“Eucalyptus is not only the Universal Australian, it is the Ideal Australian - versatile, tough, sardonic, contrary, self-mocking, with a deceptive complexity amid the appearance of massive homogeneity”

(Stephen J. Pyne *The Burning Bush* 1992, pp. 22)

This natural F1 hybrid between *Eucalyptus stoatei* and *E. tetraptera* was grown from seed collected in the Jerdacuttup area, south east of Ravensthorpe in Western Australia by Dean Nicolle. The photograph was taken at Dean’s Currency Creek Arboretum, in South Australia.
Exotic gene flow from plantation to native eucalypts

Matthew James Larcombe

B.Sc. (Hons)

Submitted in fulfilment of the requirements for the degree of Doctorate of Philosophy

School of Biological Sciences, University of Tasmania

June 2014
Declarations

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

The publishers of the papers comprising Chapters 3 to 5 hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining non published content of the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Signature: Date:

The Candidate Matthew James Larcombe

We the undersigned agree with the above declarations, and submission of the final corrected version of this thesis:

Brad M. Potts Anthony Koutoulis
Supervisor Head of School
School of Biological Sciences School of Biological Sciences
University of Tasmania University of Tasmania

Signed: ________________ Signed: ________________
Date: ________________ Date: ________________

René E. Vaillancourt
Supervisor
School of Biological Sciences
University of Tasmania

Signed: ________________
Date: ________________
Statement of co-authorship/contribution

The following people contributed to the publication of work undertaken as part of this thesis:

Matthew J. Larcombe, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia = **Candidate**

Brad M. Potts, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia = **Author 1**

René E. Vaillancourt, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia = **Author 2**

Rebecca C. Jones, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia = **Author 3**

Robert C. Barbour, School of Plant Science, University of Tasmania, Hobart, Australia = **Author 4**

Joaquim S. Silva, Centre for Applied Ecology, Technical University of Lisbon, Portugal = **Author 5**

Below are the details of which papers the co-authors contributed to, the contributions they made and where the papers can be found in the thesis:

Chapter 3 is published:


Candidate was the primary author, undertook all analysis, field and lab work. The Candidate as well as authors 1, 2 and 3 contributed to developing the idea, and approach. Authors 1, 2 and 3 assisted with refining the text.
Chapter 4 is published:


Candidate was the primary author, undertook most field, glasshouse, lab work and analysis. Author 4 established the paired hybrid trial in 2006, author 1 assisted with analysis. The candidate as well as author 1, and 2 contributed to developing the main idea, and approach. The paired hybrid trial was the idea of Author 4. Author’s 1 and 2 assisted with refining the text.

Chapter 5 is published:


Candidate was the primary author and was involved in all field work. Field work undertaken in Victoria was done in partnership with author 5. The candidate undertook most analysis with assistance from authors 1 and 5. The candidate as well as author 1, and 5 developed the main ideas, and approach. Author’s 1, and 2 and 5 assisted with refining the text.

The following people made contributions to the unpublished Chapter 2:

The candidate was the lead author, undertook some analysis, all field work, glasshouse work and project specific lab work (with the exception of the generating the DArT data see below).

Barbara R. Holland, School of Maths and Physics, University of Tasmania. Barbara gave statistical advice and assisted with analysis, particularly the computer code (R code) used to undertake the modelling.

Rebecca C. Jones, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia – provided the species-level DArT dataset and commented on the text.

Dorothy A. Steane, School of Plant Science, National Centre for Future Forest Industries, University of Tasmania, Hobart, Australia – provided the section-level DArT dataset and commented on the text.
Dean Nicolle, Currency Creek Arboretum, South Australia – provided access to the Arboretum where most of the crossing was undertaken. He also contributed his unpublished classification of *Eucalyptus*, and commented on the text.

We the undersigned agree with the above stated “proportion of work undertaken” for the published chapters 3, 4 and 5, and the unpublished Chapter 2 of this thesis:

<table>
<thead>
<tr>
<th>Brad M. Potts</th>
<th>Anthony Koutoulis</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Supervisor</em></td>
<td><em>Head of School</em></td>
</tr>
<tr>
<td>School of Biological Sciences</td>
<td>School of Biological Sciences</td>
</tr>
<tr>
<td>University of Tasmania</td>
<td>University of Tasmania</td>
</tr>
</tbody>
</table>

Signed: __________________  
Signed: __________________

Date:  ________________  
Date:  ________________

<table>
<thead>
<tr>
<th>René E. Vaillancourt</th>
<th>Matthew James Larcombe</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Supervisor</em></td>
<td><em>Candidate</em></td>
</tr>
<tr>
<td>School of Biological Sciences</td>
<td>School of Biological Sciences</td>
</tr>
<tr>
<td>University of Tasmania</td>
<td>University of Tasmania</td>
</tr>
</tbody>
</table>

Signed: __________________  
Signed: __________________

Date:  ________________  
Date:  ________________
Abstract

The movement of species around the world by humans has created situations where “exotic” gene flow can arise between species that would not naturally co-occur. *Eucalyptus globulus* has been planted widely throughout temperate Australia over the past 15 years, with around 538,000 ha of plantations now growing, mainly outside the species native range. Concerns have been raised that these plantations could genetically contaminate natural eucalypt populations. This thesis aimed to assess the risk, and management, of pollen- and seed-mediated gene flow from *E. globulus* plantations.

The thesis initially addresses the risk of introgression through pollen-mediated gene flow from *E. globulus* plantations and hybridisation with co-occurring native species. Prior to this study there were no known complete barriers to hybridisation between *E. globulus* (as the pollen parent) and other species in subgenus Symphyomyrtus. This meant that as many as 484 species could have been considered at risk of hybridisation if they occurred within the pollen dispersal zone of *E. globulus* plantations. A controlled crossing program (where *E. globulus* pollen was applied to the stigma of 100 other eucalypts species) was undertaken to identify phylogenetically controlled barriers to hybridisation in subgenus Symphyomyrtus. This crossing suggests the presence of a complete barrier to hybridisation between *E. globulus* and more divergent groups within Symphyomyrtus, probably reducing the number of at-risk species by over 70% (to 138). Hybridisation success declined with increasing genetic distance, meaning the most at risk species were those within the same taxonomic section as *E. globulus*, Maidenaria (68 species). The results also provided new insights into the evolution of reproductive barriers in forest trees.

Because hybrid identification is vital for management of exotic gene flow and can be difficult in eucalypts, a Bayesian modelling approach to detect hybrids in at-risk species was tested. Range-wide samples from five at risk species, as well as samples from *E. globulus* (total n = 606 individuals) were genotyped at 10 microsatellite loci. The ability of Bayesian clustering to identify hybrids using this database was tested
using simulations. The technique was highly effective at identifying $F_1$ hybrids, which are currently the primary concern in the Australian $E. \text{globulus}$ estate.

The crossing study showed that species in section Maidenaria should be the focus of management attention. The frequent proximity of $E. \text{ovata}$ (Maidenaria) to plantations and its known cross-compatibility with $E. \text{globulus}$ makes it a prime candidate for exotic gene flow. However, by conducting a case study in $E. \text{ovata}$ forests around plantations, the actual risk posed was found to be low. Hybridisation was assessed in 24,322 open pollinated progeny from 142 trees in 25 native forest remnants. Although patch size and tree position affected hybridisation risk (small patches and edge trees were at highest risk), the rate of hybridisation declined very rapidly inside $E. \text{ovata}$ patches, and hybrid establishment along native forest-plantation boundaries was low. Furthermore, hybrids showed a 78% reduction in survival compared to pure $E. \text{ovata}$ after six years, making it unlikely that hybrids will reach reproductive maturity to enable backcrossing and subsequent introgression. However this study showed that pure $E. \text{globulus}$ seedlings (wildings) were establishing in far higher numbers than hybrids at the edge of plantations, raising the concern that they could pose a threat to native forests.

As well as having ecological impacts as locally exotic species, wildlings could cause introgression via hybridisation if they reach reproductive maturity. To assess the risk that wildings pose to native forests in Australia, surveys to quantify current levels of establishment were undertaken along 290 km of $E. \text{globulus}$ plantation edges. Wildling establishment was low with the vast majority occurring within the plantation disturbance zone. It also appears that current management practices, including short rotations and firebreak maintenance, are reducing the risk of wildling spread.

In conclusion this thesis has found that there are significant barriers to hybridisation between $E. \text{globulus}$ and native eucalypts that will limit the opportunity for exotic gene flow. If these barriers are overcome, avenues for management exist. While wildling establishment appears to currently be limited, the Australian plantation estate is young and on-going monitoring is warranted.
Acknowledgements

Firstly I would like to thank my supervisors Brad Potts and René Vaillancourt for their exceptional guidance and support – I really can’t imagine better supervisors, it would have been harder and less rewarding without them.

Secondly I would like to thank the people, companies and organisations who helped by providing time, resources and access to the properties to make the various projects possible; Dean Nicolle for developing such an amazing resource at Currency Creek Arboretum and providing such willing access and help; Hancock Victoria Plantations, particularly Ross Gillies, Stephen Elms, Justine Krieger, Mark Reynolds and Phil Whitman; Australian Bluegum Plantations, particularly Ben Bradshaw; Elders Forestry, particularly Nigel England; WA Plantation Resources, particularly Matthew Kovacs and Sara Mathieson; Forestry Tasmania, particularly Stephen Read and Tim Wardlaw; Arianda Pty Ltd. particularly Martyn Lavery; Seed Energy, particularly Peter Gore; Gunns Ltd; the Australian Bureau of Agricultural and Resource Economics and Sciences (Canberra), particularly Mijo Gavran, for information and providing GIS data. For providing DNA and or tissues samples for DNA extraction I thank Simon Southerton, Dean Williams, Sandra Hetherington, Kelsey Joyce, James Marthick, Susan Foster, Jules Freeman, Dorothy Steane, and Corey Hudson.

For financial support I would like to thank Forests and Wood Products Australia, the Australian Research Council, the CRC for Forestry, and the National Centre for Future Forest Industries.

For helping out along the way I would like to thank Corey Hudson, James Worth, Beck Jones, Joaquim Sande Silva, Dorothy Steane, Sascha Wise, Adam Smolenski, Grant Williamson, Paul Tilyard, Winston Smith, Guy and Simone Roussel, Catherine Jones, Clancy Carver, Jodi Noble, Mark Hovenden, Michèle Lang, Michael Oates, and Justin Bloomfield, as well as everyone else in the Euc Group.

For moral support I would like to thank Matilde Ravizza, Penny Ravizza, Con Larcombe, Katie Larcombe, Jan Larcombe my office mates and flat mates.

Finally, special thanks must go to Mum and Dad who have helped support me through my many years at school, and gave me the opportunity to do this.
List of publications and presentations arising from this thesis

Refereed journal articles:


Refereed Conference Proceedings:


Presentations:

Invited Presentations:


Other presentations:

Larcombe MJ, Steane D, Jones RC, Nicolle D, Holland B, Vaillancourt RE, Potts BM. Phylogenetic patterns of reproductive isolation in *Eucalyptus*. Presentation to Phylomania 2013, University of Tasmania, School of Maths and Physics, 6-8 November 2013.


Posters:


Format of thesis chapters

The experimental chapters of this thesis (chapters 2 – 5) have been written in paper style and chapters 3 – 5 are published. Being paper style means that some repetition of concepts and ideas was unavoidable, particularly in the introduction sections. The published papers have been reformatted (including figure and table numbering), and the acknowledgments and references have been consolidated into single sections in the thesis. Chapter 1, the general introduction, includes text that was published as a peer reviewed conference proceedings and the full paper is provided in the General Appendix (with all publications arising from the thesis) in the back of the thesis.
# Table of Contents

Chapter 1: General introduction: exotic gene flow and Australian eucalypt plantations 1

1.1 Gene flow .................................................................................................................. 1

1.2 Exotic gene flow ......................................................................................................... 4

1.2.1 Genetically modified crops ................................................................................. 4

1.2.2 Large scale biological releases ........................................................................ 5

1.3 Exotic gene flow from eucalypt plantations in Australia ............................. 8

1.3.1 Barriers to exotic gene flow provide a research and management framework ........................................................................................................... 9

1.3.2 Pre-mating barriers – before pollen reaches the flower ......................... 10

1.3.3 Post-mating barriers – after pollen reaches the flower ............................ 13

1.3.4 Risk management ................................................................................................. 16

1.3.5 A case study: E. perriniana Strickland, central Tasmania .................... 18

1.4 Exotic gene flow from *Eucalyptus globulus* in Australia ....................... 21

1.4.1 Key difference between Eucalyptus globulus and E. nitens .................. 21

1.4.2 Seed mediated gene flow from *E. globulus* – an emerging issue? .......... 22

1.5 Thesis outline .............................................................................................................. 23

1.5.1 Chapter 2: Phylogenetic patterns of reproductive isolation in *Eucalyptus* ... 23

1.5.2 Chapter 3: A microsatellite database and Bayesian modelling approach for identifying hybrids between plantation and native eucalypts. ......................... 24

1.5.3 Chapter 4: Assessing the risk of exotic gene flow from *Eucalyptus globulus* plantations to native *E. ovata* forests ................................................................. 24

1.5.4 Chapter 5: Assessing the invasive potential of *Eucalyptus globulus* in Australia: quantification of wildling establishment from plantations ......................... 24

1.5.5 Chapter 6: General discussion ............................................................................. 25
Chapter 2: Phylogenetic patterns of reproductive isolation in *Eucalyptus* 26

2.1 Abstract .................................................................................................................. 26

2.2 Introduction ........................................................................................................... 27

2.3 Materials and Methods ....................................................................................... 31

2.3.1 Crossing ............................................................................................................ 31

2.3.2 Phylogenetic and statistical methods ................................................................. 34

2.4 Discussion ............................................................................................................ 43

2.4.1 The evolution of post mating barriers ................................................................. 43

2.4.2 Hybridisation and the timing of speciation in Eucalyptus ................................. 47

2.4.3 Implications for tree breeding and the genetic risk posed by E. globulus ..... 48

2.5 Conclusion ........................................................................................................... 49

Chapter 3: Assessing a Bayesian approach for detecting exotic hybrids between plantation and native eucalypts .................................................. 78

3.1 Abstract .............................................................................................................. 78

3.2 Introduction ......................................................................................................... 79

3.3 Materials and Methods ..................................................................................... 81

3.3.1 Sample description ............................................................................................ 81

3.3.2 Molecular methods .......................................................................................... 85

3.3.3 Analytical approach ......................................................................................... 86

3.3.4 1) Assessment of genetic differentiation between species ................................. 86

3.3.5 2) Calculating detection power with simulated and pedigreed hybrids ......... 86

3.3.6 3) Classification of the putative hybrids using STRUCTURE ....................... 87

3.4 Results ................................................................................................................ 88

3.5 Discussion ......................................................................................................... 94

3.5.1 Performance of the approach ......................................................................... 94

3.5.2 Allocation of putative hybrids ......................................................................... 97

3.6 Conclusion ........................................................................................................ 98
Chapter 4: Assessing the risk of exotic gene flow from *Eucalyptus globulus* plantations to native *E. ovata* forests ......................................................... 113

4.1 Abstract ........................................................................................................ 113

4.2 Introduction .................................................................................................. 114

4.3 Materials and Methods ............................................................................... 119

4.3.1 The effect of patch size on hybridisation rate in a fragmented *E. ovata* landscape ....................................................................................................................... 119

4.3.2 Assessing the frequency of hybrid establishment in the wild ................. 121

4.3.3 Survival and fitness of *E. ovata* x *globulus* F1 hybrids in the wild ........ 122

4.3.4 Data analysis ................................................................................................ 123

4.4 Results ........................................................................................................... 125

4.4.1 The effect of patch size on hybridisation rate in a fragmented *E. ovata* landscape ....................................................................................................................... 125

4.4.2 Assessing the frequency of hybrid establishment in the wild ................. 126

4.4.3 Survival and fitness of *E. ovata* x *globulus* F1 hybrids in the wild ........ 128

4.5 Discussion ...................................................................................................... 128

4.5.1 Hybridisation and the effect of patch size, tree position and isolation ...... 128

4.5.2 Hybrid fitness ............................................................................................... 132

4.5.3 The likelihood of introgression ..................................................................... 133

4.5.4 Wildlings ....................................................................................................... 134

4.6 Conclusion ..................................................................................................... 135

Chapter 5: Assessing the invasive potential of *Eucalyptus globulus* in Australia: quantification of wildling establishment from plantations ........ 147

5.1 Abstract ........................................................................................................... 147

5.2 Introduction .................................................................................................. 148

5.3 Methods ......................................................................................................... 152

5.3.1 Survey area and method ............................................................................. 152
Chapter 1: 
General introduction: exotic gene flow and Australian eucalypt plantations

1.1 Gene flow

Gene flow is the transfer of genes between populations or species *via* migration and mating (Slatkin 1985a). Gene flow is often considered to be an intra-specific process, but inter-specific gene flow and subsequent introgression (the permanent incorporation of genes from species into the gene pool of another) can also occur where distinct species are able to hybridise (Potts *et al.* 2003). For example inter-specific gene flow has been reported in chimpanzees (Won and Hey 2005), imperial eagles (Martínez-Cruz and Godoy 2007), fruit fly (Powell 1983), white oaks (Whittemore and Schaal 1991), irises (Arnold *et al.* 1992), sun flowers (Rieseberg and Wendel 1993) and eucalypts (McKinnon *et al.* 2004a).

Both inter- and intra-specific gene flow can be important evolutionary process. In combination with natural selection and genetic drift, intra-specific gene flow helps shape the natural distribution of genetic diversity across the geographic range of a species (Gleaves 1973; Slatkin 1985a). Intra-specific gene flow is often considered to be a constraining force in evolution because it frequently counteracts natural selection and genetic drift (Slatkin 1985a; Slatkin 1987). That is, natural selection produces local adaptation, but regular gene flow between populations continually introduces new maladapted genotypes, slowing local adaptation as well as preventing the populations becoming differentiated through random genetic drift (Slatkin 1985a). However, intra-specific gene flow can also be a driving force in
evolution because it can spread highly adaptive traits that arise in one population across the species range potentially allowing more rapid evolution of the species as a whole (Slatkin 1987; Wright 1982). For example it is thought that the capacity for long-distance gene flow in forest trees could enable species to adapt to rapid climate change on a continental scale in just one generation (Kremer et al. 2012).

Inter-specific gene flow is more likely to cause differentiation and create novel evolutionary units because it involves (initially at least) the merging to two more highly diverged genomes (Mallet 2007). Although the fitness of inter-specific hybrids can be variable and is often reduced, there is also an associated increase in phenotypic variation on which natural selection can operate (Hegarty 2012). For example, hybrid zones often become established in ecotones between two parental habitats, where intermediate hybrid phenotypes might be uniquely adaptive (Arnold 1997; Rieseberg et al. 2003b). In this situation if there is some barrier that prevents backcrossing to the parental species (or there is very strong selection favouring the hybrid phenotype) then rapid speciation may occur (Arnold 1997). One of the best known examples of this occurs in sunflowers where three separate speciation events have occurred through hybridisation between the same two species, in three different and quite extreme environments (Rieseberg 1991; Rieseberg et al. 2003b). Contemporary (for example see Czypionka et al. 2012) and historical patterns of genetic diversity suggest that hybrid evolution is now, and has been important in the past (Mallet 2007). However, barriers to hybridisation are common, and the widespread co-occurrence of related species demonstrates that these barriers are usually effective at maintaining species boundaries (Ellstrand 2003; Potts et al. 2003).

In plants gene flow can arise through movement of pollen, seed or vegetative propagules between populations or species. Of these mechanisms, pollen dispersal is usually the most extensive and therefore generally contributes disproportionately to gene flow (Ellstrand 1988; Ellstrand 1992b). However, in certain systems and situations seed dispersal can be at least as important as pollen dispersal. For example in the wind-pollinated and wind-dispersed tree *Fraxinus excelsior*, seed dispersal was reported to be six times more important than pollen dispersal in maintaining
connectivity between forest remnants in fragmented landscapes (Bacles et al. 2006). Widespread seed mediated gene flow has also been important in shaping the genetic diversity of *Sorbus torminalis* across Europe (Oddou-Muratorio et al. 2001), and seeds make significant contributions to interpopulation gene flow in many *Pinus* species (Ozawa et al. 2013). These examples highlight the probably underrecognised influence of rare long-distance seed dispersal in effecting gene flow, and subsequently influencing patterns of genetic diversity in plants (Bacles et al. 2006).

Despite the importance of seed dispersal the generally higher mobility of pollen has resulted in studies of pollen mediated gene flow dominating the literature (reviewed by Ellstrand 1992b; Ellstrand 2003). Early studies of pollen dispersal from source populations lead to some misconceptions about the likely extent of gene flow (Ehrlich and Raven 1969; Levin 1984). These studies found strongly leptokurtic dispersal of pollen suggesting that gene flow might be far less extensive than previously thought (Levin 1984; Mayr 1963; Stebbins 1959). However, the advent of molecular techniques to measure gene flow into sink populations (as opposed to assessing dispersal from a source to estimate gene flow) showed that the tail of these leptokurtic distributions could result in evolutionarily significant gene flow (Ellstrand 2003; Friedman and Adams 1985).

The use of molecular markers has greatly enhanced the study of gene flow. Molecular markers are now used in two main ways, either to directly assign paternity so that minimum gene flow in a given cohort can be explicitly measured (simple paternity exclusion technique), or to estimate historical patterns of gene flow by studying the distribution of genetic variation within and between populations (Ellstrand 1992a). The simple paternity exclusion technique is effective for measuring contemporary gene flow, but is most commonly used in experimental settings because it requires that all genotypes within the population are known, so that offspring sired by immigrants can be identified, and this is often impractical in natural populations (Ellstrand 1992a). For this reason the second approach is most commonly applied in population genetic studies. It involves investigating either the distribution of rare alleles among populations (Barton and Slatkin 1986; Slatkin 1985b), or calculating the level of differentiation between populations – typically
using one of two commonly used metrics Wright’s $F_{ST}$ (Wright 1950) or Nei’s $G_{ST}$ (Nei 1973). These methods have now been widely employed across a range of taxa, and they reveal that gene flow is highly variable both within and between species (reviewed by Ellstrand 2003; Kisel and Barraclough 2010).

1.2 Exotic gene flow

1.2.1 Genetically modified crops

The fragmentation of natural populations and the movement and cultivation of plants around the world has created the opportunity for gene flow to take place between populations, varieties or species that would not naturally occur (here in referred to as exotic gene flow). The potential for exotic gene flow between wild (weed) populations and cultivated crop species was identified by Harlan (1965). He considered con-generic weeds that co-occurred with crop species to be “reservoirs of reserve germplasm” that had aided the domestication process by periodically injecting beneficial genes into the gene pool of the crop species (Harlan 1965). However, exotic gene flow became a widespread concern with the advent of genetically modified (GM) organisms (particularly crops). The potential for pollen-mediated gene flow and the subsequent escape of transgenes into natural populations has long been recognised (Brill 1985; Colwell et al. 1985). One fear is that gene flow to wild relatives could result in the evolution of more vigorous and difficult to control weeds (Colwell et al. 1985). Indeed transgenes have been detected in the weedy relatives of canola, Mazie, rice and other common GM crop species (Mannion and Morse 2013). For example, transgenes for insect resistance have been found to reduce herbivory by 79 % and increase fecundity by 44 % in advanced generation rice crop-weed hybrids (Yang et al. 2011). There is also evidence in canola, that weed populations can obtain multiple transgenes from different crops. Individuals have been found that are resistant to both Glyphosate and Glyphosate, a combination that has not been engineered, and therefore must have evolved “naturally” in the weed population as a result of gene flow from two different GM varieties (Mannion and Morse 2013). Despite these examples, introgression of a transgene will not automatically confer a fitness benefit in the weed population. For example herbicide resistance will only be advantageous to a weed if the population is
treated with a herbicide, and there could actually be a fitness cost through resource allocation in individuals carrying such a trait (Duke 2005). Thus, the potential for, and the effect of such gene flow needs to be assessed on a case by case basis (Warwick et al. 2009).

Exotic gene flow from GM crops may also erode the genetic integrity of native populations and landraces of important crop varieties. For example, Mexico is the global centre of Maize genetic diversity and origin of Maize domestication. There are a number of cross-compatible wild relatives and at least 60 domesticated landraces from which the world’s commercially important Maize variates have been developed (Pineyro-Nelson et al. 2009). The Mexican government banned the use of GM Maize because it feared that exotic gene flow could result in the introgression of transgens into culturally important landraces and native populations. Transgenic Maize has since been illegally introduced to the region, and exotic gene flow has lead to the introgression of transgens into local landraces in several locations (Pineyro-Nelson et al. 2009). A similar situation has arisen with creeping bent grass, which has become one of the most well studied examples of pollen mediated transgenes escape (Reichman et al. 2006; Watrud et al. 2004; Zapiola et al. 2008; Zapiola and Mallory-Smith 2012). Transgenic populations have been found up to 21 km from the original source (Reichman et al. 2006; Watrud et al. 2004; Zapiola et al. 2008), and gene flow has recently been detected between a feral GM creeping bent grass (Agrostis stolonifera) population and rabbit foot grass (Polypogon monspeliensis). This is the first report of inter-generic trasgene movement in plants and hybrids in both directions were found to be capable of producing viable backcrosses, making introgression a real possibility (Zapiola and Mallory-Smith 2012).

1.2.2 Large scale biological releases

Exotic gene flow from GM sources has received most attention in the literature. However, gene flow from exotic non-GM sources can have similar consequences, particularly in compromising the genetic integrity of natural populations (Byrne et al. 2011; Laikre et al. 2010). This is the case in industries that involve the release of large numbers of individuals with restricted or manipulated genetic diversity into areas where they can interact with related natural populations. Such industries
include fisheries, wildlife management, forestry and revegetation, and managing the genetic risk posed by these activities is a growing concern (Byrne et al. 2011; Laikre et al. 2010; Potts et al. 2003). These industries generally aim to improve productivity of a “natural system” by artificially increasing population sizes with translocations and introductions (Laikre et al. 2010). The number of individuals involved in these large scale biological releases is staggering (Laikre et al. 2010). In a review of the issue Laikre et al. (2010) show that over 10 billion hatchery bread fish are released into northern hemisphere rivers and oceans annually, and in Sweden more than 30 billion Norway spruce were planted in the 20th century alone. The genetic diversity of these large introduced cohorts are rarely (if ever) a good representation of the gene pool of natural populations (Charlesworth 2006; Gold et al. 2008; Laikre et al. 2010). The introduced populations may have reduced diversity, for example *P. radiata* (Moran and Bell 1987) and the extreme example of clonal *Populus* plantations (Vanden Broeck et al. 2005), or they could have much broader genetic diversity than any one local population (Jones et al. 2006). In either case subsequent introgression is likely to alter the composition of the native gene pools.

Laikre et al. (2010) identify four main ways in which the genetic integrity of natural populations can be affected by exotic gene flow from large scale biological releases: they can 1) change population structure (reduce differentiation between populations; 2) change genetic composition (genetic replacement); 3) cause the breakdown of local genetic adaptation; 4) and reduce genetic diversity. Fisheries is the only industry with a long enough history of release and where research has been undertaken on a scale large enough to detect all these effects, and they have all been found in natural populations (see Box 2 Laikre et al. 2010). For example, the massive release of salmonoids around the world have caused reduced population differentiation (Eldridge et al. 2009), the extinction of local gene pools (Araguas et al. 2009), lead to a loss of local adaptation and population fitness (McGinnity et al. 2009), and reduced genetic diversity (Peeler et al. 2006). Although fisheries is the most well studied system, exotic gene flow has now been detected in other industries involving large scale biological releases (Laikre et al. 2010).
The forestry industry is potentially a significant source of exotic gene flow, with very large plantation estates often embedded within a landscape containing cross-compatible congeners (Barbour et al. 2010; Potts et al. 2003; Vanden Broeck et al. 2005). The genus *Populus* provides a good example of the issues (Vanden Broeck et al. 2005). *Populus* species are naturally common across the northern hemisphere, and have been widely cultivated with species often grown well outside their native range (Vanden Broeck et al. 2005). They are dioecious and wind pollinated, with natural populations typically showing high levels of pollen mediated gene flow (Eckenwalder 1996; Vanden Broeck et al. 2005). Natural populations of many *Populus* species are highly fragmented, and as a result several species are now of conservation concern (Cagelli and Lefevre 1995; Smulders et al. 2008; Talbot et al. 2012). Hybridisation has been known in the genus for centuries, and exotic gene flow from introduced variates is recognised as a threat to natural populations (Cagelli and Lefevre 1995; Lexer et al. 2010). The genus is commonly used in forestry plantations to produce paper pulp in Europe, North America and China (Vanden Broeck et al. 2005). Plantations often co-occur with native stands, and are usually developed from a small number of elite hybrids through clonal propagation so that variation in tree form and performance can be minimised (Vanden Broeck et al. 2005). This propagation system produces plantations with very narrow genetic diversity (Vanden Broeck et al. 2005). This combination of biology and silvicultural regime has lead to concerns about exotic gene flow from *Populus* plantations (Benetka et al. 1999; Cagelli and Lefevre 1995; Smulders et al. 2008; Vanden Broeck et al. 2005).

A number of studies have now assessed exotic gene flow from *Populus* plantations in both Europe and North America. *Populus nigra* is one of the most threatened trees in Europe (Rathmacher et al. 2010), and is thought to be at risk of exotic gene flow from the common plantation hybrid *P. x canadensis* (= *P. deltoides* × *nigra*) (Cagelli and Lefevre 1995; Vanden Broeck et al. 2005). Several early studies showed either no, or very low levels of hybridisation in open pollinated *P. nigra* seed collected from the vicinity of *P. x canadensis* plantations (Benetka et al. 1999; Fossati et al. 2003; Heinze 1997). One study that found high levels of hybridisation investigated gene flow into a single isolated *P. nigra* female surrounded by male *P. x canadensis,
suggesting that the competitiveness of intra-specific pollen may limit exotic gene flow (Vanden Broeck et al. 2004; Vanden Broeck et al. 2005). More recently about half of the P. nigra seedlings establishing along a section of riverbank in the Netherlands were found to be of hybrid origin, indicating hybrids are being produced and they can effectively compete for establishment niches with their pure siblings (Smulders et al. 2008). In Canada, exotic gene flow into natural stands of P. deltoides was detected from three exotic species used in nearby hybrid plantations – P. nigra, P. trichocarpa, and P. maximowiczii (Meirmans et al. 2010). This study found significant variation in hybridisation rate between stands, between sites and between years (hybridisation ranged from 2 - 72%; Meirmans et al. 2010). Another Canadian study recorded exotic gene flow of 2.3% in native stands of P. balsamifera, and identified reproductively mature hybrid individuals established within native populations, meaning that advanced generation hybridisation and introgression could be occurring (Talbot et al. 2012). These examples highlight that, like natural gene flow, exotic gene flow from forestry plantations is likely to vary within and between species depending on a range of biological factors (e.g. taxonomy, sex ratio and relative population size), and will require case-by-case risk assessment and monitoring (Vanden Broeck et al. 2005).

1.3 Exotic gene flow from eucalypt plantations in Australia

Over the past two decades there has been a dramatic increase in the eucalypt plantation estate in Australia, which now totals around one million hectares (Gavran and Parsons 2011). In many cases this has involved translocating plantation species beyond their natural range, and into the range of native eucalypt species with which they do not naturally co-occur (Barbour et al. 2008a; Barbour et al. 2008b; Barbour et al. 2006a). Eucalypts (in the broad-sense) are a diverse group of plants, most of which are encompassed by the genera Eucalyptus, Corymbia and Angophora, which include around 900 taxa (Euclid 2006). There are often weak reproductive barriers between closely related species and reports of natural and artificial hybrids are common (Potts et al. 2003; Potts and Dungey 2004). The large scale of the plantation establishment and the lack of reproductive isolation has led to concerns that plantation eucalypts could lead to widespread exotic gene flow in natural
populations (Barbour et al. 2008a; Barbour et al. 2008b; Potts et al. 2003). Furthermore because eucalypts are usually keystone species, shifts in the phenotypic characteristics of native populations resulting from hybridisation and introgression, may have flow-on effects to the broader community and ecosystem (Barbour et al. 2009; Dungey et al. 2000; Whitham et al. 2006). In their global review of the issue Laikre et al. (2010) indicate that exotic gene flow has been largely neglected by the forestry sector, with the exception of Australia where it has been well studied. Following is a summary of the work undertaken in Australia to date.

1.3.1 **Barriers to exotic gene flow provide a research and management framework**

The likelihood of introgression depends on the presence and strength of a range of pre- and post-mating barriers to gene flow. Pre-mating barriers to gene flow include flowering asynchrony, and pollen dispersal distance. Post-mating barriers include a range of physical and genetic pre-zygotic barriers, such as pollen tube growth; as well as post-zygotic barriers, such as seedling fitness (Potts et al. 2003). Figure 1.1 summarises the range of barriers to exotic gene flow and their interactions, demonstrating how they may operate to prevent introgression. Research to date in Australia has focused largely on understanding the strength of key barriers to exotic gene flow, so that its probability of occurrence can be incorporated into risk management frameworks (see Barbour et al. 2008a; Barbour et al. 2008b; Byrne et al. 2011).
Figure 1.1 Framework for identifying barriers to exotic gene flow from eucalypt plantations to native forest (modified after Potts et al. 2003). Successful introgression requires that all of these barriers are overcome.

The majority of research into the potential risk of gene flow in Australia has been conducted in temperate regions where the main plantation eucalypts are *E. globulus* and *E. nitens* (Potts et al. 2003). Much of the research has been done in Tasmania, which is being used as a model for exotic gene flow research because it is a closed island system with only 30 native eucalypt species and the main plantation species *E. nitens* is locally exotic and widely planted (207,000 ha; Gavran and Parsons 2011). Additionally, *E. nitens* is in the subgenus *Symphyomyrtus*, which is reproductively isolated from the 13 Tasmanian species in the subgenus *Eucalyptus*, meaning that only 17 native eucalypt species are potentially at risk of exotic gene flow in Tasmania, further reducing complexity in the system.

1.3.2 Pre-mating barriers – before pollen reaches the flower

1.3.2.1 Pollen quantity

A primary barrier to gene flow is the quantity of pollen produced from any given plantation. Pollen output is dictated by the time to first flowering, the flower abundance in the plantation, which is in turn influenced by plantation size, tree spacing and position, as well as a range of environmental and site conditions.
Chapter 1: General introduction

(Barbour et al. 2006b; Jones et al. 2011; Potts et al. 2003; Williams et al. 2006). An assessment of the level of flowering across the *E. globulus* estate from Tasmania to Western Australia, showed that flowering was generally low in pulpwood plantations with 7.8 % of trees flowering in edge rows and 4.4 % of trees flowering in inside rows (Barbour et al. 2008b). Although, these estimates were derived from a range of plantation ages (from 1-10 years old; Barbour et al. 2008b), and higher levels have been reported when assessing harvest age plantations 9-12 years old (29% in edge rows; Larcombe et al. 2014). The majority of the Australian plantation estate is currently managed for pulp production, but transitioning parts of the estate to solid wood is being considered (Beadle et al. 2008; Forrester et al. 2010). Solid wood management may employ silvicultural techniques (e.g. longer rotations and thinning) that increase pollen production from a given plantation (Forrester et al. 2010; Williams et al. 2006).

1.3.2.2 Flowering season overlap

Flowering season overlap is an important barrier because gene flow cannot occur unless the plantation and native forest species flower at or around the same time (Potts et al. 2003). Flowering season is dictated by the interaction between environmental conditions and tree genetics (Jones et al. 2011). Variation in the duration and peak flowering time between the 17 at risk Tasmanian species and *E. nitens* has been assessed (Barbour et al. 2006b). Flowering synchrony was highly variable, with the proportion of flowering overlap between *E. nitens* and native species ranging from 3 to 97 %, and it also varied with season and altitude (Barbour et al. 2006a).

1.3.2.3 Dispersal (mainly pollen)

Eucalypts typically have limited seed dispersal (usually one to two tree heights; Cremer 1977) meaning that hybrid or pure seed from plantations is unlikely to reach a native forests to effect gene flow (Potts et al. 2003). Two possible exceptions to this are *E. camaldulensis* and *Corymbia torelliana*. The seed of *E. camaldulensis* is adapted for water dispersal and can travel long distances (Pettit and Froend 2001). However, although *E. camaldulensis* is widely grown in plantations outside
Australia (Harwood 2011), it is only a minor component of the Australian plantation estate (Gavran and Parsons 2011), so widespread seed-mediated gene flow from this species is unlikely. Conversely, *C. torelliana*, which has bee dispersed seed (Wallace and Trueman 1995), is widely used in subtropical forestry, particularly in hybrid breeding programs. Recent work on the dispersal mechanism (involving resin harvesting from the capsule by native bees) has shown that seed can be dispersed 220m by bees, but that the majority (88%) of seed is dispersed by gravity close to the tree (Wallace et al. 2008). The inheritance of this mechanism in hybrids will be important for determining the probability of seed-mediated gene flow, although the range of attributes that are required to facilitate this type of gene flow are rare (H. Wallace pers. comm. April 2012).

Pollen dispersal is thought to be the main mechanism for exotic gene flow in eucalypts and is affected by numerous factors including pollinator specificity, nectar abundance, pollen longevity, and source-sink ratios (Potts et al. 2003). The potential distance of pollen transport depends on the type of pollinator, for example insects will typically transport pollen shorter distances than birds and mammals (Southerton et al. 2004). Some *Corymbia* species may have particularly long dispersal potential because their pollinators include bats which may regularly travel upwards of 50 km in one night (Southerton et al. 2004). However, more recent work has shown that bats are only occasional visitors to *Corymbia* flowers (<2% of visitations), with birds (lorikeets and honeyeaters) being more regular (22% of visitation), and the majority of flower visitations were made by insects (mainly the introduced honey bees 73%), arguing that most pollen will be distributed close to plantations with sporadic long distance bird and bat dispersal (Bacles et al. 2009). In comparison, at least planted *E. nitens* is thought to be exclusively insect pollinated and the vast majority of pollen dispersal occurs within 200 m of the source, although rare pollination events extend to at least 1600 m (Barbour et al. 2005b). Some studies have also shown that pollen may move further in fragmented landscapes. In the insect pollinated *E. wandoo*, 65% of pollinations in isolated remnants arose from pollen that had travelled at least 1 km (Byrne et al. 2008). Trends for increased pollen dispersal in fragmented populations have also been noted in *E. nitens* (Barbour et al., 2005b) and in one of the two regions studied in *E. globulus* (Mimura et al. 2009). Understanding pollen
dispersal will be particularly important where isolated populations of rare species occur near plantations.

1.3.3 Post-mating barriers – after pollen reaches the flower

1.3.3.1 Pre-dispersal – post-zygotic

Pre-zygotic barriers are a range of physical and genetic barriers operating between the stigmatic surface and the ovaries. Style length can be a barrier; for example the pollen tube of small-flowered \textit{E. nitens} cannot reach the ovaries of the large-flowered \textit{E. globulus}, whereas the reverse cross is possible (Gore \textit{et al}. 1990). Phylogenetic relationships also seem important. Ellis \textit{et al}. (1991) showed that as phylogenetic distance between mothers and pollen parents increased pollen tube penetration down the stigma decreased. More recently, studies in \textit{Corymbia} have shown phylogenetic related pre-zygotic barriers exist in pollen adhesion to the stigma, pollen germination, pollen tube growth, and pollen penetration of the ovule (Dickinson \textit{et al}. 2012b). Thus, phylogenetic relationships between plantation species and neighbouring native populations are likely to be informative when assessing the probability of exotic gene flow.

Post-zygotic, pre-dispersal barriers that operate on developing embryos may also be important, especially in the absence of pre-zygotic barriers (Dickinson \textit{et al}. 2012b). For example embryo abortion is thought to be important for regulating self-fertilisation (Pound 2003; Pound \textit{et al}. 2002). Hybrid embryos may have impaired cellular process, including failure to divide or reduced rates of division, resulting in embryo abortion/failure or seed with fewer resources compared to intra-specific siblings at dispersal, possibly resulting in weaker seedlings (Sedgley and Granger 1996). Experiments to identify the importance of factors acting on embryo development between inter-specific crosses are time-consuming because of the need to decouple pre- and post-zygotic effects.

One method of assessing the integrated effect of post-mating, pre-dispersal barriers is to manually apply pollen from the plantation species to receptive stigma of the native species. Natural pollination can be better mimicked by allowing pollen competition. This can be done in a native forest situation by crossing without
isolating the flower of the native species (flowers are isolated after pollination in controlled crossing) and allowing the exotic pollen to compete with natural intra-specific pollinations. This “dabbing” technique requires suitable morphological, molecular or other markers to identify hybrids (Delaporte et al. 2001). Barbour et al. (Barbour et al. 2005a) used this approached to quantify the relative risk that *E. nitens* posed to the 17 potentially compatible Tasmanian eucalypt species, pollinating an average of 790 flowers on up to 20 trees per species. Hybrids were recorded in all but one species (*E. urnigera*) and the rate of hybridisation in the other species varied considerably ranging from 1 % [*E. cordata*] to 29 % [*E. rubida*] in the over 18,000 seedlings assessed. Recent work in the subtropics using isolated controlled pollinations has shown that although barriers to hybrid seed set exist in *Corymbia*, they are incomplete (Dickinson et al. 2012a; Dickinson et al. 2012b). Future studies on *Corymbia* might extend this work to hybridisation rates in natural situations, where factors such as intra-specific pollen competition are operating. Techniques for identifying hybrids will need to be developed for *Corymbia* (see Abasolo et al. 2012, and below).

