots on the left-hand side of the figure include the presence of hybrid individuals and populations. Plots on the right-hand side of the figure have had all hybrid individuals removed. The top two plots show the first two principal coordinates of Euclidean genetic distances, where squares, triangles, diamonds and circles respectively represent populations from the Otways, Tasmania, Central Victoria, and Gippsland regions of Australia. The bottom two plots indicate the effect of increasing geographic distance on genetic differentiation, with both axes transformed to allow visualisation using a linear relationship.
3.4.4 Spatial structure and population genetics in *E. regnans*

After removing all hybrids and discounting sites that had less than 10 samples, 200 individuals from 15 sites were retained. The spatial genetic structure analysis in GenAlEx identified weak but significant autocorrelation (determined by inspection of bootstrap confidence intervals) between genotypes at pairwise distances of up to 90 km, with the autocorrelation coefficient $r$ intercepting zero at 108 km. A Mantel test, comparing the logarithm of geographic distance with Nei’s genetic distance for all pairs of samples, showed that there was a relatively small but significant ($r=0.18$, $p=0.001$) positive effect of isolation-by-distance.

Within-population genetic structure at the 15 sites was very similar (Supporting Information Table S4.2), with little variation in $A$, $A_E$, $H_O$, $H_E$, or $F_{IS}$. Most sites had higher $H_O$ than $H_E$, and all sites had negative $F_{IS}$ values, implying that outbreeding is the common condition throughout the species distribution. Global $F_{ST}$ was very low (0.03), and approximately equal to the mean pairwise $F_{ST}$ between sites (0.028), despite some large (>600 km) pairwise geographic distances (Figure 3.6). Pairwise $F_{ST}$ values ranged from 0.003, between the geographically close Central Victorian sites of Powelltown and Toolangi South (39 km), to 0.063, between the geographically distant Errinundra Plateau and Gellibrand River (489 km).

Across the overall distribution of the species, *E. regnans* showed very little population structure, with the AMOVA indicating that just 3% of the genetic variation was occurring among populations, and 97% within populations. The first two coordinates of the PCoA explained 13.2% and 11.5% of the total variation, with the Tasmanian samples intermediate between sites in central-eastern Victoria and sites from the Otways region (Figure 3.6). The PCoA also identified the Errinundra Plateau site as being relatively distinct from the rest of the sites.
3.5 Discussion

Using a genome-wide SNP dataset, we investigated the extent of hybridisation between two widespread forest trees and the way in which admixture influences estimation of genetic parameters and interpretation of population genetic structure. We found that hybrids were not distributed evenly across geographic or environmental space, with some populations showing more admixture than others, and a strong association of hybrid occurrence with summer temperatures. Once hybrids were removed, we found very little evidence of population genetic differentiation or local isolation of alleles, with high levels of gene flow, slow generation times, and a lack of historical geographic isolation probably responsible for these findings.

This study has implications for our understanding of the processes involved in adaptation and assisted migration. Previous work has shown that species can not only adapt to changing environments through the development of multiple novel genetic pathways (Steane et al., 2017b), but also through a process of adaptive introgression (De La Torre et al., 2014b; Suarez-Gonzalez et al., 2018), whereby hybridisation with congeners allows for the capture of beneficial genetic components from parent species. For example, hybrid individuals between *Picea glauca* (Moench) Voss and *Picea engelmannii* Parry ex Engelm. have been shown to utilise adaptive introgression as a way to maximise fitness in a changing climate (De La Torre et al., 2014b). With environmental changes outpacing the natural ability of many species to adapt through standing genetic variation and mutation, adaptive introgression provides an alternative pathway, providing faster development of adaptive traits and rapid adaptation to novel conditions. It is in this context that we discuss our findings and explore their implications for studies of population genetics and genetically-informed conservation.

3.5.1 Hybridisation

Had we not considered the possibility of hybrid individuals occurring in our dataset, we would have included introgressed sites such as Wilsons Promontory and Bruny Island in the analyses, as well as 41 hybrid samples scattered through our other sites. Many of the hybrid individuals showed no obvious morphological differences (e.g. in bud and fruit shape or rough bark height and thickness) to pure specimens, with morphological approaches now generally regarded as a poor way of identifying hybrids (Field et al., 2009; McKinnon et al., 2010; Rhymer and Simberloff, 1996; Schwabe et al., 2015). Despite this, the Wilsons Promontory samples, which are morphologically more similar to *E. regnans* but genetically much closer to *E. obliqua*, had a different growth habit to typical *E. regnans* individuals, appearing stunted in form. This was initially thought to be the result of phenotypic plasticity.
rather than genetic architecture, as the trees do not resemble typical *E. obliqua* either. In our study, removing clear outliers using PCA or a similar method would have retained more than half of the hybrid samples. Including highly admixed samples and populations could have led us to make different inferences about patterns of genetic structure and possibly evolutionary history. As many modern studies of eucalypts (and other plant taxa) do not explicitly address the presence of hybridisation in their study species, it is plausible that many studies are influenced by this issue. We propose that future studies always be explicit in acknowledging the potential for hybridisation between their species of interest and other species, as gene flow between some species is clearly a common, and possibly evolutionarily significant, phenomenon (Gerber et al., 2014; Lepais et al., 2009; Palme et al., 2004; De La Torre et al., 2014a).

Eucalypts are typically preferentially outcrossing, open-pollinated, and often found in sympatry with multiple congeneric species, which may partly explain why more than half of all species form natural hybrid combinations (Griffin et al., 1988; Potts et al., 2003). In addition, many of these combinations can occur at relatively high frequencies within populations (Field et al., 2009; McKinnon et al., 2010); for example, at least 27% of *Eucalyptus globulus* Labill. within 450 m of *Eucalyptus cordata* Labill. show some level of admixture (McKinnon et al., 2010). However, despite the knowledge that hybridisation is not uncommon, it is often not considered in population genetic studies of eucalypts. We also note the possibility that for some species there may be multiple hybrid combinations with other species, which may vary regionally (Griffin et al., 1988). In this study we only considered admixture with a single species, however gene flow with red stringybark may regularly occur in particular sites within Victoria (Ashton and Sandiford, 1988).

In the case of *E. regnans*, hybridisation with *E. obliqua* appears to be a more pervasive phenomenon than previously realised, with all sampled sites containing at least one hybrid individual, and two sites where more than half of the samples were hybrids. The Wilsons Promontory individuals contained very high levels of *E. obliqua* ancestry, despite their greater morphological similarity to *E. regnans*. To understand why variation in the level of admixture between populations occurs, examination of the factors controlling gene flow between eucalypt species is necessary. Previous research has identified three predominant drivers of hybridisation in eucalypts, including the extent of geographic isolation, the degree of overlap in flowering times, and the level of phylogenetic divergence between species (Butcher et al., 2009; Barbour et al., 2006; Potts et al., 2003; Field et al., 2009; McKinnon et al., 2010), although strong abiotic (e.g. climatic or geological) gradients may also be explanatory in some cases (Pryor, 1976). When stands of one species are geographically isolated from conspecifics, there is a greater opportunity for pollen from another species to
successfully pollinate flowers in the isolated stand, termed ‘pollen swamping’ (Field et al., 2009; Ellstrand and Elam, 1993). Additionally, environmental conditions can influence the degree of overlap in flowering times, with temperature previously shown to be an important driver of flowering phenology in eucalypts (Hudson et al., 2010; Rawal et al., 2015; Law et al., 2000). As E. obliqua is a predominantly summer-flowering species, floral development in E. obliqua may occur earlier in sites with warmer summers, resulting in less phenological overlap with the autumn-flowering E. regnans. We therefore suggest that the primary cause of the extensive introgression of E. obliqua into E. regnans stands at Wilsons Promontory and Bruny Island is the patchy distribution of E. regnans in these E. obliqua-dominated regions, possibly assisted by environmentally-driven overlap in flowering times. The patchy distribution of E. regnans at these and other coastal sites (e.g. Tasman Peninsula) is likely due to their being located on the periphery of suitable climatic conditions for this species.

### 3.5.2 Patterns of genetic structure

The low levels of genetic population structure that we observed are not atypical for eucalypts, with a number of studies (Dillon et al., 2015; Supple et al., 2018; Broadhurst et al., 2018; Sampson et al., 2018; Gauli et al., 2014) finding that geographic structure does not contribute greatly to population differentiation. Similarly, the $F_{ST}$ values that we observed between sites are low but comparable to those found in other eucalypts (Supple et al., 2018; Sampson et al., 2018; Yeoh et al., 2012), and low geographic structuring of genetic diversity is not unusual in widespread forest trees, such as Pinus taeda L. (Eckert et al., 2010) and Quercus lobata Née (Grivet et al., 2008; Sork et al., 2010).

As gene flow in trees is often effected predominantly through pollen dispersal (Sork, 2016), there is often a clear difference in population structure between the nuclear genome (which is inherited biparentally) and genetic components that are inherited maternally, e.g. the chloroplast in angiosperms (Sampson et al., 2018). In eucalypts, seed dispersal is highly restricted, with individual trees distributing seeds tens of metres, and stand edges typically only capable of moving about 1-2 m per year (Booth, 2017). While it is difficult to ascertain the upper limit of the dispersal curve, pollen appears to regularly disperse hundreds of metres to kilometres (Sampson et al., 2018; Bloomfield et al., 2011; Broadhurst, 2013; Byrne et al., 2008). This explains why, when investigating chloroplast structure for phylogeographic purposes, Nevill et al. (2010) found a highly structured genetic pattern in E. regnans, whereas our (nuclear-based) results show very little population structuring. Chloroplast DNA is highly conserved, with genetic structure typically reflecting historical patterns of dispersal and colonisation. Further, only a very small proportion of chloroplast variation is typically...
contained within coding regions (Young et al., 2011), suggesting that there is unlikely to be substantial levels of local selection acting on the chloroplast.

Broadhurst et al. (2017) identify range disjunctions as being one of the primary predictors for genetic differentiation within the Australian flora. This holds true for many plant and animal species across the Bass Strait, with species found in Victoria and Tasmania often being identified as genetically distinct, typically to the level of separate races or subspecies (Gill and Donsker, 2018; van Dyck et al., 2013; Simmons et al., 2008). Even for species with low levels of isolation-by-distance, there are often clear Victorian and Tasmanian genetic clusters (Duncan et al., 2016). Our findings for E. regnans, with Gippsland and some Central Highlands sites more closely affiliated with Tasmanian sites than the Otways region, supports the suggestion that, for some species, the Port Phillip Bay and surrounding area has been a bigger obstacle to gene flow than the Bass Strait since the last glacial maximum (Yeoh et al., 2012).

While the pattern of very weak genetic population structure that we observed is indicative of a lack of local isolation of alleles, local adaptation may still be occurring. With the majority of genetic variation present in the seed crop (or mature trees) of a stand, selection may be acting to promote particular genotypes within a generation, as environmental conditions filter out particular alleles. There is a steep reduction in the stem density within E. regnans stands in the decades after a regeneration event (von Takach Dukai et al., 2018), which suggests that selection could easily promote genotypes that increase survival rates under local environmental conditions (Kuparinen et al., 2010). Previous studies have shown that when selection is very high, high levels of dispersal can maximise local adaptation, and when selection pressure is low (but present), local adaptation is highest under moderate levels of dispersal (Banks et al., 2017; Forester et al., 2016). This is due to migration providing the genetic diversity for selection to act upon, but also potentially overriding the effects of selective processes when selection is weak.

3.5.3 Implications for seed provenancing

Seed used for native vegetation restoration activities has historically been collected from small local geographic areas (Broadhurst et al., 2008), because of the perceived risk of introducing genotypes that are not adapted to local conditions (Hamilton, 2001). Over the past two decades, this ‘local is best’ approach has been criticised for a number of reasons, including the failure to consider changing environmental/climatic conditions and not incorporating enough genetic variability (Choi, 2007; Broadhurst et al., 2008). By using only locally-adapted genotypes, managers may be restricting the ability of populations to survive under changing conditions. To address this issue, the definition of what constitutes the most
appropriate geographic spread for seed collection to encompass a beneficial amount of genetic variation needs to be considered (Breed et al., 2013; Prober et al., 2015; Crow et al., 2018).

As eucalypts and wattles (Acacia spp.) dominate the majority of Australian ecological restoration schemes (Broadhurst et al., 2015), developing knowledge for these taxa is of critical importance. Our results, and recent work by others (Supple et al., 2018; Dillon et al., 2015; Bloomfield et al., 2011), suggests that for many Eucalyptus species, most of the total genetic variation is held within rather than among populations. This has large implications for the concept of local provenancing. For example, Breed et al. (2013) suggest that estimates of historical gene flow such as $F_{ST}$ values can be used to determine the extent of the local ($F_{ST} < 0.05$), intermediate ($0.05 < F_{ST} < 0.1$) or distant ($F_{ST} > 0.1$) provenances. Under this definition, the entire range of E. regnans would be considered local provenance. Because most of the genetic variation of the species is already contained within local geographic areas, the source of seed used for restoration activities is not likely to be as important as previously considered. While we acknowledge that reciprocal transplant experiments show a strong effect of provenance in some species (Wang et al., 2010), any negative effects of using distant seed sources must be considered in light of considerable recent evidence showing that locally-adapted traits can not only be maintained despite gene flow (Fitzpatrick et al., 2015), but also that the potential benefits of gene flow are large and often outweigh the risk of negative impacts associated with anthropogenic disturbances or novel selective pressures (Fisher et al., 2009; Harrisson et al., 2016). In many cases, the preservation of genetic uniqueness and taxonomic integrity is no longer considered scientifically justifiable (Ralls et al., 2018). Further, as rare new variants are unlikely to be the cause of beneficial adaptation to local conditions (Alberto et al., 2013; Savolainen et al., 2013), adaptive alleles are still likely to be present in seed collected from distant localities, simply at different frequencies. Reductions in the frequency of some adaptive alleles is typically of minor importance and can be naturally corrected over a small number of generations (Ralls et al., 2018; Fitzpatrick et al., 2015). Thus, we suggest that sourcing seed from trees across a range of environments (local or distant) will ensure adaptive potential for restoration into a changing and challenging future.

There are other considerations for seed sourcing that also need to be considered. For example, climate-adjusted provenancing has been suggested as a way of assisting plant populations with adaptation to climate change (Prober et al., 2015). To achieve this, seed sourcing for restoration would have to consider the future climate at a site and predict the best source populations to match that climate. However, there is one important consideration for this approach that has not been adequately considered in the recent literature: the issue of hybridisation and introgression. The possibility that some populations contain a higher degree
of admixture with another species cannot be ignored, because inter-specific gene flow is one way in which species may cope with rapidly changing abiotic environments (De La Torre *et al.*, 2014b). By selecting seed from populations that have ‘adapted’ to different climatic conditions, managers may actually be selecting seed that is simply admixed with another species. For example, our data indicates that the occurrence of *E. regnans x obliqua* hybrids is correlated with temperatures of the hottest and coolest months, so collection of seed based on temperature may inadvertently result in the collection of hybrid genotypes, or even near-pure *E. obliqua*, which is likely not the desired outcome of management actions. It should also be noted that while our sampling avoided young trees, contemporary seed crops may show different levels of admixture in some populations. For example, those in landscapes where logging and fire have modified the extent and age structure of *E. regnans* compared to *E. obliqua*, or where changes in temperature has increased overlap in flowering times. To be certain of the hybrid status at a proposed seed collection locality, genotyping of seedlings would be necessary.
3.6 Conclusion

Our population genomic analysis of *E. regnans* found widespread admixture of varying levels with a congener, suggesting regular hybridisation throughout the range of the species. As many genera of plants are known to form natural hybrids, it is critical that admixture and its role in the adaptive process is considered appropriately in population genetic studies, as introgressed populations may skew genetic analyses and affect management decisions. The combination of widespread hybridisation and high levels of gene flow in *E. regnans*, with similar results having been found for a number of other eucalypt species, suggests that introgressive adaptation through porous genomes may be a common way for this taxon to adapt to rapid environmental change in climate and fire regimes. Selection on hybrids expressing traits harvested from sympatric congeners may allow for rapid adaptive change to new conditions. Furthermore, as the occurrence of hybrid individuals was not distributed evenly across geographic or climatic space, the use of climatic variables to select genotypes for assisted migration may not be the most appropriate way to manage eucalypts for conservation purposes, and requires more detailed consideration.
3.7 Acknowledgements

This project was funded by an Australian Research Council Future Fellowship (FT130100043) to SB and a Lesslie Research Scholarship in Landscape Conservation and Ecology to BvTD. Thanks to Rebecca Taylor and Oliver King for their assistance and company during fieldwork, and Wade Blanchard for statistical advice. We also appreciate the helpful and constructive comments provided by three anonymous reviewers.
3.8 Data archiving

Sequence data for all individuals are deposited to the NCBI Sequence Read Archive (SRA). A CSV file containing location data and called genotypes from ANGSD for all individuals are available at the Dryad Digital Repository: doi:10.5061/dryad.445m9j4
Chapter 3: Pervasive admixture in eucalypts

3.9 References


Chapter 3: Pervasive admixture in eucalypts


Handy, A. D., 1918. Tall trees of Australia, Victorian Naturalist, 35: 46-55.


Chapter 3: Pervasive admixture in eucalypts


Chapter 3: Pervasive admixture in eucalypts


https://www.nature.com/articles/nature04246#supplementary-information


Ribeiro Jr., P. J. and Diggle, P. J., 2016. geoR: a package for geostatistical analysis, R package version 1.7-5.2. Available at: http://cran.r-project.org/package=geoR


https://www.nature.com/articles/nature25508#supplementary-information


Wickham, H., 2009. ggplot2: elegant graphics for data analysis, Springer-Verlag New York,


Xu, T. and Hutchinson, M. F., 2011. ANUCLIM, software, version 6.1, Fenner School of Environment and Society, Australian National University, Canberra.


### 3.10 Supporting information

Table S4.1 Sites visited across the natural distribution of Eucalyptus regnans in the Australian states of Victoria and Tasmania. The estimated age of trees, the number of samples collected for DNA extraction, the mean latitude and longitude and the mean elevation are also provided.

<table>
<thead>
<tr>
<th>State</th>
<th>Region</th>
<th>Site name</th>
<th>Estimated age of trees (years)</th>
<th>No. samples</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Altitude (m a.s.l.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victoria</td>
<td>Otways</td>
<td>Maits Rest</td>
<td>200+</td>
<td>4</td>
<td>-38.75</td>
<td>143.56</td>
<td>231</td>
</tr>
<tr>
<td>Victoria</td>
<td>Otways</td>
<td>Gellibrand River</td>
<td>77</td>
<td>19</td>
<td>-38.64</td>
<td>143.67</td>
<td>496</td>
</tr>
<tr>
<td>Victoria</td>
<td>Otways</td>
<td>Erskine Falls</td>
<td>77</td>
<td>21</td>
<td>-38.52</td>
<td>143.92</td>
<td>339</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Maroondah north</td>
<td>77</td>
<td>13</td>
<td>-37.59</td>
<td>145.61</td>
<td>678</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Maroondah south</td>
<td>77-200+</td>
<td>24</td>
<td>-37.67</td>
<td>145.66</td>
<td>748</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>O’Sshannassy</td>
<td>77-200+</td>
<td>13</td>
<td>-37.67</td>
<td>145.78</td>
<td>803</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Powelltown</td>
<td>77-90</td>
<td>25</td>
<td>-37.82</td>
<td>145.85</td>
<td>724</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Toolangi south</td>
<td>77-90</td>
<td>19</td>
<td>-37.57</td>
<td>145.55</td>
<td>775</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Toolangi north</td>
<td>50</td>
<td>20</td>
<td>-37.52</td>
<td>145.50</td>
<td>701</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Upper Yarra north</td>
<td>77</td>
<td>8</td>
<td>-37.60</td>
<td>145.94</td>
<td>841</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Upper Yarra south</td>
<td>77</td>
<td>8</td>
<td>-37.77</td>
<td>146.12</td>
<td>847</td>
</tr>
<tr>
<td>Victoria</td>
<td>Central Victoria</td>
<td>Dandenong Ranges</td>
<td>200+</td>
<td>20</td>
<td>-37.83</td>
<td>145.36</td>
<td>540</td>
</tr>
<tr>
<td>Victoria</td>
<td>Gippsland</td>
<td>Wilsons Promontory</td>
<td>50-60</td>
<td>21</td>
<td>-39.03</td>
<td>146.40</td>
<td>268</td>
</tr>
<tr>
<td>Victoria</td>
<td>Gippsland</td>
<td>Tarra Bulga</td>
<td>100-250</td>
<td>21</td>
<td>-38.43</td>
<td>146.57</td>
<td>607</td>
</tr>
<tr>
<td>Victoria</td>
<td>Gippsland</td>
<td>Mount Elizabeth</td>
<td>77</td>
<td>21</td>
<td>-37.49</td>
<td>147.93</td>
<td>871</td>
</tr>
<tr>
<td>Victoria</td>
<td>Gippsland</td>
<td>Errinundra Plateau</td>
<td>100</td>
<td>21</td>
<td>-37.28</td>
<td>148.97</td>
<td>986</td>
</tr>
<tr>
<td>Tasmania</td>
<td>North</td>
<td>Dial Range</td>
<td>50-70</td>
<td>21</td>
<td>-41.25</td>
<td>146.07</td>
<td>211</td>
</tr>
<tr>
<td>Tasmania</td>
<td>North</td>
<td>Frome</td>
<td>55-65</td>
<td>20</td>
<td>-41.16</td>
<td>147.90</td>
<td>291</td>
</tr>
<tr>
<td>Tasmania</td>
<td>Central</td>
<td>Mount Field</td>
<td>50-200+</td>
<td>21</td>
<td>-42.68</td>
<td>146.71</td>
<td>450</td>
</tr>
<tr>
<td>Tasmania</td>
<td>South</td>
<td>Tasman Peninsula</td>
<td>55-65</td>
<td>20</td>
<td>-43.10</td>
<td>147.89</td>
<td>140</td>
</tr>
<tr>
<td>Tasmania</td>
<td>South</td>
<td>Arve Valley</td>
<td>60-200+</td>
<td>7</td>
<td>-43.13</td>
<td>146.76</td>
<td>323</td>
</tr>
<tr>
<td>Tasmania</td>
<td>South</td>
<td>Bruny Island</td>
<td>50+</td>
<td>20</td>
<td>-43.39</td>
<td>147.31</td>
<td>242</td>
</tr>
</tbody>
</table>
Figure S4.1 Barplots showing proportion of *Eucalyptus obliqua* in 21 putative *E. obliqua* and 359 putative *Eucalyptus regnans* tissue samples, genotyped using genotyping-by-sequencing. Ancestry coefficients were determined by SNPRelate (top), STRUCTURE (middle) and NGSadmix (bottom). Each individual is coloured according to the site at which it was collected. All reference *E. obliqua* samples belong to the CathRange population. Individuals were initially identified as either *E. obliqua* or *E. regnans* in the field using morphological characters.
Figure S4.2 Hierarchical clustering dendrogram of 323 *Eucalyptus regnans* samples, including individuals admixed with *Eucalyptus obliqua*. Computations were done using 2474 single-nucleotide polymorphisms from across the genome. Technical replicates are identified with an x or y after the sample name and are paired together on long branches.

Chapter 3: Pervasive admixture in eucalypts
Figure S4.3 Plots showing the relationship between seven (scaled and centred) environmental variables and the predicted probability that a *Eucalyptus regnans* individual will be admixed (i.e. have an admixture coefficient with *Eucalyptus obliqua* >0.1). Results are derived from a mixed effects logistic regression. Missing data was included in the global model to rule out the possibility that it was influencing the identification of admixed individuals. Grey areas indicate 95% confidence intervals.
Table S4.2 Genetic diversity parameters (derived from 2481 single-nucleotide polymorphism markers) for 200 *Eucalyptus regnans* individuals in 15 populations from across the species geographic distribution. Mean values and standard errors are presented for each population, after removal of all admixed individuals.

