Assessing and managing pollen-mediated gene flow from locally-exotic Corymbia plantations

Myralyn Aguda Abasolo
Southern Cross University

Publication details
Copyright MA Abasolo 2013
Assessing and managing pollen-mediated gene flow from locally-exotic *Corymbia* plantations

Myralyn Aguda Abasolo (MSci)

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Southern Cross Plant Science
Southern Cross University
Lismore, NSW, Australia
(09 September 2013)
Thesis Declaration

I certify that the work presented in this thesis is, to the best of my knowledge and belief, original, except as acknowledged in the text, and that the material has not been submitted, either in whole or in part, for a degree at this or any other university.

I acknowledge that I have read and understood the University's rules, requirements, procedures and policy relating to my higher degree research award and to my thesis. I certify that I have complied with the rules, requirements, procedures and policy of the University (as they may be from time to time).

Print Name: MYRALYN ABASOLO

Signature: .................................................................

Date: .................................................................
Abstract

Assessing the risk of pollen-mediated gene flow from planted forests requires input on the taxonomy, biology and geography of the donor and recipient species. In subtropical Australia, the two situations of concern for gene flow from hardwood plantation forestry are the potential for infra-sectional gene flow due to the planting of *Corymbia citriodora subsp variegata* (Section *Maculatae*) (formerly *Eucalyptus maculata*), and, the potential for inter-sectional gene flow from the planting of the first generation hybrid of *Corymbia torelliana* (Section *Torelliana*) and Section *Maculatae*. This thesis addresses gaps in the knowledge of the genetics and flowering biology of the planted taxa, and methods for hybrid diagnosis.

I studied the degree of genetic determination in the timing of key floral development stages in *C. citriodora* subsp. *variegata*. Flowering observations were undertaken in a common garden on 1855 trees from eight regions during spring for three consecutive years, and monthly on a subset of 208 trees for 12 months. Peak anthesis time in translocated trees was stable over years and tended to be congruent with that recorded in native stands, suggesting strong genetic control of anthesis time. Two flowering races of *C. citriodora* subsp. *variegata* were identified; an early flowering northern race and a late flowering southern race. Overall, the level of flowering in the planted stand (age <10 years) was low (8 to 12% of assessed trees with open flowers in November), compared with flowering in nearby native stands.

I investigated the utility of seedling morphology for hybrid discrimination in three hybrid groups relevant to the monitoring of gene flow from plantings of *Corymbia* (L.D. Pryor & L.A.S. Johnson ex Brooker) taxa in subtropical Australia. Outbred F₁ hybrids between spotted gums (*C. citriodora* subsp. *variegata*, *C. citodora* subsp. *citriodora* and *C. henryi*) tended to more closely resemble their maternal *C. torelliana* parent and the most discriminating characters were the ratio of blade length to maximum perpendicular width, the presence or absence of a lignotuber, and specific leaf weight. Overall power to resolve among outbred F₁ hybrids from both parental taxa
was low to moderate. Advanced generation hybrids (outbred F\textsubscript{2} and outbred backcrosses) were more difficult to resolve reliably due to higher variances.

I also assessed the potential of near infrared spectroscopy for *Corymbia* hybrid determination. The NIR profiles of fresh foliage from eight-month-old seedlings and a handheld NIR instrument (980 - 1800 nm) had the highest mean assignment rates for the three hybrid groups which ranged from 76 to 90%. The assignment rates of the F\textsubscript{1} taxa were higher than those for parents at 100% and 72%, respectively. Hybrid resolution was even greater for 2\textsuperscript{nd} generation backcross hybrids. Age effects, and tissue handling, affected the NIR spectra of the seedlings and had implication for hybrid diagnosis in nursery situations.

The genetic and environmental factors influencing timing of anthesis and degree of synchrony among taxa provided in this thesis allow for refinement of the current risk assessment for planted *Corymbia*. The genetically controlled variation in peak flowering time between the two races of *C. citriodora* subsp. *variegata* and the lower fecundity of the plantations relative to the native stands moderates the otherwise high likelihood of infra-sectional gene flow. However, the impact of gene flow may be higher than what was previously thought because of potentially adaptive differences between the two races of *C. citriodora* subsp. *variegata*. 
Acknowledgment

I want to thank my supervisory team for their support and enthusiasm, Mervyn Shepherd, Carolyn Raymond and David Lee. My utmost gratitude to the CRC Forestry and IPRS of Southern Cross University for the financial support. I am very grateful to the following institutions: Forests NSW for access to their trials, the Department of Agriculture, Fisheries and Forestry Queensland for access to their trials, access to their nursery facilities, and for providing materials for the nursery experiment, and CSIRO for the use of their near infrared spectroscopy instrument. I also want to acknowledge the people who provided me technical support, to Roger Meder for his support in analysing my NIR data in Chapter 4, to Lyndon Brooks for his help on the statistical analysis in Chapter 2 and to Myrna Deseo for helpful comments in Chapter 4. I would also like to thank Troy Brown for the logistic arrangements for trial access in NSW, Terry Moody for facilitating access to the Bonalbo Common, Tracy Menzies and John Oostenbrink for the assistance with the nursery work in Gympie, Queensland and Gary Abblet for the assistance in fieldwork. I thank the journal editors and reviewers for their constructive comments on the results chapters.

My utmost gratitude to the staff and colleagues in Southern Cross Plant Science; to Graham King and Hans Wohlmuth for support in their capacity as administrators, to all the lovely people of SCPS it has been a pleasure knowing you all and sharing my PhD journey with you in one way or the other.

My friends and family had been very instrumental in my PhD journey. I thank my friends at the Southern Cross Presbyterian Church for their prayers, Peter, Vicki and David Roberts, John and Delma, Peter and Eden Caton, and Sophie Moore. I thank Mucha Mkono for her friendship, to Peter and Lexie Compton for opening their home to me, to Russell and Dion Virtue, for their love and support. To my lovely family, Willie, Aaron and Alyssa for the support and understanding that I have to be away to do this. Above all, to God be the honour and glory, who guided me through it all. Now, I am done.
# Table of Contents

Thesis Declaration ........................................................................................................... i  
Abstract ............................................................................................................................ ii  
Acknowledgment ............................................................................................................ iv  
Table of Contents ............................................................................................................ v  
List of Tables .................................................................................................................. ix  
List of Figures ................................................................................................................. xi  
Publications and other outputs from PhD candidature ................................................. xiii  
Contributions from others ........................................................................................... xv  

**Chapter 1. Introduction** ............................................................................................... 1  

1.1. Gene flow from translocated forest species in Australia and worldwide ................................................................. 1  
    1.1.1. Genetic and ecological impacts of translocation ................................................................. 1  
    1.1.2. Translocated forest species is a global issue ...................................................................... 3  
    1.1.3. Pollen mediated gene flow from Australian eucalypt plantations is well recognised ................................................................................................................................. 4  

1.2. Gene flow risk assessment of eucalypt plantations in Australia ............................................................................... 4  
    1.2.1. Gene flow risk assessment and research ......................................................................... 8  
    1.2.2. Legislative environment is affecting gene flow management in Australia ......................... 9  

1.3. Gene flow from *Corymbia* plantations in subtropical Australia .............................................................................. 10  
    1.3.1. Use of *Corymbia* in subtropical plantation forestry ............................................................. 10  
    1.3.2. Taxonomy and distribution of spotted gums (Section *Maculatae*)  
        and *Corymbia torelliana* (Section *Torellianae*) ..................................................................... 11  
    1.3.3. Potential for exotic gene flow to native *Corymbia* forest ........................................... 13  
    1.3.4. Current risk assessment for pollen-mediated gene flow in  
        *Corymbia* plantations ............................................................................................................. 14
1.4. The present study .............................................................................................................. 16
  1.4.1. Determination of the genetic variation in peak flowering and bud abundance in plantation CCV ......................................................... 16
  1.4.2. Techniques to assess gene flow .................................................................................. 18

Chapter 2. Genetic control of flowering in spotted gum, *Corymbia citriodora* subsp. *variegata* and *C. maculata* ................................................................................. 21

  2.1. Abstract ......................................................................................................................... 22
  2.2. Introduction .................................................................................................................. 23
  2.3. Methods ......................................................................................................................... 27
    2.3.1. Trial descriptions and germplasm material ......................................................... 27
    2.3.2. Flowering assessments ......................................................................................... 30
    2.3.3. Statistical analysis ............................................................................................... 32
  2.4. Results .......................................................................................................................... 39
    2.4.1. Timing, intensity, and synchronicity of anthesis in spotted gum ...................... 39
    2.4.2. The progression of flower development in spotted gum ..................................... 42
    2.4.3. Timing of anthesis in translocated populations is stable .................................... 44
    2.4.4. Testing for genetic control in the near-term potential to flower ....................... 45
    2.4.5. Tree vigour and flowering ...................................................................................... 48
  2.5. Discussion ...................................................................................................................... 52
    2.5.1. Two flowering races align with geographic and genetic groupings .................. 52
    2.5.2. Northern race flowers earlier .............................................................................. 53
    2.5.3. Northern flowering race is more fecund and disease resistant .......................... 53
    2.5.4. Genetic control of anthesis is indicated by stability in anthesis time of translocated trees ................................................................. 55
    2.5.5. Plastic regions include some initially below the threshold size for flowering .... 55
    2.5.6. Implications for risk assessment of gene flow from plantings of CCV .................. 57
  2.6. Acknowledgment .......................................................................................................... 59
Chapter 3. Identification of intersectional *Corymbia* hybrids based on seedling morphology improves with parental divergence

3.1. Abstract............................................................................................................. 61
3.2. Introduction ...................................................................................................... 62
3.3. Materials and methods...................................................................................... 66
   3.3.1. Materials.................................................................................................. 66
   3.3.2. Glasshouse methods and experimental design ........................................ 67
   3.3.3. Seedling assessment ................................................................................ 68
   3.3.4. Statistical analysis ................................................................................... 71
3.4. Results .............................................................................................................. 76
   3.4.1. Germination, survival and vigour of hybrid seedlings ............................ 76
   3.4.2. Univariate analysis to identify variables that discriminate hybrid taxa from their parents in each hybrid group ......................................................... 78
   3.4.3. Multivariate analyses .............................................................................. 83
   3.4.4. Assignment testing .................................................................................. 86
3.5. Discussion ......................................................................................................... 88
   3.5.1. Assignment and the use of morphology for hybrid identification .......... 88
   3.5.2. Genetic divergence and transgressive characters .................................... 93
3.6. Conclusion ........................................................................................................ 94
3.7. Acknowledgment .............................................................................................. 95

Chapter 4. Deviant near-infrared spectra allow identification of *Corymbia* hybrids

4.1. Abstract............................................................................................................. 97
4.2. Introduction ...................................................................................................... 98
4.3. Methods .......................................................................................................... 102
   4.3.1. Materials and experimental design........................................................ 102
   4.3.2. Collection of NIR spectra...................................................................... 103
   4.3.3. Outliers were removed before statistical analysis ............................... 104
   4.3.4. Statistical analysis ............................................................................... 106
   4.3.5. Comparison of the NIR spectra of different leaf tissue preparation, age and instruments.......................................................... 108
4.4. Results ............................................................................................................. 110
    4.4.1. Determination of a standard pre-treatment ............................................. 110
    4.4.2. Effect of leaf tissue storage ................................................................... 110
    4.4.3. Effect of leaf tissue age ......................................................................... 111
    4.4.4. Effect of spectral range ......................................................................... 114
    4.4.5. Classification power of eight-month-old fresh leaf tissues measured with NIR1 .............................................................. 114
    4.4.6. Diagnostic regions in the NIR spectra .................................................. 115
    4.4.7. Principal component analysis of fresh leaf tissue using NIR1 .............. 116
    4.4.8. PCA and assignment testing including the F_{1} and advanced generation BCT ................................................................. 118
    4.4.9. Exclusion analysis using two or four parental taxa in the calibration group ........................................................................... 119

4.5. Discussion ....................................................................................................... 120
    4.5.1. Utility and optimisation of NIR for *Corymbia* hybrid identification ................................................................. 120
    4.5.2. Recommended conditions for *Corymbia* hybrid diagnosis ................. 121
    4.5.3. Hybridisation causes deviation in NIR spectra of *Corymbia* .......... 122
    4.5.4. Application of NIR analysis in Scenario 1 – distinguishing intersectional F_{1} hybrids between CT and spotted gums ......................... 124
    4.5.5. Application of NIR in Scenario 2 – Plantations of F_{1} hybrids .......... 125
    4.5.6. Potential utility of NIR for *Corymbia* hybrid identification in a range of environments ................................................................. 126

4.6. Acknowledgements ........................................................................................ 127

Chapter 5. Conclusions/Recommendations .............................................................. 128

Appendices ................................................................................................................... 138

References .................................................................................................................... 148
List of Tables

Chapter 2

**Table 1.** List of *Corymbia citriodora* subsp *variegata* and *Corymbia maculata* provenances, grouped by regions, assessed for flowering at the Bonalbo trial. ................................................................................................................................. 29

**Table 2.** A summary of key floral development attributes for *C. c variegata* and *C. maculata* compiled from observations in native stands reported in literature and translocated materials from this study. .................. 33

**Table 3.** Pairwise synchrony index among- and within-provenances of spotted gums calculated from observations at the Bonalbo Trial in 2010-2011. ................................................................................................................................. 42

**Table 4.** Deviance Information Criterion (DIC) for the different binary logistic models used to determine the odds of trees with buds from each region in each year. ........................................................................................................... 47

Chapter 3

**Table 1.** List of the number of families, provenances and locations of *Corymbia* species and hybrids, number of seeds sown and seedlings assessed in this study. ................................................................. 70

**Table 2.** Characters measured on eight-month-old *Corymbia* seedlings with descriptions and transformations ........................................................................ 72

**Table 3.** Morphological characters used in the canonical variates analysis and the standardized coefficients (CV1 and CV2) from a discriminant analysis used to separate each taxon in the different spotted gum hybrid groups. ........................................................................ 74

**Table 4.** Assignment accuracy (%), and frequencies of transgressive characters in the F1 taxa for each of the three hybrid groups, CT-CCC, CT-CCV and CT-CH. ........................................................................................................... 87
Chapter 4

**Table 1.** Corymbia parental and hybrid taxa information for each data set used in NIR analysis. .......................................................... 105

**Table 2.** Summary statistics from principal component analysis of NIR1 spectra from eight-month-old fresh leaf tissue for the CH hybrid group comprised of the taxa *C. torelliana* (CT), *C. henryi* (CH) and the interspecific hybrid (F1H) for four spectral pre-treatments. ........................................... 111

**Table 3.** Summary statistics of PLS-DA of NIR spectra obtained from various tissue types at two NIR wavelength ranges. ...................................................... 112
List of Figures

Chapter 1

Figure 1. Gene flow risk assessment protocol. ................................................................. 6
Figure 2. Distribution of Corymbia species. ................................................................. 12

Chapter 2

Figure 1. Locations of Corymbia citriodora subsp variegata (CCV) and C. maculata (CM) provenances (grouped into regions) studied at the Bonalbo Trial. .................................................................................................................. 28
Figure 2. Reproductive structures of Corymbia species............................................ 31
Figure 3. Provenance level flowering intensities for a subset of 208 trees assessed monthly between September 2010 and November 2011. .................. 40
Figure 4. Progression of floral development throughout the year for three representative regions of CCV, Coastal QLD, Subcoastal QLD, and Richmond. ........................................................................................................... 43
Figure 5. Regional odds of trees having buds as assessed in November for three years in 2009, 2010 and 2011 at the Bonalbo trial......................... 49
Figure 6. a. Mean diameter at breast height (DBH) of each flowering effort category; b. Regional values for DBH (bars) plotted with mean flowering effort (lines). ................................................................................................. 50

Chapter 3

Figure 1. a. average percent germination on a family basis and b. average height on individual basis at age eight months of the spotted gum parental taxa and their hybrids.............................................................. 77
Figure 2. Leaf morphology of parents and hybrid taxa: Corymbia torelliana (CT), spotted gums – Corymbia citriodora subsp variegata (CCV; top row), Corymbia citriodora subsp citriodora (CCC; middle
row) and *Corymbia henryi* (CH; bottom row), and their hybrids, $F_1$
(CTxspotted gum; same order as the spotted gum column), $BCT$
(CTxCCVxCT), $BC_V$ (CTxCCVxCCV), $F_{2V}$ (CTxCCV) x (CTxCCV) ................. 80

**Figure 3.** Comparison of taxa within each hybrid group for
BLDL:MPW and SLW; a) CT-CH group ; b) CT-CCV group; c) CT-
CCC group. ..................................................................................................................... 83

**Figure 4.** A canonical variates analysis of CT, spotted gum,
interspecific hybrid, CCV backcross, CT backcross, CTxCCV $F_2$
using characters that discriminated among taxa in each hybrid group. ...................... 85

**Figure 5.** Canonical variate analysis of CCV and CH based on
thirteen characters that discriminate between taxa ......................................................... 86

**Chapter 4**

**Figure 1.** NIR1 average absorbance plot (after Gap-Segment and SNV
pre-treatment) of eight-month-old fresh tissue of the CH hybrid group................. 115

**Figure 2.** Principal component loadings plot of the first two
components derived from NIR1 (after Gap-Segment and SNV pre-
treatment) of eight-month-old fresh tissue of the CH hybrid group (*C.
torelliana, $F_{1H}$ hybrid and *C. henryi)*. ........................................................................... 116

**Figure 3.** Scores plot of the first two principal components derived
from NIR1 (after Gap-Segment and SNV pre-treatment) using eight-
month-old fresh tissue of the CH hybrid group. ........................................................... 117

**Figure 4.** Scores plot of the first two principal components derived
from NIR1 (after Gap-Segment and SNV pre-treatment) using eight-
month-old fresh tissue of the CCV hybrid group ........................................................... 118
Publications and other outputs from PhD candidature

Refereed publications


Abasolo M, Lee DJ, Brooks L, Raymond C and Shepherd M. Genetic control of flowering in Corymbia citriodora subsp. variegata. Submitted for publication to Australian Journal of Botany.

Co-authored publication

Shepherd M, Abasolo M and Lee DJ. Genetic and ecological constraints on hybridisation from translocated eucalypts in Northern New South Wales. Submitted.

Research Reports


Oral Presentations

Abasolo M, Lee DJ, Brooks L, Raymond C and Shepherd M. 2012. Flowering asynchrony between planted and native spotted gums is a barrier to gene flow. Speed talk at the Ecological Society of Australia Annual Conference. 3 December 2012. Melbourne, VIC. Australia.


Poster presentation

Abasolo M, Lee DJ, Brooks L, Raymond C, Shepherd M. 2012. Flowering in spotted gums is under genetic control. Poster presented at the CRC Forestry Annual Scientific Meeting. 6-8 March 2012. Mooloolaba, Sunshine Coast, QLD.

Contributions from others

I, Myralyn Abasolo, certify that this thesis is substantially my own work. Key contributions of others have been stated at the start of each experimental chapter. Acknowledgments are also provided on page iv of this thesis and at the end of each experimental chapter.

My principal supervisor, Mervyn Shepherd, provided extensive editing of each section of this thesis. Co-supervisors David J Lee and Carolyn Raymond, each provided invaluable feedback during the progression of the project with major contributions to the editing particularly in Chapter 4.

As Chapters 3, 4 and 5 have been published or submitted for publication, in accordance with the requirements of the Division of Research at Southern Cross University, a written statement regarding the proportion of contribution of Myralyn Abasolo is attached as Appendix 7.

PhD Candidate: ……………………………….. Date……………………………..

MYRALYN ABASOLO

Principal Supervisor: …………………………. Date……………………………..

MERVYN SHEPHERD
Chapter 1. Introduction

1.1. Gene flow from translocated forest species in Australia and worldwide

Gene flow induced by human activities may result in hybridisation between previously allopatric species (Ellstrand and Schierenbeck, 2000; Arnold, 2006; Ellstrand et al., 2010). Gene flow is defined as “the movement and the incorporation of genes from one or more populations into the gene pool of another population” (Slatkin, 1985; Futuyama, 1998). The massive translocation of plant species throughout the world for agriculture, forestry or ornamental purposes creates an opportunity for gene flow from the translocated to the native populations through spontaneous hybridisation (Rhymer and Simberloff, 1996; Wolf et al., 2001; Ellstrand et al., 2010). The reproductive barriers of closely related species or populations are often weak when geographic barriers to hybridisation are removed (Ledig, 1992; Potts et al., 2003). Interspecific hybrids (hybrids between closely related species) and intraspecific hybrids (hybrids between genetically distinct populations) are often observed in translocation sites (reviewed in Ellstrand and Schierenbeck, 2000; Schierenbeck and Ellstrand, 2009). Pollen dispersal followed by hybridisation may result in the introgression or permanent incorporation of exotic genes from translocated species to the native species (Abbott, 1992; Rieseberg and Wendel, 1993; Rhymer and Simberloff, 1996; Ellstrand et al., 1999; Vilà et al., 2000). The introgression of genes from translocated species with the closely related species or genetically distinct population may have harmful genetic and ecological consequences (Potts et al., 2003; Laikre et al., 2010; Byrne et al., 2011).

1.1.1. Genetic and ecological impacts of translocation

There are four broad categories of genetic and ecological impacts that are identified as a result of introgression of genes from translocated species (Nagy, 1997; Lesica and Allendorf, 1999; Hufford and Mazer, 2003; McKay et al., 2005; Broadhurst et al., 2008; Huff et al., 2011). First, there may be loss of genetic variation through the increase in the frequency of introduced alleles and subsequent loss of locally adapted
alleles (termed genetic swamping). Genetic swamping is more likely to occur in small populations (Allendorf et al., 2001; Hufford and Mazer, 2003; Laikre et al., 2010). Although there is no production of hybrid progenies, or the hybrids are eliminated by outbreeding depression (see below), pollen swamping can reduce the size of native population through wasted reproduction (Keim et al., 1989; Rieseberg et al., 2007; Whitney et al., 2006; Rhymer and Simberloff, 1996). At the other end of the spectrum, when invasive hybrids establish and reproduce, the genetic composition and the effective population size of the native population may be reduced (Allendorf et al., 2001).

Second, there may be a loss of fitness through outbreeding depression (hybrid breakdown). Outbreeding depression may be due to both intrinsic and extrinsic factors. Intrinsic outbreeding depression is due to mating among genotypes with genetic incompatibilities leading to low hybrid fitness (Hufford and Mazer, 2003; Laikre et al., 2010). Extrinsic outbreeding depression is due to mating of species or ecotypes each of which is adapted to a distinct environment and the resulting hybrids may perform poorly in a specific environment (Frankham et al., 2002).

Third, there may be displacement of the native species particularly if the exotic alleles can establish and hybrid progenies invade native habitat. This will result in the genetic composition of the native populations being altered (Anttila et al., 2000; Vilà et al., 2000; Olden et al., 2004). Fourth, there may be a change in the population genetic structure if genetically distinct populations hybridise (Olden et al., 2004). Changing the population genetic structure through hybridisation may narrow gene pools of native populations (Rieseberg et al., 1989; Ellstrand and Elam, 1993; Schemske et al., 1994; Levin et al., 1996). Hybrid grasses, shrubs and trees which are products of spontaneous hybridisation between the translocated species and the native species were also documented to be invading native habitats (Ellstrand and Schierenbeck, 2000; Schierenbeck and Ellstrand, 2009).

Despite the negative effects, hybridisation can also have positive effects such as increasing the genetic variation especially in rare species and in fragmented/remnant populations. This has been demonstrated in both animal and plant species whereby, introduction of alleles from exotic populations increased the genetic variation of the
Depauperate population (i.e. Burke and Arnold 2001; Ebert et al., 2001; Ingvarsson 2003; Tallmon et al., 2005; Vila et al., 2002; Duputie et al., 2007). Restored populations may benefit from the increase in genetic variation by increasing their ability to cope with climate change, and their ability to survive in general (Field 2008; Edmands 2007; Allendorf et al., 2001).

1.1.2. Translocated forest species is a global issue

In the northern hemisphere, studies on gene flow from forestry plantations are limited and most relate to concerns that transgenes might be transferred from genetically modified trees to native trees (DiFazio et al., 2004; Brunner et al., 2007; DiFazio et al., 2012). There is a rigid assessment of regulatory issues before large scale or commercial deployment of genetically modified crops (Brunner et al., 2007; Strauss et al., 2010) but such assessment is lacking for exotic species (Laikre et al., 2010). The more important issue at present is the monitoring of large scale translocation of locally exotic trees, especially where remnant or small native populations may be affected, because the genetic integrity of the native populations may be compromised through introgression (Rieseberg and Wendel 1993; Ellstrand 1992). In the northern hemisphere, poplars are the most studied species in terms of hybridisation between exotic and the native species. Hybrid poplars in Europe are colonising novel habitats (Benetka et al., 1999; Vanden Broeck et al., 2003) and narrowing native poplar gene pools (Cagelli and Lefèvre, 1995; Vanden Broeck et al., 2005; Smulders et al., 2008). However, gene flow monitoring of these translocated poplar plantations is limited (but see Talbot et al., 2012). Lack of gene flow monitoring limits the understanding of the degree of both likelihood and impact of gene flow. Gene flow, to a certain extent, may compromise the genetic diversity of the populations (Laikre et al., 2010).
1.1.3. Pollen mediated gene flow from Australian eucalypt plantations is well recognised

In Australia, eucalypts dominate forest communities (Wardell-Johnson et al., 1997). Hybridisation and introgression are natural processes in eucalypts with many species being interfertile (Pryor and Johnson, 1971; Pryor, 1976; Griffin et al., 1988). Indeed, the natural process of hybridisation shaped the evolution of eucalypts (McKinnon et al., 2003). Of the 528 species examined in a comprehensive study on natural hybridisation among eucalypts, 289 (55%) species were observed to be involved in at least one hybrid combination (Griffin et al., 1988). Natural hybridisation also appears to decline with increasing taxonomic distance between parents (i.e. inter-subgeneric < inter-sectional < inter-series < intra-series; Potts and Wiltshire 1997; Potts et al., 2003). Various authors proposed different taxonomic classification of eucalypts (i.e. Pryor and Johnson 1971; Hill and Johnson 1995; Brooker 2000; Ladiges and Udovicic 2000) but the higher order taxonomy has not changed substantially, except for Corymbia. Because of their propensity to hybridise, however, the planting of species or provenances of eucalypts in places where they do not occur naturally (locally-exotic) may allow the transfer of genes through pollen into adjacent native forests at levels that would not occur naturally, here referred to as pollen-mediated gene flow of locally-exotic species (Barbour et al., 2002; Barbour et al., 2003; Hingston et al., 2003; Barbour et al., 2005; Barbour et al., 2006; Barbour et al., 2008). In Australia, the establishment of eucalypt plantations on a broad scale, which is now estimated to cover an area of 980,000 hectares (Gavran, 2012) may result in pollen-mediated gene flow from the plantations into native eucalypt populations (Potts et al., 2003; Kanowski et al., 2005; Grimbacher, 2011).

1.2. Gene flow risk assessment of eucalypt plantations in Australia

The recognition of the potential for gene flow from the locally-exotic plantations has resulted in gene flow risk assessments being undertaken, particularly in temperate Australia where locally-exotic species or races of eucalypts have been planted extensively. Systems for assessing gene flow risk in perennial species like acacias for revegetation in a range of scenarios (Millar et al., 2012; Millar 2008) and gene flow risk
from exotic eucalypt plantations (Potts et al., 2003) have been formulated. Monitoring systems founded on well-established genetic principles have been developed and are highly transferable across forest species (Potts et al., 2003; Byrne et al., 2011). Gene flow risk assessments have been carried out for temperate eucalypts such as *Eucalyptus globulus* and *E. nitens* in Tasmania (Barbour et al., 2002; Barbour et al., 2003; Barbour et al., 2006; Barbour et al., 2010) and a range of *Eucalyptus* species and hybrids in Victoria (Hingston et al., 2003) as well as for *Corymbia* plantations in the subtropics (Barbour et al., 2008).

Gene flow risk assessment allows the identification and characterisation of the likelihood and impact of gene flow (Potts et al., 2003; Byrne et al., 2011). As the process of gene flow is characterised by several steps (Potts et al., 2003; Barbour et al., 2008), it is possible to identify a failure in one step which may prevent gene flow. A risk assessment based on a decision tree allows evaluation of risk using existing information and knowledge (Byrne et al., 2011). In this thesis, three broad criteria, and factors that may influence these criteria are presented (Figure 1). It also allows the identification of knowledge gaps and for more rigorous evaluation of the risk (Byrne et al., 2011).

**Taxonomic and genetic divergence criteria** seek to establish if the planted taxon can serve as a source of foreign genes through hybridisation (Byrne et al., 2011). The taxonomy is a good indicator of the likelihood of gene flow because barriers to hybridisation increase between more distantly related species (Griffin et al 1988; Potts et al. 2003).

The risk of gene flow is substantial when species are more genetically divergent because the transfer of mal-adapted alleles from locally-exotic plantings might have both ecological and evolutionary consequences. Ecological consequences include negative effects on the re-establishment and long term persistence of the species while evolutionary consequences include species isolation, divergence and introgression (Ellstrand and Elam 1993; Vilà et al., 2000; Ellstrand 2003). Based on the taxonomic criteria, patchy or disjunct distribution may indicate possible genetic divergence among geographically distinct populations because of isolation and genetic drift (Byrne et al.,
2011). Genetic divergence can be detected through genetic studies and can be indicated by the presence of subspecies, variants, landraces and by trait variation.

**Biological criteria** determine the potential success of pollen transfer from the translocated species to the stigmas of the native species. The criteria also determine the degree of hybrid vigour if pollen transfer is successful (Byrne et al., 2011; Figure 1). Various factors such as flowering abundance, flowering synchrony, and post-zygotic barriers such as embryo development, seed maturation, seed survival, germination and hybrid growth to maturity (reviewed in Potts et al. 2003; Barbour et al. 2008) all play a role on the level of likelihood and impact of gene flow. The likelihood of gene flow may be low if pre-mating barriers such as pollen production in the translocated species
is low and/or there is flowering asynchrony between the native and the translocated species (Potts et al., 2003; Byrne et al., 2011). For example, there was decreased pollen penetration down the stigma as the taxonomic distance between mothers and pollen parents increased (Ellis et al., 1990). In Corymbia, it has been found that there are biological pre-zygotic barriers to hybridisation such as pollen adhesion to stigma, pollen germination, pollen tube growth and pollen penetration of the ovule (Dickinson et al., 2012). Other pre-zygotic factors such as the direction of the cross, (i.e. unilateral incompatibility; de Nettancourt 1997), have also been found to be a barrier to hybridisation. For example, the pollen tube small flowered E. nitens cannot reach the ovaries of the large flowered E. globulus, preventing crossing while the opposite is viable (Gore and Potts, 1995). In a natural population of E. rubida which has larger flowers and E. aggregata with smaller flowers, crossing towards E. aggregata was more frequent (Field et al., 2010). Therefore, the likelihood of gene flow decreases with increasing taxonomic distance between the translocated and the native species.

The impact of gene flow may also be low if introgression fails, i.e. if the hybrids are not fit enough or otherwise do not reach maturity and therefore are unable to backcross with the associated native species (Potts et al., 2003; Byrne et al., 2011; Figure 1). At the other end of the spectrum, if the post-mating barriers are weak (i.e. the hybrid progenies exhibit heterosis), hybrids may establish in the native populations and if the hybrids synchronise flowering with the native population, repeated backcrossing may have high impact (Ellstrand et al., 1999). This impact is particularly relevant in remnant populations where there may be relatively fewer pure bred native species compared to the large scale plantings of exotic species. The remnant native species may lose genetic identity by assimilation of the foreign genome through genetic swamping (Haygood et al., 2003; Ellstrand et al., 1999).

**Geographic criteria** are indicators of how pollen is dispersed within the landscape (Byrne et al., 2011; Potts et al., 2003). Distance of the planting site from the native stands, patterns of pollen flow and pollinators, and the existence of a requirement to protect the native species, are all considerations under the geographic criteria. The likelihood of gene flow may be low if the native stands are not within the pollen dispersal distance of the plantation (Potts et al., 2003; Byrne et al., 2011).
The impact of gene flow might be lower in larger populations and in a widespread species and impact may be greater in small, remnant populations and/or in endangered species (Byrne et al., 2011; Potts et al., 2003). Gene flow from plantations to rare species and/or remnant populations speeds up the time to extinction (Wolf et al. 2001). This is because of the reduction in the number of pure-bred genotypes that counteracts gene flow resulting to the dilution of gene pool through introgression by the more abundant species or the larger population (Levin et al. 1996). In particular, extinction of rare species through gene flow can be brought about by demographic swamping. The production of unfit hybrids reduces seed set and seedling recruitment of the pure-bred genotypes of the rare species (Levin 1996). If hybrids are fertile, genetic swamping may occur in which the genetic composition of the rare species is assimilated by the more common congeners (Ellstrand and Elam 1993; Levin et. al., 1996). Empirical studies found that gene flow by pollen (hybridisation) has been the major cause of the extinction in a number of rare plant species (reviewed by Levin et al. 1996; Rhymer and Simberloff 1996; Rieseberg and Carney 1998).