1.3.3.2 Post dispersal – hybrid fitness

Exotic hybrids must successfully disperse, establish, survive and reach reproductive maturity for introgression to occur. Low hybrid fitness has been reported for many inter-specific eucalypt crosses, although increased vigour has also been reported (discussed in Potts et al. 2003). In Tasmania, Barbour et al. (2006a) found that *E. ovata x nitens* hybrid seed (collected from *E. ovata* trees adjacent to *E. nitens* plantations) was 61 % less likely to establish in native forest compared with pure *E. ovata*. They also found reduced growth, increased mortality and poorer plant health in hybrid seedlings compared to their pure siblings at age four (Barbour et al. 2006a), with health remaining lower in the surviving hybrids at 11 years of age (unpublished data). Reduced hybrid fitness (*E. globulus x nitens*), in terms of growth and survival, has recently been attributed to outbreeding depression associated with genetic interactions (epistasis) that operate as a post-zygotic reproductive barrier (Costa e Silva et al. 2012). Outbreeding depression was evident in first generation (F₁) hybrids and became more pronounced in second generation (F₂) and backcross
hybrids. This barrier may be important in maintaining the genetic integrity of species in nature (Costa e Silva et al. 2012), and could also operate to limit exotic gene flow.

In *Corymbia* species in the subtropics, while hybrid seed set is often reduced and some F₁ hybrid combinations show reduced germination, nursery growth (Abasolo et al. 2012), and field survival (Lee et al. 2009), vigorous *Corymbia* F₁ hybrids have been selected for commercial development (Lee et al. 2010). The deployment of selected hybrids bypasses some of the first generation barriers associated with vegetative fitness but not necessarily reproductive fitness. However, *Corymbia* F₁ hybrids have been shown to readily produce flowers and seed in plantations (Barbour et al. 2008a), and hybrid backcrosses may have capsule and seed-set rates equivalent to intra-specific crosses indicating few pre-zygotic barriers to gene flow exist (Dickinson et al. 2012a). Furthermore, recent work has shown that spontaneous advanced generation hybrids can establish and remain vigorous in the wild for at least three years (Shepherd et al. 2013). Given the possibility of such backcrossing, and that the hybrids typically involve *C. torelliana*, identifying the rate of hybridisation from hybrid plantations and possible seed dispersal rates (Wallace et al. 2008), will be important information for effective risk management in the future.

In the temperate system, convenient juvenile characters exist that allow identification of hybrids between most at-risk native species and plantation *E. nitens* and *E. globulus* at the early seedling stage (Barbour et al. 2008a; Barbour et al. 2005a). This has allowed assessments of hybridisation rate in native forests adjacent to plantations (Barbour et al. 2008b). *Corymbia* species have less distinctive juvenile characteristics making morphological identification of hybrids more difficult (Barbour et al. 2008a). Abasolo et al. (2012) have recently identified a suite of characters that can be used to differentiate F₁ hybrids with low to moderate accuracy, but differentiation of backcrosses was problematic. To assess exotic hybridisation rates, other systems such as molecular (DNA) markers will need to be used in conjunction with these morphological markers to confirm *Corymbia* hybrids, particularly backcrosses to native *Corymbia* populations (as in Shepherd et al. 2013).

It should be noted that even if introgression does not occur, hybridisation itself may affect the demography of a population if a large proportion of progeny are exotic
hybrids (demographic swamping; Ellstrand 1992b), this will be of particular concern in small isolated populations (Ellstrand 1992b). The presence of exotic hybrids in a native forest could also have effects beyond the population and affect dependent communities (Barbour et al. 2009; Laikre et al. 2010; Tovar-Sanchez and Oyama 2006). For example the intermediate chemistry of hybrids may allow parasite species to “jump” from the plantation to the native species (the hybrid bridge hypothesis; Floate and Whitham 1993). Demographic swamping and the hybrid bridge hypothesis do not involve introgression but should be considered in assessing the consequences of exotic hybridisation (Ellstrand 1992b; Laikre et al. 2010).

1.3.4 Risk management and simulation modelling

The issues outlined above deal predominantly with assessing the probability of exotic gene flow and subsequent introgression at various stages in the gene flow pathway. The issue facing forest managers is the ‘risk’ of these events, which is a function of their probability, and the consequences of their occurrence. Given the economic, social and environmental benefits of plantation forestry, managers need a method of assessing these benefits against any potential negative consequences (Byrne and Stone 2011; Byrne et al. 2011). Risk assessments and simulation modelling are two approaches that can be used to integrate empirical data with environmental and biological information to estimate risk (DiFazio et al. 2004; Byrne et al. 2011). Risk assessments provide frameworks (often in the form of decision-trees) for considering factors that influence the probability of an event occurring (i.e. exotic gene flow) and identify the consequences of that event (Byrne and Stone 2011; Byrne et al. 2011; Suter 2006). The outcome of such an assessment is to identify a risk metric, so as to maximise the benefits of the proposed activity where it poses a low risk, while minimising negative environmental effects when a high risk is identified. Simulation modelling is a tool that can be used to inform risk assessments. Simulation models incorporate empirical data with biological and environmental data using computer modelling and Geographic Information Systems (GIS), to produce projections of how gene flow will progress in complex natural systems through both space and time (DiFazio et al. 2004; DiFazio et al. 2013). Simulation modelling and risk assessments have been important tools in
understanding the risk of GM trans gene escape (Colbach et al. 2001; Meirmans et al. 2009; Paoletti et al. 2008), and have also been used to assess the risk of gene flow from *Populus* plantations in North America (DiFazio et al. 2004; DiFazio et al. 2013; Meirmans et al. 2010).

Ideally risk assessments and simulation modelling should be carried out prior to plantation establishment (Byrne and Stone 2011; DiFazio et al. 2013). Risk assessment protocols have been developed for *E. nitens* (FPA, 2009), *E. globulus* (Barbour et al. 2008b), and for assessing genetic risk in revegetation (Byrne et al. 2011) and are all highly transferable. These assessments consider biological, phylogenetic and geographic information as well as the conservation status of potentially affected native species, to identify the risk of exotic gene flow (typically low, moderate or high risk) and/or give pre- and post-establishment guidelines. The relatively long generation time of eucalypts means that practical management techniques such as hand weeding of established hybrids are feasible in many situations, and all the above risk assessments promote such adaptive management approaches when a moderate risk of exotic gene flow is identified. Simulations studies are yet to be widely used in assessing the risk of gene flow from eucalypt plantations, but would be useful for integrating empirical data and estimating the effect of low levels of gene flow over very long time scales.

The global demand for certified forest products is increasing as people become more aware of sustainability in the production systems of the wood and wood products they are purchasing. In response to this demand, many forestry companies seek certification under either the Australian Forestry Standard (AFS) or the Forest Stewardship Council (FSC). Both these standards have clauses related to containing plantation germplasm and preventing “genetic pollution” (exotic gene flow and introgression) of nearby native populations. Evidence of ‘genetic risk’ assessment and management planning is likely to be assessed for compliance under these schemes (Ross Garsden, NSC International Pty Ltd, Pers Comm).
1.3.5  A case study: E. perriniana *Strickland, central Tasmania*

Rare species are a high priority for genetic risk assessment and management (Barbour *et al.* 2010). The Eucalypt Genetics Group at the University of Tasmania has been working with the Tasmanian Forest Practices Authority (FPA) to implement a risk assessment and management framework (FPA, 2009) that integrates the research on exotic gene flow from *E. nitens* plantations undertaken in the Tasmania landscape (Barbour *et al.* 2003; Barbour *et al.* 2005a; Barbour *et al.* 2005b; Barbour *et al.* 2006a). The assessment was conducted on all 17 at risk Tasmanian species, and based on flowering overlap, cross-compatibility, geographic proximity and conservation status. *Eucalyptus perriniana* was identified as one of the species most at risk (Table 1, page 4 in FPA 2009). In Tasmania, *E. perriniana* is known from three populations totalling around 1000 individuals and is listed as rare under the *Threatened Species Protection Act 1995*. In 2003 a proposal to expand *E. nitens* plantations around the largest of its three populations at Strickland was identified by FPA to pose a high risk of exotic gene flow. A consultative committee was formed involving FPA, Eucalypt Genetics Group, the land owner, and the Threatened Species Section (Department of Primary Industries and Water) to identify possible alternatives or solutions. Based on the genetic risk assessment, planting *Pinus radiata* was considered to avoid exotic gene flow. This outcome was seen as less appropriate because of the landscape context of a large *P. radiata* plantation surrounding a small formal reserve and the associated wildling (weed) risk. Instead, an adaptive management strategy was implemented. *Eucalyptus nitens* would be established beyond a 500m buffer – based on the pollen dispersal reported in Barbour *et al.* (2005b) – with ongoing monitoring of the levels of hybridisation, and an option to extend the buffer distance at second rotation (or earlier) if a serious genetic risk was realised. At the time of plantation expansion (c. 500 ha) a smaller plantation of 60 ha was already established within 630m of the *E. perriniana* population (Fig. 1.2). Pollen flow from this plantation could be monitored to assess the likelihood of hybridisation and provide a base-line by which to assess the impact of the expansion. The monitoring involved (i) collecting and growing seed from 100 ‘sentinel’ trees in the population to detect annual exotic and natural hybridisation
rates, and (ii) survey for seedling establishment following regeneration events in the
*E. perriniana* population.

The level of hybridisation has been assessed since 2005 using morphological
markers, and molecular genetics to confirm putative hybrids (Barbour *et al.* 2010).
Levels of exotic hybridisation from the *E. nitens* plantation are low and although
they have varied they have never reach 1 % (Table 1.1; Larcombe *et al.* 2012).
Additionally, the level of natural hybridisation (with co-occurring native species) is
much higher than exotic hybridisation indicating the population may be resilient to
some inter-specific gene flow (Larcombe *et al.* 2012). This population of *E.
perriniana* occurs in an unusual habitat that is water logged for about half the year,
and dry for the other half. These extreme conditions make seedling recruitment
difficult, and seedlings are rare (Rathbone *et al.* 2007). In fact, until consecutive fires
burnt parts of the population in 2007 and 2008, no seedlings had been recorded in
the population. Despite 100’s of trees being affected by these fires only 26 seedlings
were recorded in surveys of the burnt area in 2009, most of these were establishing
around a single tree, and none were hybrids with *E. nitens* (Fig. 1.2; Larcombe *et al.*
2012). The low levels of exotic hybridisation and seedling establishment indicate
that exotic gene flow and introgression are unlikely at this stage, and monitoring will
continue as the second stage of *E. nitens* planting reaches reproductive maturity.
Figure 1.2. Left: Map showing the proximity of the *E. perriniana* population to the mature 60 ha *E. nitens* plantation. Circles show the position of the 100 study trees and red bubbles indicate trees where hybrids have been found. Right: a mature *E. perriniana* coppicing after the 2008 fire with detail of seedling establishing at its base. A total of 26 seedlings were identified 18 months after the fire, mainly around this tree.
Table 1.1. The percent of hybridisation with *E. nitens* and co-occurring native species recorded in seedlots collected from 90 to 100 *E. perriniana* trees in the Strickland population which would have resulted from flowering in the years indicated. The large planting of *E. nitens* in the surrounding landscape was not reproductive during this period and the main source of pollen was the c. 60ha plantation 630m from the population (Fig. 1.2). Results from Larcombe et al. (2012).

<table>
<thead>
<tr>
<th>Year</th>
<th>% hybridisation</th>
<th>No. Seedlings assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>E. nitens</em></td>
<td>Native species</td>
</tr>
<tr>
<td>2005</td>
<td>0.02</td>
<td>2.32</td>
</tr>
<tr>
<td>2006</td>
<td>0.36</td>
<td>1.50</td>
</tr>
<tr>
<td>2008a</td>
<td>0.05</td>
<td>2.55</td>
</tr>
<tr>
<td>2009</td>
<td>0.02</td>
<td>7.93</td>
</tr>
<tr>
<td>2010</td>
<td>0.24</td>
<td>7.53</td>
</tr>
</tbody>
</table>

\*Fires burnt approximately one third of the population in 2007 and 2008, which coincided with an increase in levels of native hybridisation, perhaps in response to a decrease in intra-specific pollen.

1.4  **Exotic gene flow from *Eucalyptus globulus* in Australia**

The 1 million ha of *Eucalyptus* plantations growing in Australia are dominated by two species, *Eucalyptus globulus* (55.3%) and *E. nitens* (24.2%) (Gavran and Parsons 2011). As noted above research on the risk of exotic gene flow from these plantations has focused largely on the case of *E. nitens* in Tasmania (but see Barbour et al. 2008a; Barbour et al. 2008b). The challenge now is to expand this research to the larger *E. globulus* estate to test the generality of the exotic gene flow patterns identified in *E. nitens*, as well as produce specific risk management protocols for *E. globulus*.

1.4.1  **Key difference between *Eucalyptus globulus* and *E. nitens***

There are a number of important differences between the *E. nitens* system in Tasmania and the situation with *E. globulus*. Of the approximately 538,000 ha of *E. globulus* plantations in Australia most occur in the southern mainland, mainly outside the species natural range (Gavran and Parsons 2011). Barbour et al. (2008b) found that 48% of *E. globulus* plantations occur beside cross-compatible native species, and unlike *E. nitens* in Tasmania (with 17 cross-compatible cogeneras) this includes hundreds of eucalypt species across the estate. *Eucalyptus globulus* and *E.
are closely related taxa, both occurring in the same section (Madeneria) in the subgenus *Symphomyrtus* (Brooker 2000), but differences in flower morphology and pollinators may result in different pollen dispersal patterns and exotic gene flow potential (Hingston *et al.* 2004a; Hingston and Potts 2005; Hingston *et al.* 2004b; Hingston *et al.* 2004c). *Eucalyptus globulus* has large single flowers (occasionally in threes) while *E. nitens* has small flowers in umbels of seven (Euclid 2006). This could influence the genetic risk posed by *E. globulus* in that larger flowers attract larger, potentially more mobile pollinators (Armstrong 1979; Hingston *et al.* 2004b; Hingston *et al.* 2004c). Additionally, because the length of pollen tube growth is related to flower size (i.e. larger flowers have longer pollen tubes; Gore *et al.* 1990), *E. globulus* may be able to pollinate species with a larger range of flower sizes than *E. nitens*. For example, *E. globulus* pollen can pollinate *E. nitens* flowers, but the reciprocal cross is not possible because the *E. nitens* pollen tube cannot reach the ovary of the larger *E. globulus* flowers (Gore *et al.* 1990). These geographic and biological features mean that understanding the risk of exotic gene from *E. globulus* plantations will require specific research targeting key stages in the exotic gene flow and introgression pathway.

1.4.2 *Seed mediated gene flow from E. globulus – an emerging issue?*

As discussed above, because pollen dispersal is typically more effective than seed dispersal, gene flow research has focuses mainly on pollen mediated gene flow (Ellstrand 1992b). However, seed dispersal is often significant on a local scale (Ellstrand 1992a; Ellstrand 1992b), and if seeds can disperse from a plantation into neighbouring native forests then exotic seed-mediated gene flow could affect native populations. There are two main consequences of such seed dispersal, initially the seedlings establishing within the native forest represent an invading exotic species, which has a range of potential ecological consequences (Richardson and Rejmánek 2011). Secondly if these seedlings become established, reach reproductive maturity and can hybridise with the native species then they could be a source of pollen-mediated gene flow and subsequent introgression. The occurrence of seed-mediated exotic gene flow has not been assessed in eucalypts (except for the special case of *C. torelliana*, see above), but recent reports of *E. globulus* becoming invasive in other
countries suggests that it should be considered. In a global review of invasive trees, *E. globulus* was reported to be invasive in a third of the regions considered (5 of 15; Richardson and Rejmánek 2011). Recent studies in Portugal and Spain have also shown that *E. globulus* plantations can spread to neighbouring communities via seed dispersal, with seedling densities in Spain predicted to be as high as 5000 plants/ha 20m from the plantation boundary (Calviño-Cancela and Rubido-Bará 2013; Silva and Marchante 2012). These reports certainly indicate a potential for seed-mediated gene flow from *E. globulus* plantations.

The potential for seed dispersal also needs to be viewed in the context of the Australian plantation industry. The reproductive capacity of eucalypts increases with age and because plantations in Australia are typically managed on 10 to 15 year rotations (Greaves et al. 2003) their pollen (and seed) production may be limited by age. Additionally given an uncertain pulpwood market in Australia, in some cases plantations may not be replanted following harvest (Gavran and Parsons 2011), possibly removing the pollen source indefinitely. However, if seed dispersal from plantations results in seedlings establishing within native forests, given that they have a potential life span of hundreds of years, these trees could become a long term source of exotic pollen and gene flow if they are not managed appropriately. Therefore research aimed at assessing the spread of *E. globulus* seedlings from plantations is needed for a complete assessment of exotic gene flow risk in Australia.

1.5 **Thesis outline**

The aim of this thesis is to provide biological information to help improve the management of exotic gene from *E. globulus* plantations. This is achieved by investigating key stages in the gene flow pathway so that barriers to gene flow can be identified and incorporated in to risk assessments, as detailed in the chapter outlines below.

1.5.1 *Chapter 2: Phylogenetic patterns of reproductive isolation in Eucalyptus*

Prior to this study, based on known barriers to hybridisation in *Eucalyptus*, any of the 484 species in subgenus *Symphyomyrtus* were thought to be potentially at-risk of exotic gene flow from *E. globulus*. This chapter uses controlled crossing between *E.
globulus and 100 native species, mainly from Symphyomyrtus, to identify patterns in
cross-compatibility. The results are then combined with genetic distances from
genome wide scans to identify the relative risk of exotic gene flow from E. globulus
to the various phylogenetic clades within Symphyomyrtus. This approach also allows
the first broad assessment of the evolution of reproductive isolation in the
Eucalyptus.

1.5.2 Chapter 3: A microsatellite database and Bayesian modelling approach for
identifying hybrids between plantation and native eucalypts.

The ability to identify hybrids is essential when investigating exotic gene flow. In
this chapter molecular markers (microsatellites) are used to develop a hybrid
identification method for six species (E. camaldulensis, E. cypellocarpa, E. globulus,
E. nitens, E. ovata and E. viminalis) that are thought to be at risk from exotic gene
flow. The technique is used in subsequent chapters for hybrid validation, and is also
now available to enable fast low cost assessment of exotic hybridisation in the field
for management purposes.

1.5.3 Chapter 4: Assessing the risk of exotic gene flow from Eucalyptus globulus
plantations to native E. ovata forests

Eucalyptus ovata is known to be cross-compatible with E. globulus and is a common
plantation neighbour, making it ideal for studying the risk of exotic gene flow in
detail. This chapter investigates key components of the exotic gene flow pathway in situ in E. ovata, specifically: the effect of the recipient patch size (the relative source
sink ratio) on hybridisation rate; the fitness of hybrid seedlings in terms of their
ability to establish; and the fitness of exotic hybrids in terms of survival.

1.5.4 Chapter 5: Assessing the invasive potential of Eucalyptus globulus in
Australia: quantification of wildling establishment from plantations

This chapter assess seed-mediated dispersal of E. globulus from plantations by
undertaking surveys for wildlings (plantation derived E. globulus seedlings)
establishing around plantations at two geographic scales across the plantation estate.
A broad-scale survey looks at the regional and landscape level factors associated
with establishment and spread, while a fine-scale survey simultaneously assesses
local microsite factors important for establishment in areas with high densities of wildling occurrence.

1.5.5 Chapter 6: General discussion

This chapter summarises the key findings, address some issues not covered in the thesis, and places the results in the context of other threats to native eucalypt forests in Australia.
Chapter 2: Phylogenetic patterns of reproductive isolation in *Eucalyptus*

2.1 Abstract

Reproductive isolation is a fundamental characteristic of speciation. This study assesses patterns of reproductive isolation in *Eucalyptus* by investigating the capacity for hybridisation within and between lineages. *Eucalyptus globulus* pollen was applied to 100 eucalypt species from 13 taxonomic sections, mainly from the most speciose and commercially important subgenus, *Symphyomyrtus*. Compatibility was assessed at two time points, one representing pre-dispersal hybrid success, and the second post-dispersal hybrid survival at 9 months (minimum). This data was combined with phylogenetic data (based on over 8000 genome-wide DArT markers) to identify genetically controlled patterns in hybridisation, and assess the relative importance of pre- and post-dispersal barriers to reproductive isolation. Crossing success was higher within the closely related clade I (which includes *E. globulus*) and clade II (92% of taxa producing hybrids) than with the phylogenetically more distal clades III and IV (10% of taxa producing hybrids). This suggests significant reproductive barriers exist between the more divergent clades within *Symphyomyrtus*. In general, hybrid compatibility declined with increasing genetic distance between parents. Patterns of hybridisation indicate that both pre- and postzygotic barriers influence post-mating isolation, and may be driven by different processes, possibly including both natural selection and drift. Comparing the results with the most recently published dated phylogeny indicates that the time taken for complete reproductive isolation to develop in *Eucalyptus* may be 21-31 million
years. The study has practical implications for hybrid breeding in *Eucalyptus*, as well as for quantifying the genetic risk that *E. globulus* plantations pose to indigenous eucalypts in Australia.

### 2.2 Introduction

Reproductive barriers that evolve to prevent hybridisation between previously cross-compatible taxa are fundamental drivers of speciation (Coyne and Orr 2004). Comparative studies that assess patterns of reproductive isolation among groups of related taxa have been important in identifying when these barriers develop, and determining the relative contribution of different barriers to reproductive isolation (Coyne and Orr 1989; Moyle *et al.* 2004; Orr and Turelli 2001; Widmer *et al.* 2009).

It has been shown that pre-mating barriers, such as mate choice or flowering time, usually develop first, are often responsible for most of the observed isolation between taxa, and because they experience steep selection gradients as a result of their direct impact on reproductive success, they tend to evolve more quickly than post-mating barriers (Coyne and Orr 2004; Lowry *et al.* 2008; Widmer *et al.* 2009). However, pre-mating barriers also tend to be “leaky”, and it is often a combination of pre- and post-mating barriers that result in actual isolation (Coyne and Orr 2004; Widmer *et al.* 2009).

Post-mating barriers include both pre- and postzygotic mechanisms that are thought to arise principally via drift in a clock-like fashion (Coyne and Orr 2004; Hogenboom and Mather 1975; Orr and Turelli 2001). Theory suggests that as populations diverge minor allelic changes develop that are neutral or adaptive in the population of origin, but cause epistatic incompatibilities when brought together in inter-population hybrids – leading to inviability or sterility (a process widely known as the Dobzhansky-Muller model; Bateson 1909; Dobzhansky 1937; Muller 1942; Orr and Turelli 2001). In animal systems there is abundant evidence for Dobzhansky-Muller incompatibilities. For example, reproductive isolation consistently increases with genetic distance (Coyne and Orr 1989; Lee 2000; Malone and Fontenot 2008; Presgraves 2002), and male hybrid sterility has been found to involve hundreds of genes (Lu *et al.* 2010), both of which are indicative of accumulating incompatibilities. However, the situation in plants is less clear-cut
Chapter 2: Reproductive isolation

(Lowry et al. 2008; Widmer et al. 2009). Some studies show a correlation between genetic distance and postzygotic isolation consistent with Dobzhansky-Muller type incompatibles (Jewell et al. 2012; Meiners and Winkelmann 2012; Scopece et al. 2007), while others studies have found no clear evidence for epistatic mechanisms (Widmer et al. 2009). For example, isolation can developed quickly as a result of selection acting on post-mating prezygotic barriers (Giraud and Gourbière 2012; Scopece et al. 2007), or genomic rearrangements and very simple gene incompatibilities can cause instantaneous and strong reproductive isolation (Nosrati et al. 2011; Ramsey et al. 2003; Sweigart et al. 2006). Additionally, allopolyploidy can lead to rapid unpredictable patterns of speciation in plants (Abbott et al. 2013; Levin 2013). It has recently been suggested that simple Dobzhansky-Muller models are inadequate for explaining all the patterns of post-mating isolation emerging, particularly in plants, and more research is needed identify how the model should be broadened (Abbott et al. 2013; Giraud and Gourbière 2012).

Another key point of difference between animal and plant systems is the speed at which pre- and post-mating isolation develops, with pre-mating barriers evolving more quickly than post-mating barriers in animals (Coyne and Orr 1989; Coyne and Orr 1997; Mendelson 2003), but not plants (Moyle et al. 2004; Scopece et al. 2007; Scopece et al. 2008). It has been suggested that there may be more barriers affecting isolation between plant taxa (Widmer et al. 2009), but it may also be related to the fact that post-mating barriers are typically assessed only at one, often early, ontogenetic stage (but see: Scopece et al. 2007; Scopece et al. 2008; Stelkens et al. 2010). This may significantly underestimate hybrid incompatibilities that are expressed later in development – affecting estimates of both their strength and rate of evolution (Stelkens et al. 2010). Related to this, studies investigating the phylogenetic basis of reproductive isolation in plants have to date focused on herbs. But, as noted recently by Levin (2012), future studies need to include trees, because long generation times extend the period when hybrid incompatibilities can be expressed prior to reproduction (often 5-10 years), which potentially represents a life-history barrier that annual flowering herbs do not experience.
Chapter 2: Reproductive isolation

Recently methods have been developed to use model fitting in order to determine support for different modes of evolution, including the Dobzhansky-Muller model (Giraud and Gourbière 2012; Gourbière and Mallet 2010). Gourbière and Mallet (2010) proposed three relationships between the log of compatibility and time since divergence that provide support for different modes of evolution. The first provides support for Dobzhansky-Muller “snowball model” where the number of incompatibilities accelerate relative to the number of mutations because of epitasis (Gourbière and Mallet 2010; Orr and Turelli 2001). The second produces a “linear model” supporting a mode of evolution where epitasis does not influenced the rate of increase in incompatibilities through time (Gourbière and Mallet 2010). The third produces a “slowdown model” which is indicative of natural selection, specifically reinforcement, acting to cause reproductive isolation, and produces the opposite pattern to the “snowball model” (Gourbière and Mallet 2010). This method has now been deployed across a range of taxa, and there is often support for the “slowdown” and “linear” models but only rarely is there evidence for the “snowball” effect (Giraud and Gourbière 2012; Gourbière and Mallet 2010).

_Eucalyptus_ is a large and diverse genus of around 700 species of trees and shrubs that are native mainly to Australia (Euclid 2006). They are often foundation species and dominate most Australian forests and woodlands, making _Eucalyptus_ one of the most ecologically important genera in Australia (Pryor and Johnson 1981). There is great morphological diversity in the genus including the world’s tallest angiosperm _E. regnans_ (99.6m), and the prostrate mountain-top shrub _E. vernicosa_ (Grattapaglia _et al._ 2012). It is a taxonomically complex group that is divided into 10 subgenera, with the most speciose being _Symphyomyrtus_ (c. 470 species) and _Eucalyptus_ (formally _Monocalyptus_, c. 108 species) (Grattapaglia _et al._ 2012). There is no evidence for polyploidy in the genus (2n = 22 in all 135 species assessed to date), making it an ideal lineage for investigating patterns of diploid speciation (Grattapaglia _et al._ 2012).

There is a well recognised complete barrier to hybridisation between the major _Eucalyptus_ subgenera (Griffin _et al._ 1988; Pryor and Johnson 1981). However, widely reported hybridisation within subgenera has led to _Eucalyptus_ being
renowned for having weak reproductive barriers between species (Field et al. 2009; Grattapaglia et al. 2012). Despite this reputation, a major review of hybridisation in the genus found that barriers clearly exist (Griffin et al. 1988). Griffin et al. (1988) showed that even within subgenera only 15% of hybrid combinations (expected based on range overlap) have actually been found, and that hybridisation within taxonomic sections was more common than between sections. There is clearly variation in the strength of barriers to hybridisation in *Eucalyptus*, making it a good candidate for investigating the phylogenetic basis of post-mating isolation.

Eucalypts are also the world’s most widely grown hardwood trees, with 20 million ha of plantations under cultivation in over 100 countries across six continents (GIT 2013). Their adaptability and fast growth means eucalypts are likely to be a prominent source of biomass for future sustainable energy production (Shepherd et al. 2011) and eucalypt plantations already reduce anthropogenic pressure on native forests and biodiversity (Bauhus et al. 2010). Despite great taxonomic diversity, the global eucalypt plantation estate is dominated by nine species and their hybrids, all of which come from the subgenus *Symphyomyrtus* (Harwood 2011). The superior pulpwood properties of *Eucalyptus globulus* have seen it become the most widely grown species in temperate zones (Stackpole et al. 2010). As a result it has also become a significant component of hybrid breeding programs to improve wood properties in the tropics and subtropics (Bison et al. 2007; Hardner et al. 2011). A better understanding of phylogenetic barriers to hybridisation will benefit hybrid breeding programs in this economically important genus.

The ecological and economic significance of *Eucalyptus* has seen the development of significant genomic recourses (Grattapaglia et al. 2012). Recent developments include diversity arrays technology (DArT), which is a marker based system providing thousands of genome-wide polymorphic loci (Petroli et al. 2012; Sansaloni et al. 2010; Steane et al. 2011). Levin (2012) noted that the small number of genes used to estimate divergence in studies of plant reproductive isolation may not be representative of divergence across the genome as a whole, possibly causing problems in interpretation. The application of DArT markers would largely overcome this problem.
Chapter 2: Reproductive isolation

This study uses over 8000 DArT markers to build a distance based phylogeny to investigate the phylogenetic basis of cross-compatibility between *E. globulus* and 100 eucalypt species belonging to 13 sections, mainly from subgenus *Symphyomyrtus*. We have chosen *E. globulus* because of its global economic importance, and because it is the most widely grown eucalypt in Australia where it is often locally exotic, and could affect the genetic integrity of indigenous eucalypt populations through hybridisation and introgression (Barbour *et al.* 2008b; Potts *et al.* 2003). We assess two stages of post-mating compatibility: the number of viable hybrids seeds produced, herein termed pre-dispersal compatibility; and F1 survival at nine months, herein termed post-dispersal compatibility (see methods for detailed definitions). The following specific questions are addressed: 1) does incompatibility increase with genetic distance, and if it does, is the increase consistent with a Dobzhansky-Muller model of isolation? 2) are patterns of pre- and post-dispersal isolation similar in strength indicating that they evolve at a similar rate?

2.3 Materials and Methods

2.3.1 Crossing

Pollen was collected from 14 unrelated *E. globulus* trees to provide a broad genetic base for the crossing program. These included two native forest trees in south east Tasmania, and 10 trees from a Seed Energy Pty Ltd seed orchid at Cambridge, Tasmania (described in Jones *et al.* 2011; also see Electronic Supplementary Material 2.1, on the CD in the back of the thesis, for more details regarding these trees). Pollen was extracted, and viability tested using the techniques of Potts and Marsden-Smedley (1989), except that 150ppm boric acid was used in the agar medium during viability testing. All individual tree pollens were kept separate to allow parentage analysis of hybrids.

In total 100 species were crossed, with *E. globulus* always used as the pollen parent. The majority of the crossing was undertaken at a specialist eucalypt research arboretum in South Australia (Currency Creek Arboretum – CCA; www.dn.com.au). Additionally, for 13 species, native forest and/or ornamental trees in Tasmania, Western Australia and South Australia were crossed (see Electronic Supplementary
Chapter 2: Reproductive isolation

Material 2.1). Crossing was undertaken between May 2010 and May 2011, with one to five trees per species crossed. Two approaches were taken in the crossing. The first approach was designed to mimic as closely as possible natural pollination, and involved simply dabbing pollen onto the receptive stigma of open flowers (“supplementary” pollination; Barbour et al. 2005a). The second approach was aimed at overcoming two potential issues, sigma receptivity, and stigma incompatibilities (Boland and Sedgley 1986; de Sousa and Pinto-Junior 1994; Oddie and McComb 1998). This involved removing the stigma with a razor blade and applying pollen directly to the surface of the cut style (“cut-style” pollination; Cauvin 1988; Patterson et al. 2004a). On each tree seven treatments were applied, each to a different branch (mean flowers/branch = 9.6, number of treated branches = 1039). Treatments 1 to 3 were supplementary pollinations each using different E. globulus pollen. Treatment 4 was cut-style pollination applied to open flowers in the same state as those pollinated in treatments 1 to 3. Treatment 5 aimed to exclude pollen from non-target species and involved cut-style with the following additions: pollination was carried out on flowers prior to anthesis by removing the operculum, cutting the style, applying the pollen, and isolating the pollinated flowers in wax paper bags for at least 3 months (unless the bag failed). Treatments 6 and 7 were open pollinated controls used to judge whether any failure to produce seed on the treatment branches was influenced by tree sterility (flower abortion), and also to provide pure species specimens for morphological comparison with hybrid progeny.

There is evidence that pollen tube growth in Eucalyptus is correlated with flower size and that it can affect the success of hybridisation between species when there are large differences in style length, particularly when small flowered pollen is applied to large flowered species (Gore et al. 1990). To test the effect of flower size on crossing success at least three flowers from each species were dissected and measurements from the stigma to the top of the ovaries, and the stigma to the base of the ovaries were made. It should be noted that no control for the act of pollination was used (i.e. dabbing the stigma without the pollen), however treatment controls were used and the pollination techniques employed in this study have been widely used in studies of eucalypts and no adverse effects of these techniques have been reported (Cauvin 1988; Patterson et al. 2004a; Barbour et al 2005). Furthermore, the
final analysis is based only on successful crosses (see below), minimising the influence of the pollination technique on seed set.

Capsules were collected when they were judged to be mature based on having well developed valves (Barbour et al. 2005a), which was between 10 and 26 months after pollination. Capsules were dried and the seed was extracted and stored with a small amount of the powdered fungicide Mancozeb. Due to the large number of seedlots (519), sowing was done in batches over 12 months between January 2012 and January 2013. Seed lots were randomised within and between batches, except that each batch was designed to have about equal representation of each taxonomic section. Seeds were sown in boxes (450 x 295 x 250 mm) on to moist potting mix and the soil surface was covered with vermiculite to reduce the possibility of seed jumps during watering prior to germination. Germination was undertaken in a controlled environment facility under a light regime of 16 hours light to 8 hours dark at 22°C and 70% humidity. Three weeks after germination the boxes were moved into a conventional glasshouse and/or outside to grow on. Hybrids were initially identified based on morphological deviation from the pure parental type, and subsequently validated with parentage analysis using 10 microsatellite loci (see Appendix 2.1 and 2.2; and Electronic Supplementary Material 2.1). There was only one cross type where the parental seedling morphology made identification of hybrids somewhat ambiguous (E. globulus x mannifera). In this case all 28 ambiguous progeny were assessed with parentage analysis (Electronic Supplementary Material 2.1). In all other situations a conservative approach was taken to morphological deviation from the pure phenotype with any suspect plants checked with parentage analysis, thus the false negative rate of the morphological assessments was assumed to be zero. If overcrowding was affecting the health of plants in a box, pure (non-hybrid) samples were removed to reduce competition as soon as they could be distinguished from hybrids. All hybrids that survived and representatives of the pure species were potted on to individual pots after scoring with hybrid and pure seedlings placed in the same size pots. All pots were randomised and moved outside with hybrid and pure seedlings always moved at the same time. Survival of the potted hybrids was assessed in September 2013. The staggered sowing meant that time since germination ranged from 9 to 21 months at
the time of the survival assessment but the taxonomically randomised sowing approach (see above) ensured there was no taxonomic bias in the survival estimates. Therefore two measures were used to assess compatibility at different life history stages: firstly the initial assessment made at the early seedling stage which assess barriers from pollination to viable seed production (pre-dispersal); and secondly hybrid survival at nine months which assess barriers to F₁ seedling survival (post-dispersal).

2.3.2 Phylogenetic and statistical methods

To investigate phylogenetic patterns in hybridisation, as well as the correlation between hybridisation rate and genetic distance, we used two datasets derived from genome-wide scans (using Diversity Arrays Technology - DArTs; Sansaloni et al. 2010). The DArT approach produces hyper-variable dominant molecular markers that assay genome wide patterns in genetic diversity (Petroli et al. 2012). The first dataset was used to calculate the average genetic distance between E. globulus and each of the 13 taxonomic sections covered in the crossing program, and is referred to as the “section-level dataset” from here on. This section-level dataset included 8,350 markers and has previously been published by Steane et al. (2011), with sample collection, DNA analysis and genotyping explained in that publication. Of the 94 species used by Steane et al. (2011), 78 belong to the taxonomic sections used in our experiment (Electronic Supplementary Material 2.1). Part of the distance matrix produced by Steane et al. (2011) that corresponded to the sections used in the crossing with E. globulus was extracted and used to produce a phylogenetic network (see below). It should be made clear that the species in this phylogenetic analysis do not correspond exactly to the species used in this study, with 22 of the 78 species being coincident (Electronic Supplementary Material 2.1). Taxonomy at the sectional level is typically well resolved in Eucalyptus (Byrne 2008) making analysis of relationships at this level reasonable.

The second DArT dataset was used to calculate genetic distance between E. globulus and the 21 species that were crossed (and produced viable seed) in five taxonomic sections that are most closely related to E. globulus (Grattapaglia et al. 2012) including Maidenaria (the section to which E. globulus belongs), Exsertaria,
Incognitae, Latoangulatae and Racemus. This dataset is here in referred to as the “species-level dataset”. The species-level dataset is a subset of a large dataset that consists of 558 samples including 191 taxa genotyped with 5050 DArT markers, which is to be published in full elsewhere (R.C. Jones in prep.), with the DNA analysis and genotyping following Steane et al. (2011). Pairwise genetic distance (Additive Dollo Distance [ADD], a measure developed specifically for DArTs; Woodhams et al. 2013) between E. globulus and the species of interest were extracted (and averaged where multiple samples were available; see Electronic Supplementary Material 2.1) from the complete genetic distance matrix.

A phylogenetic network (using ADD; Woodhams et al. 2013) was produced from the section-level dataset (of 78 species) in Splitstree4 (Huson and Bryant 2006). Clades were identified based on the topology of the network and comparisons with a similar analysis (Grattapaglia et al. 2012). The location of taxonomic sections within the clades was identified using the most recent classification of Eucalyptus (D. Nicolle in prep.). This analysis was undertaken on species, and the data for three subspecies was merged to the species level (the full subspecies data is available in Electronic Supplementary Material 2.1). All species that produced viable seed (64 species) were then assigned to positions on the phylogenetic network (i.e. their taxonomic section), to reveal patterns of hybridisation (see Fig. 2.1). Species that failed to produce viable seed were excluded from the analysis. A Chi-squared test was used to compare the number of hybridising species in the different clades (pooling both supplementary and cut style pollination techniques). Pre-dispersal crossability was calculated between E. globulus and various points on the phylogenetic network by averaging the cross success (number of hybrids/number of progeny) for all species in the clade(s) of interest. A second crossability estimate that accounted for hybrid survival was produced by multiplying the first estimate by the average hybrid survival in the clade(s). Both these estimates were adjusted by the known intra-specific cross-compatibility in E. globulus when using the same supplementary crossing approach (95% success rate for applied pollen; Patterson et al. 2004b).
A series of Generalised Linear Models (GLMs) were used to assess the relationship between genetic distance and hybridisation. Firstly, the relationship between genetic distance (ADD) and pre-dispersal hybridisation rate (the number hybrids/number plants produced) was tested considering only supplementary pollination so as to estimate the “natural” hybridisation potential (i.e. without cutting the style). Secondly, the relationship between ADD and hybrid survival was tested, considering hybrids from both supplementary and cut style pollination. Thirdly, an estimate of the combined effect of hybridisation rate and survival on the overall compatibility was compared to ADD. The combined rate was calculated by multiplying the number of hybrids produced by the probability of survival and adjusting that number by the known intra-specific cross-compatibility in *E. globulus* (95%; Patterson et al. 2004b). Quasi-binomial or quasi-Poisson distributions were used to account for overdispersion (excess of zeros) where necessary (Zuur et al. 2009). Logistic regression fits with logit link functions were used to model all three isolation levels (i.e. pre-dispersal, post-dispersal and combined) for both the section-level and species-level datasets. The average of both flower-size measurements (stigma top to ovary top, and stigma top ovary bottom) and their interaction were included as covariates in the logistic models for the species-level dataset. Model simplification was undertaken by removing non-significant terms starting with the higher order interactions.

A second round of modelling was undertaken to determine if the patterns in hybridisation provide evidence for the mode of evolution as outlined by Gourbiere and Mallet (2010). Models developed by Gourbiere and Mallet (2010) to determine support for the three modes of evolution were fitted to the crossing data here. The pre-dispersal, post-dispersal and combined estimates were again used as measures of compatibility, with both the section-level and species-level datasets tested. The models fitted were (notation follows: Giraud and Gourbière 2012; Gourbière and Mallet 2010):

**Linear:** $E(\ln C_T) = e_1T$

**Snowball:** $E(\ln C_T) = e_2T^2$
Slowdown: $E(\ln C_T) = e_3 \ln(1 + aT)/a$

where $T = \text{ADD}$ (a proxy for time since divergence); $C_T =$ compatibility; $e_1 = \Sigma(T \log \text{compatibility})/\Sigma(T^2)$; $e_2 = \Sigma(T^2 \log \text{compatibility})/\Sigma(T^4)$; $a$ is a parameter that controls the shape of the curve, and $a$ and $e_3$ are both found by optimising the least squares criterion, rather than having closed form solutions of their own (see Gourbière and Mallet 2010). The residual standard error was used to determine which model was the best fit to the data. The significance of any difference between model fits was tested by comparing the ratio of residual standard errors between the models (best versus worst) using the $F$ statistic. Model fitting was undertaken in R version 2.14.1 (R Development Core Team 2011).

**Results**

Of the 100 species crossed, 64 produced viable seed (treatment or control), and from the crossing treatments 4571 progeny were assessed. In all, 616 *E. globulus* hybrids were identified (Table 2.1) and 42 natural hybrids were found (where morphology deviated from the maternal type but was inconsistent with *E. globulus* hybrids). The cut-style pollination technique was more efficient at producing hybrids, with a 25% success rate (number of hybrids/number of plants produced) compared to the supplementary pollination at 14%. *Eucalyptus globulus* hybrids were found in 26 (41%) of the 64 species that produced seed; these were all verified using microsatellite based parentage analysis (see Appendix 2.2 for details). Hybrid survival was assessed for 215 of the hybrids, representing all 26 species. Seventy-five (35%) of the hybrids died by the time they were evaluated, and complete hybrid mortality occurred in five species.