<table>
<thead>
<tr>
<th>Site</th>
<th>N</th>
<th>A</th>
<th>A_E</th>
<th>H_E</th>
<th>H_O</th>
<th>F_IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dandenong Ranges</td>
<td>14</td>
<td>1.426</td>
<td>1.122</td>
<td>0.085</td>
<td>0.089</td>
<td>-0.053</td>
</tr>
<tr>
<td>Dial Range</td>
<td>12</td>
<td>1.363</td>
<td>1.118</td>
<td>0.081</td>
<td>0.087</td>
<td>-0.08</td>
</tr>
<tr>
<td>Errinundra Plateau</td>
<td>10</td>
<td>1.293</td>
<td>1.124</td>
<td>0.119</td>
<td>0.096</td>
<td>-0.126</td>
</tr>
<tr>
<td>Erskine Falls</td>
<td>11</td>
<td>1.341</td>
<td>1.122</td>
<td>0.098</td>
<td>0.095</td>
<td>-0.111</td>
</tr>
<tr>
<td>Frome</td>
<td>11</td>
<td>1.384</td>
<td>1.127</td>
<td>0.088</td>
<td>0.094</td>
<td>-0.085</td>
</tr>
<tr>
<td>Gellibrand River</td>
<td>13</td>
<td>1.395</td>
<td>1.133</td>
<td>0.113</td>
<td>0.103</td>
<td>-0.101</td>
</tr>
<tr>
<td>Mount Elizabeth</td>
<td>16</td>
<td>1.494</td>
<td>1.13</td>
<td>0.092</td>
<td>0.099</td>
<td>-0.059</td>
</tr>
<tr>
<td>Maroondah north</td>
<td>12</td>
<td>1.409</td>
<td>1.123</td>
<td>0.087</td>
<td>0.094</td>
<td>-0.077</td>
</tr>
<tr>
<td>Maroondah south</td>
<td>19</td>
<td>1.529</td>
<td>1.121</td>
<td>0.088</td>
<td>0.089</td>
<td>-0.03</td>
</tr>
<tr>
<td>Mount Field</td>
<td>12</td>
<td>1.397</td>
<td>1.124</td>
<td>0.086</td>
<td>0.092</td>
<td>-0.073</td>
</tr>
<tr>
<td>O'Shannassay</td>
<td>11</td>
<td>1.407</td>
<td>1.123</td>
<td>0.086</td>
<td>0.093</td>
<td>-0.078</td>
</tr>
<tr>
<td>Powelltown</td>
<td>14</td>
<td>1.463</td>
<td>1.133</td>
<td>0.104</td>
<td>0.101</td>
<td>-0.066</td>
</tr>
<tr>
<td>Tarra Bulga</td>
<td>14</td>
<td>1.453</td>
<td>1.127</td>
<td>0.09</td>
<td>0.096</td>
<td>-0.063</td>
</tr>
<tr>
<td>Toolangi south</td>
<td>15</td>
<td>1.437</td>
<td>1.126</td>
<td>0.1</td>
<td>0.095</td>
<td>-0.067</td>
</tr>
<tr>
<td>Toolangi north</td>
<td>16</td>
<td>1.484</td>
<td>1.126</td>
<td>0.094</td>
<td>0.096</td>
<td>-0.059</td>
</tr>
</tbody>
</table>

N, sample size; A, number of alleles; A_E, effective number of alleles; H_E, expected heterozygosity; H_O, observed heterozygosity; F_IS, inbreeding coefficient.
Chapter 4: The influence of fire and silvicultural practices on the landscape-scale genetic structure of an Australian foundation tree species


Keywords: Disturbance; fire; logging; mountain ash; SNP; genetic diversity
4.1 Abstract

Natural disturbance regimes in forest ecosystems are being rapidly modified by anthropogenic pressures, including silvicultural practices and climate change. Australian forests dominated by mountain ash (Eucalyptus regnans) are critically endangered, with wildfires and logging predicted to cause ecosystem collapse within the next 50 years. To investigate the influence of disturbance on patterns and extent of genetic diversity in mountain ash, we compare replicated sites with three different disturbance histories (undisturbed, burnt, and logged). We employ genetic analysis at five chloroplast microsatellite loci and 1240 nuclear single-nucleotide polymorphisms (SNPs) to estimate within- and among-population genetic diversity, and assess the extent of fine-scale spatial genetic structure among individuals, for the three disturbance treatments.

Consistent with the expectation of extensive pollen dispersal but limited seed dispersal, we detected low levels of genetic differentiation at nuclear SNPs ($F_{ST} = 0.013$), and very high levels of differentiation at cpDNA microsatellites ($F_{ST} = 0.751$). The patterns of genetic structure were similar among treatments at nuclear SNPs, although logged sites showed higher relatedness between individuals and had a lower proportion of polymorphic loci than undisturbed and burnt sites. Analysis of cpDNA revealed significantly higher levels of total and within-site genetic diversity in the logged treatment than the burnt or undisturbed treatments, with haplotypes entering the system via the use of non-local seed in the regeneration process. We suggest that artificial regeneration activities should utilise a greater number of maternal parents, which could be achieved via variable retention harvesting or utilising a regional admixture provenancing approach.
### 4.2 Introduction

Disturbance regimes are critical components of many ecosystems, affecting the patterns and extent of spatial and temporal heterogeneity of ecosystem structural and functional attributes, which in turn control population and community dynamics (Sousa, 1984; Turner, 2010a). Changes to historical disturbance regimes can therefore have substantial impacts on species abundance and community composition of ecosystems (Connell, 1978). Recently, anthropogenic activities such as land clearing, invasive species, and climate change have been implicated in changes to the frequency, intensity, and extent of disturbance events around the world (Seidl et al., 2017; Alfaro et al., 2014).

In forest ecosystems, some of the most common naturally-occurring disturbances are fires, wind storms, and insect attacks (Seidl et al., 2014; Franklin et al., 2002). However, the historical pattern of these disturbances has been, or is being, rapidly modified by anthropogenic activities (Lindenmayer et al., 2011; Seidl et al., 2016). In addition, timber harvesting and habitat degradation are modifying the spatial distribution, structure and composition of forests worldwide (Edwards et al., 2014; Foster et al., 1998).

Changes to disturbance regimes in forests not only have flow-on effects on species abundance and community composition (Clarke, 2002b), but also on the spatial distribution of genetic diversity, through either selective processes or selectively-neutral demographic processes (Banks et al., 2013). Selective processes modify genetic structure by selecting for particular behavioural or physiological traits (Clarke, 2002a), whereas selectively-neutral processes modify genetic structure through direct effects (e.g. through population bottlenecks or redistribution of individuals) or indirect effects (e.g. manipulation of habitat suitability or permeability) (Brown et al., 2013; Templeton et al., 2011). As the level of standing genetic variation in a population is likely to be a primary source of adaptive potential in a changing environment (Hoffmann and Sgrò, 2011), we need to determine how modified disturbance regimes influence the spatial distribution of genetic diversity, particularly in species that have a strong role in structuring ecosystems.

Wet forests dominated by mountain ash (*Eucalyptus regnans*), the world’s tallest flowering plant and a foundation tree species (sensu Ellison et al., 2005), are widespread in the south-eastern Australian states of Victoria and Tasmania. The two predominant types of disturbances in these forests are wildfire and logging (Lindenmayer et al., 2015). Mountain ash is a fire-dependent species, in that it is killed by intense fire, after which a new cohort regenerates from seed stored in serotinous capsules in the canopy (von Takach Dukai et al., 2018). Thus, mountain ash typically (although not exclusively) occurs as single-aged stands of trees that regenerated after the most recent wildfire (Turner et al., 2009). Historically, the
fire regime in this system was one of infrequent (75-150 year intervals), high-intensity (35,000–100,000 kW m⁻¹) fires (Gill and Moore, 1990; McCarthy et al., 1999). However, in the core of the Victorian range of mountain ash (the ‘Central Highlands’), there have been major wildfires in 1932, 1939, 1983 and 2009, and it is likely that this trend of frequent fire will continue into the future. Logging for paper pulp and timber production is prevalent in this region, with about 20% of the entire mountain ash forest estate (about 157,000 ha) located in closed water catchments or national parks, and most of the remaining 80% subjected to paper pulp and timber production (Burns et al., 2015). Together, wildfires and logging have left approximately 1% of the total area of mountain ash forest remaining as old-growth (i.e. patches where the oldest age class of trees are > 120-150 years old), which historically would have been around 30-60% (Lindenmayer, 2009a; Lindenmayer et al., 2012a).

Our aim was to identify the effect of logging and fire history on measures of genetic diversity, fine-scale genetic structure, and population genetic differentiation in mountain ash stands. Predicting the genetic effects of ecological disturbance depends on a number of factors, with dispersal distance and the number of in situ surviving individuals or propagules being particularly important (Davies et al., 2016; Banks et al., 2017). As with ecologically similar eucalypt species (Bloomfield et al., 2011), it is likely that pollen-mediated nuclear gene flow is very high in mountain ash, but analysis of range-wide cpDNA structure suggests limited seed dispersal (Nevill et al., 2010). We expected that silvicultural practices would lead to different patterns of genetic diversity in logged sites compared to burnt and naturally-regenerated sites. This is because logged sites cannot regenerate from canopy-stored seedbanks and need to be artificially sown with seed (typically via use of a helicopter) to ensure regeneration. Knowing this, there are a number of possibilities as to how genetic structure in logged sites could be changing, including (1) the invasion of non-local alleles due to seed for artificial regeneration being sourced from distant localities, leading to altered patterns of diversity and structure among logged sites (Potts et al., 2003), or (2) a reduction in genetic diversity resulting from the regeneration of a small number of (local or non-local) genotypes (Glaubitz et al., 2003a; Buchert et al., 1997; Soliani et al., 2016), due to either the use of few maternal trees for seed collection or the effects of local selective processes. Either of these scenarios could have flow-on effects to unlogged forest, whereby trees in logged sites pollinate trees in, or disperse seeds into, neighbouring stands. Regenerating stands of unlogged forest that were burnt in the 2009 fires could thus potentially be very different in genetic structure to the generations of trees that germinated prior to the proliferation of artificial seeding practices. We expected here that the effects of logging would differ between nuclear and chloroplast genomes, with the naturally higher levels of structure and lower
within-site diversity of chloroplast DNA making it more susceptible to seed provenancing strategies.

We investigated the patterns and extent of genetic diversity and genetic differentiation among populations at both chloroplast and nuclear DNA loci in forest sites subjected to timber harvesting, and compared these to (1) sites that had regenerated naturally after fires prior to 1940, had burnt in the 2009 fire, and had not experienced any known timber harvesting, and (2) sites that had regenerated naturally after fires prior to 1940 and had not experienced any major disturbance since. We use spatially- and temporally-replicated contrasts of sites with different disturbance histories as a method of maximising information about the impacts of disturbance on genetic structure (Banks et al., 2013). Additionally, whilst Nevill et al. (2010) have examined the pre-disturbance genetic structure of mountain ash using chloroplast DNA, our study provides an interesting contrast to their data while also providing new information on the nuclear DNA using next-generation sequencing and a recently-described method (Smouse et al., 2017) for investigating genetic diversity among and within strata.
4.3 Materials and methods

4.3.1 Sampling and DNA extraction

We collected tissue samples from a total of 31 sites in the Central Highlands of Victoria (Figure 4.1). Nine (“undisturbed”) sites had not experienced major disturbance (e.g. high-severity wildfire or timber harvesting) in at least 50-60 years and typically more than 75 years. Eight (“burnt”) sites had been burnt by wildfire in 2009, and were typically composed of dense, regenerating saplings. The remaining 14 (“logged”) sites had been logged between 2003 and 2008 using a clear-felling method, after which the remaining woody debris (slash) was burnt and the site was sown with mountain ash seeds (Flint and Fagg, 2007).

Due to the large spatial scale and stochastic nature of the 2009 wildfires, as well as the location of regions open to logging, we were limited in our ability to choose the location of disturbance ‘treatments’. This is a typical problem with natural experiments (Driscoll et al., 2010; Smith et al., 2016), with no simple solution. However, we consider our landscape-scale sampling design and data analysis methods appropriate for testing our particular hypotheses. We chose undisturbed sites that cover a wider geographic area than the disturbed sites to reduce the possibility of missing naturally-occurring genetic variation in the landscape.

We collected ten leaf or cambium tissue samples from each site, with a sample taken from the closest tree at every 20 m interval along a 180 m transect. As we were collecting extra samples from four of the undisturbed sites and five of the burnt sites (for a future study of dispersal patterns) we were able to incorporate an extra 72 samples into the current study. Thus, a total of 382 samples was collected for analysis, of which 117 were from undisturbed sites, 125 were from burnt sites, and 140 were from logged and seeded sites. In addition, 20 messmate (Eucalyptus obliqua) samples were collected to allow us to genetically identify hybrid individuals. During fieldwork, tissue samples were placed into small paper bags and dried in sealed zip-lock bags filled with silica gel beads.

In the laboratory, approximately 600 mg of each leaf sample was placed into a 20 mL scintillation vial and stored at -80°C overnight before mechanical disruption. Disruption and homogenisation were conducted using an automated cryogenic (-65°C) tissue grinder and dispensing system (Labman Automation Ltd). Total genomic DNA was then extracted from the powdered tissue using DNeasy Plant Mini Kits (Qiagen, Hilden, Germany) by following the manufacturer’s protocol, although we extended the incubation in lysis solution to one hour.
Figure 4.1 Map of the Central Highlands of Victoria, in Australia's southeast, showing the locations of each site sampled for mountain ash (Eucalyptus regnans). Sites have been separated into one of three disturbance categories (treatments) based on the disturbance history at each location.
4.3.2 Genotyping-by-sequencing

Approximately 400 ng of genomic DNA from each sample was used to prepare a reduced representation library for genotyping-by-sequencing (Elshire et al., 2011). This and similar next-generation sequencing methodologies are typically used to obtain thousands to millions of unbiased single-nucleotide polymorphisms (SNPs) from across the genome. Using large numbers of SNP loci allows for highly accurate estimation of genetic parameters, inference of population structure and calculation of relatedness metrics that are suitable for this study (Putman and Carbone, 2014; Fischer et al., 2017). Briefly, the library preparation consisted of (1) digestion of genomic DNA using the PstI restriction enzyme (New England BioLabs Inc.), (2) ligation of fragments with a unique barcode adapter pair, (3) individual purification and PCR amplification, (4) quantification using microfluidic capillary electrophoresis (PerkinElmer Labchip GX-II), (5) equimolar pooling using an automated robotic liquid handling machine (PerkinElmer NGS Express), and (6) a final purification step using a Genelute PCR Clean-Up Kit (Sigma-Aldrich). The pooled samples were then transferred to the John Curtin School of Medical Research (John Curtin) at The Australian National University, where they were size fractionated and a 250-450 bp gel cutout performed. Sequencing was done on multiple lanes of an Illumina HiSeq 2500 platform using a 100-base paired-end read, along with hundreds of additional mountain ash samples that were being used for separate projects.

All incubation steps and PCR amplification were done in Eppendorf Mastercycler Nexus machines. Library preparation was also done using only filter tips to minimise risk of contamination. To minimise plate effects in data, samples were randomised among plates and sequencer lanes, and technical replicates from the field were used to ensure consistency in the dataset.

4.3.3 Bioinformatic pipeline

Initial bioinformatics was undertaken by the ANU Bioinformatics Consultancy, at the John Curtin School of Medical Research of The Australian National University. Reads were demultiplexed using Axe (Murray and Borevitz, 2018), mapped to the E. grandis v2.0 reference genome (Bartholome et al., 2015; Myburg et al., 2014) using NextGenMap (Sedlazeck et al., 2013), and then converted to Binary Alignment/Map (BAM) files with sorted reads using SAMtools (Li et al., 2009).

A SNP by sample matrix was produced from BAM files using the program ANGSD (Korneliussen et al., 2014), by choosing loci that were highly likely to be polymorphic (likelihood ratio test p-value ≤0.001), were genotyped in at least 50 individuals, had an
average sequencing depth per sample ≥0.5 and ≤1000, and had a posterior genotype probability ≥0.95. Posterior genotype probabilities assumed a uniform prior. This retained 49622 SNPs across all samples, with a mean read depth per locus per sample of 15.9 and median read depth per locus per sample of 7.9. The SNP-by-sample matrix was then read into the statistical analysis software R (R Core Team, 2017), for further filtering and analysis.

4.3.4 Hybridisation analysis

To identify putative hybrids, we first filtered the SNP-by-sample matrix using call rate (≥ 30% of samples genotyped per locus) and minor allele frequency (MAF ≥ 0.01), and then calculated $F_{ST}$ values at all loci between the 20 messmate individuals and a group of 24 mountain ash individuals (randomly sampled from across the range of the species, that were sequenced for a separate project). We retained loci that were positively differentiated between species ($F_{ST}$ values ≥0.05), leaving 2778 SNPs. Five samples were removed from the dataset here due to high levels of missing data (<75% of loci genotyped). The remaining 353 mountain ash and 20 messmate samples had their ancestry coefficients estimated in the Bayesian clustering program STRUCTURE v2.3.4 (Pritchard et al., 2000; Falush et al., 2003), using a K value of 2 and a burn-in of 20000, followed by 200000 Markov-Chain Monte-Carlo iterations.

Mean values of admixture for each treatment were calculated, and a single-factor analysis of variance (ANOVA) was used to determine whether there was a significant difference in the level of admixture between treatments. Two linear mixed-effect models were also fitted; one in which admixture was predicted by treatment type and a random effect of site, and an intercept only (null) model where just the random effect of site was used. A coefficient of determination (pseudo-$R^2$) was calculated for the global model, using the $r.squaredGLMM$ function of the ‘MuMIn’ (Barton, 2016) package, to estimate the proportion of variance explained by the fixed and random effects. Akaike’s Information Criterion corrected for finite sample sizes ($AIC_c$) were compared between models to determine whether the inclusion of treatment was an important predictor of admixture.

4.3.5 GBS data analysis

After removing all samples clearly identified as hybrid individuals (admixture coefficients with messmate ≥ 0.1), 347 mountain ash samples were retained for further analysis. Using these samples, loci were filtered using call rate (> 85% of samples genotyped per locus) and minor allele (MA) count (MA genotyped ≥ 3 times), retaining 1240 loci. All samples with missing data representing more than 20% of the loci were then removed, retaining 327 samples.
Standard genetic diversity metrics, including the number of alleles ($N_A$), number of effective alleles ($N_E$), inbreeding coefficient ($F_{IS}$), expected heterozygosity ($H_E$), and observed heterozygosity ($H_O$), were calculated for each site using GenAlEx (Peakall and Smouse, 2006; Peakall and Smouse, 2012), version 6.51b2. The mean MAF and the proportion of polymorphic loci ($P_P$) were calculated for each site in R. All genetic diversity metrics were then averaged across sites within each treatment to obtain treatment level means and standard errors. A generalised linear model (GLM) was conducted with $P_P$ as a response variable and treatment type as a predictor variable. The Anova function from the ‘car’ package was then used on the model to determine whether there was a significant difference in $P_P$ between treatments. To determine whether there was a significant difference in genetic variance between treatments, an analysis of molecular variance (AMOVA) was performed using the adonis function of the ‘vegan’ (Oksanen et al., 2018) package, using treatment as a stratum, with 999 permutations of the raw data used to estimate the P value.

The SNP-by-sample matrix was converted to a genind object using the df2genind function in the ‘adegenet’ (Jombart, 2008) package, and a global $F_{ST}$ value calculated. To determine the proportion of genetic variance occurring among and within sites for each treatment, the genind object was first subset into the three separate treatments. An AMOVA was then conducted for each treatment with site as a stratum, using the poppr.amova function of the ‘poppr’ (Kamvar et al., 2014) package, with 999 permutations used to determine P values.

The genind objects for each treatment were transformed using the genind2hierfstat function of the ‘hierfstat’ (Goudet, 2005) package. Genetic differentiation, in the form of Nei’s $F_{ST}$, was then calculated independently for each treatment using the basic.stats function.

To investigate the partitioning of genetic diversity across the treatments, we used the QDiver function (Smouse et al., 2017) of GenAlEx. This approach computes standardised genetic diversity metrics partitioned into hierarchical strata, and allows for evaluation of homo/heterogeneity of within-stratum diversity components (Smouse et al., 2017). We also used GenAlEx to compare the extent of spatial autocorrelation of genotypes between treatments, via the distance-based Spatial function. Spatial autocorrelation coefficients for six distance classes (0.2, 1, 5, 10, 20, and 40 km) were used, with results for each treatment (including 95% confidence intervals) read into R and correlograms created using the ‘ggplot2’ (Wickham, 2009) package. Similarities and differences in the patterns of genetic structuring between treatments were inferred by inspecting the bootstrap-estimated 95% confidence intervals and permutation tests for departure from a ‘null’ correlogram.
4.3.6 Microsatellite genotyping

We used five chloroplast microsatellites (EMCRC60cp, EMCRC67cp, EMCRC74cp, EMCRC86cp, and EMCRC90cp) to genotype the 347 samples with less than 10% admixture (identified by the STRUCTURE analysis). These microsatellites were found to be polymorphic with strong and reproducible amplification profiles for mountain ash samples in a previous study (Nevill et al., 2010). Primers for microsatellites were purchased (GeneWorks Pty Ltd, Thebarton, Australia) using the descriptions published in Steane et al. (2005), with a universal M13 tail attached to the forward primer, to allow for multiplexing.

Total reaction volume for the polymerase chain reaction (PCR) step was 12.5 µL, including 1.5 µL dNTP mix (Qiagen) at 2 mM, 1.5 µL MgCl₂ (Qiagen), 1 µL bovine serum albumin, 1.25 µL 10 x PCR buffer (Qiagen), 0.25 µL reverse primer, 0.1 µL forward primer (with M13 tail), 0.5 µL M13 primer (with fluorescent dye label), 0.2 µL Taq DNA polymerase (Qiagen), 2 µL DNA sample, and 4.2 µL ultrapure (Milli-Q) water.

For cpDNA microsatellite PCRs, we used an Eppendorf Mastercycler Nexus thermal cycler with reaction conditions specified in Nevill et al. (2010). Afterwards, we combined amplified products and diluted them with Milli-Q water using an 8x dilution factor, and then separated fragments on an ABI 3130xl Genetic Analyser (Applied Biosystems). A number of samples were also used as technical replicates (i.e. were run multiple times on the sequencer on different plates) to ensure peak calls weren’t varying between plates.

4.3.7 Microsatellite data analysis

We called microsatellite alleles using Geneious v6.1.8 (Kearse et al., 2012), and checked technical replicates to confirm the absence of variation in peak calls between plates. Any samples with uncertain peaks or missing data were removed from the analysis, resulting in 302 samples being successfully genotyped at all five microsatellite loci. A CSV file was exported from Geneious and population genetic analysis was conducted in R.