1.2.1. Gene flow risk assessment and research

Gene flow risk assessment relies on research examining factors that influence the likelihood and impact of gene flow. For example, the initial concern of broad scale gene flow from locally-exotic eucalypts in Tasmania appears to be reduced as further biological information is gathered on factors that influence gene flow risk (Barbour et al., 2002; Barbour et al., 2003; Barbour et al., 2005; Barbour et al., 2006; Barbour et al., 2010). By evaluating the factors on a site by site and species by species basis, specific sites and species are being recognised to be potentially at risk, particularly the rarer species (Barbour et al., 2010). It has also been found that hybridisation rates between the native and the exotic species increase with increasing flowering synchrony (Barbour et al., 2002; Barbour et al., 2003). There are other factors that may reduce the likelihood and impact of gene flow as indicated above (Section 1.1.3 and Figure 1).
1.2.2. Legislative environment is affecting gene flow management in Australia

The development of strategies to minimise the likelihood and impacts of gene flow are among the requirements necessary for complying with biodiversity legislation and forest certification. Biodiversity legislation is aimed at maintaining diversity at the ecosystem, population, individual and genetic levels. The management of gene flow in production forests is one of the ways to conserve genetic diversity outlined in the Federal Environment Protection and Biodiversity Conservation Act 1999.

In NSW and QLD, there are three areas of legislation that impact gene flow from plantations. The legislation 1) provide for the conservation and management of protected areas including national parks, native vegetation, (threatened or vulnerable) and threatened ecological communities; 2) provide for the ecologically sustainable forest management, including silviculture, of all State-owned forests; and 3) aim to protect native vegetation (i.e. remnant stands in adjacent plantation and the eradication/control of noxious weeds within the property, private or public. Research on gene flow management from plantations may be used to guide management decisions for policy legislation and standards. Assessing the risk of gene flow is also a way of acquiring forest certification and thus access to the wider market for forest products (Legros, 2001; Rametsteiner and Simula, 2003). In Australia, there are two certification schemes and both have clauses that deal with germplasm containment and prevention of hybrid establishment in native forests (Australian Forestry Standard, 2007a, b; Forest Stewardship Council, 2011). 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. Gene flow from *Corymbia* plantations in subtropical Australia

1.3.1. Use of *Corymbia* in subtropical plantation forestry

Hardwood *Corymbia* plantations in subtropical Australia have been established for solid wood production (Lee et al., 2010) and potentially for pulp production (Brawner et al., 2012). Among the *Corymbia*, the major plantation species are the spotted gums (Section *Maculatae*) and in particular, *Corymbia citriodora* subsp. *variegata* (CCV) is planted extensively (Lee, 2007; Lee et al., 2009; Lee et al., 2010; Nichols et al., 2010). In recent years, the seed sources of commercial spotted gum plantations were largely based on a few select provenances of CCV from the northern end of the range of CCV, particularly Woondum, near Gympie in Queensland (QLD). This provenance has a high tolerance to *Quambalaria* shoot blight (Johnson et al., 2009; Brawner et al., 2011), a fungal disease in which the fungi attack new shoots and expanding leaves leading to a loss of apical dominance in severe cases (Pegg et al., 2011). Due to the susceptibility of the pure spotted gum taxa to shoot blight, there was an increased interest in the intersectional hybrid between spotted gum and *Corymbia torelliana* (Section *Torellianae*; CT) for forestry. The hybrid is more tolerant of the fungi that attack the pure spotted gum taxa (Lee, 2007). In addition, the hybrid possesses more desirable characters that are important to forestry than the pure spotted gum taxa. The hybrid grows more vigorously, it is more frost tolerant and has higher rooting ability (i.e. can be propagated by cuttings) compared to the pure spotted gum taxa (Lee, 2007). Small scale plantings of the *Corymbia* F$_1$ hybrid have been established in southern QLD and experimental plantings have been established in southern QLD and northern New South Wales (NSW), (Barbour et al., 2008).
1.3.2. Taxonomy and distribution of spotted gums (Section *Maculatae*) and *Corymbia torelliana* (Section *Torellianae*)

Spotted gums are a complex of closely related taxa that occur naturally along the eastern coast of Australia from around latitude 16°S in north QLD in the north to around 37°S in eastern Victoria in the south (Hill and Johnson, 1995; Figure 2). There are four taxa that are recognised based on morphology, leaf oils (Hill and Johnson, 1995), and ribosomal DNA (Parra-O. et al., 2009). Three of the four taxa occur as a latitudinal replacement series (one taxon replacing another).

*Corymbia citriodora* subsp. *citriodora* (CCC) is distributed from south-west of Cooktown to the Maryborough district, and west to the Great Dividing Range west of Springsure in QLD, with a disjunction of 300 km separating the northern and southern populations (Hill and Johnson, 1995). *Corymbia citriodora* subsp. *variegata* has a wide range of distribution from the Carnarvon and Dawes Ranges north of Monto in QLD, contracting southwards to sub-coastal regions around the upper Nymboida River and north-west of Coffs Harbour in NSW (Hill and Johnson, 1995; Shepherd et al., 2012). Intergrades and hybrids of CCC and CCV may occur in the north-east of CCV’s range.

*Corymbia henryi* (CH) tends to occur on less fertile low lying country from Brisbane in QLD in the north to near Glenreagh south of Grafton, in the south. Hybrids or intergrades of CCV and CH are expected where the species occur in sympatry (Ochieng et al., 2010). *Corymbia maculata* (CM) occurs mainly along the coast of NSW from the Manning River valley in the north to near Bega in the south, with outlying occurrences near Nowa in eastern Victoria (Hill and Johnson, 1995).

Studies of the genetic structure in the spotted gum complex have shown that genetic diversity within the three northern taxa (CCC, CCV and CH) is distributed on a geographic rather than a taxonomic basis, and that they are genetically distinct from the southernmost taxon, CM (Shepherd et al., 2008; Shepherd et al., 2012). Three populations were evident within the northern taxa: first, a northern population
composed of CCC material from north of the major disjunction in central QLD, second, a central population comprised of CCC, CCV and CH material from below this disjunction but north of the Border Ranges and third, a southern population comprised of CCV and CH from predominantly south of the Border Ranges (Shepherd et al., 2008). The genetic structuring of CM based on microsatellites also showed two groups with a northern population distinct from a southern population (Shepherd et al., 2012).
Adaptive differences appear to be prevalent among races and among provenances of eucalypts. Racial differences in adaptive characters like leaf size and bark thickness were observed in *Eucalyptus globulus* (Steane *et al*., 2006), provenance difference in stem diameter in *E. marginata, E. globulus* and *E. nitens* (Muneri and Raymond, 2000; Kube *et al*., 2001; O'Brien *et al*., 2007) and provenance difference in drought responses in *E. microtheca* (Li and Wang, 2003). Evidence is also accumulating that there are potentially adaptive differences between populations of spotted gums. Within CCV, regional differences were observed in leaf mass area (Ochieng, 2009), disease resistance (Johnson *et al*., 2009; Pegg *et al*., 2011), seedling leaf oils, (Asante *et al*., 2001) and frost tolerance (Larmour *et al*., 2000).

*Corymbia torelliana* (Section *Torelliana*; hereafter called CT) is the closest relative of the spotted gums (Hill and Johnson, 1995; Shepherd *et al*., 2008). *Corymbia torelliana* naturally occurs in the margins of tropical rainforests, in a narrow belt of some 50-80 km wide on coastal foothills and adjacent ranges of tropical north QLD from near Cooktown in the north to west of Ingham in the south (Brooker and Kleinig, 1994; Boland *et al*., 2006; Hill and Johnson 1995). *Corymbia torelliana* may occur in sympatry with the northern population of CCC around Cooktown in QLD but these two species may occupy different habitats (Brooker and Kleinig, 1994; Boland *et al*., 2006; Hill and Johnson 1995).

### 1.3.3. Potential for exotic gene flow to native *Corymbia* forest

A risk assessment of *Corymbia* plantings by Barbour *et al.* (2008) noted three main factors that may be important in considering exotic gene flow to native *Corymbia* forests. First, interspecific hybridisation within the genus is frequently observed in nature, in regions of sympatry (Hill and Johnson, 1995). Second, there may be long distance pollen dispersal, because some of the known pollen vectors of *Corymbia* like flying foxes can travel hundreds of kilometres (Eby, 1991; Southerton *et al*., 2004; Eby and Law, 2008). And third, most of the planted areas are established among native populations of the same species, or closely related species which will likely have weak barriers to hybridisation (Barbour *et al*., 2008).
There are two main situations of concern with gene flow from planting of *Corymbia* in subtropical Australia. One situation is the planting of locally-exotic spotted gum provenances. For example, northern provenances of CCV from around Gympie in QLD are widely planted in northern NSW where both CH and genetically distinct populations of CCV occur (Shepherd *et al.*, 2008a). Natural interspecific hybridisation between CCV and CH occurs in regions of sympatry, and northern and southern provenances of CCV are inter-fertile and have flowering synchrony to some extent (Hill and Johnson, 1995; Ochieng *et al.*, 2008; Ochieng *et al.*, 2010). *C. henryi* and CCV have distinct morphology and although they co-occur at some locations, they tend to occupy different positions in the landscape, suggesting adaptive differentiation (Ochieng *et al.*, 2010). This means that if gene flow occurs from locally-exotic provenances of CCV to native provenances in the south of the species range, non-adapted genes may be introduced in local CCV or CH populations and other *Corymbia* species.

The second situation of concern with planting *Corymbia* in subtropical Australia is the planting of the first generation intersectional hybrids between CT and spotted gum (F₁). The planting of the Corymbia F₁ hybrid generates a potential for backcrossing to native populations of spotted gums or other *Corymbia* species, and the hybrid may transfer weedy attributes from CT. *Corymbia torelliana* is planted widely throughout the subtropics where it is regarded as an environmental weed in some shires of northern NSW and southern QLD (Hill and Johnson, 1995; NCWAC, 2003; Kingston *et al.*, 2004). *Corymbia torelliana* naturally hybridises with spotted gums in north QLD and can be successfully crossed with spotted gums in controlled pollinations (Hill and Johnson, 1995; Dickinson *et al.*, 2010). Introgression of CT into the local gene pool may be undesirable because the same attributes of adaptability and vigour, that make the hybrids attractive for plantations, may also contribute to the ability of the hybrids to invade novel and disturbed environments (Lee, 2007; Lee *et al.*, 2009; Abasolo *et al.*, 2012; Shepherd *et al.*, submitted).

### 1.3.4. Current risk assessment for pollen-mediated gene flow in *Corymbia* plantations

14
The current gene flow risk assessment of *Corymbia* plantings was based on taxaspecific information available in 2007. Knowledge gaps were identified and an assessment of likelihood and impact of gene flow formed on the basis of available data (Barbour *et al.*, 2008). In the instance of gene flow from planting of locally-exotic spotted gum provenances, the likelihood of pollen-mediated gene flow was assessed to be high. Taxonomically, the locally-exotic and the native spotted gums belong to the same species therefore there is high likelihood that there are few crossing incompatibilities. *Corymbia citriodora* subsp. *variegata* was known to have a broad flowering period therefore some flowering synchrony with native CCV and other *Corymbia* species is likely. Because of the variety of generalist pollinators which can travel both short and long distances (Eby, 1991; Southerton *et al.*, 2004; Eby and Law, 2008), the likelihood of long distance pollen dispersal was high. The plantation sites also abutt native spotted gums and other *Corymbia* species. However, as the species concerned are widespread and abundant, the impact of any gene flow was considered to be low and the overall risk of gene flow moderate.

The risk of planting the *Corymbia* F₁ hybrid, on the other hand, was considered high. This was in part because of the scarcity of information on crossing incompatibilities and biological information, particularly flowering synchrony between the intersectional hybrid F₁ and the native taxa, at the time of assessment. The impact of gene flow was also considered high because of the potential to transfer the weedy characters of the CT parent (Barbour *et al.*, 2008).

The review also identified knowledge gaps and the research requirements necessary to further refine the risk assessment of *Corymbia* plantations. Among the research priorities that were listed were the need to assess the degree of flowering synchrony between *Corymbia* plantations and native populations, and to develop techniques for quantifying and monitoring for gene flow from *Corymbia* plantations (Barbour *et al.*, 2008). These two priority research areas are the subject of this thesis.
1.4. The present study

1.4.1. Determination of the genetic variation in peak flowering and bud abundance in plantation CCV

One of the key opportunities for managing risk is the deployment of planting stock that has genetically controlled asynchrony in flowering time with native stands as temporal isolation is an important barrier to gene flow among populations and species of eucalypts (Griffin, 1982; Potts et al., 2003). Several studies have shown that differences in flower opening time (anthesis) between taxa, provenances and families of eucalypts are under moderate to strong genetic control (Volker et al., 1988; Mullin and Pswarayi, 1990; Gore and Potts, 1995; Hayden, 1995; Wiltshire et al., 1998; Apiolaza et al., 2001; Jones et al., 2011) including differences among species of spotted gums (McDonald, 2004).

Genetically controlled variation in flowering time can reduce flowering synchrony among genotypes, races or provenances. The genetic variation in flowering time throughout the year among races of *E. globulus* resulted in flowering asynchrony between early and late flowering races (Jones et al., 2011). Similarly, races of *E. globulus* were separated in peak flowering times by as much as eight months (Gore and Potts, 1995). Mating among trees from the same provenance in a seed orchard of *E. regnans* also occurred more often than mating among trees from different provenances due to differences in peak flowering time (Burczyk et al., 2002). It is clear from the above studies that genetic variation in flowering traits can moderate gene flow from plantings by temporal partitioning in flowering time of planted and native populations of eucalypts.

In *Corymbia*, most surveys of flowering time to date have been of native stands where genetic differences are confounded by environmental factors. Peak flowering is usually during winter (D. Lee, pers com.) but summer flowering is also possible. In the Brisbane area in QLD, peak flowering of CCV was observed from November to February (Specht and Brouwer, 1975). In western QLD, peak flowering occurred from October to January (Dale and Hawkins, 1983) whereas, flowering was noted to peak
between January and March on the north coast of NSW (Law et al., 2000). These observations from the north coast of NSW were largely in accord with records from bee keepers in upper Northeast NSW where spotted gum was noted to flower between December and March (Birtchnell and Gibson, 2008; Eby and Law, 2008).

A detailed study of floral phenology in the closely related taxa, *C. maculata* (in the paper it is *E. maculata*) was undertaken on several trees on the South Coast of NSW. *Corymbia maculata* takes 14 to 18 months for bud initials to develop into a flower (Pook, 1984; Pook et al., 1997). Two peaks of flower opening may occur per year, one in winter and one in summer (Specht and Brouwer, 1975). The winter peak is smaller than the summer peak however, because when temperature drops below 16°C, photosynthates are concentrated on stems and roots thereby limiting shoot development (Specht and Brouwer, 1975).

Differences in the time of flowering between native *C. maculata* stands have also been recorded. Flowering follows a north-south direction such that populations in the north around Kempsey (now regarded as CCV or intergrade populations; Shepherd et al., 2012) begin flowering in the December/January bi-month period, in February/March in the lower-north east region (Lower Hunter) and in April/May in the South East region near Batemans Bay (Eby, 1991).

Flower abundance also influences the direction of gene flow and determines the relative strength of the plantation or the native population as a pollen source. A population with abundant flowers and therefore more likely to be pollen source, whereas a population with fewer flowers tends to become the pollen sink (Field et al., 2010; Millar et al., 2012). Flower abundance is highly influenced by the environment (Eldridge et al., 1994; Moncur and Hasan, 1994) and was found to have low heritability (McGowen, 2007). It has also been found that wide spacing in plantations promotes flowering (Williams et al., 2006) and trees at the edge of the plantation are more reproductively active than trees at the centre (Barbour et al., 2008) because flowering intensity decreases with intense competition after canopy closure (Potts et al., 2003).

The influence of tree size has not been fully studied in a plantation situation but it is likely that larger (and older) trees will have higher flowering intensity than smaller (and
younger) trees as was observed in native populations of *Eucalyptus* and *Corymbia* (Ashton, 1975; Law *et al.*, 2000; Bacles *et al.*, 2009). If tree size influences flower abundance, then older plantations may have higher flowering intensity and propensity than younger plantations. Therefore, older plantations may represent a higher gene flow risk than younger plantations, because older plantations have greater potential to act as pollen source.

### 1.4.2. Techniques to assess gene flow

Long term monitoring of gene flow involves periodic assessment of the rates of hybridisation, a process that is increasingly integral to risk management of planting of trees for forestry or restoration purposes (Potts *et al.*, 2003; Barbour *et al.*, 2010; Byrne *et al.*, 2011). Assessing impact of gene flow using biological criteria is a function of the rate of hybridisation and the rate of hybrid establishment (Potts *et al.*, 2003; Byrne *et al.*, 2011). Often, these parameters must be initially assumed, and management strategies are adaptive, putting in place monitoring processes, and utilising information to inform management responses iteratively once evidence on hybridisation rates comes to hand (B Potts Pers. Comm.). Directly measuring hybridisation rates has confirmed expectations, for example, that hybrid frequency tends to increase in populations with high flowering synchrony, and that pollen tends to move from a more abundant pollen source to populations with less pollen, i.e. fewer flowering trees, (Field *et al.*, 2008; Lepais *et al.*, 2009; Field *et al.*, 2011).

Despite the observation that even low levels of gene flow from exotic plantations may constitute a risk (Barbour *et al.*, 2008), there has been little assessment of the scale and impact of gene flow, and little monitoring continues to be undertaken (Laikre *et al.*, 2010). This, in part, may be due to the difficulty and costs associated with the screening of large numbers of individuals needed to detect low levels of gene flow and often a lack of cost-effective reliable hybrid detection tools (Barbour *et al.*, 2008).

In temperate Australia, a system based on the visual identification of hybrids using morphological differences has been effective in large-scale screening to detect hybrid eucalypt seedlings and provide direct measurements of gene flow (Potts and Reid,
The use of morphological markers has the advantage above other methods in these situations because many thousands of individuals may be screened easily and cheaply to detect even relatively low levels of gene flow provided at least a few reliable discriminating characters can be identified (Barbour et al., 2002; Barbour et al., 2003). In eucalypts, the F₁ hybrids may easily be identified using morphology but advanced generation hybrids may be difficult to identify because of trait segregation (Field et al. 2009). In the case of planting F₁ hybrids, there is a need to identify the advanced generation hybrids.

A promising method to overcome the difficulty of large scale screening of advanced generation hybrids is the use of near infrared spectroscopy (NIR). Near infrared spectroscopy has been widely used to predict chemical compositions in agricultural crops and to predict the wood chemistry of trees (e.g. Jones et al., 2006; Poke and Raymond, 2006; Maranan and Laborie, 2008; Meder and Meglen, 2012). NIR has also been shown to have potential to evaluate the nutritional value of forests in ecological studies (Foley et al., 1998; Stolter et al., 2006; Meder et al., 2007) and for selection of wood quality in tree breeding programmes (Schimleck, 2007; Meder et al., 2011; Meder and Schimleck, 2011). To a lesser extent, NIR has also been used for hybrid identification in trees with studies restricted to Eucalyptus and Betula (Atkinson et al., 1997; Humphreys et al., 2008). The main value of near infrared spectroscopy in hybrid screening is where it is difficult to identify hybrids by morphology (Tibbits, 1988; Abasolo et al., 2012) but it may offer other advantages such as reduced laboratory time and non-destructive sampling relative to other methods like oil analysis (Foley et al., 1998; Ebbers et al., 2002; Humphreys et al., 2008).

In this thesis, I aimed to study the factors influencing gene flow risk from planted Corymbia to native Corymbia forests in subtropical Australia. In particular, I addressed the following questions:

1. Is there a genetic variation in the timing of peak flowering among regions of Corymbia spp?
2. Does tree vigour influence bud production in Corymbia?
3. Is it possible to identify Corymbia hybrids from their parental taxa, spotted gums and C. torelliana, with morphology?
4. Can near infrared spectroscopy detect Corymbia hybrids?
Each data chapter (Chapters 2-4) was written as a stand-alone publishable manuscript. Chapter 2 determines if there is genetic variation in the timing of flowering and potential to flower among regions of CCV, and investigates the influence of tree vigour on bud production. Using a common garden, I compared the timing of floral development in the common garden to that observed in native populations. I also determined if flowering intensity and abundance were related to tree vigour.

Methodologies to detect hybrids for large scale screening of seedlings in order to determine hybridisation rates were also explored. In Chapter 3, I investigated the use of morphology to identify hybrids. Leaf and stem characters of parental species (C. c. torelliana as the mother; C. c. variegata, C. c. citriodora and C. henryi as pollen parents) and hybrids (outbred F1 and advanced generation hybrids) were studied in a common garden, and determined which among these characters were most useful for hybrid identification.

In Chapter 4, I determined the utility of NIR to detect hybrids. Using the same materials as in Chapter 3, I obtained NIR spectra at different ages, different tissue preparation and different NIR instruments. I determined the best combination of age, tissue preparation and NIR instrument for hybrid identification. I also discussed the use of NIR in different scenarios in which gene flow from Corymbia hybrids might occur.
Chapter 2. Genetic control of flowering in spotted gum, *Corymbia citriodora* subsp. *variegata* and *C. maculata*

This chapter has been submitted for publication 09 September 2013.


Contributions to this chapter

**Experimental design:** Myralyn Abasolo, Mervyn Shepherd and David J Lee designed the experiment.

**Field work:** Myralyn Abasolo and Mervyn Shepherd were primarily involved in the field work and Gary Ablett provided assistance.

**Statistical analysis and interpretation:** Myralyn Abasolo and Lyndon Brooks performed the statistical analyses.

**Writing:** Myralyn Abasolo, Mervyn Shepherd, David J Lee, Carolyn Raymond and Lyndon Brooks wrote the manuscript.
2.1. Abstract

Genetically determined asynchrony in anthesis is an effective barrier to gene flow between planted and native forests. The degree of genetically controlled variation in the timing of key floral developmental stages in a major plantation species in subtropical Australia, *Corymbia citriodora* subsp. *variegata* and its relative *C. maculata* was investigated. Flowering observations were made in a common garden in spring on 1855 trees from eight regions over three consecutive years, and monthly on a subset of 208 trees for 12 months. Peak anthesis time was stable over years and translocated trees tended to be congruent with the observations in native stands, suggesting strong genetic control of anthesis time. Two flowering races of *C. c. variegata* were identified, an early flowering northern race and a late flowering southern race. The near-term potential to flower was best explained by the interaction of region and year suggesting that a threshold size should be reached before a tree flowers. Partial asynchrony of peak anthesis between the two flowering races of *C. citriodora* subsp. *variegata* may moderate an otherwise high likelihood of gene flow where the northern race of *C. citriodora* subsp. *variegata* is planted in the south of its range neighbouring native stands. Other barriers to minimise gene flow should be explored.
2.2. Introduction

Gene flow via pollen from planted forests is increasingly an issue for forest managers where transplanted species can hybridise with the natives (Potts et al., 2003; Laikre et al., 2010). While hybridisation is a natural process that shapes populations and species, human-mediated hybridisation may be undesirable as it may alter the genetic structure of large populations or cause demographic swamping in small populations (Rieseberg et al., 1989; Thompson et al., 2010). Human-mediated hybridisation is a widespread problem occurring throughout the world, with many examples of translocated species hybridising with natives (Laikre et al., 2010). Hybridisation is facilitated by synchronicity in flowering time and has been documented across many groups of trees including oaks, poplars, eucalypts and pines (e.g. Meirmans et al., 2010; Millar et al., 2012; Thompson et al., 2010; Talbot et al., 2012).

In Australia, a major issue with gene flow from forest plantations is the planting of locally exotic eucalypts (Potts et al., 2003; Barbour et al., 2008) and, in the subtropics, the most important taxa are the spotted gums (Genus Corymbia Section Maculatae). Corymbia citriodora subsp. variegata (CCV) represents the single most important spotted gum taxon for hardwood plantations in subtropical Australia, with around 20 000 ha established in south east Queensland (QLD) and northern New South Wales (NSW) (Lee, 2007; Nichols et al., 2010). There are four taxa of spotted gums that have wide natural distribution in subtropical Australia. They occur as a latitudinal replacement series along the eastern coast from around Cooktown in Queensland to as far south as the Mottle Ranges in Victoria, but with regions of disjunction and sympathy (Hill and Johnson, 1995).

Recent studies of the genetic structure in the spotted gum complex have shown that genetic diversity within the three northern taxa (CCV, C. citriodora subspecies citriodora and C. henryi) is distributed on a geographic rather than a taxonomic basis, and that they are genetically distinct from the southernmost taxon, C. maculata (CM) (Shepherd et al., 2008a; Shepherd et al., 2012). The planting of CCV, mainly in southeast QLD and northern NSW, exposes native forests that contain Corymbia
species (mainly spotted gums, but also other *Corymbia* species) to potential negative consequences from gene flow (Barbour *et al.*, 2008).

Genetic risk decision trees are based on taxonomic, biological and geographic criteria (Byrne *et al.*, 2011). Based on taxonomic and geographic criteria, the likelihood and impact of gene flow from plantation *Corymbia* has been assessed (Barbour *et al.*, 2008). The likelihood of gene flow between planted and native CCV or other spotted gums in northern NSW was thought to be high (Barbour *et al.*, 2008), as a number of taxa are highly inter-fertile and likely share common pollinators, including those able to travel long distances (Bacles *et al.*, 2009). The impact of gene flow was thought to be low because of the view that widespread species with low conservation value are less vulnerable than rare species (Barbour *et al.*, 2008). But the impact may also be high when species exhibit genetic and adaptive divergence because the offspring may exhibit outbreeding depression (Costa e Silva *et al.*, 2012). Clearly biological factors such as the timing of flowering should be studied as flowering synchrony may influence the likelihood and impact of gene flow.

One of the key opportunities for managing risk from plantings is the deployment of planted stock which has genetically controlled asynchrony in anthesis time with native stands (Griffin, 1982; Potts *et al.*, 2003). Several studies have shown that differences in anthesis times between taxa, provenances and families of temperate eucalypts are under genetic control (Gore and Potts, 1995; Apiolaza *et al.*, 2001; Jones *et al.*, 2011). Differences in anthesis time have also been noted between taxa of subtropical spotted gums (McDonald, 2004). Most surveys to date, however, have been of native stands where genetic differences are confounded by environmental factors (Dale and Hawkins, 1983; Pook, 1984; Pook *et al.*, 1997; Law *et al.*, 2000). Floral development observations in a common garden are required to separate out the degree of genetic and environmental determination in these processes. Comparison of floral development in native stands with that of translocated trees is also useful for assessing the extent of genetic control, as stability of flowering processes across different environments is also an indication of genetic determination. Determining the potential of a tree to flower by examining buds, instead of open flowers, can overcome the difficulty of examining zero-inflated data, especially in cases where most trees have no flowers but have buds. Approaches that model the presence or absence of species (or reproductive structures)
are typically used (Kunhert et al., 2005; Faddy 1998). One approach is the use of logistic regression models which take into consideration the presence or absence of a species (Fletcher et al., 2005; Welsh et al., 1996). To my knowledge, this is the first study of flowering in trees using replicated samples, analysing results with the presence or absence data, and comparing them with flowering in native populations.

One of the objectives of the study was to test if there was a relationship between tree vigour and flowering in spotted gum, and whether a certain threshold size must be attained before a tree reaches maturity. Studies of eucalypts, including native stands of spotted gums have shown that larger trees tend to produce more flowers (Ashton, 1975; Pook et al., 1997; Bacles et al., 2009). In turn, tree vigour may be strongly correlated to disease or pest tolerance. If tree size influences bud abundance, then this has also implication on the pollen source/sink ratio of the plantations relative to the native stands. There may be higher risk of gene flow in plantations with larger relative population size (higher source/sink ratio) than plantations with smaller relative population size (lower source/sink ratio). Management strategies such as increasing the isolation distance from the plantation might be needed to minimise the risk in situations when the plantation is intended for long rotations (i.e. 25 years for solid wood production).

In recent years, commercial spotted gum plantations in northern NSW have been derived from one or a few select provenances of CCV from the northern end of its range, particularly Woondum, near Gympie in QLD. This provenance has been recognised as having a high tolerance to *Quambalaria shoot blight* (QSB), a fungal disease which attacks new shoots and expanding leaves leading to a loss of apical dominance in severe cases (Pegg et al., 2011). Genetics has been shown to influence disease resistance in other trees such as pines (Devey et al., 1995; Li et al., 2006) and poplars (Cervera et al., 1996; Lefèvre et al., 1998; Dowkiw and Bastien 2004; Jorge et al., 2005). There was a strong correlation between the fungal resistance and genetic material of *Eucalyptus globulus* planted in a range of environments, suggesting that genetic control of fungal damage is stable across environments (Freeman et al., 2008). Similarly, tolerance to QSB has been shown to hold up across a range of environments and seasons including at the current study site (Johnson et al., 2009). Trees tolerant of QSB are more vigorous (Brawner et al., 2011) and it is hypothesized that trees from
recognised QSB tolerant provenances in the current study will be more vigorous and thus more likely to exhibit the potential to flower than trees with lower tolerance provenances in this study.

The study addresses the following questions:
1. Is there a genetically controlled difference in floral development of spotted gum taxa both at the regional and family level?
2. Is there flowering synchrony among the different provenances of CCV?
3. Is there a relationship between tree vigour, disease resistance and flowering in spotted gums?
2.3. Methods

2.3.1. Trial descriptions and germplasm material

2.3.1.1. Bonalbo Trial

The main study was conducted at a trial located 13 km southeast of Bonalbo (28°52’S, 152°38’E; elevation 164 m asl) in northern NSW (Figure 1). The trial was an ex-grazing site with an east-facing aspect and an infertile yellow podzolic soil (Johnson et al., 2009). The long term mean annual rainfall (MAR) at the site is 1031 mm/yr (data from the Bonalbo Post Office, 9.6 km northeast of the site; (Australian Bureau of Meteorology, 2012).

The Bonalbo trial was established in 1999 as a randomised complete block design (RCBD) with five replicates of 196 open-pollinated families planted in five-tree-line plots. The trial was not thinned at the time of assessment, thus it maintains its original spacing of 4.0 x 2.5 m (nominally 1 000 stems/ha (Forests NSW)). It contains the widest collection of CCV provenances from across its natural distribution for any trial of that age in NSW.

In this study, flowering was observed on a subset of 1855 trees from 128 families of CCV and CM. The 26 CCV provenances represented are those that are typically used commercially. The five CM provenances assessed in this study may contain intergrades with CCV as they originate from just south of the southernmost distribution of CCV (Shepherd et al., 2012). Each CCV and CM provenance was represented by one to thirteen families with each family nominally having 20 trees assessed. The 31 CCV and CM provenances (Table 1) were categorised into eight eco-geographic regions based on four environmental variables: latitude, elevation, mean annual rainfall and temperature (Figure 1).
Figure 1. Locations of *Corymbia citriodora* subsp variegata (CCV) and *C maculata* (CM) provenances (grouped into regions) studied at the Bonalbo Trial.
Table 1. List of *Corymbia citriodora* subsp *variegata* and *Corymbia maculata* provenances, grouped by regions, assessed for flowering at the Bonalbo trial. Geographic coordinates, summary climatic parameters and numbers of families and trees are shown for each provenance.