The phylogenetic network of 78 species showed four main clades within subgenus *Symphyomyrtus* (Fig. 2.1). These clades had previously been identified in a similar analysis (Steane et al. 2011). Clade I is made up of species from section *Maidenaria* which includes *E. globulus*. Clade II is most closely associated with clade I and includes species from sections *Exsertaria, Incognitae, Latoangulatae* and *Racemus*. There is a relatively long branch between clades I /II and clades III/IV. Clade III includes species from sections *Adnataria, Diversities* (previously Bisectae I; Steane
et al. 2011) and Dumaria, while clade IV is made up of species from Bisectaria (previously Bisectae II; Steane et al. 2011).

**Table 2.1.** Summary of crossing between Eucalyptus globulus (pollen parent) and 100 eucalypt species from 13 taxonomic sections, mainly from subgenus Symphyomyrtus. A complete list of the species crossed and individual results are provided in Supplementary Material.

<table>
<thead>
<tr>
<th>Clade&lt;sup&gt;a&lt;/sup&gt;</th>
<th>n of sections represented&lt;sup&gt;c&lt;/sup&gt;</th>
<th>n species crossed</th>
<th>n flowers crossed</th>
<th>n species producing viable seed</th>
<th>n species producing hybrids</th>
<th>n hybrids</th>
<th>% hybrid survival&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1</td>
<td>21</td>
<td>1,391</td>
<td>13</td>
<td>12</td>
<td>544</td>
<td>80</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>11</td>
<td>1,058</td>
<td>11</td>
<td>10</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>49</td>
<td>3,330</td>
<td>31</td>
<td>3</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
<td>15</td>
<td>1,107</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>V&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3</td>
<td>4</td>
<td>171</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>7,057</td>
<td>64</td>
<td>26</td>
<td>616</td>
<td>-</td>
</tr>
</tbody>
</table>

<sup>a</sup>Clades are shown in Fig. 2.1; <sup>b</sup>species outside subgenus Symphyomyrtus; <sup>c</sup>taxonomy according to D. Nicolle (in prep.); <sup>d</sup>survival was assessed at least nine months after germination.

There was a phylogenetic pattern to hybridisation success (Fig. 2.1). Chi-square tests showed that E. globulus is significantly more likely to hybridise with taxa in clades I (its own clade) and II (21 species), than those in clades III and IV (4 species; \(\chi^2 = 23.9, P < 0.0001\); Fig. 2.1). However, there is no difference in the number species forming hybrids with E. globulus in clades I and II (\(\chi^2 = 0.17, P > 0.99\)).

When assessing pre-dispersal hybridisation rate (number of hybrids/number of plants produced) for both the section-level (Fig. 2.2a) and species-level datasets (Fig. 2.2d), there was a strong negative effect of genetic distance (Fig. 2.2; Table 2.2). This pre-dispersal relationship is consistently negative, but the rate of decrease in compatibility is high at low genetic distances, and becomes lower at higher genetic distance (Fig. 2.2a/d). Post-dispersal compatibility (hybrid survival) measured for both the section-level (Fig. 2.2c) and species-level datasets (Fig. 2.2e; Table 2.2), also declined with increasing genetic distance, but the relationship was more linear than that found in the pre-dispersal estimates. The shape of the pre-dispersal compared to post-dispersal curves in both datasets indicates that pre-dispersal barriers develop more quickly (i.e. at lower genetic distances) and are stronger than...
post-dispersal barriers. Combining the pre- and post-dispersal estimates of compatibility improved the fit for both datasets (Fig. 2.2c/f; Table 2.2). Finally, despite considerable variation in flower size, there was no detectable effect of style length ($t_{22} = -0.003$, $P = 0.99$), or style length + ovary depth ($t_{22} = 0.166$, $P = 0.87$).

**Table 2.2.** Logistic regression models explaining the relationship between crossability (pre-dispersal, post-dispersal and combined) and genetic distance between *E. globulus* and the section-level and species-level datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Reproductive isolation stage</th>
<th>Estimate (log-probability)</th>
<th>df</th>
<th>$P$</th>
<th>Deviance explained by genetic distance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sectional-level</td>
<td>Pre-dispersal</td>
<td>-3.283</td>
<td>10</td>
<td>&lt;0.0001</td>
<td>90.3</td>
</tr>
<tr>
<td>Sectional-level</td>
<td>Post-dispersal</td>
<td>-5.201</td>
<td>7</td>
<td>&lt;0.0001</td>
<td>75.4</td>
</tr>
<tr>
<td>Sectional-level</td>
<td>Combined</td>
<td>-1.492</td>
<td>10</td>
<td>&lt;0.0001</td>
<td>97.1</td>
</tr>
<tr>
<td>Species-level</td>
<td>Pre-dispersal</td>
<td>-11.326</td>
<td>22</td>
<td>0.019</td>
<td>69.3</td>
</tr>
<tr>
<td>Species-level</td>
<td>Post-dispersal</td>
<td>-1.588</td>
<td>20</td>
<td>&lt;0.0001</td>
<td>42.8</td>
</tr>
<tr>
<td>Species-level</td>
<td>combined</td>
<td>-6.071</td>
<td>22</td>
<td>&lt;0.0001</td>
<td>76.1</td>
</tr>
</tbody>
</table>

* Genetic distance = log(AAD) see methods
Figure 2.1. Top: the phylogenetic network based on 78 species and 8350 markers shows four main clades in *Symphyomyrtus*, which are annotated with the most recent sectional taxonomy (D. Nicole in prep.). The estimated crossability between *E. globulus* and various parts of the phylogenetic network are shown (arrows and dashed lines), the first number gives the probability of producing hybrids, and in parentheses is the combined probability of producing hybrids and those hybrids surviving to nine months. The histogram shows the percentage of hybrids found in 64 species pollinated with *E. globulus* pollen. The species are coloured according to their phylogenetic affinities to clades above (species are ranked by hybridisation rate, and then clades based on genetic distance to *E. globulus*). Species with an asterisk did not produce hybrids with supplementary pollination but did with cut-style pollination (see methods). † indicates complete hybrid mortality.
Table 2.3. Models assessing the mode of evolution (linear, snowball and slowdown; see text) in pre-dispersal, post-dispersal and combined reproductive compatibility. Models compared log-compatibility and genetic distance (ADD) between *E. globulus* and either the section-level or species-level datasets. The models are presented in order from best to worst fit according the residual standard error (Error). The $F$ statistic and $P$ value quantify the difference in residual standard error between the best and worst model only.

<table>
<thead>
<tr>
<th>Comparison/dataset</th>
<th>Model</th>
<th>Error</th>
<th>df</th>
<th>$F$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-dispersal/section-level</td>
<td>slowdown</td>
<td>1.279</td>
<td>10</td>
<td>2.385</td>
<td>0.0909</td>
</tr>
<tr>
<td></td>
<td>linear</td>
<td>1.788</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>snowball</td>
<td>3.051</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-dispersal/section-level</td>
<td>snowball</td>
<td>1.731</td>
<td>8</td>
<td>1.371</td>
<td>0.3319</td>
</tr>
<tr>
<td></td>
<td>linear</td>
<td>2.05</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slowdown</td>
<td>2.374</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined/section-level</td>
<td>snowball</td>
<td>1.253</td>
<td>10</td>
<td>2.830</td>
<td>0.0562</td>
</tr>
<tr>
<td></td>
<td>linear</td>
<td>2.1</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slowdown</td>
<td>3.546</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-dispersal/species-level</td>
<td>linear</td>
<td>2.222</td>
<td>23</td>
<td>1.034</td>
<td>0.4682</td>
</tr>
<tr>
<td></td>
<td>snowball</td>
<td>2.298</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slowdown</td>
<td>3.564</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-dispersal/species-level</td>
<td>snowball</td>
<td>1.274</td>
<td>21</td>
<td>1.055</td>
<td>0.4503</td>
</tr>
<tr>
<td></td>
<td>linear</td>
<td>1.298</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slowdown</td>
<td>1.345</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined/species-level</td>
<td>linear</td>
<td>2.132</td>
<td>23</td>
<td>1.057</td>
<td>0.4473</td>
</tr>
<tr>
<td></td>
<td>snowball</td>
<td>2.228</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slowdown</td>
<td>2.253</td>
<td>22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Model failed to converge

Fitting of the three alternative models (following Gourbière and Mallet 2010) showed that for the section-level dataset (Fig. 2.3a-c) the slowdown model fitted the pre-dispersal data better (Table 2.3) while the snowball model fitted the post-dispersal data better (Table 2.3), and in the combined dataset the slowdown model again fitted best (Table 2.3). While for the species-level dataset, the linear model fits the pre-dispersal data best (Table 2.3), the snowball again fitted the post-dispersal data better (Table 2.3), and the linear model fitted the combined data best (Table 2.3).
2.3). However, none of the models could be statistically separated from the alternate models at $P = 0.05$ with this data, although there was some support at the $P = 0.1$ level. For the full dataset, the slowdown model was significantly better (at the 0.1 level; Table 2.3) than the snowball for both the pre-dispersal ($F_{10,11} = 2.39, P = 0.0909$; Table 2.3) and combined comparisons ($F_{10,11} = 2.83, P = 0.0562$; Table 2.3).

**Figure 2.2.** The relationship (mean and 95% confidence intervals) between genetic distance (ADD: Additive Dollo Distance) and hybridisation rate between *E. globulus* and the section-level and species-level datasets. The two genetic distances are derived from different DArT datasets and are not comparable (because different subsets of markers from the DArT array provided phylogenetic resolution at different taxonomic levels).
2.4 Discussion

2.4.1 The evolution of post mating barriers

There is evidence that Dobzhansky-Muller incompatibilities lead to post-mating isolation in a range of taxa (Coyne and Orr 2004; Hogenboom and Mather 1975; Orr and Turelli 2001; Turelli and Moyle 2007). A key property of the Dobzhansky-Muller model, as demonstrated by Orr and Turelli (2001), is that the rate of increase for hybrid incompatibilities is exactly the square root of time since divergence.
Therefore, using genetic distance as a proxy for time, Dobzhansky-Muller incompatibilities should initially develop slowly, but accelerate relative to genetic distance because of the increasing complexity of epistatic interactions, producing a “snowball effect” (Orr and Turelli 2001). We do not have strong statistical support for any of the models testing the mode of evolution, but the general patterns and weak support for certain models, may provide some insight into the evolution of reproductive isolation in *Eucalyptus*. The relationship found here between pre-dispersal reproductive isolation and genetic divergence in *Eucalyptus* (in the section-level dataset) is best fitted by the slowdown model, with the snowball being the worst fit. This is a pattern indicative of reinforcement driven by selection (Gourbière and Mallet 2010; Hopkins 2013). Reinforcement creates an initially steep selection gradient that promotes assortative mating and prevents the formation of unfit hybrids, but because selection pressure decreases when gene flow is restricted, the rate at which incompatibilities develop also decrease (Gourbière and Mallet 2010; Hopkins 2013). This produces a “slowdown” effect consistent with the results here in *Eucalyptus* and with a range of other taxa (Giraud and Gourbière 2012; Gourbière and Mallet 2010).

It is not surprising that post-mating prezygotic barriers should be influenced by selection because they directly affect reproductive success. However, it is often thought that post-mating prezygotic barriers in plants are largely controlled by incongruity (De Nettancourt 2001; Hogenboom and Mather 1975; Hogenboom 1984). Incongruity is a specific type of Dobzhansky-Muller incompatibility that would also be expected to produce a “snowball” response (De Nettancourt 2001; Hogenboom and Mather 1975; Hogenboom 1984; Mangum and Peffley 2005). Incongruity refers to the breakdown of intercellular signalling between sporophyte and gametophyte tissues due to the accumulation of minor genetic differences through drift between diverging parental populations (Hogenboom 1984). Like reinforcement this is a prezygotic mechanism, but the epistatic nature of accumulating differences means incongruity will result in accelerating rates of reproductive isolation. This is distinct from reinforcement where rates at which incompatibilities develop decreases as gene flow is restricted (Gourbière and Mallet 2010; Hopkins 2013). There is evidence in eucalypts that post-mating prezygotic
barriers are important components of post-mating isolation (Dickinson et al. 2012b; Ellis et al. 1991; Sedgley and Granger 1996), and incongruity has been used to explain the evolution of these mechanisms (Dickinson et al. 2012b; Ellis et al. 1991; Gore et al. 1990). The results here suggest that reinforcement may also play a role in the development of pre-dispersal barriers in *Eucalyptus*.

Pre-dispersal barriers also seem to develop more quickly than post-dispersal barriers in *Eucalyptus* (Fig. 2.2) which is common in animals, but is rarely observed in plants (Abbott et al. 2013; Hopkins 2013; Widmer et al. 2009). This may be further evidence of reinforcement. In sympatric species reinforcement is expected to cause prezygotic barriers to develop more quickly than post-zygotic barriers, because the former are under relatively strong selection (Coyne and Orr 1989; Dobzhansky 1940; Hopkins 2013). Testing this prediction requires the comparison of allopatric and sympatric species (Hopkins 2013), which is not directly possible with our data. However, given the diversity and ecological dominance of *Eucalyptus*, combined with the effect of fluctuating climate on the distribution and evolution of the group (Byrne 2008), it is likely that many species have experienced phases of parapatry/sympatry during their evolution. Therefore, it is possible that barriers that evolved as a result of reinforcement with other conspecifics are expressed here when crossed with *E. globulus*.

The post-dispersal barriers assessed here appear to develop more in line with a Dobzhansky-Muller model. In both datasets, the snowball model fitted the data better than the alternative models, with the slowdown being worst fit in both cases. This seemingly opposite trend between pre- and post-dispersal isolation indicates different modes of evolution. The more clock-like evolution of post-dispersal barriers is consistent with the fact that unlike pre-dispersal barriers, they do not experience the steep selection gradients associated with sexual selection. This being said, despite a large number of studies assessing this relationship (mainly in animals), few have found a strong snowball effect, with the majority reporting some linear relationship (Coyne and Orr 2004; Coyne and Orr 1989; Coyne and Orr 1997; Malone and Fontenot 2008; Moyle et al. 2004). The low power here means that the results may also fit a linear model better, although with a simple visual assessment of
the plots it seems unlikely that the slowdown model could be the best fit. One of the possible causes of the “missing snowball” is that the genomic architecture of incompatibilities is more complicated than a simple substitution model (Coyne and Orr 2004; Gourbière and Mallet 2010; Widmer et al. 2009).

Whole genome sequencing in a range of plants is revealing the extent of chromosomal rearrangements (Lysák and Schubert 2013), and changes such as inversions, deletions and duplications could affect the expression of incompatibilities (Coyne and Orr 2004). Associated with this, simple heterozygote disadvantage resulting from rearrangements is predicted to cause a linear accumulation of incompatibilities (Coyne and Orr 2004; Giraud and Gourbière 2012; Gourbière and Mallet 2010). The genus *Eucalyptus* has a uniform karyotype number, but there is considerable genome size variation between species (Grattapaglia and Bradshaw 1994). For example, Grattapaglia and Bradshaw (1994) showed (using DNA quantification) that the *E. globulus* genome is about 20% smaller than the *Eucalyptus grandis* genome (section *Latoangulatae*, clade II). It is therefore likely that interspecific hybrids would show increased heterozygosity and possibly heterozygote disadvantage. It has also recently been revealed that *Eucalyptus* has the highest rate of gene duplication found to-date in plants (Myburg et al. in press). It is difficult to predict the effect of these duplications on the expression of Dobzhansky-Muller incompatibilities, but they could lead to more linear relationships between post-dispersal isolation and genetic distance.

Accounting for hybrid survival at one year improved the fit of the post-mating estimates, and provided addition information regarding the evolution of reproductive barriers. However, the estimates are still likely to be an underestimate of the strength of the post-dispersal incompatibility. Several studies have found selection against inter-specific eucalypt hybrids is expressed later in development (2-10 years after field planting; Barbour et al. 2006a; Lopez et al. 2000a). Even within section *Maidenaria* (clade I) this reduced juvenile fitness in hybrids is thought to limit gene flow and help maintain allopatric/sympatric species boundaries in nature (Chapter 4; Larcombe et al. 2014; Lopez et al. 2000a). The causal mechanism leading to reduced hybrids survival in *E. globulus* x *E. nitens* (*Maidenaria*) hybrids was recently
suggested to be negative additive x additive epitasis (Costa e Silva et al. 2012), which is consistent with the evolution of Dobzhansky-Muller incompatibilities and the observed post-dispersal results here. It is therefore likely that hybrid survival in more distant crosses (i.e. between clade I and II, or I and III/IV) will experience even stronger later age incompatibility than estimated here at nine months.

2.4.2 Hybridisation and the timing of speciation in Eucalyptus

In addition to the previously recognised complete barrier to hybridisation between subgenera in Eucalyptus (Griffin et al. 1988), this study has identified that significant post-mating barriers to hybridisation extended to phylogenetic clades within subgenera. These barriers become stronger as genetic divergence increases and are likely to result in a parallel increase in reproductive isolation. The results here clarify the genus-wide taxonomic patterns in hybridisation identified by Griffin et al. (1988), and inter-sectional patterns of pollen-pistil interactions found by Ellis et al. (1991). For example Ellis et al. (1991) found a pollen-pistil barrier between Bisectaria (clade III/IV) and Exsertaria (clade II) that was not present between Bisectaria and Adnataria (clade III), while Griffin et al. (1988) found higher levels of natural hybridisation between species within Maidenaria (clade I), than between co-occurring species from sections Maidenaria and Exsertaria. The phylogenetic approach taken here supports, quantifies, and clarifies these barriers identified based on taxonomy.

Eucalyptus is an ancient and diverse lineage, and this study shows that prezygotic reproductive barriers seem to develop relatively quickly between Eucalyptus species, but that when these barriers are overcome, hybridisation is still possible for some time. The time it takes for pairs of taxa to achieve complete reproductive isolation is a poorly understood, but important aspect of plant speciation (Levin 2012; Rieseberg and Willis 2007). Our ability to quantify crossability at various positions on the phylogenetic network presented in Fig. 2.1 allows us to roughly estimate the timing of reproductive isolation from dated phylogenies. The most recent molecular clock approach (Crisp et al. 2011), suggests that species divergence within sections in Symphyomyrtus occurred c. 3-10 mya, suggesting that it takes at least 3 million years for crossability to decline by half. Divergence between clades I (Maidenaria) and II
(Exsertaria) is estimated to have occurred c.10-15 mya (Crisp et al. 2011), indicating that it takes at least 10 million years for crossability to fall by about 95%. Finally, divergence between clades I/II and clades III/IV is dated to at 21-31 mya (Crisp et al. 2011), implying that complete reproductive isolation takes at least 21 million years in Eucalyptus. These comparisons are obviously approximate, but are analogous to the method used recently to produce the first wide ranging assessment of the timing of hybrid sterility in plants (Levin 2012). That study found a pronounced reduction in hybrid fertility after 4 million years, which is in-line with the 50% reduction within clade I here, and also that in trees partial cross-compatibility between species can be maintained for a very long time, up to 50 million years in Plantanus (Levin 2012).

2.4.3 Implications for tree breeding and the genetic risk posed by E. globulus

The estimates of crossability presented here will be of particular interest to tree breeders. These are the first broad estimates of crossability between sections across subgenus Symphyomyrtus, the most economically important and widely planted group in the genus (Grattapaglia et al. 2012). The estimates provide breeders with the ability to incorporate the likelihood of cross-success in selecting target species for hybrid breeding programs. Inter-specific eucalypt hybrids have been widely utilised in pulpwood breeding for 30 years (Grattapaglia and Kirst 2008), but the generation times of trees makes developing successful hybrid lines based on traditional approaches time consuming and expensive (Potts and Dungey 2004). Interestingly the crossing results from these long-term hybrid breeding programs match well with the results reported here. Hybrids between species in sections Latoangulatae and Exsertaria (both clade II, Fig. 2.1) are more widely utilised, and show fewer abnormalities than hybrids between clades I and II (Potts and Dungey 2004). Also some crosses within clade I (e.g. E. globulus x gunnii) show very few abnormalities (Potts and Dungey 2004), while others, that are now known to be more divergent (E. globulus x nitens), show morphological abnormalities (Potts and Dungey 2004) and epistatic incompatibilities (Costa e Silva et al. 2012). Therefore the patterns in hybridisation are broadly consistent with those observed between other eucalypt species in natural populations (i.e. see above; Ellis et al. 1991; Griffin
et al. 1988) and in long running hybrid breeding programs, suggesting that the
crossability estimates produced here using *E. globulus* could be indicative of general
trends between the clades.

There are around 538,000 ha of *E. globulus* plantations growing across southern
Australia, mainly outside the species native range (Gavran and Parsons 2011).
Concerns have been raised that these plantations may pose a genetic risk to
indigenous eucalypt species through pollen mediated gene flow and introgression
(Barbour et al. 2008b; Potts et al. 2003). Based on the well-known complete barrier
to hybridisation between subgenera, it was previously thought that any of the 484
*Symphyomyrtus* species that occurred within the pollen dispersal range of *E. globulus*
plantations were at risk (Barbour et al. 2008b; Potts et al. 2003). Our results suggest
that species in clades III and IV are reproductively isolated from *E. globulus*,
reducing the number of at-risk species to 138 (a 71% reduction). Furthermore
species in clade II are at a far lower risk of introgression than species within clade I,
and the crossability estimates presented will enable more informed genetic risk
assessment during plantation planning and establishment.

2.5  Conclusion

It appears that the evolution of post-mating isolation in *Eucalyptus* involves multiple
barriers that might be driven by both natural selection and drift. In particular, the
patterns of pre-dispersal hybridisation point to natural selection operating on
prezygotic barriers. In contrast, post-dispersal barriers develop in a more clock like
manner, in-line with Dobzhansky-Muller incompatibilities evolving via drift.
Consequently, endogenous post-dispersal barriers take longer to establish than pre-
dispersal barriers, which is consistent with many animal systems, but is not
commonly found in plants. The use of a phylogenetic approach has allowed the
estimation of the time it takes for various levels of reproductive isolation to evolve,
with complete isolation probably taking 21-31 million years. The study has also
shown that contrary to previous understanding there are a range of barriers operating
within subgenera in eucalypts, including some that result in practically complete
reproductive isolation. This has direct implications for the number of species at risk
from exotic gene flow in Australia, and will be valuable for tree breeding programs in forestry.
Appendix 2.1: Morphological characteristics of hybrids between *E. globulus* and 24 eucalypt species resulting from controlled crossing

The identification of hybrids in the crossing experiment relied initially upon the detection of hybrid seedlings based on morphology. *Eucalyptus globulus* has distinctive morphology with several characteristics often obvious in F₁ hybrids (Barbour *et al.* 2008; Larcombe *et al.* 2014a; Larcombe *et al.* 2014b; Lopez *et al.* 2000). These include sessile opposite leaves and a stem that is square to winged in cross section (Figure 1). Parentage analysis of all 26 species that produced hybrids was undertaken using microsatellites (see Supplementary Information 3). Below is a summary of the morphological characteristics used initially to identify 24 of the 26 species. There are no photographs of *E. nubila*, but the characteristics that led to the detection of the hybrids with this species were: opposite, sessile leaves and square stem in comparison to the pure species which had petiolate, alternate leaves and a round stem. There are no photographs of *E. viminalis* ssp. *cygnetensis* but the morphology was analogous to *E. viminalis* ssp. *viminalis*.

### Figure 1
Key juvenile characteristics of *Eucalyptus globulus* that are often useful for detecting F₁ hybrids. These characteristics are referred to in the figures bellow.
Chapter 2: Reproductive isolation

*E. globulus* x *arcana*

Figure 2. Juvenile morphological characteristics of *E. globulus* (g.), *E. arcana* (a.) and their F₁ hybrids. Note the intermediate nature of the F₁ in terms of leaf shape (middle) and stem squareness (right).
Chapter 2: Reproductive isolation

*E. globulus x botryoides*

**Figure 3.** The juvenile morphological characteristics of *E. globulus* (g.), *E. botryoides* (b.) and their F₁ hybrids. Note shortly petiolate leaves, square stem (right) and intermediate leaf shape of the F₁.
**Figure 4.** Juvenile morphological characteristics of *E. globulus* (g.), *E. camaldulensis* (c.) and their F₁ hybrids. Note the intermediate characteristics in F₁, including the mainly opposite leaves that are shortly petiolate (left), the leaf orientation and shape (middle) and stem squareness (right). These photographs are of *E. camaldulensis* ssp. *camaldulensis*. A single *E. camaldulensis* ssp. *simulata* x *E. globulus* hybrid was also found that had analogous morphology.
Chapter 2: Reproductive isolation

*E. globulus* x *cephalocarpa*

**Figure 5.** Juvenile morphological characteristics of *E. globulus* (g.), *E. cephalocarpa* (ce.) and their F₁ hybrids. Note the intermediate characteristics in F₁, including the shape (left and middle) and stem squareness (right)
Figure 6. Cotyledon morphology of *E. globulus* (g.), *E. conglobata* (con.) and their F₁ hybrids. Note the bi-lobed cotyledon shape in both the *E. globulus* and the F₁ in comparison to the reinform cotyledon in *E. conglobata* (right). After appearing healthy initially (first F₁ photo), hybrid fitness declined rapidly, as shown in the second F₁ photo.
Chapter 2: Reproductive isolation

*E. globulus x cornuta*

**Figure 7.** The juvenile morphological characteristics of *E. globulus* (g.), *E. cornuta* (c.) and their F₁ hybrids. Note intermediate characteristics in the F₁ including the opposite, shortly petiolate leaves and leaf shape.
Figure 8. The juvenile morphological characteristics of *E. globulus* (g.), *E. cosmophylla* (co.) and their F$_1$ hybrids. Note the intermediate characteristics in the F$_1$, including the shortly petiolate and opposite leaves, and the leaf shape. The painting of the *E. cosmophylla* is by Ian Roberts and is taken from Nicolle (2013).
Figure 9. Juvenile morphological characteristics of *E. globulus* (g.), *E. crenulata* (cr.) and their F₁ hybrids. Note the intermediate characteristics in F₁, including the slightly warty stem (top middle), stem squareness (top right), slightly crenulate leaf margins and leaf shape (bottom).
Chapter 2: Reproductive isolation

*E. globulus x dwyeri*

**Figure 10.** Juvenile morphological characteristics of *E. globulus* (*g.*), *E. dwyeri* (*d.*) and their F$_1$ hybrids. Note the intermediate characteristics in the F$_1$, including the mainly opposite leaves that are shortly petiolate (left and middle), the leaf orientation and shape (left and middle) and stem squareness (right).
Figure 11. Juvenile morphological characteristics of *E. globulus* (g.), *E. goniocalyx* (go.) and their F₁ hybrids. Note the intermediate characteristics in F₁, including the leaf shape and orientation (left and middle) and stem squareness (right).
Chapter 2: Reproductive isolation

_E. globulus x gunnii_

**Figure 12.** Juvenile morphological characteristics of _E. globulus_ (_g._), _E. gunnii_ (_gu._) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the branching pattern (top left), leaf shape and orientation (top left and bottom), and stem squareness (top right).
Chapter 2: Reproductive isolation

*E. globulus* x *hallii*

**Figure 13.** Juvenile morphological characteristics of *E. globulus* (g.), *E. hallii* (h.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the mainly opposite leaves that are shortly petiolate (left and middle), the leaf orientation and shape (left and middle) and stem squareness (right).
Figure 14. Juvenile morphological characteristics of *E. globulus* (*g.*), *E. kabiana* (*k.*) and their F$_1$ hybrids. Note the intermediate characteristics in the F$_1$, including the shortly petiolate leaves (left-middle), the leaf orientation and shape (left-middle) and stem squareness (right-middle).
Figure 15. Juvenile morphological characteristics of *E. globulus* (g.), *E. mannifera* (ma.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including leaf shape and orientation (left and middle), and stem squareness (right).
Figure 16. Juvenile morphological characteristics of *E. globulus* (g.), *E. michaeliana* (m.) and their F$_1$ hybrids. Note the intermediate characteristics in the F$_1$, including the often opposite leaves that are shortly petiolate (left and middle), the leaf orientation and shape (left and middle) and stem squareness (right).
Chapter 2: Reproductive isolation

E. globulus x minniritchii

Figure 17. Juvenile morphological characteristics of E. globulus (g.), E. minniritchii (mi.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the opposite leaves that are shortly petiolate (left and middle), the leaf orientation and shape (left and middle) and stem squareness (right).
Figure 18. Juvenile morphological characteristics of *E. globulus* (g.), *E. morrisbyi* (mo.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the branching pattern (left), leaf shape and orientation (left and middle), and stem squareness (right).
Chapter 2: Reproductive isolation

*E. globulus* x *nortonii*

**Figure 19.** Juvenile morphological characteristics of *E. globulus* (*g.*), *E. nortonii* (*n.*) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the leaf shape and orientation (left and middle), and stem squareness (right).
Figure 20. Juvenile morphological characteristics of *E. globulus* (g.), *E. ovata* (o.) and their F<sub>1</sub> hybrids. Note the intermediate characteristics in the F<sub>1</sub>, including the leaf shape and orientation (left and middle), and stem squareness (right).
Figures 21. Juvenile morphological characteristics of *E. globulus* (g.), *E. pulverulenta* (p.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the leaf shape and orientation (left and middle), and stem squareness (right).
Chapter 2: Reproductive isolation

_E. globulus x rudis_

Figure 22. Juvenile morphological characteristics of _E. globulus_ (g.), _E. rudis_ (r.) and their _F_1 hybrids. Note the intermediate characteristics in the _F_1, including the opposite leaves that are shortly petiolate (left and middle), the leaf orientation and shape (left and middle) and stem squareness (right).
Chapter 2: Reproductive isolation

*E. globulus x scias*

![Image showing juvenile morphological characteristics of E. globulus (g.), E. scias (s.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the opposite leaves that are shortly petiolate (middle left), the leaf orientation and shape (middle left) and stem squareness (middle right).](image_url)

**Figure 23.** Juvenile morphological characteristics of *E. globulus* (g.), *E. scias* (s.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the opposite leaves that are shortly petiolate (middle left), the leaf orientation and shape (middle left) and stem squareness (middle right).
Chapter 2: Reproductive isolation

_E. globulus x tereticornis_

*Figure 24.* Juvenile morphological characteristics of _E. globulus_ (g.), _E. tereticornis_ (t.) and their F<sub>1</sub> hybrids. Note the intermediate characteristics in the F<sub>1</sub>, including the sessile leaves (no petiole left- middle), and leaf shape. Hybrids also had a square stems in comparison to the angular or round stems of _E. tereticornis_.

74
Figure 25. Juvenile morphological characteristics of *E. globulus* (g.), *E. viminalis* (v.) and their F₁ hybrids. Note the intermediate characteristics in the F₁, including the branching pattern (left), leaf shape and orientation (left and middle), and stem squareness (right). These photos are of *E. viminalis* ssp. *viminalis*, two *E. viminalis* ssp. *cygnetensis* x *E. globulus* hybrids were also produced and showed analogous morphology.
Appendix 2.2: details regarding the molecular validation of hybrids arising from the crossing program

Hybrids arising from the crossing program (see Materials and Methods, main paper) were identified initially based on morphological markers (see Supporting Information 2). Here we describe the molecular validation of the hybrids, using 10 microsatellite loci, to determine parentage.

Materials and methods

A subset of putative hybrids from each cross combination was tested using molecular markers to validate the morphological classifications (Supporting Information 2). A parentage approach was taken, where tissue from both parents and the putative hybrids were genotyped and the alleles compared. The parentage analysis involved directly comparing (by eye) the microsatellite profiles of each parent with that of the putative progeny. Using such direct comparisons, it is clear when two alleles are the same; as such the error rate in the parentage analysis is assumed to be zero. Parentage was assumed to be correct when the hybrid had one allele from each parent at each locus (except where there was ambiguity or missing allele data; see Supporting Information 4). Total genomic DNA was extracted from frozen leaf samples using the CTAB protocol of Doyle and Doyle (1990) with the adjustments made by McKinnon et al. (2004b). The quality and quantity of DNA were assessed using gel electrophoresis and comparison with Lambda Hind III molecular weight standard. Ten microsatellite loci were used, four (EMCRC2; EMCRC7; EMCRC8; and EMCRC11) designed by Steane et al. (2001), two (EMBRA11 and EMBRA16) designed by Brondani et al. (1998), and four (EMBRA23; EMBRA30; EMBRA38; and EMBRA63) designed by Brondani et al. (2002). The primer sequences for the loci can be found in the respective references. These loci have been mapped and there is no evidence of linkage between them (J. S. Freeman, pers. comm.). In order to allow simultaneous analysis of different loci, each forward primer was labelled at its 5’ end with a fluorescent dye (NED, 6-FAM, PET, or Hex; Perkin Elmer Applied Biosystems, Foster City, CA, USA). The ten loci were multiplexed in three mixes: mix 1 included EMCRC2, EMBRA63 and EMBRA11 (using 0.2 µM of primer/reaction – forward and reverse combined); mix
Chapter 2: Reproductive isolation

2 included EMBRA10 (0.2 μM/reaction), EMCRC11 (0.4 μM/reaction), EMBRA23 (0.2 μM/reaction) and EMCRC7 (0.4 μM/reaction); and mix 3 included EMBRA30, EMBRA38 and EMCRC8 (all at 0.4 μM/reaction). PCRs were performed using a QIAGEN Multiplex PCR kit (Hilden, Germany) according to the manufacturer’s specifications for 5 μl reactions, using approximately 5 ng of genomic DNA per reaction. Thermo-cycler conditions followed Bloomfield et al. (2011) except the annealing temperatures were 59°C for mix 1, and 58°C for mix 2 and mix 3. The PCR products were diluted 1 in 10 and then 1 μl of that dilution was dried at 50°C. Fragment separation was undertaken on an AB3730 DNA analyser (Perkin Elmer Applied Biosystems) by the Australian Genome Research Facility (AGRF), Adelaide, South Australia. GENEMAPPER (version 3.7) was used for allele scoring. Allele binning and data checking followed Bloomfield et al. (2011).

Results

A total of 72 putative hybrids from 30 species were genotyped and compared to 34 mothers and 10 potential fathers. Hybrids from 26 species were found to have alleles matching the parents (a table of the allelic comparisons is given in Supporting Information 4) and were thus validated. Putative hybrids from four species (*E. crebra*, *E. gillii*, *E. scoparia* and *E. varia*) had alleles that did not match the parents and were therefore removed from the *E. globulus* hybrid dataset (Supporting Information 4).
Chapter 3: Assessing a Bayesian approach for detecting exotic hybrids between plantation and native eucalypts

This chapter is published in the *International Journal of Forestry Research*

3.1 Abstract

*Eucalyptus globulus* is grown extensively in plantations outside its native range in Australia. Concerns have been raised that the species may pose a genetic risk to native eucalypt species through hybridisation and introgression. Methods for identifying hybrids are needed to enable assessment and management of this genetic risk. This paper assesses the efficiency of a Bayesian approach for identifying hybrids between the plantation species’ *E. globulus* and *E. nitens*, and four at-risk native eucalypts. A database including range-wide DNA samples of *E. camaldulensis*, *E. cypellocarpa*, *E. globulus*, *E. nitens*, *E. ovata* and *E. viminalis*, as well as pedigreed and putative hybrids (*n* = 606), was genotyped with 10 microsatellite loci. Using a two-way simulation analysis (two species in the model at a time), the accuracy of identification was 98% for first and 93% for second generation hybrids. However, the accuracy of identifying simulated backcross hybrids was lower (74%). A six-way analysis (all species in the model together) showed that as the number of species increases the accuracy of hybrid identification decreases. Despite some difficulties identifying backcrosses, the two-way Bayesian modelling approach was highly effective at identifying *F₁*S, which, in the context of *E. globulus* plantations, are the primary management concern.
3.2 Introduction

Plants are well known for their propensity to hybridise (Levin 1979; Rieseberg and Carney 1998), and the role of hybridisation in animal systems is receiving growing attention (Randi 2008; Schwenk et al. 2008). Natural hybridisation has been widely documented in plants (Rieseberg and Carney 1998), with hybrid zones often being used to investigate the mechanisms that underlie speciation (Lexer et al. 2010; Rieseberg et al. 2003a; Rieseberg et al. 1999). These studies have demonstrated that barriers to hybridisation that evolve in allopatry are often incomplete, and hybridisation and introgression are still possible when species secondarily come into contact (Lexer et al. 2010; Rieseberg et al. 1999). Two consequences of human development have been the fragmentation of natural plant populations and the widespread movement of plant species around the world (Mack and Lonsdale 2001). In many situations this has resulted in exotic species coming into contact with cross-compatible indigenous species, leading to human mediated exotic hybridisation (Byrne and Stone 2011; Byrne et al. 2011; Laikre et al. 2010). This exotic hybridisation and potential for subsequent introgression, may threaten the genetic integrity of native species (Byrne et al. 2011; Laikre et al. 2010).

Given the genetic risk posed by exotic hybridisation, methods for detecting hybrid progeny are needed to enable quantification and management of the issue (Laikre et al. 2010). In some situations first generation (F₁) hybrids can be detected based on intermediate morphology, but in species with similar characteristics and in advanced generation hybrids (second (F₂) and backcross (BC) generations) morphological detection is often difficult and unreliable (Rieseberg and Ellstrand 1993). Over the past two decades several techniques utilising molecular markers have been developed for detecting hybrids (Anderson and Thompson 2002; Elo et al. 1997; Hewitt 2001; Pritchard et al. 2000). Early methods often depended on identifying species specific markers that could be used to identify immigrants (D’hont et al. 1995; Heath et al. 1995). This approach is highly effective in theory (Boecklen and Howard 1997), but in practice identifying species specific markers is problematic, especially in closely related taxa (Pritchard et al. 2000; Rieseberg and Ellstrand 1993).
Chapter 3: Hybrid detection

The development of highly polymorphic microsatellite markers, combined with new Bayesian statistical approaches (Beaumont and Rannala 2004), and advances in computing power, have allowed the development of model-based techniques for hybrid detection (Anderson and Thompson 2002; Pritchard et al. 2000). These approaches produce admixture estimates based on multi locus allele frequencies and Bayesian clustering (Anderson and Thompson 2002; Pritchard et al. 2000). The techniques have now been widely used to identify inter-specific hybridisation (Adams et al. 2007; Burgarella et al. 2009; Muñoz-Fuentes et al. 2007) and introgression (Barilani et al. 2007b; Sanz et al. 2009; Thompson et al. 2010; Valbuena-Carabaña et al. 2007). The two most commonly used programs are STRUCTURE (Pritchard et al. 2000) and NEWHYBRIDS (Anderson and Thompson 2002). Sanz et al. (2009) found that of four programs STRUCTURE and NEWHYBRIDS were the most effective at identifying hybrids, but STRUCTURE was the most accurate, correctly identifying 100% of simulated F1 and F2 hybrids and 96% of backcrosses, while NEWHYBRIDS misclassified 8 to 14% of F2s and 30 to 34% of backcrosses.

In this study we test a Bayesian modelling approach (using STRUCTURE) for identifying hybrids between the plantation species *Eucalyptus globulus* and five other eucalypt species – *E. camaldulensis*, *E. cypellocarpa*, *E. nitens*, *E. ovata* and *E. viminalis*. Over the last two decades there has been a major expansion of the eucalypt plantation estate in Australia, which now covers around 1,000,000 ha (Gavran and Parsons 2011). *Eucalyptus globulus* is the most widely planted species and is grown mainly outside its natural range (Gavran and Parsons 2011), raising concerns that it could pose a genetic risk to native eucalypt populations (Barbour et al. 2008b; Potts et al. 2003). Hybridisation is well documented in eucalypts (Griffin et al. 1988), and is more likely to occur between closely related species (Griffin et al. 1988; Potts et al. 2003). Hybrids have been reported based on morphology, with F1s typically being intermediate between the parental taxa (Barbour et al. 2008b; Griffin et al. 1988; Potts and Reid 1985). The distinctive juvenile morphology of many species also makes identification of hybrids possible at an early age, and this characteristic has been widely used in studies investigating exotic gene flow in eucalypts (Abasolo et al. 2012; Barbour et al. 2008b; Barbour et al. 2007).
Chapter 3: Hybrid detection

As in other groups, identifying eucalypt hybrids can be problematic where species have similar morphology, or when there are advanced generation hybrids that resemble the backcross parent (Barbour et al. 2007). Additionally, in the context of *E. globulus* plantations, there are a range of native species that have similar seedling morphology to *E. globulus*, which could hybridise with other native species and produce hybrid seedlings resembling exotic hybrids. There are also at least 36 native eucalypt species that grow adjacent to *E. globulus* plantations (Barbour et al. 2008b), and these species have a wide range of juvenile characteristics. This diversity can make it difficult to distinguish between the intermediate morphology of a hybrid, and the morphology of an unfamiliar species.

There is evidence that exotic hybridisation is occurring from *E. globulus* plantations (Barbour et al. 2008b). Barbour et al. (2008b) found that 35% of *E. globulus* plantations are in close proximity to cross compatible native species and they detected low levels of hybridisation in open pollinated native seedlots, as well as hybrids establishing in native forest beside one plantation. In order to enable identification and management of exotic hybridisation from *E. globulus* plantations, an approach is needed to validate and/or identify hybrids between *E. globulus* and at-risk species. The species selected here have been chosen because: they are common plantation neighbours; are known to hybridise with *E. globulus*; and in some cases their seedling morphology is similar enough to that of *E. globulus*, that distinguishing their hybrids from *E. globulus* hybrids would be difficult (*E. cypellocarpa* and *E. nitens*). The specific aims of the study are to assemble range wide molecular databases for each species using 10 microsatellite loci, and then using Bayesian admixture analysis test the ability of those marker sets to detect hybrids with *E. globulus* using simulated, pedigreed and putative hybrid samples.