The genotypes in microsatellite allele size format were first concatenated to create haplotypes and the dataset converted into a genind object using the df2genind function with a ploidy of one, and a global $F_{ST}$ value calculated. Minimum spanning networks were used to visualise the relationships between haplotypes (Supplementary Figures S4.1 and S4.2), via the imsn function of the ‘poppr’ (Kamvar et al., 2014) package, with pairwise distances calculated using Bruvo’s genetic distance (Bruvo et al., 2004) and the number of dissimilar alleles between haplotypes. A Euclidean distance matrix was then produced using the dist function, and the matrix converted to binary format. As for the SNP dataset, an AMOVA was performed across all samples using the adonis function with treatment as a stratum and 999
permutations, to determine whether there were significant differences in genetic variance between treatment types. The genind was then subset by treatment to create a separate object for each treatment. To determine the proportion of genetic variance among versus within sites, an AMOVA was conducted for each treatment using the *adonis* function, with site as the stratum and 999 permutations. To investigate the partitioning of genetic diversity across the treatments, we used the *QDiver* function of GenAlEx, as for the nuclear SNP dataset. We also analysed fine-scale genetic structure between all samples, via the *Spatial* function in GenAlEx, to allow for visual inspection of differences between the chloroplast and nuclear datasets.

The genind objects for each treatment were transformed using the *genind2hierfstat* function of the ‘hierfstat’ (Goudet, 2005) package. Genetic differentiation, in the form of Nei’s $F_{ST}$, was then calculated independently for each treatment using the *basic.stats* function. Due to the uneven collection of samples between sites and treatments, we standardised our haplotype diversity using the allelic (haplotype) richness (HR) technique (El Mousadik and Petit, 1996). We calculated HR for each site using the *allel.rich* function in the ‘PopGenReport’ (Adamack and Gruber, 2014) package, and then obtained the mean and standard deviation of HR within each treatment for comparison. Haplotypes were also mapped across the landscape by treatment to allow for visual inspection of landscape-scale genetic structure.
4.4 Results

4.4.1 Hybridisation

The mean level of admixture per individual for undisturbed, burnt, and logged treatments were 0.022, 0.014, and 0.009 respectively. Single-factor ANOVA showed that the level of admixture differed in a marginally significant manner between treatments ($F(2,349) = 3.041$, $P = 0.049$). Post-hoc pairwise comparisons using Tukey’s test identified the undisturbed treatment as being significantly different to the logged treatment. After removing a single highly admixed individual from the undisturbed treatment, the significant effect of treatment on admixture disappeared ($F(2,348) = 2.702$, $P = 0.069$).

The global model had a very low pseudo-$R^2$ estimate (0.017), suggesting that treatment did not explain a substantial portion of the variance. The AICc of the intercept-only model (1008.03) was lower than the AICc of the global model (1010.85), suggesting that treatment did not substantially influence the level of admixture.

4.4.2 Genotyping-by-sequencing

Within-population genetic structure for all sites was similar, with just minor variations in $N_A$, $N_E$, MAF, $H_O$, $H_E$, $P_F$, and $F_{IS}$ (Supplementary Table S4.1). Observed heterozygosity was equal to or higher than $H_E$ at all sites, with negative $F_{IS}$ values suggesting that outbreeding is the standard condition. The mean proportion of polymorphic loci within sites was significantly lower in the logged treatment ($0.43\pm0.01$) than in the undisturbed ($0.48\pm0.02$) or burnt ($0.49\pm0.02$) treatments (GLM, $F=3.968$, df=2, $P=0.03$), suggesting that logged sites have experienced a small loss of alleles (Table 4.1). The level of genetic differentiation within treatments was similar but significantly different (GLM, $F=6.797$, df=2, $P=0.001$), being highest in the logged treatment ($0.015\pm0.001$) followed by the undisturbed ($0.013\pm0.001$) and burnt ($0.011\pm0.001$) treatments. This is despite the undisturbed sites having the greatest geographic spread, and logged sites having the lowest geographic spread.

Among the three disturbance treatments, the AMOVA, conducted on a matrix of pairwise genetic distances, identified small but significant differences in genetic variance (adonis $F = 1.678$, $R^2=0.01$, $P=0.001$) (Figure 4.2). Within each treatment, AMOVA similarly revealed a very small but significant level of genetic differentiation among sites (adonis, $F=1.34$ to 1.44, $R^2 = 0.09$ to 0.14, $P=0.001$), with the among-site variance accounting for between 1.4 and 1.9% of the total genetic variance.

The QDiver function of GenAlEx computed standarised genetic diversity metrics partitioned into hierarchical strata (Smouse et al., 2017). Partitioning total diversity between
treatments showed that all three treatments had similar levels of overall diversity ($\sigma = 0.135$ to 0.137) (Table 4.2). As for the AMOVA (variance-based) results, most genetic diversity within treatments was held at the level of the site ($\alpha = 0.131$ to 0.134), with just small amounts of diversity found among sites ($\beta = 0.004$ to 0.007). The Bartlett’s test of within-site homogeneity showed that genetic diversity within sites was not significantly different within any of the treatments ($BC=0.2704$ to 0.433, $P=0.395$ to 0.945). Similarly, the level of diversity within sites was not significantly different between treatments ($BC=1.99$, $P=0.391$). As with the variance-based results, the logged sites had the highest among-site diversity, however the difference was not significant ($P=0.798$).

The fine-scale spatial structure analysis revealed that the logged sites exhibited very similar correlograms to the undisturbed and the burnt sites (Figure 4.3), despite being artificially regenerated. All three treatments showed the pattern of short distance (hundreds of metres) significant positive spatial genetic structure at the first distance class of 0.2 km, indicating a non-random distribution of related genotypes at the local site level, whereas at all larger distance classes autocorrelation values were not significantly different from zero. However, differences in the magnitude of the autocorrelation coefficient $r$ at the first distance class (i.e. within sites) are of interest. Inspection of the bootstrap-estimated 95% confidence intervals (CI) about $r$ revealed that the logged treatment (CI = 0.031 to 0.046) showed a significantly higher $r$ value than the burnt treatment (CI = 0.017 to 0.029). Heterogeneity testing in GenAlEx across the three treatment types at all distance classes showed that all treatments had similar levels of $r$ ($P > 0.1$ in most cases), with the exception of the first distance class ($P = 0.001$). The estimated $r$ in the logged treatment was also higher than that found in the undisturbed treatment (CI = 0.023 to 0.043), but with overlapping CIs. Together, these results indicate that the artificially reseeded sites typically contain more related individuals than observed at the other two types of sites.
Table 4.1 Means and standard errors (SE) of genetic parameters, including number of alleles ($N_A$), number of effective alleles ($N_E$), minor allele frequency (MAF), expected heterozygosity ($H_E$), observed heterozygosity ($H_O$), proportion of polymorphic loci ($P_P$), genetic differentiation ($F_{ST}$) and inbreeding coefficient ($F_{IS}$), for stands of mountain ash (*Eucalyptus regnans*) subjected to three different disturbance treatments.

<table>
<thead>
<tr>
<th></th>
<th>$N_A$</th>
<th>$N_E$</th>
<th>MAF</th>
<th>$H_E$</th>
<th>$H_O$</th>
<th>$P_P$</th>
<th>$F_{ST}$</th>
<th>$F_{IS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisturbed</td>
<td>Mean</td>
<td>1.48</td>
<td>1.22</td>
<td>0.097</td>
<td>0.13</td>
<td>0.18</td>
<td>0.48</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.02</td>
<td>0.00</td>
<td>0.001</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Burnt</td>
<td>Mean</td>
<td>1.49</td>
<td>1.21</td>
<td>0.095</td>
<td>0.13</td>
<td>0.17</td>
<td>0.49</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.02</td>
<td>0.00</td>
<td>0.001</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Logged</td>
<td>Mean</td>
<td>1.43</td>
<td>1.22</td>
<td>0.098</td>
<td>0.13</td>
<td>0.18</td>
<td>0.43</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.01</td>
<td>0.00</td>
<td>0.001</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Figure 4.2 Summary of analysis of molecular variance results, showing the percentages of genetic variance found within (light grey) and among (dark grey) stands of mountain ash (*Eucalyptus regnans*). The top row of charts shows the results for five microsatellites located in the chloroplast DNA, and the bottom row shows the results for 1240 single-nucleotide polymorphisms (SNP) from across the genome. Plots marked with (A), (B) and (C) represent undisturbed, burnt, and logged site treatments respectively. A significant ($P=0.001$) proportion of the genetic variance in both the nuclear and chloroplast genomes was found among sites in all treatments, based on 999 random permutations of sample matrices.
Table 4.2 Diversity metrics for mountain ash (*Eucalyptus regnans*), where sites have been separated into three forest categories each representing different disturbance histories. Diversity metrics include grand total (GT), within treatment (WR), among sites within treatments (AP/WR), and within sites within treatments (WP/WR). Values were calculated in GenAlEx with the QDiver function (including interpolation) using 1240 single-nucleotide polymorphisms (SNPs). *P*-values of Bartlett's tests for differences among treatments are also provided.

<table>
<thead>
<tr>
<th>Diversity parameter</th>
<th>Undisturbed</th>
<th>Burnt</th>
<th>Logged</th>
<th><em>P</em>-value (Bartlett’s test)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nuclear SNPs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = Q(GT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ = Q(WR)</td>
<td>0.137</td>
<td>0.135</td>
<td>0.137</td>
<td>0.925</td>
</tr>
<tr>
<td>β = Q(AP/WR)</td>
<td>0.004</td>
<td>0.005</td>
<td>0.007</td>
<td>0.798</td>
</tr>
<tr>
<td>α = Q(WP/WR)</td>
<td>0.134</td>
<td>0.131</td>
<td>0.131</td>
<td>0.391</td>
</tr>
<tr>
<td><strong>Chloroplast haplotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = Q(GT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ = Q(WR)</td>
<td>0.443</td>
<td>0.699</td>
<td>0.789</td>
<td>0.001</td>
</tr>
<tr>
<td>β = Q(AP/WR)</td>
<td>0.395</td>
<td>0.660</td>
<td>0.712</td>
<td>0.001</td>
</tr>
<tr>
<td>α = Q(WP/WR)</td>
<td>0.079</td>
<td>0.115</td>
<td>0.266</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Figure 4.3 Spatial autocorrelation of mountain ash (*Eucalyptus regnans*) genotypes, identified using 1240 single-nucleotide polymorphisms, in the Central Highlands of Victoria, Australia. Three forest categories, representing differing disturbance histories, were investigated. Error bars indicate 95% confidence intervals.
4.4.3 Chloroplast microsatellites

Across all 285 individuals and 28 sites, we observed a total of 15 haplotypes. Six of these haplotypes were found in more than 10 individuals. Four haplotypes were private to single sites, and two of those were private to the same site. All sites with private haplotypes were logged sites.

Two of the four private haplotypes were found in more than one individual. Undisturbed, burnt, and logged treatments had five, seven and 10 haplotypes respectively. Using rarefaction to account for uneven sample sizes resulted in logged sites having the highest mean (± standard error) HR (1.85 ± 0.25), with the undisturbed and burnt sites having nearly equal HR (1.36 ± 0.17 and 1.34 ± 0.12 respectively) (Figure 4.4).

Within each of the three treatments, the AMOVA analysis revealed extensive genetic differentiation among sites, with the among-site variance accounting for between 66.7 and 83.7% of the total genetic variance (adonis, $F = 15.9$ to 78.7, $R^2 = 0.67$ to 0.84, $P = 0.001$) (Figure 4.2). Furthermore, the variation in the magnitude of genetic differentiation was significantly different among the three treatments (adonis, $F = 81.4$, $R^2 = 0.13$, $P = 0.001$). The logged treatment had the highest level of within-site genetic variance (33.3%), suggesting the seed used to regenerate the sites is typically sourced from multiple localities.

The QDiver analysis revealed a significant difference ($BC = 11.105$, $P = 0.743$) in the level of total diversity between the three treatments, with the logged treatment ($\sigma = 0.789$) containing more diversity than the undisturbed ($\sigma = 0.443$) and burnt ($\sigma = 0.699$) treatments (Table 4.2). As in the AMOVA (variance-based) results, a large proportion of the genetic diversity within treatments was found among sites ($\beta = 0.395$-0.712). The Bartlett’s test of within-site homogeneity showed than genetic diversity within sites was significantly different between treatments ($BC = 37.044$, $P = 0.001$), with the logged sites having higher within-site diversity ($\alpha = 0.266$) than the burnt ($\alpha = 0.115$) or undisturbed ($\alpha = 0.079$) treatments.

The fine-scale individual-by-individual spatial autocorrelation analysis of the chloroplast haplotypes revealed a very different pattern to that found at the nuclear SNP loci (Figure 4.5). Unlike the SNPs, for the chloroplast haplotypes, significant and large positive $r$ values (0.25-0.6, $P < 0.001$) were found at the distance classes 0.2, 1 and 5 km.

Mapping haplotype frequencies within sites across the landscape showed clear spatial patterns of haplotype distribution for the undisturbed and the burnt treatments (Figure 4.6). Geographic structuring of haplotypes was weakest in the logged treatment, where there was no clear pattern of haplotype distribution among sites. Mapping the spatial distribution of haplotypes among all unlogged sites did not reveal any strong or obvious patterns with respect to landscape topography, suggesting a complex history of recolonisation from refuge habitats,
although large areas of suboptimal low- or high-elevation habitat may have played a role in historical dispersal patterns (Figure 4.7).
Figure 4.4 Mean haplotype richness per site for three categories of mountain ash (Eucalyptus regnans) forest with different disturbance histories. Haplotypes were derived from five chloroplast microsatellites. Logged sites were clear-cut and sown with mountain ash seed. Burnt sites were subjected to wildfire in 2009. Undisturbed sites had regenerated from wildfires prior to 1960. Error bars represent the standard error of the mean values across sites within each treatment.
Figure 4.5 Comparison of the decay of spatial autocorrelation of mountain ash (*Eucalyptus regnans*) genotypes from two datasets: one made of haplotypes derived from five chloroplast microsatellites and another made of 1240 nuclear single-nucleotide polymorphisms. Data covers 31 sites in the Central Highlands of Victoria, Australia. Error bars indicate 95% confidence intervals.
Figure 4.6 Spatial distribution of chloroplast microsatellite haplotypes identified in mountain ash (*Eucalyptus regnans*) across the Central Highlands of Victoria. Each panel shows the locations of sampled sites with a particular disturbance history, representing undisturbed (A), burnt (B), and logged (C) scenarios.
Figure 4.7 Spatial distribution of chloroplast microsatellite haplotypes identified in mountain ash (*Eucalyptus regnans*) across the Central Highlands of Victoria, where stands of trees have not been subjected to timber harvesting. Landscape topography is indicated by elevation (m).
4.5 Discussion

We investigated patterns of genetic diversity and structure in the chloroplast and nuclear genomes of a foundation tree species, comparing sites subjected to different disturbance histories. We found strong differences in patterns of landscape genetic structure between chloroplast haplotypes and the nuclear genome, with the chloroplast exhibiting much more genetic differentiation among sites. Our results suggest that artificial regeneration of logged sites is having small effects on the nuclear genome, including a reduction in the proportion of polymorphic loci, increased genetic differentiation between sites, and an increase in the average relatedness between individuals within sites (indicated by lower \(F_{IS}\) values and higher spatial autocorrelation of multilocus genotypes). In addition, the natural pattern of landscape genetic structure of the chloroplast haplotypes is being substantially modified by current management practices, with alleles probably entering the system through artificial contribution of non-local seed to the regeneration process. The logged treatment was also found to have lower levels of genetic differentiation among sites for chloroplast markers, as well as higher haplotype richness, compared to the undisturbed or burnt treatments.

4.5.1 Genetic differentiation, gene flow and dispersal

The processes of gene flow, genetic drift and local adaptation drive patterns of spatial genetic structure and genetic differentiation among populations (Orsini et al., 2013). Many widespread forest trees have been found to exhibit a pattern of continuously-distributed genetic variation with very low levels of nuclear genetic differentiation among populations (Supple et al., 2018; Eckert et al., 2010; Sampson et al., 2018; Sork et al., 2010). Whilst this appears to be the case for many open-pollinated eucalypts with large geographic ranges, a number of studies have identified patterns of strong among-population genetic differentiation at chloroplast loci (Bloomfield et al., 2011; Nevill et al., 2014). This suggests that dispersal of nuclear alleles, which are biparentally-inherited and dispersed via pollen and seeds, tends to occur on a greater scale than dispersal of chloroplast haplotypes, which are maternally-inherited and only dispersed via seeds.

Our results for mountain ash were consistent with these previous findings, showing strong significant positive autocorrelation of chloroplast haplotypes over distances of 5 to 10 km, implying a low level of haplotype mixing across the landscape, and only very weak short-distance (hundreds of metres) autocorrelation between nuclear genotypes, indicating more random mating patterns. The strongly contrasting results between chloroplast and nuclear
markers in the QDiver analysis, along with the $F_{ST}$ values we observed, suggest that high levels of pollen gene flow minimise nuclear genetic differentiation in mountain ash across the landscape. This ensures that even small stands of trees contain most of the genetic diversity found in the landscape, and helps to explain the differences in genetic structure we observed between treatments.

### 4.5.2 Logging and genetic structure

Globally, the most simple and widespread system of silvicultural management in forests is clear cutting followed by artificial regeneration of target tree species (Ratnam et al., 2014). The effects of this type of disturbance on forest genetic resources have been studied in a number of tree species (Fageria and Rajora, 2014; Rajora and Pluhar, 2003; Glaubitz et al., 2003b; Glaubitz et al., 2003a), with mixed results. Whilst we found minor differences in a number of genetic diversity parameters and within-site spatial genetic structure, Rajora and Pluhar (2003) found that clear cutting and artificial regeneration of black spruce ($Picea mariana$) had no substantial impacts on measures of genetic diversity, with diversity maintained across the logged landscape. Similarly, Glaubitz et al. (2003a) and Glaubitz et al. (2003b) showed that clear cutting of yertchuk ($Eucalyptus consideniana$) and silvertop ash ($Eucalyptus sieberi$) had no effect on multiple genetic parameters for either species. In a more recent study, Fageria and Rajora (2014) found that plantations of white spruce ($Picea glauca$) have reduced allelic diversity and heterozygosity compared to old-growth and naturally regenerating stands.

While the influence of logging on the nuclear genome of mountain ash was relatively small, the effect on chloroplast haplotypes was very strong. All private chloroplast haplotypes were found in logged sites, which suggests that seed used to regenerate these sites after logging was sourced from outside the geographic extent of the study (the Central Highlands). This is of interest because seed sourcing practices in this region, as across most of southern Australia, are supposed to rely on local provenancing. Seed used in regeneration of logged sites should be collected either at the site to be sown or within a short distance (typically < 25 km) of that site (Fagg, 2001; VicForests, 2018; Forestry Tasmania, 2010; DEPI, 2014), in an attempt to prevent disrupting the extent of local adaptation to environmental conditions. Our results suggest that exceptions to this strategy have occurred over the past few decades. Such exceptions could be due to many reasons, but often may occur as a result of local seed shortages in some years.

The high level of standing genetic variation in the nuclear genome of many widespread forest trees means that artificial seeding of logged sites is likely to cause only minor alterations in the patterns of fine-scale genetic structure and nuclear genetic diversity of
stands of trees, unless seed is sourced from very distant locations (e.g. hundreds to thousands of kilometres) or from a very limited number of parents. Recent work in other eucalypt species has shown that, for species in which high dispersal and low range-wide population genetic differentiation has been demonstrated, the use of locally-sourced seeds for restoration plantings may be unnecessarily restrictive (Dillon et al., 2015; Supple et al., 2018; Breed et al., 2013). Our data for mountain ash is largely in agreement with these results, in that spatial genetic structure at the nuclear genome is very weak.

Despite the lack of a large effect of provenancing strategy on nuclear genetic structure, we detected a number of small differences between logged and naturally regenerated sites, including reductions in the proportion of polymorphic loci and the inbreeding coefficient, and an increase in genetic differentiation between sites. This suggests that the seed mixture used for artificial regeneration probably comes from fewer maternal parents than is necessary to maintain natural patterns of stand-level genetic diversity. This could potentially result in reduced adaptive potential through a loss of alleles, and can have flow-on implications for diversity in other components of the forest ecosystem (Hughes et al., 2008). These impacts could easily be averted by adopting a regional admixture provenancing approach, whereby seeds from multiple regional populations are mixed thoroughly to maximise genetic diversity while retaining regionally-adapted genotypes (Bucharova et al., 2019). Alternatively, increasing the amount of natural regeneration on logged sites could also mitigate loss of potentially adaptive alleles. One way to achieve this would be to shift from clear cutting to variable retention harvesting, which has the additional benefits of retaining important structural and functional attributes within sites (Lindenmayer et al., 2012b).

The invasion of alleles due to artificial movement/mixing of local haplotypes across the landscape during reseeding operations is restructuring the natural chloroplast population genetic structure in areas subjected to timber harvesting. However, because there is very limited natural population genetic structuring in the nuclear DNA, even over large distances, it appears as though this disruption is limited to the maternally-inherited components of the DNA. In addition, our chloroplast haplotype data of sites burnt in 2009 did not contain any strong evidence to suggest that seeds are dispersing from logged sites into unlogged sites, which is likely due to the small dispersal distances of eucalypt seeds and the resulting slow rate of migration (Gill, 1994; Lamb and Smyth, 2006; Booth, 2017). Conducting a more extensive study over a number of disturbance events would help determine whether such a phenomenon is occurring.
4.5.3 Phylogeographic implications

The natural genetic diversity of mountain ash chloroplast haplotypes in the Central Highlands of Victoria was previously thought to be extremely low, with one study (which used the same microsatellite loci but fewer samples) finding just four haplotypes in approximately the same geographic extent sampled for our investigation (Nevill et al., 2010). Nevill et al. (2010) used this result to suggest that mountain ash, which may have been absent from the region during the last glacial maximum (c. 23,000 to 16,000 years ago), had recolonised the Central Highlands from the warmer coastal regions where it was able to persist. However, in sites that had not experienced logging, we found considerably more haplotype diversity both within and among sites than expected, with 11 haplotypes present. Further, the levels of within- and among-site genetic variance for chloroplast haplotypes that we observed are comparable to other regions containing mountain ash (Nevill et al., 2010), as well as those found in a study of messmate (Bloomfield et al., 2011), a broadly sympatric species throughout much of the range of mountain ash.

Our results have implications for the phylogeographic history of mountain ash, with one of two historical recolonisation scenarios following the last glacial maximum possible; either (1) there were multiple recolonization events from lower-lying coastal regions, or (2) local refuges persisted in the landscape to allow later recolonization. As previous research has suggested the local existence of moist, sheltered refuges for eucalypts even at moderate altitudes (>1000 m) within the broader region (Kershaw et al., 2007), we suggest it is likely that mountain ash was able to persist in the region in refuge habitats at lower altitudes than where it is currently found.