<table>
<thead>
<tr>
<th>Region&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Provenances</th>
<th>Latitude (deg min)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Long (deg min)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Elev (m)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>MAR (mm)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>MAT (°C)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Families (n)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Number of trees assessed each year</th>
<th>Number of trees assessed monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQLD</td>
<td>Toolara Home</td>
<td>26°05' 152°43'</td>
<td>137</td>
<td>1143</td>
<td>19.70</td>
<td>1</td>
<td>15</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Toolara Wolvi</td>
<td>26°07' 152°47'</td>
<td>120</td>
<td>1148</td>
<td>19.70</td>
<td>1</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brooyar</td>
<td>26°10' 152°30'</td>
<td>90</td>
<td>1143</td>
<td>19.90</td>
<td>2</td>
<td>32</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Woondum</td>
<td>26°15' 152.82</td>
<td>400</td>
<td>1600</td>
<td>17.86</td>
<td>7</td>
<td>177</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Region total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11 179 44</td>
</tr>
<tr>
<td>IQLD</td>
<td>Wondai</td>
<td>26°22' 151°49'</td>
<td>350</td>
<td>800</td>
<td>18.30</td>
<td>3</td>
<td>48</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 48 18</td>
</tr>
<tr>
<td>SCQLD</td>
<td>Esk</td>
<td>27°18' 152°20'</td>
<td>300</td>
<td>850</td>
<td>18.20</td>
<td>2</td>
<td>29</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lockyer</td>
<td>27°28' 152°17'</td>
<td>150</td>
<td>850</td>
<td>18.90</td>
<td>1</td>
<td>11</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 40 21</td>
</tr>
<tr>
<td>RICH</td>
<td>Richmond Range</td>
<td>28°40' 152°42'</td>
<td>405</td>
<td>1233</td>
<td>16.70</td>
<td>10</td>
<td>149</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sugarloaf</td>
<td>28°56' 152°37'</td>
<td>215</td>
<td>1083</td>
<td>17.50</td>
<td>1</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mount Pikapene</td>
<td>29°03' 152°40'</td>
<td>235</td>
<td>1157</td>
<td>17.30</td>
<td>1</td>
<td>15</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mount Belmore</td>
<td>29°08' 152°45'</td>
<td>270</td>
<td>1260</td>
<td>17.00</td>
<td>1</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13 196 29</td>
</tr>
<tr>
<td>HENSW</td>
<td>Gilgarry</td>
<td>28°50' 152°16'</td>
<td>670</td>
<td>1148</td>
<td>14.90</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Girard</td>
<td>28°52' 152°18'</td>
<td>675</td>
<td>1270</td>
<td>14.75</td>
<td>2</td>
<td>27</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ewingar</td>
<td>29°03' 152°28'</td>
<td>483</td>
<td>1187</td>
<td>15.83</td>
<td>3</td>
<td>50</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Washpool</td>
<td>29°17' 152°27'</td>
<td>428</td>
<td>1147</td>
<td>15.93</td>
<td>4</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grange</td>
<td>29°28' 152°35'</td>
<td>528</td>
<td>1294</td>
<td>15.22</td>
<td>5</td>
<td>71</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ramornie</td>
<td>29°43' 152°36'</td>
<td>425</td>
<td>1222</td>
<td>15.68</td>
<td>3</td>
<td>46</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dalmorton</td>
<td>29°49' 152°24'</td>
<td>648</td>
<td>1104</td>
<td>14.35</td>
<td>2</td>
<td>23</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marara</td>
<td>29°56' 152°35'</td>
<td>590</td>
<td>1246</td>
<td>14.58</td>
<td>4</td>
<td>57</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chaelundi</td>
<td>29°56' 152°23'</td>
<td>680</td>
<td>1069</td>
<td>14.13</td>
<td>3</td>
<td>47</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27 383 35</td>
</tr>
<tr>
<td>HRNSW</td>
<td>Boundary Ck</td>
<td>29°57' 152°34'</td>
<td>481</td>
<td>1197</td>
<td>15.20</td>
<td>10</td>
<td>158</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sheas Nob</td>
<td>30°03' 152°37'</td>
<td>653</td>
<td>1315</td>
<td>14.12</td>
<td>3</td>
<td>56</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kangaroo River</td>
<td>30°05' 152°52'</td>
<td>399</td>
<td>1452</td>
<td>15.68</td>
<td>4</td>
<td>64</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bagawa</td>
<td>30°10' 152°57'</td>
<td>343</td>
<td>1563</td>
<td>16.20</td>
<td>4</td>
<td>56</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21 334 44</td>
</tr>
<tr>
<td>CNSW</td>
<td>Wedding Bells</td>
<td>30°04' 153°10'</td>
<td>82</td>
<td>1528</td>
<td>18.25</td>
<td>4</td>
<td>49</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower Bucca</td>
<td>30°10' 153°04'</td>
<td>200</td>
<td>1582</td>
<td>17.28</td>
<td>4</td>
<td>57</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Regional total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8 106 17</td>
</tr>
<tr>
<td>MAC</td>
<td>Newry</td>
<td>30°32' 152°54'</td>
<td>140</td>
<td>1716</td>
<td>17.50</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ingalba</td>
<td>30°51' 152°45'</td>
<td>166</td>
<td>1388</td>
<td>17.35</td>
<td>4</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tamban</td>
<td>30°56' 152°48'</td>
<td>113</td>
<td>1323</td>
<td>17.70</td>
<td>4</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Boonanghi</td>
<td>31°03' 152°31'</td>
<td>443</td>
<td>1551</td>
<td>15.62</td>
<td>5</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yessabah</td>
<td>31°08' 152°37'</td>
<td>232</td>
<td>1418</td>
<td>16.72</td>
<td>5</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19 253</td>
</tr>
</tbody>
</table>

*Region codes; CQLD = Coastal Queensland; IQLD = Inland Queensland; SCQLD = Subcoastal Queensland; RICH = Richmond; HENSW = High Elevation New South Wales; HRNSW = High Rainfall New South Wales; CNSW = Coastal New South Wales; MAC = Maculata. *Corymbia maculata* trees were from the MAC region while *C. c* *variegata* trees came from seven other regions indicated (See Figure 1).

<sup>a</sup>Latitude, longitude, elevation and mean annual rainfall (MAR) were from Johnson *et al.* (2009).

<sup>b</sup>Latitude, longitude, elevation and mean annual temperature (MAT) were from Johnson *et al.* (2009).

<sup>c</sup>MAT = mean annual temperature.

<sup>d</sup>Number of families assessed each year.
2.3.1.2. *Emu Creek Trial*

Flowering observations were also conducted at a second site, Emu Creek, for the purpose of estimating family level genetic control. Emu Creek (28°49’S, 152°29’E; elevation 183 m) is located 14 km northwest from the Bonalbo trial. Long term mean annual rainfall at the site is 1273 mm/yr (based on the Tabulam (Muirne) weather station; (Australian Bureau of Meteorology, 2012)) which is 9.14 km NW of the site. Environmental conditions at this trial were largely expected to be similar with Bonalbo but with slightly higher rainfall.

The Emu Creek trial was established in 2005 as a randomised incomplete block design (IBD) with 18 replicates and 17 blocks per replicate. Each block consists of 32 single tree-plots. Trees from 50 families that were in common with the Bonalbo trial were assessed in November 2010 and 2011.

2.3.2. **Flowering assessments**

There were two types of flowering surveys done, annual and monthly. In each type of survey, the number of reproductive structures (initials, buds, flowers or current season/green capsule) on each tree were counted or estimated (in cases of prolific budding/flowering), and recorded, using 10 x 40 binoculars (Figure 2). For the present study, pre-anthesis flowers were classified as initials when they were visible and up to approximately 3 mm in width, above which size they were classified as buds. Flowers were classified as open when operculum had been shed and anthers were evident. Post-anthesis structure was classified coarsely as green capsules when anthers and styles were no longer evident and distinguished from older capsules from earlier seasons by their appearance and position on the branch. Green capsules (current season) appear plump and mostly located towards the tip of the branches while older capsules appear shrivelled and concentrated on branches near the trunk of the tree. In *Corymbia*, the maturation of buds may take between nine to 24 months (See Table 2).
2.3.2.1. Annual

An intense survey of flowering was undertaken at the Bonalbo trial in November each year from 2009 to 2011 when the trees were 10, 11 and 12 years old. The timing of the assessment corresponded with what was thought to be within the flowering period of spotted gums in northern NSW (Law et al., 2000; Barbour et al., 2008; Table 2). Four replicates of each open pollinated family were assessed with two observers assessing flowering in two replicates each. The observers were equally trained and so the degree of discrepancy between two observations was deemed minimal, thus variation was not analysed. The number of initials, buds, open flowers and current season capsules present on each of the 1855 trees were assessed.

Figure 2. Reproductive structures of *Corymbia* species. A. Initials; B. Buds; C. Flowers; D. Green capsules.

To allow testing for family level genetic control, a set of 50 families common to both the Bonalbo and Emu Creek trials were also assessed in November of 2010 and 2011 to allow comparative analysis. The trees assessed were five and six years old during the time of assessments. Three hundred and eleven and 505 trees in six and 10 replicates
were assessed in 2010 and 2011 respectively. In addition to the flowering assessment, the Diameter at Breast Height over bark (DBH) of each of the 1855 trees was measured at the Bonalbo Trial during the 2009 assessment.

2.3.2.2. Monthly

To track the progress of floral development in spotted gum throughout the year and assess the degree of synchrony among provenances, monthly assessments were undertaken over a period of 12 months on a subset of 208 trees from 22 provenances in two replicates at the Bonalbo trial (Table 1).

The monthly assessments were undertaken from September 2010 to November 2011, however, no assessment was undertaken during December 2010, January 2011 and March 2011 due to heavy rainfall during this period that prevented site access. Two assessments were completed in October 2010 which was thought to be the peak anthesis time of plantation CCV in northern NSW and southern QLD (Barbour et al., 2008). The trees chosen for the monthly assessments were reproductively active (i.e. having initials, buds flowers and/or capsules) during the first assessment in November 2009 and had the largest DBH in each plot.

The timing and duration of key floral development time like bud initiation, bud development and peak anthesis time obtained in the monthly assessment of translocated trees were compared with studies of observations in native populations as presented in Table 2. The comparison of floral development between the native populations and the translocated trees were used to test for stability in these parameters.

2.3.3. Statistical analysis

2.3.3.1. Reproductive development across the 12 month-assessments

From the 208 trees assessed monthly, the trees which had any reproductive structure (initials, buds, flowers and/or green capsules), i.e. reproductively active trees, at any point during the observation period, were determined. These reproductively active trees
were used to determine the average amount of reproductive structure in each month for each region in the years 2010 and 2011.

**Table 2.** A summary of key floral development attributes for *C. c variegata* and *C. maculata* compiled from observations in native stands reported in literature and translocated trees from this study.

<table>
<thead>
<tr>
<th>Region</th>
<th>Location of Observations</th>
<th>Taxon</th>
<th>Native Stand/Translocated</th>
<th>Bud initiation (inclusive months)</th>
<th>Bud development time (no. of months)</th>
<th>Peak anthesis time (inclusive months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern subtropic</td>
<td>Bonalbo</td>
<td>CCV</td>
<td>Translocated&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Sept – Nov</td>
<td>9-13</td>
<td>Jul – Aug</td>
</tr>
<tr>
<td></td>
<td>Barakula</td>
<td>CCV</td>
<td>Native&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Nov – Feb&lt;sup&gt;d&lt;/sup&gt;</td>
<td>10 – 11</td>
<td>Nov</td>
</tr>
<tr>
<td></td>
<td>Gympie</td>
<td>CCV</td>
<td>Native&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Sept-Oct</td>
<td>10-11</td>
<td>July-Aug</td>
</tr>
<tr>
<td>Southern subtropic</td>
<td>Bonalbo</td>
<td>CCV</td>
<td>Translocated&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Feb – Mar</td>
<td>20-21</td>
<td>Oct</td>
</tr>
<tr>
<td></td>
<td>Various Nth NSW</td>
<td>Native&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Dec – May</td>
<td>15 – 24</td>
<td>Jan-Mar</td>
<td></td>
</tr>
<tr>
<td>Temperate</td>
<td>Grafton</td>
<td>CM</td>
<td>Translocated&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Oct</td>
<td>&gt;12 months</td>
<td>No anthesis observed due to bud abortion</td>
</tr>
<tr>
<td></td>
<td>Southern NSW (Kiola State Forest)</td>
<td>CM</td>
<td>Native&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Sept</td>
<td>18</td>
<td>Jun</td>
</tr>
</tbody>
</table>

<sup>a</sup>Bud development time refers to the time from initiation to anthesis.

<sup>b</sup>Present study.

<sup>c</sup>Dale and Hawkins (1983).

<sup>d</sup>Authors’ interpretation based on reported bud development and anthesis times.

<sup>e</sup>DJ Lee, unpublished anecdotal observations.

<sup>f</sup>Law *et al.* (2000).

<sup>g</sup>Abasolo *et al.* unpublished data from a separate taxa trial.

<sup>h</sup>Pook *et al.*(1997).

The provenance of trees is coarsely grouped into three zones for the purposes of this table, the Northern sub-tropical approximating the range of CCV north of the QLD-NSW border and a Southern Subtropical region encompassing the range of CCV south of the border. *C. maculata*’s origins are classified as Temperate, although it overlaps with CCV in the north of its range.
2.3.3.2. Flowering intensity and synchrony

The monthly flowering intensity and degree of flowering synchrony at the provenance level were determined using the data on the subset of 208 trees observed monthly from September 2010 to November 2011. An intensity value of zero to 5 was determined for each provenance from the sum of score values for the quantity of flowers and for the proportion of trees that flowered in each provenance (scaling system modified from Keatley et al. (2004) and Keatley and Hudson (2007)). The modification of the scaling system was necessary to fit the lower flower abundance of the trials assessed in this study compared to the native stands studied in Keatley et al. (2004) and Keatley and Hudson (2007) with which the scoring was based. The Intensity score ranged from 0 to 3 and was assigned as follows: 0 = no flowering; 0.5 = < 100 flowers; 1 = 101 – 1 000 flowers; 2 = 1 001 – 5 000 flowers and 3 = > 5000 flowers. The score for the proportion of trees that flowered in a provenance was assigned as follows: 0.5 = 0.1 – 10%; 1 = 11 – 40%; 1.5 = 41 – 60%; 2 = 61 – 100%. An intensity score of 0 indicated that no tree from the provenance flowered while a score of 5 indicated that flowering was heavy and thus the Intensity score was 3 and most of the trees in the provenance flowered (i.e. a proportion score of 2). Peak flowering was the month or months with the highest flowering intensity (Keatley et al., 2004; Keatley and Hudson, 2007). The monthly intensity values are reported below. Average intensity across years of observation are typically reported in long term flowering studies (i.e. Keatley et al., 2004; Keatley and Hudson, 2007) but averaging was not appropriate for the two-years (13 observations including two observations in October 2010) of data.

Synchrony index was calculated according to Augspurger (1983):

\[ Xi = \left( \frac{1}{n-1} \right) \left( \frac{1}{fi} \right) \sum_{j=1}^{n} ej \neq i \]

where: \( Xi \) = index of synchrony for individual \( i \); \( ej \) is the number of times both individuals \( i \) and \( j \) are flowering synchronously where \( j \neq i \); \( fi \) = number of times individual \( i \) is flowering and \( n \) is the number of individuals in the population. Individuals \( i \) and \( j \) have complete synchrony if \( X = 1.0 \) and no overlap if \( X = 0.0 \). The values for the strength of synchrony were as follows: low synchrony (≤0.19); moderate
(0.20 – 0.30); high (0.31 – 0.75); very high (0.76 – 1.00) (modified from Keatley et al. (2004)).

2.3.3.3. Assessing the near-term potential to flower by determination of the odds of a tree from a region having buds

The variation in flowering among regions (due to genetics), years and interaction effects was tested based on the November assessment over three years (2009, 2010 and 2011). The presence of buds was used as an indicator for the “potential to flower” because of the very few trees with flowers at the census time-point. The use of buds as surrogate to flowers is justified as buds and flowers have good correlation (O’Brien et al. 2007).

A high percentage of the trees within the trial did not produce buds in any given year (i.e. 92% in 2011, data not shown), leading to the dataset being highly zero-inflated. Such data is not suitable for parametric statistical analysis. The near-term potential of a tree to flower in a year, region or family was assessed using odds, calculated as the number of trees with buds divided by the number of trees without buds for each year in each replicate. Odds were used because simple quantity counts were not appropriate due to unequal sample sizes and the binary nature of the data (i.e. trees with buds, trees without buds) at any group level (i.e. family in a replicate).

A multilevel (family, replicate within family, observation within replicate within family) binary logistic regression model was used to test the significance of region (ecoregion), year and region by year interaction effects. Regions and years were the fixed effects while family, replicate and observation (tree) were treated as the random effects. A logit function was necessary to link the non-linear response variable to components of a linear model. It was also necessary to remove families from the analysis which had no trees with buds in any replicate in any year because the logarithm of zero is undefined. Of the 128 families in the original set assessed, 105 families had at least one tree with buds in each replicate in each year and were included in the analysis.

The final model fitted to determine the regional, year and interaction of region and year effects was:
Chapter 2: Genetics of flowering

Logit (π_{ijk}) = β_{0jk} + β_{Regions\_k} + β_{Year\_ijk} + β_{(Regions \times Year)\_ijk} + variance (family)_{k} + variance (replicate)_{jk} + variance (residual)_{ijk}

Where Logit (π_{ijk}) is the natural logarithm of the odds of trees having flower buds in observation \( i \) in replicate \( j \) from family \( k \); \( β_{0jk} \) is the intercept in the model (Region 1 in year 1); \( β_{Regions\_k} \) is a set of dummy variables representing the eight regions; \( β_{Year\_ijk} \) is a set of dummy variables representing the three years of observations; \( β_{(Regions \times Year)\_ijk} \) is a set of dummy variables representing the interaction of regions and years, and variance (family)\_k, variance (replicate)\_jk, and variance (residual)\_ijk were the residual variances at the family, replicate and observation levels.

A set of models with this random structure and different fixed effects was fitted and compared using the deviance information criterion (DIC, Spiegelhalter et al., 2002) and the model with the lowest DIC was chosen for interpretation. The DIC is a measure of a model’s overall fit, adjusted for parsimony – it decreases as additional explanatory variables reduce the model deviance and increases as the number of fitted parameters increases.

Five models were compared for the odds of trees having buds with the random structure, family, replicate within family, and observation within replicate within family, using the DIC criterion: namely 1) Intercept only; 2) Intercept + Region; 3) Intercept + Year; 4) Intercept + Region + Year; 5) Intercept + Region + Year + Region x Year (Table 4).

The models were estimated by the Markov Chain Monte Carlo (MCMC) method in MLwiN software v2.24 (Rasbash et al., 2000). Each model was run with 500 cycles of burn in and a monitoring chain length of 175,000 iterations.

For descriptive purposes, the logit scale estimates from the models were back-transformed to estimates of the odds of trees with buds for each region in each year, and odds ratios for comparing each region between years and comparing regions in each year. For example, the odds ratio comparing the odds of trees with buds in region 1 in year 1 to the odds of trees with buds in region 2 in year 1 equals (the number of trees with buds from region 1 in year 1 / the number of trees without buds from region 1 in...
year 1) divided by (the number of trees with buds from region 2 in year 1 / the number of trees without buds from region 2 in year 1).

Odds ratios for multiple pairwise comparisons were tested using single degree of freedom Wald chisquare tests ($\chi^2 = \text{estimate/se}^2$). No adjustment was made for the number of comparisons as the power for the tests was low due to the small proportions of trees with buds.

### 2.3.3.4. Vigour assessments

It was determined whether tree size can also be related to flowering using the 1855 trees with DBH data. Several tests were used to relate flowering effort to DBH. First, the difference between the mean DBH of the trees that were reproductively active against the mean DBH of trees which were not, was determined. A one tailed independent sample t-test was used to determine the statistical significance between these two groups. Second, flowering effort was used as a measure of the ability of a tree to flower and compared it to DBH.

Flowering effort was a measure of the capacity of a tree to bear any reproductive structure. It was a measure of the total number of reproductive structure across the three years of study. To avoid double counting, only initials were added to the total amount of materials evident in the first year of assessment, 2009, hence the formula to estimate flowering effort was:

\[
\text{Flowering effort} = \text{Initials}_{2009} + \text{Buds}_{2009} + \text{Flowers}_{2009} + \text{Green capsules}_{2009} + \text{Initials}_{2010} + \text{Initials}_{2011}
\]

As there was a very large variation in the total flowering effort, the total amount to scores were converted as follows:

- 0 = no flowering
- 1 = 1 - 10
- 2 = 11 – 100
- 3 = 101 – 1 000
The differences in mean DBH among the six different scores of flowering effort (i.e. 0, 1, 2, 3, 4 and 5) were determined using a univariate ANOVA. A LSD was used to determine the significant differences among the mean DBH of trees belonging to each group with the respective flowering scores. The mean DBH and mean flowering score of each region was also determined and compared the tree vigour (DBH) of each region to its fecundity (flowering effort). A univariate ANOVA and LSD were used to determine the significance of differences in DBH and mean flowering effort of each region. All analyses were performed using SPSS v19 (IBM).

2.3.3.5. Family level genetic control

The expectation was used that if there is genetic control at the family level, there would be a relationship in the average proportion of trees with flowers in a family between the two study sites. Observations from a set of 50 families common to both the Bonalbo and Emu Creek sites were evaluated in this study. This consisted of 1143 and 1130 trees at the Bonalbo Trial in 2010 and 2011 respectively, and 311 and 505 trees at the Emu Creek in 2010 and 2011, respectively. A goodness of fit test was constructed to test for concordance, comparing the expected class sizes of families that flowered in both trials to those that only flowered at one site under the null hypothesis of no genetic effect (where equal class sizes are expected in the one or two site categories This analysis was performed using the Contingency Table Analysis (Crosstabs) routine in SPSS v19 (IBM). Test for concordance was used to test genetic effects instead of the classical broad heritability measure because of the limited number of flowering trees within a family.
2.4.  Results

2.4.1. Timing, intensity, and synchronicity of anthesis in spotted gum

The monthly flowering assessments between September 2010 and November 2011 suggested that CCV flowered from winter (June) to early summer (November) at the Bonalbo trial (Figure 3a-c). Trees from some regions may have also flowered in December and January of those years (no observations were made in December 2010 and January 2011) but the flowering season appeared to be finished by February 2011 as no trees were in flower at this time. The assessment was based on 10 out of the 22 provenances in the study as four provenances (Dalmorton, Ewingar, Lower Bucca and Kangaroo River) had only a single tree that flowered during the assessment period, and no tree from the other eight provenances flowered. Hence, the provenances with only one tree that flowered and those that did not flower were excluded from the following analysis.

Overall, the flowering intensity was low at the site (flowering intensity is the sum of the score value of the number of flowers in the population and the score value of the proportion of trees that flowered) with provenance values ranging between 1 and 2, out of a maximum possible score of 5 (Figure 3). There was a trend, however, where more northerly provenances generally had higher flowering intensity scores than provenances from the south. The northern provenances, namely: Toolara Home, Toolara Wolvi, Woondum, Brooyar and Wondai generally flowered annually in contrast to the more southerly provenances, namely: Esk, Lockyer, Boundary Creek, Wedding Bells and Richmond Range which tended to flower in alternate years (Figure 3a-c). For example, trees from Wedding Bells and Boundary Creek flowered in 2010 whereas trees from Richmond Range did not, but in the following year, trees from Richmond Range flowered, but trees from both Boundary Creek and Wedding
Figure 3. Provenance level flowering intensities for a subset of 208 trees assessed monthly between September 2010 and November 2011 (no assessments in Dec 2010 and Jan 2011) at the Bonalbo Trial. a. Provenances from the Coastal QLD region; b Inland and Subcoastal QLD regions and c. NSW regions, including Richmond Region, Coastal NSW and High Rainfall NSW. Flowering intensity is the total of a score for the quantity of flowers and a score of the proportion of trees that flowered in a provenance (Scaling modified from Keatley et al. (2004) and Keatley and Hudson (2007)).
Bells did not (Figure 3c). The flowering window of the northern provenances also appeared to be earlier than those from the south. The flowering of the northern provenances spanned from June to October whereas flowering in southern provenances spanned from October to November. Thus to characterise flowering in CCV, there appeared to be regional differences where those from the north had an earlier, broader, and more stable peak of anthesis, relative to those from the south.

Overall synchronicity levels between or within provenances were low to high (range from 0 to 0.41, Table 3), but as expected, within provenance synchrony was generally higher than between provenances. The provenances from the north, Woondum and Toolara Wolvi, had the highest within provenance synchrony, both had 0.35, while both Brooyar and Boundary Creek had 0.25 (Table 3). The within-provenance synchrony of these provenances was considered moderate to high relative to other provenances in the study, while the within-provenance synchrony of Lockyer (0.16), Richmond Range (0.15) and Esk (0.13) were considered low. Both Wedding Bells and Toolara Home provenances had only two trees each that flowered, but these two trees did not flower synchronously, therefore their within-provenance synchrony indices were zero. Other provenances had only one tree flowering at any given time, therefore their synchrony values were also zero.

Provenances within the same region, and therefore geographically close, tended to have higher between-provenance synchronicity than provenances from different regions (Table 3). There was also a trend where northern provenances tended to be more synchronised among themselves than southern provenances (Table 3). For example, the highest between-provenance synchrony observed was between two northern provenances from the same region (CQLD), Toolara Wolvi and Brooyar (0.41). This was higher when compared with the synchronicity between Toolara Wolvi and (CQLD) and Wondai (IQLD), which was 0.22. However, this trend was not observed among provenances from the southern regions. There were only three provenances that flowered from the southern regions and none of them belong to the same region. Furthermore, there was no trend for geographically close provenances from southern regions to have higher synchronicity than more distal provenances (Table 3; p>0.05).
Table 3. Pairwise synchrony index among- and within-provenances of spotted gums calculated from observations at the Bonalbo Trial in 2010-2011. The synchrony index was based on Keatley et al (2004). Provenance codes are; TOOH = Toolara Home; TOOW = Toolara Wolvi; BROO = Brooyar; WOON = Woondum; WOND = Wondai; LOCK = Lockyer; RRNG = Richmond Range; BOUN = Boundary Creek; WEDD = Wedding Bells.

<table>
<thead>
<tr>
<th>Regions/ Prov</th>
<th>CQLD</th>
<th>Regions/ Prov</th>
<th>IQLD and SCQLD</th>
<th>NSW Regions (RICH, CNSW and HRNSW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOOH</td>
<td></td>
<td>TOOH</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOOW</td>
<td>0.23</td>
<td>TOOW</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>BROO</td>
<td>0.20</td>
<td>BROO</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>WOON</td>
<td>0.25</td>
<td>WOON</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>WOND</td>
<td>0.20</td>
<td>WOND</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>ESK</td>
<td>0.12</td>
<td>ESK</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>LOCK</td>
<td>0.15</td>
<td>LOCK</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>RRNG</td>
<td>0.14</td>
<td>RRNG</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>BOUN</td>
<td>0.06</td>
<td>BOUN</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>WEDD</td>
<td>0.06</td>
<td>WEDD</td>
<td>0.18</td>
<td></td>
</tr>
</tbody>
</table>

2.4.2. The progression of flower development in spotted gum

Regional variation was observed in the floral development of spotted gums. Data was pooled into regions for this analysis because regional effects had been identified (See Figure 5a-c and Section 2.4.4.1) and there were too few trees with reproductive structure at the provenance level for reliable estimates. Northern regions refer to all QLD regions (CQLD, IQLD and SCQLD) and southern regions refer to those from NSW, including RICH which is closest to Bonalbo trial.

One of the most notable effects was regional differentiation in initiation of buds (Figure 4 and Appendix 1), where northern provenances tended to possess more early initiating genotypes. The two most northerly regions CQLD and SCQLD were observed to have floral initials during the first four observation times from 8 Sept 2010 till 25 Nov 2010, whereas no tree from any other region (exemplified by RICH in Figure 4) were observed with initials during this period. Instead, all the southern regions were observed to be initialising later in February 2011 (Figure 4 and Appendix 1). Although observations were not taken in November and December 2010, and the commencement of initiation might have been missed, it was still observed that northern regions initialised earlier than southern regions.
Figure 4. Progression of floral development throughout the year for three representative regions of CCV, Coastal QLD, Subcoastal QLD, and Richmond. Log scale values of estimates of counts were as follows: 1 = 1 to 10; 2 = 11-100; 3 = 101-1 000; 4 = 1 001 – 10 000; 5 = > 10 000.
Some regions exhibited cycles of floral development (i.e. from initiation to anthesis) throughout the observation period, most notably CQLD (Figure 4). The trees from this region were observed to be initialising in September 2010. These initials matured over the period February to May until they began to undergo anthesis in June then converted into capsules (Figure 4). Thus the period between initiation and anthesis for CQLD could be nine months (September 2010 – June 2011). A progression from initials to anthesis was evident in trees from some other regions such as SCQLD, but SCQLD differed to CQLD in that it did not progress to anthesis during the observation period (Figure 4). This indicated that anthesis of trees from SCQLD either occurred later than the last observation, or not at all (See Supplementary Material 1). All the southern regions initialised later in February 2011 (Figure 4 and Supplementary Material 1) and did not progress to anthesis by November 2011.

Another effect that was evident was the apparent low conversion rate from buds to open flowers in some regions (See CQLD and SCQLD for example in Figure 4). Here the decline in the amount of buds did not seem to be accounted for by the progression of buds to open flowers or capsules (marked decline in the amount of buds between October and November 2011 levels in CQLD, Figure 4). This sudden decline in bud number without an accompanying increase in the flower or capsules count was also evident in the SCQLD region.

**2.4.3. Timing of anthesis in translocated populations is stable**

A review of the model of flowering patterns in light of observations from this study along with earlier records of flowering (Table 2) indicated that the timing and duration of key reproductive developmental stages for transplanted spotted gums tended to be congruent with those recorded for their native location. This suggests timing of anthesis was largely genetically determined and translocation did not affect the timing in the key reproductive development stages of spotted gums over the scales investigated.
The bud initiation time and bud development time (i.e. duration from initiation to anthesis) of the translocated trees were consistent with the observations in native stands (Table 2). For example, CQLD, which is within the Gympie region, initiated buds in September to November in this study. This was similar to the bud initiation time of CCV in the Gympie region which occurs from September to October (DJ Lee, unpublished anecdotal observation). Bud development time was also similar for translocated trees from CQLD which took 9-13 months for initials to progress to anthesis while coastal (Gympie region) and more inland (Barakula) native populations take 10-11 months to undergo bud development (Dale and Hawkins, 1983; DJ Lee unpublished anecdotal observation). In this study, bud development time of the translocated southern CCV was longer (20-21 months) than the translocated northern CCV but may be in accord with the observation of the longer bud development of native populations of CCV in the south of its range (Law et al., 2000). Similar congruence in the key floral development phases was evident for translocated and native CM.

### 2.4.4. Testing for genetic control in the near-term potential to flower

A detailed annual assessment of two trials (Bonalbo (n= 1855) in 2009, 2010 and 2011 and Emu Creek (n=311 in 2010 and n=505 in 2011)) showed that, overall, the level of reproductive structures (i.e. percentage of trees with buds, flowers or green capsules) present was low at both sites during the November censuses. At the Bonalbo trial, for example, the percentage of trees with reproductive structures was 10.45%, 7.33% and 18.98% in 2009, 2010 and 2011 respectively. The percentage of assessed trees with buds was 7.44%, 3.07% and 8.30% representing approximately ½ of the reproductively active trees in 2009, 2010 and 2011, respectively.

### 2.4.4.1. Regional differences in the potential to flower: a year by year assessment

A model that explains more of the variation will have a lower DIC and the change in DIC value indicated the degree of increased fit (Spiegelhalter et al., 2002). Five models were evaluated; 1) Intercept only model, 2) Intercept + Region; 3) Intercept + Year, 4) Intercept + Region + Year; 5) Intercept + Region + Year plus Region x Year (Table 4).
An initial exploration of the importance of factors that explained model variation indicated that the region by year interaction (Model 5) was the most important factor (DIC = 2214.76; Table 4). The addition of year and region to the model (Model 4) had the second largest fit (DIC = 2310.95; Table 4), whereas the addition of region (Model 2) gave the lowest improvement (DIC = 2376.74; Table 4).

The addition of the interaction term (Model 5) gave the largest improvement to the fit of the model, indicating it was the most significant term in explaining the observed variation. Therefore, results based on this model are presented below, namely, 1) the odds of trees with buds from each region in each year, and the odds ratios comparing 2) each region between years and 3) between regions in each year. The estimated odds of trees with buds for each region with 95% confidence intervals are plotted for each year (Figure 5).

Single-year analyses were undertaken to simplify the investigation of regional effects (ANOVA not shown). Significant differences were found amongst regions within each year (Figure 5 a-c). As an example, regional effects in 2011 are considered, as the largest differences among regions were evident in this year (Figure 5c). Together, MAC and CNSW regions recorded the highest odds for any region at any time during the three years, 0.31 and 0.25, respectively. The MAC region had significantly higher odds of trees with buds than all other regions except CNSW (p<0.05; Figure 5c). The odds ratio between MAC and three other southern regions, HENSW, HRNSW and RICH were 9.97, 7.68 and 4.63, respectively (p<0.001; Figure 5c). The MAC region was almost 176 times more likely to have trees with buds than CQLD region (odds ratio = 175.91; p<0.001), and more than four times more likely than IQLD region (odds ratio = 4.46; p<0.05) and SCQLD (odds ratio = 4.45; p<0.05). The CNSW region was also 128 more likely to have had trees with buds compared to CQLD, its coastal counterpart in the north (odds ratio = 127.88; p<0.001). It was the large change in the odds for the MAC and CNSW regions in 2011 relative to earlier years that drove the significance of the Region x Year interaction effect (Table 4).
Table 4. Deviance Information Criterion (DIC) for the different binary logistic models used to determine the odds of trees with buds from each region in each year, Z – and p – values for the random factors from the best-fitting model (*). The estimates were obtained by MCMC with a burn in length of 500 and a monitoring chain length of 175 000.

<table>
<thead>
<tr>
<th>Binary logistic model</th>
<th>DIC values</th>
<th>Z values</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept only</td>
<td>2378.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept + Region</td>
<td>2376.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept + Year</td>
<td>2322.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept + Region + Year</td>
<td>2319.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept + Region + Year + Region x Year*</td>
<td>2214.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Factors</td>
<td></td>
<td>1.24</td>
<td>0.107</td>
</tr>
<tr>
<td>Family</td>
<td></td>
<td>3.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Replicate within Family</td>
<td></td>
<td>7.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.4.4.2. Yearly differences in the potential to flower, a region by region assessment

The relatively high rate of budding in the final year (2011) could largely be attributed to two regions with high odds of trees with buds, CNSW and MAC (Figure 5c). The high likelihood of buds on trees from MAC and CNSW regions in 2011 was in contrast to their low odds in 2010 (odds = 0.013 and 0.026, respectively (Figure 5a-c). The MAC region was almost seven times more likely to have trees with buds in 2011 compared to 2009 (odds ratio = 6.93; p<0.001) and 48 times more likely compared to 2010 (odds ratio = 48.61; p<0.001; Figure 5a-c). The CNSW region on the other hand was three times more likely to have trees with buds in 2009 compared to 2010 (odds ratio = 3.04, p<0.001) and nine times more likely to have trees with buds in 2011 compared to 2010 (odds ratio = 9.03; p<0.001; Figure 5a-c). The odds of a tree having buds in the other regions were relatively stable in the three years compared to CNSW and MAC (Figures 5 a-c).