3.3 Materials and Methods

3.3.1 Sample description

Collections of range-wide samples of *E. camaldulensis*, *E. cypellocarpa*, *E. globulus*, *E. nitens*, *E. ovata* and *E. viminalis* were assembled from various sources (Table 3.1; Fig. 3.1). A range of pedigreed and putative hybrid samples were also
collected for assessing and testing the ability of the modelling approach (see below) to detect hybrids. All hybrids referred to as “pedigreed” have either been validated with parentage analysis using molecular markers (to be published elsewhere), or are from controlled crossing. Samples validated with molecular markers include the following: *E. camaldulensis* x *globulus* (*n* = 2), *E. viminalis* x *globulus* (*n* = 3), and *E. ovata* x *globulus* (*n* = 2). Samples produced through controlled crossing from an advanced generation hybrid trial between *E. globulus* and *E. nitens* were also used. Details of the crossing approach and trial establishment can be found in Costa e Silva et al. (2012), and validation of the cross types with Near Infrared Reflectance Spectroscopy is explained in O’Reilly-Wapstra et al. (2013). Samples used from this trial were 12 F$_1$, four F$_2$ and 12 *E. globulus* backcross hybrids (BC$_{glob}$).

Unconfirmed putative hybrid samples were also used to test the method. Four putative *E. camaldulensis* x *globulus* F$_1$ hybrids, identified from morphology in open pollinated seed from *E. camaldulensis* trees beside an *E. globulus* plantation in South Australia. Seven putative *E. cypellocarpa* x *globulus* hybrids from Victoria: one was collected from a mature native tree in mixed *E. globulus/cypellocarpa* forest; two were tentatively identified as “possible hybrids?” (with a high degree of uncertainty based on morphology) from beside an *E. globulus* plantation; and four were from a population that has been speculated to be a phantom hybrid zone (Kirkpatrick et al. 1973; Parsons and Kirkpatrick 1972). Four putative *E. viminalis* x *globulus* samples were collected in Tasmania, one identified on the basis of seedling morphology in open-pollinated progeny from a native *E. globulus* tree; the other samples were collected from mature native trees with intermediate bud and capsule morphology. Details of the six pure species are given in Appendix 3.1, and Figure 3.1 shows distribution maps for each species, the collection location of samples in this study and the distribution of *E. globulus* plantations in Australia.
Figure 3.1. The natural distributions (in orange) of *Eucalyptus camaldulensis*, *E. cypellocarpa*, *E. globulus*, *E. nitens*, *E. ovata* and *E. viminalis*, with the distribution of samples used (yellow circles), the hybrids used (red crosses) and of *E. globulus* plantations (green polygons). The box within the *E. camaldulensis* map shows the extent of the maps for the other species (species distributions reproduced from: Euclid 2006; Nicolle 2013).
**Chapter 3:** Hybrid detection

**Table 3.1.** Summary of samples used to create range wide microsatellite databases for six *Eucalyptus* species in order to develop a model for identifying hybrids with *E. globulus*. All the species are members of the subgenus Symphyomyrtus, and their lower level taxonomy is given according to Euclid (2006); see Appendix 1 for more details. Also given for each species is: the number of pure samples; the number of pedigreed and putative hybrid samples with *E. globulus*; their co-occurrence with *E. globulus*; reported hybrids with *E. globulus*; the similarity in their seedling morphology to *E. globulus*; and the source of samples.

<table>
<thead>
<tr>
<th>Species</th>
<th>Taxonomy (section series)</th>
<th>n pure samples</th>
<th>n pedigreed hybrids</th>
<th>n putative hybrids</th>
<th>Naturally co-occur?</th>
<th>Hybrids with <em>E. globulus</em></th>
<th>Seedling morphology</th>
<th>Sample source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. camaldulensis</em></td>
<td>Exsertaria Rostratae</td>
<td>97</td>
<td>2</td>
<td>4</td>
<td>no</td>
<td>natural &amp; manipulated</td>
<td>different</td>
<td>(Butcher et al. 2009); this study</td>
</tr>
<tr>
<td><em>E. cypellocarpa</em></td>
<td>Maidenaria Globulares</td>
<td>97</td>
<td>-</td>
<td>7</td>
<td>yes</td>
<td>natural</td>
<td>similar</td>
<td>this study</td>
</tr>
<tr>
<td><em>E. globulus</em></td>
<td>Maidenaria Globulares</td>
<td>87</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Freeman et al. 2001); (Steane et al. 2006); (Foster et al. 2007); (Hudson 2012); this study</td>
</tr>
<tr>
<td><em>E. nitens</em></td>
<td>Maidenaria Globulares</td>
<td>88</td>
<td>28</td>
<td>-</td>
<td>no</td>
<td>manipulated</td>
<td>similar</td>
<td>(Hudson 2012); (Humphreys et al. 2008)</td>
</tr>
<tr>
<td><em>E. ovata</em></td>
<td>Maidenaria Foveolatae</td>
<td>100</td>
<td>2</td>
<td>-</td>
<td>yes</td>
<td>natural &amp; manipulated</td>
<td>different</td>
<td>(Marthick 2005); this study</td>
</tr>
<tr>
<td><em>E. viminalis</em></td>
<td>Maidenaria Viminalis</td>
<td>87</td>
<td>3</td>
<td>4</td>
<td>yes</td>
<td>natural &amp; manipulated</td>
<td>different b</td>
<td>(Marthick 2005); this study</td>
</tr>
<tr>
<td>Hybrids total</td>
<td></td>
<td>-</td>
<td>50</td>
<td>35</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>606</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* a Natural *E. camaldulensis* x *globulus* hybrids have been found where *E. globulus* co-occurs as an exotic (Barbour et al. 2008b)

b The seedling morphology of *E. viminalis* is easily distinguishable from *E. globulus*, but several key characters are similar enough to make hybrid identification problematic.
3.3.2 Molecular methods

A total of 606 samples were genotyped, and 27 samples were repeated to enable assessment of the accuracy and repeatability of allele calling. For the samples collected in this study, total genomic DNA was extracted from the frozen leaf samples using the CTAB protocol of Doyle and Doyle (1990) with the adjustments used by McKinnon et al. (2004b). The quality and quantity of DNA was assessed using gel electrophoresis and comparison with Lambda HindIII molecular weight standard. Additionally because of quarantine restrictions preventing the importation of eucalypt material from New South Wales to Tasmania, the 45 E. cypellocarpa samples collected in New South Wales were sent fresh to the Australian Genome Research Facility, South Australian for DNA extraction and quantification. Ten microsatellite loci were used for genotyping, four (EMCRC2; EMCRC7; EMCRC8; and EMCRC11) designed by Steane et al. (2001), two (EMBRA11 and EMBRA16) designed by Brondani et al. (1998), and four (EMBRA23; EMBRA30; EMBRA38; and EMBRA63) designed by Brondani et al. (2002); the primer sequences for all loci can be found in their respective references. These loci have been mapped and there is no evidence of linkage between them (J. S. Freeman, pers. comm.). In order to allow simultaneous analysis of different loci, the forward primers were labelled at their 5’ end with the fluorescent dyes NED, 6-FAM, PET, or Hex (Perkin Elmer Applied Biosystems, Foster City, CA, USA). The ten loci were multiplexed in three mixes, mix 1 included EMCRC2, EMBRA63 and EMBRA11 (using 0.2µM of primer/reaction – forward and reverse combined); mix 2 included EMBRA10 (0.2µM/reaction), EMCRC11 (0.4 µM/reaction), EMBRA23 (0.2 µM/reaction) and EMCRC7 (0.4 µM/reaction); and mix 3 included EMBRA30, EMBRA38 and EMCRC8 (all at 0.4 µM/reaction). The PCRs were performed using a QIAGEN Multiplex PCR kit (Hilden, Germany) according to the manufacturer’s specifications for 5 µl reactions using approximately 5 ng of genomic DNA per reaction. Thermocycler conditions followed Bloomfield et al. (2011) with annealing temperatures of 59°C for mix 1, and 58°C for mix 2 and mix 3. The PCR product was diluted 1 in 10 in H2O and then 1 µl of that dilution was dried at 50°C. Fragment separation was undertaken on an AB3730 DNA analyser (Perkin Elmer Applied Biosystems) by the
Chapter 3: Hybrid detection

Australian Genome Research Facility, South Australia. Allele scoring followed Bloomfield et al. (2011). The assigned genotypes of the 27 repeated samples (which were scored blindly) were compared at each locus to obtain a measure of repeatability (number of allelic errors/number of alleles compared).

3.3.3 Analytical approach

3.3.4 1) Assessment of genetic differentiation between species

Deviation from Hardy-Weinberg-equilibrium within species was assessed in GENEPOP version 4.2 (Raymond and Rousset 1995; Rousset 2008). Pairwise $F_{ST}$ and $D_{EST}$ (a standardised version of GST with a range from 0 to 1; Jost 2008) were calculated for each species pair in GENALEX (Peakall and Smouse 2012) using Analysis of Molecular Variance (AMOVA), which also tests the statistical significance of the $F_{ST}$ comparisons. This analysis used the default parameters for the AMOVA function except that the “interpolate missing data” function was switched on, and the number of permutations increased to 9999.

All 559 pure individual genotypes were run in STRUCTURE using the admixture model (which was used in all analyses) without a priori population information. This analysis used a burn-in of 50,000 Markov chain Monte Carlo (MCMC) iterations, followed by run of 100,000 data generating MCMC iterations, with all other program parameters set to default. A range of $K$ from 1 to 10 were used, with each analysis repeated 10 times. STRUCTURE HARVESTER (Earl 2012) was used to calculate the mean likelihood of $K$ (Pritchard et al. 2000), and the log likelihood method, $\Delta K$ (Evanno et al. 2005) to determine the most appropriate number of genetic clusters ($K$).

3.3.5 2) Calculating detection power with simulated and pedigreed hybrids

The program HYBRIDLAB (Nielsen et al. 2006) was used to generate a series of simulated hybrid generations to assess the accuracy of the STRUCTURE technique. For each parental combination (i.e. $E. globulus$ and any of the other five species [parent-2]), 300 individuals were simulated, including 50 each of the following: Parent-$globulus$, Parent-2, $F_1$, $F_2$, BC$_{globulus}$, BC$_{parent-2}$. In order to check that the number of hybrid samples did not affect the accuracy of assignment, the pairs with
the highest and lowest level of differentiation (E. globulus x camaldulensis and E. globulus x cypellocarpa respectively) were also analysed using only 10 replicates of each simulated hybrid generation (keeping 50 simulated parental samples). This had very little effect on the hybrid assignments, so only the data for \( n = 50 \) are presented.

In Tasmania and Gippsland native populations of E. globulus and E. ovata occur adjacent to E. nitens plantations and distinguishing E. globulus x ovata from E. nitens x ovata juveniles based on morphology would be very difficult. Therefore a three way simulation was run involving E. globulus x ovata (F1s), E. nitens x ovata (F1s) and the three pure parental populations, so as to assess the ability to detect hybrid parentage between these three species in the field.

Each pair of parental and simulated hybrid populations, and any pedigreed hybrids samples were then analysed in STRUCTURE using \( K = 2 \). After checking the species groups were correctly identified with no a priori information (Pritchard et al. 2000), the parental species were used to define the two genetic clusters (the USEPOPINFO method) and the genotype membership (\( q \)) of hybrid samples were allocated using the admixture model. This analysis was run five times/combinations.

The default parameters were used except that the ‘allele frequencies updated using individuals with POPFLAG=1 ONLY’ option was selected. A burn-in of 50,000 MCMC iterations was followed by 100,000 data generation runs. The program CLUMPP (Jakobsson and Rosenberg 2007) was used to merge the five STRUCTURE runs and that data was used for hybrid allocation. A \( q \) cut-off of \( \geq 0.2 \) was used to identify hybrids. This cut-off means that if \( q \geq 0.2 \) in both \( K \) clusters the individual is classified as a hybrid, and if \( q < 0.2 \) in one cluster, then the individual is indistinguishable from the parental species (i.e. the cluster with \( q > 0.8 \)). The proportion of simulated individuals correctly assigned as hybrids at \( q > 0.2 \) was used to estimate the hybrid detection power.

### 3.3.6 3) Classification of the putative hybrids using STRUCTURE

The STRUCTURE protocol above was used to classify the putative hybrid samples for comparison to the pedigreed and simulated results. Finally, to test the model when maternity is completely unknown, all pure samples, the simulated parental samples, simulated F1s, and the pedigreed and putative hybrids were all run together
in a six species analysis. Classification of hybrids given \( K = 6 \) is slightly more complicated. Classification as a hybrid was considered correct if the two true parental \( q \) values summed to at least 0.67 (i.e. more than two thirds the total possible \( q \)), both true parents contributed \( q > 0.2 \), and no other species contributed more than either the parent. The program DISTRUCT (Rosenberg 2004) was used to produce individual genotype membership plots for comparison of pedigreed, simulated and putative hybrid samples.

### 3.4 Results

In the 606 individuals genotyped at 10 loci we found 344 different alleles. There was 1.5\% missing data, and repeatability was 93\%. *Eucalyptus camaldulensis* had the highest average number of alleles per locus (25.2), and more private alleles than any other species (28). *Eucalyptus ovata* and *E. nitens* had the lowest genetic diversity (\( H_e = 0.82 \), and 0.83 respectively), while *E. cypellocarpa* had the highest (\( H_e = 0.91 \)). The hybrid group had the lowest number of private alleles and the highest observed heterozygosity (full population genetic details are in Appendix 3.2). As expected given our range wide sampling of the species, there was significant departure from Hardy-Weinberg-Equilibrium (HWE) at several loci, but just under half the loci-population combinations (24 out of 60) were in HWE, and there were no clear patterns in departure between species. STRUCTURE assumes HWE, and although other authors have found the program to be robust to deviations from HWE (Curtu et al. 2007), this departure from the assumptions makes the simulation analysis particularly important for determining the accuracy of our hybrid allocations. According to the AMOVA, most variation in the dataset was within species (92\%) with just 8\% partitioned between species. The pairwise \( F_{ST} \) estimates were low (ranging from 0.027 to 0.112), but were all highly significant (Table 3.2). \( D_{EST} \) showed more intermediate levels of differentiation (0.26 to 0.71) than \( F_{ST} \), but the pattern of pair-wise differentiation was similar between the two estimates. Under both measures the lowest levels of molecular differentiation were between *E. cypellocarpa* and *E. viminalis*, and between *E. cypellocarpa* and *E. globulus*, while the most well differentiated species pairs were *E. ovata* and *E. nitens*, and *E. camaldulensis* and *E. ovata* (Table 3.2).
Chapter 3: Hybrid detection

Table 3.2. Two measures of pairwise genetic differentiation between six eucalypt species genotyped at 10 microsatellite loci. Below the diagonal is Wright’s $F_{ST}$, and above is $D_{EST}$. All pairwise $F_{ST}$ comparisons were significant at $P < 0.0001$ using AMOVA.

<table>
<thead>
<tr>
<th></th>
<th>E. camaldulensis</th>
<th>E. cypellocarpa</th>
<th>E. globulus</th>
<th>E. nitens</th>
<th>E. ovata</th>
<th>E. viminalis</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. camaldulensis</td>
<td>-</td>
<td>0.62</td>
<td>0.59</td>
<td>0.64</td>
<td>0.71</td>
<td>0.64</td>
</tr>
<tr>
<td>E. cypellocarpa</td>
<td>0.063</td>
<td>-</td>
<td>0.35</td>
<td>0.64</td>
<td>0.51</td>
<td>0.26</td>
</tr>
<tr>
<td>E. globulus</td>
<td>0.069</td>
<td>0.040</td>
<td>-</td>
<td>0.63</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>E. nitens</td>
<td>0.092</td>
<td>0.087</td>
<td>0.097</td>
<td>-</td>
<td>0.60</td>
<td>0.67</td>
</tr>
<tr>
<td>E. ovata</td>
<td>0.106</td>
<td>0.074</td>
<td>0.093</td>
<td>0.112</td>
<td>-</td>
<td>0.43</td>
</tr>
<tr>
<td>E. viminalis</td>
<td>0.070</td>
<td>0.027</td>
<td>0.059</td>
<td>0.097</td>
<td>0.067</td>
<td>-</td>
</tr>
</tbody>
</table>

The ability of STRUCTURE to distinguish between the six species was tested by using no a priori species information. The mean likelihood of $K$ indicated a plateau in the likelihood surface corresponding with $K = 6$ (suggesting six groups), while the $\Delta K$ method of Evanno et al. (2005) showed a major genetic split in the data at $K = 4$, with secondary peaks at $K = 5$ and 6 (Fig. 3.2). At $K = 4$ E. cypellocarpa and E. globulus clustered together, as did E. ovata and E. viminalis; while at $K = 5$, E. cypellocarpa clustered with E. viminalis. However, at $K = 6$ the clusters corresponded to the species groups, which is the most biologically meaningful result (see Appendix 3.2). This indicates that although higher levels of structure seem to exist (at $K = 4$ and 5; which was also evident in the low $F_{ST}$ measures between some pairs of species), the dataset does differentiate the six unique species groups at $K = 6$.

In the two-way STRUCTURE analysis using a priori species information the accuracy of detecting both simulated and pedigreed F$_1$s was high, with 98% for each (Table 3.3). In three of the simulated and three of the pedigreed F$_1$ combinations 100% of hybrids were detected (Table 3.3). Detectability of simulated F$_2$s was slightly lower at 93% (Table 3.3). The overall accuracy of detecting simulated parental individuals was slightly lower again (91%), which was due mainly to difficulty in detecting simulated parents in the E. cypellocarpa x globulus and E. viminalis x globulus combinations (Table 3.3), which were also the least well differentiated species in terms of $F_{ST}$ (Table 3.2). The lowest detectability in both simulated and pedigreed hybrids was in the backcross generations, where the accuracy fell to just 33.4% for detecting the only pedigreed backcross combination.
(E. nitens x globulus; BC globulus; Table 3.3; Fig.3.3). Figure 3.3 shows that although there was some variation in the simulated groups, the patterns across the different generations and species combinations are consistent with theoretical expectations. This was clear in the group means (presented in Appendix 3.3). For example in the E. camaldulensis x globulus combination the mean of the F1 assignments were 0.53 (E. camaldulensis cluster) and 0.47 (E. globulus cluster), while the means for the BC camal were 0.74 and 0.26. These are very close to the theoretical allele frequencies expected for F1 and BC generations (i.e. 0.5 to 0.5 and 0.75 to 0.25 respectively), and similar theoretically consistent mean q values were obtained for all combinations (Appendix 3.3).

**Figure 3.2.** Methods for estimating the most appropriate number of genetic clusters (K) testing K = 2 to K = 10 for 556 individuals from six Eucalyptus species. a) The mean likelihood of K (±SD) showing a plateau in the likelihood surface at K = 6, and b) ΔK methods showing significant genetic groupings at K = 4, 5 and 6. This genetic clustering of multi locus genotypes was undertaken in STRUCTURE with no a priori species information.
Table 3.3. The overall accuracy of assignment of simulated parental and hybrid individuals and pedigreed hybrids between *Eucalyptus globulus* and five other species, based on Bayesian cluster analysis using $q > 0.2$ to identify hybrid individuals (See Supplementary Material 1 for more detail). Each simulated generation consisted of 50 individuals.

<table>
<thead>
<tr>
<th>Model combination</th>
<th>Generation</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. camaldulensis</em> × <em>globulus</em></td>
<td>Simulated <em>E. globulus</em></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated <em>E. camaldulensis</em></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;glob&lt;/sub&gt;</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;camal&lt;/sub&gt;</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Pedigreed F₁ ($n = 2$)</td>
<td>100</td>
</tr>
<tr>
<td><em>E. cypellocarpa</em> × <em>globulus</em></td>
<td>Simulated <em>E. globulus</em></td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Simulated <em>E. cypellocarpa</em></td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;glob&lt;/sub&gt;</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;cypel&lt;/sub&gt;</td>
<td>80</td>
</tr>
<tr>
<td><em>E. nitens</em> × <em>globulus</em></td>
<td>Simulated <em>E. globulus</em></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated <em>E. nitens</em></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;glob&lt;/sub&gt;</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;nit&lt;/sub&gt;</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Pedigreed F₁ ($n = 12$)</td>
<td>91.7</td>
</tr>
<tr>
<td></td>
<td>Pedigreeed F₂ ($n = 4$)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Pedigreeed BC&lt;sub&gt;glob&lt;/sub&gt; ($n = 12$)</td>
<td>33.4</td>
</tr>
<tr>
<td><em>E. ovata</em> × <em>globulus</em></td>
<td>Simulated <em>E. globulus</em></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated <em>E. ovata</em></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;glob&lt;/sub&gt;</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;ovat&lt;/sub&gt;</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Pedigreeed F₁ ($n = 2$)</td>
<td>100</td>
</tr>
<tr>
<td><em>E. viminalis</em> × <em>globulus</em></td>
<td>Simulated <em>E. globulus</em></td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Simulated <em>E. viminalis</em></td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;glob&lt;/sub&gt;</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;vina&lt;/sub&gt;</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Pedigreeed F₁ ($n = 3$)</td>
<td>100</td>
</tr>
<tr>
<td>Generation means</td>
<td>Simulated <em>E. globulus</em></td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Simulated parent-2</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Simulated F₁</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Simulated F₂</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;globulus&lt;/sub&gt;</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Simulated BC&lt;sub&gt;parent-2&lt;/sub&gt;</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Pedigreeed F₁ ($n = 19$)</td>
<td>98</td>
</tr>
</tbody>
</table>
Of the 15 putative hybrid samples assessed, 12 were classified as hybrids. The remaining three putative hybrids were indistinguishable from their pure parents (Table 3.4). The $q$ values of all 12 samples assigned as hybrids were outside the 95% confidence intervals (CIs) of both parents, indicating they are unlikely to be miss classified parental samples. However, one *E. camaldulensis* x *globulus* $F_1$ sample showed stronger affinities to *E. globulus*, with $q$ values of 0.335 (95% CIs = 0.108,0.587) to *E. camaldulensis* and 0.665 (95% CIs = 0.413,0.892) to *E. globulus*, but the 95% CIs did include the simulated $F_1$ means (*E. camaldulensis* cluster mean = 0.530; *E. globulus* cluster mean = 0.470). The remaining three *E. camaldulensis* x *globulus* $F_1$s had c. 0.5 assignment to each parent. The three individuals that were indistinguishable from one parent had overlapping 95% CIs with that parent.

**Table 3.4.** The assignment of 15 putative eucalypt hybrid samples based on their mean $q$ values, with $q > 0.2$ to indicating hybrid status under the two-way and six-way STRUCTURE analyses’ (hybrid’ = putative hybrid identified based on morphology from mature native forest trees, generation unknown; see materials and methods for more detail).

<table>
<thead>
<tr>
<th>Putative hybrid samples</th>
<th>Original classification</th>
<th>Assignment with two-way analysis</th>
<th>Assignment with six-way analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. camaldulensis</em> x <em>globulus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- #1</td>
<td>$F_1$</td>
<td>hybrid</td>
<td><em>E. globulus</em></td>
</tr>
<tr>
<td>- #2</td>
<td>$F_1$</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #3</td>
<td>$F_1$</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #4</td>
<td>$F_1$</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td><em>E. cypellocarpa</em> x <em>globulus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- #1</td>
<td>hybrid’</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #2</td>
<td>hybrid’</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #3</td>
<td>hybrid’</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #4</td>
<td>hybrid’</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>- #5</td>
<td>hybrid’</td>
<td>hybrid</td>
<td><em>E. cypellocarpa</em></td>
</tr>
<tr>
<td>- #6</td>
<td>$F_1$</td>
<td><em>E. cypellocarpa</em></td>
<td><em>E. cypellocarpa</em></td>
</tr>
<tr>
<td>- #7</td>
<td>$F_1$</td>
<td><em>E. cypellocarpa</em></td>
<td><em>E. cypellocarpa</em></td>
</tr>
<tr>
<td><em>E. viminalis</em> x <em>globulus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- #1</td>
<td>hybrid’</td>
<td>hybrid</td>
<td><em>E. globulus</em></td>
</tr>
<tr>
<td>- #2</td>
<td>hybrid’</td>
<td>hybrid</td>
<td><em>E. globulus</em></td>
</tr>
<tr>
<td>- #3</td>
<td>hybrid’</td>
<td>hybrid</td>
<td><em>E. globulus</em></td>
</tr>
<tr>
<td>- #4</td>
<td>hybrid’</td>
<td><em>E. globulus</em></td>
<td><em>E. globulus</em></td>
</tr>
</tbody>
</table>
Figure 3.3. Proportion of genotype membership ($q$) of simulated, pedigreed and putative hybrid individuals based on Bayesian cluster analysis. The real parental samples (not shown) were used to define the genetic clusters for each combination and showed between 98 and 100% affinity to their defined group. Each simulated generation consists of 50 individuals in each combination. Each individual is represented by a single vertical line that is partitioned based on its genotype affinities to the parental groups. Two species analyses are shown in a) to e), with three and six species analysis in f) and g) respectively. Abbreviations in g): P/P hybrids = pedigreed and putative hybrids; Sim = simulated; camal = E. camaldulensis; cypell = Eucalyptus cypellocarpa; glob = E. globulus; nit = E. nitens; ova = E. ovata; and vim = E. viminalis. In b) the bar with the “M” indicates samples from the Mallacoota population that are referred to in the text.
The three-way simulation involving *E. ovata* x *E. nitens* and *E. ovata* x *E. globulus* showed that the marker set could accurately identify parental contributions when there are multiple parent-hybrid combinations in the dataset (Fig. 3.3). Figure 3.3 (f and g) shows that there is an increase in the noise of the admixture assignments as the number of genetic clusters increases. This is more apparent in the six way analysis (Fig. 3.3), where for some species, in particular *E. cypellocarpa* and *E. viminalis*, the accuracy of assignment is markedly reduced (Fig. 3.3). This being said the assignment of F1 pedigreed samples is still relatively efficient at 83.3%. The reduction in efficiency of the six-way analysis can also be seen in the classification of the putative hybrid samples with four samples that classified as hybrids under the two-way analysis now being indistinguishable from one parental species (Table 3.4).

3.5 Discussion

3.5.1 Performance of the approach

This study successfully developed and tested a microsatellite database and Bayesian modelling approach for a group of six species, so as to enable the detection of exotic F1 hybrids from *E. globulus* (and possibly *E. nitens*) plantations when the maternal native species is known. This required sufficient differentiation in allele frequencies so that individuals with genotypes that are admixed between two species could be detected. Yet, the level of differentiation between species estimated from the inbreeding coefficient $F_{ST}$ is low in comparison to a range of other eucalypts (Byrne 2008; Jones et al. 2012; McGowen et al. 2001; Potts and Wiltshire 1997), particularly considering the comparison is between taxonomically distinct species. However, because the range-wide sampling strategy used here covered geographically discrete populations of these widespread species (Fig. 3.1), it probably captures significant within species differentiation, which could subsequently reduce $F_{ST}$ values between species (Heller and Siegismund 2009; Jost 2008; Meirmans 2006; Meirmans and Hedrick 2011). Nevertheless, all pairwise $F_{ST}$ values were highly significant and the standardised metric $D_{EST}$ indicated intermediate species differentiation. The level of differentiation in the dataset was sufficient for STRUCTURE to distinguish between the species, and enabled the accurate detection of hybrids. Patterns in hybrid detection were consistent with
patterns of $F_{ST}$ between pairs of taxa. This is consistent with theoretical expectations (Pritchard et al. 2000), and empirical modelling (Vähä and Primmer 2006) in that species with the highest differentiation (e.g. *E. camaldulensis* and *E. globulus*) showed the highest accuracy in hybrid detection and *vice versa*.

For all species combinations in the two-way analysis, the accuracy and likelihood of identifying simulated and pedigreed F$_1$ hybrids was high, and comparable with similar studies in forest trees (Burgarella et al. 2009; Curtu et al. 2007; Lepais et al. 2009). Parental and F$_2$ generations were also accurately identified, although the success in detecting backcross hybrids was lower. Several studies have had similar problems identifying backcross generations using the same approach (Barilani et al. 2007b; Burgarella et al. 2009; Randi 2008). For example Lepais et al. (Lepais et al. 2009) reported that 32% of simulated oak backcrosses were misclassified as pure species despite strong species differentiation and the use of a lower $q$ cut-off ($q > 0.1$). If $q > 0.1$ was used here, then the efficiency in detecting simulated *E. camaldulensis* x *globulus* backcrosses would increase from 70% to 96%, however, there would be a parallel increase in the number of pure parents incorrectly identified as hybrids – rising from 1% to 36%. This trade off has been documented by several authors (Barilani et al. 2007a; Barilani et al. 2007b; Lepais et al. 2009; Vähä and Primmer 2006) and simulation studies show that the number of markers necessary for accurate and efficient identification of backcross hybrids could be as high as 48 microsatellite loci given an $F_{ST}$ of 0.21 (Vähä and Primmer 2006). Vähä and Primmer (2006) tested a range of $q$ values and found that $q > 0.2$ most effectively balanced efficiency and accuracy. However, detection of exotic backcross hybrids is not currently necessary in the *E. globulus* system. The Australian *E. globulus* estate is young, with most plantations nearing the end of their first 10 – 15 year rotation (Gavran and Parsons 2011; Greaves et al. 2003) making it highly unlikely that mature exotic F$_1$ hybrids exist to produce backcrosses.

If the detection of backcrosses does become necessary, more loci could be added to the existing set, or a different marker systems could be used (Barilani et al. 2007b; Vähä and Primmer 2006). For example, Diversity Array Technology (DArT) has recently been developed for a range of eucalypt species including *E. globulus*, *E.
Chapter 3: Hybrid detection

camaldulensis and E. nitens. The DArT system produces hyper variable dominant markers (Steane et al. 2011), with over 5,000 polymorphic loci currently available in eucalypts (Sansaloni et al. 2010; Steane et al. 2011). Despite the lower information content per marker due to dominance, using such large marker datasets and similar methods to those used here would presumably lead to highly accurate assignments. Indeed much smaller DArT datasets (1122 markers) have been shown to outperform similar sized microsatellite (8 loci) datasets in other studies employing Bayesian clustering (Bouchet et al. 2012). Alternatively, with so many markers it may be possible to identify subsets of species specific loci or alleles that could differentiate hybrid generations with greater power than microsatellite based systems, without the need to assay all 5,000 loci. For example, Boecklen and Howard (1997) found that four or five independent species specific markers can accurately identify first generation backcrosses. However, development of such marker systems is time consuming and expensive and their deployment will depend on a trade-off between cost, time and the required detection power (Avise 2004). The system developed here is effective and cost efficient (lab costs are approximately $15AU/new sample, including DNA extraction and microsatellite assay, but not technician time) given the current requirement for identifying F₁ hybrids.

In eastern and southern Tasmania as well as parts of Gippsland in Victoria, E. nitens plantations occur within the native range of both E. globulus and E. ovata. Because of the very similar juvenile morphology of E. globulus and E. nitens, this could result in situations where the parentage of hybrids detected with E. ovata would be ambiguous. Exotic hybrids between E. ovata and both plantation species are well known (Barbour et al. 2008b; Barbour et al. 2003) and do show similar morphology. The three way simulation (E. globulus, E. nitens and E. ovata) here showed the utility of the microsatellite based approach in overcoming ambiguity in this situation, and it could be an important management tool for distinguishing between exotic and natural hybridisation where E. nitens grows within the native range of E. globulus.

The six-way analysis is in some ways the ultimate test of the approach, and will be particularly useful for assessing putative hybrids collected in the field where maternity is unclear. The accuracy clearly decreased as the number of species in the
model increased (i.e. from 2 – 3 – 6). However, despite this reduction, the six-way model could still identify over 80% of F₁ hybrids. Most other published studies assess two (Cullingham et al. 2013; Muñoz-Fuentes et al. 2007; Muranishi et al. 2013; Ortego and Bonal 2010; Randi 2008; Vähä and Primmer 2006), three (Barilani et al. 2007a; Barilani et al. 2007b), or occasionally four (Lepais et al. 2009) species when investigating hybrid parentage, and in reality it is unlikely that a situation will arise where all six species from this model are potential parents. An assessment of the species growing where the hybrid was collected would probably narrow down the number of potential parents. Also most exotic plantation hybrids identified in the field are found among pure seedlings of the native species (Barbour et al. 2008b; Barbour et al. 2003) likely enabling the identification of a single putative maternal species. In an operational context where a putative exotic hybrid has been identified in the field, an effective approach may be to run a full six-way analysis to rule out other species contributing pollen, which can travel long distances (Barbour et al. 2005b), then reduce the number of species to those found in the vicinity of the putative hybrid.

3.5.2 Allocation of putative hybrids

The putative hybrid samples assessed here came from a range of situations including native forests where the generation of the putative hybrids was unknown. This is a more challenging problem than identifying hybrids around E. globulus plantations where any hybrids can currently be assumed to be F₁s. However, by incorporating additional information, including demographic details at the collection site, and morphology, it is possible to improve allocation confidence, and estimate the hybrid generation of the samples. Of the 15 putative hybrids 12 were classified as hybrids and three could not be distinguished from their pure parents in the two-way analysis. Of the three indistinguishable samples, the two putative E. cypellocarpa F₁s are probably correctly classified as pure E. cypellocarpa. These two putative were collected from seedlings beside a 10 year old plantation – ruling out the possibility that they are backcrosses. The samples were only tentatively classified as “possible” F₁s (with a low degree of certainty) based on the degree of glaucousness, which can be a variable trait (Euclid 2006). The putative E. viminalis x globulus hybrid that was
classified as *E. globulus* is less clear-cut. This sample was identified in open pollinated *E. globulus* seed, collected from a native forest with no other cross-compatible eucalypts nearby. The sample showed distinctively intermediate morphology on multiple traits, consistent with known hybrids between these species (Griffin et al. 1988). The likelihood of random morphological deviations on multiple traits resulting in intermediate characteristics is low, and is more easily explained by inter-specific hybridisation (Hopper 1978; Potts and Dungey 2004). Therefore, considering the model inaccuracy when identifying backcrosses, it is possible that this sample is actually a first or perhaps later generation backcross.

Incorporating additional site and demographic information also indicates that several samples identified as hybrids might actually be advanced generation hybrids. For example, the putative *E. cypellocarpa x E. globulus* hybrids collected at Mallacoota come from one of the first reported examples of a phantom hybrid zone (Kirkpatrick et al. 1973; Parsons and Kirkpatrick 1972; Watson et al. 1987). All trees in the population appear to be intermediate to varying degrees between *E. cypellocarpa* and *E. globulus*, but despite *E. cypellocarpa* occurring nearby, the nearest native *E. globulus* tree is 6.4 km away – hence the “phantom” hybrid zone (Kirkpatrick et al. 1973). After morphological and chemical analysis, a previous study concluded that the population is most likely of hybrid origin and represents a genetic remnant of the past distribution of *E. globulus* (Kirkpatrick et al. 1973). The current population is at sea level and it was hypothesised that the *E. globulus* source population was probably flooded when sea level rose after the last glacial maximum (Kirkpatrick et al. 1973). This situation would result in the population being made up mainly of backcrosses (to *E. cypellocarpa*) or F₂ hybrids. The four samples analysed from this population do appear to fit this expectation with the one sample being consistent with an F₁, F₂ or a backcross, and the other three being most similar to backcrosses towards *E. cypellocarpa* (Fig 3.3; and see Appendix 3.3 for more detail and a discussion of all putative hybrid samples).

### 3.6 Conclusion

The marker set and Bayesian modelling approach implemented here accurately identified simulated and pedigreed first generation hybrids, which was the aim of the
study. The system was tested with more challenging scenarios from mature native forests that possibly included advanced generation hybrids. Despite this, it was concluded that 14 of the 15 unknown samples were correctly allocated, and one somewhat ambiguous sample was possibly an advance generation hybrid. The approach highlighted the power of using multiple lines of evidence, including morphology, the demographic setting of the native forest where the sample was collected, and molecular data, in classifying putative hybrids. The combined evaluation undertaken here has provided validation of natural advanced generation hybrids between *E. globulus* and *E. cypellocarpa*, and *E. globulus* and *E. viminalis*. It also provided confirmation of exotic F₁ hybridisation between *E. globulus* plantations and native *E. camaldulensis*. The database is now available for deployment in the detection of exotic hybrids from plantations in Australia, and in the future could be built upon to include other species and used for comparison with other hybrid systems.
Appendix 3.1: Species descriptions and sample details

_Eucalyptus camaldulensis_ (subgenus _Symphyomyrtus_, section _Exsertaria_, series _Rostratae_).

_Eucalyptus camaldulensis_ is the most naturally widespread eucalypt species, and one of the most widely cultivated outside Australia (Butcher _et al._ 2009). It is a common, often dominant, water course and flood plain tree covering a range of climatic zones across continental Australia (Fig. 3.1; Euclid 2006). There are seven recognised subspecies that are more-or-less geographically structured, and intergrade populations are known to occur where the distributions meet (Euclid 2006). Generally _E. camaldulensis_ is a large tree with smooth bark, seven flowered inflorescences and has fruit with an ascending disc and very exert valves (Euclid 2006). There is considerable variation in leaf and bud morphology across the seven subspecies, but distinctively, all have yellow-brown double-coated seeds (Euclid 2006). The species is common around _E. globulus_ plantations in western Victoria and the Green Triangle. In this area the natural distribution of _E. camaldulensis_ is highly fragmented, with remnant populations often occurring directly adjacent to, and embedded within _E. globulus_ plantations. Of the 97 _E. camaldulensis_ DNA samples used in this study 92 were from Butcher _et al._ (2009), and cover the full geographic range of the species; an additional two samples were collected from open pollinated progeny collected from trees within the plantation zone in the Green Triangle, two samples were collected from Currency Creek Arboretum (CCA), with one sample sourced from existing collections from Petford in Queensland.

_Eucalyptus cypellocarpa_ (subgenus _Symphyomyrtus_, section _Maidenaria_, series _Globulares_, subseries _Remanentes_)

_Eucalyptus cypellocarpa_ is common and widespread in coastal and inland ranges from northern New South Wales, south to eastern and central Victoria, with outlying populations in the Grampians and Otway Ranges in western Victoria (Euclid 2006). It is typically a tall wet forest tree to 65m, and also occasionally occurs as a mallee at the extremes of its range (Euclid 2006). It has smooth bark, long lanceolate leaves and inflorescences are in umbels of seven with buds and fruit often having a
longitudinal ridge. The juvenile foliage is striking, with large opposite sessile leaves on square to winged stems (Euclid 2006), and can resemble the juvenile foliage of *E. globulus* and *E. nitens*. In fact the three species are all occur in the same series (Euclid 2006), and given their similar juvenile characteristics, hybrids between these three species would be difficult to detect based on juvenile morphology. *Eucalyptus cypellocarpa* and *E. globulus* naturally co-occur and hybrids between the two have been reported (Kirkpatrick *et al.* 1973; Parsons and Kirkpatrick 1972). *Eucalyptus cypellocarpa* occurs in the vicinity of industrial *E. globulus* plantations mainly in the Strzelecki Ranges in Gippsland Victoria, where it is a common component of native forests adjacent to plantations. Ninety-eight individuals from 26 populations across the range of *E. cypellocarpa* were sampled in this study. Leaf tissue and herbarium specimens were collected from three to five individuals per population.

*Eucalyptus globulus* (subgenus *Symphyomyrtus*, section *Maidenaria*, series *Globulares*, subseries *Euglobulares*)

*Eucalyptus globulus* is common in coastal and sub-coastal and inland forests below 700m in eastern Tasmania, southern Victoria and the Bass Strait Islands with outlying populations in western Tasmania (Euclid 2006). The species is often tall (up to 90 m) and has smooth bark, large falcate leaves and usually single budded influences (occasionally in 3s; Jordan *et al.* 1993) that are large, warty and ribbed (Euclid 2006). The juvenile foliage is conspicuous with large sessile opposite and highly glaucous leaves on square and winged stems (Euclid 2006). This characteristic juvenile foliage has been exploited as morphological marker for identifying juvenile hybrid seedlings involving *E. globulus* (Barbour *et al.* 2008b; Lopez *et al.* 2000b). It is one the most economically important temperate hardwood species in the world and is widely planted across southern mainland Australia in industrial plantations (Gavran and Parsons 2011; Potts *et al.* 2004). *Eucalyptus globulus* plantations occur within its native range in Gippsland and south-eastern Tasmania (Fig. 3.1). Of the 87 *E. globulus* DNA samples used in this study, 79 were provided by Hudson (2012), and cover the full geographic range of the species. The remaining eight samples were sourced from existing UTAS collections.
Chapter 3: Hybrid detection

*Eucalyptus nitens* (subgenus *Symphyomyrtus*, section *Maidenaria*, series *Globulares*, subseries *Remanentes*)

*Eucalyptus nitens* has a disjunct natural distribution occurring in scattered populations in highland wet forests in Victoria and New South Wales (Euclid 2006). It is a tall tree (to 70 m) with mainly smooth bark, glossy green lanceolate to falcate leaves, small buds in umbels of seven that are angular (Euclid 2006). *Eucalyptus nitens* has not been as widely cultivated as *E. globulus*, but is particularly important in Tasmania where its superior frost resistance is exploited at high altitudes (Beadle *et al.* 1996; Gavran and Parsons 2011; Potts *et al.* 2011). Like *E. globulus* the juvenile foliage is conspicuous with large sessile opposite and glaucous leaves on square and winged stems, but *E. nitens* juveniles can be distinguished from *E. globulus* by their fused apical buds (Euclid 2006). The species does not naturally occur within the main *E. globulus* planting zone but *E. nitens* plantations do occur within the native distribution of *E. globulus* and *E. cypellocarpa* in Gippsland and *E. ovata*, *E. viminalis* and *E. globulus* Tasmania. The 94 *E. nitens* DNA samples used in this study were provided by Hudson (2012), and cover the full geographic range of the species.