4.5.4 Additional impacts of logging

While we found minor differences in fine-scale genetic structure and patterns of genetic diversity between treatments, is seems likely that current silvicultural practices in mountain ash forests will not have large impacts on the population genetic structure or adaptive potential of the species. However, there is a range of structural and ecological repercussions that can also result from timber harvesting. While some authors have previously found that disturbance by clearcutting is broadly equivalent to disturbance by wildfire (Attiwill, 1994; Baker et al., 2004), others have more recently noted differences in species richness and abundance of functional groups or specific taxa (Blair et al., 2016; Lindenmayer et al., 2018; Bowd et al., 2018). Additionally, the combined effects of logging and fire in mountain ash forests is reducing the average age class of the landscape, creating a feedback loop of disturbance that is likely to result in the collapse of multiple ecosystem functions.
(Lindenmayer and Sato, 2018). Thus, to conserve forest ecosystems, we need to consider the full array of biotic and abiotic impacts that anthropogenic disturbances are having on the system. This not only requires an understanding of how the distribution of genetic diversity is being impacted, but also how structural components are changing across the landscape and through time, and how interactions between foundation species and the species that depend on them are being altered.
4.6 Conclusions

Our study of nuclear and chloroplast genetic markers in mountain ash has provided insights into the similarities and differences between the effects of wildfire and logging in the forests of south-eastern Australia. We found that current silvicultural practices are having minor impacts on the fine-scale genetic structure and genetic diversity of mountain ash stands, and are not likely to substantially impact the nuclear genetic resources of the species. We suggest that a regional admixture provenancing strategy would mitigate the minor impacts of artificial regeneration on the genetic structure of logged sites, as would adopting the practice of variable retention harvesting. We also discovered large differences in genetic parameters between the nuclear and chloroplast genomes, with chloroplast haplotypes exhibiting highly significant differences in genetic differentiation and within-site genetic diversity between logged and unlogged treatments. Thus, the current methods of artificial reseeding in logged sites is altering the natural distribution of chloroplast haplotypes across the landscape. Our detailed chloroplast haplotype data also allowed us to identify new possibilities for the phylogeographic history of forest trees in the Central Highlands of Victoria, a topic that will require further investigation before it can be fully resolved.
4.7 Acknowledgements

This study was supported by an Australian Research Council Future Fellowship (FT130100043) to SB, a Lesslie Research Scholarship in Landscape Conservation and Ecology to BvTD, and an Ignition Grant from the Centre for Biodiversity Analysis to BvTD. Thanks to Oliver King, Joe Salmona and Donna Belder for assistance with fieldwork.
4.8 References


Barton, K., 2016. MuMIn: multi-model inference, R package version 1.15.6. Available at: http://cran.r-project.org/package=MuMIn


Chapter 4: Wildfires, logging and genetic structure


Fagg, P., 2001. *Eucalypt sowing and seedfall, native forest silviculture guideline no. 8*, Department of Natural Resources and Environment, Melbourne, Australia.


Chapter 4: Wildfires, logging and genetic structure


Chapter 4: Wildfires, logging and genetic structure


Chapter 4: Wildfires, logging and genetic structure


https://www.nature.com/articles/nclimate3303#supplementary-information


### 4.9 Supplementary material

Table S4.1 Standard genetic parameters for 31 populations of mountain ash (*Eucalyptus regnans*) in the Central Highlands of Victoria, Australia. Parameters were calculated using 1240 single-nucleotide polymorphisms from across the genome. Mean values (with standard errors in parentheses) are given for the number of alleles ($N_A$), number of effective alleles ($N_E$), minor allele frequency (MAF), expected heterozygosity ($H_E$), observed heterozygosity ($H_O$), and the inbreeding coefficient ($F_{IS}$). The proportion of polymorphic loci ($P_P$) at each site is also provided.

<table>
<thead>
<tr>
<th>Site</th>
<th>Treatment</th>
<th>$N_A$ (0.01)</th>
<th>$N_E$ (0.01)</th>
<th>MAF (0.00)</th>
<th>$H_O$ (0.01)</th>
<th>$H_E$ (0.01)</th>
<th>$F_{IS}$ (0.01)</th>
<th>$P_P$ (0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PowePt 56</td>
<td>Logged</td>
<td>1.41</td>
<td>1.21</td>
<td>0.096</td>
<td>0.18</td>
<td>0.12</td>
<td>-0.29</td>
<td>0.41</td>
</tr>
<tr>
<td>PowePt 58</td>
<td>Logged</td>
<td>1.42</td>
<td>1.22</td>
<td>0.099</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.28</td>
<td>0.42</td>
</tr>
<tr>
<td>PowePt 62</td>
<td>Logged</td>
<td>1.38</td>
<td>1.21</td>
<td>0.094</td>
<td>0.17</td>
<td>0.12</td>
<td>-0.29</td>
<td>0.38</td>
</tr>
<tr>
<td>PowePt 64</td>
<td>Logged</td>
<td>1.43</td>
<td>1.22</td>
<td>0.101</td>
<td>0.19</td>
<td>0.13</td>
<td>-0.3</td>
<td>0.43</td>
</tr>
<tr>
<td>PowePt 65</td>
<td>Logged</td>
<td>1.47</td>
<td>1.23</td>
<td>0.102</td>
<td>0.19</td>
<td>0.14</td>
<td>-0.25</td>
<td>0.47</td>
</tr>
<tr>
<td>PowePt 67</td>
<td>Logged</td>
<td>1.45</td>
<td>1.23</td>
<td>0.103</td>
<td>0.19</td>
<td>0.14</td>
<td>-0.28</td>
<td>0.45</td>
</tr>
<tr>
<td>PowePt 80</td>
<td>Logged</td>
<td>1.41</td>
<td>1.21</td>
<td>0.095</td>
<td>0.17</td>
<td>0.13</td>
<td>-0.28</td>
<td>0.41</td>
</tr>
<tr>
<td>ToolPt 18</td>
<td>Logged</td>
<td>1.46</td>
<td>1.22</td>
<td>0.100</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.26</td>
<td>0.46</td>
</tr>
<tr>
<td>ToolPt 20</td>
<td>Logged</td>
<td>1.40</td>
<td>1.21</td>
<td>0.095</td>
<td>0.18</td>
<td>0.12</td>
<td>-0.3</td>
<td>0.40</td>
</tr>
<tr>
<td>ToolPt 22</td>
<td>Logged</td>
<td>1.41</td>
<td>1.22</td>
<td>0.098</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.28</td>
<td>0.41</td>
</tr>
<tr>
<td>ToolPt 24</td>
<td>Logged</td>
<td>1.41</td>
<td>1.21</td>
<td>0.093</td>
<td>0.17</td>
<td>0.13</td>
<td>-0.27</td>
<td>0.41</td>
</tr>
<tr>
<td>ToolPt 27</td>
<td>Logged</td>
<td>1.47</td>
<td>1.22</td>
<td>0.099</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.25</td>
<td>0.47</td>
</tr>
<tr>
<td>ToolPt 32</td>
<td>Logged</td>
<td>1.47</td>
<td>1.23</td>
<td>0.102</td>
<td>0.19</td>
<td>0.13</td>
<td>-0.25</td>
<td>0.47</td>
</tr>
<tr>
<td>ToolPt 78</td>
<td>Logged</td>
<td>1.40</td>
<td>1.22</td>
<td>0.101</td>
<td>0.19</td>
<td>0.13</td>
<td>-0.33</td>
<td>0.40</td>
</tr>
<tr>
<td>Dandenong</td>
<td>Undisturbed</td>
<td>1.56</td>
<td>1.22</td>
<td>0.101</td>
<td>0.18</td>
<td>0.14</td>
<td>-0.19</td>
<td>0.56</td>
</tr>
<tr>
<td>MarooNorth</td>
<td>Undisturbed</td>
<td>1.49</td>
<td>1.22</td>
<td>0.100</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.23</td>
<td>0.49</td>
</tr>
</tbody>
</table>

(continues)
<table>
<thead>
<tr>
<th>Location</th>
<th>Forest Type</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
<th>Value 5</th>
<th>Value 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>MarooSouth</td>
<td>Undisturbed</td>
<td>1.62</td>
<td>0.100</td>
<td>0.14</td>
<td>-0.15</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>OShan</td>
<td>Undisturbed</td>
<td>1.51</td>
<td>0.100</td>
<td>0.13</td>
<td>-0.22</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Powl110</td>
<td>Undisturbed</td>
<td>1.46</td>
<td>0.096</td>
<td>0.13</td>
<td>-0.23</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>PowlStarOld</td>
<td>Undisturbed</td>
<td>1.43</td>
<td>0.098</td>
<td>0.13</td>
<td>-0.28</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Tool15</td>
<td>Undisturbed</td>
<td>1.4</td>
<td>0.088</td>
<td>0.12</td>
<td>-0.22</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Tool277</td>
<td>Undisturbed</td>
<td>1.41</td>
<td>0.093</td>
<td>0.12</td>
<td>-0.24</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>UpperYarra</td>
<td>Undisturbed</td>
<td>1.46</td>
<td>0.098</td>
<td>0.13</td>
<td>-0.23</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Camb2</td>
<td>Burnt</td>
<td>1.55</td>
<td>0.093</td>
<td>0.13</td>
<td>-0.14</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>CathRange</td>
<td>Burnt</td>
<td>1.54</td>
<td>0.095</td>
<td>0.13</td>
<td>-0.17</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Mary911</td>
<td>Burnt</td>
<td>1.56</td>
<td>0.093</td>
<td>0.13</td>
<td>-0.16</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Mary912</td>
<td>Burnt</td>
<td>1.52</td>
<td>0.098</td>
<td>0.13</td>
<td>-0.18</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Powe11</td>
<td>Burnt</td>
<td>1.44</td>
<td>0.094</td>
<td>0.12</td>
<td>-0.20</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>PowlStarYoung</td>
<td>Burnt</td>
<td>1.39</td>
<td>0.095</td>
<td>0.13</td>
<td>-0.26</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Tool340</td>
<td>Burnt</td>
<td>1.49</td>
<td>0.094</td>
<td>0.13</td>
<td>-0.19</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Tool914</td>
<td>Burnt</td>
<td>1.40</td>
<td>0.096</td>
<td>0.13</td>
<td>-0.28</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>
Figure S4.1 Minimum spanning network of mountain ash (*Eucalyptus regnans*) haplotypes. Each node represents a haplotype, where node size is scaled by the number of individuals with that haplotype. Individuals have been separated into one of three disturbance treatments based on site history. Haplotypes were identified using five chloroplast microsatellites and distances between nodes were calculated using Bruvo’s genetic distance.
Figure S4.2 Minimum spanning network of mountain ash (*Eucalyptus regnans*) haplotypes. Each node represents a haplotype, where node size is scaled by the number of individuals with that haplotype. Individuals have been separated into one of three disturbance treatments based on site history. Haplotypes were identified using five chloroplast microsatellites and distances between nodes were calculated using counts of dissimilar alleles.
Chapter 5: Scale-dependent signatures of local adaptation in a widespread foundation tree species

Citation: von Takach Dukai, B., Lindenmayer, D. B., and Banks, S. C., in review. Scale-dependent signatures of local adaptation in a widespread foundation tree species. *Molecular Ecology*.

Keywords: Mountain ash; local adaptation; microgeographic adaptation; *Eucalyptus regnans*; SNP; gene flow; genome scan; outlier loci; genotype-environment
5.1 Abstract

Understanding the patterns and drivers of local adaptation over different spatial scales is important for developing effective management strategies under rapidly changing environmental conditions. Here, we compared signatures of local adaptation across the genome of mountain ash (*Eucalyptus regnans*) at two spatial scales: (1) across the range of the species, in relation to coarse-scale climatic and topographic variation, and (2) within an intensively-sampled sub-region exhibiting fine-scale climatic and edaphic variation.

We used genotyping-by-sequencing to genotype 292 individual trees, retaining about 2000 single-nucleotide polymorphisms at each spatial scale. We quantified patterns of spatial genetic structure among populations across the range of the species, then utilised three methods to identify loci under selection.

Genetic differentiation across the range of the species was low, with similar patterns observed for genic and non-genic loci. Genome scans identified no loci showing significant departure from background levels of differentiation at either spatial scale. Latent factor mixed models and redundancy analysis together identified 40 loci that were significantly correlated with environmental predictors, of which 30 and 10 were unique to the range-wide and geographically-restricted datasets respectively. However, two loci identified in the range-wide dataset were in linkage disequilibrium with nearby loci in the geographically-restricted dataset.

Our results suggest that local adaptation in mountain ash mostly occurs through small-effect loci acting on polygenic traits, and that adaptive genetic components can differ substantially depending on the spatial scale of environmental gradients. This highlights the importance of studying multiple spatial scales for understanding patterns of adaptive genetic variation across landscapes.
5.2 Introduction

Globally, anthropogenic modifications to landscapes and climates are resulting in changing frequencies and intensities of disturbance events such as wildfires, pest outbreaks, and extreme weather events (Abatzoglou and Williams, 2016; Seidl et al., 2017). In the context of these rapid contemporary environmental changes, understanding the patterns and drivers of local adaptation is important for developing practical and effective management strategies (Sinclair et al., 2015; Hoffmann and Sgrò, 2011). For example, if there is evidence that patterns of local adaptation vary across different spatial scales, this (1) could have implications for our understanding of the risk to populations or species under emerging new environmental conditions, and (2) assist management actions designed to ensure that such populations or species have the adaptive potential to persist (Holliday et al., 2017; Alfaro et al., 2014).

Advancements in sequencing technologies and methods of data analysis over the past decade have led to increasingly detailed insights into the way that genetic diversity varies across landscapes (Manel and Holderegger, 2013; Hendricks et al., 2018). Researchers can now utilise thousands to hundreds of thousands of single-nucleotide polymorphisms (SNPs) from across the genome to investigate patterns of demography and evolutionary history in unprecedented detail (Bragg et al., 2015). Such datasets can be used to identify specific loci at which variation is being driven by local adaptation, through the exploration of associations between phenotypes, genotypes and environmental variables (Evans et al., 2014; Ahrens et al., 2018; Rellstab et al., 2015). By understanding the rate of change in allele frequencies as a result of selection, we can even begin to place into context the ways in which particular genetic pathways are likely to contribute to rapidly changing environments (Jordan et al., 2017; Rellstab et al., 2016).

Much of the current landscape genomics literature focuses on large-scale latitudinal or longitudinal gradients in environmental variables, as simple climate change predictions focus on shifts in bioclimatic conditions on this scale (Dillon et al., 2014; Bradbury et al., 2010; Hecht et al., 2015; Jordan et al., 2017). However, bioclimatic conditions can also be influenced by topographic and edaphic heterogeneity at fine spatial scales (Lind et al., 2017; Wang et al., 2010), such that selective processes can be driven on a large scale by latitude or longitude and on a small scale by microgeographic variation in topographic or edaphic factors. Previous work has shown that local adaptation can influence genetic composition on microgeographic scales (i.e. where dispersal neighbourhoods extend across environmental gradients) when the homogenising effect of gene flow is exceeded by divergent selection (Banks et al., 2017; Richardson et al., 2014; Eckert et al., 2015; Skelly, 2004). Thus, even
when dispersal is high and gene flow across environmental gradients is strong, selection can potentially still lead to local adaptation by filtering of individuals within each generation (Vizcaíno-Palomar et al., 2014).

There is a clear gap in the research literature around the detection of adaptive genetic variation across different spatial scales, although there is evidence that similar adaptive phenotypes (convergent traits) can develop as a result of multiple different genetic/biochemical pathways both within and between species (Pascoal et al., 2014; Martin and Orgogozo, 2013; Yeaman et al., 2016), suggesting that the genetic architecture of local adaptation could vary according to the scale at which selective processes operate. In Australia, research on eucalypts has found that traits showing local adaptation to aridity across three species are probably derived largely from independent genetic components (Steane et al., 2017b). Together, these studies suggest that the scale at which sampling is conducted may be important for understanding patterns of local adaptation, as well as the genetic architecture of adaptive traits, because selective processes can differ between regions, populations and species (Anderson et al., 2015; Richardson et al., 2014).

Here, we investigate the genetic components underlying local adaptation at two spatial scales in a widespread foundation species, and one of the world’s tallest trees, mountain ash (Eucalyptus regnans). Mountain ash grows in mid-elevation forests of the south-eastern Australian states of Victoria and Tasmania, often in monotypic stands. It has a patchy distribution across a large area (700 km by 500 km), with substantial geographical barriers to gene flow existing between patches. Importantly, the Bass Strait separates the Victorian and Tasmanian parts of the species range, with a distance of at least 200 km likely preventing any gene flow between these regions for at least ten thousand years (Duncan et al., 2016; Lambeck et al., 2014). Within each state, the distribution is further divided by inhospitable low-lying habitats, which have led to levels of genetic differentiation between regions that are approximately equal to that between the states (von Takach Dukai et al., 2019). Forests dominated by mountain ash are also considered critically endangered, with wildfire and logging (in conjunction with specific life history traits) causing substantial changes to the age structure and structural characteristics of the forest (von Takach Dukai et al., 2018; Burns et al., 2015; Lindenmayer et al., 2015).

Our primary aims here are to identify and describe signatures of local adaptation in the genome of mountain ash, and determine the extent to which patterns of local adaptation vary across two geographic scales; (1) the full range of the species and (2) within the spatially-restricted region of the Victorian Central Highlands (VCH). The VCH is considered a suitable small-scale region for study as it is topographically complex, and mountain ash is more or
less continuously distributed throughout the region with very low levels of genetic differentiation (von Takach Dukai et al., In press). We utilise three methods to identify loci under putative selection, including (1) genetic differentiation ($F_{ST}$) outlier analysis, (2) latent-factor mixed modelling, and (3) redundancy analysis, in order to maximise information content and allow for comparison between analytical methods. We predict that low levels of genetic differentiation between mountain ash populations ensures that a high level of standing genetic variation is available in most populations, with low levels of allelic isolation suggesting that many of same alleles will be available for selection to act on across both large- and fine-scale environmental gradients. In this context, we hypothesise that many of the loci under putative selection at the range-wide spatial scale will also be identified at the small spatial scale. However, we also suggest that fundamental differences between spatial scales (such as the extent of variation in environmental covariates), as well as the large array of potential genetic pathways available for adaptive change, will lead to the identification of more putatively adaptive loci than would be found if a single spatial scale was investigated. Thus, we suggest that by examining the similarities and differences between these two spatial scales we will gain a greater understanding of the extent and distribution of adaptive genetic components across the genome.
5.3 Materials and methods

5.3.1 Sample collection

We collected 442 cambium tissue samples of mountain ash from across its natural distribution in Victoria and Tasmania (Figure 5.1). All samples were from stands older than about 50 years to ensure that modern methods of artificial reseeding were not impacting natural levels of genetic diversity and population differentiation. We collected 313 of the 442 samples from 19 widespread sites (mean of 16.5 samples per site), typically by taking a sample from the nearest tree at every 20 m interval along a 180 m transect. Linear transects and exactly even sample sizes were not possible across all sites as the combined effects of logging and wildfire has reduced the size and density of old trees (Lindenmayer et al., 2016) in many parts of the distribution. At nine sites, we also collected a second sample from one of the trees, to act as a biological replicate from the field. The remaining 128 samples were collected from across the Central Highlands region of Victoria (VCH, Figure 5.2), with variable distances between samples and at a range of elevations and topographic conditions. As messmate (Eucalyptus obliqua) commonly hybridises with mountain ash, we collected 20 messmate samples to allow for the calculation of ancestry coefficients and exclude admixed individuals from the dataset (von Takach Dukai et al., 2019). All samples were placed in individual brown paper bags for air-drying and then stored at 4°C prior to DNA extraction.
Figure 5.1 Location of all mountain ash (Eucalyptus regnans, blue circles) and messmate (Eucalyptus obliqua, red circle) tissue samples collected in southeastern Australia for genotyping-by-sequencing. Species distributions are derived from records of each species in the Atlas of Living Australia. Number of samples collected at each site is represented by point size. Distributions of mountain ash and messmate are shown in green and grey respectively.
Figure 5.2 Map showing location of all mountain ash (*Eucalyptus regnans*) tissue samples collected in the Victorian Central Highlands. Elevation (m) is shown in greyscale and the distribution of mountain ash is shown in light green. Sampling locations are indicated by blue circles.
5.3.2 Sample preparation and sequencing

We homogenised samples using an automated cryogenic tissue grinding machine (Labman Max Planck Cryogenic Grinder Dispenser, Labman Automation & Custom Robotics), and stored the ground samples at -18°C. We extracted whole genomic DNA in plate format, using a randomised order of samples and following the kit (Stratec Invisiorb DNA Plant HTS 96) instructions.

Library preparation for genotyping-by-sequencing was conducted as per von Takach Dukai et al. (2019), and included (1) digestion using PstI restriction enzyme digestion (New England BioLabs Inc.), (2) fragment ligation with a unique barcode adapter pair, (3) purification and PCR amplification using two GBS primers (Integrated DNA Technologies), (4) post-purification quantification using microfluidic capillary electrophoresis (PerkinElmer Labchip GX-II), (5) equimolar pooling using an automated robotic liquid handling machine (PerkinElmer NGS Express), and (6) purification using a Genelute PCR Clean-Up Kit (Sigma-Aldrich). The pooled samples were then size fractionated and a 250-450 bp gel cutout was performed, with sequencing conducted on multiple lanes of an Illumina HiSeq 2500 platform (along with hundreds of additional mountain ash samples used for separate projects) using a 100-base paired-end read.

5.3.3 Bioinformatic pipeline and hybrid exclusion

We demultiplexed sequencing data using exact matches and combinatorial index mode with Axe (Murray and Borevitz, 2018), with about 7% of reads unable to be assigned to a sample. Adapters were removed and reads quality-trimmed (Phred score Q=30) at both ends using BBduk, with NextGenMap (Sedlazeck et al., 2013) used to align reads to the E. grandis v2.0 reference genome (Bartholome et al., 2015; Myburg et al., 2014). We then used SAMtools (Li et al., 2009) to create sorted Binary Alignment/Map (BAM) files for each sample.

To obtain a sample-by-SNP matrix, we used the ANGSD software package (Korneliussen et al., 2014) to calculate GATK genotype likelihoods (McKenna et al., 2010). The likelihoods were then used to call initial genotypes using loci that (1) were highly likely to be polymorphic (likelihood ratio test p-value ≤0.001), (2) were genotyped in at least 50 individuals, (3) had a minimum average sequencing depth per sample of 0.5, and (4) had a maximum average sequencing depth per sample of 1000. We only called genotypes where the posterior genotype probability (assuming a uniform prior) was at least 0.95. This produced a sample-by-SNP matrix containing 462 samples and 25867 SNPs, which was read into R (R Core Team, 2017) for further filtering and analysis.
To exclude hybrids from the dataset, we first filtered out SNPs with low (< 0.5) call-rates (proportion of samples in which the site was genotyped) or minor allele frequencies (< 0.05), retaining 1606 SNPs. We also removed samples in which less than 50% of loci were genotyped, retaining 394 samples. The resulting matrix was read into the ‘SNPRelate’ package (Zheng et al., 2012; Zheng and Weir, 2016), and we used the snpgdsAdmixProp function to investigate individual ancestries via an eigen-analysis approach. Any samples with more than 10% admixture with messmate were removed from the analysis, which retained 301 samples.

Once hybrids had been removed, the initial 25867 SNPs were filtered on call rate (≥ 0.5) and minor allele frequency (MAF ≥ 0.05). We did this for two sample sets; the complete range-wide dataset and a reduced dataset containing only samples collected from the VCH. We created hierarchical clustering dendrograms of these datasets using the hclust function on a Euclidean distance matrix, and inspected them visually to ensure that all biological replicates were closely paired, indicating data reliability and lack of contamination. Biological replicates were then removed from the analysis, retaining 292 samples for the range-wide dataset and 139 samples for the VCH dataset. These two datasets represent different geographical scales at which we were able to search for signs of local adaptation, via genome scans for loci exhibiting significantly greater genetic differentiation than background levels, as well as by searching for loci that were correlated with environmental variables.

### 5.3.4 Genomic position and local selection

Using the method of Steane et al. (2017b), we obtained all gene positions in *E. grandis* v2.0 (Bartholome et al., 2015; Myburg et al., 2014) and determined whether each of the 1960 SNPs was genic (located within a gene), near-genic (within 5000 bp of a gene), or nongenic (> 5000 bp from the start or end points of a gene). While the extent of linkage disequilibrium (LD) can vary across the genome (Gion et al., 2016; Silva-Junior and Grattapaglia, 2015), plotting the decay of LD by physical distance between all loci (Supplementary material Figure S5.1) suggested that 5000 bp was a reasonable distance category to use for this study.