It was also evident that the trees from the local region (RICH, close to the Bonalbo trial) had less within region variation in budding across the three years, than trees from regions that were translocated over greater distances (Figures 5 a-c). The MAC region which was the farthest from the Bonalbo trial (and represented the greatest latitudinal change and transfer from a temperate to subtropical region), had the highest variation in
the odds of trees with buds (confidence interval = 1.92 to 5.92), while the RICH region which is closest (and closest in latitude) to the trial, had the lowest within region variation (confidence interval = 1.67 to 2.61; Figures 5 a-c).

2.4.4.3. **Family level genetic effects**

In 2010 and 2011 there was a tendency whereby the number of families that flowered simultaneously at both sites was higher than expected class size, but chi-square tests on contingency tables were not significant for either year. In 2010, 15 families flowered in at least one site, with 10 of these flowering at both sites, the remaining five flowered only at either Bonalbo or Emu Creek. The deviation in these observed numbers of families flowering at both or either site, from the expected values of 7.5, was not significant ($\chi^2 = 1.667$, $p=0.197$). In 2011, only 5 out of the 50 families flowered at one or more sites, four of these families flowered at both sites, and only one flowered at a single site. Again the disturbance of these values from the expected value of 2.5 for the null hypothesis of no genetic effect was not significant ($\chi^2 = 1.80$; $p=0.180$), nonetheless, the disturbance that was evident, was in a direction that suggested a genetic influence.

2.4.5. **Tree vigour and flowering**

Among the 1855 trees with DBH measurements, there were only 421 trees (22.69%) that had reproductive structures. The mean DBH of trees that were reproductively active (Mean ± SE; 16.15 ± 0.21 cm, n=421) was significantly higher than the mean DBH of trees that were not active during the observation period ((Mean ± SE; 11.85 ± 0.12 cm, n=1434) (t test $p<0.001$).
Figure 5. Regional odds of trees having buds as assessed in November for three years in 2009, 2010 and 2011 at the Bonalbo trial. Lines indicate the extent of the lower and upper 95% confidence interval on the odds value estimates. Different letters indicate significant differences at p<0.05.

Odds of trees having buds analysed by a multi-level logistic regression model in MLWin v2.24 (Rasbash et al 2000).
The mean DBH was also tested against the score of flowering effort. In this study, flowering effort was the total amount of reproductive structure that a tree bore across the three years of study. The overall F test in a univariate ANOVA indicated that there were significant differences among the mean DBH of trees categorised by flowering effort scores (p<0.001; Figure 5a). The mean DBH of trees that did not produce any reproductive structure during the period (mean DBH = 11.84 cm, flowering effort score = 0) was compared with that of the group of trees that were reproductively active. The pairwise comparison showed that the trees that were reproductive active were significantly larger (p<0.001). However, there was no significant differences among the categories of the trees that had made some flowering effort (i.e. among categories 1, 2, 3, 4 and 5; p>0.05; Figure 6a). These results suggest that a threshold size, (i.e. 14.20 cm; Figure 6b) must be reached to permit flowering. But once that size is obtained, tree size is not a good predictor of the frequency of flowering for trees of the age assessed.

![Figure 6. a. Mean diameter at breast height (DBH) of each flowering effort category. b. Regional values for DBH (bars) plotted with mean flowering effort (lines). Significant differences in DBH are indicated by different small letters, and capital letters for mean flowering effort as determined by least significance difference (LSD). Significant differences are indicated by different letters. All comparisons were determined by LSD using SPSS v19 (IBM).](image)

Although not significantly different, CQLD (mean DBH = 13.70 cm) and IQLD (mean DBH = 13.08) tended to have larger diameters than trees from the southern regions i.e. RICH (mean DBH = 12.98 cm), HENSW and MAC (both with mean DBH = 12.14 cm; p>0.05; Figure 6b). SCQLD had a significantly lower mean DBH (11.16 cm) than any
other region (p<0.05; Figure 6b). It also followed that the northern region, IQLD (mean flowering effort = 0.688) had significantly higher mean flowering effort than HENSW (mean flowering effort = 0.359), and MAC (mean flowering effort = 0.168; p<0.05; Figure 6b). CQLD (mean flowering effort = 0.518) had a significantly higher mean flowering effort than MAC (p<0.05) and CQLD also tended to have a higher mean flowering effort than trees the southern CCV regions, although not significantly more so (Figure 6b; p>0.05).
2.5. **Discussion**

The objectives of the study were to describe the timing in the key floral development in planted spotted gum taxa, compare this to patterns in native stands, and determine the degree to which elements of the floral development process are under genetic control. Results indicate that trees need to reach a threshold size before they will become reproductively active. Once past this size, flowering effort would appear to be independent of tree size and both timing and intensity of flowering are affected by genetic and environmental factors. Anthesis time would appear to be less influenced by environmental factors (site and climate) and thus under stronger genetic control than duration of bud development. The tests using common families, although not significant, also tended to support genetic regulation of anthesis at the family level, as members of the same families tended to flower synchronously when planted at two locations in the Bonalbo region of NSW.

A review of these and other key findings are provided below, and attempt to rationalise the flowering patterns observed in CCV in terms of underlying genetic, environmental and interaction factors. The model of floral development for CCV and CM is revisited and refined, based on observations from the present study, and conclude by reviewing gene flow management implications.

2.5.1. **Two flowering races align with geographic and genetic groupings**

In the present study, two flowering races of CCV has been found, each race is a group of regions with distinct flowering time from the other race: an early flowering northern race and a late flowering southern race. The distribution of these flowering races tended to accord with the genetic structuring previously identified using genetic markers. Trees from southeast QLD localities (northern) were distinguished by allele frequency differences at microsatellite marker loci when compared with trees from northern NSW (southern) (Shepherd et al., 2008a; Shepherd et al., 2012). Genetic structure revealed by molecular markers does not necessarily infer adaptive variation, because markers are generally selectively neutral, however, congruent patterns may arise where populations
have undergone historical isolation, facilitating genetic drift at neutral loci and differentiation at adaptive traits. This has been observed in other widespread east coast eucalypts (Butcher et al., 2009; Shepherd et al., 2010) as well as conifers (Li and Adams, 1989).

2.5.2. Northern race flowers earlier

The difference in flowering time in CCV is hypothesized to be adaptive and has evolved in response to environmental factors associated with latitude which affects both vegetative growth and reproduction (Olsson and Ågren, 2002; Hall and Willis, 2006). One possible driver of adaptation is the avoidance of anthesis during high rainfall months to maximise the number of pollinators like bees and other invertebrates (Elzinga et al., 2007). In summer months when there is high rainfall, pollinator activity is reduced. In the Gympie area where mean annual rainfall is 1066 mm, 50% of it falls in the summer months from December to March. Flowering earlier in the season may also be a strategy to avoid damage from herbivores and diseases, which tend to be more abundant in summer. For example, almost half of the crown of a dominant CM was destroyed by cup moth larvae before flowering in the summer months (Pook et al., 1998). Herbivory may result in fewer flowers being produced by trees that flower later in the season than those that undergo anthesis earlier.

2.5.3. Northern flowering race is more fecund and disease resistant

A pronounced trend of higher fecundity and greater vigour of trees from the northern flowering race relative to the southern race may be due to adaptive variation of higher tolerance to QSB (Brawner et al., 2011; Pegg et al., 2011). It has been shown that populations from the higher rainfall region around Gympie (QLD), including the provenances assessed in this study belonging to CQLD, have higher tolerance to QSB relative to the NSW provenances, which leads to higher early plantation productivity of trees from the CQLD region (Brawner et al., 2011). Shoot blight damage caused by QSB has a major effect on the crowns of planted CCV because the lesions and distortion of new shoots, including stems and expanding leaves, can cause a massive
reduction in foliage, or loss of apical dominance in severe cases (Pegg et al., 2011). Damage caused by QSB was recorded in the Bonalbo trial in an earlier study and as observed in other studies, northern populations tended to be more tolerant (Johnson et al., 2009). Trees from CQLD and IQLD were also found to be more vigorous (i.e. had larger DBH) and tended to have greater flowering effort (abundance) than the other regions. It has been shown that larger trees tend to produce more flowers (Ashton, 1975; Pook et al., 1997; Bacles et al., 2009), therefore it is hypothesized that the greater fecundity of the northern flowering race in this study may be attributed to its greater tolerance to QSB during early growth, leading to greater vegetative growth and reproductive capacity, relative to other trees at QSB affected locations. With the trees affected with QSB, it may be possible that reproductive output is compromised in favour of growth and survival, which is expected under the cost of reproduction (Stearns, 1989; Partridge and Harvey 1988; Roff 1992).

Other factors may also contribute to the observed pattern of relatively high flowering levels in the northern flowering race when planted at Bonalbo. There is a well-recognised trend for more profuse and earlier flowering in plantings of trees that are regarded as “off-site” (planting outside the native range), relative to their native environment, which is utilised by tree breeders in choosing sites for seed orchard locations (Eldridge et al., 1994). This effect has been documented for Melaleuca alternifolia in which a NSW north coast population flowers up to 18 months earlier, and more heavily, when planted in southern NSW outside its native environment (Baskorowati et al., 2010). Such an effect may also be evident in CCV translocated from CQLD when planted at Bonalbo but unlike the M. alternifolia example, growth rates of CQLD trees generally exceed native trees, suggesting they are growing vigorously, experiencing less stress (i.e. suffering a lower level of QSB damage).
2.5.4. Genetic control of anthesis is indicated by stability in anthesis time of translocated trees

Despite transfer over two degrees of latitude southward, the northern flowering race maintained its peak anthesis during winter. Records for the northern flowering race in its native range show that trees may flower at any time of the year but the peak period is in winter, which coincided with the observed timing of its peak anthesis at Bonalbo. Similarly, trees from the southern flowering race maintained stability of anthesis period when transferred. Native CCV in northern NSW was recorded to be flowering in the summer to autumn months (Law et al., 2000) which coincided with anthesis of trees from Richmond Range, Wedding Bells and Boundary Creek in this study. This correspondence in anthesis time between translocated trees and trees in their native environments has also been observed in other eucalypts. In a common garden study in Tasmania, the peak anthesis of early flowering genotypes of *E. globulus* subsp. *globulus* from Tasmania was observed to be in late October, while late flowering genotypes originating from the Strzelecki Ranges in Victoria underwent anthesis in mid-November to late December (Jones et al., 2011). In their native environment, the anthesis time of Tasmanian population is from September to November (Barbour et al., 2006) while the Victorian population flowers from mid-November to late December (Sasse et al., 2003). This congruence in anthesis time between translocated trees and native populations of temperate eucalypts has a genetic basis and is supported by various studies (i.e. (Dutkowski and Potts, 1999; Jones et al., 2011). Similarly, the congruence in anthesis time of spotted gums in translocations and in their native environments as observed in this study suggests that there is also a genetic basis to the timing of anthesis in spotted gums.

2.5.5. Plastic regions include some initially below the threshold size for flowering

The region by year interaction effect had the largest impact on the near-term potential to flower in this study, and was largely due to the large increase in the potential to flower in 2011 relative to the two earlier years for trees from the CNSW and MAC regions. For the MAC region, one interpretation of this result may be that trees first attained the threshold size for flowering in 2011. The MAC region had the lowest average DBH
(mean DBH = 12.14 cm; Figure 6b) of any in the study, and the range of tree sizes in this region was below the average DBH of 14.20 cm of for trees which exhibited the lowest level of flowering effort (i.e. 0.168; Figure 6b). Poor growth of MAC may be attributed to its low tolerance to QSB disease, as provenances from this region tended to be rated among the most susceptible in an assessment of the Bonalbo trial at 25 months (Johnson et al., 2009). The years 7-10 is often a critical age in the maturation in eucalypts with many species noted as commencing to flower at about this age (Eldridge et al., 1994; p. 186). It was possible that trees from the MAC region may have suppressed growth relative to trees from other regions but attained threshold size in the final year of observations in 2011, at age 12 years post-planting.

The response of trees from the CNSW region may not be explained in the same way as the MAC region. CNSW was one of the better growing regions and the trees from this region may have exhibited adaptive response to higher rainfall. Trees from CNSW may have reached their threshold size for flowering in earlier years, and the increase in likelihood of trees with buds from CNSW in 2011 might have been a response to the higher rainfall recorded at the trial in 2011 (mean annual rainfall = 1090 mm), compared to previous years (2009 and 2010 had 870 and 885 mm, respectively). The long term MAR for this region is above 1500 mm, hence, provenances from near the coast in the CNSW region may require higher rainfall to promote flowering.

In contrast to the plastic regions, some southern regions, i.e. RICH, HENSW and HRNSW, showed relatively stable and low near-term potentials to flower over the three years and may suggest that many trees from these regions had yet to attain a threshold size to flower. Again, this may be a result of slower growth of trees from these regions due to QSB, because QSB has been shown to impact southern provenances heavily (Johnson et al., 2009; Brawner et al., 2011). Alternatively, these regions may be exhibiting flowering cycles that are longer than one year. It was observed, for example, that CCV in native populations from northern NSW may only flower once every five years (Law et al., 2000). Trees from the RICH region may be considered local to the study site and tended to exhibit flowering pattern similar to the native spotted gums growing near the trial at Bonalbo which had little or no flowering in 2010 and 2011 (Abasolo et al. unpublished data).
The varying responses evident in this study for trees from different regions exemplify the complexity of unravelling the factors contributing to flowering. Despite a genetic contribution to the potential to flower, genetic differences are overlaid with, and interact with, tree health, age, year and seasonal environmental factors. Many of these environmental/tree health factors were confounded in this study making it difficult to identify those contributing to variation in flowering. This highlights the challenges in studying factors controlling flowering in trees as well as the difficulties faced by forest managers where they may need to predict flowering patterns in the forests they manage.

2.5.6. Implications for risk assessment of gene flow from plantings of CCV

The identification of genetically determined partial asynchrony in peak flowering of CCV suggests that the likelihood of gene flow may be lower than previously thought in situations where trees from the northern race are planted within the range of the southern geographic race. The likelihood may be further moderated by the observation of low levels of pollen production in plantations of young age (i.e. less than 14 years old) relative to nearby native stands, so that the source to sink ratio may be low.

While the likelihood of gene flow may be low, the impact may be higher than previously thought. This is due to the increasing recognition that within CCV there may be distinct geographic races with adaptive differences. For example, there is accumulated evidence that the northern population is distinct from the southern population in microsatellite data (Shepherd et al., 2008) and for potentially important adaptive traits like disease tolerance (Johnson et al., 2009; Pegg et al., 2011), frost tolerance (Larmour et al., 2000), terpene chemistry (Asante et al., 2001) and peak flowering time (this study). If these traits and genetic differences are indeed important for survival and fitness in the local environment, then there may be ecological and evolutionary consequences for the transfer of mal-adapted alleles from plantings. This suggests that the impact of gene flow between the CCV races may be greater than previously thought.
The recognition of two geographic races (exhibiting both adaptive differences in anthesis time and disease tolerance) within CCV presents a challenge for forest managers where trees from the northern race are planted in the south of CCV’s natural range. This study revealed two factors that mitigate gene flow between these races: partial asynchrony in anthesis time and the overall low levels of flowering in planted stands of CCV. These factors may reduce the overall risk profile of such planting that otherwise may be high due to biological factors such as high cross-compatibility due to a lack of many pre- (structural incompatibilities) and post-zygotic (genetic compatibility) reproductive barriers.

Due to high stocking rate at the study sites, flower abundance might have been underestimated. Thus, further study is required to include the influence of stocking rate on flower abundance. Although the observed overall abundance of flowering in the 12-year old trial was low (from 8 to 12% of trees flowered in the November census, with relatively few flowers compared to native trees), it has been shown that flowering increases with increasing DBH and suggests that flowering may increase in stands that are managed for solid wood production that are subject to rotations of 20 years or more, and are routinely thinned to allow greater diameter growth and greater light penetration into the crown. A separate study in a wide age range of planted CCV (up to 14 years old) showed that although flowering increases with age, the abundance of flowers produced by a tree is still relatively low compared to large trees in native populations (Abasolo et al. in prep). Furthermore, there has been a trend in recent years towards planting of the more fecund Woondum provenance, therefore, a pure planting of this provenance is expected to flower more heavily than the trials planted with many provenances studied here.

Long term flowering studies are also required across a range of sites to better understand the impact of environment on flowering in CCV. The yearly variation in anthesis time in both native and planted populations may impact on the degree of synchrony between races in any one year. This study showed that there was limited synchrony between the two flowering races of spotted gums. For example, the highest inter-flowering race synchrony observed was between Woondum and Wedding Bells (synchrony index = 0.29; moderate), but this was based on only two years of data. Thus, the limited data may not represent the overall environmental spectra that the trees are
exposed to. Further, the low flowering in the trials limited the observations to the few trees that flowered within the observation period, and they may have represented a subset of the trees that are deployed in commercial plantations. Planting of the early flowering northern race near native stands of southern race may moderate gene flow if it reduces synchrony between the two races. This knowledge can be used by plantation managers who seek to minimise hybridisation between genetically distinct populations of CCV.

2.6. Acknowledgment

This study is a part of the PhD research, “Assessing and managing gene flow from locally exotic spotted gums”, of M Abasolo supported by the CRC Forestry and the Southern Cross University International Postgraduate Research Scholarship. The authors thank the Forest NSW for granting access to the trial sites, G Ablett for assistance in field work, and G Luker for preparing the map in Figure 1.
Chapter 3. Identification of intersectional *Corymbia* hybrids based on seedling morphology improves with parental divergence

This chapter has been published as:


Contributions to this chapter

**Experimental design:** Myralyn Abasolo, David J Lee and Mervyn Shepherd designed the experiment.

**Data Collection:** Myralyn Abasolo collected the data with assistance from Tracey Menzies and John Oostenbrink.

**Statistical analysis:** Myralyn Abasolo performed the statistical analyses.

**Writing:** Myralyn Abasolo Mervyn Shepherd and David J Lee wrote the manuscript.
3.1. Abstract

Differences in morphology have provided a basis for detecting natural interspecific hybridisation in forest trees for decades but have come to prominence again more recently as a means for directly measuring gene flow from planted forests. Here the utility of seedling morphology is examined for hybrid discrimination in three hybrid groups relevant to the monitoring of gene flow from plantings of *Corymbia* (L.D. Pryor & L.A.S. Johnson ex Brooker) taxa in subtropical Australia. Thirty leaf and stem characters were assessed on 907 eight-month old seedlings from four parental and six hybrid taxa grown in a common garden. Outbred F<sub>1</sub> hybrids between spotted gums (*C. citriodora* subspecies *variegata*, *C. citriodora* subspecies *citriodora* and *C. henryi*) tended to more closely resemble their maternal *C. torelliana* parent and the most discriminating characters were the ratio of blade length to maximum perpendicular width, the presence or absence of a lignotuber, and specific leaf weight. Assignment of individuals into genealogical classes based on a multivariate model limited to a set of the more discriminating and independent characters was highest in the hybrid group where parental taxa were genetically most divergent. Overall power to resolve among outbred F<sub>1</sub> hybrids from both parental taxa was low to moderate, but this may not be a limitation to its likely major application of identifying hybrids in seedlots from native spotted gum stands. Advanced generation hybrids (outbred F<sub>2</sub> and outbred backcrosses) were more difficult to resolve reliably due to the higher variances of hybrid taxa and the tendency of backcrosses to resemble their recurrent parents. Visual assessments of seedling morphology may provide a filter allowing screening of the large numbers needed to monitor gene flow, but will need to be combined with other hybrid detection methods to ensure hybrids are detected.
3.2. Introduction

Worldwide, forest plantations are being established on an enormous scale to meet demands for wood products. The annual rate of increase in forest plantation establishment was five million hectares per year over the 1990-2000 period (FAO, 2010) which leads to a continuous translocation of species (Allendorf et al., 2001). Hybridisation of the translocated species with the native populations may decrease genetic variation, or alter population structure and composition of the native species (Arnold et al., 1999). In general such translocations are unmonitored (Laikre et al., 2010), but in Australia Potts et al (2003) outlined the need for monitoring and assessment of gene flow from planted exotic hardwoods nearly a decade ago. As a result, monitoring systems are now in place in temperate Australia where hardwoods, typically eucalypts, have been planted extensively (Gavran and Parsons, 2010).

Long term monitoring of gene flow at plantation boundaries involves periodic assessment of the rates of hybridisation, a process that is increasingly integral to risk management for planting of trees for forestry or restoration purposes (Potts et al., 2003; Barbour et al., 2010; Byrne et al., 2011). Risk management usually requires evaluating taxonomic, biological and geographic criteria within a decision tree framework to quantify the risk associated with each planting (Byrne et al., 2011). The major biological criteria include the likelihood of hybridisation and the rate that hybrids established (Potts et al., 2003; Byrne et al., 2011). Often these parameters must be initially assumed, and management strategies are adaptive, putting in place monitoring processes, and utilising information to inform management responses iteratively once evidence on hybridisation rates comes to hand (Pers. Comm. B Potts). Directly measuring hybridisation rates has confirmed expectations, for example, that hybrid frequency tends to increase in populations with high flowering synchrony, and that pollen tends to move from a more abundant pollen source to the populations with less pollen, i.e. fewer flowering trees, (Field et al., 2008; Lepais et al., 2009; Field et al., 2011).
In temperate Australia, a system based on the visual identification of hybrids using morphological differences has been effective in large-scale screening to detect hybrid eucalypt seedling and provide direct measurements of gene flow (Potts and Reid, 1985; Potts and Reid, 1988; Barbour et al., 2003). The use of morphological markers has the advantage above other methods in these situations because many thousands of individuals may be screened easily and cheaply to detect even relatively low levels of gene flow provided at least a few reliable discriminating characters can be identified (Barbour et al., 2002; Barbour et al., 2003). The effect of environment on morphological traits may obscure accurate identification by morphology (Rieseberg 1995; McDade 1990) however, traits that are highly heritable or common garden conditions may be required to use morphology for identification (Barbour et al., 2003; Field et al., 2009). If necessary, where high surety is required, morphological markers can also be combined with molecular markers like isozymes or microsatellites, to achieve higher confidence in hybrid detection (Barbour et al., 2003; Field et al., 2009).

Study of hybrid morphology also has ecological significance. In many hybrids, including eucalypts, $F_1$ hybrid morphology is intermediate relative to its parents, but sometimes the hybrid may resemble one parent, or exceed the range of its parents (e.g. (Barbour et al., 2003; Meddings et al., 2003; Rieseberg et al., 1999; Edmands 2002). Advanced generation hybrids are more difficult to distinguish by morphology because of trait segregation (Anderson and Hubricht 1938). Although the fitness of hybrids can be depressed, surviving transgressive individuals may have greater fitness in marginal or novel environments than their parents (Arnold and Hodges, 1995; Rosenthal et al., 2002; Rieseberg et al., 2007). This may be important in extending adaptability of a taxon and creating a new ecological space to support hybrid speciation (Rosenthal et al., 2002), increase its weedy potential, or allow it to act as a bridge for gene flow (Whitney et al., 2006; Whitney et al., 2010).

In subtropical Australia, spotted gums (Genus Corymbia Section Maculatae) and their hybrids are the principal taxa of interest for hardwood plantation forestry (Nichols et al., 2010). Plantations are often established in close proximity to native forests where spotted gums and other bloodwoods (Corymbia sp) are endemic. As these plantations mature, management of gene flow is an increasing concern for forest managers (Barbour et al., 2008).
One concern is the planting of locally exotic spotted gum provenances. For example, northern provenances of *Corymbia citriodora* subsp. *variegata* (CCV) from around Gympie (QLD) are widely planted in northern NSW where both *Corymbia henryi* (CH) and genetically distinct populations of CCV occur (Shepherd *et al.*, 2008a). Natural interspecific hybridisation between CCV and CH occurs in regions of sympatry and northern and southern provenances of CCV are inter-fertile and have flowering synchrony to some extent (Hill and Johnson, 1995; Ochieng *et al.*, 2008; Ochieng *et al.*, 2010). *Corymbia henryi* and CCV are distinct in their morphology and although they co-occur at some locations, they tend to occupy different positions in the landscape, suggesting adaptive differentiation (Ochieng *et al.*, 2010). Southern and northern provenances of CCV also show differential adaptability to frost (Larmour *et al.*, 2000), tolerance to pests and diseases (Johnson *et al.*, 2009), and leaf mass area, an ecologically important trait due to its role in tolerance to water stress (Fonseca *et al.*, 2000). This suggests that if gene flow occurs between exotic provenances of CCV when planted in the south of its range, non-adapted genes may be introduced in local CCV or CH populations.

The second concern is with the planting of spotted gum - *Corymbia torelliana* (CT) hybrids. These hybrids have proved to have many benefits for plantation timber production in the subtropics (Lee, 2007). The concern with the use of hybrids for plantations is that backcrossing and transfer of weedy attributes from CT to native spotted gums may occur. *Corymbia torelliana* is a native of north Queensland but has been planted widely throughout the subtropics where it is regarded as an environmental weed in some shires of northern NSW and southern QLD (Hill and Johnson, 1995; NCWAC, 2003; Kingston *et al.*, 2004). *Corymbia torelliana* naturally hybridises with spotted gums in north QLD and can be easily crossed with spotted gums in controlled pollinations (Hill and Johnson, 1995; Dickinson *et al.*, 2010). The ability of CT to invade disturbed environments and the adaptability and vigour of its hybrids (Lee *et al.*, 2005; Lee, 2007; Lee *et al.*, 2009; Nahrung *et al.*, 2011) raise concerns about the translocation of CT and use of hybrids in plantations (Barbour *et al.*, 2008).

In this study, the potential of morphology was evaluated to resolve among genealogical classes by a comparative analysis of seedling attributes in *Corymbia* hybrids with their
parental taxa. Seedlings of up to 10 families in each taxon, from usually two provenances, were grown in a common garden and assessed for a range of leaf and stem characters, and the presence or absence of lignotubers. Three hybrid groups (either Corymbia citriodora subsp. citriodora (CCC, CCV or CH) of interest to plantings of CT or its hybrid in three situations in subtropical Australia were studied. In the context of gene flow from plantations, the transfer of pollen from planted sources onto native spotted gums is the cross direction of most concern. However, crosses with a spotted gum maternal parent were not available, hence the study was limited to crosses in the reciprocal direction, i.e. with CT maternal parents. A recent study of seed germination showed that controlled crosses with CCV maternal parents are as viable as the reciprocal crosses (Dickinson et al., 2012), strengthening evidence for the possibility for gene flow in both directions. For the CT-CCV hybrid group, advanced generation (outbred F$_2$ and backcross) hybrids were also examined in addition to the F$_1$.

In this chapter, the following questions were specifically addressed: 1) How accurate could hybrids be identified using morphology and what are the distinguishing features that can be used to identify them?; 2) Do assignment rates differ among different F$_1$ hybrid groups?; 3) Are advanced generation hybrids identifiable by morphology?; 4) Are the advanced generation hybrids more vigorous than their parents and the F$_1$ hybrids?; 5) Are there transgressive traits in the hybrids that may have ecological significance?; 6) What are the implications of these transgressive traits on the planting of hybrids?; 7) How can morphological markers be used for direct gene flow assessment in scenarios of planting pure and hybrid Corymbia?
3.3. Materials and methods

3.3.1. Materials

The materials used in this study were from two sections of the Genus *Corymbia*. The first was from Section *Maculatae* (i.e. the spotted gums) represented by three of the four recognised taxa (CCV, CCC and CH; *Corymbia maculata* was not studied) and the second was from section *Torelliana* (CT) (Hill and Johnson, 1995; Parra-O. *et al*., 2009).

A total of 1,834 seeds from 11 parental or hybrid taxa were sown in March 2010 at the Department of Employment, Economic Development and Innovation (DEEDI) glasshouse facilities at Gympie, QLD (Table 1). The parental taxa, CCC, CCV, CH, and CT were represented by open-pollinated families from different provenances: CCC from Kirrima, QLD (17°39’S, 146°5’E) and Yeppoon, QLD (23°07’S, 150°44’E); CCV from Richmond Range, NSW (28°40’S, 152°42’E) and Woondum, QLD (26°15’S, 152°49’E); CH from Lockyer, QLD (27°30’S, 152°04’E), Nerang, QLD (27°59’S, 153°20’E) and Myrtle, NSW (29°08’S, 152°05’E); CT from Helenvale, North Queensland (15°43’S, 145°14’E) and Gympie Landrace (no latitude and longitude available but molecular studies attribute this material to be of Kuranda origin (McVey, 2004). Because the parental taxa were open-pollinated, the possibility of inadvertently including hybrids cannot be ruled out especially for species that occur in sympatry. Inbreds may likewise have been included in the parental taxa. The hybrid taxa studied were from controlled pollinations and consisted of outbred F1 interspecific hybrids, CTxCH (F1H), CTxCCV (F1V), CTxCCC (F1C), outbred backcrosses, CTxCCVxCCV (BCV), CTxCCVxCC (BCH), CTxCCVxCF (BTC) and an outbred F2 (CTxCCV) x (CTxCCV) (F2V). Controlled-pollinated seed was obtained by using the “one-stop pollination” technique of Harbard *et al*., (1999), modified such that the entire inflorescence was bagged instead of covering the style with a tubing. A mix of pollen (pollen polymix) from each individual paternal parent was used for crossing (Dickinson *et al*., 2012; Dickinson *et al*., 2010). Because pollen viability can last for only 4-5 days (Geoff Dickinson, personal communication), the pollen was stored in gel capsules at 4°C until required, to preserve the viability (Dickinson *et al*., 2012; Dickinson *et al*., 2010).
Each parental taxon was represented by 10 families. Where possible, parental control populations were composed of open-pollinated seedlots of trees from the same provenance as the parents used for hybrid crosses but they were not the same individuals (Table 1). In the case of CT, all the maternal parents used in hybrid crosses were from the Gympie Landrace, hence for the CT parental control, open-pollinated seedlots were chosen from individuals in the Gympie Landrace. For the spotted gum parental taxa, open-pollinated seedlots were obtained from individuals from the same provenance as the individuals used in the hybrid crosses or a nearby provenance where possible.

The numbers of parents involved in the hybrid crosses differed in each hybrid group, for example the $F_{1H}$ had seven maternal and six pollen parents, the $F_{1V}$ had 10 maternal and 11 pollen parents, and the $F_{1C}$ had five maternal and three pollen parents. The number of families in the hybrid taxa was lower and more variable as fewer hybrid families were available for study (Table 1).

### 3.3.2. Glasshouse methods and experimental design

Seeds were initially sown in 8x5-celled hyco trays (70 ml) in the glasshouse in March 2010 using a potting media consisted of 50% pine bark fines (0-10 mm), 25% pine bark peat, 25% coarse perlite, a mix of 12-14 month slow release Osmocote fertilisers at a rate of 4 kg m$^{-3}$-1, gypsum (1kg m$^{-3}$-1), Micromax (1kg m$^{-3}$-1) and a granular wetting agent Hydroflo2 (1kg m$^{-3}$-1). In July 2010, approximately four months after sowing, seedlings from each family/treatment were transplanted into individual pots with a 13.5 cm diameter (1.25l) with potting mix as described above. Survival was variable, but where possible, 15 seedlings were retained for each family, otherwise all remaining seedlings in a family were retained. The $BC_H$ was not included in further analysis because all seedlings died. There were 929 seedlings that were transplanted and these were assessed at four months of age.

The experiment was established and analysed using an Incomplete Block Design (IBD) with five replicates and 45 blocks. Using family as a treatment, individuals from a family were randomly assigned to replicates and blocks by the computer software
CycDesigN Version 1.2 (Whitaker et al., 2006) and seedlings in individual pots were rearranged into blocks on the nursery bench. Each replicate contained nine blocks and each block was arranged into four rows and six columns. Each row contained two families (treatments) so that each block had eight treatments. A treatment was represented by a maximum of three seedlings in each replicate where possible.

3.3.3. Seedling assessment

3.3.3.1. Germination, survival and vigour

Average percentage germination on a family basis was noted as the proportion of seedlings that survived 30 days after sowing. Average survival was computed on a family basis as the proportion of seedlings that survived to age eight months. Total height was also recorded for each seedling at eight months after planting as part of the morphological assessment. In total, 907 seedlings were assessed in November 2010 (Table 1).

3.3.3.2. Morphology

A total of 30 characters that have been previously used for taxonomic classification or in previous studies of morphological differentiation in the group were assessed (Hill and Johnson, 1995; Larmour et al., 2000; Nahrung et al., 2009; Ochieng, 2009). The characters assessed included leaf shape and dimensions, stem colours and the presence of hairs and lignotubers, on all the individuals in July 2010 (four months after sowing) and again in November 2010 (eight-month old; See Table 2 for a full list of characters and how they were assessed). Characters were either assessed visually in the glasshouse or a leaf from the 8th node (four-month old) and 10th node (eight-month old) of each seedling taken for further analysis by scanning with WINFOLIA image analyser (Regent, Quebec, Canada) to allow estimation of leaf dimensions and calculation of ratios.
The leaf from the 10th node was also used for specific leaf weight (SLW) measurement of eight-month old seedlings. The following methods were used for the calculation of SLW. Where possible, 10 leaf discs were punched out from fresh leaves collected at Node 10 from each seedling with a paper punch and oven dried overnight at 108±3°C. Each disc had an area of 28.27 mm$^2$ and therefore specific leaf weight could be determined by the formula:

$$SLW = \frac{ODW}{(\text{no. of leaf discs} \times \text{leaf disc area})}$$

Where:
- $SLW$ = specific leaf weight (mg mm$^{-2}$-1)
- $ODW$ = oven dry weight (mg)
- Leaf disc area (28.27 mm$^2$)
### Table 1. List of the number of families, provenances and locations of *Corymbia* species and hybrids and number of seeds sown and seedlings assessed in this study.