*Eucalyptus ovata* (subgenus *Symphyomyrtus*, section *Maidenaria*, series *Foveolatae*)

*Eucalyptus ovata* is widespread and common in poorly drained sites across south-eastern Australia, from Kangaroo Island in the west to the Southern Tablelands of New South Wales in the east, and south to Tasmania (Euclid 2006). It is a small to medium sized tree (6-25 m) that is sometimes multi-stemmed. The bark can be smooth throughout or have loose rough slabs extending up the trunk (Euclid 2006). Buds are in umbels of seven, adult leaves are lanceolate to ovate, and juvenile leaves are petiolate and alternate by node four to six (Euclid 2006). There are two subspecies recognised, the widespread subsp. *ovata* and subsp. *grandiflora* which has larger buds and fruit and prominent oil glands in the adult leaves, and is restricted to far south-west Victoria and the south-east tip of South Australia (Nicolle 2013). *Eucalyptus ovata* co-occurs with *E. globulus* in Tasmania and Victoria and naturally occurring hybrids between the species were among the first recognised in
Chapter 3: Hybrid detection

_Eucalyptus_ (McAulay 1937). The species is also very common in plantation landscapes, especially in the Green Triangle, where it commonly occurs as remnant forest patches adjacent to and embedded within _E. globulus_ plantations (Fig. 3.1; Barbour _et al._ 2008b). Of the 100 _E. ovata_ DNA samples used in this study, 83 were provided by Marthick (2005), which covered the full range of the species, and 17 additional samples were collected from populations within the plantation zone in the Green Triangle and Gippsland in Victoria.

_Eucalyptus viminalis_ (subgenus _Symphyomyrtus_, section _Maidenaria_, series _Viminalaes_, subseries _Lanceolatae_)

_Eucalyptus viminalis_ is widespread in wet or seasonally wet sites across south eastern Australia from the Eyre Peninsula through South Australia, Victoria, Tasmania, and extending as far north as the Northern Tablelands of New South Wales (Euclid 2006). It is a tree with incredible variation in form, from a small coastal mallee to a tall forest tree to 90 m (Euclid 2006). The bark is smooth and white with a persistent stoking of rough bark at the base, sometimes extending up the trunk. The inflorescences are in threes or sevens (see below), adult leaves are lanceolate to falcate, while juvenile leaves are lanceolate opposite and sessile for many pairs (Euclid 2006). There are four subspecies, of which two (subsp. _viminalis_ and subsp. _cygnetensis_) are common in the _E. globulus_ plantation zone in south-eastern mainland Australia (Fig. 3.1; Euclid 2006). Subspecies _cygnetensis_ occurs mainly in South Australia and south-western Victoria and has fruits usually in sevens, with rough bark extending further up the trunk than subspecies _viminalis_ (Euclid 2006). _Eucalyptus viminalis_ naturally co-occurs with _E. globulus_ and occasional hybrids have been reported between the two species (Potts and Wiltshire 1997). Of the 89 _E. viminalis_ DNA samples used in this study, 87 were from Marthick (2005) and covered most of the species range including the main plantation growing areas. An additional two samples were also collected from Tinderbox in southeast Tasmania.
Appendix 3.2: Additional results

Table A.3.2.1. Genetic diversity parameters for six *Eucalyptus* species genotyped at 10 microsatellite loci. Parameters: \( n \) = number of individuals genotyped; \( N_a \) = average number of alleles per locus; \( P_a \) = total number of private alleles; \( H_e \) = expected heterozygosity; \( H_o \) = observed heterozygosity.

<table>
<thead>
<tr>
<th>Species</th>
<th>( n )</th>
<th>( N_a )</th>
<th>( P_a )</th>
<th>( H_e )</th>
<th>( H_o )</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. camaldulensis</em></td>
<td>97</td>
<td>25.2</td>
<td>28</td>
<td>0.89</td>
<td>0.78</td>
</tr>
<tr>
<td><em>E. cypellocarpa</em></td>
<td>97</td>
<td>22.8</td>
<td>13</td>
<td>0.91</td>
<td>0.82</td>
</tr>
<tr>
<td><em>E. globulus</em></td>
<td>87</td>
<td>19.3</td>
<td>6</td>
<td>0.87</td>
<td>0.78</td>
</tr>
<tr>
<td><em>E. nitens</em></td>
<td>88</td>
<td>17.2</td>
<td>4</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td><em>E. ovata</em></td>
<td>100</td>
<td>20.9</td>
<td>5</td>
<td>0.82</td>
<td>0.68</td>
</tr>
<tr>
<td><em>E. viminalis</em></td>
<td>87</td>
<td>23.2</td>
<td>8</td>
<td>0.89</td>
<td>0.82</td>
</tr>
<tr>
<td>Hybrids combined</td>
<td>50</td>
<td>18.3</td>
<td>3 ( ^a )</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>606</td>
<td>21.0</td>
<td>66</td>
<td>0.87</td>
<td>0.78</td>
</tr>
</tbody>
</table>

\(^a\) it is inappropriate to calculate \( H_e \) across multiple species

Table A.3.2.2. Comparison of the accuracy of two- and six-way Bayesian cluster analysis for identifying hybrids. Comparisons are the average across 50 simulated parental and 50 simulated F\(_1\) generations for each combination between *Eucalyptus globulus* and five other species.

<table>
<thead>
<tr>
<th>Combination</th>
<th>Two-way analysis</th>
<th>Six-way analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. camaldulensis x globulus</em></td>
<td>99</td>
<td>85</td>
</tr>
<tr>
<td><em>E. cypellocarpa x globulus</em></td>
<td>84</td>
<td>65</td>
</tr>
<tr>
<td><em>E. nitens x globulus</em></td>
<td>99</td>
<td>92</td>
</tr>
<tr>
<td><em>E. ovata x globulus</em></td>
<td>97</td>
<td>83</td>
</tr>
<tr>
<td><em>E. viminalis x globulus</em></td>
<td>94</td>
<td>69</td>
</tr>
</tbody>
</table>
Figure A.3.2.1. Proportion of genotype membership of 556 individuals from six eucalyptus species based on Bayesian cluster analysis (in STRUCTURE) at $K = 6$, using no *a priori* species information. All individuals were genotyped at 10 microsatellite loci. Each individual is represented by a single vertical line that is partitioned based on its genotype affinities to one of six genetic clusters (represented by the six colours), and samples are ordered based on their morphological species classification. The genotype affinities of the samples clearly correspond to the species classifications indicating that STRUCTURE differentiates the six species at $K = 6$. 
Appendix 3.3: Detailed discussion regarding the classification putative hybrids, and detailed sample allocation data

Of the 15 putative hybrids 12 were classified as hybrids and three could not be distinguished from their pure parents in the two-way analysis. Of the three indistinguishable samples, the two putative *E. cypellocarpa* F1s are probably correctly classified as pure *E. cypellocarpa*. These two samples were collected from seedlings growing among native *E. cypellocarpa* beside a 10 year old *E. globulus* plantation in Gippsland in the absence of native *E. globulus* – ruling out the possibility that they are backcrosses (i.e. no mature F1s available to facilitate backcrossing). As noted previously the seedling morphology of these two species is very similar and can be difficult to distinguish, and these two samples were only tentatively classified as “possible” F1s (with a low degree of certainty) based on the degree of glaucousness, which can be a highly variable trait, even within species (Euclid 2006). Therefore given that they overlap the 95% confidence intervals of simulated pure *E. cypellocarpa*, and there was low confidence in the morphological assessment, it is likely that they are correctly allocated as pure *E. cypellocarpa*.

The situation with the putative *E. viminalis* x *globulus* hybrid that was classified as *E. globulus* is less clear. This sample was identified in open pollinated *E. globulus* seed, collected from native forest in southern Tasmania, where the species co-occur in the absence of other cross-compatible eucalypts in close proximity. The sample showed distinctively intermediate morphology, with linear, non-glaucous leaves compared to the broadly-linear, glaucous leaves of *E. globulus* and the narrowly-linear non-glaucous leaves of *E. viminalis*. The sample could be a morphologically unusual pure *E. globulus* that is coincidently intermediate between the species. However, the likelihood of random morphological deviations on multiple traits resulting in intermediate characteristics is low, and is more easily explained by inter-specific hybridisation (Hopper 1978; Potts and Dungey 2004). Therefore, considering the model inaccuracy when identifying backcrosses, the native forest setting of the mother, and the morphology, it is possible that this sample is actually a first or perhaps later generation backcross.
Chapter 3: Hybrid detection

Of the four putative *E. camaldulensis* x *globulus* samples collected from open pollinated seed, three were within the 95% confidence intervals of the mean of the simulated F$_1$’s, providing good evidence they are F$_1$ hybrids between *E. camaldulensis* and *E. globulus* Table A.3.3.1. This supports Barbour et al. (2008b) who also identified putative *E. camaldulensis* x *globulus* F$_1$’s in seedlots from the same region based on morphology. The fourth sample from this seedlot had mean $q$ values outside the 95% confidence intervals of the simulated F$_1$’s, but its 95% confidence intervals did include the mean of the simulated F$_1$’s (Table A.3.3.1). These four samples were classified as putative F$_1$’s because the age of plantation neighbouring the mothers (10 years) made it unlikely that a mature F$_1$ hybrid could occur to facilitate backcrossing. Moreover, the $q$ values of the fourth sample resemble a backcross to *E. globulus* (Table A.3.3.1), which is impossible given that it came from an *E. camaldulensis* seedlot. The sample could be a hybrid with another species, although the six-way analysis shows that none of the other species in the model were involved, despite *E. ovata* and *E. viminalis* occurring in the area where the seed was collected. Therefore given the somewhat divergent nature of this sample it could be an exotic F$_1$, or it could be a hybrid with an unknown species.

The putative *E. viminalis* x *globulus* hybrids (generation unknown) were mature trees identified on the basis of intermediate capsule morphology within native forests, and were classified as hybrids in the two-way model. Both natural and manipulated hybridisation between the species has been reported (Potts and Wiltshire 1997). All three samples here have $q$ values that are more consistent with backcrosses towards *E. globulus* than F$_1$’s or F$_2$’s (Table A.3.3.5). They were collected from native forests that were dominated by *E. globulus* with scattered *E. viminalis*. This demographic ratio might be conducive to the formation of hybrids due to pollen swamping by the more numerous species (Field et al. 2008), and in such a situation backcross hybrids towards *E. globulus* would be common. Patterns of asymmetrical gene flow in native eucalypts have also arisen through differences in flower size (Field et al. 2011). It has been shown that in many cases the pollen tube of small flowered species cannot reach the ovaries of large flowered species, while the reverse cross can work (Gore et al. 1990). *Eucalyptus globulus* has significantly larger flowers than *E. viminalis* (Euclid 2006) again making backcrosses towards *E.
Chapter 3: Hybrid detection

globulus more likely. Therefore these samples are most likely backcrosses to E. globulus as suggested by the admixture analysis, rather than F₁ or F₂ hybrids.

The five putative E. cypellocarpa x globulus hybrids (generation unknown) have q values that are consistent with F₁s/F₂s or backcrosses towards E. cypellocarpa (Table A.3.3.2). The sample from a mature tree in native forest at Moonlight Head has q values that fall within the 95% confidence intervals of the simulated F₁s. It was found in forest where both species were common, and was unique among the trees inspected in that it had intermediate capsule morphology. Natural F₁ hybrids between these species have been reported from mixed stands elsewhere in Victoria (Kirkpatrick 1974; Kirkpatrick et al. 1973) and the simplest explanation is probably that this sample is an F₁ that has survived to reproductive maturity.

The hybrid status of the population where the four Mallacoota samples were collected has been the focus of considerable study (Kirkpatrick et al. 1973; Parsons and Kirkpatrick 1972; Watson et al. 1987). All trees in the population appear to be intermediate to varying degrees between E. cypellocarpa and E. globulus. Despite nearby sources of E. cypellocarpa the nearest E. globulus is 6.4 km away. This led to the population being one of the first reported examples of a phantom hybrid zone (Kirkpatrick et al. 1973; Parsons and Kirkpatrick 1972). There are several other well documented examples of hybrid swarms between the two species in river valleys to the west of Mallacoota where both species are present (Kirkpatrick 1974; Kirkpatrick et al. 1973; Parsons and Kirkpatrick 1972). The trees in the Mallacoota population occur basically at sea level and are large, probably ranging in age from tens to hundreds of years old. Morphometric analysis showed that none of the trees could be definitively classified as E. cypellocarpa or E. globulus, but two trees in particular were very E. globulus-like (Kirkpatrick et al. 1973). However, terpene analysis found all trees to be indistinguishable from E. cypellocarpa (Kirkpatrick et al. 1973). It was concluded that the population is probably of hybrid origin and represents a genetic remnant of the past distribution of E. globulus that was flooded when sea level rose after the last glacial maximum (Kirkpatrick et al. 1973). Although we only analysed four samples from this population, they do appear to fit with the Kirkpatrick et al. (1973) hypothesis, with one sample being consistent with
an F₁ or F₂ (although its 95% CIs did include the simulated backcross means; Table A.3.3.2) and the other three being more similar to backcrosses towards *E. cypellocarpa* (Table A.3.3.2).

**Table A.3.3.1.** Assignment of simulated, pedigreed and putative hybrid samples between *E. camaldulensis* and *E. globulus* obtained from STRUCTURE using a *q* cut-off of 0.2.

<table>
<thead>
<tr>
<th>Simulated samples</th>
<th>n</th>
<th>% correctly assigned at q&gt;0.2</th>
<th>E. camaldulensis cluster</th>
<th>E. globulus cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>mean q</td>
<td>95% CI (+ -)</td>
</tr>
<tr>
<td>Simulated <em>E. camaldulensis</em></td>
<td>50</td>
<td>98</td>
<td>0.902</td>
<td>(0.910,0.891)</td>
</tr>
<tr>
<td>Simulated <em>E. globulus</em></td>
<td>50</td>
<td>100</td>
<td>0.084</td>
<td>(0.093,0.076)</td>
</tr>
<tr>
<td>Simulated F1</td>
<td>50</td>
<td>100</td>
<td>0.530</td>
<td>(0.560,0.501)</td>
</tr>
<tr>
<td>Simulated F2</td>
<td>50</td>
<td>90</td>
<td>0.526</td>
<td>(0.573,0.478)</td>
</tr>
<tr>
<td>Simulated BCc</td>
<td>50</td>
<td>60</td>
<td>0.738</td>
<td>(0.770,0.706)</td>
</tr>
<tr>
<td>Simulated BCg</td>
<td>50</td>
<td>80</td>
<td>0.305</td>
<td>(0.339,0.270)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pedigreed and putative hybrids</th>
<th>n</th>
<th>assignment at q&gt;0.2</th>
<th>mean q</th>
<th>95% CI (+ -)</th>
<th>mean q</th>
<th>95% CI (+ -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.506</td>
<td>(0.247,0.768)</td>
<td>0.495</td>
<td>(0.232,0.753)</td>
</tr>
<tr>
<td>Pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.580</td>
<td>(0.308,0.853)</td>
<td>0.421</td>
<td>(0.147,0.692)</td>
</tr>
<tr>
<td>Putative F1*</td>
<td>1</td>
<td>hybrid</td>
<td>0.335</td>
<td>(0.108,0.587)</td>
<td>0.665</td>
<td>(0.413,0.892)</td>
</tr>
<tr>
<td>Putative F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.536</td>
<td>(0.271,0.799)</td>
<td>0.464</td>
<td>(0.201,0.729)</td>
</tr>
<tr>
<td>Putative F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.583</td>
<td>(0.303,0.854)</td>
<td>0.417</td>
<td>(0.146,0.697)</td>
</tr>
<tr>
<td>Putative F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.416</td>
<td>(0.171,0.685)</td>
<td>0.584</td>
<td>(0.315,0.829)</td>
</tr>
</tbody>
</table>

* Divergent sample referred to in the text
### Table A.3.3.2. Assignment of simulated, pedigreed and putative hybrid samples between *E. cypellocarpa* and *E. globulus* obtained from STRUCTURE using a $q$ cut-off of 0.2.

Putative hybrid' = putative hybrid samples collected from mature trees in native forest where the generation is unknown.

<table>
<thead>
<tr>
<th>Simulated samples</th>
<th>n</th>
<th>% correctly assigned at $q&gt;0.2$</th>
<th>$E. cypellocarpa$ cluster</th>
<th>$E. globulus$ cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>mean $q$</td>
<td>95% CI (+ -)</td>
</tr>
<tr>
<td>Simulated <em>E. cypellocarpa</em></td>
<td>50</td>
<td>70</td>
<td>0.815</td>
<td>(0.842,0.789)</td>
</tr>
<tr>
<td>Simulated <em>E. globulus</em></td>
<td>50</td>
<td>72</td>
<td>0.165</td>
<td>(0.184,0.148)</td>
</tr>
<tr>
<td>Simulated F1</td>
<td>50</td>
<td>98</td>
<td>0.490</td>
<td>(0.524,0.456)</td>
</tr>
<tr>
<td>Simulated F2</td>
<td>50</td>
<td>100</td>
<td>0.488</td>
<td>(0.526,0.450)</td>
</tr>
<tr>
<td>Simulated BCc</td>
<td>50</td>
<td>80</td>
<td>0.677</td>
<td>(0.714,0.639)</td>
</tr>
<tr>
<td>Simulated BCg</td>
<td>50</td>
<td>86</td>
<td>0.339</td>
<td>(0.378,0.299)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>300</strong></td>
<td></td>
<td><strong>84.4</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pedigreed and putative hybrids</th>
<th>n</th>
<th>assignment at $q&gt;0.2$</th>
<th>mean $q$</th>
<th>95% CI (+ -)</th>
<th>mean $q$</th>
<th>95% CI (+ -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>putative hybrid'</td>
<td>1</td>
<td>hybrid</td>
<td>0.563</td>
<td>(0.155,0.940)</td>
<td>0.437</td>
<td>(0.060,0.845)</td>
</tr>
<tr>
<td>putative hybrid'</td>
<td>1</td>
<td>hybrid</td>
<td>0.669</td>
<td>(0.322,0.960)</td>
<td>0.331</td>
<td>(0.040,0.678)</td>
</tr>
<tr>
<td>putative hybrid'</td>
<td>1</td>
<td>hybrid</td>
<td>0.628</td>
<td>(0.278,0.941)</td>
<td>0.372</td>
<td>(0.059,0.722)</td>
</tr>
<tr>
<td>putative hybrid'</td>
<td>1</td>
<td>hybrid</td>
<td>0.726</td>
<td>(0.347,0.983)</td>
<td>0.274</td>
<td>(0.017,0.653)</td>
</tr>
<tr>
<td>putative hybrid'</td>
<td>1</td>
<td>hybrid</td>
<td>0.494</td>
<td>(0.127,0.868)</td>
<td>0.506</td>
<td>(0.132,0.873)</td>
</tr>
<tr>
<td>putative F1</td>
<td>1</td>
<td><em>E. cypellocarpa</em></td>
<td>0.864</td>
<td>(0.616,0.996)</td>
<td>0.136</td>
<td>(0.004,0.384)</td>
</tr>
<tr>
<td>putative F1</td>
<td>1</td>
<td><em>E. cypellocarpa</em></td>
<td>0.848</td>
<td>(0.568,0.995)</td>
<td>0.152</td>
<td>(0.005,0.432)</td>
</tr>
</tbody>
</table>

* Samples from the putative phantom hybrid zone at Mallacoota discussed it the text
Table A.3.3.3. Assignment of simulated, pedigreed and putative hybrid samples between *E. nitens* and *E. globulus* obtained from STRUCTURE using a \( q \) cut-off of 0.2.

<table>
<thead>
<tr>
<th>Simulated samples</th>
<th>( n )</th>
<th>% correctly assigned at ( q&gt;0.2 )</th>
<th>( E. nitens ) cluster</th>
<th>( E. globulus ) cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>mean ( q )</td>
<td>95% CI (+-)</td>
</tr>
<tr>
<td>Simulated <em>E. nitens</em></td>
<td>50</td>
<td>98</td>
<td>0.917</td>
<td>(0.927,0.907)</td>
</tr>
<tr>
<td>Simulated <em>E. globulus</em></td>
<td>50</td>
<td>98</td>
<td>0.086</td>
<td>(0.096,0.075)</td>
</tr>
<tr>
<td>Simulated F1</td>
<td>50</td>
<td>100</td>
<td>0.504</td>
<td>(0.527,0.482)</td>
</tr>
<tr>
<td>Simulated F2</td>
<td>50</td>
<td>88</td>
<td>0.520</td>
<td>(0.567,0.474)</td>
</tr>
<tr>
<td>Simulated BCn</td>
<td>50</td>
<td>66</td>
<td>0.728</td>
<td>(0.760,0.696)</td>
</tr>
<tr>
<td>Simulated BCg</td>
<td>50</td>
<td>68</td>
<td>0.256</td>
<td>(0.289,0.223)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>300</td>
<td><strong>86.3</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pedigreed hybrid samples</th>
<th>( n )</th>
<th>assignment at ( q&gt;0.2 )</th>
<th>mean ( q )</th>
<th>95% CI (+-)</th>
<th>mean ( q )</th>
<th>95% CI (+-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.6562</td>
<td>(0.408,0.869)</td>
<td>0.3438</td>
<td>(0.131,0.592)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.442</td>
<td>(0.153,0.721)</td>
<td>0.558</td>
<td>(0.279,0.847)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.5108</td>
<td>(0.267,0.743)</td>
<td>0.4892</td>
<td>(0.257,0.733)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.5187</td>
<td>(0.274,0.753)</td>
<td>0.4813</td>
<td>(0.247,0.726)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.4818</td>
<td>(0.191,0.764)</td>
<td>0.5182</td>
<td>(0.236,0.809)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.3872</td>
<td>(0.150,0.645)</td>
<td>0.6128</td>
<td>(0.355,0.850)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1659</td>
<td>(0.011,0.402)</td>
<td>0.8341</td>
<td>(0.598,0.989)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.4274</td>
<td>(0.169,0.700)</td>
<td>0.5726</td>
<td>(0.300,0.831)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.4327</td>
<td>(0.164,0.694)</td>
<td>0.5673</td>
<td>(0.306,0.836)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.2366</td>
<td>(0.041,0.491)</td>
<td>0.7634</td>
<td>(0.509,0.959)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.4423</td>
<td>(0.178,0.711)</td>
<td>0.5577</td>
<td>(0.289,0.822)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>1</td>
<td>hybrid</td>
<td>0.427</td>
<td>(0.175,0.688)</td>
<td>0.573</td>
<td>(0.312,0.825)</td>
</tr>
<tr>
<td>pedigreed F2</td>
<td>1</td>
<td>hybrid</td>
<td>0.6063</td>
<td>(0.353,0.833)</td>
<td>0.3937</td>
<td>(0.167,0.647)</td>
</tr>
<tr>
<td>pedigreed F2</td>
<td>1</td>
<td>hybrid</td>
<td>0.4833</td>
<td>(0.218,0.736)</td>
<td>0.5167</td>
<td>(0.264,0.782)</td>
</tr>
<tr>
<td>pedigreed F2</td>
<td>1</td>
<td>hybrid</td>
<td>0.5528</td>
<td>(0.271,0.814)</td>
<td>0.4472</td>
<td>(0.186,0.729)</td>
</tr>
<tr>
<td>pedigreed F2</td>
<td>1</td>
<td>hybrid</td>
<td>0.3944</td>
<td>(0.137,0.659)</td>
<td>0.6056</td>
<td>(0.341,0.863)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1713</td>
<td>(0.012,0.404)</td>
<td>0.8287</td>
<td>(0.596,0.988)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1615</td>
<td>(0.011,0.388)</td>
<td>0.8385</td>
<td>(0.612,0.989)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1613</td>
<td>(0.018,0.383)</td>
<td>0.8387</td>
<td>(0.617,0.982)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1434</td>
<td>(0.009,0.346)</td>
<td>0.8566</td>
<td>(0.654,0.991)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.185</td>
<td>(0.022,0.406)</td>
<td>0.815</td>
<td>(0.594,0.978)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td>hybrid</td>
<td>0.2244</td>
<td>(0.028,0.466)</td>
<td>0.7756</td>
<td>(0.534,0.972)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td>hybrid</td>
<td>0.2247</td>
<td>(0.051,0.435)</td>
<td>0.7753</td>
<td>(0.565,0.949)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td>hybrid</td>
<td>0.2204</td>
<td>(0.038,0.445)</td>
<td>0.7796</td>
<td>(0.555,0.962)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td>hybrid</td>
<td>0.2713</td>
<td>(0.062,0.527)</td>
<td>0.7287</td>
<td>(0.473,0.938)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1317</td>
<td>(0.004,0.356)</td>
<td>0.8683</td>
<td>(0.644,0.996)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.1866</td>
<td>(0.028,0.410)</td>
<td>0.8134</td>
<td>(0.590,0.972)</td>
</tr>
<tr>
<td>pedigreed BCg</td>
<td>1</td>
<td><em>E. globulus</em></td>
<td>0.115</td>
<td>(0.010,0.284)</td>
<td>0.885</td>
<td>(0.716,0.990)</td>
</tr>
</tbody>
</table>
### Chapter 3: Hybrid detection

#### Table A.3.3.4. Assignment of simulated, pedigreed and putative hybrid samples between *E. ovata* and *E. globulus* obtained from STRUCTURE using a $q$ cut-off of 0.2.

<table>
<thead>
<tr>
<th>Simulated samples</th>
<th>$n$</th>
<th>$%$ correctly assigned at $q&gt;0.2$</th>
<th>$E. ovata$ cluster</th>
<th>$E. globulus$ cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated <em>E. ovata</em></td>
<td>50</td>
<td>100</td>
<td>0.925 (0.933,0.917)</td>
<td>0.075 (0.083,0.067)</td>
</tr>
<tr>
<td>Simulated <em>E. globulus</em></td>
<td>50</td>
<td>98</td>
<td>0.090 (0.102,0.078)</td>
<td>0.910 (0.922,0.898)</td>
</tr>
<tr>
<td>Simulated F1</td>
<td>50</td>
<td>94</td>
<td>0.477 (0.514,0.440)</td>
<td>0.523 (0.560,0.486)</td>
</tr>
<tr>
<td>Simulated F2</td>
<td>50</td>
<td>98</td>
<td>0.502 (0.544,0.461)</td>
<td>0.498 (0.539,0.456)</td>
</tr>
<tr>
<td>Simulated BCg</td>
<td>50</td>
<td>74</td>
<td>0.278 (0.313,0.242)</td>
<td>0.708 (0.744,0.673)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>300</strong></td>
<td><strong>89.3</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Pedigreed samples**

<table>
<thead>
<tr>
<th>Pedigreed samples</th>
<th>hybrid assignment at $q&gt;0.2$</th>
<th>$E. ovata$ cluster</th>
<th>$E. globulus$ cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>pedigreed F1</td>
<td>hybrid</td>
<td>0.5365 (0.266,0.804)</td>
<td>0.4635 (0.196,0.734)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>hybrid</td>
<td>0.5316 (0.273,0.788)</td>
<td>0.4684 (0.212,0.727)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>300</strong></td>
<td><strong>86.7</strong></td>
<td></td>
</tr>
</tbody>
</table>

#### Table A.3.3.5. Assignment of simulated, pedigreed and putative hybrid samples between *E. viminalis* and *E. globulus* obtained from STRUCTURE using a $q$ cut-off of 0.2. Putative hybrid = putative hybrid samples collected from mature trees in native forest where the generation is unknown.

<table>
<thead>
<tr>
<th>Simulated samples</th>
<th>$n$</th>
<th>$%$ correctly assigned at $q&gt;0.2$</th>
<th>$E. viminalis$ cluster</th>
<th>$E. globulus$ cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated <em>E. viminalis</em></td>
<td>50</td>
<td>88</td>
<td>0.868 (0.886,0.849)</td>
<td>0.132 (0.151,0.114)</td>
</tr>
<tr>
<td>Simulated <em>E. globulus</em></td>
<td>50</td>
<td>88</td>
<td>0.133 (0.151,0.115)</td>
<td>0.867 (0.885,0.849)</td>
</tr>
<tr>
<td>Simulated F1</td>
<td>50</td>
<td>100</td>
<td>0.507 (0.544,0.470)</td>
<td>0.493 (0.530,0.456)</td>
</tr>
<tr>
<td>Simulated F2</td>
<td>50</td>
<td>90</td>
<td>0.487 (0.537,0.438)</td>
<td>0.513 (0.562,0.462)</td>
</tr>
<tr>
<td>Simulated BCv</td>
<td>50</td>
<td>78</td>
<td>0.717 (0.749,0.688)</td>
<td>0.281 (0.312,0.251)</td>
</tr>
<tr>
<td>Simulated BCg</td>
<td>50</td>
<td>76</td>
<td>0.284 (0.314,0.254)</td>
<td>0.716 (0.746,0.685)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>300</strong></td>
<td><strong>86.7</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Pedigree and putative hybrids**

<table>
<thead>
<tr>
<th>Pedigree and putative hybrids</th>
<th>hybrid assignment at $q&gt;0.2$</th>
<th>$E. viminalis$ cluster</th>
<th>$E. globulus$ cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>pedigreed F1</td>
<td>hybrid</td>
<td>0.6896 (0.378,0.944)</td>
<td>0.3104 (0.056,0.622)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>hybrid</td>
<td>0.5601 (0.276,0.840)</td>
<td>0.4399 (0.160,0.724)</td>
</tr>
<tr>
<td>pedigreed F1</td>
<td>hybrid</td>
<td>0.4754 (0.200,0.759)</td>
<td>0.5246 (0.241,0.800)</td>
</tr>
<tr>
<td>putative hybrid</td>
<td>hybrid</td>
<td>0.3296 (0.104,0.591)</td>
<td>0.6704 (0.409,0.896)</td>
</tr>
<tr>
<td>putative hybrid</td>
<td>hybrid</td>
<td>0.2058 (0.019,0.464)</td>
<td>0.7942 (0.536,0.981)</td>
</tr>
<tr>
<td>putative hybrid</td>
<td>hybrid</td>
<td>0.3522 (0.083,0.654)</td>
<td>0.6478 (0.346,0.917)</td>
</tr>
<tr>
<td>putative hybrid</td>
<td><em>E. globulus</em></td>
<td>0.1137 (0.002,0.355)</td>
<td>0.8863 (0.645,0.998)</td>
</tr>
</tbody>
</table>
Chapter 4:
Assessing the risk of exotic gene flow from *Eucalyptus globulus* plantations to native *E. ovata* forests

This chapter is published in *Forest Ecology and Management*

4.1 Abstract

The recent rapid expansion of *Eucalyptus* plantations in Australia has raised concern that exotic gene flow could pose a genetic risk to indigenous eucalypt species. The most widely used plantation species *Eucalyptus globulus*, now covers about 538,000 ha of southern Australia. *Eucalyptus ovata* is a common plantation neighbour, and this paper assesses the likelihood of exotic gene flow from *E. globulus* plantations to remnant *E. ovata* forests by assessing three key processes. Firstly, the effect of patch size on the rate of F₁ hybridisation was investigated. Open pollinated seed was collected off native *E. ovata* trees from five patch-size classes in and around plantations (142 trees from 25 patches). Hybridisation was then assessed in 24,322 open pollinated progeny. The overall rate of exotic hybridisation was 1.62%, with small patches having higher rates than larger patches, as did trees on the edges of patches and paddock trees in open pasture 50 – 200 m from the plantation edge. Secondly, natural hybrid establishment was investigated along 3.9 km of plantation-native forest boundary that was classified as being high risk for hybridisation and establishment. Of the 745 seedlings found, only 1% of seedlings were identified as *E. ovata x globulus* F₁ hybrids, with 73% being pure *E. ovata* and 26% pure *E. globulus* seedlings (wildlings). Finally, evidence for selection against hybrids was
found when assessing the survival of naturally established \( E. \ globulus \times ovata \) \( F_1 \)’s that were paired with \( E. \ ovata \) in fire affected native forest beside a plantation. After six years the exotic hybrids were 78 % less likely to survive than the pure \( E. \ ovata \) seedlings. It is concluded that hybrid fitness appears to be an important barrier to exotic gene flow and introgression between \( E. \ globulus \) and \( E. \ ovata \). However, \( E. \ globulus \) wildlings are establishing along the boundaries of native forest, and could have a greater impact on the integrity of \( E. \ ovata \) populations close to plantations.

4.2 Introduction

Genetic contamination of natural populations has received considerable attention, especially in the context of genetically modified (GM) crops. Pollen mediated gene flow from GM crops has resulted in the movement of transgenes into native populations of maize, soybean wheat and canola (Liu et al. 2010; Pineyro-Nelson et al. 2009; Rieben et al. 2011; Wang and Li 2012), and management of the issue has been a primary research focus (Devos et al. 2012; Husken et al. 2010; Middelhoff et al. 2011; Tricault et al. 2011). Genetic contamination of native populations may also occur through gene flow from exotic non-GM taxa that are cross compatible. This issue is receiving growing attention in agriculture (crop-wild gene flow), as well as in industries that involve the large-scale biological releases of individuals with exotic, manipulated or restricted genetic diversity, as can occur in fisheries, game management, revegetation and forestry (Byrne and Stone 2011; Byrne et al. 2011; Friedman and Adams 1985; Laikre et al. 2010; Potts et al. 2003). For example, in order to bolster game stocks in Europe, millions of partridges (including the exotic hybrid \( A. \ rufa \times A. \ chucker \)) are released annually, and this has resulted in the wide spread introgression of exotic \( A. \ chucker \) genes into natural \( A. \ rufa \) populations (Baratti et al. 2005; Friedman and Adams 1985; Negro et al. 2001; Tejedor et al. 2007).

Exotic gene flow may arise from hybridisation between related species (inter-specific), or between genetically differentiated subspecies, variants or populations within species (intra-specific). In certain situations exotic gene flow could be advantageous (Broadhurst et al. 2008; Weeks et al. 2011). Such gene flow can increase genetic diversity in fragmented, inbred populations and increase their
resilience in the face of climate change (Broadhurst et al. 2008; Weeks et al. 2011). Indeed some authors suggest that mixing of population gene pools is actually an under-utilised conservation tool (Weeks et al. 2011). On the other hand, exotic gene flow may threaten the genetic integrity of species and populations (Byrne et al. 2011). For example, exotic gene flow can cause outbreeding depression through disruption or loss of locally adaptive gene complexes, reduce the genetic distinctiveness of unique populations or species, and in extreme cases cause population replacement through genetic swamping (Byrne et al. 2011; Ellstrand et al. 1999; Laikre et al. 2010; Levin et al. 1996).

Plantation forestry is a potentially significant source of exotic germplasm, particularly where large and fecund plantations are embedded in a landscape with cross compatible native species (Barbour et al. 2008b; Laikre et al. 2010; Potts et al. 2003; Vanden Broeck et al. 2005). If pollen can disperse from these plantations and successfully hybridise with native species there is a risk to the genetic integrity of the native populations. For example, the native European black poplar (*Populus nigra*) is threatened by gene flow from hybrid plantations of American cottonwoods (*P. deltoides* and *P. trichocarpa*) (Vanden Broeck et al. 2005). Similar genetic risks to native poplars also occur in North America, but in this case from both locally exotic (American species from outside their natural range) and introduced European species and hybrids (Talbot et al. 2012; Vanden Broeck et al. 2005).

In Australia there are around 1 million ha of eucalypt plantations, and the risk of exotic gene flow from these plantations to the native eucalypt flora has been the focus of significant research (Barbour et al. 2002; Barbour et al. 2008a; Barbour et al. 2008b; Byrne et al. 2011; Laikre et al. 2010; Potts et al. 2003). In particular, the case of *E. nitens* on the island of Tasmania has been used as a model system to study exotic gene flow from *Eucalyptus* plantations (Barbour et al. 2003; Barbour et al. 2005a; Barbour et al. 2005b; Barbour et al. 2006a; Potts et al. 2003). This research has identified (among other things): that exotic hybrids are establishing in the wild; that understanding pollen dispersal and flowering overlap is essential for predicting the extent of gene flow; and that poor F$_1$ hybrid fitness is at least a partial barrier to introgression (Barbour et al. 2003; Barbour et al. 2005b; Barbour et al. 2006a).
Chapter 4:  

The challenge now is to expand this research to the *E. globulus* estate, which is more than twice as large as the *E. nitens* estate and occurs mostly on mainland Australia, where the majority of the 898 native eucalypt taxa occur (Euclid 2006). There are currently around 538,000 ha of *E. globulus* plantations in Australia, occurring mainly outside the species natural range (Fig. 4.1a; Gavran and Parsons 2011). Barbour *et al.* (2008b) showed that 48% of *E. globulus* plantations were growing adjacent to potentially cross-compatible native eucalypt species. Research is now needed to identify situations where *E. globulus* plantations pose a significant genetic risk to these neighbouring species.

Australian *E. globulus* plantations typically occur within a landscape of fragmented remnant native forest and or farm land (Lindenmayer and Hobbs 2004). These native forest remnants are of conservation significance, providing habitat and other ecosystem services, and are often protected by legislation (Lindenmayer and Hobbs 2004). They are also useful for studying the factors that affect hybridisation in complex landscapes and assessing the genetic risk posed to remnant native forests by plantations. Pollen dispersal curves have been generated for plantation *E. nitens* (Barbour *et al.* 2005b) and for native *E. globulus* (Mimura *et al.* 2009) and they show classic leptokurtic distributions, where most pollen (> 90%) is deposited close to the plantation (within 200 m), but with low levels of much longer distance dispersal – to at least 1.6 km in *E. nitens* (Barbour *et al.* 2005b). However, several studies have shown that distance alone is not necessarily a good predictor of gene flow potential (Byrne *et al.* 2008; Cresswell and Osborne 2004; Ellstrand 1992b; Ellstrand 2003; Ellstrand and Marshall 1985; Llorens *et al.* 2012; Richards *et al.* 1999). This is because source-sink relationships (the relative proportion of exotic to native pollen) are often more important in predicting the likelihood of gene flow (Cresswell and Osborne 2004; Ellstrand 2003; Field *et al.* 2008). Given the fragmented mosaic of native forests in plantation landscapes in Australia, these source-sink relationships are likely to be important in determining the risk of exotic gene flow from *E. globulus*.

Another factor influencing genetic risk is hybrid fitness (Potts *et al.* 2003). Introgression requires hybridisation followed by backcrossing, meaning the ability of
first generation (F₁) exotic hybrids to reach reproductive maturity is critical. Hybrid fitness in eucalypts is likely to decrease as genetic divergence between the parental taxa increase (Potts et al. 2003). There are a range of stages where hybrid fitness can be compromised – starting from seed development and maturity through to seedling establishment, growth, and reproductive capacity (Costa e Silva et al. 2012; Dickinson et al. 2012b; Potts et al. 2003). Indeed hybrid fitness was identified by Barbour et al. (2006a) as an important potential barrier to gene flow from plantation *E. nitens* in Tasmania. Hybrid fitness is also likely to be important in determining the risk of introgression between *E. globulus* and neighbouring native species.

One of the most common native species neighbouring *E. globulus* plantations in south eastern Australia is *E. ovata* (Fig. 4.1a). It is a widespread species, but its common occurrence in lowland agricultural landscapes has resulted in the species becoming highly fragmented as a result of deforestation, often surviving on road sides or as remnant patches in open farmland. Consequently, several forest communities that are characterised by the presence of *E. ovata* are now listed as threatened under both state and federal legislation (Commonwealth of Australia 1999; Tasmanian Government 1995). Barbour et al. (2008b) identified *E. ovata* as one of the species most at risk of gene flow from *E. globulus* because of its close taxonomic relationship, flowering synchrony, cross-compatibility and frequent proximity to plantations. In this paper we study the risk of gene flow from *E. globulus* plantations to neighbouring *E. ovata* populations by investigating in situ the following three questions: 1) what effect does patch size and proximity to the plantation have on F₁ hybridisation rate; 2) at what rate are F₁ hybrids establishing in the wild; and 3) what is the relative fitness of F₁ hybrids in of terms survival, compared to *E. ovata*. 
Figure 4.1. a) the study area in south eastern Australia showing the natural distribution of Eucalyptus globulus (orange) and the distribution of E. globulus plantations in green. The red triangles show the location of hybrid establishment survey sites in Tasmania, Gippsland and the Green Triangle. The patch size and hybrid fitness studies were undertaken at the same location in the Green Triangle. b) Aerial photograph of a section of the plantation landscape used in the patch-size study showing examples of the five remnant classes (RC 1 – RC 5; see text and Table 4.1). Note that RC 1 classes actually sampled did not have any trees between them and the plantation as in this example. Circles show the typical distribution and location of trees from which seed was collected within the patches (orange circles = edge trees; yellow circles = central trees; see text).
Chapter 4: \textit{E. ovata} case study

4.3 Materials and Methods

4.3.1 \textit{The effect of patch size on hybridisation rate in a fragmented \textit{E. ovata} landscape}

A landscape in the Green Triangle region of south-western Victoria was identified where \textit{E. ovata} remnant forest formed a mosaic with harvest-age \textit{E. globulus} plantations (8 to 12 years old), and farmland (Fig. 4.1b). The study landscape covered approximately c.7550 ha, of which c. 2850 ha consisted of \textit{E. globulus} plantation, c.1150 ha of remnant native forest and c. 3550 ha of pasture. The remnant native forest occurred in large patches (> 100 ha), and also as smaller remnants adjacent to plantations, surrounded by pasture, and embedded within the plantation matrix (Fig. 4.1b). In order to determine the effect of patch size on hybridisation rate, five remnant classes were identified: 1) small remnants with one to 10 trees in paddocks 50 to 200 m from a plantation edge (paddocks in Australia are dominated by wind pollinated monocot grasses that do attract pollinators); 2) small remnants with 1 to 30 trees (but usually 10 or less) totally surrounded by plantation (embedded); 3) small remnants of around 50 trees embedded within a plantation; 4) large remnants with over 100 trees embedded within a plantation; 5) continuous native forest (>1000 trees) adjacent to a plantation (not embedded but adjoining the plantation) (Fig. 4.1; Table 4.1).