To compare the influence of genomic position on standard genetic parameters, we calculated the means and standard errors of the inbreeding coefficient ($F_{IS}$), fixation index ($F_{ST}$), expected heterozygosity ($H_e$), and observed heterozygosity ($H_0$) across all loci (for the range-wide dataset only) and plotted the results as barcharts using ‘ggplot2’. As genic SNPs are more likely to be under positive selection than nongenic SNPs, they often make up a greater proportion of SNPs with high $F_{ST}$ values (Barreiro et al., 2008; Allendorf et al., 2013). We calculated the $F_{ST}$ for each SNP and created kernel smoothed histogram (density) plots...
for genic and nongenic loci to visually inspect whether the patterns of range-wide $F_{ST}$ differed among genic and non-genic SNPs.

### 5.3.5 Geographic variation in genetic diversity

As an initial investigation into geographic patterns of genetic diversity, we calculated $H_E$, $H_O$, $F_{IS}$, and the percentage of polymorphic loci for each population (where $n > 8$) using GenAlEx v6.503 (Peakall and Smouse, 2006; Peakall and Smouse, 2012). We then determined whether individual heterozygosity ($H_I$, the proportion of heterozygous loci in a sample’s genotype) was associated with any environmental variables, by fitting a linear mixed effects model with $H_I$ as the response variable. If environmental covariates are associated with $H_I$ this would potentially indicate topographically-driven influences on genetic diversity. For example, the direction of prevailing winds across the landscape may drive pollinator or seed immigration into a site or region. To obtain predictor variables, we uploaded the latitude and longitude of each individual into the Spatial Portal of the Atlas of Living Australia (ALA, [http://spatial.ala.org.au/](http://spatial.ala.org.au/)). We extracted the values for 28 environmental variables at each location, including 10 precipitation, 3 radiation, 8 temperature, 3 soil attribute, and 4 topographic variables. Environmental variables used in the ALA were collated or derived from multiple sources (Williams et al., 2010; De Vries, 2009; Xu and Hutchinson, 2013; Xu and Hutchinson, 2011; Williams et al., 2006). As topographic aspect was in degrees, we transformed this into two new variables for each individual, consisting of an east-west component and a north-south component, producing a total of 29 predictor variables (Supplementary material Table S5.1). We then calculated pairwise correlations between the 29 variables and removed them one-by-one until there were no more pairwise correlations $> 0.7$, which retained 13 variables (Supplementary material Figure S5.2). We also checked multicollinearity using variance inflation factors (VIFs), with VIF $> 10$ considered too high for a variable to be used in the analysis. We fitted a mixed effects model using all 13 predictor variables, as well as a random effect of site, using the ‘lme4’ package (Bates et al., 2015). The ‘MuMIn’ package (Barton, 2016) was then used to perform model selection, with spatial autocorrelation investigated in the model residuals. We then fitted all possible sub-models of the global model and ranked them by Akaike’s information criterion corrected for finite sample sizes ($AIC_c$). Of the top-ranked models (i.e. those with $\Delta AIC_c < 2$), the most parsimonious model was considered the most informative, with the model parameters inspected and the effect size of each predictor variable on $H_I$ investigated.
5.3.6 Spatial genetic structure and $F_{ST}$ outliers

Next, we used TESS3R (Caye et al., 2016) to generate maps of modelled genetic structure across the range of the species. TESS3R applies a model of genetic structure featuring a discrete number (K) of ancestral populations and estimates the K value that best represents the data using cross-entropy metrics. The best-supported scenario was then mapped across the landscape. We conducted a genome scan using the \textit{pvalue} function of TESS3R, to identify any SNPs with significant genetic differentiation ($F_{ST}$) outliers. These outliers indicate loci that are potentially under selection by identifying strong allele frequency differences between populations. The genome scan was then repeated for the VCH-only dataset.

Selective processes at a site may only be responsible for small shifts in allele frequencies, which means that $F_{ST}$ outlier approaches could result in false negatives, or at least low overlap with more complex models (de Villemereuil et al., 2014; Rellstab et al., 2015). Thus, to further search for signatures of local adaptation in mountain ash, we also used two alternative methods of genotype-environment association (GEA) to identify loci that were associated with environmental variables.

5.3.7 Genotype-environment association

After imputing all missing genotype data, we implemented two GEA methods for both the range-wide dataset and the VCH-only dataset. We first used latent factor mixed models (LFMM) in the LEA package to (Frichot and François, 2015) to identify allele frequencies that were correlated with the first principal component (PC) of the environmental variables. This method controls for population structure using a value of K (the number of latent factors) derived from a cross-entropy plot similar to the TESS3R package. We adjusted the p-values using a reduced value (0.4) of the genomic inflation factor to improve the shape of the histogram (Martins et al., 2016), and then used a Benjamini-Hochberg algorithmic correction (Benjamini and Hochberg, 1995) to ensure a low rate of false discovery (corrected to 1 in 1000 SNPs). A Manhattan plot showing the positions of all loci under putative selection was created, and each SNP was identified as being genic, near-genic or non-genic as described above.

The second method of GEA we used was redundancy analysis (RDA), which summarises genotypes into sets of covarying markers and identifies markers that are most strongly correlated with individual environmental variables (Forester et al., 2018). We used the 13 environmental variables identified for the mixed model creation above, to allow for correlations with individual environmental variables to be identified. The \textit{rda} function of the
‘vegan’ (Oksanen et al., 2018) package was used to conduct the RDA, and the proportion of the variance explained by the environmental variables estimated using the RsquareAdj function. The significance of the linear relationship between each constrained axis and the environmental variables was then assessed using the anova.cca function. Outlier loci for each significant axis were then identified using a standard deviation of 3 (two-tailed p-value = 0.0027). The predictor variable that each outlier locus was most correlated with, along with the strength of the correlation and whether the locus was genic, near-genic or non-genic, was then determined and tabulated.

To determine whether the loci that were identified as being significantly correlated with environmental variables in either the LFMM or RDA analyses showed greater levels of genetic differentiation than the background level across the genome, we plotted the adjusted p-values of all loci used in the genome scan and inspected the locations of the putatively adaptive loci.

5.3.8 Spatial structure in candidate SNPs

Following identification of a set of loci that were putatively under selection based on the GEA and outlier analyses, we explored patterns of spatial genetic structure in these loci relative to the entire SNP panel. We used the spatial autocorrelation analysis function in GenAlEx to determine the strength of spatial autocorrelation between genotypes at eight distance classes. Theoretically, candidate loci should have a stronger pattern of spatial genetic structure, with the spatial scale of heterogeneity in environmental variables determining the strength of spatial autocorrelation. Thus, we compared spatial autocorrelation among individual genotypes using genetic distances calculated across all 1960 loci and across the 30 candidate loci, as well as a locus-by-locus analysis for the 30 candidate loci. Significant autocorrelation at all distance classes was determined by plotting the results on a correlogram and inspecting whether either of the bootstrap confidence intervals of the spatial autocorrelation coefficient $r$ intercepted zero on the y-axis.
5.4 Results

5.4.1 Filtering

The mean and median read depths per site per sample for genotypes called in ANGSD were 17.4 and 8.1, which we considered to be suitable for subsequent analyses. Further, inspection of the hierarchical clustering dendrograms showed that all biological replicates were closely paired, indicating that contamination was not an issue (Supplementary material Figure S5.3).

Of the 442 putative mountain ash samples, 66 were removed due to having > 10% admixture with messmate, and 53 were removed due to having > 50% missing data. This retained a total of 292 mountain ash samples in the range-wide dataset and 139 samples in the VCH-only dataset. After filtering out SNPs with call-rates of < 0.5 and MAFs < 0.05, 1960 SNPs were retained in the range-wide dataset and 2069 SNPs in the VCH-only dataset, of which 1746 SNPs were common to both datasets.

5.4.2 Genomic position

Of the 1960 SNPs identified using the range-wide dataset, 669 were genic, 829 were nongenic, and 462 were near-genic. Population genetic parameters varied slightly between genic and non-genic datasets, with genic loci showing higher $H_O$ and $H_S$, and lower $F_{IS}$, than non-genic loci (Figure 5.3). The mean level of genetic differentiation was very similar between genic and non-genic loci, with density plots of $F_{ST}$ values showing no clear differences between the two groups, and no bump in the upper tail of genic loci (Supplementary material Figure S5.4). Only one locus, a non-genic SNP located on an unanchored scaffold, had an $F_{ST}$ value > 0.6. Taking the top 5% of the $F_{ST}$ distribution of SNPs resulted in the retention of 98 loci. Of these, 33 were genic, 41 were nongenic, and 24 were near-genic, respectively 4.9%, 4.9% and 5.1% of the total number of loci in each category. This suggests that loci with high $F_{ST}$ values were no more likely to be genic than non-genic.
Figure 5.3 Comparison of genetic parameters, including the inbreeding coefficient ($F_{IS}$), fixation index ($F_{ST}$), expected heterozygosity ($H_E$), and observed heterozygosity ($H_O$), for 292 mountain ash (*Eucalyptus regnans*) individuals. Parameters were calculated from genome-wide SNPs separated based on whether they were located inside genes (n=669) or outside genes (n=829). A SNP was defined as being outside a gene if it was more than 5000 bp from the start or end point of a known *Eucalyptus grandis* gene, with SNPs < 5000 bp from the start or end points not included in the analysis. Error bars represent standard errors of the mean.
5.4.3 Geographic variation in genetic diversity

Population genetic parameters did not vary substantially between sites or regions. Mean site-level values (±SE) of $H_E$, $H_O$, and $F_{IS}$ were 0.283 (0.001), 0.362 (0.002), and -0.211 (0.002) respectively, with selective elimination of inbred genotypes likely responsible for the heterozygote excess (Griffin et al., 2019). Genetic differentiation between populations was low with a mean (±SE) pairwise $F_{ST}$ of 0.038 (0.001). The percentage of polymorphic loci was high in all populations (mean 87.7%), with the highest percentage (91.7%) in Tarra Bulga NP and lowest in the Errinundra Plateau and Erskine Falls populations (80%).

Modelling showed that aspect was a significant predictor of $H_I$, with both east-west aspect and north-south aspect retained in the most parsimonious model. While significant effects were identified for these variables, particularly with respect to east-west aspect (with higher $H_I$ on westerly aspects), the effect size was very small, with only minor differences in heterozygosity (e.g. 0.02) observed between extreme differences in aspect. Other variables were found in some of the top-ranked models; however none were considered important predictors of $H_I$. To further investigate the significance of aspect, we attempted to determine whether there was some directionality to gene flow (e.g. through prevailing winds assisting travel of pollinators or seeds) by running spatial autocorrelation analyses in GenAlEx. Pairwise geographic distances between all samples were separated based on whether the bearing between pairwise samples fell into one of four bearing axes (N-S, NE-SW, E-W, and SE-NW), and a spatial autocorrelation analysis was run on each of the four groups. No differences were observed in spatial autocorrelation between any of the axes of directionality in either the range-wide dataset or the VCH-only dataset, suggesting no evidence of directional gene flow.

5.4.4 Spatial genetic structure and $F_{ST}$ outliers

Investigation of the scree-plot of cross-validation scores (created using the TESS3R package) suggested there were five main components in the range-wide genetic dataset. While the overall level of genetic differentiation in the range-wide dataset was low ($F_{ST} = 0.02$), a small amount of geographic structuring was present (Figure 5.4). This is typical of long-lived forest trees such as mountain ash, which often have large effective population sizes, predominantly outcrossing mating systems, and high levels of gene flow and fecundities (Lind et al., 2018). After adjusting for the false discovery rate, no candidate SNPs were identified in the genome scan for $F_{ST}$ outliers in either the range-wide or VCH-only dataset.
Figure 5.4 Population structure in mountain ash (*Eucalyptus regnans*), identified using 1960 genome-wide SNPs. Colours indicate ancestry coefficients, and black points indicate sampling locations. The actual range of the species is considerably smaller than the extent of colouring and is shown in Figure 5.1.
5.4.5 Genotype-environment association

The PCA analysis showed that elevation and solar radiation variables were negatively correlated with temperature variables (Supplementary material Figure S5.5). Variables relating to temperature, precipitation, radiation, and elevation contributed most to the first principal component (PC1), with temperature, radiation and some precipitation variables contributing most to the second principal component (PC2).

Using the range-wide dataset, 10 loci were identified by the LFMM as being associated with PC1 (Figure 5.5). Of these 10, four were genic, four were near-genic, and two were non-genic. Using the VCH-only dataset, four loci were identified by the LFMM as being associated with PC1. Of these four, two were non-genic, one was genic and one was near-genic. No loci were identified in both the range-wide and VCH LFMM analyses.

For the range-wide RDA, thirteen environmental predictor variables were retained (Figure 5.6). The RDA model had an adjusted $R^2$ value of 0.01, which is not unexpected considering that most SNPs are likely to be neutral. Of the 1960 SNPs investigated, 23 loci were significantly correlated with environmental variables, of which three pairs were in linkage disequilibrium (Table 5.1). Ten of these were genic, four were near-genic, and nine were non-genic. Eight loci were correlated most strongly with soil phosphorous; six with annual rainfall; three with annual temperature; two with annual radiation; and one with each of elevation, maximum temperature of the warmest month, topographic relief, and topographic slope.

Loci that showed significant associations with environmental variables in either the LFMM or RDA analyses did not exhibit significantly greater levels of $F_{ST}$ than the background level across the genome, with many other loci showing substantially more departure from background levels (Figure 5.7).

Of the 23 loci identified by the RDA, three were also identified in the LFMM analysis (Table 5.1), including two genic SNPs (one each on chromosome four and chromosome five), and one non-genic SNP (on scaffold 124). As these loci were identified using two separate approaches, we consider these the strongest candidates. Of the three, RDA showed that one candidate was correlated most with annual radiation, one with maximum temperature of the warmest month, and one with annual rainfall. Of the total 12 genic loci identified, gene expression products were available for nine, and related to various proteins, enzymes and transporters (Table 5.1).

For the VCH dataset, 10 predictor variables were retained. The RDA had a very low adjusted $R^2$ (0.003), but there was still a significant influence of the predictor variables on the genetic dataset ($p = 0.041$). Of the 2069 SNPs investigated, six loci were identified as being
correlated with environmental variables; two with topographic relief, and one with each of annual rainfall, summer rainfall, elevation, and north-south aspect. Of the six loci, two were genic and in linkage disequilibrium (90 bp apart on chromosome six). The remaining four loci were non-genic.

No loci were identified in both the range-wide and VCH RDA analyses; however, there were two instances in which an outlier locus identified in the range-wide analysis was in linkage disequilibrium with an outlier locus identified in the VCH analysis ($D' = 1$ in both pairs). In both cases the SNP pairs were non-genic and located less than 15 base pairs apart, with one pair located on chromosome three and the other on chromosome seven. In both instances each locus in the pair was correlated most strongly with differing environmental predictors (Table 5.1).
Figure 5.5 Manhattan plot of adjusted and transformed p-values, which were used to determine whether the allele frequency for each of 1960 genome-wide SNPs in mountain ash (*Eucalyptus regnans*) was significantly correlated with the first principal component of 29 environmental variables. Loci highlighted in blue are those that were significantly correlated after adjusting for a false discovery rate of 1 in 1000 SNPs. Loci are ordered numerically by chromosome and position, with unanchored scaffolds pooled on the far right.
Figure 5.6 Triplot showing the relationship that 1960 SNPs (grey points) and 292 mountain ash (*Eucalyptus regnans*) individuals (coloured circles) had with the first two ordination axes resulting from linear combinations of 13 environmental variables (blue). Individuals are been coloured according to four broad geographic regions of the species distribution.
Figure 5.7 Manhattan plot showing adjusted and transformed p-values used to determine whether each of 1960 SNPs had a significantly different level of genetic differentiation ($F_{ST}$) to the background rate of differentiation across the genome of mountain ash (Eucalyptus regnans). The 30 SNPs highlighted in blue are those that were identified as being significantly associated with environmental variables across the range of the species by either latent factor mixed modelling or redundancy analysis. Loci are ordered numerically by chromosome and position, with unanchored scaffolds pooled on the far right.
Table 5.1 Loci showing signatures of local adaptation in mountain ash (*Eucalyptus regnans*). Latent factor mixed models (LFMM), using the first principal component of 29 environmental variables, and redundancy analysis (RDA), using 13 environmental variables, were used to identify signatures of adaptation. Where the chromosome number has an asterisk, that SNP is likely to be linked to the following SNP. Gene products from the *E. grandis* v2.0 genome annotation are shown for genic SNPs, unless a description was not available (N/A). Two sample sets, collected on different geographic scales, were used to identify SNPs showing signs of selection, including a range-wide set (RW) and a central Victorian set (VCH).

<table>
<thead>
<tr>
<th>Chromosome/scaffold #</th>
<th>Position</th>
<th>Genic/non-genic</th>
<th>Gene product</th>
<th>Correlated predictor</th>
<th>RDA correlation</th>
<th>Method used</th>
<th>Dataset used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr 1</td>
<td>4764980</td>
<td>Non-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 1</td>
<td>13500062</td>
<td>Genic</td>
<td>N/A</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 1*</td>
<td>21167390</td>
<td>Genic</td>
<td>N/A</td>
<td>Annual rainfall</td>
<td>0.26</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 1</td>
<td>21167439</td>
<td>Genic</td>
<td>N/A</td>
<td>Annual rainfall</td>
<td>0.27</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 2</td>
<td>9845247</td>
<td>Non-genic</td>
<td>-</td>
<td>Relief</td>
<td>0.3</td>
<td>RDA</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 2*</td>
<td>19592021</td>
<td>Near-genic</td>
<td>-</td>
<td>Annual rainfall</td>
<td>0.15</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 2</td>
<td>19592169</td>
<td>Near-genic</td>
<td>-</td>
<td>Soil phosphorus</td>
<td>0.19</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 2</td>
<td>27676135</td>
<td>Non-genic</td>
<td>-</td>
<td>Summer rainfall</td>
<td>0.23</td>
<td>RDA</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 2</td>
<td>33134767</td>
<td>Non-genic</td>
<td>-</td>
<td>Annual temperature</td>
<td>0.18</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 3</td>
<td>8832847</td>
<td>Near-genic</td>
<td>-</td>
<td>Annual rainfall</td>
<td>0.19</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 3*</td>
<td>20400828</td>
<td>Non-genic</td>
<td>-</td>
<td>Soil phosphorus</td>
<td>0.24</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 3</td>
<td>20400841</td>
<td>Non-genic</td>
<td>-</td>
<td>Elevation</td>
<td>0.24</td>
<td>RDA</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 3</td>
<td>59334637</td>
<td>Genic</td>
<td>Protein kinase family protein</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 3</td>
<td>60193946</td>
<td>Genic</td>
<td>Casein kinase</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 3</td>
<td>70930232</td>
<td>Non-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 4</td>
<td>38012081</td>
<td>Genic</td>
<td>Abietadienol hydroxylase</td>
<td>Annual radiation, PC1</td>
<td>0.21</td>
<td>LFMM, RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 5</td>
<td>574454</td>
<td>Near-genic</td>
<td>-</td>
<td>Elevation</td>
<td>0.23</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 5</td>
<td>598887</td>
<td>Genic</td>
<td>Cyclic nucleotide-gated ion channels</td>
<td>Warmest month maximum temperature, PC1</td>
<td>0.28</td>
<td>LFMM, RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 5</td>
<td>35514365</td>
<td>Near-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 5</td>
<td>48354994</td>
<td>Non-genic</td>
<td>-</td>
<td>Annual rainfall</td>
<td>0.25</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 5</td>
<td>61162241</td>
<td>Non-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 6*</td>
<td>7150146</td>
<td>Genic</td>
<td>N/A</td>
<td>North-south aspect</td>
<td>0.23</td>
<td>RDA</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 6</td>
<td>7150236</td>
<td>Genic</td>
<td>N/A</td>
<td>Annual rainfall</td>
<td>0.17</td>
<td>RDA</td>
<td>VCH</td>
</tr>
</tbody>
</table>

(continues)
<table>
<thead>
<tr>
<th>Chromosome/scaffold #</th>
<th>Position</th>
<th>Genic/non-genic</th>
<th>Gene product</th>
<th>Correlated predictor</th>
<th>RDA correlation</th>
<th>Method used</th>
<th>Dataset used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr 6</td>
<td>26031624</td>
<td>Near-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 6</td>
<td>31294868</td>
<td>Genic</td>
<td>N/A</td>
<td>Topographic relief</td>
<td>0.17</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 6</td>
<td>55918684</td>
<td>Near-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 7</td>
<td>9425605</td>
<td>Genic</td>
<td>EamA-like transporter family</td>
<td>Annual temperature</td>
<td>0.18</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 7*</td>
<td>17615405</td>
<td>Non-genic</td>
<td>-</td>
<td>Topographic relief</td>
<td>0.32</td>
<td>RDA</td>
<td>VCH</td>
</tr>
<tr>
<td>Chr 7</td>
<td>17615410</td>
<td>Non-genic</td>
<td>-</td>
<td>Annual temperature</td>
<td>0.2</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 8</td>
<td>32792928</td>
<td>Non-genic</td>
<td>-</td>
<td>Soil phosphorus</td>
<td>0.21</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 9</td>
<td>5868764</td>
<td>Genic</td>
<td>Protein brassinosteroid insensitive-1</td>
<td>Soil phosphorus</td>
<td>0.22</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 9</td>
<td>5885154</td>
<td>Genic</td>
<td>Protein brassinosteroid insensitive-1</td>
<td>Soil phosphorus</td>
<td>0.22</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 9</td>
<td>9591035</td>
<td>Near-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 9</td>
<td>9698767</td>
<td>Near-genic</td>
<td>-</td>
<td>PC1</td>
<td>-</td>
<td>LFMM</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 10</td>
<td>360261</td>
<td>Non-genic</td>
<td>-</td>
<td>Annual radiation</td>
<td>0.3</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 10</td>
<td>28242055</td>
<td>Genic</td>
<td>Lysine histidine transporter 1-related</td>
<td>Soil phosphorus</td>
<td>0.22</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 11*</td>
<td>5302886</td>
<td>Non-genic</td>
<td>-</td>
<td>Soil phosphorus</td>
<td>0.19</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Chr 11</td>
<td>5302926</td>
<td>Non-genic</td>
<td>-</td>
<td>Soil phosphorus</td>
<td>0.2</td>
<td>RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Scaffold 124</td>
<td>18913</td>
<td>Non-genic</td>
<td>-</td>
<td>Annual rainfall, PC1</td>
<td>0.54</td>
<td>LFMM, RDA</td>
<td>RW</td>
</tr>
<tr>
<td>Scaffold 215</td>
<td>141678</td>
<td>Genic</td>
<td>Sugar transporter ERD6-like 14-related</td>
<td>Topographic slope</td>
<td>0.29</td>
<td>RDA</td>
<td>RW</td>
</tr>
</tbody>
</table>
5.4.6 Spatial structure in candidate SNPs

We found considerably stronger spatial autocorrelation when using the 30 candidate loci, compared with the full 1960 loci. Within sites (<200 m), the spatial autocorrelation coefficient $r$ was greater than 0.15 for the candidate loci, but less than 0.05 for the full dataset (Figure 5.8). The intercept of $r$ with $y=0$ occurred at 90 km in the full dataset and 158 km when using only the candidate loci. Together these results suggest that candidate loci exhibit not only stronger spatial autocorrelation, but also autocorrelation across greater distances than would be expected from a random sample of SNPs, which is likely being driven by the scale at which environmental conditions are varying. Investigation of autocorrelograms for individual candidate loci showed that some loci exhibit very strong ($r > 0.25$) spatial structuring, indicative of strong divergent selection.
Figure 5.8 Spatial autocorrelograms showing the difference in spatial structure observed between SNPs with signatures of local adaptation (“Candidate SNPs”, n = 30) and SNPs from across the genome (“All SNPs”, n=1960). Candidate SNPs have been analysed together (A) and individually (B). Analyses were done using 292 mountain ash (*Eucalyptus regnans*) from across the species distribution. Note the different y-axis scales between the plots.
5.5 Discussion

We compared the signatures of local adaptation across the genome of a well-known and economically important foundation tree species at two spatial scales, identifying numerous loci under putative selection. Importantly, a large proportion of these loci exhibited signatures of selection at one spatial scale but not both. Our research adds to a growing body of literature that investigates genome-wide SNP loci for signatures of local selection (Lind et al., 2017; Yang et al., 2017; Hecht et al., 2015), and builds upon recent Australian studies that have identified an array of genotype-environment associations in eucalypts (Steane et al., 2014; Steane et al., 2017a; Jordan et al., 2017; Ahrens et al., 2019). Our results support the concept that by examining selective processes at multiple spatial scales we can gain a more complete understanding of the genetic architecture of local adaptation (Anderson et al., 2015; Richardson et al., 2014).