<table>
<thead>
<tr>
<th>Taxon Code</th>
<th>Taxon Code</th>
<th>No. of families</th>
<th>Provenance</th>
<th>No. of seeds sown</th>
<th>No. of seedlings assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCV</td>
<td><em>Corymbia citriodora subsp citriodora</em></td>
<td>4</td>
<td>Richmond Range</td>
<td>91</td>
<td>60</td>
</tr>
<tr>
<td>CCV</td>
<td><em>Corymbia citriodora subsp variegata</em></td>
<td>6</td>
<td>Woondum</td>
<td>151</td>
<td>89</td>
</tr>
<tr>
<td>CH</td>
<td><em>Corymbia henryi</em></td>
<td>3</td>
<td>Myrtle</td>
<td>69</td>
<td>30</td>
</tr>
<tr>
<td>CH</td>
<td><em>Corymbia henryi</em></td>
<td>3</td>
<td>Lockyer</td>
<td>67</td>
<td>45</td>
</tr>
<tr>
<td>CH</td>
<td><em>Corymbia henryi</em></td>
<td>4</td>
<td>Nerang</td>
<td>93</td>
<td>59</td>
</tr>
<tr>
<td>CCC</td>
<td><em>Corymbia citriodora subsp citriodora</em></td>
<td>5</td>
<td>Yeppoon</td>
<td>141</td>
<td>55</td>
</tr>
<tr>
<td>CT</td>
<td><em>Corymbia torelliana</em></td>
<td>5</td>
<td>Kirrima</td>
<td>139</td>
<td>71</td>
</tr>
<tr>
<td>CT</td>
<td><em>Corymbia torelliana</em></td>
<td>5</td>
<td>Helensvale (NQ)</td>
<td>121</td>
<td>75</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>3</td>
<td>Woondum</td>
<td>66</td>
<td>44</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>1</td>
<td>Home</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>1</td>
<td>Brooyar</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>1</td>
<td>Wolvi</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>1</td>
<td>Richmond Range</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>1</td>
<td>Kangaroo River</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>2</td>
<td>Grange</td>
<td>53</td>
<td>27</td>
</tr>
<tr>
<td>CTxCCV</td>
<td>1V</td>
<td>2</td>
<td>Unknown</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>CTxCH</td>
<td>1H</td>
<td>2</td>
<td>Bom-bom</td>
<td>53</td>
<td>20</td>
</tr>
<tr>
<td>CTxCH</td>
<td>1H</td>
<td>2</td>
<td>Grafton</td>
<td>41</td>
<td>16</td>
</tr>
<tr>
<td>CTxCH</td>
<td>1H</td>
<td>1</td>
<td>Myrtle</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td>CTxCH</td>
<td>1H</td>
<td>2</td>
<td>Unknown</td>
<td>52</td>
<td>26</td>
</tr>
<tr>
<td>CTxCH</td>
<td>1H</td>
<td>1</td>
<td>Devines</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>CTXCCC</td>
<td>1C</td>
<td>5</td>
<td>Unknown</td>
<td>107</td>
<td>59</td>
</tr>
<tr>
<td>F2v</td>
<td>(CTxCCV) x (CTxCCV)</td>
<td>2</td>
<td>Unknown</td>
<td>49</td>
<td>5</td>
</tr>
<tr>
<td>BCv</td>
<td>CTxCCVXCCV</td>
<td>3</td>
<td>Woondum</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>BCt</td>
<td>CTxCCVXCT</td>
<td>4</td>
<td>Kuranda</td>
<td>70</td>
<td>17</td>
</tr>
<tr>
<td>BCt</td>
<td>CTxCCVXCT</td>
<td>2</td>
<td>Unknown</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>BCH</td>
<td>CTxCCVXCH</td>
<td>4</td>
<td>Unknown</td>
<td>51</td>
<td>0</td>
</tr>
</tbody>
</table>

| Total      | 80          | 1834          | 907         |

*aCrosses = maternal parent x paternal parent. If provenance is provided in this case it is the provenance of the paternal parent.*

Only the data from the eight-month old seedlings are presented. Preliminary testing of data from the assessment of four-month old seedlings using canonical variate analysis of all characters (see below) indicated hybrids were not resolvable from the maternal parent, CT, whereas this initial evaluation of the data from eight-month old seedlings
were much more promising. Poor resolution of hybrids from the maternal parent at age four months was due to the sharing of the “red hair” characteristic and the lack of development of other, later expressed differences, such as with BLDL:MPW, SLW, and development of the lignotuber.

### 3.3.4. Statistical analysis

#### 3.3.4.1. Germination, survival and vigour

Comparisons for germination, survival and vigour (height) were initially made across the entire experiment. A non-parametric test Kruskal-Wallis Analysis of variance (ANOVA) was used to test for differences in taxa mean ranks in both germination and survival. Difference among taxa for height was assessed using standard ANOVA. Post hoc comparison of taxa in each hybrid group was carried out using Mann Whitney U tests or least significant squares (LSD) for Kruskal-Wallis test and ANOVA, respectively. All these analyses were performed using SPSS v19 (IBM SPSS Statistics, IBM Corporation, NY).

#### 3.3.4.2. Univariate analyses of morphology

For each taxon, the means and standard deviations (SD) of the 26 quantitative characters were obtained (See Appendix 2). Each character was tested for normality and homogeneity of variance at the level of genealogical classes (spotted gum parents, CT and hybrids) and transformed to achieve normality if necessary (Table 2).

Each genealogical class was assigned to a hybrid group, where a hybrid group is composed of the parental taxa, CT and spotted gums, and their corresponding hybrid or hybrids. The purpose of the groupings was to compare the different hybrid classes ($F_1$, BCv, BCt, $F_2$) with their corresponding parental taxa in terms of the different characters in succeeding analyses. There were four hybrid groups: 1) CT-CH group was composed
### Table 2. Characters measured on eight-month old *Corymbia* seedlings with character descriptions and transformations (Trans.)

<table>
<thead>
<tr>
<th>Character</th>
<th>Abbreviation</th>
<th>Description</th>
<th>Trans</th>
<th>Wald/d.f.</th>
<th>CT -CCV-</th>
<th>CT -CCV-</th>
<th>CT -CCC hybrids</th>
<th>CT -CH hybrids</th>
<th>CH and CCV between provenances</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QUANTITATIVE CHARACTERS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CT -all</td>
<td>CT -F_2</td>
<td>CT hybrids</td>
<td>CT hybrids</td>
<td></td>
</tr>
<tr>
<td>Internode length</td>
<td>(INTL10)</td>
<td>the distance between node 9 and node 10 measured with a standard ruler (mm)</td>
<td>none</td>
<td></td>
<td>3.44 **</td>
<td>8.12***</td>
<td>7.65***</td>
<td>14.50***</td>
<td>0.61 ^m</td>
</tr>
<tr>
<td>No of branches</td>
<td>(NBRN)</td>
<td>count of the branches from the main stem</td>
<td>none</td>
<td></td>
<td>38.71***</td>
<td>91.31***</td>
<td>1.24 ^m</td>
<td>63.38***</td>
<td>5.97***</td>
</tr>
<tr>
<td>Diameter at node 10</td>
<td>(DIA10)</td>
<td>measured below Node 10 with a calliper (mm)</td>
<td>none</td>
<td></td>
<td>12.06***</td>
<td>28.74***</td>
<td>44.53***</td>
<td>44.27***</td>
<td>4.65***</td>
</tr>
<tr>
<td>Leaf thickness</td>
<td>(LFTCKNS)</td>
<td>measured at Node 10 leaf with a calliper (mm)</td>
<td>none</td>
<td></td>
<td>1.14 ^n</td>
<td>1.32 ^n</td>
<td>2.77 ^n</td>
<td>9.88***</td>
<td>8.25***</td>
</tr>
<tr>
<td>Vertical length</td>
<td>(VERL)</td>
<td>the total length of the leaf at Node 10 including the petirole (cm)</td>
<td>log</td>
<td></td>
<td>1.72</td>
<td>3.34^*</td>
<td>0.63 ^m</td>
<td>15.09***</td>
<td>6.82***</td>
</tr>
<tr>
<td>Petiole length</td>
<td>(PETL)</td>
<td>measured with a standard ruler (mm)</td>
<td>log</td>
<td></td>
<td>0.22</td>
<td>0.11^*</td>
<td>2.36 ^m</td>
<td>2.30 ^m</td>
<td>351.15***</td>
</tr>
<tr>
<td>Specific leaf weight</td>
<td>SLW</td>
<td>see text for description (mg/mm²)</td>
<td>sqrt</td>
<td></td>
<td>4.42***</td>
<td>7.85***</td>
<td>5.03**</td>
<td>8.94***</td>
<td>3.63**</td>
</tr>
<tr>
<td>Total height</td>
<td>THGT</td>
<td>seedling height from the base to the tip of the stem (cm)</td>
<td>sqrt</td>
<td></td>
<td>11.25***</td>
<td>22.56***</td>
<td>0.45 ^m</td>
<td>75.60***</td>
<td>9.93***</td>
</tr>
<tr>
<td>Scan from WINFOLIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form Coefficient</td>
<td>(FORM)</td>
<td>4πA/P^2 where A= leaf area, P = leaf perimeter; value is between 0 and 1,</td>
<td>none</td>
<td></td>
<td>21.52***</td>
<td>43.44</td>
<td>570.63***</td>
<td>31.87***</td>
<td>6.77***</td>
</tr>
<tr>
<td>Angle of leaf tip1</td>
<td>(AGLTP1)</td>
<td>being a perfect circle and 0 a filiform object</td>
<td>none</td>
<td></td>
<td>3.79**</td>
<td>0.58^*</td>
<td>39.2***</td>
<td>16.23***</td>
<td>4.79**</td>
</tr>
<tr>
<td>Angle of leaf tip2</td>
<td>(AGLTP2)</td>
<td>Angle formed at 97% of the blade length (°)</td>
<td>none</td>
<td></td>
<td>2.47</td>
<td>0.52^*</td>
<td>56.86***</td>
<td>14.19***</td>
<td>4.72**</td>
</tr>
<tr>
<td>Leaf width max</td>
<td>(LLWM)</td>
<td>length from the leaf base to the maximum perp width (cm)</td>
<td>none</td>
<td></td>
<td>4.87 ***</td>
<td>10.54***</td>
<td>11.86***</td>
<td>17.59***</td>
<td>3.43**</td>
</tr>
<tr>
<td>Lobe angle1</td>
<td>(LOBE1)</td>
<td>Angle formed from the leaf base to 3% of the blade length (°)</td>
<td>none</td>
<td></td>
<td>38.64***</td>
<td>82.25***</td>
<td>215.99***</td>
<td>33.016***</td>
<td>5.91***</td>
</tr>
<tr>
<td>Lobe angle2</td>
<td>(LOBE2)</td>
<td>Angle formed from the leaf base to 5% of the blade length(°)</td>
<td>none</td>
<td></td>
<td>61.59***</td>
<td>135.07***</td>
<td>415.58***</td>
<td>58.13***</td>
<td>10.26***</td>
</tr>
<tr>
<td>Blade length</td>
<td>(BLDL)</td>
<td>Length from the leaf base to the leaf tip (cm)</td>
<td>sqrt</td>
<td></td>
<td>2.50</td>
<td>5.64**</td>
<td>2.05^m</td>
<td>21.16***</td>
<td>8.04***</td>
</tr>
<tr>
<td>Max Perp Width</td>
<td>(MPW)</td>
<td>widest width of the leaf (cm)</td>
<td>sqrt</td>
<td></td>
<td>14.93***</td>
<td>31.38***</td>
<td>131.51***</td>
<td>0.41 ^m</td>
<td>18.04***</td>
</tr>
<tr>
<td>Aspect ration</td>
<td>(ASRA)</td>
<td>Ratio of blade width to the blade length</td>
<td>sqrt</td>
<td></td>
<td>1.7</td>
<td>22.72***</td>
<td>58.52***</td>
<td>8.05***</td>
<td>5.31***</td>
</tr>
<tr>
<td>Perpendicular Width1</td>
<td>(PRPWD1)</td>
<td>Width measured perpendicular to the length at Position 1 (positions pre-defined) (cm)</td>
<td>sqrt</td>
<td></td>
<td>4.05***</td>
<td>2.38^*</td>
<td>46.01***</td>
<td>8.99***</td>
<td>3.65**</td>
</tr>
<tr>
<td>Perpendicular Width2</td>
<td>(PRPWD2)</td>
<td>Width measured perpendicular to the length at Position 2 (positions pre-defined) (cm)</td>
<td>sqrt</td>
<td></td>
<td>3.23**</td>
<td>0.15^*</td>
<td>29.37***</td>
<td>8.22***</td>
<td>2.34 ^m</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.005, ***p<0.001
of CT, CH and F_{1H}; 2) CT-CCC group- CT, CCC, F_{1C}; 3) CT-CCV-F_{1} – CT, CCV, F_{1V}; and 4) CT-CCV-all – CT, CCV, F_{1V}, BC_{T}, BC_{V} and F_{2V}. In addition to the four hybrid groups, CCV-CH group was formed by the two spotted gum taxa, CCV and CH.

In each of the hybrid groups and the CCV-CH, a two-step univariate analysis was used to determine which among the 30 characters differentiated hybrids from their parental taxa and between CCV and CH. First, a restricted maximum likelihood (REML) analysis which handles unbalanced designs was carried out in GENSTAT Discovery Edition 3 (Buysse et al., 2007) treating taxon and block as fixed and random effects respectively. The significance of taxon effect was tested with the default Wald statistic and chi-squared test provided by GENSTAT. Second, post hoc tests, either Mann-Whitney U or LSD, depending on whether the character was treated as a qualitative or quantitative character, were used to test for pairwise differences among hybrid groups. This was done for all characters where the Wald statistics was significant in step one. Post hoc testing was conducted using SPSS v19 (IBM SPSS Statistics, IBM Corporation, NY). When a hybrid was significantly different from at least one parent for a character, that particular character was considered important to identify hybrids in a univariate analysis.

Characters were also classified as intermediate or transgressive. Transgression was defined on a population rather than cross-type level i.e. where the mean of hybrid taxon exceeded the parental population means either in a positive or negative direction (e.g. Rosenthal et al., 2002). Independent t-tests were conducted on the means of each genealogical class for the set of characters that had significant Wald values by univariate analysis.

An independent t-test was conducted on each of the three hybrid groups with the parents and the F_{1} only. The type of character (intermediate or transgressive) was the independent variable and the Wald/d.f. was the dependent variable. The equality of variances was also tested using Levene’s test of homogeneity of variances. These analyses were performed using the independent t-test sub-option on SPSS v19 (IBM SPSS Statistics, IBM Corporation, NY).
Table 3. Morphological characters used in the canonical variates analysis and the
standardized coefficients (CV1 and CV2) from a discriminant analysis used to
separate each taxon in the different spotted gum hybrid groups.

<table>
<thead>
<tr>
<th>Character</th>
<th>CT-CCV-ALL</th>
<th>CT-CCV-F1</th>
<th>CT-CH</th>
<th>CT-CCC</th>
<th>CCV-CH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CV1</td>
<td>CV2</td>
<td>CV1</td>
<td>CV2</td>
<td>CV1</td>
</tr>
<tr>
<td>LIGNO</td>
<td>3.72</td>
<td>1.27</td>
<td>3.74</td>
<td>1.57</td>
<td>2.78</td>
</tr>
<tr>
<td>BLDL:MPW</td>
<td>4.91</td>
<td>0.62</td>
<td>4.65</td>
<td>0.04</td>
<td>3.18</td>
</tr>
<tr>
<td>SLW</td>
<td>0.02</td>
<td>2.05</td>
<td>1.14</td>
<td>-0.20</td>
<td>3.15</td>
</tr>
<tr>
<td>STMHRCLR</td>
<td>-0.11</td>
<td>0.11</td>
<td>-0.11</td>
<td>-0.10</td>
<td>-0.07</td>
</tr>
<tr>
<td>TPCLR</td>
<td>0.56</td>
<td>0.29</td>
<td>0.60</td>
<td>-0.35</td>
<td>0.50</td>
</tr>
<tr>
<td>DIA10</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>-0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STHR</td>
<td>-0.22</td>
<td>0.32</td>
<td>-0.08</td>
<td>-0.40</td>
<td>-1.71</td>
</tr>
<tr>
<td>NBRN</td>
<td>0.09</td>
<td>0.25</td>
<td>-0.35</td>
<td>-0.66</td>
<td>0.09</td>
</tr>
<tr>
<td>THGT</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.07</td>
<td>-</td>
</tr>
<tr>
<td>MPW:PRPWD2</td>
<td>-0.63</td>
<td>1.35</td>
<td>-0.35</td>
<td>-0.66</td>
<td>-</td>
</tr>
<tr>
<td>INTL10</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>PETL</td>
<td>0.09</td>
<td>0.27</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LTT2:AGLTP2</td>
<td>-</td>
<td>-</td>
<td>1.21</td>
<td>2.42</td>
<td>2.37</td>
</tr>
<tr>
<td>LFTCKNS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.34</td>
<td>3.99</td>
</tr>
<tr>
<td>VERL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BLDL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LLWM:BL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MPW</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PL:BL</td>
<td>-</td>
<td>-</td>
<td>0.70</td>
<td>0.05</td>
<td>-</td>
</tr>
<tr>
<td>PRPWD1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LOBE1:AGLTP1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\[a\] CT-CCV-ALL = parents, F1V, BCv, BCT, F2v.

\[b\] CT-CCV-F1 = parents and F1v

3.3.4.3. Multivariate analysis of morphology

Multivariate analysis was conducted to identify which characters were more
discriminating in each hybrid group and to examine the degree of interdependence
among characters. Initially, correlation values were estimated for all pair-wise
combinations using all 30 characters. If a pair of characters had an \( r > 0.30 \), the character
with the higher Wald value (Wald statistics/d.f.) was selected for use in further multivariate models. This method was done to reduce redundancy and allow the generation of models based on fewest and most discriminating variables. This is because higher Wald values indicated characters that contributed to greater discrimination among classes (among parents and hybrids within a hybrid group). Models with eight, 11, 11 and 12 variables were used in the analysis of CT-CCC, CT-CCV-F$_1$, CT-CH and CT-CCV-all hybrid groups respectively (See Table 3). Thirteen variables were used in the analysis of the CCV-CH group.

Using the character sets identified above, canonical variate analyses were run separately on each hybrid group to examine the degree of clustering of the individuals into taxa when arranged in a discriminant space. The analysis was implemented using the Multivariate Analysis option and the Canonical Variate sub-option. The same procedure was carried out for CCV-CH group to examine inter-taxon and inter-provenance resolution. All the analyses were carried out using GENSTAT Discovery Edition 3 (Buysse et al., 2007).

### 3.3.4.4. Assignment testing

Assignment tests were used to determine the power of models based on the least number of independent characters for classifying individuals into their respective genealogical class. Re-allocation of individuals in the training set (a priori grouping) was based on Mahalanobis squared distance. Here, the distance between the individual and the group means for a set of characters were calculated from the canonical variate scores and each individual was assigned to the group where it had the smallest Mahalanobis squared distance (Buysse et al., 2007) although the assignment may not necessarily be where that particular individual belongs (Manly, 1994). Assignment accuracy was expressed as the percentage of correct individual assignment for 100 bootstrap permutations.
3.4. Results

3.4.1. Germination, survival and vigour of hybrid seedlings

In general, pair-wise comparisons of germination, survival or vigour (height) among F₁ and their parental taxa were not significantly different (Figure 1). For germination, only the F₁H was significantly lower than its parents (p<0.05) while both the F₁C and F₁V were not significantly different to their parental taxa (Figure 1a). For vigour, only the F₁H taxon was significantly lower than either of its parents (p<0.01) (Figure 1b). The F₁V taxon was significantly shorter than its CT parent (p<0.001) but not significantly different to its CCV parent (p>0.05). The tallest F₁ taxon, F₁C, was not significantly different from either of its parents (p>0.05) at age eight months. There were few differences among taxa for early survival (data not shown). Survival of F₁H was high (95.41%) and not significantly different from both parents (data not shown). The F₁V, however, had significantly lower survival (94.8%) than its parental CT (100%; p<0.05) but was not different to its CCV parent (99.3%; p>0.05). The F₁C on the other hand, had 100% survival and was also not different to either parental taxon.

The advanced generation hybrids tended to have lower germination than their parental taxa or F₁s (Figure 1a). The BC₇ taxon had the lowest germination of any taxon, as no seedling germinated in any family which prevented us from doing further analysis on this taxon. All individuals in the two families of the BC₇ taxon and one family from each of the F₂ and BC₅ taxa also failed to germinate. Survival analysis could only be conducted on the BC₇ taxon, therefore, because only one family from each of the F₂ and BC₅ taxa germinated. For the BC₇ taxon, survival was 94.76% and was not significantly different to the F₁V and CCV taxa, but significantly lower than parental taxon CT (data not shown). In terms of vigour, the average height of the BC₇ taxon was also not significantly different from the F₁V (p<0.05). The BC₅ and F₂ taxa were not significantly different in height from any other taxa.
Figure 1. **a.** average percent germination on a family basis and **b.** average height on individual basis at age eight months of the spotted gum parental taxa and their hybrids. A Kruskal-Wallis test was used to determine significant differences in germination between taxa while a univariate ANOVA was used for total height analysis. Post-hoc testing was carried out using Mann-Whitney U Test and LSD for Kruskal-Wallis and ANOVA respectively. Different letters indicate significant differences between taxa (p<0.05).
3.4.2. Univariate analysis to identify variables that discriminate hybrid taxa from their parents in each hybrid group

Of the 26 quantitative characters assessed, 13 characters had approximately normal distributions when taxa were pooled into genealogical classes. The thirteen characters with non-normal distributions were transformed to achieve approximate normality (Table 2).

The qualitative characters had significant taxa effect on all the four hybrid groups based on the Wald statistic. Of the 26 quantitative characters, 24 in the CT-CH group, 22 in the CT-CCV group, 20 in both the CT-CCC and CT-CCV-\(F_1\) groups, had significant taxa effects (p<0.05) based on the Wald statistic (Table 2). Visually, hybrid taxa tended to be distinct in leaf shape and dimension when compared to their parents (Figure 2).

Below are the results of post hoc comparisons that detailed the characters in which the hybrids were significantly different from at least one parent and therefore were useful to identify the hybrids in a univariate analysis.

3.4.2.1. \(F_1\) hybrids

Post hoc testing of characters that had significant Wald/d.f. indicated that the \(F_1\) taxon (\(F_{1H}\)) from the hybrid group with the most genetically distinct parental taxon had the highest number (18/26) of discriminating quantitative characters (e.g. Figure 3 for BLDL:MPW and SLW).

In the CT-CH hybrid group, each taxon was significantly different from the other taxa in the group for 9 of the 26 quantitative characters (AGLTP1, AGLTP2, LFTCKNS, LOBE1, LOBE2, BLDL:MPW, THGT, DIA10 and NBRN, (p<0.05) (i.e. Figure 3a for BLDL:MPW). Also, the \(F_{1H}\) hybrid was significantly (p<0.5) different from CT in the following three characters: LLWM:MPW, PRPWD1 and LLWM but not to the CH parent. In addition, there were six characters (SLW, LLWM:BL, MPW:PRPWD2, PRPWD2, INTL and LOBE1:AGLTP1) where the \(F_{1H}\) hybrid was significantly different from both parents (p<0.05) but the parents were not significantly different
from each other (p>0.05). For the qualitative characters the $F_{1H}$, CT and CH were significantly different in LIGNO and TPCLR (p<0.05).

Among the 18 quantitative characters where the $F_{1H}$ was significantly different to at least one parent taxon, the $F_{1H}$ had either higher or lower means relative to the parental taxa in most (12) characters (SLW as an example in Figure 3a). In the case of four characters (PRPWD1, PRPWD2, LLWM:BL and MPW:PW2) the $F_{1H}$ taxon had higher means than its parents whereas it had lower means compared to the parents in eight characters. The $F_{1H}$ taxon was intermediate to parental means for the remaining six characters (LOBE1, LOBE 2, BLDL:MPW, NBRN, LLWM:MPW and LLWM).

The hybrid group with parents of intermediate genetic distance (CT-CCV-$F_1$) showed the next highest number (15/26) of discriminating characters. Pair-wise comparisons showed that the intersectional hybrid $F_{1V}$ was significantly different from both of its parents in 7 out of the 26 quantitative characters assessed (NBRN, FORM, LOBE2, MPW, ASRA, BLDL and BLDL:MPW) (BLDL:MPW and SLW as examples in Figure 3b). For four characters, THGT, DIA10, PRPWD1 and INTL10, the $F_{1V}$ taxon was significantly different to its CT parent only. For a further four characters (LLT2:AGLTP2, PL:BL, SLW and MPW:PRPWD2) the $F_{1V}$ taxon was significantly different from its CCV parent only. Among the 15 aforementioned characters that discriminated $F_{1V}$ from at least one of its parental taxa, five were transgressive. The transgression was positive compared to its parents for two character (SLW and PL:BL), negative for three characters (LLT2:AGLTP2, BLDL and PRPWD1), and was intermediate to its parents for the remaining 10 characters. Two qualitative characters, TPCLR and LIGNO, were significantly different among all three taxa in this hybrid group based on Mann Whitney U Test.
Figure 2. Leaf morphology of parents and hybrid taxa: *Corymbia torelliana* (CT), spotted gums – *Corymbia citriodora* subsp variegata (CCV; top row), *Corymbia citriodora* subsp *citriodora* (CCC; middle row) and *Corymbia henryi* (CH; bottom row), and their hybrids, $F_1$ (CTxspotted gum; same order as the spotted gum column), $B_C_T$ (CTxCVxCCVxCT), $B_C_V$ (CTxCVxCCVxCCV), $F_2V$ (CTxCCV) x (CTxCCV)
The hybrid group with the genetically most-similar parental taxa (CT-CCC) showed the fewest number (10/26) of discriminating quantitative characters i.e. where at least the F₁ hybrid could be resolved from at least one of its parents in a univariate analysis (e.g. Figure 3c). The F₁C taxon was significantly different from both of its parents in seven characters (FORM, LOBE1 and LOBE2, SLW, PL:BL, DIA10 and BLDL:MPW). For a further three characters (PETL, MPW:PRPWD2 and MPW) the F₁C was significantly different from its CCC parent only. The F₁C taxon exceeded its parental taxa means in three characters (PETL, MPW:PRPWD2 and MPW) and intermediate in the remaining seven characters. None of the qualitative characters distinguished the F₁C taxon from its parents, but the parents differed from each other by three characters (STMCLR, LIGNO and TPCLR).

### 3.4.2.2. Advanced generation hybrids

Overall, advanced generation hybrids were difficult to resolve from either the F₁ hybrids or the parental taxa, CT and CCV using univariate analysis (e.g. Figure 3b and Appendix 2). The BCₜ was significantly different from CT, CCV, F₁V and BCᵥ (p<0.05) but not the F₂ for the character LOBE1. The BCₜ taxon was also significantly different from both the BCᵥ and F₁V taxa for the characters LLWM:BLDL and LOBE2. It was also significantly different to BCᵥ in LIGNO and to both BCᵥ and F₁V for STHR (p<0.05). The BCₜ exhibited positive transgression for the characters LLWM:BLDL and LOBE2.

The BCᵥ taxon was significantly different to both the F₁V and BCₜ taxa for the characters THGT and LIGNO, and to BCₜ for STHR (p<0.05).

The outbred F₂ taxon (F₂ᵥ) was significantly different from its parental taxa and BCᵥ but not F₁V or the BCₜ taxon for TPCLR, therefore no character could uniquely resolve the F₂ᵥ.
Figure 3. Comparison of taxa within each hybrid group for BDL:MPW and SLW; a) CT-CH group; b) CT-CCV group; c) CT-CCC group. A general linear model (GLM) was used to identify pair wise significant differences between taxa in each hybrid group, different letters indicate significant difference (P<0.05). Number of individuals used in the analysis: CT:n=144; CH:n=121; CCV:n=133; CCC:n=122; F_{1H}:n=75; F_{1V}:n=144; F_{1C}:n=54; BC_{T}:n=33;
Chapter 3: Hybrid Morphology

3.4.2.3. Corymbia citriodora subsp variegata and Corymbia henryi

Most (23) of the 26 quantitative characters significantly differentiated the two spotted gum taxa, CH and CCV, when provenances were pooled within each taxa (Table 2). Eight of these were highly significantly discriminating (Wald statistics p<0.001) (BLDL:MPW, LFTCKNS, MPW:PRPWD2, PETL, VERL, LLT2:AGLTP2, BLDL and SLW).

3.4.3. Multivariate analyses

The canonical variate analyses were run on each CT-spotted gum hybrid combination separately (Figure 4 A-D). The first canonical variate explained most of the variation in seedling morphology in all three hybrid groups, 91.88%, 94.91% and 98.77% in the CT-CH, CT-CCV-F_1 and the CT-CCC hybrid groups respectively (Figures 4A-C). The second canonical variate (CV2) explained the remaining variation in seedling morphology. In the CT-CCV-all model, the CV1 explained 95.17% while CV2 explained 3.29% of the variation in seedling morphology (Figure 4D). There was a general trend in all hybrid groups where CV1 clearly separated out the spotted gum taxa from the CT taxa and their respective hybrid taxon, while CV2 provided moderate separation in a second dimension that tended to resolve hybrid taxa from their CT parent (Figures 4A-D).

Canonical coefficients (loadings) were examined to identify characters with a major influence on a canonical variate. This indicated that LIGNO, BLDL:MPW and SLW were important in both CV1 and CV2 and were common to all hybrid groups (Table 3). In addition to these three characters, there were other characters that had high loadings on CV2, i.e. MPW:PRPWD2 and PETL in the CT-CCC hybrid group, LFTCKNS and LLT2:AGLTP2 in the CT-CH group, and MPW:PRPWD2 in the CT-CCV group (Table 3). These characters were important for the moderate separation of the F_1 from the CT in their respective hybrid groups.
While good separation of spotted gums from their hybrids and the CT parental taxon was evident in each hybrid group, separation of CT from the F₁₈ and advanced generation hybrids tended to be poor (Figures 4A-D).

Some individuals of the spotted gum taxa from some hybrid groups tended to be intermediate between the main grouping of spotted gums and CT. This was particularly so for CCC (Fig. 4C). The CT-CCC hybrid group was examined more closely, and found that the CCC individuals intermediate on CV1 were all from Yeppoon provenance and derived from across all five families that were tested from this provenance. It was inferred a provenance, rather than family level effect induced the deviance of these individuals. When examined for the two most influential characters underlying CV1 (Table 3), these individuals had lower leaf ratios (BLDL:MPW) and were classified as not having a lignotuber, both conditions which would make them more closely resemble CT than other CCC individuals (Figure 4C).
The two closely related spotted gum taxa CCV and CH were clearly separated in discriminant space with CV1 contributing 97.14% of the variation and CV2 contributing 1.43% (Figure 5). Among the 13 characters used in the canonical variate analysis, BLDL:MPW, VERL and LFTCKNS had high loadings at CV1 which contributed significantly to the separation of CCV and CH (Table 3). There was moderate clustering of individuals belonging to the same provenance but provenances within a taxon were not well resolved (Figure 5).
3.4.4. Assignment testing

In all three hybrid groups, the spotted gum individuals were assigned to their correct group with a higher success rate than the CT taxon (Table 4). The CH taxon had the highest success rate at 96% followed by CCV (95%), then CCC (92%). The assignment rate for the CT taxon was 93%, 63% and 52% for the CT-CH, CT-CCC, CT-CCV groups respectively (Table 4).

Following the high assignment rates for individuals in the parental taxa in the CT-CH hybrid group, the $F_{1H}$ individuals were assigned correctly with the highest accuracy (76%) of any of the $F_1$ hybrid (Table 4). Both the $F_{1V}$ and $F_{1C}$ had 60% assignment accuracy.
Chapter 3: Hybrid Morphology

Table 4. Assignment accuracy (%), and frequencies of transgressive characters in the F\textsubscript{1} taxa for each of the three hybrid groups, CT-CCC, CT-CCV and CT-CH.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>CT-CCC</th>
<th>CT-CCV</th>
<th>CT-CH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spotted gum parent\textsuperscript{a}</td>
<td>92</td>
<td>95</td>
<td>96</td>
</tr>
<tr>
<td>CT</td>
<td>63</td>
<td>52</td>
<td>93</td>
</tr>
<tr>
<td>F\textsubscript{1}</td>
<td>60</td>
<td>60</td>
<td>76</td>
</tr>
<tr>
<td>BC\textsubscript{V}</td>
<td>na</td>
<td>25</td>
<td>na</td>
</tr>
<tr>
<td>BC\textsubscript{T}</td>
<td>na</td>
<td>9.5</td>
<td>na</td>
</tr>
<tr>
<td>F\textsubscript{2V}</td>
<td>na</td>
<td>33</td>
<td>na</td>
</tr>
<tr>
<td>No. of transgressive characters in the F\textsubscript{1}</td>
<td>3</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>

\textsuperscript{a}CT-CCC = Corymbia citriodora subsp citriodora, CT-CCV = Corymbia citriodora subsp variegata, CT-CH = Corymbia henryi.