Open pollinated seed capsules were collected from 32 \textit{E. ovata} trees in every remnant class except class 1 where only 14 suitable trees could be located (Table 4.1). Seed was collected from up to four trees (remnant classes 1 and 2 often had less than four trees) at the edge of each remnant. In order to detect patterns of pollen movement within remnants, capsules were also collected from trees 50 m inside the remnants in classes 3 (two trees/remnant), 4 and 5 (four trees/remnant). Capsules were always collected from the canopy side closest to the plantation and from as high in the canopy as possible (c. 3 – 5 m). The reproductive output of the adjacent plantation was assessed for classes 2 – 5 (class 1 were isolated trees with no adjoining plantation). These assessments involved using binoculars to visually estimate capsule abundance on a logarithmic scale (see Barbour et al (2008b) for details). Assessments of reproductive capacity were carried out on \textit{E. globulus} trees...
on the plantation edge, with 15 adjacent trees assessed for remnant classes 2, 3 and 4; and 20 trees for class 5.

Table 4.1. Description of the five remnant classes used to detect the influence of patch size on the likelihood of exotic gene flow from _Eucalyptus globulus_ plantations to native _E. ovata_ populations. The description summarises the number of _E. ovata_ trees in each remnant class and their position in the landscape relative to the plantation. The average percentage of plantation trees with capsules is also shown for the different remnant classes (na = not assessed; see text for details).

<table>
<thead>
<tr>
<th>Remnant class</th>
<th>Description</th>
<th># of remnants</th>
<th># of trees sampled</th>
<th>% of plantation trees reproductive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-10 trees in open pasture 50-200 m from plantation edge</td>
<td>3</td>
<td>14</td>
<td>na</td>
</tr>
<tr>
<td>2</td>
<td>1-30 trees embedded in the plantation (fully surrounded)</td>
<td>9</td>
<td>32</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>c. 50 trees embedded in the plantation</td>
<td>5</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>&gt;100 trees embedded in the plantation</td>
<td>4</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>&gt;1000 trees adjoining plantation</td>
<td>4</td>
<td>32</td>
<td>30</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>25</strong></td>
<td><strong>142</strong></td>
<td><strong>29</strong></td>
</tr>
</tbody>
</table>

After drying, the seed was extracted from the capsules and weighed. Each individual _E. ovata_ seedlot was randomized and sown in separate Styrofoam vegetable boxes (450 x 295 x 250 mm; Appendix 4.1 Fig. A.4.1.5) onto wet potting medium that was nine parts composted pine bark, to one part perlite, with a low level of slow release fertilizer. Once the seed was sown the surface of the potting mix was covered with vermiculite to reduce the probability of seed jumps during watering before germination occurred. Seeds were germinated in the dark at 26°C, for 7 days before being moved out to a shade house for two weeks hardening off and finally moved outside for the duration of the experiment.

The seed lots collected from _E. ovata_ trees were assessed at 11, 16 and 35 weeks after sowing for the presence of seedlings with _E. ovata x globulus_ hybrid morphology. _F_1 hybrids in eucalypts typically show morphology that is intermediate between the parental taxa (Griffin et al. 1988). _F_1 hybrids between _E. globulus_ and _E. ovata_ are well-known (Barbour et al. 2008b; Lopez et al. 2000a) and have
distinctive seedling morphology allowing for rapid screening of large numbers of seedlings. The morphological markers used to identify putative F₁ hybrids here were: petiole absent at node six; leaves opposite at node six; leaves oblong; stem square to winged in cross section. The alternate traits in the pure *E. ovata* half-sibs seedlings were: petiole present at node two or three; leaves alternate before node six; leaves ovate to lanceolate; stem round to angular not winged (see Appendix 4.1 for pictures demonstrating these characters). A subset of putative hybrids were also checked using molecular markers and Bayesian assignment, all of which were confirmed to be hybrids (Fig. 4.2); the details of the assessment including materials and methods, data analysis and results can be found in Appendix 4.2. Pure *E. ovata* seedlings were thinned from the boxes to reduce competition and allow suppressed seedlings to grow. Every seedling was categorised as pure *E. ovata*, putative *E. ovata* x *globulus* F₁ hybrid, or native hybrid (if the morphology differed from the pure seedlings but was not characteristic of the exotic hybrid; Appendix 4.1, Fig. A.4.1.7). It should be noted that because the plantations in the study landscape are in their first rotation and only 8 – 12 years old, there has been no previous opportunity for hybridisation, meaning all *E. ovata* x *globulus* hybrids can be assumed to be F₁’s.

4.3.2 Assessing the frequency of hybrid establishment in the wild

Hybrid establishment was assessed by undertaking detailed surveys along the boundary between *E. globulus* plantations and native forests. Because of the inherent difficulties of surveying for rare events, and in order to obtain a conservative estimate of hybrid establishment in the wild (i.e. the worst case scenario), a survey was designed to target “high risk” sites. High risk sites were those where: 1) an *E. globulus* plantation was growing directly adjacent to native *E. ovata* forest; and 2) there was recruitment with seedling of similar sizes from both the native *E. ovata* and the plantation *E. globulus*, indicating recent reproductive activity at about the same time, with conditions suitable for establishment. Additionally, so that all hybrids identified could be assumed to be exotic, sites were not considered where native *E. globulus* or *E. cypellocarpa* (a close relative of *E. globulus* that has similar juvenile morphology) were present. Vehicle based surveys were undertaken in Gippsland, the Green Triangle and Tasmania covering a total of 216 km of plantation
boundary. Four high risk sites were located, one in Tasmania, two in Gippsland and one in the Green Triangle (Fig. 4.1a). Detailed surveys were undertaken at each of these sites covering a linear distance of 3.9 km and a total area of 11.7 ha. The native forest and plantation were typically separated by a road or fire break, although the plantation boundaries sometimes had occasional native E. ovata growing amongst the E. globulus along the plantation edge. The surveys were conducted parallel to the boundary and covered the area 15 m either side of the road (i.e. from the edge of the road 15 m towards the plantation, and from the other edge 15 m towards the native forest). The distance between the road edge and either boundary was always less than 15 m (usually about 10) meaning that all the open area most suitable for establishment along the boundary was assessed. Every seedling identified along each boundary was categorised as pure E. ovata, pure E. globulus or putative E. ovata x globulus hybrid.

![Figure 4.2](image)

**Figure 4.2.** Proportion of genotype membership for 204 individuals allocated to the predefined Eucalyptus ovata and E. globulus genetic groups, based on Bayesian cluster analysis in STRUCTURE. The figure shows the genotype affinities of putative F1 hybrids identified in open pollinated seed collected from native E. ovata trees growing adjacent to E. globulus plantations. The hybrids are clearly distinguished from the pure E. ovata, and have genotype affinities that are most consistent with F1 hybrids between E. ovata and E. globulus. The samples were genotyped at 10 microsatellite loci. Each individual is represented by a single vertical line that is partitioned based on its genotype affinities to the two pure species groups. Individuals are grouped based on their morphological classification as E. ovata, putative E. ovata x globulus hybrids or E. globulus. See Appendix 4.2 for full details regarding this analysis and a comprehensive discussion of the results.

### 4.3.3 Survival and fitness of E. ovata x globulus F1 hybrids in the wild

Whilst undertaking surveys for naturally establishing exotic hybrids, Barbour et al. (2008b) identified a site where a reproductively mature E. globulus plantation and
adjoining *E. ovata* native forest had been burnt in a wildfire approximately one and a half years before the survey in 2006. Eighty putative *E. ovata* x *globulus* hybrids were identified amongst naturally establishing *E. ovata* seedlings in an area 150 x 30 m within the native forest. In 2006 each of the hybrids were paired with the closest pure *E. ovata* of equal size, with members of a pair typically less than 30 cm apart. As a measure of initial fitness, leaf area loss was recorded from all 160 seedlings. The measurement of leaf area loss was a subjective visual estimation of the proportion of leaf area lost through herbivore damage, necrosis and disease compared to the assumed full undamaged leaf area. In 2011 the pairs were reassessed for survival and fitness. In the five years since the plots were setup, the site became overgrown with dense early successional vegetation. As a result only 34 of the original 80 pairs could be relocated. Survival and leaf area loss were recorded from these pairs.

### 4.3.4 Data analysis

In order to test for differences in the level of hybridisation between the five patch size classes we used a series of Generalized Linear Models (GLM). The GLM approach allows the incorporation of non-normal response distributions, which are often useful for analysing count data (Zuur et al. 2009). In particular our dataset was zero inflated with more than half of the trees not producing exotic hybrids. In order to account for the overdispersion associated with the extra zeros we used negative binomial distribution with a log link function (Zuur et al. 2009). When fitting categorical variables in a GLM (as we do here, see below) the model produces an analysis of deviance, which is analogous to analysis of variance (ANOVA) (Zuur et al. 2009). Model fit was assessed by inspection of standard residual, normality and leverage plots.

The dependent variable in all models testing differences between patch-size classes was the percent of hybrids per tree (Hybrids). We firstly tested overall differences between the patch size classes regardless of tree position (Hybrids ~ Patch). Because the small patches (class 1 and 2) had no central trees (i.e. all trees are on the edges of these patches), we wanted to test if differences between patches could be seen when considering only the trees closest to the edge in all patches. To do this we used a
subset of the data with only edge trees and ran the same model (Hybrids ~ Patch). To test for differences between the edge and central trees we again used a subset of the data which included only the patch size classes that had both edge and central trees (classes 3, 4 and 5). This model tested for differences between patch size classes, tree position and their interaction (Hybrids ~ Patch + Position + Patch*Position). The full model was then simplified by removing non-significant terms using analysis of deviance (drop 1 command; Zuur et al. 2009). To test the difference between the spatially equivalent class 1 and central trees in classes 3, 4 and 5 (equivalent in terms of distance to the plantation edge) we used a reduced dataset including all class 1 trees and central trees from classes 3, 4 and 5 (Hybrids ~ Patch). To test for differences in reproductive output of the plantations beside the embedded remnant classes (classes 2 – 5) we converted capsule scores to the proportion of capsule bearing trees (which was normally distributed) and used this as the dependent variable, testing the differences between classes with one way ANOVA. For the native hybrids, the sample size and level of zero inflation precluded any formal statistical test of variation between patch size classes or tree position (Fig. 4.3).

To test whether or not the level of observed hybrid establishment was equal to what would be expected given the rate of hybridisation that we identified in the open pollinated seed lots, we used chi-squared ($\chi^2$) test. The establishment surveys were conducted along native forest boundaries meaning that the appropriate rate of hybridisation to use in calculating the expected rate is that found in the edge trees.

For the paired hybrid and native *E. ovata* trial, paired t-tests were used to assess differences between hybrids and pure *E. ovata* seedlings in terms of leaf area loss in 2006. Because of mortality there were not enough complete pairs left in 2011 to use paired tests, so an un-paired t-test was used. To determine if there were differences in survival between the hybrids and pure *E. ovata* seedlings we used a $\chi^2$ test. Survival was also converted to relative fitness ($W$) using the formula $W = W_{abs(\text{hybrid})}/W_{abs(\text{pure})}$, where $W_{abs}$ is the absolute fitness, calculated as the proportion of either cross type (hybrid or pure) surviving (Conner and Hartl 2004). Relative fitness was calculated in Microsoft Excel, and all other analysis was conducted in R.
Chapter 4:  

version 2.14.1 (R Development Core Team 2011) with the packages MASS (Venables and Ripley 2002), and lmtest (Zeileis and Hothorn 2002).

4.4 Results

4.4.1 The effect of patch size on hybridisation rate in a fragmented E. ovata landscape

From the 142 open pollinated E. ovata seedlots a total of 24,322 seedlings were assessed. Of these 394 (1.62 %) were identified as putative E. ovata x globulus F₁ hybrids based on morphological markers (Appendix 4.1), and 42 (0.17 %) as putative native hybrids (Appendix 4.1, Fig. A.4.1.7). Exotic F₁ hybrids were found in 63 of 142 seedlots (44 % of trees). The maximum rate of hybridisation in an individual tree was 42.2 %. That tree was in a remnant class 2 that had just two conspecifics. This patch also showed the highest hybridisation rate of any patch in the study (30.4 %).

Overall, we found that the two small patch size classes had higher rates of exotic hybridisation than the three larger patch classes. Remnant class two (the smallest embedded remnant) had higher rates of hybridisation than the larger classes three (Z = -3.09, P = 0.002), four (Z = -2.86, P = 0.004), and five (Z = -3.91, P = < 0.001). However, there was no difference in hybridisation rate between the small paddock remnants 50-200m from the plantation edge (class one) and the small embedded class two remnants (Z = -0.73, P = 0.47). This pattern of reduced hybridisation in the larger patches was also evident when comparing the trees directly adjacent to the plantation edge in the different size classes (Fig 4.3), and was significant between class two and three (Z = -2.59, P = 0.0097), and class two and five (Z = -2.53, P = 0.012). Although the difference between class two and four was no longer significant when looking at only the edge trees (Z = -1.43, P = 0.15), there were lower rates of hybridisation in class four, and there was no significant difference between classes three, four and five (χ² = 1.7, df = 2, P = 0.42) indicating the pattern of reduced hybridisation probably does hold beyond a threshold somewhere between class two and three. There was also a strong effect of tree position with central trees (50 m inside remnant classes 3, 4 and 5) having significantly lower levels of hybridisation than edge trees (Z = 3.8, P = 0.0002; Fig 4.2). The central trees also had significantly
lower levels of hybridisation than the spatially analogous class 1 trees, which were also 50 m (or more) from the plantation edge but were in open paddocks ($Z = 3.7$, $P = 0.0002$; Fig 4.3). Figure 4.3 shows there were some similarities in the rate of native and exotic hybridisation; in particular the small remnant classes 1 and 2 also showed higher rates of native hybridisation than the larger patches. However, in contrast to the exotic hybrids, there was no consistent difference between edge and central trees in the rate of native hybridisation (Fig. 4.3). There were no significant differences in reproductive output of the plantations adjoining the embedded remnant classes (classes two – five; $F = 0.15$, df = 3, $P = 0.92$; Table 4.1). There were no plantations adjoining class one remnants so they were not tested.

4.4.2 Assessing the frequency of hybrid establishment in the wild

At the four “high risk” sites where detailed boundary surveys were undertaken 745 naturally established seedlings were assessed in a total area of 11.7 ha ($E. globulus = 26\%$; putative $E. ovata \times globulus \ F_1$ hybrids = 1\%; $E. ovata = 73\%$; Fig. 4.4). $F_1$ hybrids are likely to be from native $E. ovata$ trees because a unilateral crossing barrier prevents $E. globulus$ being the maternal parent (Gore et al. 1990), and the ratio of exotic hybrid seedlings to the native $E. ovata$ was 1.3 %. This is significantly lower than the number of hybrids found in seedlots from the edge trees in the crossing study (2.59 \%; $\chi^2 = 3.9$, df = 1, $P = 0.048$). Putative hybrids were only located in Tasmania (n = 2) and the Green Triangle (n = 5). In contrast to the hybrids, $E. globulus$ seedlings (wildlings) were more common, making up 25.6 \% of the observed seedlings.
Figure 4.3. a) The mean percentage of exotic hybridisation (*Eucalyptus ovata* x *globulus*) (a), and native hybridisation (b), identified in open pollinated seedlots from different size *E. ovata* remnant patches in and around *E. globulus* plantations. Remnant class one (left of the dotted line) are paddock trees/patches 50-200 m from the plantation edge, all other classes (right of the dotted line) are either embedded within or they directly adjoin the plantation matrix (see Table 4.1 and Figure 4.1). In (a) means with the same letter are not significantly different (*P* > 0.05; see results for details), and (b) shows the mean ± standard error. The figures are based on the assessment of hybridisation in 24,322 open pollinated progeny collected from 142 trees from 25 patches.
4.4.3 Survival and fitness of E. ovata x globulus F\textsubscript{1} hybrids in the wild

At the time of trial establishment in 2006 the putative exotic hybrids had significantly higher levels of leaf area loss, as a result of herbivory or necrosis, (22.5 % ± 1.3) than the pure E. ovata (16 % ± 1.1; \( t = 3.8, \text{ df} = 79, P = 0.0003 \)). In 2011, 28 of a possible 68 trees (from the 34 relocated pairs) were alive, with significantly fewer hybrids than pure E. ovata seedlings surviving (5 compared to 23; \( \chi^2 = 28.3, \text{ df} = 1, P < 0.0001 \)). When converted to relative fitness (\( W = 0.22 \)) this equates to a 78 % reduction in the fitness of the hybrids compared to the pure E. ovata seedlings with which they were competing. There were not enough complete pairs remaining in 2011 to do a paired analysis of leaf area loss (three pairs), but a straight comparison of the two groups showed no significant difference in leaf area loss (pure = 26.1 % ± 5; F\textsubscript{1} = 27.6 % ± 9.7; \( t = 0.52, \text{ df} = 4, P = 0.63 \)).

![Figure 4.4](image)

**Figure 4.4.** The number of Eucalyptus ovata, E. ovata x globulus hybrids, and E. globulus seedlings identified in surveys of high-risk plantation-native forest boundaries (see methods). The surveys were undertaken at four sites and covered total area of 11.7 ha.

4.5 Discussion

4.5.1 Hybridisation and the effect of patch size, tree position and isolation

The overall rate of exotic F\textsubscript{1} hybridisation identified in this study (1.62 \%) is low, but similar to the level identified in native E. ovata where it naturally co-occurs with E. globulus (≈ 1 \%; Lopez et al. 2000a). The rate is also consistent with eucalypts generally (Griffin et al. 1988; Potts et al. 2003). For example, Potts et al. (2003) found that 1.62 \% was also the average rate of natural hybridisation across 13 Tasmanian eucalypt species. Two other studies that have investigated exotic inter-
specific gene flow into native *E. ovata* populations found contrasting levels of hybridisation, with 7.2 % from *E. nitens* (Barbour *et al.* 2005b) and 0.3 % from *E. globulus* (Barbour *et al.* 2008b). However, the first study targeted only *E. ovata* trees that flowered synchronously with the neighbouring *E. nitens* plantation, making direct comparisons difficult. While the second study was undertaken in much younger plantations (5 to 6 years old compared to 10 or 11 here), meaning that an increasing reproductive output with age might explain the higher rate of hybridisation here.

The rate of exotic hybridisation varied considerably between individual trees (0 – 42.2 %) and between patches (0 – 30.4 %), which is consistent with that observed in naturally hybridising eucalypt species (Field *et al.* 2008; Potts and Wiltshire 1997). Despite this variation, it was still possible to detect higher levels of exotic hybridisation in small embedded native remnants, in native trees close to the plantation boundary, and in native paddock trees/patches away from the plantation edge. We also found higher levels of native hybridisation in small and isolated patches. Higher levels of hybridisation in small patches, is predicted in theory (Ellstrand and Elam 1993; Hadley and Betts 2012) and has been identified across a range of taxa (Cresswell and Osborne 2004; Ellstrand and Marshall 1985; Kunin 1997; Richards *et al.* 1999), including in other eucalypt species (Field *et al.* 2008; Potts *et al.* 2003; Potts and Wiltshire 1997).

There are two main theoretical explanations for increased hybridisation in small patches. Firstly, given an equal amount of pollen in the environment, the relative proportion of inter- to intra-specific pollen increases as patch size decreases, resulting in more inter-specific hybridisation (Ellstrand and Elam 1993). Secondly, in comparison to large patches, pollinators spend less time in small patches due to low resource availability, meaning they move more often between patches, which facilitates more inter-patch and inter-specific pollinations (Ellstrand and Elam 1993). In the plantation context, this second effect could be complicated by pollinator preference and flowering synchrony between the plantation and native species, possibly resulting in small patches becoming resource beacons leading to higher visitation rates and longer stays (Hadley and Betts 2012). These explanations are
both important drivers in source-sink dynamics, but their relative importance can be difficult to determine (Ellstrand and Elam 1993). However, there is evidence that the relative portion of intra- to inter-specific pollen drives source-sink dynamics in some naturally hybridising eucalypts (Field et al. 2008).

There were three patterns in hybridisation rate here that demonstrate the complexity of sources-sink relationships, and warrant particular explanation. Firstly, the more than fourfold reduction in exotic hybridisation rate over just 50 m in the large embedded remnant patches. This reduction is independent of patch size, indicating that the relative ratio of inter- to intra-specific pollen is not affecting hybridisation rate or pollen penetration into these larger patches. This rate of reduction in hybridisation is consistent with the exponential drop off in pollen dispersal shown in E. nitens (Barbour et al. 2005b) and E. globulus (Mimura et al. 2009). The effect of distance could be a consequence of low pollen carryover, where pollen from the plantation is intercepted (screened out) by trees near the edge of the patch. Low pollen carryover is consistent with both experimental (Price and Waser 1982; Thomson and Plowright 1980) and field based studies (Robertson 1992), and is supported by the absence of a reduction in native hybridisation between edge and central trees here – because the hybridising native species’ co-occur within the patch.

The second noteworthy pattern was the high rates of hybridisation identified in small embedded patches (remnant class 2; Fig. 4.3) compared to edge trees in larger embedded patches. In contrast to the larger patches above (pattern one), this result might be most simply explained by pollen swamping (Petit et al. 2004; Petit et al. 1997). The very low intra- to inter-specific pollen ratio in the small patches may result in an increase in hybridisation on a local scale, compared to the edges of larger patches where more intra-specific pollen is available. This local effect would be consistent with the generally high outcrossing rates maintained in eucalypts through self incompatibility (Griffin et al. 1987; Potts and Savva 1988; Pryor 1976). Similar patterns of exotic hybridisation have also been attributed to pollen swamping in small, compared to intermediate and large Populus remnants (Meirmans et al. 2010). The overall spatial proximity of trees in small patches to the plantation edge may also influence their relative exposure to inter- versus intra-specific pollen. Trees in
small patches can receive exotic pollen from all directions, whereas trees on the edges of larger patches only receive exotic pollen from the plantation in one direction, receiving intra-specific pollen from within the patch from the other directions. Therefore with nearest neighbour pollination, inter-specific hybridisation will be more likely in small patches (Turner et al. 1982).

The third point is the almost 10 fold increase in hybridisation rate in isolated paddock trees when compared to spatially equivalent trees 50 m inside the larger embedded remnant classes (central trees; Fig. 4.3). The potential for increased dispersal and outcrossing rates in fragmented landscapes has long been recognised (Heinrich and Raven 1972), and has been noted in several eucalypts (Barbour et al. 2005b; Byrne et al. 2008; Mimura et al. 2009). It requires generalist pollinators in a landscape where the between patch matrix is not a strong barrier to the pollinator movement (Hadley and Betts 2012; Heinrich and Raven 1972). The results here suggest that pollinators have no trouble crossing the treeless paddock matrix. The relatively high hybridisation rate might therefore be a consequence of improved pollen carryover due to an absence of trees to intercept the pollen in the paddock matrix (Hadley and Betts 2012). A second contributing factor could be that these isolated patches may be acting as beacons or refugia in the largely unusable paddock matrix, with pollinators being attracted to and spending longer in the isolated patches because of a lack of adjacent habitat. There is theoretical and empirical evidence for this beacon effect (Hadley and Betts 2012), and studies investigating the use of isolated trees in Australian agricultural landscapes, do show that these trees are focal points for birds and insects (Law et al. 2000; Lumsden and Bennett 2005).

The patterns identified here in exotic hybridisation rate have two obvious management consequences. Firstly, there have been two regularly suggested methods for mitigating gene flow from eucalypt plantations: buffer zones, where a space is left between the plantation and native forest of conservation significance to reduce pollen flow (Barbour et al. 2008a; Barbour et al. 2005b); and edge plantings of non-flowering or alternate plantation species, or low-risk native forest around the border of a plantation to act as a filter, minimising pollen escape (Potts et al. 2003; Vanden Broeck et al. 2005). The results here suggest that edge plantings or leaving
low-risk native forest (e.g. non-crosscompatible native species) around plantations will be more successful than buffer zones without trees. Secondly, revegetation strategies should avoid collecting seed close to plantations boundaries and from small embedded patches.

4.5.2 Hybrid fitness

The ability of F₁ hybrid seed to geminate, establish and reach reproductive maturity within the competitive native forest environment is vital for successful gene flow (Potts et al. 2003), and could be a significant barrier to introgression (Heinze 2011). The patch size study has measured the rate of post-zygotic, pre-dispersal hybridisation in this system. That is, the rate of F₁ hybridisation after overcoming all pre-zygotic barriers (e.g. flowering time or pollinator preferences), and any post-zygotic barriers in seed development, but prior to barriers that could reduce the rate of hybridisation post dispersal in the wild (Potts et al. 2003). The results of the establishment surveys and paired hybrid trial indicate that post-dispersal barriers do exist. Despite targeting high-risk sites, F₁ hybrid establishment was low in comparison to both pure species, and even if hybrids do establish, the paired trial showed clear evidence for selection against the F₁’s. Selection against the F₁’s may be further evidenced by hybrid establishment being lower than expected given the rate of pre-dispersal hybridisation found in edge trees in the patch size study, although caution is needed here because the comparison is confounded by site. Lower establishment has been reported in exotic E. ovata x E. nitens F₁’s compared to E. ovata seedlings at 16 months (Barbour et al. 2006a), while other studies have shown selection acting later in development. For example, in manipulated E. ovata x globulus F₁ hybrids Lopez et al. (2000a) found equal survival at age 2, but by age 4 hybrid mortality was significantly higher than the pure species, and at age 10 hybrid mortality was even higher than selfs of the pure species.

In the situation studied here E. globulus is locally exotic, but the two species do naturally co-occur, for example in southern Tasmania where they are often codominant, and have overlapping flowering times (Lopez et al. 2000a). Although occasional natural hybrids occur and low rates of hybridisation have been observed in open pollinated seed, hybrid swarms are not known and there is little evidence of
introgression (Lopez et al. 2000a). This suggests that barriers to introgression maintain the species boundaries where they have evolved in sympatry. The results here suggest that these barriers also occur where the species have evolved in allopatry, making widespread introgression unlikely.

4.5.3 The likelihood of introgression

We found that although hybridisation is occurring, there appear to be significant postzygotic barriers to establishment and survival, which result in a low likelihood of introgression. However, the fact that a few hybrids are surviving mean that the risk cannot be totally discounted. This being said, theoretical modelling studies suggest that in the presence of even moderate negative selection the probability of introgression falls to nearly zero (Meirmans et al. 2009). Also, even if F$_1$ hybrids can reach reproductive maturity, previous studies suggest further barriers might exist to prevent introgression. Firstly, while both parental species overlap in flowing time, manipulated E. ovata x globulus F$_1$’s show asynchronous flowering to both parents (Lopez et al. 2000a), which would provide a strong barrier to backcrossing. Secondly, even if this flowering barrier is overcome the fitness of the advanced generation hybrids may present a final barrier to introgression. Costa e Silva et al. (2012) found that F$_2$ and backcross generations (between E. globulus and E. nitens) displayed reduced fitness in terms of growth and survival due to a breakdown of epistasis (outbreeding depression).

Our study focused on the temperate E. globulus system in Australia and the results are likely to be transferable to related temperate taxa (e.g. E. nitens; Barbour et al. 2006a). Some results are also likely to be transferable to the emerging tropical and subtropical plantation industry in northern Australia, (Lee et al. 2009; Lee 2007), although the use of Corymbia species in this system will necessitate some additional research. In particular the effects of patch size are likely to be transferable, although possibly scale dependent because of potentially greater pollinator mobility in Corymbia (Bacles et al. 2009; Southerton et al. 2004). Hybrid establishment and fitness are likely to be more taxon specific and are likely to require additional research in the tropical/subtropical system – particularly because: there is evidence that Corymbia species may hybridise more readily than Eucalyptus (Dickinson et al.
2012a; Dickinson et al. 2012b; Griffin et al. 1988); inter-specific hybrids involving the main plantation taxa are known to reach reproductive maturity (Barbour et al. 2008a) and; in some situations elite Corymbia F₁ hybrids are deployed as plantation stock (Lee et al. 2010), circumventing the initial F₁ fitness barrier.

More broadly, our results are in contrast to a number of studies in other (non-eucalypt) systems that have found that introgression as a result of exotic gene flow is likely (Meirmans et al. 2009; Poppy 2004; Wilkinson et al. 2003a; Wilkinson et al. 2003b). These systems usually have few fitness barriers to hybridisation, and even when they do (e.g. see Laikre et al. 2010), the barriers are incomplete and hybrids commonly contribute to reproduction (Laikre et al. 2010). Therefore, strong negative selection that reduces the likelihood of hybrids reaching reproductive maturity is probably a key point of difference in this system.

4.5.4 Wildlings

In contrast to the low levels of exotic hybrid establishment, we found that pure-bred E. globulus seedlings (wildlings) were more common. Given that our survey conditions required the presence of E. globulus wildlings, we cannot comment on the scale of establishment (but see Larcombe et al. 2013). However, seed mediated gene flow is certainly a potential risk to native forests, and we have shown that E. globulus wildlings are establishing. Eucalyptus globulus seed is dispersed by gravity (Cremer 1977), sometimes inside the capsule (Calviño-Cancela and Rubido-Bará 2013), meaning it must have a lower capacity for dispersal into the native forest than pollen as shown by Barbour et al. (2003) in E. nitens. This is also supported by studies in native eucalypts that show pollen dispersal is more important than seed dispersal (Bloomfield et al. 2011; Byrne et al. 2008). Nevertheless, if wildlings along the boundary are able to reach reproductive maturity they could become a secondary source of pollen and seed mediated gene flow. Indeed recent work in Spain shows that very high densities of E. globulus wildlings can establish directly adjacent to plantations (Calviño-Cancela and Rubido-Bará 2013). Plantation companies in Australia typically operate under forest certification schemes that require the management of offsite effects such as wildling spread (Dare et al. 2011). Calviño-Cancela and Rubido-Bará (2013) suggest that clearing wildlings within 15
Chapter 4:  

E. ovata case study

m of plantations will significantly reduce the risk, and similar recommendations have been made for Australia (Larcombe et al. 2013). Therefore, wildlings will mainly be a concern if plantations and/or the associated management are removed. For example in 2010 around 20,000 ha of E. globulus plantations were written off by managed investment schemes in Australia, and although some are likely to be converted back to agriculture the fate of others is unknown (M. Gavran pers. com). Eucalyptus globulus is long lived and shows increasing reproductive output with age, meaning that wildlings have the potential to become significant long term sources of exotic pollen and seed if left unmanaged.

4.6 Conclusion

In conclusion, hybridisation is occurring at rates that are similar to those observed where E. ovata and E. globulus naturally co-occur, but are higher in small patches, edge trees and paddock trees. Barriers exist to F₁ hybrid survival, which are consistent with those that maintain the species boundaries where they naturally co-occur and, as in nature, these barriers are likely to limit introgression between plantation E. globulus and native E. ovata. Finally, while pollen dispersal is likely to be a more important genetic risk over long distances, pure E. globulus wildlings seem to pose a greater potential risk to the integrity of native forests close to the plantation boundary.
Appendix 4.1: Summary of the morphological markers used to identify hybrids between *Eucalyptus globulus* and *E. ovata*

First generation (F$_1$) hybrids in eucalypts typically show morphology that is intermediate between the parental taxa (Griffin *et al*. 1988). F$_1$ hybrids between *E. globulus* and *E. ovata* are well known (Barbour *et al*. 2008b; Lopez *et al*. 2000a) and have distinctive seedling morphology allowing for rapid screening of large numbers of seedlings as in this study. The morphological markers used to identify putative hybrids here were: petiole absent at node six; leaves opposite at node six; leaves oblong; stem square to winged in cross section. The alternate traits in the pure *E. ovata* seedlings are: petiole present at node two or three; leaves alternate before node six; leaves ovate to lanceolate; steam round to angular not winged. Below is a series of photographs that summarise the morphological markers (Fig. A.4.1.1 - 4) and some examples from the nursery grown open pollinated progeny from the patch size study (Fig. A.4.1.5 – .7). Also see the main paper and Appendix 4.2 for molecular validation of the morphological markers.

*Figure A.4.1.1.* Apical-tip segments of *Eucalyptus globulus*, *E. ovata* and *E. ovata* x *globulus* F$_1$ hybrid showing the intermediate nature of the hybrid morphology with its
distinctive opposite and sessile leaves compared to the alternate and petiolate leaves of the pure *E. ovata*.

**Figure A.4.1.2.** Close up of node segments of *Eucalyptus globulus* (after node 10), *E. ovata* (c. node 7) and *E. ovata x globulus* F$_1$ hybrid (c. node 7) showing the showing distinctive opposite and sessile leaves compared to the alternate and petiolate leaves of the pure *E. ovata*. 
Figure A.4.1.3. Leaves of *Eucalyptus globulus* (after node 10), *E. ovata* (c. node 7 to 9) and *E. ovata x globulus* F₁ hybrid (c. node 7 to 9) showing the intermediate morphology of the hybrids linear to ovate leaf shape and sessile leaves compared to the petiolate, and ovate leaves of the pure *E. ovata*.
Figure A.4.1.4. Stems (top) and stem cross sections (bottom) of *Eucalyptus globulus*, *E. ovata* and *E. ovata x globulus* $F_1$ hybrid showing the intermediate morphology of the hybrids square to winged stem compared to the round or slightly angular *E. ovata* stem (ruler is graduated in mm).
Figure A.4.1.5. Two *E. ovata* x *globulus* F₁ hybrids (red arrows) growing among pure *E. ovata* at 16 weeks just prior to thinning of pure seedlings from the open pollinated progeny arrays. Note the square, winged stems and sessile opposite leaves of the hybrids compared to the pure *E. ovata* with alternate and petiolate leaves (blue arrow). Even at this relatively early stage hybrids were obvious.
Figure A.4.1.6. An *E. ovata x globulus* $F_1$ hybrid, and pure *E. ovata* seedlings at the final assessment at 35 weeks.
Figure A.4.1.7. A native hybrid at 35 weeks, the morphology of this plant differs from both the *E. ovata* and the *E. ovata x globulus* F₁ hybrid in that the leaves and stems are round, and it has a spreading branching pattern, whereas both the *E. ovata x globulus* F₁ hybrid and the pure *E. ovata* are usually upright with minimal lateral branching at this age.
Appendix 4.2: Details regarding the molecular validation of hybrids between *Eucalyptus globulus* and *E. ovata*

In the main chapter morphological markers are used to identify F₁ hybrids as well as molecular markers to validate a subset of hybrids. Here the materials and methods for the molecular analysis, and a more detailed description of the results of the hybrid validation are given.

*Materials and methods*

A subset of the putative hybrids was tested using molecular markers to validate the morphological classifications. Leaf samples were collected from 15 randomly selected putative hybrids found in the patch size study and two samples located in the establishment study. Additionally 100 *E. ovata* samples and 87 *E. globulus* samples were collected or obtained from various pre-existing DNA stocks (see Electronic Supplementary 3 for sample information). Nine samples were included twice to quantify repeatability of the technique. Total genomic DNA was extracted from frozen leaf samples using the CTAB protocol of Doyle and Doyle (1990) with the adjustments used by McKinnon et al. (2004b). The quality and quantity of DNA was assessed using gel electrophoresis and comparison with Lambda *HindIII* molecular weight standard. Ten microsatellite loci were used, four (EMCRC2; EMCRC7; EMCRC8; and EMCRC11) designed by Steane et al. (2001), two (EMBRA11 and EMBRA16) designed by Brondani et al. (1998), and four (EMBRA23; EMBRA30; EMBRA38; and EMBRA63) designed by Brondani et al. (2002). The primer sequences for the loci can be found in their respective references. These loci have been mapped and there is no evidence of linkage between them (J. S. Freeman, pers. comm.). In order to allow simultaneous analysis of different loci, the forward primers were labelled at their 5’ end with the fluorescent dyes NED, 6-FAM, PET, or Hex (Perkin Elmer Applied Biosystems, Foster City, CA, USA). The ten loci were multiplexed in three mixes: mix 1 included EMCRC2, EMBRA63 and EMBRA 11; mix 2 included EMBRA16, EMCRC11, EMBRA23 and EMCRC7; and mix 3 included EMBRA30, EMBRA38 and EMCRC8. PCRs were performed using a QIAGEN Multiplex PCR kit (Hilden, Germany) according to the manufacturer specifications for 5 µl reactions using approximately 5 ng of genomic DNA per
Chapter 4:  

E. ovata case study

reaction. Thermo-cycler conditions followed Bloomfield et al. (2011) except the annealing temperatures were 59°C for mix 1, and 58°C for mix 2 and mix 3. The PCR products were diluted 1 in 10 and then 1 µl of that dilution was dried at 50°C. Fragment separation was undertaken on an AB3730 DNA analyser (Perkin Elmer Applied Biosystems) by the Australian Genome Research Facility (AGRF), Adelaide, South Australia.

Allele scoring procedure followed Bloomfield et al. (2011). The assigned genotypes of the nine repeated samples (which were scored blindly) were compared at each locus to obtain a measure of repeatability (number of alleles scoring errors error/number of alleles). The basic dataset statistics, number of alleles, number of private alleles, and the observed heterozygosity, were obtained using GenAlEx (Peakall and Smouse 2012). We then used Bayesian clustering (implemented in STRUCTURE) to investigate genotype membership of each putative hybrid to the parental clusters. Firstly, in order to determine the power of the molecular dataset to distinguish between the two pure species, we ran only the pure species samples using a full admixture model in STRUCTURE. The admixture model uses no prior sample/species information to allocate the individual samples to one of $K$ genetic clusters. We tested $K = 1$ to 5 and using the method of Evanno et al. (1992) it was clear that two clusters was by far the most appropriate grouping ($\Delta K = 803.6$) with next best being three clusters ($\Delta K = 1.5$). At $K = 2$ the genetic clusters matched the species groups perfectly. The dataset including the putative hybrids was then used in an analysis where the species information for E. ovata and E. globulus was used (USEPOPINFO) to define the two genetic clusters, and the hybrid samples were then allocated membership to one, or a proportion of both clusters using the admixture model.

Results

We identified 266 alleles from the 204 individuals (see Electronic Supplementary 4.1 for sample details) genotyped at 10 microsatellite loci (26.6/locus with a range of 23-29 alleles per locus; Table A.4.2.1). There was a total of 1.4 % missing data as a
result of either non-amplification or scoring ambiguity. We calculated repeatability to be 95 % based on the blind scoring of 9 repeated samples.

Table A.4.2.1. Genetic diversity parameters for *Eucalyptus ovata*, *E. globulus* and putative *E. ovata x globulus* F1 hybrids genotyped at 10 microsatellite loci. Parameters: \( n = \) number of individuals genotyped; \( N_a = \) average number of alleles per locus; \( P_a = \) total number of private alleles; \( H_o = \) observed heterozygosity.

<table>
<thead>
<tr>
<th>Group</th>
<th>( n )</th>
<th>( N_a )</th>
<th>( P_a )</th>
<th>( H_o )</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. ovata</em></td>
<td>100</td>
<td>20.9</td>
<td>53</td>
<td>0.68</td>
</tr>
<tr>
<td><em>E. globulus</em></td>
<td>87</td>
<td>19.3</td>
<td>37</td>
<td>0.78</td>
</tr>
<tr>
<td><em>E. ovata x globulus</em>  hybrids</td>
<td>17</td>
<td>12.1</td>
<td>2</td>
<td>0.90</td>
</tr>
<tr>
<td>Total</td>
<td>204</td>
<td>17.4</td>
<td>92</td>
<td>0.79</td>
</tr>
</tbody>
</table>

The hybrid group had fewer private alleles (2 compared to 53 and 37 respectively) and higher observed heterozygosity (0.90 compared to 0.68 and 0.78 respectively; Table A.4.2.1) than the pure species, which is consistent with theoretical expectations for hybrid populations (Ellstrand 1992b). A Principal Coordinate’s Analysis also showed that when comparing PCo 1 and PCo 2 the hybrids formed a group collinear with, and between the two main species clusters (Fig A.4.2.1).

In the admixture analysis, the genotype allocations for all putative hybrids were partitioned more or less equally between the *E. ovata* and *E. globulus* genetic clusters (Figure 4.2 main chapter). On average the hybrids showed 50.5 % genotypic affiliation to *E. ovata*, and 49.5 % to *E. globulus*, with a range of 39.6 – 60.4 %. These results provide strong support for our morphological classification of *E. ovata x globulus* hybrids.
Figure A.4.2.1. Axis 1 and 2 of a Principal Coordinate Analysis of individual multilocus genotypes from 100 *Eucalyptus ovata*, 87 *E. globulus* and 17 putative *E. ovata x globulus* F<sub>1</sub> hybrids.
Chapter 5: Assessing the invasive potential of *Eucalyptus globulus* in Australia: quantification of wildling establishment from plantations

This chapter is published in *Biological Invasions*

5.1 Abstract

*Eucalyptus globulus* is one of the most widely planted temperate hardwood species in the world, and in Australia there are 538,000 hectares growing in plantations. Although it has been reported as invasive, quantification of *E. globulus* invasion is rare. We conducted surveys at two geographic scales to assess the level of, and factors influencing, wildling establishment from industrial *E. globulus* plantations in Australian. We surveyed 290 km of plantation boundary, both within (22%) and outside (78%) the species native range. In areas of relatively high establishment, a density triggered paired plot approach (plots with and without wildlings) was used to assess fine-scale factors influencing establishment. We recorded 4,939 wildlings (17/km), 98 % of which occurred within 10 m of the plantation edge (maximum 175 m). Establishment varied between regions, ranging from 1.2 to 39.6 wildlings/km. Generalized linear models showed that the probability of a wildling being present increased with plantation age, that wildling abundance was higher along burnt transects, as well as sites that received regular, relatively high rainfall and had lower mean annual temperatures. The only fine-scale/local factor influencing wildling presence was the reproductive output of the plantation. The current level of *E.
**globulus** establishment in Australia is low in comparison to other invasive forestry trees. However, given the relatively young age of the Australian estate, local and regional variation in establishment, and potential future changes in plantation management, monitoring is warranted. Implications for assessing the general invasiveness of *Eucalyptus* and possibilities for *E. globulus* wildling control are discussed.