5.5.1 Spatial scale and local adaptation

We utilised two GEA methods to search for loci that are associated with environmental variables, on both a large spatial scale (range-wide sample set) and a small spatial scale (Central Highlands of Victoria). Our prediction that analyses at finer spatial scales will allow us to discover additional loci under putative selection, as a result of microgeographic environmental heterogeneity causing divergent selection, was correct. The LFMMs and RDAs found signatures of selection in 30 loci in the range-wide dataset and 10 loci in the VCH dataset, suggesting that soil phosphorus and rainfall variables play a major role in local adaptation. Such results are similar to those published for other eucalypts, with Steane et al. (2014) identifying 43 outlier markers in red ironbark (Eucalyptus tricarpa) that were strongly associated with the first axis of 15 least-intercorrelated climatic variables, and Steane et al. (2017a) finding secondary adaptations associated with soil phosphorous and maximum temperature in smooth-barked York gum (Eucalyptus loxophleba ssp. lissophloia) of Western Australia. In a study of grey box (Eucalyptus microcarpa), Jordan et al. (2017) found that the majority of 81 putatively adaptive SNPs were variously associated with rainfall, temperature, and aridity. Most recently, Ahrens et al. (2019) examined environment associations in 9593 SNPs across the genome of marri (Corymbia calophylla) and found 537 that were putatively adaptive to climate via small patterns of allelic turnover.

While there was no exact overlap in loci under putative selection between our two spatial scales, two loci identified in the range-wide dataset were in linkage disequilibrium with loci identified in the VCH dataset. These results indicate there are genomic regions involved in local adaptation at both spatial scales, but there are many more separate genetic architectures...
that contribute to adaptation via different genetic pathways. The redundancy analyses support the idea that selective forces can differ between spatial scales, with a large proportion (~30%) of candidate loci in the range-wide dataset most strongly correlated with soil phosphorous, whereas for the VCH-only dataset just one candidate SNP had a correlation with soil phosphorous greater than 0.15, and even that locus was more strongly correlated with topographic relief.

Genotype-environment association studies that directly compare patterns of divergent selection on small and large spatial scales are not common in the literature. However, similar results to ours were obtained in a study of the Amazonian forest tree Eperua falcata, where no overlap in outlier loci between regional and local spatial scales was found (Brousseau et al., 2018), and a larger number of outlier loci were identified on the regional scale. A study of great tits (Parus major) in France found that spatial scale was an important determinant of the significance of urbanisation on genetic differentiation (Perrier et al., 2018). Studies of salmonid fishes have also shown that geographically distant populations experience markedly different selection regimes compared to more proximate populations (Fraser et al., 2011), which is likely to give rise to different genomic architectures relating to adaptation. Despite the dearth of such comparative methods in the primary literature, numerous studies have found evidence of microgeographic adaptation (Richardson and Urban, 2013; Antonovics, 2006; Skelly, 2004; Eckert et al., 2015), with selective processes in these local populations likely to differ substantially from broader patterns of adaptation across geographic space.

### 5.5.2 Implications of low genetic differentiation

Our genome scan for $F_{ST}$ outlier loci found no SNPs in either dataset that showed significant departure from background levels of genetic differentiation. Further, our candidate loci were not consistently found in the upper tail of the $F_{ST}$ distribution, but appeared to be more or less randomly distributed throughout the range of $F_{ST}$ values. This is notable, but not entirely unexpected when conducting searches for signs of selection in large SNP datasets (de Villemereuil et al., 2014). Rather than implying there is no local adaptation, it suggests that locally-adapted loci do not show strong departures from background levels of genetic differentiation, and that local adaptation to environmental variables in mountain ash is not strongly linked to isolation of alleles (Evans et al., 2014). Similarly, in a study of loblolly pine (Pinus taeda) in southeastern USA, Eckert et al. (2010) found that $F_{ST}$ values of loci showing significant correlations with environmental variables were within the background level of genetic differentiation. They also found that $F_{ST}$ outliers were located in different regions of the genome to, and showed no meaningful correlations with, loci associated with environmental variables. Such results are not overly surprising in forest trees, considering
that many species show high levels of within-population genetic variation and low levels of allelic isolation (Lind et al., 2018). With pollen flow occurring over long distances, alleles are typically present in most populations, and while selection can filter alleles within each generation, it is unlikely to completely remove alleles from a population (von Takach Dukai et al., 2019). Thus, large differences in allele frequencies between populations tend not to occur and \( F_{ST} \) values remain low. Further, most traits that contribute to local adaptation are polygenic (Savolainen et al., 2013; Lind et al., 2018), and \( F_{ST} \) values tend to decline as the number of loci involved increases, because there are many allelic combinations that can produce optimal trait values (Josephs et al., 2017). This means that the small-effect loci involved in polygenic adaptation often exhibit only subtle allele shifts, and are less likely to be found through standard \( F_{ST} \) outlier sweeps (Berg and Coop, 2014; Kemper et al., 2014).

### 5.5.3 Limitations and future directions

While genotyping-by-sequencing (and similar reduced representation methodologies) produce large, highly informative genomic datasets, they often also result in moderate to high levels of missing data (Supple et al., 2018). As complete datasets are typically required for the GEA analyses, there is a necessary data imputation step, which may reduce the strength of correlations between candidate loci and predictor variables. As a result, the number of false negatives could potentially be inflated. With rapid development in both sequencing techniques and analysis methodologies, this should become less of an issue in the future.

Our LFMM analyses were done using the first principal component of the 29 environmental variables, as this axis is typically the most informative (Steane et al., 2014). However, investigation of other principal components may result in identification of additional locally-adapted loci. Further, it may be useful to separate environmental variables into climatic or geographical groupings prior to conducting the PCA (Dillon et al., 2014), although the development of RDA methods has largely negated the benefit of doing this.

Finally, while local adaptation in forest trees is presumed to occur as a result of selective processes modifying allele frequencies across generations, it is possible that strong environmental selection filters out less fit genotypes mostly within each generation, and high gene flow ‘resets’ the allele frequencies of locally-adapted loci in the seedbank prior to dispersal. A simple method of testing this hypothesis would be to compare allele frequencies at loci showing signs of putative selection in both the seedbank and mature trees within stands. If local adaptation in forest trees such as mountain ash occurs as a result of within-generational filtering of genotypes, there are large implications for planning restoration and assisted migration strategies. With most alleles present in the seedbank at most sites, collection of seeds based on the environmental conditions of source sites would not
necessarily result in a greater level of local adaptation than using seeds collected from random localities, meaning there would be little benefit to such resource-intensive strategies.
5.6 Conclusion

We found evidence of local adaptation across the genome of mountain ash at two different spatial scales, with almost no affiliation between outlier loci at different scales. This shows that while local adaptation can be produced by broad latitudinal and longitudinal variation in environmental conditions, fine-scale environmental gradients (caused by topographic and edaphic factors) can also lead to substantial levels of adaptation, despite much higher levels of gene flow. Loci that are adapted to these fine-scale gradients are likely to go undiscovered without intensive sampling across microgeographic environmental gradients. By studying the processes involved in local adaptation at multiple spatial scales, we gain a deeper understanding of how genetic diversity is maintained within and among populations. This allows us to make genetically-informed management actions that will better conserve or maximise adaptive potential, ensuring that species persist into a changing and uncertain future.
5.7 Acknowledgements

This study was funded by an Australian Research Council Future Fellowship (FT130100043) to SB and a Lesslie Research Scholarship in Landscape Conservation and Ecology to BvTD. Thanks also to Rebecca Taylor and Oliver King for assistance with fieldwork.
5.8 References


https://www.nature.com/articles/ng.78#supplementary-information


Chapter 5: Spatial scale and local adaptation


https://www.nature.com/articles/nclimate3303#supplementary-information


Xu, T. and Hutchinson, M. F., 2011. ANUCLIM, software, version 6.1, Fenner School of Environment and Society, Australian National University, Canberra.


5.9 Supplementary material

Figure S5.1 Decay of linkage disequilibrium as a function of physical distance between loci in mountain ash (Eucalyptus regnans). Vertical error bars represent the standard error of the mean for each distance class.
Figure S5.2 Bivariate scatter plots (below the diagonal), histograms (on the diagonal), and Pearson correlations (above the diagonal) for all pairwise combinations of 13 environmental variables used for latent factor mixed models and redundancy analyses in mountain ash (Eucalyptus regnans). The absolute value of the Pearson correlation is also represented by font size.
Figure S5.3 Hierarchical clustering dendrogram of 394 *Eucalyptus regnans* samples, including hybrid individuals with *Eucalyptus obliqua*. Computations were done using 1606 single-nucleotide polymorphisms from across the genome. Biological replicates are identified with an ‘x’ or ‘y’ after the sample name and are paired together on long branches.
Figure S5.4 Kernel-smoothed histogram (density) plot of SNPs with positive $F_{ST}$ values in 292 mountain ash (*Eucalyptus regnans*) individuals sampled from across the natural geographic range of the species. SNPs have been separated based on whether they were found inside genes (blue, n=669) or outside genes (red, n=829). A SNP was defined as being outside a gene if it was more than 5000 bp from the start or end point of a known *Eucalyptus grandis* gene, with SNPs < 5000 bp from the start or end points not included in the analysis.
Figure S5.5 Biplot of a principal component analysis conducted on 29 environmental variables from across the natural distribution of mountain ash (*Eucalyptus regnans*), showing the contribution that each variable has on the first two principal components (PC1 and PC2).
Table S5.1 Environmental variables used to search for signals of selection in genome-wide SNP datasets for mountain ash (*Eucalyptus regnans*). Latent-factor mixed models (LFMM) and redundancy analyses (RDA) were used. RDA\(^1\) indicates that the variable was used only in the range-wide dataset for that analysis. RDA\(^2\) indicates that the variable was used only in a dataset of samples collected in the Victorian Central Highlands. All variables except for north-south aspect and east-west aspect were downloaded from the Atlas of Living Australia's (ALA) spatial portal ([http://spatial.ala.org.au/](http://spatial.ala.org.au/)). The two aspect variables were derived from a single aspect variable used in the ALA.

<table>
<thead>
<tr>
<th>Full name</th>
<th>Short name</th>
<th>Analysis used</th>
<th>Descriptive information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Precipitation (mm)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean annual precipitation (Bio12)</td>
<td>Annual rainfall</td>
<td>LFMM, RDA</td>
<td>Sum of all 12 monthly precipitation estimates (mm y(^{-1}))</td>
</tr>
<tr>
<td>Precipitation of driest period (Bio14)</td>
<td>Driest period rainfall</td>
<td>LFMM, RDA(^1)</td>
<td>The precipitation of the driest week or month, depending on the time step</td>
</tr>
<tr>
<td>Precipitation of driest quarter (Bio17)</td>
<td>Driest quarter rainfall</td>
<td>LFMM</td>
<td>The driest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated</td>
</tr>
<tr>
<td>Seasonality of precipitation (Bio15)</td>
<td>Rainfall seasonality</td>
<td>LFMM, RDA(^2)</td>
<td>The coefficient of variation is the standard deviation of the weekly precipitation estimates expressed as a percentage of the mean of those estimates</td>
</tr>
<tr>
<td>Mean spring precipitation</td>
<td>Spring rainfall</td>
<td>LFMM</td>
<td>Average precipitation in Sep, Oct and Nov based on 109 years of rainfall data</td>
</tr>
<tr>
<td>Mean summer precipitation</td>
<td>Summer rainfall</td>
<td>LFMM, RDA(^2)</td>
<td>Average precipitation in Dec, Jan and Feb based on 109 years of rainfall data</td>
</tr>
<tr>
<td>Mean autumn precipitation</td>
<td>Autumn rainfall</td>
<td>LFMM</td>
<td>Average precipitation in Mar, Apr and May based on 109 years of rainfall data</td>
</tr>
<tr>
<td>Mean winter precipitation</td>
<td>Winter rainfall</td>
<td>LFMM</td>
<td>Average precipitation in Jun, Jul and Aug based on 109 years of rainfall data</td>
</tr>
<tr>
<td>Precipitation of wettest quarter (Bio16)</td>
<td>Wettest quarter rainfall</td>
<td>LFMM</td>
<td>The wettest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated</td>
</tr>
<tr>
<td>Precipitation of wettest period (Bio13)</td>
<td>Wettest period rainfall</td>
<td>LFMM</td>
<td>The precipitation of the wettest week or month, depending on the time step</td>
</tr>
<tr>
<td><strong>Temperature (°C)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continues)
### Table S5.1 (continued)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean annual temperature (Bio01)</td>
<td>Annual temperature</td>
<td>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Mean of all weekly mean temperatures</td>
</tr>
<tr>
<td>Minimum temperature of the coldest period (Bio06)</td>
<td>Coldest period minimum temperature</td>
<td>LFMM</td>
<td>The lowest minimum temperature in all weeks of the year</td>
</tr>
<tr>
<td>Mean temperature of the coldest quarter (Bio11)</td>
<td>Coldest quarter mean temperature</td>
<td>LFMM</td>
<td>The coldest quarter of the year is determined (to the nearest week), and the mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>temperature of this period is calculated</td>
</tr>
<tr>
<td>Minimum temperature of the coldest month</td>
<td>Coldest month minimum temperature</td>
<td>LFMM</td>
<td>The minimum temperature of any monthly minimum temperature</td>
</tr>
<tr>
<td>Mean temperature of the warmest quarter (Bio10)</td>
<td>Warmest quarter mean temperature</td>
<td>LFMM</td>
<td>The warmest quarter of the year is determined (to the nearest week), and the mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>temperature of this period is calculated</td>
</tr>
<tr>
<td>Mean maximum temperature of the warmest period (Bio05)</td>
<td>Warmest period maximum temperature</td>
<td>LFMM</td>
<td>The highest maximum temperature in all weeks of the year</td>
</tr>
<tr>
<td>Minimum temperature of the warmest month</td>
<td>Warmest month minimum temperature</td>
<td>LFMM</td>
<td>The maximum temperature of any monthly minimum temperature</td>
</tr>
<tr>
<td>Mean maximum temperature of the hottest month</td>
<td>Warmest month maximum temperature</td>
<td>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>The maximum temperature of any monthly maximum temperature</td>
</tr>
</tbody>
</table>

### Radiation (MJ m⁻² day⁻¹)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean annual solar radiation (Bio20)</td>
<td>Annual radiation</td>
<td>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>The mean over the whole year of all the weekly solar radiation estimates</td>
</tr>
<tr>
<td>Mean solar radiation of the coldest quarter (Bio27)</td>
<td>Coldest quarter radiation</td>
<td>LFMM</td>
<td>The coldest quarter of the year is determined (to the nearest week), and the average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>solar radiation over this period is calculated</td>
</tr>
<tr>
<td>Mean solar radiation of the warmest quarter (Bio26)</td>
<td>Warmest quarter radiation</td>
<td>LFMM</td>
<td>The warmest quarter of the year is determined (to the nearest week), and the average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>solar radiation over this period is calculated</td>
</tr>
</tbody>
</table>

### Soil attributes

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-European plant-available nitrogen</td>
<td>Soil nitrogen</td>
<td>LFMM, RDA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Consists of the organic nitrogen in litter and soil plus the mineral plant-available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>nitrogen</td>
</tr>
</tbody>
</table>

(continues)
Table S5.1 (continued)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Method</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-European plant-available phosphorus</td>
<td>LFMM, RDA</td>
<td>Excludes phosphorus that is chemically-bound to soil matrix</td>
</tr>
<tr>
<td>Topographic wetness index</td>
<td>LFMM, RDA</td>
<td>Potential soil moisture. Calculated as $\log$ (specific catchment/local slope)</td>
</tr>
<tr>
<td>Topographic attributes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-south aspect</td>
<td>LFMM, RDA</td>
<td>Calculated as $\cos(\pi \ast \text{aspect}/180)$</td>
</tr>
<tr>
<td>East-west aspect</td>
<td>LFMM, RDA</td>
<td>Calculated as $\sin(\pi \ast \text{aspect}/180)$</td>
</tr>
<tr>
<td>Elevation</td>
<td>LFMM, RDA</td>
<td>Height above sea-level (metres)</td>
</tr>
<tr>
<td>Topographic slope (degrees)</td>
<td>LFMM, RDA</td>
<td>Slope in degrees (°)</td>
</tr>
<tr>
<td>Topographic relief</td>
<td>LFMM, RDA</td>
<td>Range of the 9-second DEM elevation values in each 36 second grid cell (metres)</td>
</tr>
</tbody>
</table>
Chapter 6: Synthesis

6.1 Findings and implications

This thesis aimed to provide novel insights into the demographic and genetic impacts that modified disturbance regimes have on an ecologically and economically important foundation tree species, through investigations of vital rates, hybridisation, population genetic structuring, and local adaptation. To achieve this, I have undertaken a number of quantitative studies that together provide a wealth of new data into the biological processes operating in mountain ash forests. Each of the chapters presented in this thesis has implications for the conservation and management of mountain ash-dominated landscapes, and more generally, forest tree species and forested landscapes worldwide. These implications should help conservation practitioners and management agencies to maximise genetic diversity and adaptive capacity in order to assist population persistence in a changing and uncertain future. The primary findings of these papers are thus summarised here as well as placed into a broader context.

6.1.1 Chapter 2

To investigate the variability between mountain ash populations in their susceptibility to increasing fire frequency, I characterised the response of growth rates and age to maturity in young stands. I found that the time taken to produce seed varied geographically, and that mean growth rates were greater in areas receiving higher levels of solar radiation, a trend that became stronger with tree age. Tree size and tree age had the strongest influence on the production of fruit capsules. Fruit capsules were found in trees as young as 11 years old, but stands may not develop reproductively viable seed crops until they are more than 21 years old.

My results show that environmental factors influence the primary juvenile period of a keystone obligate seeder, in turn affecting the time taken for a population to develop a reproductively viable amount of seed. Reduced fire return intervals may therefore constrain the species’ realised niche (and geographic distribution) to areas where it can tolerate shorter fire return intervals due to faster growth and maturation. I suggest that populations of obligate seeders that reach reproductive viability faster are more likely to persist when exposed to multiple fires in short succession, whereas populations experiencing conditions that slow
plant growth or otherwise extend the time to reproductive viability will be more susceptible to declines (Scheele et al., 2017). Intra-stand variation in seed crops suggests that selection could also act on rapidly-maturing individuals, resulting in some populations exhibiting high levels of precocious reproductive activity. Differences such as these developing between populations has been an important component of plant evolution historically, and thus is of importance to both taxonomic classification and conservation management (Conway and Poethig, 1993; Jones, 1992; Wiltshire, 1991; Potts and Wiltshire, 1997).

6.1.2 Chapter 3

Here, I used over 2400 informative SNP loci from across the genome to examine the extent of hybridisation with messmate stringybark throughout the distribution of mountain ash. I also investigated the population genetic structure of mountain ash when hybrids were included in or excluded from the analysis. I found hybrid individuals at all sites, and two highly introgressed populations. Hybrid individuals were not distributed evenly across environmental gradients, with logistic regression identifying hybrids as being associated with temperature. Removal of hybrids resulted in increases in genetic differentiation ($F_{ST}$), expected heterozygosity, observed heterozygosity and the inbreeding coefficient, and different patterns of isolation-by-distance. After removal of hybrids and introgressed populations, mountain ash showed very little population genetic structure, with a small effect of isolation-by-distance, and very low global $F_{ST}$ (0.03).

These results have a number of implications for the planning of conservation strategies and forest restoration practices. Firstly, the finding that hybrids are associated with environmental variables indicates that seed collection for assisted migration may be confounded by admixture between species. Collection of seeds based solely on the climatic conditions in which they occur, may inadvertently result in the collection of hybrid genotypes rather than locally-adapted mountain ash genotypes. This is not necessarily a negative outcome of management actions, but is unlikely to be the desired outcome.

Once hybrids had been removed from the analysis, mountain ash showed very little genetic differentiation and very little isolation of alleles between populations. This suggests that conservation strategies such as climate-adjusted provenancing may have little benefit, as most populations already contain almost all the genetic diversity present in the species, and selection may simply act within rather than between generations (discussed further in section 6.3). Further, as the risk of outbreeding depression is low, and most adaptive alleles are likely
to be present in seed collected from even distant localities, local provenancing is of little benefit to restoration plantings.

6.1.3 Chapter 4

Here, I examined whether wildfires or silvicultural practices are modifying patterns of within- and among-population genetic diversity, and fine-scale spatial genetic structure, of mountain ash across the landscape. Consistent with the expectation of extensive pollen dispersal but limited seed dispersal, I detected low levels of genetic differentiation at nuclear SNPs ($F_{ST} = 0.016$), and very high levels of differentiation at cpDNA microsatellites ($F_{ST} = 0.751$). The patterns and extent of genetic diversity and genetic differentiation among populations at nuclear SNPs were similar among disturbance histories, although there were minor differences in genetic diversity parameters and fine-scale genetic structure. Analysis of cpDNA revealed significantly higher levels of total and within-site genetic diversity in the logged treatment than the burnt or undisturbed treatments, and suggested that haplotypes are entering the landscape via the use of non-local seed in the silvicultural regeneration process.

These results imply that dispersal of nuclear alleles, which are biparentally-inherited and dispersed via pollen and seeds, tends to occur on a greater scale than dispersal of chloroplast haplotypes, which are maternally-inherited and dispersed only via seeds. It also suggests that high levels of pollen gene flow minimise nuclear genetic differentiation, ensuring that even small stands of trees contain most of the genetic diversity found in the landscape. The high level of standing genetic variation suggests that artificial seeding of logged stands is likely to cause only minor alterations in the patterns of fine-scale genetic structure and genetic diversity of stands of trees, unless seed is sourced from very distance locations (hundreds to thousands of kilometres) or from a very limited number of parents. This could easily be mitigated by adopting either a regional admixture provenancing strategy or utilising variable retention harvesting methods. In addition, the presence of multiple chloroplast haplotypes across the region implies a different phylogeographic history to that proposed by Nevill et al. (2010), who suggested that mountain ash had colonised the region from the south after the last glacial maximum. My data suggest that there were either multiple colonisation events or that local refuges allowed eucalypts to persist in some parts of the landscape.

6.1.4 Chapter 5

To further our understanding of the processes involved in local adaptation to climatic and edaphic variables, I utilised three methods to search the genome of mountain ash for
signatures of local adaptation. This was done at two spatial scales: (1) across the range of the species, in relation to coarse-scale climatic and topographic variation, and (2) within an intensively-sampled sub-region, in relation to fine-scale environmental drivers of climatic and edaphic conditions.