Among the advanced generation hybrids in the CT-CCV hybrid group, the F\textsubscript{2V} had higher assignment success at 33\% than either of the backcrosses (Table 4). The BC\textsubscript{V} had 25\% assignment success, while BC\textsubscript{T} had the lowest assignment rate of any taxon tested (9.5\%). Using the multivariate models, individual spotted gum seedling could be assigned correctly to their parental taxa groups with a high degree of success; CH was assigned to its correct taxon with 100\% success while CCV was assigned with 97\% accuracy (data not shown).
3.5. Discussion

3.5.1. Assignment and the use of morphology for hybrid identification

In this study, the reliability of morphology was examined to distinguish among genealogical classes of spotted gums in three scenarios relevant to the plantings of spotted gums and their hybrids with *Corymbia torelliana* (CT) in subtropical Australia. In order to assess and manage gene flow from plantation spotted gums, ideally a rapid and reliable tool to measure gene flow should be in place before the plantations reach maturity. Direct measurement of gene flow is thought to be more effective in detecting low levels of gene flow but may involve screening of many thousands of seedlings (Barbour *et al.*, 2002; Barbour *et al.*, 2003). Morphological identification could facilitate the screening of large numbers of individuals provided that at least a few reliable discriminating characters can be identified.

Overall, seedling morphology only offered moderate reliability to distinguish hybrids from parental taxa. It was only possible to assign an F\(_1\) hybrid with 60-76% accuracy however, only low correct assignment could be achieved for advanced generation hybrids (9.5-33%). Therefore, in all hybrid cases, there was only moderate assignment accuracy using morphology. The best case was in the hybrid group with the most divergent parental pair, CT and CH, where individuals from all three taxa could be assigned accurately at greater than 76%. Prospects for using morphology to distinguish among advanced generation hybrids like backcrosses and F\(_2\)S were even lower. As expected, using morphological characters that are used to taxonomically discriminate the species, seedling leaf morphology could reliably distinguish among the spotted gum taxa CCV and CH. It has been shown, however, that there was little variation in morphology among provenances in each of these species even though geographically distinct provenances were evaluated and, as a consequence, the two taxa were well separated in discriminant space.

There were three characters common to the three hybrid groups which showed greatest potential for use to screen CT hybrids, BLDL:MPW, SLW and LIGNO. The ratio of leaf blade length to maximum perpendicular width, which describes leaf shape, was the
most highly discriminating character in all three hybrid groups. The intermediate nature of the leaf shape of hybrids observed here conforms to the general observation among eucalypts where hybrid morphology is intermediate to their parents (Potts and Dungey, 2004). Although SLW was intermediate in the F_{1C}, it was transgressive in both F_{1V} and F_{1H}, and was highly discriminating in both the univariate and multivariate analyses.

The presence or absence of lignotuber (LIGNO) was also among the more powerful characters for separating hybrids from CT and has potential in rapid visual screening. Lignotubers are obligatory on CCV, absent on CT (Shepherd et al., 2007) but were found to segregate in a large outcross F_{2} family where the greater majority of hybrids possessed lignotubers (Shepherd et al., 2008b). In the present study, the segregation observed in control pollinated F_{1}, BC and F_{2} families was congruent with this conclusion but the small family sizes, and the young age of seedling (with the potential for mis-classification) did not allow reliable estimates of segregate class sizes therefore data was equivocal in regard to the genetic model. Furthermore, in my detailed investigation on the CT-CCC hybrid group, I found deviant “intermediate” individuals were classified as having no lignotubers. I believe that the assignment of a no lignotuber class to these individuals was due to mis-classification, a difficulty associated with determining this attribute state on these still relatively young seedlings. It is believed that these individuals would eventually develop lignotubers. The lignotuber state does not appear to be definitive for assessing seedlings for hybridity at this age, but the presence of a lignotuber on offspring from a CT individual, nonetheless, would be strong evidence for infusion of CCV background.

Analysis of deviant CCC individuals also showed that families from the Yeppoon provenance had wide variances in their leaf ratios (BLDL:MPW), which is believed to be due to variation for this character in this population. Yeppoon is from the southern genetic group of CCC (Shepherd et al., 2008a) where it intergrades with CCV, a taxon with a higher leaf ratio. Alleles underlying a higher leaf ratio are likely to be more frequent in CCC from Yeppoon. These population differences among the parental taxa highlight the need for caution, and the need to test a range of provenances in this study, before inferences with respect to taxa-wide effects are made.
3.5.1.1. Scenario 1- Distinguishing intersectional F₁s hybrids of CT x spotted gums

Scenario 1 requires discrimination of the intersectional F₁ hybrid from its parental spotted gums and CT taxa. This may be necessary in situations where CT has been planted as an amenity tree near native spotted gum, for example. In this scenario, full resolution of all three genealogical classes would only be possible with modest reliability, in the case of the two most genetically divergent parents, CH and CT. In the CT-CH group, each genealogical class could be assigned at a rate of better than 76%, the lowest assignment rate was for the F₁ taxon, but the parents were assigned at a rate >93%. In all three hybrid groups, the assignment accuracy of distinguishing the spotted gum parent from the F₁ or the CT parent was high (>92% accuracy), but assignment accuracy for CT parent or the F₁s was between 52% and 93%.

It is possible therefore, that F₁ hybrid identification based on seedling morphology might improve assignment rates in all three hybrid groups to around at least twice the rate of random assignments but highly reliable determinations seem unlikely, especially for the F₁C and F₁V taxa. Determinations with moderate confidence levels, nonetheless, may be sufficient for some applications. For example, screening with morphology might be a useful initial step to screen thousands of seedlings to a manageable size which can then be verified by other methods. This approach was found to be valuable in quantifying gene flow from a Eucalyptus nitens orchard into surrounding native stands of Eucalyptus ovata and Eucalyptus viminalis (Barbour et al., 2002; Barbour et al., 2003). A high congruence between morphology and isozyme analysis was found in determining F₁ eucalypt hybrids which indicates the utility of this 2-stage approach (Barbour et al., 2002; Barbour et al., 2003; Field et al., 2009). A similar application of morphological screening, in conjunction with molecular marker screening, may be useful for identifying Corymbia hybrids with CT when high assignment accuracy is needed.

The relatively low levels of hybrid assignment accuracy using morphology found in this study were comparable to that obtained by Field et al. (2009). One reason for the difficulty in discriminating hybrids from parental taxa in young seedlings may be due to maternal effects which may cause the F₁ to resemble its maternal parent, but it may not be the case in this study. Maternal effects have been proposed to explain the
resemblance of early age hybrid seedlings to their maternal parents in a number of studies of eucalypts (e.g. (de Assis, 2000; Delaporte et al., 2001; Lopez et al., 2003) but this inference is sometimes made without testing reciprocal crosses. Maternal effects are thought to be a consequence of the greater contribution of the maternal parent to the endosperm including maternally derived tissues such as seed coat and integuments and are most important at the embryo stage and early seedling vigour (Roach and Wulff, 1987; Lopez et al., 2003). Their influence is thought to decline rapidly as the seedlings age and to be minimal for assessing performance on surviving seedlings in the nursery (Griffin et al., 2000; Harbard et al., 2000). In this case, the $F_1$s were most similar to the maternal parent (CT) and this was particularly so for younger seedling (i.e. four-month old seedlings, data not shown). Although maternal effects may cause the similarity between hybrids and their maternal parent as observed, this may also be a consequence of inheritance, and this issue needs clarification with a study of the reciprocal crosses.

### 3.5.1.2. Scenario 2 – Distinguishing advanced generation hybrids

In the second scenario, it was sought to discriminate among advanced generation hybrids (backcrosses and $F_2$s) and their parental taxa. This type of determination may be necessary in situations where the $F_{1V}$ is planted experimentally or operationally in close proximity to native spotted gums. Distinguishing among the advanced generation hybrids, $F_1$ and their parents was more challenging than the discrimination of $F_1$ from parental taxa because there was more overlap in phenotypes among the taxa. Large within –family variation was often evident in advanced generation hybrids because of the segregation of species differences (Shepherd et al., 2005). This proved to be the case in this study where the variances in the advanced generation families for characters such as VERL and BLDL exceeded the variances evident in parental taxa and the $F_1$ hybrids. The wide variances in these families would have exasperated the problem of resolving among genealogical classes and contributed to the low assignment rates for the hybrid taxa (between 9.5% and 33%; Table 3). Again, the difficulty occurred in distinguishing among the hybrid taxa ($F_1$, BC and $F_2$), and the CT maternal parent; only CCV could be distinguished with reasonable certainty from other taxa (95%), similar to the first scenario. However, a morphological assay should still have important practical application because the situation of most interest is the detection of pollen-mediated
gene flow in seedlots from native spotted gum stands, where the critical test is the discrimination of offspring with *C. torelliana* background from pure spotted gums.

### 3.5.1.3. Scenario 3 – Distinguishing CCV from CH

The third situation where it may be necessary to identify interspecific hybrids arises where locally exotic species or provenances of spotted gums are planted in close proximity to native spotted gums. This situation is common in northern NSW where both CCV and CH are native and several coastal provenances of CCV from around Gympie, QLD are planted extensively because they have higher tolerance to *Quambalaria* shoot blight (QSB) (Johnson *et al.*, 2009; Brawner *et al.*, 2011). In a previous study, it was shown that CCV and CH are difficult to resolve using molecular markers and there was more variation within species than between species leading to a suggestion that they could be regarded as morphotypes or ecotypes rather than species (Shepherd *et al.*, 2008a; Ochieng *et al.*, 2010). Neutral molecular markers such as microsatellites are not useful for taxa discrimination in such cases as they usually do not associate with phenotypes that are adaptive or account for morphological differences (McKay and Latta, 2002; Steane *et al.*, 2006). Morphological markers may be more useful than neutral genetic markers in discriminating hybrids in such cases.

The results show that the characters used for taxonomic classification provided reliable distinction of CCV from CH. This study was also consistent in this regard with two earlier studies of seedling morphology of these two taxa (Larmour *et al.*, 2000; Ochieng, 2009). It has also been found that there was relatively little morphological variation between provenances within each taxon despite geographically divergent provenance pairs were selected for testing. The selection of geographically divergent provenances in each taxon was intended to maximise the likelihood that they were also genetically divergent. This suggests that seedling morphology could be used with confidence to distinguish between taxa but it is unlikely that it could be used to distinguish among provenances within each of the taxa.
Chapter 3: Hybrid Morphology

The power of seedling morphology to discriminate hybrids between these two spotted gum taxa is of interest in monitoring for gene flow. However, hybrids of these taxa were not available for this study. A previous study suggested that F$_1$ of these two taxa are intermediate and significantly different in leaf mass area, and BLDL: MPW (Ochieng, 2009). Thus, seedling leaf mass area and BLDL:MPW may be good candidates for identification of hybrids between CCV and CH.

3.5.2. Genetic divergence and transgressive characters

In this study, attention was accorded to the frequency of transgressive characters in hybrids in this study because of their potential to indicate a taxon’s wider environmental adaptability. The Corymbia hybrids are known for high environmental plasticity (Lee, 2007; Lee et al., 2009; Nahrung et al., 2011). In a tree improvement context, this is valuable as a way of extending the planting range of taxa or introducing added resilience into planted forests for abiotic and biotic tolerance. However, transgressive characters have been associated with the ability of an individual to colonise novel environments (Rieseberg et al., 2007) and contribute to weediness (Wallace et al., 2005). Although hybrids are often less fit than their parents, if genotypes with extreme characters do survive, they may have elevated fitness and adapt to novel and disturbed habitats (Arnold and Hodges, 1995). The observation that F$_1$ hybrids exceeded the parental taxa means for some characters in this study indicates there may be a need to monitor hybrid fitness in the advanced generations where extreme characters may even be more extensive.

In the present study, SLW was transgressive in the F$_{1v}$ and F$_{1h}$. This character is an indicator of water use efficiency and has been related to water availability at the place of seed origin in eucalypts; such that a decrease in SLW and leaf size accords with a decrease in availability of water or nutrients to the plant (Fonseca et al., 2000; Warren et al., 2006; Poorter et al., 2009). In a hybrid, extending the range of SLW beyond that observed in parental taxa may confer an advantage in areas with low moisture and has been speculated as one factor that may have contributed to the origin of hybrid species of sunflowers (Rosenthal et al., 2002).
In this study, there was a trend whereby $F_1$ crosses between more genetically divergent parents tended to be better resolved by morphology (i.e. the $F_{1H}$ was the most reliably resolved $F_1$ class). In the present study, the number of characters found to have extreme values in the hybrids increased with genetic divergence of the parents i.e. $F_{1C}$ (3), $F_{1V}$ (5) and $F_{1H}$ (12). *Corymbia henryi* has been shown to be genetically the most divergent spotted gum to CT by molecular markers, then CCV, then CCC (Shepherd *et al.*, 2008a). This increase in frequency of transgressive characteristics with increasing genetic distance between parents has been noted in other studies of interspecific hybrids of plants and animals (Stelkens and Seehausen, 2009). Complementary gene action is thought to most often be the genetic basis for transgressive segregation (Rieseberg *et al.*, 2007) but this would not fit in this case as this predicts transgression will be less frequent in phenotypically divergent parents (Stelkens and Seehausen, 2009).

### 3.6. Conclusion

In this study, it has been found that morphology could provide low to moderate confidence identifying $F_1$ hybrids from their parental taxa. Hybrids were resolved more reliably from their paternal spotted parent than their maternal parent CT. The greater difficulty resolving hybrids from the maternal parent in this case may not be a major limitation to the practical use of morphology for monitoring gene flow from planted *Corymbia* hybrids into native forests because the situation of most interest in detection of hybrid wildings in native forests of spotted gum. Leaf shape, specific leaf weight and the presence or absence of lignotuber can be used to determine frequency of hybridisation in plantation boundaries and aid in understanding the dynamics of hybridisation based on flowering synchrony and relative frequency of trees serving as source of pollen. Assessment of morphological characteristics may also help in detecting hybrids between CT and spotted gums in large scale screening of spotted gum seedlings but is unlikely to provide highly reliable screening in itself. Hybrids that have less discriminating morphological characters (i.e $F_{1C}$) may be undetected when using morphology thus have limited use in monitoring gene flow. Other methods for directly monitoring gene flow such as near infra-red spectrophotometry (NIRS) and molecular markers are now being assessed on the materials used in the current study to allow comparative analysis of their effectiveness.
3.7. Acknowledgment

This research forms part of M Abasolo’s PhD research supported by the Cooperative Research Centre for Forestry and International Postgraduate Research Scholarship. I thank the Department of Employment, Economic Development and Innovation, QLD, for the use of their nursery and for providing the seeds used in the experiment, Tracy Menzies and John Oostenbrink for assistance in the nursery and data collection.
Chapter 4. Deviant near-infrared spectra allow identification of *Corymbia* hybrids

This chapter has been published as:


Contributions to this chapter:

Experimental design:

Myralyn Abasolo, David J Lee and Mervyn Shepherd

Field/Laboratory work:

Myralyn Abasolo performed nursery and laboratory work. Assistance from Tracey Menzies and John Oostenbrink in the nursery work is acknowledged.

Statistical analysis:

Myralyn Abasolo and Roger Meder performed the statistical analysis.

Writing:

Myralyn Abasolo, Mervyn Shepherd, Roger Meder, David J Lee and Carolyn Raymond wrote the manuscript.
4.1. Abstract

Near infrared reflectance spectroscopy (NIR) is used extensively to predict physico-chemical properties of plant tissues including wood, wood products and grains but may also be used for rapid and non-destructive identification of hybrids. Here, I examine the potential of NIR for field diagnosis of hybrids to estimate gene flow rates from plantings of locally exotic Corymbia (formerly Eucalyptus) species in subtropical Australia. NIR profiles were generated by scanning foliage from a total of 383 hybrid and 533 parental seedlings grown in a common garden and partial least square discriminant analysis was used to test three-way model power to accurately assign individuals to their appropriate taxon; either a parental or F₁ hybrid class. Using the optimised conditions, fresh foliage from eight-month-old seedlings and a handheld NIR (980 - 1800 nm), the mean assignment rates for three hybrid groups ranged from 76 to 90%. Hybrid-parent contrast of NIR spectra deviated more so than parent-parent contrast. The F₁ taxon assignment rates were usually higher than those for parents at 100% and 72%, respectively. Hybrid resolution was even greater for 2nd generation backcross hybrids. This was unexpected given the typically intermediate status of F₁ hybrids and the tendency for backcross hybrids to resemble their recurrent parent when assessed on morphological grounds. Similar to studies of morphology, taxon assignments tended to be more accurate for hybrid groups in which the parental taxa were more divergent. The practical application of this technique for hybrid diagnosis of seedling in the nursery will require careful attention to control of environmental factors as tissue age and storage effects influenced the ability of NIR to identify hybrids. The technique may also necessitate the generation of comparable reference populations, although exclusion approaches to analysis may circumvent the need for reference populations. The application of NIR in field diagnosis will be further complicated by the need to generate global models across environments but such models have been obtained for reliable prediction of chemistries in other situations.
4.2. Introduction

Gene flow from forest plantations may have negative consequences for the genetic diversity, structure, and composition of native tree populations (Arnold et al., 1999). Despite the observations that even low levels of gene flow from exotic plantations may constitute a risk (Barbour et al., 2008), there has been little assessment of the scale and impact of gene flow, and little monitoring continues to be undertaken (Laikre et al., 2010). This, in part, may be due to the difficulty and costs associated with the screening of large numbers of individuals needed to detect low levels of gene flow and often a lack of cost-effective reliable hybrid detection tools (Barbour et al., 2008).

The spotted gums (Genus Corymbia Section Maculatae) are the major hardwood plantation taxon in subtropical Australia (Lee, 2007; Lee et al., 2009; Lee et al., 2010; Nichols et al., 2010). The four taxa (C. citriodora subsp. citriodora (CCC), C. c. variegata (CCV) and C. henryi (CH)) occur naturally as a replacement series that extends along the east coast from around Cairns (QLD) in the north, to Bega (Forests NSW) in the south (Hill and Johnson, 1995). The hybrid between spotted gum and Corymbia torelliana (Section Torellianae; CT), a native of tropical north Queensland (QLD), is also of interest to forestry and is being planted experimentally. Corymbia torelliana and spotted gum hybridise naturally in north Queensland where they co-occur, and controlled crossing experiments show that they can be easily crossed in both directions and have early viability (Hill and Johnson, 1995; Dickinson et al., 2010; Dickinson et al., 2013).

In addition to its use in forestry, CT has been widely planted as an amenity tree, but in recent years this has declined as CT was recognised as a weed in some shires of northern New South Wales and southern QLD (Hill and Johnson, 1995; NCWAC, 2003; Kingston et al., 2004). Plantings of CT stock give rise to two situations where gene flow monitoring from plantings may be warranted. Firstly, where CT is planted as an amenity tree near native spotted gum because there is potential for introgression of CT into native forest gene pools. And second, where the Corymbia F₁ hybrid (CT x spotted gums, F₁, sensu Lee (2007)) is planted near native spotted gum, and hence there is potential for backcrossing and further introgression of CT into native gene pools. This introgression may be undesirable because the same attribution of adaptability and
vigour, that make the hybrids attractive for plantations, may also contribute to the
ability of the hybrids to invade novel and disturbed environments (Lee, 2007; Lee et al.,
2009; Abasolo et al., 2012; Shepherd et al., submitted).

Near infrared spectroscopy has been widely used to predict chemical compositions in
agricultural crops and to predict the wood chemistry of trees (e.g. Jones et al., 2006;
Poke and Raymond, 2006; Maranan and Laborie, 2008; Meder and Meglen, 2012). NIR
has also been shown to have potential for evaluation the nutritional value of forests in
ecological studies (Foley et al., 1998; Stotler et al., 2006; Meder et al., 2007) and for
selection of wood quality in tree breeding programmes (Schimleck, 2007; Meder et al.,
2011; Meder and Schimleck, 2011). To a lesser extent, NIR has also been used for
hybrid identification in plants but its potential in this regard, as far as these authors are
aware, has been restricted to fresh leaves of F₁ hybrids of *Eucalyptus* and *Betula*
(Atkinson et al., 1997; Humphreys et al., 2008). The application of NIR to previously
stored tissue is limited (i.e. Cynkar et al 2009) and there is no report of NIR application
to advanced generation hybrids. The power of NIR to predict eucalypt hybrids using
two sets of NIR wavelengths is still to be explored. The main value of near infrared
spectroscopy in hybrid screening is where it is difficult to identify hybrids by
morphology (Tibbits, 1988; Abasolo et al., 2012) but it may offer other advantages such
as reduced laboratory time and non-destructive sampling relative to other methods
(Foley et al., 1998; Ebbers et al., 2002; Humphreys et al., 2008).

The principle behind NIR is similar to that of our human perception of colour in the
visible light region (400 – 700 nm) of the electromagnetic spectrum. We see an object
as being green, because all the other wavelengths of light, except green (wavelength ca.
510 nm), are absorbed by the object and only green is reflected back to our eyes. In the
same manner, infrared radiation (700 nm – 1 mm) incident on an object is either
absorbed or reflected at various wavelengths. Near infrared spectrometers operate in the
near and short wavelength range of the infrared region, typically between ca. 800 and
2,500 nm dependent on the detector technology used. Absorption of incident energy
occurs due to interaction of the incident radiation with the variety of molecular bonds
present in the object matrix, such as the stretching and bending modes of bonds
between carbon, hydrogen, oxygen, nitrogen and other elements present in organic
materials (Givens et al., 1997). It is the variety in these bond arrangements (e.g. the C-
H bond of a methylene differs slightly from the C-H bond of a methyl and is completely
different to a C-O bond) that in turn gives rise to differences in infrared absorption so
that different chemical constituents will have different absorbance in the NIR spectra
(Foley et al., 1998; Tsuchikawa, 2007; Schwanninger et al., 2011).

One of the limitations that must be addressed in the application of NIR to predict
chemical composition (or in taxonomic classification), however, is the requirement for
calibration and validation of an empirical chemical model (Foley et al., 1998; Richardson et al., 2004). A global model, based on a large sample size is ideal because
such a model should be broadly applicable (Wallis and Foley, 2003). For some
applications, such as varietal identification, the NIR spectral differences among samples
are often tested against a previously developed global model, for example in coffee
(Bertrand et al., 2006; Posada et al., 2009) and oranges (Cen et al., 2007). It can be
difficult to identify a particular compound that explains the difference between samples
as the NIR profile corresponds to a functional group (Foley et al., 1998), in which the
compounds comprising the functional group have overlapping NIR spectra (Curran,
1989). Global models exist for limited wood properties, specifically the Kraft pulp yield
of hardwoods (Downes et al., 2009; Downes et al., 2010; Downes et al., 2011) and
softwoods (Hodge and Woodbridge, 2010). Addressing these points is necessary for a
wider application of NIR (i.e. application in different environments and different ages)
for hybrid identification without the need of a reference population growing in the same
environment and the same age as the taxon to be classified.

In this study, the potential of NIR analysis to resolve among genealogical classes was
examined by comparing the NIR spectra of the Corymbia hybrids with their parental
taxa. Seedlings of up to 10 families in each taxon were grown in a common garden and
assessed at two ages, (four months and eight months), to evaluate the impact of age on
our ability to discriminate between different hybrid classes. Three hybrid groups (CT xCCC (F_{1C}), CT x CCV (F_{1V}), and CT x CH (F_{1H})) with the greatest potential for gene
flow from plantation were evaluated. The transfer of pollen from planted sources onto
native spotted gums is the cross direction of interest because the interest is gene flow
from plantings. However, crosses with a spotted gum maternal parent were not
available at the time of this study, and thus was limited to crosses in the reciprocal
direction, i.e. with CT maternal parents. A recent study of seed germination showed that
controlled crosses with CCV maternal parents are as viable as the reciprocal crosses (Dickinson et al., 2012), strengthening evidence for the possibility for gene flow in both directions. For the CCV hybrid group, advanced generation hybrids (outbred backcross, BCT) were also examined in addition to the F\textsubscript{1}. In addition to plant age, differences between fresh and thawed leaf tissues and the influence of NIR instrument wavelength range were examined to determine the impact of these factors on our ability to discriminate between taxa.

In this chapter, the following specific questions were answered: 1) Do leaf tissue storage, age and spectral ranges affect the ability to identify hybrids by NIR; 2) What is the optimum condition (tissue storage, tissue age and NIR spectral range) in which NIR can best identify hybrids?; 3) What are the diagnostic regions in the NIR spectra that could differentiate hybrids from their parents; 4) Are advanced generation hybrids identifiable by NIR? Do they have higher or lower classification power than the F\textsubscript{1} hybrids?
4.3. Methods

4.3.1. Materials and experimental design

The materials used in this study were from two sections of the Genus *Corymbia*. The first was from Section *Maculatae* (i.e. the spotted gums) represented by three of the four recognised taxa (CCV, CCC and CH; *Corymbia maculata* was not studied) and the second was from section *Torellianae* (CT) (Hill and Johnson, 1995; Parra-O. *et al.*, 2009).

The seedlings used for this study were sown and grown at the Department of Agriculture, Fisheries and Forestry glasshouse facility at Gympie, QLD as detailed in Abasolo *et al.* (2012). The parental taxa, CCC, CCV, CH, and CT were represented by 10 open-pollinated families of: CCC (ex Kirrima, QLD [17°39'S, 146°5'E] and Yeppoon, QLD [23°07'S, 150°44'E]); CCV (ex Richmond Range, NSW [28°40'S, 152°42'E] and Woondum, QLD [26°15'S, 152°49'E]); CH (ex Lockyer, QLD [27°30'S, 152°04'E], Nerang, QLD [27°59’S, 153°20’E] and Myrtle, NSW [29°08’S, 152°05’E]); CT (ex Helenvale, North Queensland [15°43’S, 145°14’E] and a landrace from the Gympie region [no latitude and longitude available but molecular studies indicate that this material was from Kuranda, QLD [(16°5’S, 145°36’E)] (McVey, 2004)).

The hybrid taxa studied were from controlled pollinations and consisted of outbred F$_1$ interspecific hybrids (F$_{1H}$, F$_{1V}$, F$_{1C}$), and outbred backcross to *C. torelliana* (BCT). Where possible, parental control populations were composed of open-pollinated seedlots of trees from the same provenance (or nearby provenances) as the parents used for hybrid crosses but they were not the same individuals (data not shown; see Abasolo *et al* (2012)). All analyses were based on hybrid groups as follows: 1) CT, F$_{1C}$, CCC for the CCC group; 2) CT, F$_{1V}$, BCT and CCV for the CCV group; and 3) CT, F$_{1H}$, CH for the CH group. The number of families in the hybrid taxa was lower and more variable than the parental taxa as fewer hybrid families were available for study (Table 1).
The experiment was established and analysed using an Incomplete Block Design with five replicates. Individuals from each family were randomly assigned to replicates and blocks using CycDesigN Version 1.2 (Whitaker et al., 2006) with each replicate containing nine blocks and each block arranged into four rows and six columns. Each row contained two families (treatments) with each block containing eight treatments. A treatment was represented by a maximum of three seedlings in each replicate where possible.

4.3.2. **Collection of NIR spectra**

Whole, intact, fresh leaves of each individual seedling were scanned using a handheld near infrared spectrometer (Polychromix Phazir Model 1018, now Thermo Scientific, Tewksbury, MA, USA, www.thermoscientific.com) operating in the nominal wavelength range (950-1,800 nm) on four- and eight-month-old seedlings.

Thawed leaf samples of eight-month-old seedlings were scanned using the original and a second handheld NIR (microPhazir RX, Model 1624, Thermo Scientific – Portable Optical Analysis Tewksbury, MA, USA, www.thermoscientific.com) operating at a higher but overlapping wavelength range (nominally 1,600 – 2,400 nm). These wavelength ranges were truncated in order to remove absorbance artefacts that appeared at either end of the spectra. Thus for all analyses, the instrument with the lower wavelength range designated as NIR1 had 81 nm truncated at the beginning and 44 nm at the end of the spectra leaving a range of 1,031 – 1,756 nm. The instrument with the higher wavelength range designated as NIR2 had 39 nm and 40 nm truncated at the beginning and the end of the spectra, respectively, leaving a range of 1,639-2,360 nm for analysis. A total of four data sets were collected and are summarised in Table 1. Data set 1; spectra from four-month-old fresh leaf tissues measured with NIR1; Data set 2; eight-month-old fresh leaf tissues measured with NIR1; Data sets 3 and 4; eight-month-old thawed leaf tissues measured with NIR1 and NIR2, respectively. Note that the instrument used to collect NIR2 spectra was obtained later in the study and was not available when the spectra of the fresh leaf tissues were obtained, thus there was no collection of NIR2 spectra from fresh leaf tissues in this study. Also note that the Data sets 3 and 4 (thawed tissues) were frozen for about one and a half years at -15°C before
both NIR1 and NIR2 spectra were acquired. The NIR spectrum of each leaf was obtained from the centre of the adaxial surface of the first fully expanded leaf from the top of the main stem avoiding the midrib and any damaged areas of the leaf. Two spectra per sample were collected and these were averaged to represent a single NIR spectrum per individual.

### 4.3.3. Outliers were removed before statistical analysis

Two methods were employed to determine obvious outliers, which were removed from further analyses. Firstly, the NIR spectra of all samples in a taxon were plotted and samples whose spectra differed significantly from the mean spectrum (i.e. putative hybrids from collections of open pollinated taxa or products of crossing between parental taxa that were not pure), or were excessively noisy (typically were spectra of air), were considered outliers and were not included in subsequent analyses. Secondly, 19 F₁ individuals which had only one allele from one parent based on 12 microsatellite loci (Shepherd et al., in preparation), and thus likely to be selfed individuals based on microsatellite data (data not shown) were removed from this study.
Table 1. *Corymbia* parental and hybrid taxa information for each data set used in NIR analysis. Data sets varied for parameters of age, instrument and leaf tissue preparation.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Age</th>
<th>Instrument(^a)</th>
<th>Leaf tissue preparation</th>
<th>Taxon(^b)</th>
<th>Number of samples (n)(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Four months</td>
<td>NIR1</td>
<td>Fresh</td>
<td><em>C. torelliana</em></td>
<td>147</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. citriodora</em></td>
<td>115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. variegata</em></td>
<td>141</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. henryi</em></td>
<td>127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1C})</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1V})</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1H})</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BCT</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td><strong>837</strong></td>
</tr>
<tr>
<td>2</td>
<td>Eight months</td>
<td>NIR1</td>
<td>Fresh</td>
<td><em>C. torelliana</em></td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. citriodora</em></td>
<td>115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. variegata</em></td>
<td>141</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. henryi</em></td>
<td>127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1C})</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1V})</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1H})</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BCT</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td><strong>840</strong></td>
</tr>
<tr>
<td>3</td>
<td>Eight months</td>
<td>NIR1</td>
<td>Thawed</td>
<td><em>C. torelliana</em></td>
<td>109</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. citriodora</em></td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. variegata</em></td>
<td>81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. henryi</em></td>
<td>107</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1C})</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1V})</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1H})</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BCT</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td><strong>519</strong></td>
</tr>
<tr>
<td>4</td>
<td>Eight months</td>
<td>NIR2</td>
<td>Thawed</td>
<td><em>C. torelliana</em></td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. c. citriodora</em></td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. x. variegata</em></td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>C. henryi</em></td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1C})</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1V})</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(_{1H})</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BCT</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td><strong>383</strong></td>
</tr>
</tbody>
</table>

\(^a\)NIR1 - 950-1,800 nm; NIR2 - 1,600-2,400 nm.

\(^b\)Number of families assessed were as follows: all parental taxa (*C. torelliana, C. c. citriodora, C. c. variegata* and *C. henryi*) = 10; F\(_{1C}\) = 5; F\(_{1V}\) = 12; F\(_{1H}\) = 8; BCT = 4. Provenances of each taxon were provided in Abasolo et al. (2012). Cross types: F\(_{1C}\) = *C. torelliana* x *C. c. citriodora*; F\(_{1V}\) = *C. torelliana* x *C. c. variegata*; F\(_{1H}\) = *C. torelliana* x *C. henryi*; BCT = *C. torelliana* x *C. c. variegata* x *C. torelliana*.

\(^c\)The number of samples used varied based on the availability of samples that were appropriate for NIR scanning at the time of data collection.
4.3.4. Statistical analysis

4.3.4.1. Determination of standardised data pre-treatment

NIR spectra are often mathematically pre-treated before performing any multivariate analysis to minimise spectral variation due to systematic errors such as particle size, sample presentation or structural variation among samples (Foley et al., 1998). In this study, one of the initial goals was to identify a standard mathematical pre-treatment to be used in all hybrid groups, in all four data sets. Data set 2 was used for this component of the study (See Table 1) and limited the data to the subset of the CH hybrid group (n=355). This Data set was chosen because fresh samples are generally more amenable for NIR analysis than thawed samples which may have undergone degradation, thus fresh samples better represent the actual chemical component of a material and this is reflected in their NIR spectra. The CH hybrid group was chosen because preliminary principal component analysis (PCA) indicated that this hybrid group had the best separation of the F\textsubscript{1} hybrids in the three hybrid groups.