### 5.2 Introduction

There are around 264 million hectares of cultivated forests worldwide and approximately 25% of these are planted with locally exotic species (FAO 2010). The social and environmental benefits of forestry are immense, but there are environmental consequences that need to be managed, with many forestry species now considered invasive in at least part of their introduced range (Dodet and Collet 2012; Essl *et al.* 2010). A large body of research has developed around two invasive forestry tree genera, *Acacia* and *Pinus* (Ledgard 2001; Nuñez and Medley 2011; Richardson 1998; Richardson and Higgins 2000; Richardson and Rejmánek 2011; Richardson *et al.* 1994; Wilson *et al.* 2011). These two genera are being used as model systems for understanding invasion dynamics in forestry trees (Dodet and Collet 2012; Essl *et al.* 2010; Richardson *et al.* 1994; Wilson *et al.* 2011). These systems suggest that some forestry species may be particularly effective invaders, partly because of the characteristics that make them attractive for forestry (Essl *et al.* 2010). Selection for fast growth often infers associated traits such as low shade tolerance, early maturity and high propagule production, which contribute to invasive capacity (Dodet and Collet 2012; Essl *et al.* 2010). Other issues include, introduction effort (propagule pressure) which is often large in the forestry context (Pysek *et al.* 2009), biological interactions with the recipient community (e.g. predator release and empty niche theories; Theoharides and Dukes 2007), and the existence of a lag phase or required residence time between introduction and invasion (Pysek and Jarosfk 2005). Richardson and Rejmánek (2011) identify a future need for more taxon specific studies in groups other than *Pinus* and *Acacia*, to test the generality of invasion syndromes, including in groups such as *Eucalyptus*. 
Native mainly to Australia, the genus *Eucalyptus* is widely used in forestry around the world, and has also been extensively introduced for horticulture, agro-forestry and amenity purposes (Booth *et al.* 1987; Doughty 2000). Over 200 species have been tested for forestry outside Australia (Kunin 1997; Lexer *et al.* 2010), with many more introduced for ornamental and horticultural purposes. For example 374 species have been introduced to California alone (Ritter and Yost 2009). The threat of invasion from such introductions has recently been assessed in the United States, with 14 of the 38 eucalypt species considered, found to pose a high risk of invasion (Lexer *et al.* 2010). A number of *Eucalyptus* species have been reported as naturalising or invasive in various places around the world (Kirkpatrick 1977; Rejmánek and Richardson 2011; Richardson and Rejmánek 2011; Ritter and Yost 2009), however, given the intensive cultivation of the genus across a wide range of environments (from the tropics to the temperate zone), for almost 200 years (Doughty 2000), eucalypts are intriguingly rare on invasive species lists. Conversely, studies in *Pinus* and *Acacia* have found invasiveness is linked to species with repeated introductions, high propagule pressure (planting intensity) and long residence time – characteristics of many *Eucalyptus* species (Ahuja 2011; Ghosh *et al.* 2012; Procheş *et al.* 2012; Wilkinson *et al.* 2003a). Understanding factors that may be reducing invasiveness in *Eucalyptus* should help in predicting the invasive potential of the genus, and may also provide insight into the general roles of propagule pressure, residence time and other invasion processes (Richardson and Rejmánek 2011).

*Eucalyptus globulus* is one of the most widely planted and economically important eucalypt species in temperate regions of the world (Potts *et al.* 2004). It was first exported from Australia to Europe and the Americas as an ornamental species in the early 1800s (Doughty 2000), and was quickly recognised for its wood quality and fibre characteristics. The species now forms the basis of hardwood forestry industries in Australia, Chile, Ecuador, Ethiopia, Portugal, Spain and Uruguay, and is important in many other countries including China (Potts *et al.* 2004). In 2004 there were estimated to be over 2.3 million hectares of *E. globulus* planted worldwide (Potts *et al.* 2004), and that area has almost certainly increased. In Australia, *E. globulus* plantations are grown largely outside the species native range in southern
temperate parts of the continent (Fig. 5.1; Gavran and Parsons 2011). In the last 10 years the Australian *E. globulus* estate has increased by more than 150%, now standing at 538,000 hectares, although growth has slowed since 2009 (Gavran and Parsons 2011).

Despite the history and extent of *E. globulus* cultivation around the world (Doughty 2000; Potts *et al.* 2004), and its listing as invasive in five of the fifteen geographic regions considered by Richardson & Rejmánek (2011) in their global review of invasive trees and shrubs, actual quantification of its spread is surprisingly rare. Kirkpatrick (1977) described eucalypts (including *E. globulus*) as weak invaders in southern California noting that disjunct habitats and limited seed dispersal were barriers to invasion. A more recent study has shown that *E. globulus* has since become widely naturalised in many parts of California, but that its broad distribution and prevalence are predominately a result of intentional plantings and localised recruitment (Ritter and Yost 2009). It was also assessed to pose the highest risk of invasion in the USA of 38 *Eucalyptus* species considered in a recent risk assessment (Lexer *et al.* 2010). In Europe there are reports of the species becoming naturalised in France, Spain and Italy but it is not considered strongly invasive in those countries (Rejmánek and Richardson 2011). Perhaps the most well-documented example of establishment and spread comes from Portugal where there are large areas of non-managed *E. globulus* where the species is becoming naturalised (Silva and Marchante 2012; Silva *et al.* 2011). Silva and Marchante (2012) report average seedling densities of 0.88 plants/m² mainly in abandoned ex-forestry land that was burnt in Portugal’s 2005 wildfires. In Australia *E. globulus* has been recorded as a weed outside its natural range in Tasmania, Victoria, South Australia and New South Wales (Lazarides *et al.* 1997). These examples are typically qualitative, noting the potential of *E. globulus* to establish and spread, or they quantify its establishment on a local scale. We have found no study that quantifies the extent of current spread from plantations at or above a regional level or any study that investigates factors influencing the likelihood of future escape of *E. globulus* from cultivation.

The effect of *E. globulus* escape in countries other than Australia could be many and varied. For example, it has been reported that the secondary metabolites in *E.*
*Eucalyptus globulus* leaf litter inhibit microbial activity in the soil, reducing nitrogen availability and subsequently altering ecosystem function in Spain (Castro-Diez *et al*. 2010). *Eucalyptus globulus* is also highly adapted to fire and thought to increase the likelihood of fire in forests where it occurs in Portugal (Randi 2008; Silva *et al*. 2011). Conversely, there are reports that *E. globulus* can have positive ecosystem effects outside its natural range. In California, the red-shouldered hawk, which has experienced significant population decline due to urban development, preferentially nest in *E. globulus* with higher clutch success than in native *Populus* species (Rottenborn 2000).

The possible consequences of escape in parts of Australia where *E. globulus* is planted outside its natural range (Fig 5.1) are similar to those in other countries, although some negative ecosystem effects could be reduced through similarities with the native eucalypt flora. For example, most Australian eucalypt forests are already highly adapted to fire, meaning that the presence of *E. globulus* should have a less dramatic influence on fire frequency than it would in a Portuguese oak forest. There is however an additional risk in Australia that is absent in other countries – the potential for exotic gene flow and subsequent genetic contamination of indigenous eucalypt species gene pools (see Barbour *et al*. 2008b; Potts *et al*. 2003). *Eucalyptus* species often have weak reproductive barriers and the potential genetic risk associated with gene flow from plantations has been the focus of a large amount of research and management attention (Barbour *et al*. 2008a; Barbour *et al*. 2008b; Byrne *et al*. 2011; Laikre *et al*. 2010; Potts *et al*. 2003), however, the genetic risk posed by *E. globulus* seed-mediated invasion/escape has not yet been considered.

From an international perspective, the Australian *E. globulus* estate is young, with the majority of plantations in the first or second rotation (Gavran and Parsons 2011; Potts *et al*. 2004). This means that any escape or spread is likely to be in the early seedling establishment and survival stage (i.e. C1 in Blackburn *et al*. 2011). Understanding the factors that facilitate establishment may allow the development of management strategies to prevent the transition to later, more problematic, stages of maturity and recruitment in Australia, and possibly help identify factors that could be useful in containing *E. globulus* in countries like Portugal. Here we use a two scaled
survey technique to, firstly, assess the level of wildling establishment (i.e. plants which have naturally spread from cultivated *E. globulus*) adjacent to plantations across the main Australian *E. globulus* growing regions (both within, but predominantly, outside the species native range; Figure 5.1; Table 5.1), and secondly, investigate the local and broad-scale factors that influence *E. globulus* wildling establishment. A better understanding of establishment and the invasive potential of *E. globulus* might also shed light on, and help guide future research into, the general under-representation of *Eucalyptus* on international invasive species lists.

5.3 Methods

5.3.1 Survey area and method

*Eucalyptus globulus* is endemic to Tasmania and southern Victoria, and industrial plantations of the species are grown largely outside its native range, with the exception of Gippsland in Victoria and southern and eastern Tasmania (Figure 5.1; Table 5.1). We undertook a car based survey to assess wildling establishment from plantations across all main growing regions including Tasmania, Gippsland, the Green Triangle, and south west Western Australia (Figure 5.1; Table 5.1). Plantations assessed were mainly outside the native range of *E. globulus* [78%] but plantations were also assessed within the native range in Gippsland [22%] (Figure 5.1; Table 5.1). We targeted plantations of reproductive age (Barbour *et al.* 2008b), no younger than 6 years old or second rotation (i.e. second planting or coppice plantation; this information was obtained from the plantation companies), across a range of topographic, edaphic, geographic and geological situations as well as both burnt and unburnt plantations. The burnt plantations were mainly a result of the catastrophic 2009 Victorian Bushfires (Teague *et al.* 2010) and were easily identifiable.

Our survey was conducted at two geographic scales, firstly aiming to identify broad-scale factors (landscape level and above) associated with wildling establishment. Secondly we used a fine-scale survey that was triggered when the wildling density reached a certain threshold (see below) to determine local and microsite factors that might be important for wildling establishment.
Chapter 5: Wildlings

Figure 5.1. The distribution of *Eucalyptus globulus* plantations in Australia (green polygons (dark grey in print)), and the native (endemic) range of the species (yellow shading (light grey in print)). Regions surveyed for wildling establishment are circled, and the position of transects within the regions are shown as dots. The Grampians was the only region where surveys were conducted within the native range of *E. globulus*. The Green Triangle planting zone (GT) is indicated by the dashed line surrounding Grampians, Penola and Portland. Albany and Manjimup are in the Western Australian planting zone (WA), while Gippsland and Tasmania are planting zones (GIP and TAS respectively).

5.3.2 Broad-scale survey

Transects were established along fire breaks and roads directly adjacent to plantation boundaries in order to visually detect wildlings. We surveyed 269 transects (59 within, and 210 outside the native range of *E. globulus*; Table 5.1) along 290 km of plantation edge mainly by car at walking pace, and occasionally on foot if vehicle access was not possible. The number of wildlings along each transect was counted and the position of each 10 m segment of transect with at least one wildling present was recorded with a GPS, as was the start and end of each transect. We always had two observers, one primary observer focusing on the plantation edge and a second
observer scanning for rare wildlings further afield. The distance of any wildling found further than 10 m from the plantation edge, which could not have originated from another source (i.e. a second plantation or paddock tree), was recorded in order to obtain a measure of maximum dispersal distance. Assuming an average minimum scanning distance of 20 m (10 m inside the plantation and 10 m outside), we converted the number of wildlings/km to the number per ha (wildlings further than 10 m from the plantation edge were not included), in order to compare our results to those of other studies using a common scale (e.g. Ledgard 2001; Richardson 1998; Richardson and Brown 1986).

5.3.3 Fine-scale survey

In order to detect variation in fine-scale factors between areas with and areas without wildlings, we used a paired plot approach. When the density of wildlings along a transect (in the broad-scale survey) reached 5 wildlings in a 10 x 10 m square straddling the plantation boundary, a fine-scale survey was initiated involving a paired plot sampling regime. The first (positive) plot was established using the first of the five wildlings that triggered the sampling as the first edge of a 10 x 10 m plot (including the 5 + n wildlings). The plot straddled the plantation edge, splitting it into two equal 5 x 10 m rectangles, one inside, and one outside the plantation running parallel to the plantation edge. The second (control) plot was established under conditions that were as similar as possible to the positive plot within the same plantation, avoiding wherever possible differences in fire history or any other obvious landscape level factors or other differences (however, there were nine pairs that could not be matched perfectly and these were removed from the paired analysis, but they were used in other analyses assessing only positive plots). The control plot was established (in either direction along the transect) in the first available wildling free area at least 100 m from the positive plot and 50 m from any wildlings. Two positive plots were always at least 200 m apart and no more than three sets of pairs were established in a single plantation. In some instances the outside plot dimensions were < 5 m x 10 m (e.g. 4.5 m x 10 m) because of the proximity of roads to the plantation edge, in this situation wildling counts and ground cover scores were corrected to that of the full plot.
The following data were collected from each fine-scale plot (positive and control): presence or absence of fire; soil cover (% bare ground, % litter, % shrubs, % herbs); aspect of the plantation edge (N, S, E, W); topographical aspect (N, S, E, W); slope (degrees); and capsule abundance in the plantation measured on a log categorical scale (0 = 0 capsules, 1 = 1 – 10 capsules, 2 = 11 – 100 capsules, 3 = 101 – 1000 capsules, and 4 = > 1000 capsules). Capsule assessment was undertaken on up to 10 trees per plot (the closest 5 trees to a maximum of 25 m either side of the plot centre), using binoculars. Each wildling in a positive plot was assigned to one of four size classes to identify variation in recruitment age (0 - 0.5 m, 0.5 - 1.3 m, 1.3 - 3.0 m, and > 3.0 m). In total seventy-one pairs of plots were surveyed.

5.3.4 Datasets

GIS layers were generated, from the field recorded GPS coordinates, containing digitised transect routes and all positive wildling records. These layers were intersected with a 2011 *E. globulus* plantation layer (provided by the Australian Bureau of Agricultural, Resource Economics and Sciences) so that plantation age and plantation edge aspect could be extracted for each transect. The centroid of each transect was calculated, and a 90 m digital elevation model (DEM; NASA Shuttle Radar Topography Mission 2010) was used to generate slope and aspect for each transect. All GIS analysis was performed in ArcGIS (version 9.3 and version 10.0) using the WGS84 projected coordinates system.

All 35 bioclimatic variables were extracted from the ANUCLIM 6.1 software package (Xu and Hutchinson 2011) for each of the transect centroids. This software uses algorithms to combine information from regional weather stations and topographic surfaces in order to generate climatic predictions for point locations based on weekly (operator defined) averages for the 30 years from 1976 to 2005 (Xu and Hutchinson 2011). Additionally the Atlas of Living Australia (<spatial.ala.org.au>) was used to obtain site fertility data (inherent rock fertility rating; source CSIRO Ecosystem Services) for each transect centroid.
5.3.5 Broad-scale analysis

In order to reduce the complexity of the modelling approach described below and because large sets of bioclimatic variables typically show high levels of colinearity (Zuur et al. 2009) we selected a subset of the 35 bioclimatic variables. Variables were selected by firstly comparing pair-wise scatter plots and their correlation coefficients and selecting the five least correlated variables. Colinearity in this set of five variables was tested by calculating their variance inflation factors (VIF). Using a VIF threshold of 3 (Zuur et al. 2007, page 267) we removed one variable that showed colinearity resulting in a final set of four bioclimatic variables: mean annual temperature (degrees C); temperature seasonality; mean annual precipitation (ml); and precipitation seasonality; (Table 5.2). The seasonality measures are the coefficient of weekly variation in rainfall and temperature (see Xu and Hutchinson 2011), with high seasonality indicative of a pronounced season (e.g. a distinct rainfall season, or large seasonal differences in temperature), and low seasonality indicating consistency in the variable throughout the year (i.e. regular rainfall/temperatures all year).

In order to identify factors that influence wildling presence and abundance across the estate in the broad-scale survey we used a series of Generalized Linear Models (GLM). The GLM approach allows the incorporation of non-normal response distributions, which are often useful for analysing count data (Zuur et al. 2009). The presence or absence of wildlings (Models 1 and 2) and the number of wildlings per km per transect (Models 3, 4 and 5) were used as dependent variables in the modelling. Transect length varied considerably within and between regions and was included as a covariate in all models to weight the dependent variable by survey effort. It also allowed for length constrained model inference (i.e. to predict models for a convenient transect length of 1 km). The factors Burnt and Rotation were not equally represented across regions and were removed from the main dataset, leaving 198 transects all from first rotation, unburnt plantations; this dataset was used in Models 1, 2 and 3. The effect of Burnt and Rotation were tested separately within the regions where there was appropriate variation in those factors (Models 4 and 5). To test for non-linear relationships in the four bioclimatic variables and site fertility, we
fitted splines in Generalized Linear Mixed Models (GLMM). The dependent variables presence/absence of wildlings and wildlings/km were fitted against the fixed effects of Transect length, Region, Plantation age and the variable of interest, with the variable-spline as a random effect. The splines did not have a significant effect on any of the models. In the main GLM analysis the relatively large number of possible interactions between explanatory variables and Region (Models 2, 3 and 5) resulted in problems with model convergence. We dealt with this by fitting interactions one by one against each full model and using likelihood ratio tests to identify the interactions that had the most influence on model significance. As many of these significant interactions as possible (to allow convergence) were then included (in order of significance) in the full model prior to model simplification. We used a backward step model simplification process with nested likelihood ratio tests to identify the final best models (Zuur et al. 2009, pages 120-122). Because all models contained factors with multiple levels (e.g. ‘Region’ has 7 levels), the significance of each factor was determined by removing the factor and comparing the subsequent reduced model to the final model using nested likelihood ratio tests, producing discrete $\chi^2$ and $P$ values for each factor (Zuur et al. 2009, page 284).

There was a large number of transects where no wildlings were recorded so we initially tested factors associated with the presence or absence of wildlings using GLM with a binomial distribution and logit link function (Models 1 and 2: Table 5.2). We had an a priori hypothesis that Plantation age might be an important predictor of wildling presence. Model 1 (Table 5.2) tested the effects of Plantation age, Region and their interaction on the presence or absence of wildlings. Model 2 expanded this analysis to incorporate the effects of all bioclimatic, site and environmental variables (except Burnt and Rotation) including the interactions Region*Plantation age and Region*Precipitation seasonality on the presence and absence of wildlings (Table 5.2).
Chapter 5: Wildlings

Table 5.1. Summary of the regions surveyed for wildling establishment showing survey statistics, the raw broad-scale results, key model predictions, and the number of fine-scale plots undertaken to identify local factors associated with wildling establishment (see methods for more detail).

<table>
<thead>
<tr>
<th>Region (planting zone)</th>
<th>No. transects surveyed</th>
<th>No. burnt transects</th>
<th>No. 2nd rotation transects</th>
<th>Plantation age range</th>
<th>Distance surveyed (km)</th>
<th>Average transect length (m)</th>
<th>Total No. wildlings</th>
<th>No. wildlings/km</th>
<th>Predicted probability of wildlings ( d )</th>
<th>Predicted mean No. wildlings/km ( e )</th>
<th>No. fine-scale pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albany (WA)</td>
<td>31</td>
<td>0</td>
<td>8</td>
<td>3&quot; - 13</td>
<td>56.1</td>
<td>1811</td>
<td>1274</td>
<td>22.7</td>
<td>0.50</td>
<td>6.9 (12.2, 6.2)</td>
<td>22</td>
</tr>
<tr>
<td>Manjimup (WA)</td>
<td>33</td>
<td>0</td>
<td>24</td>
<td>8 - 18</td>
<td>21.5</td>
<td>650</td>
<td>851</td>
<td>39.6</td>
<td>0.40</td>
<td>44.9 (71.5, 40.4)</td>
<td>9</td>
</tr>
<tr>
<td>Grampians (GT)</td>
<td>19</td>
<td>2</td>
<td>0</td>
<td>9 - 12</td>
<td>17.2</td>
<td>907</td>
<td>624</td>
<td>36.3</td>
<td>0.61</td>
<td>4.9 (4.6, 4.2)</td>
<td>5</td>
</tr>
<tr>
<td>Penola (GT)</td>
<td>49</td>
<td>0</td>
<td>0</td>
<td>11 - 14</td>
<td>64.2</td>
<td>1310</td>
<td>75</td>
<td>1.2</td>
<td>0.28</td>
<td>0.6 (0.3, 0.5)</td>
<td>3</td>
</tr>
<tr>
<td>Portland (GT)</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>10 - 14</td>
<td>44.5</td>
<td>1113</td>
<td>1483</td>
<td>33.3</td>
<td>0.90</td>
<td>34.6 (9.8, 27.6)</td>
<td>20</td>
</tr>
<tr>
<td>Gippsland (GIP)</td>
<td>59</td>
<td>37</td>
<td>0</td>
<td>8 - 18</td>
<td>60.4</td>
<td>1023</td>
<td>525</td>
<td>8.7</td>
<td>0.39</td>
<td>1.3 (1.8, 1.2)</td>
<td>12</td>
</tr>
<tr>
<td>Tasmania (TAS)</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>6 - 14</td>
<td>26.5</td>
<td>697</td>
<td>107</td>
<td>4.0</td>
<td>0.42</td>
<td>0.8 (1.6, 0.8)</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>269</td>
<td>39</td>
<td>32</td>
<td>3&quot; - 18</td>
<td>290.4</td>
<td>1073</td>
<td>4939</td>
<td>17.0</td>
<td>-</td>
<td>-</td>
<td>71</td>
</tr>
</tbody>
</table>

\( a \) Planting zones are defined in Figure 5.1

\( b \) Note that Gippsland is the only region where plantations were surveyed within the native range of *Eucalyptus globulus*, see Figure 5.1

\( c \) Three year old plantations were second rotation

\( d \) The predicted probability of wildling presence along 1 km of plantation edge, for the mean within region plantation age, under Model 1

\( e \) Predicted mean number of wildlings (± standard error) along 1 km of plantation edge, for the mean within region plantation age, under Model 3
Chapter 5: Wildlings

Models 3, 4 and 5 fitted GLMs using the number of wildlings/km per transect as the dependent variable. In all three of these models a negative binomial distribution with a log link function was used to account for overdispersion associated with the relatively large number of zeros in the data. The value of one outlying transect (transect 260: Manjimup; 18 years old; first rotation = 1523 wildlings/km) was more than three orders of magnitude larger than any other, with large leverage and influence, and was replaced with the dataset mean (Barnett and Lewis 1994; Komsta and Komsta 2011). Model 3 tested the effect of all possible explanatory variables including the interactions Region*Plantation age and Region*Mean annual precipitation (Table 5.2) on the number of wildlings per km. Model 4 used a reduced dataset including only transects from Gippsland (n = 59) because this was the only region with enough burnt transects to test the effect of fire (Table 5.1). The model tested the effect of all variables (except Region and Rotation) including the interaction between Burnt*Plantation age. Model 5 also used a reduced dataset including only the regions Albany and Manjimup (both from the Western Australian planting zone) to test the effect of Rotation. Again, this model tested all possible variables (except Burnt) including the interaction between Rotation*Plantation age (Table 5.2).

5.3.6 Fine-scale analysis

There was insufficient variation to analyse Topographic aspect and Slope in the fine-scale analysis. Linear mixed models (LMM) were used to test variation in the remaining explanatory variables between fine-scale pairs. Nine pairs were removed because they were unbalanced for Burnt or Plantation age, leaving 62 pairs. The following transformations were made to improve normality: Bare ground, Litter and Herbs were all arcsine transformed, and Shrubs were square root transformed. Each variable was tested as the dependent variable separately, fitting the fixed effects of Region, Plot type (positive or control) and their interaction, with Pair nested within Region as the random term. Looking at only positive plots (n = 71), variation in wildling numbers between the inside and outside halves of each plot was tested (LMM) using the number of wildlings, adjusted for outside plots that were undersize because of the proximity of roads. To determine if any variation between wildling
numbers on the inside and wildling numbers on the outside was associated with different age groups, a cohort analysis was undertaken where wildling size classes two, three and four were merged (Old), and compared to size class one (Young). Linear mixed models were used to test the effect of inside versus outside, fitting region, the wildling measure (Total, Old or Young) and the interaction as fixed effects with Pair nested within Region as the random effect. The ground cover classes were also tested between the inside and outside of the plots using the same model structure.

Model fitting in the broad-scale analysis was undertaken using ASREML 2.0 (Gilmour et al. 2006), and R version 2.14.1 (R Development Core Team 2011) with the packages MASS (Venables and Ripley 2002), AED (Zuur 2010), lmtest (Zeileis and Hothorn 2002), lattice (Sarkar 2008) and outliers (Komsta and Komsta 2011). The broad-scale modelling often followed Zuur et al. (2009). Fine-scale analysis was undertaken in SAS™ (version 9.2) using PROC MIXED.

5.4 Results

A total of 4,939 wildlings were observed along the 290 km of surveyed plantation edge, giving a study-wide average of 17 wildlings per km (or c. 8/ha). The overwhelming majority of wildlings were juveniles, and of the few wildlings bearing adult leaves none were observed to be reproductively mature, although reproduction was not specifically assessed. There was considerable inter-regional variation in the number of wildlings/km, ranging from 1.2 in Penola to 39.6 in Manjimup (Table 5.1). Of the 269 transects surveyed, wildlings were recorded on 154 (57 %). Where wildlings were recorded, most were found close to or just within the plantation with wildlings recorded more than 10 m from the plantation edge on only 14 transects. Across these 14 transects a total of 80 wildlings (1.6% of the total recorded; see also Appendix 5.1) were found further than 10 m from the plantation boundary, equating to an average of 0.28 wildlings per km of boundary surveyed. Of these 80 wildlings, 66 (1.3 % of total recorded) were established in remnant native vegetation. While most were within 50 m of the boundary, at one location wildlings were found up to 175 m from the plantation edge. It should be noted that our surveys were conducted parallel to the boundaries, directly adjacent to the plantation edge making it possible
that we may have underestimated long distance dispersal. However, we only recorded 19 wildlings establishing between 10 m and 15 m from the plantation edge (in full view of the observers), and if long distance dispersal was more common, we would have expected to see many more wildlings in this zone. Furthermore the characteristic juvenile foliage of *E. globulus* makes wildlings highly detectable even at some distance (see Appendix 5.1). Although it is impossible to rule out missing occasional long distance dispersal events, given their rarity, the ready detectability of juveniles of this species, and the fact that we did detect several, we believe our results give an accurate impression of wildling dispersal from the plantation edge.

### 5.4.1 Broad-scale survey

Across all models (except Model 4 where it was not tested) there was a highly significant effect of Region (Table 5.2). The predicted probability of wildling occurrence along 1 km of plantation boundary under Model 1 ranged from 0.28 (Penola) to 0.9 (Portland) with most other regions close to the middle of this range (Region effect Model 1: $\chi^2 = 46.9$, df = 6, $P = < 0.0001$; Table 5.1). Model 1 also indicated that as plantation age increased (study range was 6 to 18 years) so did the predicted probability of wildling occurrence (Fig 5.2; $\chi^2 = 7.7$, df = 1, $P = 0.0055$). This effect was consistent across regions, and became more significant under Model 2 ($\chi^2 = 44.9$, df = 6, $P = < 0.0001$; Table 5.2; Appendix 5.2). Although there appears to be duplication in Model 1 and 2, Model 1 has been retained because we believe its simplicity makes it potentially useful for plantation managers (i.e. its ability to predict the probability of wildling presence based solely on Region and Plantation age, which is information plantations managers are likely to have). Model 2 indicated that there was also a significant effect of Region*Precipitation seasonality on the probability of wildling presence ($\chi^2 = 32.3$, df = 6, $P = < 0.0001$; Table 5.2). This interaction arose from a positive relationship between Precipitation seasonality and wildling presence in Albany and Tasmania and no, or a very slight negative (Penola), relationship in the other regions. Although the effect was relatively minor in Albany, Penola and Tasmania (and at best only marginally significant based on inspection of 95% confidence intervals) the inclusion of the interaction did improve the model fit,
with Model 1 explaining 17.2 % of the total variation and Model 2 explaining 30.3 %.

**Table 5.2.** All variables considered in the five broad-scale models and their significance. Models 1 and 2 are Generalized Linear Models (GLMs) with binomial distribution used to identify factors explaining the presence and absence of wildlings. Models 3, 4 and 5 are GLMs with negative binomial distributions used to assess which factors are associated with the abundance of wildlings. Models 1, 2 and 3 use a reduced data set with only first rotation and unburnt transects ($n = 198$). Model 4 use only transects in Gippsland ($n = 59$) to assess the effect of fire. Model 5 use only transects in Western Australia ($n = 64$) to assess the effect of rotation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Presence-absence models</th>
<th>Abundance models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Transect length</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Region</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Slope</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Aspect</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Burnt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plantation age</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Plantation edge aspect</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rotation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site fertility</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mean annual temperature</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Temperature seasonality</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mean annual precipitation</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Precipitation seasonality</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Region*Plantation age</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Region* Mean annual precipitation</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Region*Precipitation seasonality</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Burnt*Plantation age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotation*Plantation age</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Total variation explained (%) | 17.22 | 30.33 | 40.80 | 47.46 | 50.81 |

*Regions Albany and Manjimup only
b Not tested separately, see results

Symbols: ✓ = variable included in the full model but dropped during model simplification; ✓ = non-significant variable included in the final model; ✓ = marginally non-significant = P > 0.05, < 0.06; asterisks (*) = variable included in the final model and significant at * = P < 0.05, ** = P < 0.01, and *** = P < 0.001.
Fig. 5.2 Selected factors and their influence on the presence (a) or abundance (b and c) of *Eucalyptus globulus* wildlings according to broad-scale Generalized Linear Models (GLMs) based on surveys of 290 km of plantation edge: (a) the predicted probability of wildling occurrence along 1 km of plantation edge in the Grampians under Model 2; (b) the predicted effect of the three significant climatic variables (Mean annual precipitation (b-i), Mean annual temperature (b-ii), and Precipitation seasonality (b-iii)) on the number of wildlings/km of plantation edge in Portland under Model 3; (c) the predicted effect of fire on the number of wildlings/km in Gippsland by plantation age under Model 4. The solid lines represent the predicted means and dashed lines are 95% confidence intervals. None of the factors were affected by interactions.

When investigating the factors influencing wilding abundance in first rotation, unburnt plantations (Model 3) there was a significant Region*Plantation age interaction ($\chi^2 = 32.3, df = 6, P = 0.0005$; Table 5.2). This interaction was the result
of the number of wildlings/km increasing with Plantation age in Manjimup, Gippsland and the Grampians, while there was no significant effect of Plantation age in the other regions (Appendix 5.2). The effect of Plantation age on wildling abundance was complicated by differences in plantation age between regions (Table 5.1; Appendix 5.2) making it difficult to generalise across the study. In order to account for this and produce regional predicted mean numbers of wildlings/km from Model 3 (Table 5.1) we made model predictions using regional levels of the explanatory variables. We did this by predicting values across the range of plantation ages actually observed in each region, while holding the climatic variables at their regional means, and Transect length at 1 km. This model prediction produced trends similar to the observed number of wildlings per km for most regions but the predictions were usually slightly lower than the observed number of wildlings/km (Table 5.2). This may be partly explained by the reduced dataset used in Model 3. One exception in which the prediction was much lower than the observed data was the Grampians region that had an observed number of wildlings/km of 36.3, while the predicted mean under Model 3 was 4.9. Inspection of the raw data revealed that 71% of the wildlings in the Grampians were found along a single transect and real outliers (not errors) of this magnitude are likely to be explained by (difficult-to-model) local phenomenon (Barnett and Lewis 1994). This outlier was retained in the analysis because its influence and leverage were not as high as the outlier in Manjimup noted above. When this transect was removed the observed number of wildlings/km fell to 12.6, bringing it in line with the variation between observed and predicted values for the other regions (Table 5.1).

Model 3 shows that wildling abundance decreases with increasing Mean annual temperature and Precipitation seasonality, and increases with Mean annual precipitation (Table 5.2; Fig. 5.2). Low precipitation seasonality indicates rain occurs regularly throughout the year, for example in temperate-maritime climates. Therefore, Model 3 predicts that more wildings will occur in areas that receive regular, relatively high rainfall and have lower temperatures.

In Gippsland, there were significantly higher numbers of wildlings/km on burnt than on unburnt transects (\(\chi^2 = 12.0, \ df = 1, \ P = 0.0005\)) and the number increased
consistently with plantation age (Model 4, Fig. 5.2). In Western Australia, ‘rotation’ also affected the number of wildlings/km but this effect was strongly dependent on plantation age ($\chi^2 = 24.4$, df = 3, $P < 0.0001$; Model 5; Table 5.2). First rotation transects showed a marginally significant increase in the number of wildlings/km with plantation age while second rotation plantations had high numbers of wildlings in young plantations but numbers decreased with increasing age. Chi-square and $P$ values could not be obtained separately for the effects of plantation age and rotation in Model 5 because the removal of the interaction term prevented model convergence (Table 5.2).

![Figure 5.3](image)

**Figure 5.3.** a) The average number of young and old wildlings inside and outside the plantations in the positive plots (the 10 x 10m plots straddled the plantation edge so that half was inside and half was outside the plantation). b) Differences between the ground cover classes inside and outside the positive plots. Different letters represent significant differences between inside and outside for each individual factor i.e. comparisons are not between factors (a-b = $P < 0.0001$, c-d = $P < 0.01$, see results for details).

**5.4.2 Fine-scale survey**

The only variable that showed a significant difference between positive plots (with wildlings) and control plots (without wildlings) in the paired analysis was the reproductive output of the plantation, with wildlings occurring in areas where higher
proportions of trees had capsules (Wald’s test: $F_{1,50} = 67.3$, $P < 0.0001$). When considering only positive plots, there was significantly greater number of wildlings on the inside compared with the outside of the plantation (Wald’s test: $F_{1,54} = 23.3$, $< 0.0001$). This difference was the result of more old wildlings inside plantations (Wald’s test: $F_{1,55} = 20.3$, $< 0.0001$), as there was no difference in the number of young wildlings between inside or outside halves of the plots (Wald’s test: $F_{1,50} = 0.5$, $P = 0.4763$; Fig. 5.3). In terms of ground cover, there was significantly more bare ground outside plantations (Wald’s test: $F_{1,55} = 94.5$, $< 0.0001$), more litter inside (Wald’s test: $F_{1,55} = 41.1$, $P < 0.0001$), more shrubs inside (Wald’s test: $F_{1,55} = 12.0$, $P = 0.0011$), but no difference in herb cover (Wald’s test: $F_{1,55} = 1.8$, $P = 0.1807$; Fig. 5.3).

5.5 **Discussion**

5.5.1 *Establishment*

With the recent, rapid expansion of *E. globulus* plantations in Australia (Gavran and Parsons 2011), and the species listing as invasive in other countries (Rejmánek and Richardson 2011), we initially aimed to assess wildling establishment in order to determine the risk of invasion in Australia. We found that the spread of *E. globulus* wildlings is currently limited, and there is a low risk of serious invasions from industrial plantations. That is, under the unified framework for biological invasions (Blackburn *et al.* 2011), most plantations would fall within category C1 (introduced or naturalised but not spreading) because wildlings are locally recruited from cultivated plants and yet to reproduce. We did identify one example of 175 m dispersal, which may be considered a D1 situation (the first invasive stage, involving populations spreading), although there was still an absence of obvious reproduction, making its allocation somewhat difficult (Blackburn *et al.* 2011). Either way, in the vast majority of situations where wildlings occur, they would not be considered invasive (i.e. established or naturalised; Blackburn *et al.* 2011).

Overall the observed study wide average of 17 wildlings per km (8/ha) is low compared to well known invasive forestry species (e.g. *Pinus* species, Richardson *et al.* 1994). *Pinus* species have been widely invasive in the southern hemisphere, and
are spreading from plantations in densities of thousands of plants per ha in Australia, New Zealand and South Africa (Ledgard 2001; Richardson et al. 1994; Williams 2007). Other *Eucalyptus* species (in particular *E. camaldulensis* and *E. grandis* in South Africa) have also been recorded spreading from agroforestry plantings in high densities along waterways (Forsyth et al. 2004). However, a confounding factor in these comparisons is the young age of the Australian estate which is predominantly in its first rotation (Gavran and Parsons 2011; Table 5.1). In contrast, *Pinus* has been under intensive cultivation in the southern hemisphere since the early 1900’s (Richardson and Higgins 2000), as have eucalypts in South Africa (Forsyth et al. 2004).

A lag phase (or required residence time) between introduction and invasion is a common feature of plant invasions, although it is a poorly understood phenomena because most invasions are identified retrospectively (reviewed by Pysek and Jarosík 2005; Pysek et al. 2009). Consequently, few studies investigate the introduction/establishment phase of a potential invasion as we have here (Pysek and Jarosík 2005), raising the possibility that the low level of *E. globulus* establishment is a product of inadequate residence time (6 and 18 years in the plantations studied; Table 5.1). Williams (2007) found that a residence time of 40 years (rotation not specified) was required for *Pinus radiata* invasions from plantations in New South Wales, and younger plantations showed virtually no signs of invasion with levels of establishment similar to those in this study. Ritter & Yost (2009) also found that *Eucalyptus conferruminata* has recently become invasive after 50 years of widespread planting in California, and prior to its recent spread showed virtually no recruitment. We can find no direct evidence in the literature of a lag phase in *E. globulus* invasions in other countries. However, because invasions are identified retrospectively it cannot be discounted. *Eucalyptus globulus* is considered invasive in many areas where it has been a long time resident (Table 1 in Rejmánek and Richardson 2011), and there is some evidence that it is becoming increasingly invasive in Portugal after many decades of cultivation (Marchante et al. 2009; Silva and Marchante 2012). Residence time is thought to be influenced by propagule pressure, and biological interactions with the recipient community. For example, da Silva et al. (2009) found that *E. grandis* could not invade from plantations in Brazil.
because of unsuitable habitat provided by the adjacent native vegetation, in particular animal predation and competition with native plants. In Australia, biological similarities between *E. globulus* and native eucalypts probably mean that barriers within the recipient community are less likely to contribute to a lag phase. Also, Australian plantations are predominantly managed for pulp wood production on short 10 to 15 year rotations, which is likely to restrict the development of propagule pools. However, initial invasions require a suite of environmental factors to coincide to result in dispersal, establishment and persistence of satellite populations, and the likelihood of this occurring increases with time (Richardson *et al.* 2000). Therefore caution should be taken in declaring a species non-invasive (Pysek and Jarosík 2005), particularly if it is relatively recently introduced and planted on a large scale – as *E. globulus* is outside its native range in Australia.

5.5.2 Dispersal

The capacity for dispersal is an important predictor of invasiveness, with the potential to establish new populations over 100 m (often much further) from source plantings within 50 years suggested as minimum threshold for invasive trees (Richardson *et al.* 2000). We observed very limited dispersal in this study with less than 2 % of wildlings recorded further than 10 m from the plantation edge. This means that over 98 % of the wildlings observed are likely to be within the harvesting disturbance zone and are unlikely to reach reproductive maturity. Several authors have noted that limited seed dispersal may be important in preventing eucalypt invasions (Forsyth *et al.* 2004; Kirkpatrick 1977; Rejmánek and Richardson 2011; Richardson and Rejmánek 2011). In a study investigating the dispersal potential of eucalypt seed, Cremer (1977) showed that, of the 19 species tested, *E. globulus* had the heaviest seed and highest terminal velocity resulting in the lowest predicted dispersal distance. This study also showed that the majority of seeds from eucalypt trees fall within a horizontal distance equal to the height of the mother tree (Cremer 1977). Given an estimated average canopy height of 10 to 15 m in the plantations surveyed here, this prediction seems to correspond well with the dispersal distances that we identified. Examples of more effective dispersal in other invasive forestry species include *Pinus radiata*, which has been found at densities of over 2000
plants/ha up to 500 m from plantations in New South Wales, with occasional wildlings located up to 2.5 km from the plantation edge (Williams and Wardle 2005). *Eucalyptus camaldulensis* (river red gum) can also be dispersed long distances in streams and rivers and can be invasive outside Australia (Forsyth *et al.* 2004; Kirkpatrick 1977; Pettit and Froend 2001).

Despite seemingly limited dispersal, there were 14 transects in this study (80 wildlings) where establishment occurred further than 10 m from the plantation edge, with over two thirds of these wildlings establishing in remnant native vegetation. At one site wildlings were found up to 175 m from the plantation edge along an obvious drainage line, indicating that water dispersal may have been involved. *Eucalyptus globulus* seed has no particular adaptation for water or wind dispersal (e.g. floats or wings). The potentially invasive *E. camaldulensis* has double coated seed which enables the seed to float for up to 14 days allowing effective water transport (Pettit and Froend 2001). Although *E. globulus* seed is single coated, morphologically similar seed from other eucalypts has been shown to float for up to eight days (Pettit and Froend 2001), making some level of sporadic water transport feasible. Given the young age of the plantations studied here, these rare dispersal events have the potential to cause localised spread of *E. globulus* in the future. However, with such limited dispersal capacity and the characteristic juvenile foliage of *E. globulus*, control programs are likely to be successful (Appendix 5.1).

### 5.5.3 Regional variation

The variation observed between regions indicates establishment may be more likely in certain areas. There were high numbers of wildlings observed in Portland, Albany, the Grampians and Manjimup, with the latter two regions also having outlier transects with very high numbers of wildlings/km. One transect in Manjimup recorded the study-wide maximum of 1523 wildlings/km which is certainly in the realm of densities observed for other invasive trees (Williams 2007). In contrast, regions within or near the native range of *E. globulus* in Gippsland and Tasmania (although it should be noted that *E. globulus* does not naturally occur in the areas surveyed in Tasmania i.e. the far north west of the island; Fig. 5.1), had some of the lowest levels of establishment along with Penola. A key difference between regions
in this study was the range of plantation ages, and given that we found that the probability of wildling occurrence increased with plantation age this difference may explain some of the observed variation. The most-likely explanation for the effect of plantation age is that reproductive output is likely to increase with age, and we did find that wildling presence was associated with higher reproductive output. In a study covering the same planting zones investigated here, Barbour et al. (2008b) also found that *E. globulus* reproductive output generally increased with age, and in addition found considerable variation between regions within planting zones. Although there were no significant differences between planting zones, they did find generally lower levels of reproductive output in Tasmania, Gippsland and Penola, with the highest levels recorded in plantations in Western Australia and other parts of the Green Triangle (differences in plantation age between regions were not accounted for).