I found low levels of genetic differentiation across the range of the species ($F_{ST} = 0.02$), and no differences in $F_{ST}$ values between genic and non-genic regions of the genome. The genome scan identified no loci in either dataset that showed significant departure from background levels of genetic differentiation. The LFMM and RDA together identified 40 loci that were significantly correlated with environmental predictors after correcting for a strict false discovery rate. Of these 40 loci, 30 were specific to the range-wide dataset, and 10 were specific to the geographically-restricted dataset. However, two loci identified in the range-wide dataset were in linkage disequilibrium with loci identified in the geographically-restricted dataset.

These results suggest that local adaptation in mountain ash is occurring through either weak selective processes or multiple small-effect loci acting on polygenic traits, of which the latter has been noted in a number of other forest tree species (Lind et al., 2018; Aitken and Bemmels, 2016). Thus, while local adaptation is present in mountain ash, high levels of gene flow and strong selection for heterozygous individuals (during stand development) ensure that diversity is maintained within populations and minimise the occurrence of large allele frequency differences between populations. The results also indicate that there are genomic regions involved in local adaptation at both spatial scales, but there are many more separate genetic architectures that contribute to adaptation via different genetic pathways. Thus, the investigation of multiple spatial scales can lead to a greater understanding of patterns of local adaptation.
6.2 Consequences of disturbance

Changes to disturbance regimes can have an array of structural and functional impacts on forest systems. Studies from many regions of the world have shown that timber harvesting and increasing fire frequency influence the species richness and abundance of many plant and animal taxa across a range of forest types (Burivalova et al., 2014; Cochrane and Schulze, 1999; Makana and Thomas, 2006; Gerber et al., 2012; Dixon et al., 2018). In the case of mountain ash forests, some authors have previously suggested that the impacts of logging are broadly equivalent to disturbance by wildfire (Attiwill, 1994; Baker et al., 2004), but more recent studies have noted significant differences in species richness and abundance of various functional groups or specific taxa (Blair et al., 2016; Lindenmayer et al., 2018; Bowd et al., 2018). Additionally, logging and wildfire both operate to reduce the average age class of the landscape, creating a feedback loop of disturbance that is likely to result in the collapse of multiple ecosystem functions (Lindenmayer and Sato, 2018; Zylstra, 2018). Building on this work, my first quantitative chapter found that, due to the influence of topographic factors on growth rates, shorter intervals between wildfires are likely to have differential impacts on populations of obligate seeders such as mountain ash. Faster maturing stands in areas with high growth rates may be more likely to persist, although topographic factors may also partly determine fire occurrence (Wood et al., 2011). Further, the high level of variability in the timing of seed production between both individuals and stands suggests that selection on early-maturing individuals could be an important component of adaptation to frequent disturbance in obligate seeders.

Despite considerable research into the structural and functional changes in mountain ash forests, very little study has been conducted into the impacts that contemporary disturbance regimes can have on population genetic structure. Further, many studies that have attempted to do so did not have access to the large genomic datasets that are now readily available as a result of the advent of next-generation sequencing (Fageria and Rajora, 2014; Rajora and Pluhar, 2003; Glaubitz et al., 2003a; Glaubitz et al., 2003b). While I found differences between disturbance treatments in the population genetic structure of both the nuclear and chloroplast genomes, the impacts that these differences may have on adaptive capacity and local adaptation are not as easily quantified, and require some theoretical interpretation. Like mountain ash, many long-lived forest trees have large effective population sizes, predominantly outcrossing mating systems, and high levels of gene flow and fecundities (Lind et al., 2018). This is an important realisation for understanding how disturbance patterns impact genetic structure, because it means that large reductions in population size need to occur before genetic diversity is substantially impacted (Dillon et al., 2015). Thus,
for many forest tree species, population bottlenecks resulting from disturbance events are unlikely to reduce gene flow and effective populations sizes to levels where genetic diversity will be adversely impacted. This is largely in agreement with my findings in Chapter 4, where just small differences in patterns of population genetic diversity in the nuclear genome were identified between disturbance treatments. It also implies that there is little risk of outbreeding depression as a result of using non-local seed to artificially regenerate stands disturbed by logging or multiple fires. While I did find large differences between treatments in the chloroplast genome, more study is needed to determine whether this is likely to have any meaningful phenotypic impacts on mountain ash. As chloroplast genomes are typically small and show low levels of variability in coding regions (Young et al., 2011), substantial phenotypic impacts are considered unlikely.
6.3 Forest trees and climate change

According to the literature, forest trees should mitigate the impacts of changing climatic conditions in three primary ways, including (1) acclimatisation and phenotypic plasticity, (2) local adaptation through changing allele frequencies, and (3) migration to more suitable climates (Bussotti et al., 2015). In addition to this, I suggest that many forest trees may also adapt to changing climates through adaptive introgression, which has not yet been studied in depth in the primary literature. As the previous chapters of this thesis have primarily related to population genetic diversity and local adaptation, I will primarily discuss here how these factors are likely to influence the ability of forest tree species to adapt to changing environments, however I also discuss the potential for adaptive introgression in section 6.4.1.

In the case of phenotypic plasticity, its ability to contribute to the persistence of populations under changing environments is typically limited at the margins of species distributions and environmental niches. Further, if the plastic response does not match the new favoured phenotypic optimum, directional selection is likely to cause adaptive divergence anyway (Ghalambor et al., 2007). In the case of migration, many species may not be able to migrate quickly enough to match the speed at which the climate is changing, and will require anthropogenic intervention to speed this process up. Translocation of species into new habitable locations outside of their existing range, termed assisted migration, has been proposed where extinction due to climate change is likely (McLane and Aitken, 2012). This approach will be most effective in regions where human activities have created large barriers to migration routes, and for species in which seed dispersal only occurs over short distances. However, it is likely that, for many regions, the in-situ conservation of foundation species (such as forest trees often are) is the only plausible management action to prevent substantial losses in species richness or diversity from ecological communities.

It has been suggested that ecological restoration activities should utilise seed collected from locations where plants have adapted to projected climatic conditions (Prober et al., 2015). While this may be suitable for some species, it may in fact have negligible benefits for many widespread forest trees. Slow generation times and a low level of genetic differentiation (caused by high gene flow) means that even isolated populations typically contain a large proportion of the total allelic diversity in the species (Aitken and Bemmels, 2016; Holliday et al., 2010). My results in Chapter 3, and work by others (Supple et al., 2018; Dillo et al., 2015; Bloomfield et al., 2011), suggests that this holds true for many Eucalyptus species in Australia. These findings have large implications for the concepts of local and climate-adjusted provenancing.
For species in which most of the total genetic variation is contained within local geographic areas, the source of seed used for restoration activities is not likely to be as important as previously considered. Rare new allelic variants are unlikely to be the cause of beneficial adaptation to local conditions (Alberto et al., 2013; Savolainen et al., 2013), and adaptive alleles are likely to be present in seed collected from almost any locality. While loci under selection may exhibit different allele frequencies between populations, there are two reasons that this may not be of great importance. Firstly, reductions in the frequency of some adaptive alleles resulting from the use of non-local seed can be naturally corrected over a small number of generations (Ralls et al., 2018; Fitzpatrick et al., 2015). More critically, however, it is possible that ‘locally-adapted’ loci are simply having their allele frequencies changed over the course of a single generation, with strong selection filtering out less suitable individuals/genotypes. In this scenario, high gene flow due to pollen dispersal results in seed crops containing a mixture of genotypes, and selection after seed dispersal then modifies allele frequencies through time, but is unlikely to completely remove alleles from the population. This may partially explain my findings in Chapter 5, as well as a number of other studies (Lind et al., 2018; Savolainen et al., 2013), where loci showing signatures of selection do not have high levels of genetic differentiation. Together, these studies show that local adaptation in many forest trees is likely to be occurring predominantly through small-effect loci, whereby large numbers of genes contribute to trait expression through varying degrees of nonadditive interactions (Erickson, 2005). It is even possible that, for species with high levels of standing genetic variation, the genetic architecture of locally-adapted traits could vary through time, with phenotypic expression of traits being maintained despite changing frequencies of adaptive alleles (Yeaman, 2015). This is due to the large number of potential genetic and physiological pathways by which local trait adaptation can occur. In practical terms, this suggests that maximising genetic diversity in any given population is more important for maintaining adaptive capacity than attempting to mimic the distribution of allele frequencies in another population.

Finally, further support for the idea that specialised provenancing strategies are less relevant for some forest tree species comes from my data in Chapter 4, where novel chloroplast haplotypes were found only in logged stands. This suggests that forestry operations have been reseeding logged mountain ash sites using seed sourced from outside the region. So, despite many of these logged sites being seeded from multiple distant localities, there has been no observed negative impacts on growth rates or survival to date. Building on such indirect evidence could involve the use of targeted studies such as provenancing trials or reciprocal transplant experiments, which would help to provide strong evidence for or against this idea.
While the research presented here contributes to our understanding of the demographic and genetic impacts of changing disturbance regimes and environmental conditions, it also suggests a number of opportunities for future research avenues. In fact, we are just beginning to identify and address knowledge gaps around many landscape genomic questions (Aitken and Bemmels, 2016), particularly the ways in which species are responding and adapting to changing environments. Understanding how adaptive genetic variation is spatially distributed is a critical component of evolutionary biology and population genetics (Storfer et al., 2018), with two fundamental questions (Manel and Holderegger, 2013): (1) how is neutral and adaptive variation being impacted by changes in disturbance regimes and climate; and (2) what adaptive responses will species exhibit as a result of these changes? The genetic structure of high gene flow forest trees presents a unique set of advantages and challenges to answering these questions, and I suggest that there are two lines of research that show considerable promise in this topic area; (1) the importance of interspecific gene flow in adaptation to changing environments (adaptive introgression), and (2) furthering our understanding of the genetic architecture of local adaptation and the processes that modify allele frequencies at loci under selection.

### 6.4.1 Adaptation through porous genomes

Hybridisation among species is a common phenomenon, with approximately 10-30% of multicellular animal and plant species known to regularly hybridise (Abbott et al., 2013). Considering that novel genetic mutations are very rare, hybridisation thus actually presents a method by which species can obtain large amounts of potentially adaptive genetic variation. For example, studies of Darwin’s finches have suggested that hybridisation has introduced new additive genetic variation at two to three orders of magnitude greater than that introduced by mutation (Grant and Grant, 1994). Further, interspecific gene flow does not necessarily indicate a breakdown of species integrity (De La Torre et al., 2014a; Sullivan et al., 2016). However, despite the growing body of evidence suggesting that hybridisation has been an important component in the evolutionary history of many plant and animal taxa (Trier et al., 2014; Lavrenchenko, 2014; Soltis et al., 2014; Abbott et al., 2013; Alix et al., 2017), very little research has been done that utilises genetic or genomic datasets in combination with methods of estimating fitness of hybrids and parent species, to investigate adaptive processes (Payseur and Rieseberg, 2016).

In Australia, more than half of all eucalypt species form natural hybrid combinations (Griffin et al., 1988; Potts et al., 2003), and many of these combinations can occur at
relatively high frequencies within populations (Field et al., 2009; McKinnon et al., 2010, and data presented in Chapter 3). As a result, this group is ideal for studying the processes involved in adaptive introgression and whether anthropogenic modifications to natural environmental conditions are influencing these processes. This research could potentially build on the style of recent work on spruce (Picea spp.) in North America, which has demonstrated that species can not only adapt to changing environments through a process of adaptive introgression (De La Torre et al., 2014b; Suarez-Gonzalez et al., 2018), whereby hybridisation with congeners allows for the capture of beneficial genetic components from parent species. Hybrid individuals between white spruce (Picea glauca) and Engelmann spruce (Picea engelmannii) have been shown to utilise adaptive introgression as a way to maximise fitness in a changing climate (De La Torre et al., 2014b).

My data in Chapter 3 showed that hybridisation between mountain ash and messmate stringybark occurs in most or all populations. One promising topic of a future study could be to investigate whether such hybrids are more or less fit under warmer and drier climates, or whether hybrid individuals are more likely to survive wildfires (e.g. by having thicker bark or a stronger resprouting response). With environmental changes outpacing the natural ability of many species to adapt through standing genetic variation and mutation (Jordan et al., 2017; Kuparinen et al., 2010), adaptive introgression could provide species such as mountain ash with an alternative pathway, leading to faster development of adaptive traits and rapid adaptation to novel conditions.

6.4.2 Local adaptation in forest trees

Numerous studies have shown that local adaptation occurs in forest trees despite very high gene flow and very low levels of allelic isolation (Lind et al., 2017; Lind et al., 2018; Budde et al., 2014; Eckert et al., 2015; Roschanski et al., 2016). The reasoning behind this is that selection acts to modify allele frequencies across generations, however I suggest that it is possible that selection is modifying allele frequencies within each generation, and that pollen flow acts to ‘reset’ the allele frequencies of locally-adapted loci in the seedbank prior to dispersal. After dispersal, strong environmental selection filters out less fit genotypes. There is already evidence for strong selection in developing stands of mountain ash (Griffin et al., 2019), with homozygous individuals selectively eliminated from stands through time, suggesting that selection for locally-adapted genotypes is a strong possibility. Both the between- and the within-generation scenarios are able to explain the fact that the genetic differentiation of loci under selection is typically not significantly different from background rates across the genome, as selective processes are unlikely to completely remove alleles from a site if alleles are being reintroduced in each generation. As many forest tree species have
long generation times, natural selection acting across generations is less efficient, because there may be substantial lag times between shifting environmental characteristics and changes in allele frequencies. It is more logical that high gene flow and strong selection for heterozygous individuals maintains genetic diversity within populations, allowing the environment to filter out less fit genotypes from a large, highly diverse offspring cohort. This fits well with the results obtained in Chapters 3 and 5, which showed that high levels of genetic diversity are held within populations and that local adaptation likely occurs through small shifts in allele frequency.

There are a number of methodological approaches that studies could use to test whether allele frequencies are changing between generations or whether genotypes are simply being filtered out within each generation as a result of strong environmental selection. As generation times in many forest trees are so long, the utilisation of replicated stands of different ages within geographically-restricted areas is likely to be a practical approach. Additionally, next-generation sequencing data from logged and re-seeded sites could be used to determine whether, despite seed being sourced from multiple and/or distant localities, local adaptation within logged sites is occurring at the same loci as unlogged sites. And a similar genomic dataset could be used to determine whether allele frequencies at loci under selection in naturally-regenerated sites are substantially different to frequencies in seedbanks, building on work undertaken on the Mediterranean maritime pine (Pinus pinaster) (Vizcaíno-Palomar et al., 2014). The results of such studies would have big implications for restoration plantings and assisted migration strategies, because, if local adaptation is occurring within each generation, and most allelic diversity is present in all sites, there may be very little benefit in planning conservation strategies around the environmental conditions of seed sources.
6.5 Concluding remarks

Collectively, the chapters presented in this thesis provide a substantial contribution to the literature around the impacts of contemporary disturbances on the population demographics and genetics of forest tree species. By utilising mountain ash as a study species, I have (1) examined the likely impacts of contemporary disturbance regimes on maturation rates and stand persistence under different topographic conditions, (2) investigated the consequences of hybridisation and range-wide population genetic structure for conservation of adaptive potential and assisted gene flow, (3) the impacts that wildfire and logging have on population genetic structure and genetic diversity, and (4) the patterns and drivers of local adaptation, and how these vary across different spatial scales. The results of these studies have implications for a variety of conservation practitioners and management agencies, particularly regarding the use of provenancing strategies for restoration and planting, as well as climate change mitigation efforts. Further, the data and conclusions presented in this thesis are relevant to a variety of forest systems around the world, with many forest tree species having similar life-histories, functional roles, and population genetic structuring. Many forest ecosystems are thus being impacted by anthropogenic disturbances in similar ways to those experienced in mountain ash forests, with frequent wildfires, climate change and timber harvesting causing analogous changes to forests globally. Finally, it is my personal and professional hope that any negative impacts of these disturbances on the world’s forests can be both managed and minimised, and that the findings presented in this thesis can provide some of the knowledge needed to assist with such honourable endeavours.
6.6 References


Chapter 6: Synthesis


Wiltshire, R. J. E., 1991. 'Heterochrony and heteroblasty in the Eucalyptus risdonii Hook.f./E. tenuiramis Mq. complex'.


Broadhurst, L. M., Breed, M., Lowe, A., Bragg, J., Catullo, R., Coates, D., Encinas-Viso, F., Gellie, N., James, E., Krauss, S., Potts, B., Rossetto, M., Shepherd, M. and Byrne, M.,


Fagg, P., 2001. *Eucalypt sowing and seedfall, native forest silviculture guideline no. 8*, Department of Natural Resources and Environment, Melbourne, Australia.


https://www.nature.com/articles/hdy2010167#supplementary-information


to conservation planning, *Molecular Ecology*, 17(1): 139-156. Available at: <Go to ISI>://WOS:000251740500012


Hew, H., Hatterjee, M., Ebi, H., Hingston, A. B., Gartrell, B. D. and Pinchbeck, G., 2004a. How specialized is the plant-pollinator association between Eucalyptus globulus ssp. globulus and the swift parrot Lathamus...

Hingston, A. B., Potts, B. M. and McQuillan, P. B., 2004b. Pollination services provided by various size classes of flower visitors to Eucalyptus globulus ssp. globulus (Myrtaceae), *Australian Journal of Botany*, 52(3): 353-369. Available at: 10.1071/BT03002


Australia, 


Lindenmayer, D. B., Incoll, R. D., Cunningham, R. B. and Donnelly, C. F., 1999. Attributes of logs on the floor of Australian Mountain Ash (Eucalyptus regnans) forests of different ages, 

Lindenmayer, D. B. and Laurance, W. F., 2016. The ecology, distribution, conservation and management of large old trees, 

Lindenmayer, D. B. and Sato, C., 2018. Hidden collapse is driven by fire and logging in a socioecological forest ecosystem, 

Liu, Z., Yang, J., Chang, Y., Weisberg, P. J. and He, H. S., 2012. Spatial patterns and drivers of fire occurrence and its future trend under climate change in a boreal forest of Northeast China, 

Loyn, R. H., 1985. Bird populations in successional forests of mountain ash Eucalyptus regnans in Central Victoria, 


Lutze, M. T., Campbell, R. G. and Fagg, P. C., 1999. Development of silviculture in the native state forests of Victoria, 

Mace, B., 1996. Mueller - Champion of Victoria’s Giant trees, 

Maiden, J. H., 1905. A critical revision of the genus Eucalyptus, 


Manel, S. and Holderegger, R., 2013. Ten years of landscape genetics, Trends in Ecology & Evolution, 

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach, 

Martin, A. and Orgogozo, V., 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation, 

http://biostor.org/reference/120887 

http://doi.org/10.1016/j.tree.2013.05.012 


https://doi.org/10.1111/1365-2486.2012.02649.x 

https://doi.org/10.1111/brv.12290 

https://doi.org/10.1073/pnas.1110245108 

https://doi.org/10.1016/j.tree.2013.05.012 

https://doi.org/10.1080/00049158.1999.10674788 

http://biostor.org/reference/120887 

http://doi.org/10.1016/j.tree.2013.05.012 


https://doi.org/10.1111/1365-2486.2012.02649.x 

https://doi.org/10.1111/brv.12290 

https://doi.org/10.1073/pnas.1110245108


https://www.nature.com/articles/nclimate3303#supplementary-information


https://www.nature.com/articles/nature25508#supplementary-information


Eucalyptus globulus, *Tree Genetics & Genomes*, 8(4): 695-708. Available at: <Go to ISI>://WOS:000307291800007


Yu, G., 2018. Package 'scatterpie', *R package version 0.1.2*.


Appendix 1: Testing contemporary genotyping methods in fire-killed eucalypts

A1.1 Abstract

Accurate genotyping of individuals is a fundamental component of conducting genetic studies. While modern genotyping methods typically require high-quality samples, there are a range of scenarios in which only low-quality samples are available. To directly measure dispersal distances using parent-offspring analyses in mountain ash, genotyping of fire-killed parent trees is likely to be required. Here, we attempt to genotype sapwood tissue samples collected from fire-killed adult trees on the premise that a parentage analysis could be undertaken. We utilise two methods of genome complexity reduction followed by next-generation sequencing, and also attempt to amplify nuclear microsatellites. While all of these methods were successful on leaf tissue samples, it proved difficult to obtain any reliable data from the sapwood samples. It is likely that most of the DNA extracted from the fire-killed trees was derived from microorganisms such as fungi or bacteria, and that very little eucalypt DNA was present in the extraction product. We suggest that future studies on this topic utilise either (1) stands with living multi-aged cohorts (for easy DNA extraction and genotyping but with many assumptions) or (2) substantially improve amplification methods for target loci/microsatellites. Either approach should considerably improve our understanding of the dispersal processes in obligate seeding tree species under natural regeneration conditions.
A1.2 Introduction

Accurate genotyping of individuals is a fundamental component of conducting genetic studies, with a multitude of genotyping methods now readily available (Ebert and Peakall, 2009; Elshire et al., 2011; Sansaloni et al., 2010; Sansaloni et al., 2011; De Barba et al., 2017; Baird et al., 2008). Advances in sequencing technology have led to massive increases in the number of genetic markers used in analyses, with modern studies often utilising thousands to hundreds of thousands of single-nucleotide polymorphisms (SNPs) from across the genome (Hudson et al., 2015; Gaughran et al., 2018; Harvey et al., 2017; Hand et al., 2015). Such large genetic datasets can provide high levels of detail into fundamental processes involved in gene flow between populations and species, and even allow researchers to pinpoint loci that may be involved in local adaptation (Ahrens et al., 2018). These sequencing methods are typically performed after collection and DNA extraction from samples containing high quality and quantities of genomic DNA, to minimise missingness in the datasets and reduce risk of contamination and other sources of error. However, there are a range of situations in which high quality samples are not available. For example, the identification of dried and cut timber products is frequently required to confirm its origins and ensure sustainable harvesting of forests, with genetic approaches not only able to identify the species but also the region where it was harvested (Tang et al., 2011; Finkeldey et al., 2010; Rachmayanti et al., 2009). Similarly, low-quality samples may be the only source of DNA available when studying dispersal patterns in serotinous obligate seeders, in which the parent generation is killed by fire prior to dispersal and establishment of offspring plants.

Dispersal is a crucially important component of understanding the population dynamics of a species (Begon et al., 2006). It not only contributes to the distribution of genetic structure among populations, but also determines whether and how fast species can recolonise areas of local extinction, and allows species to modify their range in response to changing environmental factors (Ouborg et al., 1999; Allendorf et al., 2013). For example, climate change is already modifying the distributions of many organisms (Parmesan and Yohe, 2003; Chen et al., 2011), and the seed dispersal distances of plants will strongly influence the rate at which they can spatially track climates. Where a specialist species is dependent on a foundation species (e.g. a dominant overstorey tree), the rate at which the foundation species migrates will directly affect the ability of the specialist species to shift its range (Schweiger et al., 2008; Gilman et al., 2010). With the rise of genetic approaches to ecological investigations came numerous techniques for determining gene flow in plant populations (Hamrick and Trapnell, 2011; Leblois et al., 2003; Vekemans and Hardy, 2004).