Three methods of first derivative transformation were tested to determine the best pre-treatment as indicated by model fit, namely: Gap (with a gap size of 3); Gap-Segment (with gap size of 3 and segment size of 3) and Savitzky-Golay (1964) using a 7-point smoothing window and 2\textsuperscript{nd} order polynomial fit. A first derivative transformation was chosen because it has been previously used in NIR analysis of fresh and freeze-dried eucalypt leaves (Ebbers et al., 2002; Humphreys et al., 2008) and it helps reduce baseline variation and enhance spectral features (Barnes et al., 1989; Conzen, 2003). A standard normal variate (SNV) pre-treatment was applied for scatter correction after first derivative conversion in each of the three ways tested. A de-trend scatter correction was used in addition to SNV after Gap-Segment first derivative conversion (Dury et al., 2001). For each pre-treatment tested, a PCA was run to extract the first 20 principal components. These principal components were used in a partial least-square discriminant analysis (PLS-DA) to determine the most appropriate pre-treatment, based on highest $r^2$ (correlation coefficient) and lowest standard error of calibration (cross validation mode, SECV) as described in the next section. The best pre-treatment was
used for testing all the hybrid groups in all cases (whether thawed or fresh samples, different instruments and different ages of leaf tissues).

4.3.4.2. Principal component analysis (PCA) and Partial least square discriminant analysis (PLS-DA)

Multivariate analyses, namely PCA and PLS-DA were undertaken using The Unscrambler v10.2 (CAMO A/S, Oslo, Norway, www.camo.com). PCA is a mathematical procedure in which a collection of data (here NIR spectra) can be reduced to a smaller number of principal components and used to identify regions that have the greatest variation within the sample spectra (Wold et al., 1987; Martens and Næs, 1989). In this study, PCA was used to address three objectives. Firstly, the PCA scores plot for the individual NIR spectra was used to determine the natural grouping of the samples and the segregation of each taxon based on the first two principal components. Secondly, the loadings plot was used to determine the most important NIR spectral regions that accounted to the variation among the samples. The wavelength regions that had the largest contribution to the PCA were the regions that were within the lower and upper bound of the correlation loadings, which show the correlation between the PCs and the original variable. Lastly, PCA was used to extract the first 20 principal component (PC) scores for each individual within the hybrid group. These scores were then used to build a model to discriminate between the three taxa (i.e. the two parental taxa and the hybrid of these taxa) using PLS-DA (explained below). The use of the principal components instead of the original spectra reduces the dimensionality of the data set and improves the stability of the model (Næs et al., 2002; Brereton, 2003).

The PLS-DA calibration model was developed to assign the samples to a taxon within the hybrid group. In this analysis, each taxon (Y-vector) was assigned an arbitrary integer value: CT = 0, hybrid = 1 and spotted gum = 2. These arbitrary integer values assigned to each taxon were regressed against the 20 components from the PCA via a PLS regression (Martens and Næs, 1989). Segmented cross-validation was used to optimise the number of PLS factors used in each model. The number of factors provides an indication of the model’s predictive potential (Richardson et al., 2004). A cross-validation procedure was undertaken using a Monte-Carlo simulation in which the
population was arbitrarily divided into a small number of groups and values in one group was predicted based on calibrations developed from the remaining groups (Foley et al., 1998). In the case of this study, each validation set comprised 11 samples. The ability of the model to predict the removed calibration samples was indicated by the SECV, where models with relatively lower SECV values had better fit which determines the level of accuracy of the model in predicting samples within the calibration set (Wold et al., 1987; Meglen, 2005). The coefficient of determination ($r^2$) describes the correlation between the variation in the response variable (taxa) and the predictor variables (PC loadings) and was also used as an indicator of model fit. Using the best mathematical pre-treatment, samples were assigned to each taxon as follows: samples with predicted Y-values ($Y_p$) ≤0.5 were assigned to CT, >0.5 ≤ $Y_p$ ≤1.5 were assigned to hybrid, and >1.5 <$Y_p$ < 2.5 were assigned to spotted gum. The accuracy rate for a taxon was reported as the percentage of individuals within a taxon that were correctly predicted by the PLS-DA models. The assignment rate for a hybrid group (marginal mean) was the mean assignment rate of the three taxa within a hybrid group.

4.3.5. Comparison of the NIR spectra of different leaf tissue preparation, age and instruments

The three hybrid groups, CCV, CCC and CH were used in all three comparisons, namely: different instruments, NIR1 and NIR2; different tissue preparations, fresh and thawed; and different ages, four and eight-month-old. Three-way NIR models were constructed to discriminate the hybrids from the parents, but note that the CCV hybrid group included an advanced generation hybrid BCT, which was evaluated separately with both the parents in PCA and PLS-DA. Each model represents a hybrid group as follows: 1) CT, $F_{1C}$ and CCC for the CCC group; 2) CT, $F_{1V}$ (or BCT) and CCV for the CCV group; 3) CT, $F_{1H}$ and CH for the CH group. All analyses were undertaken after the standardised mathematical pre-treatment identified earlier, was applied. The assignment rate of each taxon and model fit (indicated by $r^2$ and SECV values) were evaluated to determine which treatment, i.e. fresh or thawed, was better to discriminate the hybrid and parental taxa. The hybrid group that had the highest assignment accuracy by using the standard mathematical pre-treatment was also identified. In addition, the key wavelength regions within the spectra that discriminated the hybrids and the
parental taxa were identified for each hybrid group by examining the correlation loadings plot in the PCA, and reporting the total explained variation at the first two PCs, in which these important regions contributed.

Data sets 2 and 3 (with NIR1) were used to evaluate the power of fresh and thawed leaf tissues to discriminate between the hybrids and parents (Table 1). Data sets 1 and 2, the NIR profiles of eight-month-old and four-month-old leaf tissues, were used to determine which age is best for identifying hybrids (See Table 1). The NIR spectra of the lower wavelength (NIR1) and the higher wavelength (NIR2) were compared using eight-month-old thawed leaf tissues (Data sets 3 and 4; Table 1).
4.4. Results

4.4.1. Determination of a standard pre-treatment

Four spectral pre-treatments (Gap/SNV, Gap-Segment/SNV, Gap-Segment/SNV/Detrend and Savitzky-Golay/SNV) were compared using the CH hybrid group from Data set 2 to determine which pre-treatment provided the best model fit ($r^2$ and SECV) and highest assignment rates. The four pre-treatments had very similar results, with $r^2$ values in the range from 0.74 to 0.78, SECV in the range from 0.41 to 0.45, and mean assignment rate in the range from 87% to 88% (Table 2) indicating that there was not much difference between the pre-treatment methods. However, the Gap-Segment/SNV pre-treatment had the best model fit, i.e. had highest $r^2$ (0.78) and lowest SECV (0.41), hence this pre-treatment was adopted in subsequent analyses.

4.4.2. Effect of leaf tissue storage

The PLS-DA model of the CH hybrid group based on NIR1 of fresh leaf tissues (Data set 2) had a better model fit (Table 3, $r^2 = 0.78$; SECV = 0.41) and classification rate (mean assignment = 86%; SD = 12) than the model for thawed leaf tissues (Data set 3; $r^2 = 0.70$; SECV = 0.49; mean assignment = 83%; SD = 15). Results were similar for the two other hybrid groups, CCC and CCV (Table 3). Because the number of samples assessed in Data set 2 was larger ($n = 840$) than Data set 3 ($n = 519$), the fresh tissue analysis was re-run using the subset of individuals used in the Data set 3 analysis. In this case, even using the smaller sample, the PLS-DA model of the CH hybrid group had a similar model fit ($r^2 = 0.79$; SECV = 0.41; data not shown) to the analysis based on thawed leaf tissue (see above). All PLS-DA models of the three hybrid groups using spectra obtained from fresh tissue were better than those based on thawed tissue (Table 3).
Chapter 4: Hybrid NIR

Table 2. Summary statistics from principal component analysis of NIR1 spectra from eight-month-old fresh leaf tissue for the CH hybrid group comprised of the taxa *C. torelliana* (CT), *C. henryi* (CH) and the interspecific hybrid (*F*$_{1H}$) for four spectral pre-treatments (see text for pre-treatment conditions). The assignment rates of the different taxa obtained by partial least square discriminant analysis are also presented. The optimal pre-treatment given in bold was chosen as standard.

<table>
<thead>
<tr>
<th>Pre-treatment</th>
<th>$r^2$</th>
<th>SECV (a)</th>
<th>Assignment (%)</th>
<th>Assignment rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CT (n=150)</td>
<td><em>F</em>$_{1H}$ (n=78)</td>
</tr>
<tr>
<td>Gap/ SNV(^a)</td>
<td>0.75</td>
<td>0.44</td>
<td>81</td>
<td>100</td>
</tr>
<tr>
<td>Gap-Segment/SNV</td>
<td>0.78</td>
<td>0.41</td>
<td>81</td>
<td>100</td>
</tr>
<tr>
<td>Gap-Segment/SNV/De-trend</td>
<td>0.75</td>
<td>0.44</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Savitzky-Golay/ SNV(^c)</td>
<td>0.74</td>
<td>0.45</td>
<td>81</td>
<td>100</td>
</tr>
</tbody>
</table>

4.4.3. Effect of leaf tissue age

The PLS-DA model of the CH hybrid group based on spectra acquired using NIR1 from the fresh eight-month-old leaf tissue (Data set 2) had a better model fit (see above) than for the spectra acquired from the fresh, four-month-old leaf tissues (Data set 1; $r^2 = 0.46$; SECV = 0.65; mean assignment rate = 63%; SD = 27; Table 3). The PLS-DA models of the CCC and the CCV hybrid groups based on Data set 2 also provided a better model fit and higher assignment rates for the CCC and CCV hybrid groups than the PLS-DA models based on Data set 1 (Table 3).
Table 3. Summary statistics of PLS-DA of NIR spectra obtained on various tissue types at two different NIR wavelength ranges. Model fit (number of factors, correlation coefficient ($r^2$) and SECV) and correct assignment rates of each taxon in the different hybrid groups were obtained by PLS-DA using the PCA scores of the first 20 components in a PCA as the X-matrix and arbitrary integer value of each taxon as the Y-vector (CT = 0, hybrid = 1 and spotted gum = 2). Note that in the CCV hybrid group, an advanced generation hybrid BCT was included. The data set in bold was used in all final interpretations of assignment rates of each taxon because it had the highest model fit.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Method</th>
<th>Hybrid group</th>
<th>Factors</th>
<th>$r^2$</th>
<th>SECV</th>
<th>Assignment rate</th>
<th>Mean</th>
<th>SD</th>
<th>Number of samples (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fresh, four months, NIR1</td>
<td>CH</td>
<td>6</td>
<td>0.46</td>
<td>0.65</td>
<td>48</td>
<td>94</td>
<td>48</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCC</td>
<td>7</td>
<td>0.61</td>
<td>0.51</td>
<td>34</td>
<td>93</td>
<td>81</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV</td>
<td>7</td>
<td>0.44</td>
<td>0.61</td>
<td>59</td>
<td>98</td>
<td>35</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV BCT</td>
<td>7</td>
<td>0.47</td>
<td>0.69</td>
<td>59</td>
<td>100</td>
<td>41</td>
<td>67</td>
</tr>
<tr>
<td>2</td>
<td>Fresh, eight months, NIR1</td>
<td>CH</td>
<td>6</td>
<td>0.78</td>
<td>0.41</td>
<td>82</td>
<td>100</td>
<td>77</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCC</td>
<td>5</td>
<td>0.75</td>
<td>0.46</td>
<td>71</td>
<td>72</td>
<td>86</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV</td>
<td>10</td>
<td>0.81</td>
<td>0.36</td>
<td>86</td>
<td>99</td>
<td>84</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV BCT</td>
<td>8</td>
<td>0.82</td>
<td>0.4</td>
<td>84</td>
<td>100</td>
<td>87</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Mean for Data set 2\textsuperscript{a}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80</td>
<td>90</td>
<td>82</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>SD for Data set 2\textsuperscript{a}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>16</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Thawed, eight months, NIR1</td>
<td>CH</td>
<td>9</td>
<td>0.7</td>
<td>0.49</td>
<td>73</td>
<td>100</td>
<td>75</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCC</td>
<td>9</td>
<td>0.55</td>
<td>0.57</td>
<td>74</td>
<td>41</td>
<td>67</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV</td>
<td>10</td>
<td>0.51</td>
<td>0.58</td>
<td>69</td>
<td>100</td>
<td>38</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH</td>
<td>9</td>
<td>0.7</td>
<td>0.49</td>
<td>73</td>
<td>100</td>
<td>75</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCV BCT</td>
<td>10</td>
<td>0.54</td>
<td>0.63</td>
<td>70</td>
<td>100</td>
<td>49</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH</td>
<td>7</td>
<td>0.74</td>
<td>0.45</td>
<td>84</td>
<td>88</td>
<td>74</td>
<td>82</td>
</tr>
</tbody>
</table>
Table 4. Summary statistics of PLS-DA of NIR spectra obtained on various tissue types at two different NIR wavelength ranges. Model fit (number of factors, correlation coefficient ($r^2$) and SECV)) and correct assignment rates of each taxon in the different hybrid groups were obtained by PLS-DA using the PCA scores of the first 20 components in a PCA as the X-matrix and arbitrary integer value of each taxon as the Y-vector (CT = 0, hybrid = 1 and spotted gum = 2). Note that in the CCV hybrid group, an advanced generation hybrid BCT was included. The data set in bold was used in all final interpretations of assignment rates of each taxon because it had the highest model fit.

<table>
<thead>
<tr>
<th></th>
<th>CCC</th>
<th>CCV</th>
<th>CH</th>
<th>CCV BCT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Thawed, eight months, NIR2</td>
<td>0.77</td>
<td>0.86</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.39</td>
<td>0.37</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td>Mean for Taxon$^b$</td>
<td>80</td>
<td>88</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>SD for Taxon$^b$</td>
<td>9</td>
<td>19</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Assignment rates of F$_1$ only

$^b$ does not include four-month-old data.
4.4.4. Effect of spectral range

The PLS-DA model of the CH hybrid group based on spectra acquired using NIR2 from thawed leaf tissues (Data set 4) had a similar fit and assignment rate ($r^2 = 0.74$; SECV = 0.45, 82%; SD = 7; Table 3) to the model based on spectra acquired using NIR1 (Data set 3; $r^2 = 0.70$; SECV = 0.49; mean assignment rate = 83%; SD = 15). PLS-DA models of the CCC and CCV hybrid groups based on Data set 4 had better fit than Data set 3 (Table 3), suggesting that NIR2, with longer wavelength is better at assigning taxon. However, NIR2 was not available when the NIR spectra of the fresh leaf tissues were collected, and hence the comparison cannot be tested with the available data here. Based on the above comparisons (leaf tissue preparation, age effects and spectral range effects) the NIR2 with the eight-month-old fresh leaf tissues may offer the highest power to diagnose hybrids.

4.4.5. Classification power of eight-month-old fresh leaf tissues measured with NIR1

Mean assignment rates for a hybrid group (i.e. average assignment rates for the two parental taxa and F$_1$ hybrid) were used to assess the hybrid group in which taxa diagnosis could be performed most reliably. The CCV hybrid group had the highest mean assignment rate (90%; SD = 8; Table 3), followed by CH (86%; SD= 12) whereas the CCC hybrid group had the lowest mean assignment rate (76%; SD = 8).

Looking at taxon mean assignment rates, the F$_1$ hybrids had the highest mean assignment rate (90%; SD = 16; Table 3), the spotted gums second (mean assignment rate = 82%; SD = 5) while the CT had the lowest (mean assignment rate = 80%; SD = 8) indicating that the CT was the most difficult taxon to resolve.

Among the F$_1$ hybrids, the F$_{1H}$ had the highest assignment rate (100%; Table 3), followed by F$_{1V}$ (99%), whereas F$_{1C}$ had the lowest assignment rate (72%). The assignment rates of the spotted gum parents ranged from 77 to 86% while the CT parent ranged from 71 to 86% depending on the hybrid group involved (Table 3).
Figure 1. NIR1 average absorbance plot (after Gap-Segment and SNV pre-treatment) of eight-month-old fresh tissue of the CH hybrid group (*C. torelliana* (n = 150), Hybrid (n=78), *C. henryi* (n = 127)).

### 4.4.6. Diagnostic regions in the NIR spectra

Average absorbance plots for the CH hybrid group using Data set 2 (fresh eight-month-old tissue) showed subtle differences between the three taxa at around 1395 nm and 1495 (Figure 1) with the hybrids F<sub>1H</sub> showing the lowest absorbance in this region. These absorbances are associated with the C-H second combination band of methylene and the O-H stretch (1<sup>st</sup> overtone) of hemicellulose respectively (Schwanninger *et al.*, 2011; Workman Jr and Weyer, 2012). No specific assignment of these absorbances is made with respect to the leaf tissue. The CCC and CCV hybrid groups also had their largest deviations in these same regions (Appendix 3 a-c). For the PCA loadings
(Figure 2) there was a major peak associated with water at 1428 nm. This peak was also evident in the other hybrid groups (Appendix 3a-c).

![Principal component loadings plot](image)

**Figure 2.** Principal component loadings plot of the first two components derived from NIR1 (after Gap-Segment and SNV pre-treatment) of eight-month-old fresh tissue of the CH hybrid group (*C. torelliana*, F$_{1H}$ hybrid and *C. henryi*).

### 4.4.7. Principal component analysis of fresh leaf tissue using NIR1

A PCA was used to determine the variation contributed by the first two principal components and to determine the relative differences and similarity between the taxa and individuals using principal component scores based on NIR spectra. Principal component analysis of eight-month-old fresh leaf tissue measured with NIR1 on the CH hybrid group (Figure 3) indicated that the first principal component explained 91% of the variation and the second component explained an additional 7% of the variation. For PC1, the F$_1$ individuals were grouped together and well-separated from the parental taxa. There was also a moderate separation of the individuals belonging to the CT and the CH taxa but there was some overlap (Figure 3), hence the greater difficulty in
accurately assigning parents. The CT parent was closer to the hybrid than the CH parent in the principal component space at PC1. For PC2, the hybrids formed a tight cluster (less variable), but the parental taxa, CT and CH were more scattered, indicating that they are more variable than the hybrids.

![Scores plot of the first two principal components derived from NIR1 (after Gap-Segment and SNV pre-treatment) using eight-month-old fresh tissue of the CH hybrid group (Blue = *C. torelliana* (n=150), Red = Hybrid (n=78), Green = *C. henryi* (n=127)).](image)

Other hybrid groups exhibited a similar pattern, with the first principal component explaining most of the variation (Appendix 5). While the F$_{1H}$ was well separated from the parental taxa, the F$_{1V}$ and the F$_{1C}$ were closer to one or other of their parents, with a tendency for overlap (Appendix 5). The F$_{1V}$ individuals tended to be closer to their spotted gum parent than their CT parent, whereas the F$_{1H}$ and F$_{1C}$ individuals tended to be closer to the CT parent. There was also a trend for both CCC and the CCV to form distinct groupings, separated from CT, but, as in the case of the CH hybrid group, there was some overlap. For PC2, the hybrids tended to form a tighter cluster, more so than the parental taxa, which exhibited more variability (Appendix 5).
4.4.8. PCA and assignment testing including the F$_{1}$ and advanced generation BCT

To determine the potential to separate advanced generation hybrids from the parental species as well as first generation hybrids, a PC analysis was conducted on the CT, CCV, F$_{1V}$ and BCT. Visual examination of the PCA scores plot (Figure 4) indicated that on PC1, the BCT and the F$_{1V}$ were close together while the parental taxa, CT and CCV were close together. The F$_{1V}$ was between CCV and BCT on PC1 (Figure 4).

![Figure 4. Scores plot of the first two principal components derived from NIR1 (after Gap-Segment and SNV pre-treatment) using eight-month-old fresh tissue of the CCV hybrid group (Blue = C. torelliana (n=150), Red = F$_{1V}$ hybrid (n=144), Green = backcross hybrid BCT (n=30) and Brown = C. c. variegata (n=141)).](image-url)

A PLS-DA analysis was also conducted on this combined data set. The BCT and the F$_{1V}$ had high assignment rates, 100% and 80%, respectively; whilst the parental taxa had low assignment rates, 39% and 29% for CT and CCV, respectively (data not shown).
4.4.9. Exclusion analysis using two or four parental taxa in the calibration group

The use of three-way models requires specific inclusion of known hybrids, which means that models need to be generated for each possible hybrid combination. But there may be instances in which the interest is to identify an individual whether it belongs to a hybrid or the parental taxa regardless of the specific hybrid class. Using PLS-DA calibration models generated without including the hybrid taxa in the calibration set resulted in three models with two parents in the calibration group (CT-CH, CT-CCC and CT-CCV) and one model with four parents in the calibration group (CT-CH-CCC-CCV). In each case, the CT spectra were assigned the integer value 0 and the remaining taxa were assigned the value 2. These models resulted in the hybrid predictions shown in Table 4. The results show that the CT-CH calibration model identified all the $F_1$ hybrid combinations as being model outliers and they are hence assigned as hybrids, without specifically identifying which hybrid. The other extreme was observed for the CT-CCV calibration model where only 27.7 % of all the $F_1$ hybrids are model outliers and consequently assigned as hybrids. However, when the PLS-DA calibration is developed using all four parent taxa, the model identified 90% of the $F_{1H}$, 91% of the $F_{1C}$ and 95% of the $F_{1V}$ hybrid. This suggests that hybrids can be selected by exclusion from calibrations of pure parental taxa rather than specific inclusion into the calibration model, thereby improving the power of the combined model to identify any possible hybrid combination.
4.5. Discussion

4.5.1. Utility and optimisation of NIR for *Corymbia* hybrid identification

The present study found that multivariate analysis (PCA and/or PLS-DA) of NIR spectra acquired from leaf tissue shows considerable promise for hybrid diagnosis of *Corymbia* taxa of interest for gene pool management in subtropical Australia. These hybrids are difficult to distinguish from parental taxa using morphology (Abasolo *et al.*, 2012) and near infrared reflectance spectroscopy may be a more reliable method for *Corymbia* hybrid identification than morphology in many situations. Using optimised conditions and three-way models, mean rates of assignment for taxa in a hybrid group (i.e. the parents or the F$_1$) ranged between 76 and 90%. For the F$_1$ hybrids, the mean assignment rates ranged from 72 to 100% depending on the hybrid group (CH, CCV or CCC) and were higher than that based on an optimised morphological character set (60% to 76%) (Abasolo *et al.*, 2012). It is likely however, that morphology and NIR assessment can be combined to improve taxon diagnosis. This combination could be a cost effective method to screen large numbers of *Corymbia* seedlings, both in a field or nursery situation.

In the discussion below, recommended conditions were derived for further development of NIR screening for hybrid diagnosis in *Corymbia* as well as providing chemical or genetic rationale for observed effects. It also proposes further developments required for practical implementation of the technology for use in field or nursery situations such as the development of global prediction models for different environments. In addition, other major findings are considered, that hybridisation tended to deviate from the parental taxa, with differences being magnified even more so in advanced generation hybrid than F$_1$. This raises the challenge to explain how the process of hybridisation gives rise to such an effect and whether it is linked to the disruption of biogenic pathways which might be anticipated with hybrid incongruence. It is also noted that, as with morphological markers, hybrids from more genetically divergent parents tended to be more reliably distinguished, and again, this divergence can be explained by hybrid
genetics in which the hybrids are outside the range of the parents and are spatially discriminated by NIR. The application of NIR to hybrid diagnosis is also considered, in two scenarios relevant to the practise of gene pool management of *Corymbia* in subtropical Australia, the planting of CT as an ornamental tree, and the deployment of F1 hybrids in plantations.

### 4.5.2. Recommended conditions for *Corymbia* hybrid diagnosis

The study considered the factors of leaf age, leaf storage and spectral range upon the efficacy of classification. Although unavailable for data acquisition in this study, it is proposed that the higher wavelength NIR instrument (NIR2) for acquiring spectra of eight-month-old fresh tissue may be the best combination of instrument and leaf age to provide the highest assignment accuracy of *Corymbia* taxa. Nonetheless, most of the comparative analysis of the study was based on the NIR1 instrument because it was the only one available at the time of fresh leaf assessment. In the future, it is recommended that there will be further evaluation of the NIR2 spectral range for use on fresh tissues as well as an evaluation of the power of combined (concatenation) of the NIR1 and 2 data.

The use of fresh foliage provided higher assignment rates than thawed leaf tissues (Table 3). This may be due to the production of common degradation products arising in all samples due to the freezing and thawing process that masks other diagnostic differences. For example, oxidation increased the levels of total phenolics in grapes following long periods of freezing (Cynkar *et al.*, 2009). In the samples there was a change in the pigmentation of the leaves when thawed and this may be attributed to the oxidation of the phenolics and chlorophyll. Although the use of fresh tissue may be more optimal, there may be instances where the use of thawed tissues is unavoidable. In these cases, hybrid identification can still be reasonably successful, i.e. 81-82% mean assignment rates for the three hybrid groups.
In this study, the discrimination of hybrid seedlings was more reliable on the older seedling, eight rather than four months of age. This may be related to changes in the chemical profile associated with ontogenetic development. Changes in terpene composition associated with ontogenetic development are well documented in *Eucalyptus* and *Melaleuca* (Li *et al*., 1995, 1996; Russell and Southwell, 2002; Russell and Southwell, 2003) and are likely to have an adaptive or ecological basis (Loney, 2007; O’Reilly-Wapstra *et al*., 2011). It may be that specialisation in leaf chemistry must occur before reliable diagnosis of taxa will be possible.

**4.5.3. Hybridisation causes deviation in NIR spectra of *Corymbia***

An examination of the genetic difference between CT and the three spotted gum taxa, CCC, CCV and CH based on genetic markers, revealed that CH and CT were the most genetically divergent taxa, while CCC and CT were the closest (Shepherd *et al*., 2008a). Using NIR1, the F$_{1H}$ (the hybrid with the most genetically divergent parents) was distinguished with the highest accuracy (100%) of any F$_1$ in the three hybrid groups. On the other hand, the F$_{1C}$ taxon which had the most genetically similar parents had the lowest assignment accuracy, 76%. The geographic proximity of CT and CCC creates a potential for gene flow in nature and may explain the greater difficulty in resolving F$_{1C}$. It is possible that some parents used in the crosses were not pure, and classification into F$_1$ class may have included some advanced generation hybrids. This may have expanded the space occupied by the F$_{1c}$ cluster in the PCA, making the F$_{1C}$ more difficult to resolve. A similar trend, where hybrids of more divergent parents were more easily resolved was also found in a recent study of *Corymbia* seedling morphology (Abasolo *et al*., 2012). The relative power to resolve the hybrid taxon in each hybrid group reflected the visibly assessable divergence in the average spectral profiles, with those of the F$_{1H}$ showing greater amplitude divergence at key diagnostic regions from its parents than the F$_1$ in other groups (Figure 1 and Appendix 3). This may indicate that somehow a larger genetic distance between parents is translated into greater divergence in morphology and NIR spectra, and in the case of NIR spectra, presumably due to divergence in the underlying chemistry.
In three-way comparisons, the hybrid was usually the best resolved taxon (72 - 100% assignment rates), with parental assignment rates tending to be lower (71 - 86%). For all hybrid groups, the absorbance plots of the parental taxa overlapped across many wavelengths while the hybrid clearly diverged at specific locations (Figure 1 and Appendix 3a-c). Also, on the PCA scores plots, the hybrid taxa were distinctive whilst some individuals from the parental taxa overlapped particularly on the PC1 axis (Figure 3 and Appendix 5a-c). Hence, the parental taxa appeared to be more similar to each other in their NIR profile than to their respective F1 hybrid. Perhaps hybridisation causes disruption to biogenic pathways for the synthesis of terpenes, phenolics, or other compounds that the NIR spectra respond to. Certainly, hybridisation leads to distinctive terpene profiles in eucalypts (Shepherd, 1998). In the case of eucalypts, high proportions of novel terpenes not detected in the parents were present in the hybrid (gamma-terpinene), as well as large qualitative changes in other components. These dramatic changes in terpene profiles were thought to be due to genetic incongruence in the hybrid (Shepherd et al., 1999). Disruption to biogenic pathways for terpenes or other secondary metabolites due to hybridisation may account for NIR spectral divergence relative to parental pairs.

Differences in NIR spectra may be more effective for hybrid diagnosis than direct chemical analysis. For example, Hayes et al. (2013) found that CT, F1H and F1C cannot be differentiated by their terpene profiles. These chemical differences may be very small and difficult to detect using wet chemistry but, nonetheless, they may have great impact on the NIR spectra. Near infrared spectroscopy models were also more robust than oil models in identifying hybrids between E. globulus and E. nitens (Humphreys et al., 2008). Perhaps NIR profiles are more responsive to subtle chemistry changes that are noise in compositional analysis.

Hybrid incongruence (or outbreeding depression) may be more pronounced in second or later generation hybrids (Frankham et al., 2002) and this may explain the more extensive deviation of the BCT from its parents on the PC1 axis. The shape of the absorbance curves for F1 and the BCT were similar but there was more deviation of the BCT. These results may indicate over-expression or under-expression due to hybridisation resulting in a breakdown in the biosynthetic pathway of a compound or group of compounds. The better resolution of the BCT compared with the F1 hybrid
was unexpected given the tendency for backcrosses to more closely resemble their recurrent parent than the $F_1$ (Potts and Dungey, 2004).

### 4.5.4. Application of NIR analysis in Scenario 1 – distinguishing intersectional $F_1$ hybrids between CT and spotted gums

The extensive planting of CT as ornamental or amenity trees throughout south east QLD and northern NSW is one situation where discrimination of $F_1$ hybrids may be desirable for gene flow management purposes. Spontaneous production of $F_1$ and advanced generation hybrids has been shown to occur in these situations (Shepherd et al., submitted). Reliable diagnosis of hybrids was not possible using field morphology and studies of controlled cross hybrid families show that the hybrids are intermediate but tending toward the CT parent for many morphological attributes (Abasolo et al., 2012). Although genetic markers can be used reliably to identify hybrids (Shepherd et al., submitted), and quantification of hybridisation rates may be possible on a research scale, the use of markers on a large scale analysis is still likely to be precluded because of high costs. A cheap, reliable and field-based hybrid identification method would be highly desirable in such situations.

Because the present study has shown that taxa from hybrid groups can be discriminated reliably using three-way NIR models (mean assignment rates were between 76 and 90% depending on the hybrid group) this would appear to be a valuable way to detect hybrids and estimate hybridisation rates. But how would it be deployed in practise? The approach requires reference spectra for each taxon before classification can be undertaken. In the above example, it would require reference samples from each parental taxon and a reference hybrid, of the same physiological age and growing in the same environment. Genotyped individuals from controlled crosses could be used as reference population in using NIR to identify hybrids. The number of individuals to be used as a reference population may vary depending on the genetic make up of the species and growing conditions, and it may range from several hundreds to thousands. A reference population is needed because NIR is sensitive to change in chemical composition brought by change in growing environment (i.e. Li et al., 1995; 1996).
This change has been demonstrated by analysing NIR of three and eight month old seedlings.

Depending on the situation, exclusion analysis based on NIR may help circumvent the need for a hybrid reference population. It has been shown that the use of a calibration model which only includes parental taxa can be highly effective. This method is less specific as it may not identify a particular hybrid taxon but it may be useful when the only interest is to identify a hybrid regardless of the hybrid class (i.e. whether $F_{1C}$, $F_{1V}$, etc). Exclusion approaches using either two parental taxa or four parental taxa in the calibration group were tested as these have potential to be used as more general method to identify hybridisation and the model using four parental taxa resulted in more robust prediction of the hybrids than the model using two parental taxa (Table 4). Identification of putative hybrids can then be performed by identifying samples whose PLS-DA predicted values lie outside the models range (either high or low). These samples can be viewed as being potential hybrids, but it is not possible to identify the specific hybrid class by this exclusion method, unless other information is available to exclude or include parents.

Another way that NIR may find application in these situations is to raise seedlings collected from trees that have putatively undergone cross pollination and to grow them alongside reference populations in the nursery. Even then, care must be taken to ensure that environmental differences are minimised. One way to overcome the difficulty with environmental interaction and age effects may be the development of a sufficiently robust global model (see last section) which may allow application of NIR to field diagnosis without the need to generate new reference populations for each situation. Provided that such libraries are generated, NIR may become a valuable practical aid to the field diagnosis of hybrids where both parents are present.

4.5.5. Application of NIR in Scenario 2 – Plantations of $F_1$ hybrids

A second situation where it may be necessary to determine hybridisation rates on a large scale for gene flow monitoring purposes is where the $F_1$ hybrid is planted experimentally or deployed operationally. In this case, it may be necessary to identify
backcross (BCT) and $F_2$ from the parental taxa, the $F_1$ hybrid and the relevant spotted gum parent. Such situations exist presently around experimental plantations of the hybrid in NSW and where amenity plantings of CT have produced mature $F_1$ progeny, and thus are able to cross pollinate. Again, until global models can be generated, the more immediate applications of NIR may be limited to the collection of seedlots from trees that are candidates for cross pollination, and raising the seedlings alongside reference populations in a nursery under standardised environmental conditions.