Genetics may also influence reproductive output. *Eucalyptus globulus* is a genetically well differentiated species (Dutkowski and Potts 1999) and has genetically controlled differences in precocity (onset of first flowering) within and between sub-races (Jordan *et al.* 1999). Selection in breeding programs may result in the deployment of germplasm with different levels of precocity in different areas, and this could result in regional variation in propagule pressure over time. For example selection in breeding programs has resulted in the early onset of flowering in some *E. camaldulensis* breeding populations in India (Varghese *et al.* 2009). Fecundity is also under genetic control in *E. globulus*, and a study investigating the reproductive output of 10 sub-races showed that the Strzelecki Ranges sub-race had the lowest rates of seed production and the smallest proportion of large seed (McGowen *et al.* 2004). The Strzelecki Ranges sub-race is a key component of the pulpwood breeding program in Australia (Jones *et al.* 2006) and is widely used in plantations, so it is possible that genetic variation in plantation fecundity could be influencing wildling establishment. These reproductive factors coupled with microsite requirements, plantation management and disturbance regimes make isolating the causal differences between regions difficult. However, the variation in establishment that we have identified does provide information on the regional risk
of wildling establishment, which could now be integrated into regional management/monitoring programs (Appendix 5.1).

5.5.4 Environmental variables

_Eucalyptus globulus_ naturally occurs in areas with frequent, intermediate to high rainfall, and low temperatures in Tasmania and southern Victoria (Fig. 5.1; Appendix 5.3). We found that wildling abundance increased in areas that had environmental conditions similar to those in the native range of _E. globulus_. That is, in areas with high Mean annual precipitation, low Precipitation seasonality (regular rainfall) and low Mean annual temperature (when accounting for Region and Plantation age under Model 3). Many exotic species establish well in areas with climates that are similar to their native range, and this is likely to reflect physiological adaptations to the environment in which they evolved (Richardson and Thuiller 2007). A study investigating the seed germination conditions of 415 eucalypt species showed that optimal germination temperature was associated with the temperature in the region of origin (Boland et al. 1980). Another characteristic of _Eucalyptus_ seed that possibly influences its need for moist germination conditions is the lack of an endosperm, with the energy for early growth and root development provided by the photosynthetic cotyledons (Boland et al. 1980). This trait is thought to result in high seedling mortality as a result of water stress prior to effective root development (Cremer 1977), and may account for the higher levels of _E. globulus_ establishment in areas with frequent relatively high rainfall. Although environmental conditions are important in enabling establishment and consequent invasions, they interact with other extrinsic factors (e.g. introduction effort and biological interactions) to contribute to invasion potential. For example, although the home range climatic conditions of 12 _Pinus_ species were found to predict their invasion success, it did not predict failed invasions (i.e. areas where the species had been long-time residents but had not become invasive) suggesting the importance of other factors (Nuñez and Medley 2011). This highlights the need to consider multiple interacting factors in predicting and managing invasion risk. Our Model 3, incorporating region, plantation age and the climatic factors is our best
approximation of establishment risk in *E. globulus* and should help in developing monitoring protocols across the Australian estate (Appendix 5.1).

One factor not incorporated in Model 3 is the effect of fire. *Eucalyptus* is one of the most fire adapted floras in the world (Crisp *et al.* 2011), and fire had a strong positive influence on wildling establishment in the only region where we could test its effect (Gippsland). Fire has been shown to promote eucalypt seedling establishment in many ways. It causes the release of a large canopy stored seed-bank, opens up patches of bare ground, provides a nutrient rich ash-bed, reduces competition from understory species, and increases light availability at the forest floor – all of which contribute to successful regeneration from seed in natural eucalypt forests (Attiwill 1994). Furthermore, in contrast to other countries (where fire often increases soil water repellency: Debano 2000), fires seems to reduce the water repellency of Australian soil (as a result of fire intensity in eucalypt dominated forests) creating microsites with greater water availability that promote seedling establishment and survival (Bailey *et al.* 2012; Granged *et al.* 2011). Although we could not test the effect of fire outside of Gippsland, the occurrence of fire is likely to promote wildling establishment across the *E. globulus* estate in Australia. Fire is also likely to influence establishment and invasion in other countries, particularly in areas with relatively frequent wildfires where *E. globulus* is widely planted, such as California and the Iberian Peninsula (Rejmánek and Richardson 2011; Silva *et al.* 2011). Indeed, high wildling densities have been reported following large wildfires in Portugal in 2005 (Silva and Marchante 2012). This suggests that plantation management aimed at excluding fire is indirectly reducing the risk of invasion. If plantations do burn then the fire could also act as a management trigger to implement future surveys (Appendix 5.1).

5.5.5 Management

Our surveys were undertaken predominantly along firebreaks (including roads). These firebreaks are maintained on an annual or biannual basis, typically involving surface scrapping, slashing or the application of herbicide to remove vegetation. Therefore the higher numbers of older *E. globulus* wildlings just inside the plantation may be a consequence of wildlings being removed from outside the plantation during
firebreak maintenance. This side effect of existing management may be acting to reduce the likelihood of establishment and spread from the plantation edge in Australia, particularly given the limited dispersal that we have identified.

In Australia, plantations are managed predominantly for pulp-wood on 10 to 15 year rotations (Greaves et al. 2003). These relatively short rotations effectively maintain a low residence time for individual plantations, which may help prevent the build up of the propagule pool that comes with age. There is currently interest in transforming parts of the existing Australian *E. globulus* estate into saw logs plantations, which could see the rotation time increase to 40 to 50 years (Beadle et al. 2008; Forrester et al. 2010), significantly increasing the potential residence time. Saw-log plantations also employ different silvicultural techniques such as thinning, which could increase the reproductive output of plantations (Williams et al. 2006). Another potential future management issue is plantation abandonment. During the rapid expansion of the plantation estate in Australia over the past decade, some plantations were established on marginal sites and these have performed poorly in many cases (Sudmeyer and Simons 2008). If harvesting in these areas proves uneconomical there is a possibility that plantations could be abandoned (left unharvested and unmanaged), removing current management and age related barriers to invasion.

Gavran and Parsons (2011) note that some plantations previously owned by managed investment schemes in Australia were written-off in 2010 (c. 20,000 ha; M. Gavran pers. comm). Some of these plantations have been sold for conversion back to agriculture, but the fate of the others is unknown (M. Gavran pers. comm). If these plantations are left intact and unmanaged they may become long term sources of propagules in the future, this would be of particular concern in areas adjacent to, or embedded within, native forest. In Portugal plantation abandonment and the subsequent lack of management is thought to have contributed to landscape transitions towards *E. globulus* dominated, and *E. globulus* mixed communities; as well as increasing fire risk and therefore the probability of further establishment and spread (Silva and Marchante 2012; Silva et al. 2011). In Australia, the possibility of changing plantation management highlights the need for continued monitoring of wildling establishment to enable early detection of any future problems.
Chapter 5: Wildlings

5.5.6 The general invasiveness of *Eucalyptus*

If poor dispersal is a general characteristic of *Eucalyptus* species it may limit their ability to invade. We identified poor dispersal capacity in *E. globulus*, and Callaham Jr. et al. (2001) recently reported similar results from a detailed local study of a range of *Eucalyptus* species (not including *E. globulus*) in plantations in south-eastern USA. These two studies across regions and taxa provide *in situ* quantitative support to the suggestion by several authors that dispersal limits *Eucalyptus* invasions (Forsyth *et al.* 2004; Kirkpatrick 1977; Kunin 1997; Rejmánek and Richardson 2011; Richardson and Rejmánek 2011).

The obvious link identified between reproductive output and establishment may help explain the generally low invasiveness of *Eucalyptus*. Industrial *Eucalyptus* plantations are grown on very short rotations, so their ability to develop a significant propagule pool is diminished. The average rotation time in Australia is 10 – 15 years, whereas in Brazil, which has the largest exotic *Eucalyptus* estate in the world (4.2 m ha; Kunin 1997), plantations turnover every 6 - 7 years (Ghosh and Haccou 2010). The effect can also be seen in increasing numbers of wildlings beside older plantations, a pattern also observed in *Eucalyptus* plantations in the USA (Fishman and Willis 2001). Thus, short rotations probably prevent the build up of propagule pools, limiting the number of invasions relative to the size of the global *Eucalyptus* plantation estate.

Finally, we found that climate matching may play a role in invasiveness in *E. globulus*. Climate is thought to limit the invasiveness of several *Eucalyptus* species in the USA (Kunin 1997; Lexer *et al.* 2010). However, with such widespread planting across a range of environments and deliberate selection for forestry, good climate matches are probably quite common, and other factors such as propagule pressure and residence time are also likely to be important. This same pattern of multiple drivers has been shown in *Pinus* (Essl *et al.* 2010; Nuñez and Medley 2011) and *Acacia* (Procheș *et al.* 2012), with climate matching representing one aspect of invasive potential.
5.6 Conclusion

We have identified that there are currently low levels of establishment of *E. globulus* from managed industrial plantations, and the species poses a low risk of invasion in Australia. However, the young age of the estate means that the risk could increase in the future. We have identified regional patterns, as well as climatic and environmental variables that should assist in developing targeted monitoring programs and risk assessment protocols to reduce the likelihood of invasion from *E. globulus* plantations in the future (for detailed discussions on the implementation of weed risk assessments in Australia, see Byrne and Stone 2011; Stone and Byrne 2011). This will be particularly important if plantation management techniques, which currently seem to be reducing wildling establishment, change in the future. The factors shown to promote establishment, particularly the influence of plantation age, reproductive output, fire and climate, may prove useful for managers in other countries where *E. globulus* is invasive. Finally, our study has shed some light on the general invasive potential of *Eucalyptus*. In particular that limited dispersal and low reproductive output under the short rotations employed in most industrial plantation may act to restrict the invasiveness of the genus.
Appendix 5.1: Implications and recommendations for the management of *Eucalyptus globulus* wildlings

Minimising the offsite effect of plantation forestry (including wildling spread) is one of the key indicators of sustainable forest management in Australia (Powell 1983), and will also be important in forest certification schemes (e.g. Forest Stewardship Council), which most Australian plantation companies now operate under (Dare *et al.* 2011). We found 4,939 wildlings (17 wildlings/km) in this study, and only 80 of those were established further than 10m from the plantation edge. Assuming that all wildlings within 10m of the plantation edge are within the harvesting disturbance zone and unlikely to survive, these 80 wildlings are of most concern from a management perspective. To put this number of 80 in the context of the entire estate, it is estimated, from the plantation GIS layer, that there are approximately 300,000 km of *E. globulus* plantation edge in Australia, meaning there could be 84,000 wildlings requiring management attention. Plantation boundaries are regularly managed in Australia, with vegetation free fire breaks maintained around plantation edges. By integrating wildling management into existing fire break and weed management protocols, the added costs to plantation managers for control of wildlings are likely to be low. The distinctive juvenile characteristics of *E. globulus* make wildlings highly detectable (Fig. A.5.1.1) meaning these integrated approaches are likely to be effective. Furthermore because boundaries are managed on a regular basis (one to two years), there are likely to be multiple opportunities for detection and control over the life of a plantation. Below is a point form summary of the wildling control recommendations arising from this study:

*Maintaining fire breaks helps control wildling spread* – seed dispersal seems to be limited with the vast majority of wildlings occurring within 10 m of the plantation edge. This means that maintaining a vegetation free fire break (preferably 10 m wide; Fig. A.5.1.1) around plantation boundaries will strongly limit wildling establishment potential.

*Integrate spot wildling control into existing management activities* – because plantation boundaries are regularly managed in Australia and because the density of wildlings requiring control is low, an effective method of control would be to
undertake spot weeding wherever a wildling (outside the plantation disturbance zone) is identified during other management activities (e.g. fire break maintenance). Control methods that are likely to be most effective are hand pulling of small wildlings and cut stump herbicide application for larger plants (see www.wildingconifers.org.nz for an excellent guide to wildling tree control techniques). This would mean that wildling control kits would need to be carried in all field vehicles.

*Spot weeding locations should be recorded with a GPS* – to allow quantification of success rates (possibly the following year or at harvest) and long term trends in wildling establish and control success to be monitored.

*In the case of fire, implement targeted surveys* – because fire is such a strong trigger for recruitment we recommend undertaking targeted surveys around burnt plantations that are reproductively mature (3 years or older: Barbour et al. 2008b), two or three years post fire to identify areas requiring control. This could be integrated into GIS databases so that areas requiring surveys are flagged when surveys are due.

*Target areas with recruitment* – because recruitment is related to reproductive output it is likely to be patchy so if wildlings are detected close to the plantation edge this should be a trigger to inspect further afield (i.e. outside the plantation disturbance zone).

*Record significant flowering events* – because establishment is strongly linked to reproductive output, large scale regional or sub-regional flowering events should be recorded so that those plantations can be flagged as high risk in future assessments.

*Plantation age, climate and region* – our models have identified that older plantations, receiving regular relatively high rainfall with low temperatures are most likely to have wildlings establishing. Also the regional establishment probabilities shown in Table 5.1 (main text) should help highlight the higher risk regions. This information could all be used to flag the areas that are most at risk on work schedules for on-ground staff working in these areas.
Figure A.5.1.1. The distinctive juvenile foliage of *Eucalyptus globulus* (foreground right and arrows) enables easy detection of wildlings establishing outside the plantation disturbance zone. Also note the vegetation free fire break between the plantation left and the neighbouring remnant vegetation right. The fire break here is 15 m wide.
Appendix 5.2: Additional supporting information for the broad-scale modelling of *Eucalyptus globulus* wildling establishment

**Table A.5.2.1** Summary of Model 2 used to assess factors associated with the presence of *E. globulus* wildlings in first rotation and unburnt plantations. This model explains 30.33 % of the total variation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transect Length</td>
<td>1</td>
<td>3.82</td>
<td>0.0507</td>
</tr>
<tr>
<td>Region</td>
<td>6</td>
<td>44.87</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Plantation age</td>
<td>1</td>
<td>13.69</td>
<td>0.0002</td>
</tr>
<tr>
<td>Precipitation seasonality</td>
<td>1</td>
<td>3.61</td>
<td>0.0576</td>
</tr>
<tr>
<td>Region*Precipitation seasonality</td>
<td>6</td>
<td>32.32</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

**Table A.5.2.2** Summary of Model 3, which was used to assess factors associated with *E. globulus* wildling abundance in first rotation and unburnt plantations. This model explains 44.80 % of the total variation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transect Length</td>
<td>1</td>
<td>1.89</td>
<td>0.1691</td>
</tr>
<tr>
<td>Region</td>
<td>6</td>
<td>56.70</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Plantation age</td>
<td>1</td>
<td>4.93</td>
<td>0.0264</td>
</tr>
<tr>
<td>Mean annual temperature</td>
<td>1</td>
<td>10.94</td>
<td>0.0009</td>
</tr>
<tr>
<td>Mean annual precipitation (log)</td>
<td>1</td>
<td>22.25</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Precipitation seasonality</td>
<td>1</td>
<td>10.08</td>
<td>0.0015</td>
</tr>
<tr>
<td>Region*Plantation age</td>
<td>6</td>
<td>32.32</td>
<td>0.0005</td>
</tr>
</tbody>
</table>
Fig. A.5.2.1 The predicted number of wildlings/km by plantation age (with 95% confidence intervals) for all regions under Model 3. The interaction between plantation age and region (Table 5.2, main text) is the result of a significant increase in wildling numbers with plantation age in Manjimup, Grampians and Gippsland, and no significant variation with plantation age in the other regions (note the different age ranges).
Appendix 5.3: Summary of three key bioclimatic variables across the native range of *Eucalyptus globulus*

We identified that *E. globulus* wildling abundance increased in areas with high Mean annual precipitation, low Mean annual temperature and low Precipitation seasonality (regular rainfall). In order to estimate an average value for these three bioclimatic variables across the native range of *E. globulus*, we used the geographic location of 5 to 29 native trees from each of the 13 recognised races across the species range (Dutkowski and Potts 1999; Table A.5.3.1) to estimate the values of the three variables from ANUCLIM (Xu and Hutchinson 2011). The race averages were calculated and these values then used to estimate an average species wide value for the climatic variables.

The estimated average Mean annual temperature (12.5 degrees C) and Precipitation seasonality (22.1 coefficient of variation, see paper) across the native range are low, while Mean annual Precipitation (1010 mm) is intermediate (Table A.5.3.1; and Fig 5.3 main paper). This indicates that values of the bioclimatic variables that were found to promote wildling establishment, when accounting for Region and Plantation age (Model 3), are similar to the values of those variables where *E. globulus* grows naturally, particularly for Mean annual temperature and Precipitation seasonality. The trend with Mean annual precipitation was not as strong; however this factor is highly variable, ranging from 725 mm (Southern Furneaux) to 1845 mm (Western Tasmania).
Table A.5.3.1 Summary of Mean annual precipitation, Mean annual temperature and Precipitation seasonality for the 13 races of *E. globulus* across the species natural range. The Australian state in which the race occurs (VIC = Victoria, TAS = Tasmania), and the number of trees/locations within the geographic range of the race that were used to calculate the race averages are shown.

<table>
<thead>
<tr>
<th>Race (Eucalyptus globulus)</th>
<th>State</th>
<th>n trees/locations</th>
<th>Mean annual precipitation (mm)</th>
<th>Mean annual temperature (degrees C)</th>
<th>Precipitation seasonality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Plain (Gippsland)</td>
<td>VIC</td>
<td>14</td>
<td>889</td>
<td>14.3</td>
<td>20.1</td>
</tr>
<tr>
<td>Eastern Otways</td>
<td>VIC</td>
<td>9</td>
<td>950</td>
<td>13.4</td>
<td>30.2</td>
</tr>
<tr>
<td>Strzelecki Ranges</td>
<td>VIC</td>
<td>6</td>
<td>1128</td>
<td>12.6</td>
<td>20.0</td>
</tr>
<tr>
<td>Western Otways</td>
<td>VIC</td>
<td>10</td>
<td>1093</td>
<td>13.4</td>
<td>30.2</td>
</tr>
<tr>
<td>Flinders Island</td>
<td>TAS</td>
<td>6</td>
<td>784</td>
<td>13.2</td>
<td>20.3</td>
</tr>
<tr>
<td>King Island</td>
<td>TAS</td>
<td>15</td>
<td>936</td>
<td>13.3</td>
<td>38.2</td>
</tr>
<tr>
<td>North-eastern Tasmania</td>
<td>TAS</td>
<td>9</td>
<td>873</td>
<td>11.2</td>
<td>13.4</td>
</tr>
<tr>
<td>Recherche Bay</td>
<td>TAS</td>
<td>10</td>
<td>1232</td>
<td>11.5</td>
<td>20.2</td>
</tr>
<tr>
<td>South-eastern Tasmania</td>
<td>TAS</td>
<td>16</td>
<td>848</td>
<td>10.4</td>
<td>16.3</td>
</tr>
<tr>
<td>Southern Furneaux</td>
<td>TAS</td>
<td>27</td>
<td>725</td>
<td>13.4</td>
<td>17.6</td>
</tr>
<tr>
<td>Southern Tasmania</td>
<td>TAS</td>
<td>29</td>
<td>1017</td>
<td>10.8</td>
<td>19.7</td>
</tr>
<tr>
<td>St Helens</td>
<td>TAS</td>
<td>17</td>
<td>808</td>
<td>13.2</td>
<td>14.8</td>
</tr>
<tr>
<td>Western Tasmania</td>
<td>TAS</td>
<td>5</td>
<td>1845</td>
<td>11.8</td>
<td>26.2</td>
</tr>
<tr>
<td><strong>Total/average</strong></td>
<td></td>
<td><strong>173</strong></td>
<td><strong>1010</strong></td>
<td><strong>12.5</strong></td>
<td><strong>22.1</strong></td>
</tr>
</tbody>
</table>
Chapter 6:
General discussion – the risk of exotic gene flow from *Eucalyptus globulus* plantations in Australia

The rapid expansion of *E. globulus* plantations in southern Australia has resulted in the potential for exotic gene flow, which could affect the integrity of natural eucalypt gene pools (Barbour *et al.* 2008b; Potts *et al.* 2003). This thesis investigated the risk of such gene flow in order to better understand its scope and identify possible management options to limit any negative effects. This general discussion summarises the key findings of the thesis, considers some aspects not previously addressed, place the risk of exotic gene flow in the context of other environmental challenges including those resulting from rapid climate change, and concludes with an assessment of the importance of exotic gene flow as risk factor in native forest management in Australia.

6.1 *Summary of key findings*

Introgression depends on the capacity of species to reproduce and generate viable hybrids (Anderson 1951; Heiser 1949). Given the phylogenetic diversity of *Eucalyptus* (Byrne 2008), understanding phylogenetic barriers to reproduction is of central importance for identifying which species are at risk of exotic gene flow from *E. globulus* plantations. Prior to this study any of the 484 *Symphyomyrtus* species (which form four main clades; Steane *et al.* 2011) that occurred within the pollen dispersal zone of *E. globulus* plantations were thought to be potentially at risk of exotic gene flow (Barbour *et al.* 2008b; Potts *et al.* 2003). However, by combining
data from crossing experiments and detailed molecular phylogenies it is obvious that barriers to hybridisation exist within *Symphyomyrtus* (Chapter 2). Specifically there appears to be a virtually complete barrier to hybridisation between *E. globulus* and species in clades 3 and 4 (Chapter 2, Fig. 2.1). This finding alone reduces the number of species at risk of exotic gene flow by over 70% (to 138 species; Chapter 2). Barriers within *Symphyomyrtus* also extend to the more closely related clades I and II, with the probability of successful hybridisation between *E. globulus* and species in clade II being less than 5% post matting (Chapter 2, Fig. 2.1). The species at greatest risk were those within clade I, which includes only species from section *Maidenaria* (68 species). The overall trend for a reduction in compatibility with increasing genetic distance means that there is likely to be variation in compatibility within *Maidenaria*, but most species are likely to be at least partially cross-compatible.

Chapter 2 also provided insights into the nature of reproductive isolation in *Eucalyptus*. It appears that both pre- and postzygotic mechanisms may act to influence post-mating reproductive isolation. There is some evidence that pre-dispersal isolation (including prezygotic barriers) is influenced by natural selection (reinforcement) in the full dataset, while patterns of post dispersal isolation are more consistent with genetic drift in both datasets. Consistent with these patterns, pre-dispersal barriers seem to be responsible for a greater proportion of incompatibility than post-dispersal barriers. Although as shown in chapter 4 and by several other authors (Barbour *et al.* 2006a; Costa e Silva *et al.* 2012; Lopez *et al.* 2000a) the estimate of hybrid survival at nine months probably underestimates the true strength of post-dispersal barriers (for example see Costa e Silva *et al.* 2012). Finally, by comparing the results to the most recently published dated phylogeny (Crisp *et al.* 2011), it was possible to give some broad estimates of how long it takes for complete reproductive isolation to evolve in *Eucalyptus*, which appears be 21-31 million years.

The crossing program has revealed a significant reduction in the number of at risk species, but there are still 138 species (clades 1 and 2) that are exposed to some level of risk. Given the size of the *E. globulus* estate and the diversity of morphological
characteristics of the remaining at risk species, an approach for hybrid identification/validation is need to enable effective monitoring of this genetic risk (Byrne et al. 2011; Laikre et al. 2010). In Chapter 3, a Bayesian modelling approach using multi-locus microsatellite genotypes to identify hybrids between *E. globulus* and five at risk species from clades I and II was tested. By using simulations, I showed that the technique was highly effective at identifying F1 hybrids, which are currently the primary concern given the age of the Australian *E. globulus* estate. The microsatellite dataset compiled for this study could also be expanded to include extra species or samples if a specific threat was recognised.

It is clear from Chapter 2 that management attention needs to focus on species from clades I and II, particularly those in sections *Maidenaria* (clade I; Fig. 2.1 Chapter 2). Chapter 4 looked in detailed at the threat posed to one such species. *Eucalyptus ovata* (*Maidenaria*) seems to be a prime candidate for introgression from *E. globulus* plantations (Barbour et al. 2008b): it is of conservation significance; a common plantation neighbour; is known to hybridise with *E. globulus* where they naturally co-occur; and exotic hybrids have been identified in open pollinated seed and establishing in native forest adjacent to plantations (Chapter 4). However by assessing multiple stages in the introgression pathway *in situ*, I found that the actual risk posed to this species is low (Chapter 4). The level of exotic hybridisation (1.62%) was consistent with that found amongst co-occurring native eucalypts, (Potts et al. 2003). Landscape context affected the likelihood of hybridisation, with small patches of native forest and trees on the edge of large patches most at risk of hybridisation, but the rate of hybridisation declined very rapidly inside native forest patches (Fig. 4.2, Chapter 4). The rate of hybrid establishment was also low despite surveys targeting high-risk sites (Fig. 4.3, Chapter 4). However, the key barrier operating after hybrids are produced appears to be hybrid survival, with naturally established hybrids showing a 78% reduction in survival compared to their pure native siblings in a competitive native forest environment (Chapter 4). It therefore seems likely that few if any hybrids will reach reproductive maturity to enable backcrossing and introgression. What this study did find however was that pure *E. globulus* wildlings were establishing in higher numbers than hybrids (Fig. 4.3,
Chapter 6: General discussion

Chapter 4), raising the concern that they could pose a threat to native forests around plantations.

The spread of *E. globulus* wildlings could initially be viewed as an invasive weed issue, but could also lead to exotic gene flow and introgression when the invading wildlings reach reproductive maturity (Chapter 5; Larcombe *et al.* 2013). In Chapter 5, quantification of wildling spread across the Australian *E. globulus* plantation estate showed that establishment is relatively low with the vast majority of wildlings occurring within the plantation disturbance zone. Wildlings were more common around older plantations, plantations with high reproductive output, those that had been burnt in wildfire and in areas with environmental conditions similar to those where *E. globulus* occurs naturally (Chapter 5, Fig. 5.2). It also appears that current management practices, including short rotations and firebreak maintenance, are reducing the risk of wildling spread (Chapter 5). This being said, the Australian plantation estate is young and there has recently been moves to shift towards longer rotations for solid wood production which could increase reproductive capacity and wildling establishment (Barbour *et al.* 2008b). Additionally, some sites planted during the recent rapid expansion of the Australian estate have proved unsuitable (Sudmeyer and Simons 2008), and if these sites are abandoned (left unharvested and unmanaged) they could lead to wildling management issues in the future. Therefore continued monitoring of wildling spread is warranted and should be integrated into existing weed management protocols (Chapter 5; Larcombe *et al.* 2013).

The possibility that seed mediated gene dispersal could pose a greater threat to the integrity of native forests than pollen mediated gene flow was an unexpected finding. However, it is in line with a recent study by Wheeler *et al.* (2013) that found no evidence of exotic gene flow between the rare *E. gomphocephala* and introduced *E. cladocalyx* in Kings Park in Perth Western Australia, but concluded that invasion of pure *E. cladocalyx* seedlings was a more serious threat to remnant forest in Kings Park (Wheeler *et al.* 2013). This is the latest in a growing list of studies that have found significant barriers to inter-specific gene flow in *Eucalyptus* (Barbour *et al.* 2006a; Barbour *et al.* 2006b; Barbour *et al.* 2010; Costa e Silva *et al.* 2012; Gore *et al.* 1990; Lopez *et al.* 2000a; Wheeler *et al.* 2013).
6.2 Other potential issues

Although this thesis has shown that there are significant barriers to gene flow from *E. globulus* plantations, the risk cannot be entirely dismissed. For example hybrids surviving in the paired trial (Chapter 4) could yet reach reproductive maturity and backcross to native species. Although there is evidence of additional later generation barriers between *E. globulus x ovata* (Costa e Silva *et al.* 2012; Lopez *et al.* 2000a), other crosses involving *E. globulus* and species from section *Maidenaria* seem to show few fitness problems (Potts *et al.* 2003). For example crosses between *E. globulus* and *E. gunnii* produce large numbers of F\textsubscript{1} hybrids with few abnormalities (Potts and Dungey 2004), and *E. globulus x cypellocarpa* hybrids seem to readily reach reproductive maturity in the wild (Chapter 3; Kirkpatrick *et al.* 1973). Therefore there may be specific hybrid combinations that are more likely to lead to introgression because the barriers identified in this thesis do not prevent gene flow. Furthermore, even if the level of gene flow in these situations is low, it can still have major consequence for the recipient population, particularly over evolutionary time scales (Arnold *et al.* 1999). In situations where there are thought to be few post-mating barriers there will be two key considerations in determining the likelihood and the consequences of gene flow from *E. globulus* plantations: 1) the presence of pre-mating barriers, particularly flowering time or flower size differences, that could prevent mating; and 2) the conservation status of the at-risk populations.

If species are considered to be at risk because barriers identified in this thesis do not prevent gene flow, then an understanding of effective pre-mating barriers will be important for quantifying the likelihood of introgression. Two key pre-mating barriers that have proven to be important for preventing gene flow from *E. nitens* plantations in Tasmania are flower size and flowering synchrony (Potts *et al.* 2003). Both these factor can be measured by undertaking flowering surveys (Barbour *et al.* 2006b) and experimental pollinations (Barbour *et al.* 2005a; Gore *et al.* 1990) if a high risk situation is identified. For example exotic hybridisation from an *E. nitens* plantation in Tasmania was detected in open pollinated seed lots from neighbouring *E. ovata* but not *E. viminalis* despite both being cross compatible (Barbour *et al.* 2005a; Barbour *et al.* 2006b). A flowering survey showed that the difference was
likely to be a result of asynchronous flowering between the plantations and *E. viminalis*, while there was considerable overlap with *E. ovata* (Barbour *et al.* 2006b). It is also thought that flower size is strong barrier to gene flow between exotic *E. nitens* (small flowers) and both *E. urnigera* and *E. globulus* (large flowers) in Tasmania (Barbour *et al.* 2005a; Gore *et al.* 1990). The mechanism driving this barrier is that the pollen tubes of small flowered species cannot reach the ovaries of large flowered species (Gore *et al.* 1990). *E. globulus* has intermediate seized flowers so only species with very large flowered are likely to be reproductively isolated via this mechanism, and no effect of flower size was detected in the crossing study here (Chapter 2).

Understanding the conservation significance of at-risk populations is important for quantifying the consequences of gene flow if it does occur (Barbour *et al.* 2010; Byrne *et al.* 2011; Laikre *et al.* 2010). Rare species with small populations are particularly vulnerable to the negative genetic effects of exotic gene flow (Ellstrand and Elam 1993), and there is evidence of demographic and genetic erosion of small populations by more widespread or introduced congeneres in other plant and animal systems (Ellstrand and Elam 1993; Hwang *et al.* 2012; Roberts *et al.* 2010). Therefore preventing gene flow into populations of rare eucalypt species in section *Maidenaria* should be a high priority when designing new *E. globulus* plantations, because the consequences for such populations are likely to be more severe than they would be for widespread species with large populations (Barbour *et al.* 2010; Larcombe *et al.* 2014; Potts *et al.* 2003). Other values such as cultural and natural heritage will also need to be considered when assessing genetic risk (Barbour *et al.* 2010; Potts *et al.* 2003). For example widespread species are of greater cultural and conservation significance when growing within conservation areas such as National Parks (Potts *et al.* 2003), meaning the risk *E. globulus* poses to certain species may vary depending on the landscape context. Barbour *et al.* (2010) investigated the risk of gene flow from exotic plantations to all listed rare *Eucalyptus* species in Australia, identifying 22 high-risk species (those found within 10km of plantations), eight of which were of particular concern occurring within 1 km of plantations. They found that a detailed understating of the breeding system of the species involved could identify unexpected barriers to gene flow (e.g. *E. perriniana* Chapter I; Barbour *et
al. 2010), and identify high priority areas/species requiring conservation attention (Barbour et al. 2010).

6.3 Intra-specific gene flow

Although I have shown there are considerable barriers to inter-specific gene flow in *Eucalyptus*, exotic intra-specific gene flow remains largely unquantified in the genus. In one of the only studies to assess this issue, Sampson and Byrne (2008) found extensive gene flow between revegetation plantings of the widespread and common *E. loxophleba* ssp. *lissophloia* and the rare endemic *E. loxophleba* ssp. *supralaevis* in western Australia. Intra-specific introgression is likely to be more common because of an absence of reproductive barriers, making gene pool homogenisation a real possibility (Sampson and Byrne 2008). For example the local endemic *E. loxophleba* ssp. *supralaevis*, is a tree while the common *E. loxophleba* ssp. *lissophloia* is a mallee, so genetic swamping could lead to the loss of the tree form, causing a loss in morphological and genetic diversity (Sampson and Byrne 2008). However, such gene flow could also have a positive influence if the small remnant population is suffering the negative effects of inbreeding (Ellstrand and Elam 1993), or if the gene flow enables the population to better adapt to rapid climate change (see below and Aitken and Whitlock 2013).

*Eucalyptus globulus* is a genetically well-differentiated species (Dutkowski and Potts 1999) with considerable genetic-based morphological and functional trait variation. Intra-specific gene flow from *E. globulus* plantations to native populations could affect the distribution of this morphological and genetic structure. Plantation germplasm is typically derived from genetically improved seed orchards (Jones et al. 2006). Even though there are likely to be few post-mating barriers to hybridisation across the range of *E. globulus*, there is significant flowering time variation between races, which could provide a pre-mating barrier to gene flow (Jones et al. 2011). This variation could be exploited at the plantation design stage as a management approach to minimise intra-specific gene flow – although validation of the source of plantation stock might be difficult. Negative consequences of intra-specific gene flow are likely to be less severe than the consequences of inter-specific gene flow. However, detection of intra-specific hybrids is likely to be difficult. Molecular
approaches such as the one used in Chapter 3, would require more detailed sampling of races to determine if adequate differentiation existed to identify hybrids. A recent study was able to identify major lineages within *E. globulus* using a similar approach to the one in Chapter 2, except a more detailed population sampling was used (Jones *et al.* 2012). Even if hybrids could be detected with such methods, expensive mass screening of progeny arrays would be required to identify gene flow because of a lack of morphological characters at the seedling stage to visually detect putative hybrids. In the absence of clear flowering time differences, the safest approach might be to assume that gene flow will occur if native *E. globulus* forests neighbour *E. globulus* plantations. In this situation an assessment of the conservation significance of the native population could then guide management.

6.4 **Future issues**

An emerging issue in biodiversity conservation is that climate change may result in species becoming maladapted to the environmental conditions of their current range (Hughes 2003). Concerns that species may not be able to adapt *in situ* or migrate quickly enough to respond to such climate change has raised considerable debate as to whether assisted migrations may be required to ‘rescue’ species and ecosystems (Burbidge *et al.* 2011; Hewitt *et al.* 2011; Weeks *et al.* 2011). This debate is certainly ongoing for forest trees where natural long distance (intra-specific) gene flow may facilitate adaptation to new environments (Kremer *et al.* 2012). However, there are also suggestions that ‘assisted gene flow’ may be important for preventing maladaptation to the rapidly changing climate (Aitken and Whitlock 2013). Species that display adaptive variation across a broad geographic range are likely to have populations that would be more suited to future climates than others (Aitken and Whitlock 2013). Under an assisted gene flow scenario, individuals from the pre-adapted populations would be transferred to maladapted populations to improve that population’s adaptive capacity (Aitken and Whitlock 2013). Assisted gene flow is distinct from assisted migration in that individuals are moved within their extant range, whereas assisted migration involves moving species beyond their current range. Therefore, the risks associated with the two scenarios are in some ways analogous to those discussed above for inter- and intra-specific gene flow from
plantation forestry. It is likely that similar approaches to those used in this thesis, aimed at understanding barriers to reproduction, combined with simulation modelling would be useful for identifying the genetic risks of assisted gene flow and migration. However, if such large scale translocations of species are being considered as an adaptive response to climate change, then the risk of exotic gene flow may need to be re-considered (Aitken and Whitlock 2013; Weeks et al. 2011). Under such scenarios the whole concept of unwanted introgression may be relatively unimportant because exotic gene flow may be unavoidable, or even the aim of conservation.

6.5 Conclusion

This thesis has shown that the risk of exotic gene flow and subsequent introgression from *E. globulus* plantations to neighbouring native eucalypt species is low. The number of species at risk is lower than previously thought, and considerable barriers to introgression also extend to more high risk, closely related species. If these barriers are overcome, an assessment of conservation status and flowering synchrony and size compatibility will provide a good basis quantifying the risk posed by a particular plantation. The techniques developed in Chapter 3 will aid in the identification of high-risk situations and effective management action if gene flow is detected. There are several published genetic risk assessments (Barbour et al. 2008a; Barbour et al. 2008b; Byrne and Stone 2011; Byrne et al. 2011) that use decision tree approaches to integrate information on phylogenetic relatedness, pollen dispersal capacity, conservation status, flowering synchrony and compatibility, to determine a genetic risk rating. These protocols typically provide recommendations for plantation (or crop) establishment that range from “no management required”, to “monitoring”, “regular management required” or “do not plant or use alternative species” (e.g. Byrne et al. 2011). The results of this thesis will contribute significantly to the decision making process when assessing the genetic risk of new *E. globulus* plantations in Australia. It is also likely that integrating data developed in this thesis into simulation models would further enhance gene flow management, as it has in other systems forestry systems (DiFazio et al. 2012).
I believe that the risk of exotic gene flow should continue to be a consideration in native forests management, particularly where small threatened populations occur. However, I think that this genetic risk should be viewed in context with the array of other processes that threaten eucalypt forests and ecosystems in Australia (Lindenmayer et al. 2012). For example: fragmentation and isolation of forest remnants is likely to reduce genetic diversity and gene flow and diminish population sustainability (Byrne et al. 2008; Ellstrand and Elam 1993; Fischer and Lindenmayer 2007; Mimura et al. 2009); tree decline is causing the loss of eucalypt trees from native forests across large parts of Australia, particularly in dry agricultural landscapes (Bailey et al. 2012; Granger et al. 1994; Landsberg 1985; Landsberg 1988); invasive species are changing ecosystem dynamics and affecting biodiversity in eucalypt forests (Dawson et al. 1979; Weste 1986; Williams and Wardle 2005; Williams and Wardle 2007); and climate change is likely to interact with all these factors to affect the distribution of native forests and how they are managed in the future (Burbidge et al. 2011; Calder and Kirkpatrick 2008; Hughes 2003). Finally, there seems to be a contradiction between preventing gene flow from plantation forestry as a conservation mechanism, and advocating the large-scale translocation of species to improve genetic resilience in the face of climate change (Burbidge et al. 2011). The relative importance of exotic gene flow from plantation forestry may require re-evaluation if large scale assisted migrations become a reality, because what is now viewed as a negative impact may become a minor concern, or even a benefit (Aitken and Whitlock 2013). These issues are biologically, ethically and politically complex and will no doubt be the subject of much debate (Burbidge et al. 2011; Lindenmayer et al. 2010).
References


Ahuja MR (2011) Fate of transgenes in the forest tree genome. Tree Genetics & Genomes 7(2), 221-230.


Boland DJ, Brooker MIH, Turnbull JW (1980) 'Eucalyptus seed.' (CSIRO publishing: Canberra)


De Nettancourt D (2001) 'Incompatibility and Incongruity in Wild and Cultivated Plants.' (Springer: Berlin)

de Sousa VA, Pinto-Junior JE (1994) Stigmatic receptivity in the reproductive process of *Eucalyptus dunnii*. *IPEF Instituto de Pesquisas e Estudos Florestais* 47(44-47).


Dobzhansky T (1937) 'Genetics and the Origin of Species.' (Columbia University Press)

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. American Naturalist 74, 312-321.

Dodet M, Collet C (2012) When should exotic forest plantation tree species be considered as an invasive threat and how should we treat them? Biological Invasions 14(9), 1765-1778.


Euclid (2006) EUCLID Eucalypts of Australia 3rd edition. (Centre for Plant Biodiversity Research)


FPA (2009) Management of gene flow from plantation eucalypt species In 'Flora Technical Note 12.' (Forest Practices Authority: Hobart)


Hogenboom NG (1984) Incongruity: Non-functioning of intercellular and intracellular partner relationships through non-matching information. In


208


Lopez GA, Potts BM, Tilyard PA The performance of F1 hybrids between *E. ovata* and *E. globulus* ssp. *globulus* compared to inbred and outcrossed controls. In 'Hybrid Breeding and Genetics of Forest Trees', 9-14 Apr. 2000 2000b, Noosa, Queensland, Australia. (Eds HS Dungey, MJ Dieters and DG Nikles), pp. 457-461


Marchante E, Freitas H, Marchante H (2009) 'Guia prático para a identificação de plantas invasoras de Portugal continental.' (Imprensa da Univ. de Coimbra: Coimbra)


Mayr E (1963) 'Animal Species and Evolution.' (Harvard University Press: Massachusetts)


Muller H Isolating mechanisms, evolution and temperature. In 'Biological Symposium', 1942, pp. 71-125.


Nicolle D (2013) 'Native Eucalypts of South Australia.' (Lane Print & Post: Adelaide)


Nuñez MA, Medley KA (2011) Pine invasions: Climate predicts invasion success; something else predicts failure. *Diversity and Distributions* 17(4), 703-713.


Proceedings of the National Academy of Sciences of the United States of America 94(18), 9996-10001.

Petroli CD, Sansaloni CP, et al. (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. Plos One 7(9), e44684.


Pryor LD (1976) 'The Biology of Eucalypts.' (Edward Arnold: London)


Rathbone DA, McKinnon GE, Potts BM, Steane DA, Vaillancourt RE (2007) Microsatellite and cpDNA variation in island and mainland populations of a


Xu T, Hutchinson M (2011) ANUCLIM version 6.1 user guide. The Australian National University, Fenner School of Environment and Society, Canberra.


General Appendix: Copies of publications and posters arising from candidature