Here, we trial three methods of genotyping fire-killed mountain ash (Eucalyptus regnans) trees, on the premise that this would allow us to directly measure dispersal distances
under natural conditions using parent-offspring analyses. Parent-offspring analyses utilise genetic relatedness metrics to identify the most likely pair of parents for established offspring plants, using biparentally-inherited molecular genetic markers such as SNPs or microsatellites. In hermaphroditic species this method does not provide information on which parent is maternal or paternal; however, as pollination often occurs over longer distances than seed dispersal, the nearer parent is assumed to be the source of the seed (Dow and Ashley, 1996). Despite this and other assumptions, parent-offspring analyses are able to provide considerable insight into dispersal processes under natural conditions and so are an important step towards understanding ecological processes (Johnson et al., 2017). To determine whether this approach is possible for mountain ash, we utilised two methods of genome complexity reduction followed by next-generation sequencing, and one target amplification method whereby nuclear microsatellites are subjected to polymerase chain reaction and then sized. As the parent trees are dead, we predict that the quantity of DNA present in the samples will be low and/or degraded and that this could prevent accurate genotyping using next-generation sequencing methods. We further suggest that as a targeted PCR approach, microsatellite amplification and sizing will be the most viable way of genotyping mountain ash, to allow for estimation of dispersal distances using parent-offspring analyses.
Figure A1.0.1 Dispersal distance of seed falling from mountain ash (*Eucalyptus regnans*) under atypical regeneration scenarios. The y-axis is standardised between studies by recording the amount of seed directly underneath the canopy as 100%. Data collated from multiple sources (Cremer, 1966; Drangsholt, 1956; Cunningham, 1960).
A1.3 Materials and methods

A.1.3.1 Study area

We visited a subset of the Australian National University’s long-term monitoring sites in the Central Highlands of Victoria, which have been studied for over two decades (Lindenmayer et al., 2015). This consisted of four sites that were subjected to high-severity fire in the 2009 ‘Black Saturday’ wildfires. Prior to the 2009 fires, three of these sites (Marysville 1, Marysville 2, and Toolangi) had been predominantly composed of mountain ash that regenerated after the 1939 fire, whereas one site (Cambarville) was dominated by mountain ash that had regenerated after a fire in the early 18th century.

A.1.3.2 Study species

In Australia, eucalypts are often foundation species in natural vegetation communities, particularly in the continent’s south-eastern and south-western forests. Genetic exchange in this group is the result of both pollen flow and seed dispersal, with most authors suggesting that pollen flow plays a greater role because it is likely to occur over larger distances than seed is expected to travel (Barber, 1965; Potts, 1990; Potts and Wiltshire, 1997). This generalisation is typically supported by modern genetic analyses (Sampson et al., 2018), despite the tail end of the pollen dispersal curve being very difficult to quantify (Jones et al., 2008).

Mountain ash, a commercially important species and one of the tallest trees in the world, is a smooth-barked eucalypt that grows in the montane wet forests of the Australian states of Victoria and Tasmania. It is also a serotinous obligate seeder, meaning that it is typically killed by fire and regenerates only from seed stored in the canopy of the trees, and thus often forms stands composed of a single age cohort. Seed dispersal from mountain ash is typically considered as being very limited (Figure A1.0.1), up to the equivalent of about one tree height (Gill, 1994; Lamb and Smyth, 2006). Similarly, Grose (1960) states that seed dispersal of the ecologically similar alpine ash may, with strong winds, extend up to about 121 m from the maternal parent, although he considers such events unimportant for adequately restocking disturbed areas. Cremer (1966) further extends this concept, and notes that seed dispersal distances depend on multiple factors, such as stand density, wind speed and direction, and topography of the landscape, and suggests that under certain conditions seeds could be distributed up to 150 m from the maternal parent. These distances are comparable to other tree species with no specialised dispersal mechanisms (Thomson et al., 2011).

Studies of seed dispersal in mountain ash are limited by a two main factors, including (1) basing estimates on annual seed fall rather than during regeneration events, and (2)
measuring seedfall from the edges of patches or isolated trees rather than within the forest itself. Importantly, (Potts and Wiltshire, 1997) suggest that seed dispersal kernels could be different during regeneration events, which may involve strong winds or convectional fire updraughts (Kirkpatrick, 1977). After a moderate- to high-intensity fire in mountain ash forest, the ability of wind to penetrate the stand is likely to be far greater due to the absence of midstorey and understorey species, and (potentially) overstorey foliage. The dispersal distance in this situation will, ceteris paribus, be greater than observed from the leeward edge of an unburnt stand because the wind speed is typically much higher.

**A.1.3.3 Sample collection**

In many parent-offspring studies, there is an assumption that all potential parents can be found and genotyped, which may be very unlikely for some species. Whilst the density of trees prevents this assumption from being met in mountain ash forests, we sampled and attempted to genotype as many saplings and surrounding parent trees as possible without inflating costs to prohibitive levels (Johnson et al., 2017). At each site we walked eight parallel transects, approximately 15 metres apart, each 110 m long, to produce an area surveyed of approximately 120 m by 110 m. Along each transect, we collected a sample of mountain ash tissue every 10 m, alternating between fire-killed (but still standing) mature adult trees and immature saplings. This produced 96 samples per site, of which 44 were adults and 44 were saplings.

As the adult trees were dead and had no canopy foliage, we used a sharp, clean axe (wiped down with 80% ethanol before each use) to access the sapwood of the trees. We then placed approximately 2 g of material into a labelled paper bag, and stored the sample at room temperature prior to DNA extraction. A battery-powered drill was initially tested but did not have the power required to drill into the very hard wood, and the heat generated during the drilling may have further degraded any DNA present (Rachmayanti et al., 2009). For sapling trees, we collected two fresh, growing leaves from each individual, which were placed in a labelled paper bag and dried in the field using silica gel.

The density of adult trees at each site was also calculated, by counting the number of fire-killed adult trees to 20 m on either side of a 180 m transect and extrapolating outwards to determine the number of trees per hectare prior to the 2009 fires.

**A.1.3.4 DNA extraction**

All samples (n=384) were placed into 20 mL scintillation vials and stored at -80°C for at least 12 hours before mechanical disruption. Disruption and homogenisation were carried out using an automated cryogenic (-65°C) plant grinder and dispensing system (Labman
Appendix 1: Genotyping fire-killed eucalypts

Automation Ltd). Despite being broken into small fragments by hand prior to grinding, sapwood samples needed to be subjected to a longer grind than leaf samples due to their hardness.

Leaf samples (n=192) had total genomic DNA extracted from 20 mg of powdered material using DNeasy Plant Mini Kits (Qiagen, Hilden, Germany) by following the manufacturer’s protocol, although we extended the lysis incubation to one hour to improve DNA yield.

Sapwood samples proved more difficult to work with in the lab, and considerable efforts were made to increase DNA yield during the extraction process. Initial attempts at extraction using standard protocol of the DNeasy Plant Mini Kits resulted in negligible quantities of DNA (quantified using a Qubit Fluorometer). Extending the lysis period (5 hrs, 12 hrs, or 15 hrs) did not noticeably improve the yield. Using a modified protocol lysis buffer that included 2.6% (w/v) polyvinylpyrrolidone (PVP) (Rachmayanti et al., 2006; Rachmayanti et al., 2009) also did not improve yield. As the addition of N-phenacylthiazolium bromide (PTB) has been shown to improve DNA yield (Tang et al., 2011), we next tried incubating samples in a 1.4 mL lysis buffer composed of cetyltrimethylammonium bromide (CTAB), PTB and EDTA. This markedly improved the DNA yield, and we proceeded to extract all samples using a protocol modified from Tang et al. (2011). Approximately 110 mg of powdered tissue was subjected to a seven-hour lysis incubation at 65°C, where the lysis buffer included 700 µL 0.1 M PTB (Prime Organics, Woburn, MA, USA), 900 µL CTAB buffer (2% CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris–HCl, pH 8.0), and 100 µL 0.5 M EDTA. Following lysis, the solution was centrifuged for 10 minutes at 20,000 g, and the lysate placed into a new 2 mL microcentrifuge tube. Approximately 0.325 volumes of buffer P3 was added, with incubation on ice for five minutes. DNA was then purified following the standard protocol of the DNeasy Plant Mini Kits, beginning with step five. After the final extraction step, samples were quantified using a Qubit Fluorometer.

A.1.3.5 Genotyping-by-sequencing

All leaf samples and three sapwood samples were prepared for genotyping-by-sequencing (GBS). The majority of sapwood samples were not included because it was considered unlikely that the extracted DNA would be of high enough quality or quantity for successful preparation and sequencing of the samples, and including all samples would require re-extraction from homogenised tissue for other genotyping methods.

Library preparation included (1) digestion using a high-fidelity restriction enzyme (PstI, New England BioLabs Inc.), (2) ligation using T4 DNA ligase (New England BioLabs Inc.), (3) purification using a Qiagen MinElute 96-well PCR purification kit, (4) PCR amplification
with GBS primers (Integrated DNA Technologies), (5) quantitation using microfluidic capillary electrophoresis (PerkinElmer Labchip GX-II), (6) pooling of samples using an automated robotic liquid handling machine (PerkinElmer NGS Express), and (7) purification using a Genelute PCR Clean-Up Kit (Sigma-Aldrich).

To determine the success of the library preparation, the electrophoresis results from the quantitation step were first visually inspected for each sample, to ensure digestion resulted in a smooth smear of fragment sizes. Next, the concentration of DNA in the fragment size range of 250-450 was calculated. This size range was used as it is typically where the highest DNA concentration occurs and has proven to be appropriate for GBS of eucalypts (Nicola Aitken, pers. comm. 2016).

A.1.3.6 Diversity Arrays sequencing

A selection of 16 samples, consisting of eight sapwood and eight leaf samples, were submitted for sequencing at the Diversity Arrays Technology Pty Ltd (DArTSeq) laboratory (Sansaloni et al., 2011), to trial whether this sequencing method would work on sapwood samples well enough to identify large numbers of SNPs for use in parentage assignment of mountain ash. Sample preparation involved normalisation (as much as was possible) of DNA concentrations across samples prior to submission. The mean amount of DNA per sample submitted was 379 ng (range 206-515 ng) for leaf tissue and 126 ng (range 0.85-396 ng) for sapwood tissue.

All subsequent sample preparation and sequencing was undertaken at the DArTSeq laboratory using their proprietary genome complexity reduction and optimised barcoded next-generation sequencing protocols. After restriction enzyme digestion, adapter ligation and PCR amplification, samples were run on a single lane of an Illumina HiSeq2500, at a sequencing density of about 1.25 million reads per sample. Read assembly, quality control and SNP calling were conducted using DArTSeq proprietary analytical pipelines, which include filtering of poor-quality sequences, stringent selection criteria for identifying reads to individuals, and a robust SNP calling algorithm.

A.1.3.7 Microsatellite amplification

As we were uncertain whether the above two methods would be successful, we tried to amplify selected fragments of nuclear DNA in order to reduce the likelihood of microorganisms such as bacteria and fungi influencing the data. There have been numerous studies (Glaubitz et al., 2001; Steane et al., 2001; Ottewell et al., 2005; Brondani et al., 2006; Nevill et al., 2008) that have identified and/or utilised microsatellites from chloroplast and nuclear DNA of eucalypts. Of these, Nevill et al. (2008) identifies 10 nuclear microsatellites
that are easily amplified and produce readily scorable bands in mountain ash. We tested primer pairs for these ten microsatellites (GeneWorks Pty Ltd, Thebarton, Australia) on seven leaf and five sapwood samples, to determine whether amplification of targeted nuclear DNA sites would be of possible use in parentage assignment. Sapwood samples were replicated three times each to ensure reproducibility.

PCRs were undertaken in 10 µL reaction volumes, consisting of 5 µL of HotStarTaq Master Mix (Qiagen), 0.2 µL forward primer (2 µM, with M13 tail), 0.2 µL reverse primer (2 µM), 2 µL DNA, 2.1 µL ultra-pure (Milli-Q) water, and 0.5 µL M13 primer (with fluorescent dye label, i.e. 6-FAM, NED, VIC or PET). PCRs were carried out in an Eppendorf Mastercycler Nexus, using the following reaction conditions: 15 minutes at 95°C for initial denaturation and activation of Taq in the master mix, 39 cycles of 30 s at 94°C, 90 s at 59°C, and 30 s at 72°C, followed by an extension of 60°C for 30 m.

Following PCR, six microsatellites for two leaf and two sapwood samples were run on a 2% agarose gel and imaged using a Vilber E-Box VX2 to indicate whether amplification was likely to have been successful or not. Amplified products from all 12 samples were then run on an ABI 3130xl Genetic Analyser (Applied Biosystems) to determine fragment sizes, with peaks called using Geneious v6.1.8 (Kearse et al., 2012).
A1.4 Results

A.1.4.1 Stem density

Mean adult tree density at the three sites (Marysville 1, Marysville 2, and Toolangi) that regenerated in 1939 was 109.5 per ha. As there is a substantial degree of natural stand-thinning over time in mountain ash forests, and adult trees at Cambarville had germinated over 150 years prior to the other sites, Cambarville had a much lower density of adult trees, with 45.5 per ha (Figure 1.2).

A.1.4.2 DNA extraction

The mean concentration of extracted genomic DNA for leaf samples was 5.75 ng/µL, compared to just 0.93 ng/µL for sapwood samples, resulting in approximately 1122 ng DNA for leaf samples and 185 ng DNA for wood samples. Twenty-three (13%) of sapwood samples had DNA concentrations after extraction of less than 0.005 ng/µL, indicating that the extractions had failed to produce meaningful quantities of any DNA.

Despite extractions from sapwood samples producing lower quantities of DNA, the quantity produced was considered reasonable for two downstream genotyping methods; DArTSeq and targeted amplification of microsatellites. To test whether genotyping-by-sequencing would be viable, three wood samples were prepared for sequencing using this method.

A.1.4.3 Genotyping-by-sequencing

Library preparation for the vast majority of leaf samples was successful, with 12 ng of DNA for almost all samples able to be pooled in preparation for gel cutout and sequencing. As expected, the library preparation for the three sapwood samples appeared to fail, with visual inspection of the quantification results indicating poor digestion and amplification of the same DNA fragments as for the leaf samples. In addition, calculation of concentration at the size range of 250-450 bp showed that there was a mean of 0 ng/µL in the sapwood samples, compared with a mean of 2.2 ng/µL in the leaf samples (Figure A1.0.2). The reasons for failure in the sapwood samples are uncertain, but probably relate to the quantity and quality of mountain ash DNA present in the extraction product. It is possible that the extraction product may have been largely composed of DNA from microorganisms, further reducing the concentration of mountain ash DNA.
Figure A1.0.2 Gel electrophoresis image showing DNA concentration of six mountain ash (*Eucalyptus regnans*) samples after library preparation for genotyping-by-sequencing. Darker colouration represents higher DNA concentration. Fragment lengths in base pairs (bp) are indicated on the left- and right-hand sides of the figure. The first three columns represent sapwood tissue samples, and show very little to no DNA in the fragment size range of 250-450 bp. The last three columns represent randomly selected leaf tissue samples, with moderate to high concentrations of DNA in the fragment size range of 250-450 bp.
A.1.4.4 Diversity Arrays sequencing

Library preparation at the DArTSeq facility appeared to be successful, with quality control steps indicating no clear differences between the sapwood and the leaf samples. All of the leaf samples sequenced well with large numbers of reads and high reproducibility scores, and three of the eight sapwood samples failed to sequence well (i.e. no useful amount of data was obtained). The failures did not appear to be related to the concentration of total genomic DNA in the extracted product. For samples that sequenced well, acceptable numbers of reads for both types of tissue material were produced, with most sapwood samples having hundreds of thousands to more than one million reads. However, preliminary analysis of pairwise genetic distances between samples showed that the sapwood samples were very distantly related to leaf samples (mean genetic distance 0.71). Principal coordinates analysis of the genetic distance matrix clearly shows that the sapwood samples were strongly differentiated from the leaf samples, with 87.5% of the variance explained by the first coordinate (Figure A1.0.3). Technical staff from DArTSeq suggested this indicated that the extracted DNA from the sapwood samples was derived primarily from microorganisms such as bacteria or fungi that had colonised the dead trees. As we were unlikely to be able to improve the extraction process to reduce these contaminant sources, we did not continue with this approach.
Figure A1.0.3 Principal coordinates analysis of pairwise genetic distances between mountain ash (*Eucalyptus regnans*) tissue samples after reduced representation sequencing of whole-genomic DNA. Circles represent sapwood tissue collected from dead standing trees, and triangles represent leaf tissue collected from living trees.
A.1.4.5 Microsatellite genotyping

Gel electrophoresis indicated that amplification of nuclear microsatellites was mostly successful for the leaf samples, with five out of six microsatellites showing clear, bright bands at appropriate sizes, with little sign of primer dimer (Figure A1.0.4). Unfortunately, the gel electrophoresis suggested that the microsatellites had largely failed to amplify in the wood samples, with most wells showing either no bands at all or strong stuttering at various fragment sizes, and a noticeable level of primer dimer.

After running all samples on the sequencer, peaks were called in Geneious. Peak-calling in the leaf samples was easily done, with little ambiguity of peak locations, and biologically realistic size differences between samples. For the sapwood samples, peak calling was not possible in many cases, and, where peaks were present, they were at drastically different sizes between samples, suggesting stochastic amplification of non-target regions. Repeating the analysis using 10x dilutions of the extraction product prior to amplification had no impact on the results. Thus, despite there being reasonable quantities of DNA in the extraction product from sapwood samples, no practically useful quantity of mountain ash nuclear DNA was deemed to be present.
Figure A1.0.4 Gel electrophoresis showing results of nuclear microsatellite amplification in mountain ash (Eucalyptus regnans). Each row starts with a ladder indicating fragment size, followed by six microsatellites, and then a second ladder followed by the same six microsatellites in a different sample. The top row of the gel shows the successful amplification of five out of six microsatellites in samples extracted from leaf tissue. The bottom row shows the largely unsuccessful amplification of the same microsatellites in samples extracted from sapwood tissue.
A1.5 Discussion

The dispersal kernel of obligate seeding eucalypts under natural conditions is poorly known, as the conditions during wildfire events are likely to cause different dispersal patterns to those created by annual seed rain. To conduct parentage assignments, for directly measuring dispersal patterns in these species under natural conditions, genotyping of fire-killed adult trees (or pre-fire knowledge of genotypes) is required, because successful regeneration requires moderate- to high-intensity fires that kill mature trees. To determine whether genotyping fire-killed eucalypts was possible, we sampled sapwood tissue of trees killed in the 2009 Black Saturday wildfires, and used three methods to try and genotype individuals for use in a parentage assignment study.

After extracting whole-genomic DNA and testing three approaches of genotyping samples, we were unable to successfully sequence mountain ash DNA fragments nor amplify and size nuclear microsatellite loci from sapwood tissue. We suggest that our inability to obtain mountain ash DNA from dead standing trees is likely a combination of two things, (1) the naturally low level of living cells (e.g. ray parenchyma) in sapwood, and (2) gradual degradation of DNA in the years since tree death, as a result of exposure to environmental conditions and microorganism activity (Deguilloux et al., 2002). While our extraction method was considerably modified from standard procedures, the quantity of mountain ash DNA could probably be increased by further increasing the amount of sample material (and overall extraction volume). We suggest that unless substantial improvements are able to be made with regards to DNA extraction or genotyping methodology, it is unlikely that parent-offspring analyses can be conducted using this method.

Contamination by foreign DNA has previously been shown to be a common occurrence when attempting to genotype wood samples (Deguilloux et al., 2002). While large extraction volumes may produce enough eucalypt DNA for microsatellite amplification and sizing, or even next-generation genotyping, the presence of large amounts of contaminant DNA sources would still need to be addressed. Aligning reads to a reference genome (Myburg et al., 2014; Bartholome et al., 2015) and using only SNPs identified from leaf tissue samples may be one way around this issue. Amplification of microsatellites may be a more plausible method of genotyping dead standing trees, as this avoids the issue of requiring large quantities of DNA. Laboratory methods are continuously being improved, and recent advancements in environmental DNA approaches could be beneficial here, for example by using multiple PCR replicates (Piggott, 2016). If genotyping of parent individuals could be consistently achieved using microsatellites, there is a good chance that useful data on dispersal distances could be obtained from our sampling strategy. For example, Jones et al. (2008) used eight highly variable nuclear microsatellite loci in flooded gum (Eucalyptus grandis) to assign 34.7% of
seeds collected from maternal parent trees to a single pollen parent and another 19.5% to two possible pollen parents with reasonable (>80%) certainty.

Another way to improve genotyping success when dealing with low-quality samples is to utilise a technique such as hybridisation capture of restriction site-associated DNA (hyRAD) (Suchan et al., 2016; Linck et al., 2017). Briefly, this method involves using the size-selected library fragments, created in the preparation of high-quality leaf samples for GBS, as the basis for a set of random genomic probes that can be used to capture loci in other samples. Once the loci have been captured, they are enriched and indexed for pooled sequencing. This could be a potential next step for our samples, as the extractions have already been undertaken.

There are some remaining areas of mountain ash forest that contain multiple age cohorts, and this is even considered a typical structural characteristic of old-growth forest (Lindenmayer et al., 2000). These stands would make it easy to identify dispersal distances because the parent trees are still alive and tree cohorts are easily identified by their stem diameters. However, the issue with this approach is that the fires that produced these multiple age cohorts are often of lower intensities or smaller spatial scales than the large wildfires that are responsible for much of the regeneration events that occur across the landscape. Further, while they are not necessarily stand-replacing, they are likely to be partially stand-replacing (Lindenmayer et al., 2000), and the degree to which adult trees have been removed from the stand due to one or more fires is unknown. While such factors make this approach less than ideal, it would still provide insight into the dispersal kernel of the species under natural conditions and would make a useful comparison for future studies.

For some plant species it is possible to match the offspring directly to the maternal parent, with such maternal-matching analyses generally utilising the maternally-inherited seed coat to obtain maternal DNA (Godoy and Jordano, 2001; Jordano et al., 2007). This is very difficult to achieve for mountain ash, because obtaining seeds under natural regeneration conditions requires seed collection in areas very recently burnt at high intensity, by infrequent and unpredictable wildfires. Another possibility for maternal-matching would be via analysis of chloroplast DNA (cpDNA), as cpDNA is uniparentally and maternally-inherited in Eucalyptus (Byrne and Moran, 1994; McKinnon et al., 2001). However, it may not be possible to reliably identify individuals using cpDNA due to the highly conserved nature of the chloroplast genome, so this method would require likely require discovering new microsatellites and testing their reliability and variability in mountain ash.
A1.6 Conclusion

In this study we attempted to genotype parent and offspring mountain ash individuals in order to genetically-determine dispersal distances. As parent trees were killed by intense wildfires in 2009, this required successful extraction and genotyping of dead standing trees. We found that sapwood of parent trees contained very little eucalypt DNA, with next-generation sequencing suggesting that the extracted DNA was largely derived from microorganisms such as fungi or bacteria. We suggest that future studies on this topic utilise either (1) stands with living multi-aged cohorts (easy DNA extraction and genotyping but with many assumptions) or (2) substantially improve amplification methods for target loci/microsatellites. Either approach should considerably improve our understanding of the dispersal processes in obligate seeding tree species under natural regeneration conditions.
A1.7 Acknowledgements

The authors wish to thank Oliver King for his assistance with fieldwork. BvTD was supported by a Lesslie Research Grant in Landscape Conservation and Ecology and an Ignition Project grant from the Centre for Biodiversity Analysis. SB was supported by an Australian Research Council Future Fellowship (FT130100043).
**A1.8 References**


Cunningham, T. M., 1960. The natural regeneration of *Eucalyptus regnans*, *University of Melbourne School of Forestry Bulletin, 1*.


Sansaloni, C., Petoli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D. and Kilian, A., 2011. Diversity Arrays Technology (DArT) and next-generation sequencing combined:

Schweiger, O., Settele, J., Kudrna, O., Klotz, S. and Kühn, I., 2008. Climate change can cause spatial mismatch of trophically interacting species, Ecology, 89(12): 3472-3479. Available at: 10.1890/07-1748.1