### 4.5.6. Potential utility of NIR for *Corymbia* hybrid identification in a range of environments

The utility of NIR for *Corymbia* hybrid identification has the potential to be extended beyond growing plants in a common environment so as to broaden the use of NIR for assessing gene flow. This could be attained by building a global model which entails collection of NIR data in a range of environments and a range of ages to make the model robust. NIR analysis is largely used in wood property determination (Raymond and Schimleck, 2002; Downes *et al*., 2010; Meder *et al*., 2011), where the NIR spectra are collected across a range of sites that represents the variation in the wood property of interest. The NIR spectra are correlated with a laboratory standard so that a sample with unknown wood property could be predicted. In this case, the robustness of the NIR model to predict a sample depends on the amount of variation captured for that wood property. The same principle could be applied in acquiring a global model for *Corymbia* parental taxa and hybrids beyond four-month-old seedlings. The amount of NIR variation across environments and across ages might be less because of the high heritability in eucalypt chemistry (Shepherd *et al*., 1999; Apiolaza *et al*., 2005; Poke *et al*., 2006; Stackpole *et al*., 2011), thus sampling a limited number of sites may potentially capture total variation in chemistry which makes it easier to build a global model. The wider application of NIR for *Corymbia* hybrid identification will provide a non-destructive, rapid, cost-effective and accurate tool that could be used in the field or in the nursery for gene flow assessment.
4.6. Acknowledgements

This research is a part of the PhD project of M Abasolo on Gene flow Management funded by the Cooperative Research Centre for Forestry and the International Postgraduate Research Scholarship. I thank CSIRO Plant Industry for the loan of the handheld NIR equipment used in this research and Queensland DAFF for providing the seeds used in the experiment and the use of their Gympie nursery where the seedlings were raised. I also thank John Oostenbrink and Tracy Menzies for nursery maintenance and assistance in data collection, and Myrna Deseo for critical comments on the earlier version of the manuscript.
Chapter 5. Conclusions/Recommendations

This thesis reports upon the patterns of phenology and the development of reliable methods that will allow better risk assessment and management of gene flow between locally exotic plantings and native CCV. In particular, it demonstrated that the peak flowering time in CCV is under genetic control and that the amount of pollen produced in young, closely spaced plantations is low relative to neighbouring native forests (Chapter 2). The studies reported in this thesis also evaluated the potential for morphology (Chapter 3) and near infrared reflectance spectroscopy (Chapter 4) properties to identify intersectional hybrids between CT and the spotted gums, and/or among species of spotted gums. I consider how these new developments can refine a risk assessment for the two main situations of concern for gene flow in the subtropics, then look at how tools of hybrid diagnosis might be best deployed in practice, then finish with recommendations for further research.

The new data on genetic and environmental factors influencing timing of anthesis and degree of synchrony among taxa allow for some refinement of the risk assessment for plantations of Corymbia spp. in the subtropics beyond that of Barbour et al. (2008). At the time of Barbour et al.’s (2008) assessment, the likelihood of gene flow from planted CCV to native CCV was thought to be high based on their close taxonomic affinity and broad flowering period, but that the overall risk level was reduced by the view that any gene flow to a widespread species would have less impact than gene flow to a rare species (Barbour et al., 2008; Barbour et al., 2010). In the case of planting the Corymbia F<sub>1</sub> hybrid with CT as the female parent, the overall risk was thought to be high (Barbour et al., 2008). Gene flow risk may be high because although the likelihood of introgression was thought to be lower, the potential impact of gene flow on ecology was thought to be high because of the weedy nature of CT. Each of these two situations will be reviewed in light of my studies and other studies that have occurred in the interim.
Risk of gene flow in Scenario 1 – planting of northern CCV race in the south of its range

The recognition of genetically determined partial asynchrony in peak flowering of CCV suggests that the likelihood of gene flow may be lower than previously thought in situations where trees from the northern race are planted within the range of the southern geographic race. The likelihood may be further moderated by the observation of low levels of pollen production in plantations of young age (i.e. less than 14 years old) relative to nearby native stands, and the apparent low conversion of buds to flowers, so that the source to sink ratio may be low.

While the likelihood of gene flow may be low, the impact may be higher than previously thought. This is due to the increasing recognition that within CCV there may be distinct geographic races with adaptive differences. For example, there is accumulated evidence that the northern population is distinct from the southern population in microsatellite data (Shepherd et al., 2008) and for potentially important adaptive traits like disease tolerance (Johnson et al., 2009; Pegg et al., 2011), frost tolerance (Larmour et al., 2000), terpene chemistry (Asante et al., 2001) and peak flowering time (this study). If these traits and genetic differences are indeed important for survival and fitness in the local environment, then there may be ecological and evolutionary consequences for the transfer of mal-adapted alleles from plantings. This suggests that the impact of gene flow between the CCV races may be greater than previously thought.

The high environmental plasticity in bud abundance makes gene flow risk assessment difficult when considering year to year and site to site variation. It has been shown in this study that bud abundance is highly influenced by the interaction of the region of origin and year, suggesting the need to study each site over a number of years to determine annual variation in bud abundance and rate of conversion of buds into open flowers. In the scarcity of data such as annual and environmental variation in bud abundance and conversion rates of buds into flowers, the likelihood of gene flow must be assumed high until results show otherwise.
Overall, this study indicated a lower likelihood of gene flow from planted CCV than previously thought because of the genetically controlled variation in peak flowering time between the geographic races of CCV, the lower abundance of bud crop from the plantations (aged <14 years old), and apparent low conversion of buds to flowers in the plantations. The impact of gene flow may be higher than previously thought due to the increasing evidence that the genetic profile and adaptive traits differ between the two geographic races of CCV. Such gene flow may have negative ecological and evolutionary consequences (Allendorf et al., 2001; Ellstrand and Elam, 1993; Ellstrand et al., 1999). The overall gene flow risk appears to be moderate with low likelihood and greater impact. Long term monitoring of the locally exotic CCV is needed because climate change and other environmental factors may shift flowering periods and this may increase the level of flowering synchrony, thus increasing the likelihood of gene flow.

**Risk of gene flow in Scenario 2 – Planting F₁ hybrids**

The earlier assessment of gene flow risk due to the planting of the *Corymbia* intersectional F₁ hybrid (CT x spotted gums) can be re-evaluated in light of data on the flowering synchrony between F₁ and the parental taxa in northern NSW (Appendix 6), recent studies on the crossability and viability of seedlings in the intersectional hybrids (Dickinson et al., 2012; Dickinson et al., 2013), and inferences on longer term viability of spontaneous advanced generation hybrids (Shepherd et al., submitted).

The review of Barbour et al., (2008) indicated that the risk of gene flow from F₁ to native spotted gums may be low because of possible outbreeding depression, but the impact may be high because of the weedy nature of CT. The high impact is moderated by the low likelihood which lowers the overall risk. However the likelihood of gene flow between the F₁, CT and the native spotted gums may be higher than previously thought as there was an observed flowering synchrony between the F₁ and the CT (Appendix 6), but not between the F₁ and the spotted gum parental species, probably due to small sample size. However, backcrosses to spotted gum parents were found in a molecular analysis of the offspring of a spontaneous F₁ hybrid, indicating that backcrossing to both parental taxa can occur in some years (Shepherd et al., submitted).
The likelihood of gene flow from the F$_1$ to different Corymbia sections, and the southern race of CCV may be lower than previously thought because of crossing incompatibilities which increase with increasing taxonomic distance between the parents (Dickinson et al., 2012). Backcrosses of an F$_1$ to CCC, (CT x CCC) x CCC, were also observed to have higher seed viability than backcrosses to CCV, (CT x CCV) x CCV (Dickinson et al., 2013; Figure 1B p. 76), which is expected because CCV is more genetically divergent from CT than CCC (Shepherd et al., 2008a). The crossing incompatibilities between more genetically divergent Corymbia species observed by Dickinson et al. (2013) indicates that the risk of gene flow to other Corymbia sections like C. gummifera (Section Gummiferae; (Parra-O. et al., 2009)), C. intermedia (Section Intermediae; (Parra-O. et al., 2009)) and to CCV, which naturally occur in northern NSW where F$_1$ is planted experimentally is probably lower than to CCV.

The overall risk of gene flow from hybrids may be mitigated by outbreeding depression if the hybrids fail to survive and/or produce viable flowers. Shepherd et al. (submitted) noted the absence of mature advanced generation hybrids at a gene flow study site in northern NSW, despite observations of early vigour in backcross and F$_2$ hybrids grown from seeds collected from a spontaneous F$_1$ hybrid at the site. The early vigour of the advanced generation hybrids, combined with the apparent absence of mature trees, suggests a high mortality for the hybrids, possibly because of poor long term survival due to outbreeding depression. Outbreeding depression in the advanced generation hybrids may be a strong barrier to gene flow mitigating the tendency for the F$_1$ to serve as a genetic “bridge” to facilitate the introgression of genes into the native stands as outbreeding depression may prevent the persistence of advanced generation hybrids. Taking all the information at hand; flowering synchrony, crossing incompatibilities and outbreeding depression, the best estimate of the overall risk of gene flow from planting F$_1$ hybrids may be low to moderate (Potts et al., 2003; Byrne et al., 2011).

This study highlights the importance of gathering as much biological information as possible to assess gene flow and how this information can be used to refine risk assessment. In particular, this study shows how biological information such as the knowledge of population structure in widespread species, the importance of adaptive differentiation, genetically controlled flowering asynchrony, crossing incompatibilities
and outbreeding depression, may be used to refine the existing gene flow risk assessment for *Corymbia*.

**The use of morphology and NIR for hybrid identification in Scenario 2**

Direct measurement of gene flow by detecting hybridisation is valuable for assessing the validity of risk assessments based on inferences from other factors such as flowering synchrony and for determining the need for remedial action. Ideally, hybrid identification would use methods which are cheap, easy to perform, and highly reliable. This would allow assessment of thousands of seedlings which is often necessary to detect low levels of gene flow (Barbour *et al.*, 2002; Barbour *et al.*, 2003).

In Scenario 2, the advanced generation hybrids are the hybrids of interest. This study found that morphology cannot reliably detect advanced generation hybrids because of trait segregation (Chapter 3). However, NIR could identify advanced generation hybrids with high reliability because of deviant spectra (Chapter 4). Therefore, the use of NIR is discussed in relation to gene flow monitoring in Scenario 2 (planting of F\textsubscript{1} hybrids).

NIR technology for *Corymbia* hybrid identification could be implemented using two approaches. Firstly, if there is observed flowering synchrony between the planted F\textsubscript{1} hybrids and the native spotted gums, a nursery based approach may be used for monitoring the level of gene flow into the native stands. Secondly, for routine plantation management and culling of hybrids a field based approach is required.

A nursery based approach is useful for determining priorities for gene flow monitoring and management. Rigorous gene flow monitoring should be implemented within zones where hybridisation rates are expected to be high. The width of such a zone can be determined by constructing a pollen dispersal curve (plotted as distance (x-axis) versus hybridisation rates (y-axis)). It is expected that the pollen dispersal curve will be “fat with a long tail” and characterised by a majority of short dispersal distance and the rarity of long dispersal distance typical of eucalypts (Potts and Wiltshire, 1997).
Conclusions

Therefore the high priority zone may extend less than a kilometre from the plantations (Barbour et al., 2002; Barbour et al., 2003; Potts et al., 2003).

The nursery based approach involves collecting seeds from native mother trees at varying distances from the plantation. The resulting seedlings may be a mixture of open pollinated spotted gums and hybrids and the aim is to separate the hybrids from the pure taxa. In the case of intersectional hybrids between CT and spotted gum, a reference population consisting of the parental taxa (CT and spotted gum) needs to be grown together with the seedlings from native mother trees. The exclusion method described in Chapter 4 could be used to identify the taxa class, i.e. whether a hybrid or a spotted gum. To apply this method in the nursery, a calibration model, based on spectra of the reference population (CT and spotted gum), is used to predict the identity of the seedlings collected from the native trees. From Chapter 4, it is known that the spectra of the parental taxa are more similar than the spectra of any hybrid. Thus, any seedling whose spectra lie outside the model space of the parental taxa in a multi-dimensional space will be considered a hybrid and any seedling whose spectra are within the model space of the parental taxa is considered an open pollinated spotted gum.

A field based approach to using NIR may be more useful than the nursery approach for detecting hybrids in routine plantation management and for determining the rate of hybrid establishment in the wild. As part of active plantation management, a field crew could use a handheld NIR to scan the seedlings established in the native forest adjacent to a plantation and cull the individuals found to have the spectra of a hybrid. An exhaustive survey within the high priority zone (described above) may be required, but inclusion of areas outside this zone may also be necessary especially when the level of gene flow is low.

The field based approach can also be used to determine the number and geographic extent of hybrids in an adjacent forest. The establishment rate for hybrids may be used as an indicator as to whether further management is needed to mitigate gene flow from $F_1$ plantations. For example, if there is low level of establishment, a routine culling may be enough to eliminate hybrids. But if there is high level of establishment, further management may be needed to reduce gene flow such as the establishment of buffer zones consisting of trees that do not interbreed with the native stands (Eldridge et al.,
1994), increasing the isolation distance between the planted *Corymbia* \( F_1 \) hybrids and the native stands than what is currently maintained (i.e. less than one kilometre) or, alternatively, a decision should be made not to plant *Corymbia* \( F_1 \) hybrids at the site.

A field based approach to hybrid identification using NIR would involve scanning fresh leaves in the field and determining whether an individual seedling is a hybrid or a spotted gum. A prerequisite would be the construction of a global calibration model based on spectra of spotted gum, CT and hybrids collected from various environments and a range of ages (described in Chapter 4). Once the global calibration is built, hybrids could be identified by scanning leaf material in the field and determining whether the spectra fell within the model space. Such a system would facilitate large scale screening of seedlings in the field and would allow the full potential of NIR for gene flow monitoring to be realised.

**Recommendations for future research**

1. **Environmental factors related to bud production and bud loss.** In this study, the strongest influence on bud abundance was the interaction of the region of origin and year (Chapter 2). The rate of bud abortion was generally high and variable across years. As a consequence, there was a low conversion of buds to flowers. This implies that long term studies of flowering across the range of sites prescribed for planting spotted gums and *Corymbia* \( F_1 \) hybrids are required to understand the underlying mechanisms. The data in this study cover only two to three years of observation at four sites (see Appendix 6) and precludes any generalisation on environmental factors that may explain bud production and bud abortion. Long term flowering observations on a range of sites would allow a predictive model to be built based on the environmental factors that are important for *Corymbia* bud production, bud development and bud abortion. Such a predictive model may be useful for forest managers seeking to moderate gene flow from *Corymbia* plantings. For example, the work of Keatley and colleagues incorporating 30 years of flowering data enabled them to predict environmental factors that determine bud production on various eucalypts (Hudson *et al.*, 2003; Keatley *et al.*, 2004; Hudson *et al.*, 2011). Such a model
may be adopted for prediction of bud production in *Corymbia*. The two to three years of data obtained from this study is not sufficient to build a predictive model as there is insufficient data to cover a range of seasons and the weather patterns in the two to three years of study were unusual (i.e. more rainfall than usual). Therefore, the data gathered in this study do not reflect the usual conditions experienced at these sites. However, the data obtained here can serve as baseline information for future research on flowering in a range of sites and seasons.

2. **Development of field tools for interspecific hybrid identification.** Tools for reliably monitoring interspecific gene flow between CT and spotted gums and/or their hybrids appear to be possible (Chapters 3 and 4). However, these tools need further development to realise their practical use. For example, global models for NIR need to be developed in order to maximise its use for gene flow monitoring and management (so NIR can be applied in the field, in different environments and to seedlings of different ages). Also, a guide for morphological identification of hybrids in the form of a brochure may be useful, similar to one developed for gene flow management in Tasmania (Barbour, 2010). These developments would better equip researchers and forest managers to monitor gene flow.

3. **Tools for inter-racial hybrid detection.** Tools for identifying the inter-racial hybrid between planted northern CCV and the native southern race are also needed. There is evidence of genetic and adaptive variation between the northern and the southern CCV races (Chapter 2) and these differences may result in negative consequences in the event of inter-racial gene flow. This study found partial flowering synchrony between these two CCV races (Chapter 2) indicating the potential for inter-racial gene flow in northern NSW.

Ideally, gene flow monitoring by hybrid assessment should be done at the seedling stage as it will be easiest to manage by removal prior to maturation when there is risk for potential introgression. At present, there is very little information on the characteristics of inter-racial hybrids of CCV other than on mature natural hybrids of the parental species found in sympatry. Hybrids
between the two races have intermediate leaf mass area (Ochieng, 2009) and other parental differences are evident, such as mature crown architecture (the southern race has a larger crown) and may be useful for hybrid identification in the future.

Inter-racial hybrids cannot be distinguished by morphology especially at seedling stage (Chapter 3). For large scale screening, NIR may be the best option for identifying hybrids. This has yet to be tested but the feasibility of using NIR to identify inter-racial hybrids could be tested by building on the methods in this thesis.

Molecular markers may also aid in detection of inter-racial hybrids. Microsatellite loci developed for Corymbia (Jones et al., 2001; Shepherd et al., 2006) reveal allelic frequency differences between the two geographic races (Shepherd et al., 2008a; Shepherd et al., 2012). The molecular sub-structuring within CCV presents an opportunity to screen for loci with high differentiation between the races (i.e. high $F_{st}$) that may be more informative for inter-racial discrimination. If for a given locus, there are more alleles in the northern than in the southern race (or vice versa), an individual’s genotype can be given a probability of belonging to one or the other race.

However, the scope for using microsatellites for hybrid determination between CH and CCV seems limited because these two taxa were not differentiated at microsatellite loci (Shepherd et al., 2008a; Ochieng et al., 2010). There was more variation within each taxon across sites than between taxa, indicating these are not “good” species. Other methods are needed to detect gene flow between CCV and CH in populations in cases where CCV is planted near native CH populations.

Detailed study of controlled crosses of these hybrids, i.e. CCV (northern race) x CCV (southern race), and CH x CCV (northern race) class would be required to evaluate the utility of morphology, NIR and genetic markers for hybrid identification.
4. **Measuring gene flow rates at plantation boundaries.** This thesis established that morphology can identify hybrids with moderate accuracy and NIR with high accuracy. These tools may be used to detect realised gene flow at plantation boundaries in areas where gene flow is suspected to be occurring (i.e. where there is flowering synchrony between the planted and the native stands) and flowering is abundant. During the conduct of this thesis, the existing plantations were young (less than 12 years since planting). Like the Bonalbo trial (Chapter 2), the plantations flowered at very low levels (See Appendix 6). Managing gene flow from plantations will be required as these plantations mature because high levels of flowering are expected.

The refined gene flow risk assessment and the methods for monitoring gene flow studied will help forest managers to decide on how to manage gene flow from *Corymbia* plantations. The biological information from this study may be used as a guide to assess the risk of gene flow, and the methods for hybrid identification will be useful to validate the assessed risk and monitor gene flow on a site by site basis. Site specific information on gene flow levels will equip forest managers to take actions in order to manage gene flow at specific sites.
## Appendices

### Appendix 1

**Chapter 2: Supplementary Table.** Log scale of the average number of units of each reproductive unit of each region. The average number was based on reproductively active trees on each assessment date.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Year</th>
<th>Date</th>
<th>Initials</th>
<th>Buds</th>
<th>Flowers</th>
<th>Green capsules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal QLD</td>
<td>2010</td>
<td>8 Sep</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Inland QLD</td>
<td>2010</td>
<td>8 Sep</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Subcoastal CQLD</td>
<td>2010</td>
<td>8 Sep</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
## Chapter 2: Supplementary Table

Log scale of the average number of units of each reproductive unit of each region. The average number was based on reproductively active trees on each assessment date.

<table>
<thead>
<tr>
<th>Region</th>
<th>Year</th>
<th>Date</th>
<th>Initials</th>
<th>Buds</th>
<th>Flowers</th>
<th>Green Capsules</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Richmond Region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>8 Sep</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>High Elev NSW</strong></td>
<td></td>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Sep</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>High Rainfall NSW</strong></td>
<td></td>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Sep</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Chapter 2: Supplementary Table. Log scale of the average number of units of each reproductive unit of each region. The average number was based on reproductively active trees on each assessment date.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Year</th>
<th>Date</th>
<th>Initials</th>
<th>Buds</th>
<th>Flowers</th>
<th>Green capsules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal NSW</td>
<td>2010</td>
<td>8 Sep</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13 Oct</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Oct</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>25 Feb</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 Apr</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 May</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Jun</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Jul</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Aug</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Sep</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Oct</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 Nov</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
## Appendix 2

### Supplementary Data

### Chapter 3

Untransformed means and standard deviations of the 26 quantitative characters used to identify *Corymbia torelliana*, *C. variegata*, *C. henryi*, *C. citriodora* and their hybrids with *C. torelliana* as the maternal parent.

<table>
<thead>
<tr>
<th>Character</th>
<th>C. torelliana</th>
<th>C. variegata</th>
<th>CTxCCV</th>
<th>CTxCCV x CTxCCV</th>
<th>CTxCCV x CT</th>
<th>CTxCCV x F2</th>
<th>C. henryi</th>
<th>CTxC</th>
<th>C. citriodora</th>
<th>CTxCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBRN</td>
<td>149 1.248 1.959 149 5.148 2.502</td>
<td>143 3.503 2.960</td>
<td>5 2.800 1.924</td>
<td>35 7.380 3.068</td>
<td>5 4.260 0.636</td>
<td>121 3.488 0.103</td>
<td>5 3.005 0.938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIA10 (mm)</td>
<td>146 4.320 1.119 137 3.430 1.038</td>
<td>140 3.653 0.940</td>
<td>5 3.520 1.173</td>
<td>33 3.767 0.877</td>
<td>5 4.260 0.636</td>
<td>121 3.468 0.103</td>
<td>5 3.005 0.938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFTCKNS (mm)</td>
<td>142 0.165 0.053</td>
<td>135 0.154 0.051</td>
<td>5 0.172 0.051</td>
<td>32 0.160 0.064</td>
<td>5 0.202 0.019</td>
<td>5 3.750 3.068</td>
<td>5 4.260 0.636</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLW (mg/mm²)</td>
<td>144 0.075 0.029</td>
<td>133 0.066 0.028</td>
<td>5 0.082 0.027</td>
<td>33 0.062 0.028</td>
<td>5 0.075 0.024</td>
<td>121 0.082 0.024</td>
<td>5 0.075 0.024</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FORM</td>
<td>144 0.716 0.047</td>
<td>133 0.623 0.100</td>
<td>5 0.673 0.067</td>
<td>33 0.737 0.052</td>
<td>5 0.674 0.100</td>
<td>121 0.652 0.097</td>
<td>5 0.720 0.064</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGLTP1 (°)</td>
<td>144 74.067 24.594</td>
<td>133 77.606 30.808</td>
<td>5 72.135 37.908</td>
<td>37 38.280 15.782</td>
<td>5 26.128 10.996</td>
<td>121 38.501 4.450</td>
<td>5 26.128 10.996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGLTP2 (°)</td>
<td>144 87.095 22.844</td>
<td>133 90.095 28.218</td>
<td>5 87.230 33.411</td>
<td>37 38.280 15.782</td>
<td>5 26.128 10.996</td>
<td>121 38.501 4.450</td>
<td>5 26.128 10.996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLWM (cm)</td>
<td>144 2.376 0.421</td>
<td>133 2.093 0.675</td>
<td>5 2.631 0.467</td>
<td>33 2.360 0.531</td>
<td>5 2.391 0.694</td>
<td>121 2.020 0.572</td>
<td>5 2.094 0.562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOBE1 (°)</td>
<td>144 64.007 3.563</td>
<td>133 57.564 6.009</td>
<td>5 60.680 3.480</td>
<td>33 46.800 3.537</td>
<td>5 58.394 6.175</td>
<td>121 58.394 6.175</td>
<td>5 58.394 6.175</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOBE2 (°)</td>
<td>144 47.229 3.460</td>
<td>133 37.617 7.170</td>
<td>5 45.611 4.156</td>
<td>33 49.000 4.100</td>
<td>5 46.000 4.422</td>
<td>121 40.504 6.433</td>
<td>5 45.093 4.965</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPW (cm)</td>
<td>144 3.967 0.421</td>
<td>133 2.914 0.675</td>
<td>5 2.631 0.467</td>
<td>33 2.360 0.531</td>
<td>5 2.391 0.694</td>
<td>121 2.020 0.572</td>
<td>5 2.094 0.562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASRA (W/L)</td>
<td>144 0.557 0.117</td>
<td>133 0.467 0.154</td>
<td>5 0.527 0.110</td>
<td>33 0.559 0.088</td>
<td>5 0.548 0.059</td>
<td>121 0.512 0.142</td>
<td>5 0.589 0.155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRPWD1 (cm)</td>
<td>144 0.836 0.370</td>
<td>133 0.807 0.364</td>
<td>5 0.750 0.358</td>
<td>33 0.529 0.357</td>
<td>5 0.548 0.059</td>
<td>121 0.512 0.142</td>
<td>5 0.589 0.155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRPWD2 (cm)</td>
<td>144 0.460 0.239</td>
<td>133 0.461 0.250</td>
<td>5 0.471 0.239</td>
<td>33 0.581 0.327</td>
<td>5 0.614 0.232</td>
<td>121 0.551 0.396</td>
<td>5 0.659 0.386</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2 Continued

Chapter 3. Untransformed means and standard deviations of the 26 quantitative characters used to identify Corymbia torelliana, C. variegata, C. henryi, C. citriodora and their hybrids with C. torelliana as the maternal parent.

<table>
<thead>
<tr>
<th>Character</th>
<th>C. torelliana</th>
<th>C. variegata</th>
<th>CTxCCV</th>
<th>CTxCCVxCCV</th>
<th>CTxCCVxCT</th>
<th>CTxCCV F2</th>
<th>C. henryi</th>
<th>CTxCH</th>
<th>C. citriodora</th>
<th>CTxCCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
<td>n Mean SD</td>
</tr>
<tr>
<td>LLWM:MPW</td>
<td>144 2.376 0.421 132 2.093 0.675</td>
<td>144 2.317 0.494</td>
<td>5 2.631 0.467</td>
<td>33 2.360 0.531</td>
<td>5 2.391 0.694</td>
<td>121 2.020 0.572</td>
<td>75 2.094 0.561</td>
<td>122 0.297</td>
<td>0.101</td>
<td>54 0.565 0.106</td>
</tr>
<tr>
<td>BLDL:MPW</td>
<td>144 1.701 0.230 133 2.485 0.631</td>
<td>144 1.812 0.258</td>
<td>5 1.947 0.546</td>
<td>33 1.618 0.233</td>
<td>5 1.813 0.413</td>
<td>121 2.212 0.516</td>
<td>75 1.802 0.323</td>
<td>122 3.683</td>
<td>1.088</td>
<td>54 1.837 0.361</td>
</tr>
<tr>
<td>MPW:PRPWD2</td>
<td>144 10.221 4.601 133 7.681 4.352</td>
<td>144 10.120 4.875</td>
<td>5 8.524 3.450</td>
<td>33 8.341 3.064</td>
<td>5 6.785 2.083</td>
<td>121 10.581 7.178</td>
<td>75 7.743 3.797</td>
<td>122 0.063</td>
<td>0.023</td>
<td>54 0.079 0.033</td>
</tr>
<tr>
<td>LLT2:AGLTP2</td>
<td>144 0.080 0.034 133 0.082 0.034</td>
<td>144 0.073 0.033</td>
<td>5 0.108 0.064</td>
<td>33 0.095 0.047</td>
<td>5 0.100 0.037</td>
<td>121 0.101 0.052</td>
<td>75 0.108 0.054</td>
<td>122 2.019</td>
<td>0.779</td>
<td>54 2.272 0.545</td>
</tr>
<tr>
<td>PL:BL</td>
<td>144 0.344 0.154 133 0.297 0.157</td>
<td>144 0.345 0.136</td>
<td>5 0.336 0.218</td>
<td>33 0.387 0.136</td>
<td>5 0.294 0.123</td>
<td>121 0.276 0.194</td>
<td>75 0.351 0.170</td>
<td>122 0.229</td>
<td>0.090</td>
<td>54 0.297 0.153</td>
</tr>
<tr>
<td>LLWM:BLDL</td>
<td>144 0.258 0.051 133 0.219 0.079</td>
<td>144 0.250 0.047</td>
<td>5 0.211 0.050</td>
<td>33 0.276 0.053</td>
<td>5 0.250 0.050</td>
<td>121 0.247 0.064</td>
<td>75 0.284 0.058</td>
<td>122 0.169</td>
<td>0.077</td>
<td>54 0.256 0.050</td>
</tr>
</tbody>
</table>
Chapter 4 Supplementary Material 1. NIR1 average absorbance plots (after Gap-Segment and SNV pre-treatments) of eight-month-old fresh tissue of the different hybrid groups. A. CCC hybrid group; B. CCV hybrid group with F1V, and C. CCV hybrid group with BCT. Blue = C. torelliana, Red = Hybrid, Green = spotted gum.
Chapter 4 Supplementary Material 2. Principal component loadings plot of the first two components derived from NIR1 (after Gap-Segment and SNV pretreatments) of eight-month-old fresh tissue. A. CCC hybrid group, B. CCV hybrid group with F₁V and C. CCV hybrid group with backcross hybrid BCT.
Appendix 5

Supplementary Figure

Chapter 4 Supplementary Material 3. Scores plot of the first two principal components derived from NIR1 (after Gap-Segment and SNV pre-treatments) using eight-month-old fresh tissue of the different hybrid groups. A. CC hybrid group, B. CCV hybrid group with F1V and C. CCV hybrid group with BCT.
Appendix 6

Flowering intensity index of *Corymbia* taxa at four sites over 3 flowering seasons in 2009 to 2011: a – Bonalbo Common, b – Morpeth Park; c – Bonalbo Trial and d = Grafton Trial. Months marked with * indicates that there was no observation. Flowering intensity scores of different *Corymbia* taxa at four sites studied. Note that the F₁ hybrid had flowering overlap with the *C. torelliana* at Bonalbo Common.
Appendix 7. Relative contributions of candidate to chapter/s which was/were published or submitted for publication before thesis submission.

In accordance with the PhD thesis requirements of the Division of Research at Southern Cross University, this appendix details the proportion of contribution of Myralyn Abasolo to the experimental chapters (Chapters 2-4).

Chapter 2. Genetic control of flowering in spotted gum, *Corymbia citriodora* subsp. *variegata* and *C. maculata*.

<table>
<thead>
<tr>
<th></th>
<th>M. Abasolo</th>
<th>Co-Authors Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conception of ideas</td>
<td>30%</td>
<td>70%</td>
</tr>
<tr>
<td>Experimental design</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Field/Laboratory Work</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td>Data analysis</td>
<td>70%</td>
<td>30%</td>
</tr>
<tr>
<td>Writing</td>
<td>40%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Chapter 3. Identification of intersectional *Corymbia* hybrids based on seedling morphology improves with parental divergence.

<table>
<thead>
<tr>
<th></th>
<th>M. Abasolo</th>
<th>Co-Authors Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conception of ideas</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td>Experimental design</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td>Field/Laboratory Work</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>Data analysis</td>
<td>80%</td>
<td>20%</td>
</tr>
<tr>
<td>Writing</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Chapter 4. Deviant near-infrared spectra allow identification of *Corymbia* hybrids.

<table>
<thead>
<tr>
<th></th>
<th>M. Abasolo</th>
<th>Co-Authors Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conception of ideas</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td>Experimental design</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td>Field/Laboratory Work</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>Data analysis</td>
<td>80%</td>
<td>20%</td>
</tr>
<tr>
<td>Writing</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Principal PhD Supervisor

Signature: …………………………………… Date:……………………

DR. MERVYN SHEPHERD

School Director for Higher Degrees by Research Training

Signature:………………………………….. Date:……………………

DR. HANS WOHLMUTH
References


References


Hingston, A.B., Potts, B.M., Vaillancourt, R.E., 2003. The risk of genetic pollution of native eucalypts from plantations and farm forestry in Victoria: a scoping study. In. Cooperative Research Centre for Sustainable Production Forestry and School of Plant Sciences, University of Tasmania, Tasmania, p. 79.


References

Jones, M., Stokoe, R., Cross, M., Scott, L., Maguire, T., Shepherd, M., 2001. Isolation of microsatellite loci from spotted gum (*Corymbia variegata*), and cross-species amplification in *Corymbia* and *Eucalyptus*. Molecular Ecology Notes 1, 276-278.


References


References


References


hybridization between cultivated poplars and their wild relatives: evidence and 
consequences for native poplar populations. Annals of Forest Science 62, 601-
613.

Vila C., Sandqvist A.K., Flagstad O. et.al., 2002. Rescue of severely bottlenecked wolf 
(*Canis lupus*) population by a simple immigrant. Proceedings Biological 
Sciences, 270, 91-97.


orchard of *Eucalyptus globulus*. In, Research Working Group No.1 of the 
Australian Forestry Council, Forest Genetics. Australian Forestry Council, 
Canberra.

hardwood forestry - without the weediness. In: Lee, D. (Ed.), Corymbia 
Research Meeting: Underpinning development of a profitable hardwood 
plantation industry in northern Australia by research into Corymbia species 
and hybrids. Queensland Government, Department of Primary Industries and 
Fisheries, Gympie, Queensland, pp. 7-8.

estimate the potential intake of *Eucalyptus* foliage by folivorous marsupials. 

biogeography and contemporary distribution of eucalypts. In: Williams, J.L., 
Woinarski, J.C.Z. (Eds.), Eucalypt ecology: individuals to ecosystems. 

acclimation of leaf traits to rainfall in 29 species of 16-year-old *Eucalyptus* at 


